MMBA123CCT

کاروباری تجزیات اور شماریات

(Business Analytics and Statistics)

ماسٹر آف بزنس ایڈمنسٹریشن-ایم بی اے

دوسالہ پروگرام

پہلا سمسٹر

مرکز برائے فاصلاتی و آن لائن تعلیم

مولانا آزاد نیشنل اردو یونیورسٹی

حیدرآباد-32، تلنگانہ، بھارت

Copyright © 2025, Maulana Azad National Urdu University, Hyderabad
All right reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronically or mechanically, including photocopying, recording or any information storage or retrieval system, without prior permission in writing form the publisher (registrar@manuu.edu.in)

ISBN	:	978-81-994080-0-5
Course	:	Business Analytics and Statistics
First Edition	:	September 2025
Copies	:	1875
Price	:	180/- (The price of the book is included in admission fee of distance mode students)

Course Coordinator

Prof. Mohammed Abdul Azeem, Head, Department of Management & Commerce, MANUU

Editorial Board/Editors

Prof. Mohammed Abdul Azeem, Department of Management & Commerce, MANUU
Prof. Maqsood Khan, Ex-Director Sydenham Institute of Management Studies, Mumbai
Dr. Md. Rashid Farooqi Assist. Prof. Department of Management & Commerce, MANUU
Dr. Syed Hamid Mohiuddin Quadri, Assist Prof. (Contract) Department of Management & Commerce, MANUU
Dr. Md Jafar, Assistant Professor/Guest Faculty, CDOE, MANUU. (Language Editor)

Production

Prof. Nikhath Jahan, Professor (Urdu), CDOE MANUU	Mr. P Habibulla, Assistant Registrar, Purchase & Stores Section, MANUU	Dr. Mohd Akmal Khan, Assistant Professor (C), CDOE MANUU
Mohd Abdul Naseer, Section Officer, CDOE MANUU	Shaik Ismail, UDC, CDOE, MANUU	Syed Faheemuddin, LDC, Purchase & Stores Section, MANUU

On behalf of the Registrar, Published by:

Centre for Distance and Online Education

Maulana Azad National Urdu University
Gachibowli, Hyderabad-500032 (TG), India
Director: dir.dde@manuu.edu.in *Publication:* ddepublication@manuu.edu.in
Phone number: 040-23008314 Website: manuu.edu.in

CRC Prepared by

Printed at : Karshak Print Solutions Private Limited, Hyderabad

فہرست

پیغام		وائس چانسلر، مولانا آزاد نیشنل اردو یونیورسٹی	**5**
پیغام		ڈائرکٹر، مرکز برائے فاصلاتی وآن لائن تعلیم	**6**
کورس کا تعارف		کورس کوآرڈی نیٹر	**7**
اکائی نمبر	**اکائی کا نام**	**مصنف**	**صفحہ نمبر**
1۔	کاروباری تجزیات کا تعارف	ڈاکٹر سید حامد محی الدین قادری، شعبۂ مینجمنٹ وکامرس، مانو، حیدرآباد	9
2۔	ڈیٹا کے لیے فریم ورک	ڈاکٹر سید حامد محی الدین قادری، شعبۂ مینجمنٹ وکامرس، مانو، حیدرآباد	25
3۔	تجزیات کی صلاحیتی ترقی	ڈاکٹر سید حامد محی الدین قادری، شعبۂ مینجمنٹ وکامرس، مانو، حیدرآباد	55
4۔	شماریات اور فیصلہ سازی	ڈاکٹر محمد عابد علی، مرکز برائے فاصلاتی وآن لائن تعلیم، مانو، حیدرآباد	81
5۔	ڈسپرزن کی پیائش	ڈاکٹر محمد عابد علی، مرکز برائے فاصلاتی وآن لائن تعلیم، مانو، حیدرآباد	127
6۔	احتمال اور احتمال کی تقسیم	جناب محمد اطہر، شعبۂ مینجمنٹ وکامرس، مانو، حیدرآبا	148
7۔	کورلیشن کا تجزیہ	ڈاکٹر محمد عابد علی، مرکز برائے فاصلاتی وآن لائن تعلیم، مانو، حیدرآباد	168
8۔	رگریشن کا تجزیہ	ڈاکٹر محمد عابد علی، مرکز برائے فاصلاتی وآن لائن تعلیم، مانو، حیدرآباد	207
9۔	ٹائم سیریز کا تجزیہ	ڈاکٹر محمد عابد علی، مرکز برائے فاصلاتی وآن لائن تعلیم، مانو، حیدرآباد	234
10۔	سیمپلنگ تھیوری اور سگنیفیکے نس کا ٹیسٹ	ڈاکٹر ایم ڈی راشد فاروقی، شعبۂ مینجمنٹ وکامرس، مانو، حیدرآباد	255

11۔	کائی اسکوائر کا تجزیہ	ڈاکٹر ایم ڈی راشد فاروقی، شعبۂ مینجمنٹ و کامرس، مانو، حیدرآباد	266
12۔	ویریانس (variance) کا تجزیہ	ڈاکٹر ایم ڈی راشد فاروقی، شعبۂ مینجمنٹ و کامرس، مانو، حیدرآباد	277
	نمونہ امتحانی پرچہ		288

پیغام

مولانا آزاد نیشنل اردو یونیورسٹی (MANUU) 1998 میں پارلیمنٹ کے ایک ایکٹ کے ذریعے قائم کی گئی۔ یہ ایک مرکزی جامعہ ہے جس نے این اے اے سی کی جانب سے گریڈ +A حاصل کیا ہے۔ اس جامعہ کے قیام کے مقاصد ہیں: (1) اردو زبان کا فروغ، (2) پیشہ ورانہ اور تکنیکی تعلیم کو اردو میڈیم میں قابل رسائی اور دستیاب بنانا، (3) روایتی اور فاصلاتی طریقۂ تعلیم کے ذریعے تعلیم فراہم کرنا، اور (4) خواتین کی تعلیم پر خصوصی توجہ دینا۔ یہ وہ نکات ہیں جو اس مرکزی جامعہ کو دیگر تمام مرکزی جامعات سے ممتاز کرتے ہیں اور اسے ایک انفرادیت بخشتے ہیں۔ قومی تعلیمی پالیسی 2020 میں بھی مادری زبانوں اور علاقائی زبانوں میں تعلیم حاصل کرنے پر زور دیا گیا ہے۔

اردو کے ذریعے علم کے فروغ کا مقصد یہی ہے کہ اردو جاننے والے طبقے کے لیے عصری علوم اور مضامین تک رسائی آسان بنائی جائے۔ ایک طویل عرصے تک اردو میں درسی مواد کی کمی رہی ہے۔ اردو یونیورسٹی کے پاس اب اردو میں 350 سے زیادہ کتابوں کا ذخیرہ موجود ہے اور ہر سمسٹر کے ساتھ اس تعداد میں اضافہ ہو رہا ہے۔

اردو یونیورسٹی این ای پی 2020 کے وژن کے مطابق مادری/گھریلو زبان میں تعلیمی مواد فراہم کرنے کے قومی مشن کا حصہ بننے کو اپنے لیے ایک اعزاز سمجھتی ہے۔ مزید یہ کہ اردو بولنے والا طبقہ اردو میں مطالعہ کے مواد کی عدم دستیابی کے سبب نئے اُبھرتے شعبوں اور جدید تر معلومات کے موجودہ میدانوں میں تازہ ترین معلومات واطلاعات کے حصول سے محروم نہیں رہے گا۔ مذکورہ بالا میدانوں میں مواد کی دستیابی کی بدولت حصولِ معلومات کا نیا شعور بیدار ہوا ہے جو یقیناً اردو داں طبقے کی دانشورانہ ترقی پر اثر انداز ہوگا۔

فاصلاتی اور آن لائن طلبہ کے لیے تعلیم وتدریس کے عمل کو سہل بنانے کے لیے یونیورسٹی کا سینٹر فار ڈسٹنس اینڈ آن لائن ایجوکیشن (CDOE) اردو اور متعلقہ مضامین میں خود اکتسابی مواد (SLM) کی تیاری کو یقینی بناتا ہے۔

مولانا آزاد نیشنل اُردو یونیورسٹی (MANUU) فاصلاتی اور آن لائن لرننگ کے طلبہ کے لیے SLM بلا معاوضہ فراہم کرتی ہے۔ یہ مواد اُن تمام افراد کے لیے برائے نام قیمت پر دستیاب ہے جو اُردو کے ذریعے انتظام وتجارت (Management and Business Administration) کے میدان میں علم حاصل کرنے میں دلچسپی رکھتے ہیں۔ تعلیم تک رسائی کے دائرے کو مزید پھیلانے کے مقصد سے، اردو/ہندی/انگریزی/عربی میں eSLM یونیورسٹی کی ویب سائٹ پر مفت ڈاؤن لوڈ کے لیے دستیاب رکھا گیا ہے۔

مجھے بے حد خوشی ہے کہ متعلقہ فیکلٹی کی محنت اور مصنفین کے مکمل تعاون کی بدولت ایم۔بی۔اے (فاصلاتی وآن لائن لرننگ) پروگرام کی کتابوں کی اشاعت کا عمل بڑے پیمانے پر شروع ہو گیا ہے۔ فاصلاتی اور آن لائن لرننگ کے طلبہ کی سہولت کے لیے خود اکتسابی مواد (SLM) کی تیاری اور اشاعت کا عمل یونیورسٹی کے لیے نہایت اہمیت رکھتا ہے۔ مجھے یقین ہے کہ ہم اپنے خود تعلیمی مواد کے ذریعے اردو جاننے والے ایک بڑے طبقے کی تعلیمی اور پیشہ ورانہ ضروریات کو پورا کرنے کے قابل ہوں گے، یونیورسٹی کے مقصدِ قیام کو پورا کریں گے، اور اپنے ملک میں اپنی موجودگی کو مزید مضبوط اور بامعنی بنا سکیں گے۔

نیک تمناؤں کے ساتھ!

پروفیسر سید عین الحسن
شیخ الجامعہ، مانو

پیغام

موجودہ دور میں فاصلاتی تعلیم کو دنیا بھر میں ایک نہایت مؤثر اور مفید طریقۂ تعلیم کے طور پر تسلیم کیا جاتا ہے اور بڑی تعداد میں لوگ اس طریقۂ تعلیم سے فائدہ اٹھا رہے ہیں۔ مولانا آزاد نیشنل اردو یونیورسٹی نے بھی اردو زبان بولنے والے عوام کی تعلیمی ضروریات کو مد نظر رکھتے ہوئے اپنے قیام کے وقت سے ہی فاصلاتی تعلیم کا طریقہ متعارف کرایا۔ مولانا آزاد نیشنل اردو یونیورسٹی نے 1998 میں ڈائریکٹوریٹ آف ڈسٹنس ایجوکیشن (نظامتِ فاصلاتی تعلیم) کے ساتھ کام کا آغاز کیا اور 2004 سے باقاعدہ پروگرام شروع ہوئے، اس کے بعد مختلف شعبہ جات قائم کیے گئے۔

یو جی سی نے ملک میں نظام تعلیم کو مؤثر طور پر منظم کرنے میں ایک اہم کردار ادا کیا ہے۔ اوپن اینڈ ڈسٹنس لرننگ (ODL) موڈ کے تحت چلنے والے مختلف پروگرام، جو سینٹر فار ڈسٹنس اینڈ آن لائن ایجوکیشن (CDOE) میں چل رہے ہیں، یو جی سی-ڈی ای بی کے منظور شدہ ہیں۔ یو جی سی-ڈی ای بی نے فاصلاتی اور باقاعدہ تعلیم کے نصاب کو ہم آہنگ کرنے پر زور دیا ہے تاکہ فاصلاتی تعلیم حاصل کرنے والے طلبہ کے معیار کو بہتر بنایا جاسکے۔ چونکہ مولانا آزاد نیشنل اردو یونیورسٹی ایک دُہرے طرز (ڈوئل موڈ) کی یونیورسٹی ہے جو فاصلاتی اور روایتی دونوں طریقۂ تعلیم کی خدمات فراہم کرتی ہے، اس لیے اپنے مقاصد کو یو جی سی-ڈی ای بی کے رہنما خطوط کے مطابق حاصل کرنے کے لیے اس نے چوائس بیسڈ کریڈٹ سسٹم (CBCS) متعارف کرایا گیا اور خود اکتسابی مواد (Self Learning Materials) نئے سرے سے تیار کیے جارہے ہیں، جو بی۔اے۔ کے پروگراموں کے لیے 6 بلاکس پر مشتمل ہیں جن میں 24 یونٹس شامل ہیں اور ایم۔اے۔ کے پروگراموں کے لیے 4 بلاکس پر مشتمل ہیں جن میں 16 یونٹس شامل ہیں۔ سینٹر فار ڈسٹنس اینڈ آن لائن ایجوکیشن (CDOE) کل انیس (19) پروگرام پیش کرتا ہے جن میں یو جی، پی جی، بی ایڈ، ڈپلومہ اور سرٹیفکیٹ پروگرام شامل ہیں۔ اس کے ساتھ ساتھ تکنیکی مہارتوں پر مبنی پروگرام بھی شروع کیے جارہے ہیں۔ سی ڈی او ای نے جولائی 2025 سے این ای پی-2020 کے مطابق چار سالہ یو جی پروگرام کا آغاز کیا ہے۔ بی اے، بی ایس سی اور بی کام کے آنرز پروگراموں کو این سی ایف کے مطابق ڈیزائن کیا گیا ہے جس سے طلبہ کو آنرز ڈگری حاصل کرنے میں مدد ملے گی۔ سال 2025-2026 سے ایم بی اے پروگرام او ڈی ایل موڈ میں متعارف کرایا گیا ہے۔

مانو نے طلبہ کی سہولت کے لیے نو ریجنل سنٹرز (بنگلورو، بھوپال، دربھنگہ، دہلی، کولکاتا، ممبئی، پٹنہ، رانچی اور سری نگر) اور چھ سب ریجنل سنٹرز (حیدرآباد، لکھنؤ، جموں، نوح، وارانسی اور امراؤتی) کا ایک وسیع نیٹ ورک قائم کیا ہے۔ اس کے علاوہ وجے واڑا میں ایک ایکسٹینشن سنٹر بھی قائم کیا گیا ہے۔ ان ریجنل اور سب ریجنل سنٹروں کے تحت ایک سو ساٹھ سے زیادہ لرنر سپورٹ سنٹر (LSCs) اور بیس پروگرام سنٹر بیک وقت چلائے جارہے ہیں تاکہ طلبہ کو تعلیمی اور انتظامی سہولیات فراہم کی جاسکیں۔ سینٹر فار ڈسٹنس اینڈ آن لائن ایجوکیشن اپنی تعلیمی اور انتظامی سرگرمیوں میں آئی سی ٹی کا بھرپور استعمال کرتا ہے اور اپنے تمام پروگراموں میں صرف آن لائن موڈ کے ذریعے ہی داخلے فراہم کرتا ہے۔ طلبہ کے لیے سیلف لرننگ میٹیریل (SLM) کی سوفٹ کاپیاں سینٹر فار ڈسٹنس اینڈ آن لائن ایجوکیشن (Centre for Distance and Online Education) کی ویب سائٹ پر دستیاب کرائی گئی ہیں۔ ساتھ ہی، ایم۔بی۔اے (فاصلاتی و آن لائن لرننگ) پروگرام کے مختلف مضامین سے متعلق آڈیو و ویڈیو لیکچرز کے لنکس بھی ویب سائٹ پر فراہم کیے گئے ہیں۔ طلبہ کی سہولت کے لیے ای۔ میل اور واٹس ایپ گروپ کی سہولت بھی فراہم کی گئی ہے، جن کے ذریعے انہیں پروگرام کے مختلف پہلوؤں جیسے کورس رجسٹریشن، اسائنمنٹ، کاؤنسلنگ، امتحانات وغیرہ کے بارے میں بروقت مطلع کیا جاتا ہے۔

باقاعدہ کاؤنسلنگ کے علاوہ، طلبہ کے تعلیمی معیار کو بہتر بنانے کے لیے زائد تدارکی (Remedial) آن لائن کاؤنسلنگ کا اہتمام بھی کیا گیا ہے، جو اُن طلبہ کے لیے نہایت مفید ثابت ہوئی ہے جو کسی موضوع میں مزید رہنمائی چاہتے ہیں۔ امید کی جاتی ہے کہ ایم۔بی۔اے (فاصلاتی و آن لائن لرننگ) پروگرام انتظامیہ اور کاروبار (Management and Business) کی تعلیم کو اُردو کے ذریعے زیادہ قابل رسائی اور بامقصد بنانے میں اہم کردار ادا کرے گا اور مختلف پس منظر سے تعلق رکھنے والے طلبہ کو اعلیٰ تعلیمی معیار کے ساتھ عملی و پیشہ ورانہ مہارت حاصل کرنے کے مواقع فراہم کرے گا۔

پروفیسر محمد رضا اللہ خان

ڈائرکٹر، سی ڈی او ای، مانو

کورس کا تعارف

عزیز طلبا!

MBA کا خود اکتسابی مواد آپ کی خدمت میں پیش ہے۔ یہ جامع انداز میں تیار کردہ مواد آپ کو خود سے سمجھنے اور سیکھنے میں مددگار ہوگا۔ زیر نظر خود اکتسابی مواد جدید تعلیمی اصولوں اور MBA کے کورس کے تقاضوں اور ضروریات کو مد نظر رکھتے ہوئے تیار کیا گیا ہے تاکہ فاصلاتی تعلیم (ODL) سے وابستہ طلبا آسانی سے سمجھ سکیں۔ یہ مواد ایک مؤثر انداز میں اس طرح ترتیب دیا گیا ہے کہ نصاب کے مطابق معلومات جامع انداز میں طلبا کو مہیا ہوں گی۔

ہر بلاک میں یونٹ رکھے گئے ہیں تاکہ مرحلہ وار اور وضاحت کے ساتھ مواد پیش ہو۔ موضوعات کو قابل فہم اور بہتر طور پر سمجھنے میں مواد کی ترتیب کا خاص خیال رکھا گیا ہے۔ ہر یونٹ کی ابتدا تمہید سے ہوتی ہے، جس کے بعد اکائی کے مقاصد بیان کر کے اصل موضوع پر بات کی جاتی ہے۔ موضوع کے اختتام کے بعد اکتسابی نتائج میں اکائی کی تمام معلومات کو جامع انداز میں یکجا کیا جاتا ہے تاکہ طلبا اکائی کے مقاصد میں طے شدہ نکات کو سمجھ سکیں۔ اس کے بعد نمونہ امتحانی سوالات میں مختلف قسم کے سوالات سے طلبا خود کا محاسبہ (Self-assessment) کر سکتے ہیں۔

تیار کردہ مواد مختلف مضامین کے امتحانات کی تیاری میں طلبا کے لیے مددگار ہوگا۔ علاوہ ازیں تجارتی میدان میں درکار مہارتوں کو فروغ دینے میں بھی معاون ہوگا۔

یقیناً یہ خود اکتسابی مواد (SLM) تجارتی و کاروباری میدان سے متعلق درکار مہارتوں کے فروغ میں مددگار ہوگا اور ساتھ ہی ساتھ تنقیدی سوچ، خود اعتمادی اور فیصلہ سازی کی صلاحیتوں کو پروان چڑھانے میں میں مؤثر ہوگا۔

اس خود اکتسابی مواد (SLM) کو تیار کرنے میں اساتذہ اور اس میدان کے ماہرین کی ٹیم نے مدد کی ہے۔ جس کے لیے مرکز برائے فاصلاتی اور آن لائن تعلیم (CDOE) ان کا شکر گزار ہے۔ امید ہے کہ یہ کتاب طلبا و طالبات اور اساتذہ کرام کے لیے مفید ثابت ہو گی۔

ایم بی اے کے دو سالہ پروگرام میں کل چار سمسٹر ہیں۔ یہ زیر نظر پرچہ "کاروباری تجزیات اور شماریات" پہلے سال کے پہلے سمسٹر میں شامل ہے۔

کورس کوآرڈی نیٹر

کاروباری تجزیات اور شماریات

(Business Analytics and Statistics)

اکائی 1۔ بزنس اینالیٹکس کا تعارف

(Intorduction to Business Analytics)

اکائی کے اجزاء	Structure
1.0 تمہید	Introduction
1.1 مقاصد	Objectives
1.2 بزنس اینالیٹکس کا معنی	Introduction of Business Analytics
1.3 بزنس اینالیٹکس کی اہمیت	Importance of Business Analytics
1.4 بزنس اینالیٹکس کے اقسام	Types of Business Analytics
1.5 بگ ڈیٹا	Big Data
1.6 بگ ڈیٹا کے اقسام	Types of Big Data
1.7 بگ ڈیٹا کی خصوصیات	Features of Big Data
1.8 سوشل میڈیا اینالیٹکس	Social Media Analytics
1.9 مختلف سوشل میڈیا پلیٹ فارم پر سوشل میڈیا اینالیٹکس کا اطلاق	Application of Social Media Analytics on different Social Media Platforms
1.10 اکتسابی نتائج	
1.10 کلیدی الفاظ	Keywords
1.11 نمونہ امتحانی سوالات	Model Examination Questions

1.0 تمہید (Introduction)

آج کے ڈیٹا سے چلنے والے کاروباری ماحول میں، باخبر اور موثر فیصلے کرنے کے لیے بڑی مقدار میں معلومات اکٹھا کرنا، تجزیہ کرنا اور اس کی تشریح کرنا ضروری ہے۔ یہ وہ جگہ ہے جہاں بزنس اینالیٹکس آتا ہے۔ بزنس اینالیٹکس ایک ایسا شعبہ ہے جو اعداد و شمار کے تجزیہ، ڈیٹا مائننگ، مشین لرننگ، اور دیگر تکنیکوں کو ڈیٹا سے معلومات حاصل کرنے اور کاروباری حکمت عملی کو چلانے کے لیے استعمال کرتا ہے۔

بزنس اینالیٹکس سے مراد باخبر کاروباری فیصلوں کو آگے بڑھانے اور تنظیمی کارکردگی کو بڑھانے کے لیے ڈیٹا، شماریاتی تجزیہ اور

مقداری طریقوں کا منظم استعمال ہے۔ اعداد و شمار، انفارمیشن ٹیکنالوجی اور مینجمنٹ سے ٹولز اور تکنیکوں کو یکجا کر کے، بزنس اینالیٹکس تنظیموں کو خام ڈیٹا سے بامعنی آگاہی حاصل کرنے میں مدد کرتا ہے۔ یہ بصیرتیں رجحانات کی نشاندہی کرنے، پیچیدہ مسائل کو حل کرنے، کارکردگی کو بہتر بنانے اور اسٹریٹجک منصوبہ بندی کی حمایت کے لیے اہم ہیں۔

بزنس اینالیٹکس کی اہمیت فیصلہ سازی کے لیے ڈیٹا پر مبنی بنیاد فراہم کرنے کی صلاحیت میں مضمر ہے۔ یہ کاروبار کو ان کی موجودہ کارکردگی کو سمجھنے، مستقبل کے نتائج کا اندازہ لگانے اور متحرک ماحول میں مسابقتی رہنے کے لیے فعال اقدامات کرنے میں مدد کرتا ہے۔ ڈیجیٹل ٹیکنالوجی کی تیز رفتار ترقی کے ساتھ، کاروبار اب بہت زیادہ ڈیٹا کے سامنے آ رہے ہیں، جسے اکثر "بگ ڈیٹا" کہا جاتا ہے، جس میں ویب پلیٹ فارمز اور سوشل میڈیا چینلز سمیت مختلف ذرائع سے جمع کی گئی ساختہ اور غیر ساختہ دونوں معلومات شامل ہیں۔

بزنس اینالیٹکس کو وسیع طور پر تین اہم اقسام میں تقسیم کیا جا سکتا ہے: وضاحتی تجزیات، جو ماضی کی کارکردگی کو سمجھنے کے لیے تاریخی ڈیٹا کی جانچ پر توجہ مرکوز کرتا ہے پیشین گوئی کے تجزیات، جو مستقبل کے ایونٹس کا اندازہ لگانے کے لیے شماریاتی ماڈل اور پیشین گوئی کی تکنیکوں کا استعمال کرتا ہے۔ اور نسخہ جات کے تجزیات، جو پیش گوئی کرنے والی بصیرت کی بنیاد پر بہترین اقدامات تجویز کرتا ہے۔ مزید برآں، صارفین کے رویے اور مارکیٹ کے رجحانات کو حقیقی وقت میں سمجھنے کے لیے ویب اور سوشل میڈیا تجزیہ تیزی سے اہم ہو گیا ہے۔

آج کی ڈیٹا سینٹرک دنیا میں، بزنس اینالیٹکس میں مہارت حاصل کرنا ان تنظیموں کے لیے ضروری ہے جو اختراعات، چست رہنے، اور معتبر شواہد پر مبنی فیصلے کرنے کی خواہاں ہیں۔

1.1 مقاصد (Objectives)

اس اکائی کا مطالعہ کرنے کے بعد آپ:

- جدید فیصلہ سازی میں بزنس اینالیٹکس کے معنی، تعریف اور اہمیت کو سمجھ سکیں گے۔
- تجزیات کی اقسام کی شناخت اور ان میں فرق کریں: وضاحتی، پیشین گوئی، اور نسخہ کا جائزہ لے سکیں گے۔
- بزنس اپلی کیشن کے تناظر میں بگ ڈیٹا کے تصور، اقسام اور خصوصیات کی وضاحت کر سکیں گے۔
- صارفین کی مشغولیت اور مارکیٹنگ کے فیصلوں کو چلانے میں ویب اور سوشل میڈیا اینالیٹکس کے کردار کا تجزیہ کر سکیں گے۔
- بزنس اینالیٹکس کے مختلف کاروباری افعال میں ٹولز اور حقیقی دنیا کی اپلی کیشن کو پہچان سکیں گے۔

1.2 بزنس اینالیٹکس کے معنی (Meaning of Business Analytics)

بزنس اینالیٹکس سے مراد کاروباری فیصلوں کو مطلع کرنے کے لیے اعداد و شمار سے بامعنی بصیرت حاصل کرنے کے لیے مقداری طریقے استعمال کرنے کی مشق ہے۔ یہ فیلڈ اعداد و شمار، انفارمیشن ٹیکنالوجی اور کاروبار کی مہارتوں کو مربوط کرتا ہے پیچیدہ مسائل کو حل کرنے اور کمپنی کی کارکردگی کو بہتر بنانے کے لیے ڈیٹا سے چلنے والی تکنیکوں پر کرتا ہے کرتا ہے۔ بزنس اینالیٹکس میں رجحانات کی پیشن گوئی، عمل کو

بہتر بنانے اور اسٹریٹجک منصوبہ بندی کو بڑھانے کے لیے ڈیٹا کی وسیع مقدار کو جمع کرنا، پروسیسنگ کرنا اور تجزیہ کرنا شامل ہے۔

اس میں مختلف طریقوں کو شامل کیا گیا ہے، بشمول وضاحتی تجزیات (ماضی کی کارکردگی کی جانچ کرنا)، پیش گوئی کرنے والے تجزیات (مستقبل کے منظر ناموں کی پیشن گوئی)، اور نسخے کے تجزیات (پیش گوئیوں کی بنیاد پر اقدامات کی تجویز)۔ بزنس اینالیٹکس میں نار مل طور پر استعمال ہونے والے ٹولز میں ڈیٹا ویژولائزیشن سافٹ ویئر، شماریاتی تجزیہ کے ٹولز، اور بزنس انٹیلی جنس پلیٹ فارم شامل ہیں۔ یہ نقطہ نظر نہ صرف کاروباری اداروں کو ان کی موجودہ حالت کو سمجھنے میں مدد کرتا ہے بلکہ مستقبل کے مواقع اور چیلنجوں کا بھی اندازہ لگاتا ہے۔

1.3 بزنس اینالیٹکس کی اہمیت (Importance of Business Analaytics)

بزنس اینالیٹکس کئی وجوہات کی بناء پر ضروری ہے:

1. بہتر فیصلہ سازی Improved Decision Making

بزنس اینالیٹکس تنظیموں کو آگاہی فراہم کرتا ہے جو باخبر فیصلے کرنے میں ان کی مدد کر سکتی ہے۔ ڈیٹا کا تجزیہ کر کے، تنظیمیں پیٹرن اور رجحانات کی شناخت کر سکتی ہیں، گاہک کے رویے کو سمجھ سکتی ہیں، اور مستقبل کے نتائج کی پیشین گوئی کر سکتی ہیں، اور انہیں حکمت عملی کے فیصلے کرنے کے قابل بناتی ہیں جو ان کی کارکردگی کو بہتر بناتے ہیں۔

2. مسابقتی فائدہ Competitive Advantage

آج کے ڈیٹا سے چلنے والے کاروباری ماحول میں اپنے ڈیٹا کو مؤثر طریقے سے استعمال کرنے والی تنظیموں کو ایک اہم مسابقتی فائدہ حاصل ہے۔ بزنس اینالیٹکس تنظیموں کو اس قابل بناتا ہے کہ وہ اپنے صارفین کو بہتر طور پر سمجھ سکیں، ان کے کاموں کو بہتر بنائیں، اور مارکیٹ کے نئے مواقع کی نشاندہی کریں، جس سے انہیں بازار میں برتری حاصل ہوگی۔

3. لاگت میں کمی Cost Reduction

بزنس اینالیٹکس تنظیموں کو ان کے کاموں میں نقصان اور ضائع ہونے کی نشاندہی کرنے میں مدد کر سکتے ہیں، جس سے وہ اپنے عمل کو بہتر بنا سکتے ہیں اور اخراجات کم کر سکتے ہیں۔ اعداد و شمار کا تجزیہ کر کے، تنظیمیں ان علاقوں کی نشاندہی کر سکتی ہیں جہاں وہ اخراجات کو کم کر سکتی ہیں، جیسے سپلائی چین کی اصلاح، انوینٹری مینجمنٹ، اور پروسیس آٹومیشن۔

4. بہتر کسٹمر کا تجربہ Enhanced Customer Experience

کسٹمر ڈیٹا کا تجزیہ کر کے، تنظیمیں اپنے صارفین کی ترجیحات، رویے، اور ضروریات کو سمجھ سکتی ہیں، انہیں بہتر مصنوعات اور خدمات فراہم کرنے کے قابل بناتی ہیں جو ان کے صارفین کی توقعات پر پورا اترتی ہیں۔ بدلے میں، گاہکوں کے اطمینان اور بھروسے میں اضافہ ہوتا ہے۔

5. رسک مینجمنٹ Risk Management

بزنس اینالیٹکس تنظیموں کو خطرات کی شناخت اور is سے بچنے میں مدد کرتا ہے۔ تنظیمیں ڈیٹا کا تجزیہ کر کے اور ان کے اثرات کو کم کرنے کے لیے احتیاطی تدابیر اختیار کر کے ممکنہ خطرات کا اندازہ لگا سکتی ہیں۔

1.4 بزنس اینالیٹکس کے اقسام (Types of Buesiness Analytics)

بزنس اینالیٹکس کی اقسام مندرجہ ذیل ہیں:

A. وضاحتی تجزیات Descriptive Analytics

وضاحتی تجزیات بزنس اینالیٹکس کی سب سے آسان قسم ہے جو نمونوں، رجحانات اور بے ضابطگیوں کی بصیرت حاصل کرنے کے لیے تاریخی ڈیٹا کو بیان کرنے یا خلاصہ کرنے کے لیے استعمال ہوتی ہے۔ یہ تجزیہ ماضی میں کیا ہوا ہے اس کا ایک جائزہ فراہم کرتا ہے، تنظیموں کو بہتری کے لیے جگہوں کی نشاندہی کرنے اور ڈیٹا پر مبنی فیصلے کرنے کی اجازت دیتا ہے۔

مثال: ایک ریٹیل چین پچھلے سال کے سیلز ڈیٹا کا تجزیہ کرتا ہے تاکہ خریداری کے عروج کے موسموں کی نشاندہی کی جاسکے۔

اسٹیٹ: پبلک ریٹیل لیگ کے مطابق، 2023 میں چھٹیوں کے موسم کے دوران امریکہ میں خوردہ سودے 1.3$ ٹریلین سے تجاوز کر گئے۔

Tools for Descriptive Analytics:

- Microsoft Excel
- Tableau
- Power BI

وضاحتی تجزیات میں استعمال ہونے والی درج ذیل معیاری تکنیکیں (standard techniques) یہ ہیں:

1. ڈیٹا ویژولائزیشن اور ڈیش بورڈز Data Visualization and Dashboards

ڈیٹا ویژولائزیشن میں ڈیٹا کی تصویری نمائندگی کی تخلیق شامل ہے تاکہ تفہیم اور تشریح کو آسان بنایا جاسکے۔ اس میں چارٹس، گرافس، اور دیگر بصری امداد شامل ہیں جو ڈیٹا کے رجحانات، نمونوں اور بے ضابطگیوں کو واضح طور پر ظاہر کرتی ہیں۔ ڈیش بورڈز ڈیٹا ویژولائزیشن کی ایک قسم ہیں جو کلیدی کارکردگی کے اشارے (KPIs) اور دیگر ضروری میٹرکس کا ایک نظر میں نظارہ فراہم کرتی ہے۔

ڈیٹا ویژولائزیشن اور ڈیش بورڈز نارمل طور پر کاروباری کارکردگی کی نگرانی، اہداف کے خلاف پیشرفت کو ٹریک کرنے، اور بہتری کے لیے شعبوں کی نشاندہی کرنے کے لیے وضاحتی تجزیات میں استعمال ہوتے ہیں۔

2. شماریاتی تجزیہ اور رپورٹنگ Statistical Analysis and Reporting

شماریاتی تجزیہ میں اعداد و شمار کا تجزیہ اور تشریح کرنے کے لیے شماریاتی طریقوں کا استعمال شامل ہے۔ اس میں مرکزی رجحان، تغیر، ارتباط، مفروضے کی جانچ، اور رجعت تجزیہ کے اقدامات شامل ہیں۔ شماریاتی تجزیہ اعداد و شمار کے نمونوں اور رجحانات کی شناخت کرنے اور ان نمونوں کی اہمیت کو جانچنے میں مدد کرتا ہے۔

رپورٹنگ میں اعداد و شمار کے تجزیہ سے حاصل شدہ نتائج اور مواصلاتی آگاہی شامل ہے۔ رپورٹوں میں نارمل طور پر ٹیبلز، چارٹس اور دیگر بصری امداد شامل ہوتی ہیں جو ڈیٹا کو واضح اور اختصار کے ساتھ پیش کرتی ہیں۔ رپورٹنگ وضاحتی تجزیات کا ایک لازمی حصہ ہے، کیونکہ اس سے

نتائج کو اسٹیک ہولڈرز اور فیصلہ سازوں تک پہنچانے میں مدد ملتی ہے۔

3. تحقیقی ڈیٹا کا تجزیہ Exploratory Data Analysis

ایکسپلوریٹری ڈیٹا اینالیسس میں معلومات حاصل کرنے اور پیٹرن اور تعلقات کی شناخت کے لیے ڈیٹا کی تلاش شامل ہے۔ یہ ایک لچکدار اور تکراری عمل ہے جو ڈیٹا کو سمجھنے کے لیے متعدد تکنیکوں کا استعمال کرتا ہے۔ تحقیقی ڈیٹا کے تجزیہ میں ڈیٹا کا بصری معائنہ، شماریاتی تجزیہ اور مشین سیکھنے کی تکنیک شامل ہو سکتی ہے۔

تحقیقی ڈیٹا کا تجزیہ وضاحتی تجزیات میں مددگار ہے کیونکہ یہ تنظیموں کو اعداد و شمار کے رجحانات اور نمونوں کی شناخت کرنے کی اجازت دیتا ہے جو فوری طور پر ظاہر نہ ہوں۔ تحقیقی ڈیٹا کے تجزیے کا استعمال کرتے ہوئے، تنظیمیں اپنے ڈیٹا کی گہری سمجھ حاصل کر سکتی ہیں اور حاصل کردہ بصیرت کی بنیاد پر باخبر فیصلے کر سکتی ہیں۔

B. پیشین گوئی تجزیات Predictive Analytics

پیشین گوئی تجزیات بزنس اینالیٹکس کی ایک قسم ہے جو تاریخی ڈیٹا کا تجزیہ کرنے اور مستقبل کے ایونٹس یا نتائج کی پیشین گوئی کرنے کے لیے شماریاتی اور مشین لرننگ الگورتھم کا استعمال کرتا ہے۔ یہ تنظیموں کو باخبر فیصلے کرنے کے لیے ڈیٹا پیٹرن کی شناخت اور استعمال کرنے میں مدد کرتا ہے۔

مثال: ریئل ٹائم بصیرت: ماضی کے ڈیٹا کا تجزیہ کر کے، پلیٹ فارم سائبر پیر 2025 کے دوران 2023 کے مقابلے میں فروخت میں 20% اضافے کی پیش گوئی کرتا ہے۔

Tools for Predictive Analytics:

- SAS Advanced Analytics
- IBM Watson
- Python (libraries like Scikit-learn)

پیشین گوئی تجزیات میں استعمال ہونے والی تین نارمل تکنیکیں(common techniques) یہ ہیں:

1. مشین لرننگ اور پریڈیکٹیو ماڈلنگ Machine Learning and Predictive Modelling

مشین لرننگ مصنوعی ذہانت ہے جو کمپیوٹرز کو واضح طور پر پروگرام کیے بغیر ڈیٹا سے سیکھنے کی اجازت دیتی ہے۔ پیشین گوئی تجزیات میں، مشین لرننگ الگورتھم مستقبل کے ایونٹس یا نتائج کی پیشین گوئی کرنے کے لیے پیشین گوئی کرنے والے ماڈل بناتے ہیں۔ ان ماڈل کو تاریخی ڈیٹا پر تربیت دی جاتی ہے اور وہ اس ڈیٹا کو پیٹرن کی شناخت اور پیشین گوئیاں کرنے کے لیے استعمال کرتے ہیں۔

پیشن گوئی ماڈلنگ میں مستقبل کے ایونٹس یا نتائج کی پیشن گوئی کرنے کے لیے شماریاتی ماڈلز کا استعمال شامل ہے۔ ان ماڈلز کو تاریخی ڈیٹا پر تربیت دی جاتی ہے اور وہ اس ڈیٹا کو پیٹرن کی شناخت اور پیشین گوئیاں کرنے کے لیے استعمال کرتے ہیں۔ پیشن گوئی ماڈلنگ مختلف ایپلی کیشنز میں مددگار ہے، بشمول دھوکہ دہی کا پتہ لگانے، خطرے کی تشخیص، اور کسٹمر کی تقسیم۔

2. پیشن گوئی اور وقت کی سیریز کا تجزیہ Forecasting and Time Series Analysis

پیشن گوئی ایک ایسی تکنیک ہے جو کسی متغیر کی ماضی کی اقدار کی بنیاد پر مستقبل کی اقدار کی پیشین گوئی کرنے کے لیے استعمال ہوتی ہے۔ ٹائم سیریز کا تجزیہ وقت پر منحصر ڈیٹا کا تجزیہ اور پیشین گوئی کرنے کے لیے ایک شماریاتی تکنیک ہے۔ پیشین گوئی کے تجزیات میں، پیشن گوئی اور ٹائم سیریز کے تجزیہ کا استعمال ڈیٹا میں مستقبل کے رجحانات اور نمونوں کی پیش گوئی کرنے کے لیے کیا جاتا ہے۔

پیشن گوئی اور ٹائم سیریز کا تجزیہ مختلف ایپلی کیشنز میں مددگار ہے، بشمول سیلز کی پیشن گوئی، انوینٹری مینجمنٹ، اور مالیاتی پیشن گوئی۔

3. ڈسزن ٹری اور رینڈم فوریسٹ (Decision Trees and Random Forests)

فیصلے کے درخت ایک الگورتھم ہیں جو قواعد کی بنیاد پر فیصلے کرنے کے لیے استعمال ہوتے ہیں۔ وہ پیشین گوئی کے تجزیات میں مددگار ہیں کیونکہ ان کا استعمال ڈیٹا کے نمونوں کی شناخت اور مستقبل کے ایونٹس یا نتائج کی پیش گوئی کرنے کے لیے کیا جا سکتا ہے۔ رینڈم فارسٹس ایک جوڑا سیکھنے کا الگورتھم ہے جسے بہتر بنانے کے لیے استعمال کیا جاتا ہے۔

پیش گوئی کرنے والے تجزیات کے ماڈلز کی درستگی۔ وہ زیادہ درست پیشین گوئی کرنے کے لیے متعدد ڈسزن ٹریوں کی پیشین گوئیوں کو یکجا کرتے ہیں۔ فیصلے کے درخت اور رینڈم فوریسٹ مختلف ایپلی کیشنز میں کارآمد ہیں، بشمول گاہک کی منتھنی کی پیشن گوئی، فراڈ کا پتہ لگانے، اور خطرے کی تشخیص۔

C. نسخہ جاتی تجزیات Prescriptive Analytics

نسخہ جاتی تجزیات بزنس اینالیٹکس کی ایک قسم ہے جو تنظیموں کو تجویز کردہ اقدامات یا فیصلے فراہم کرنے کے لیے تاریخی ڈیٹا، پیشین گوئی کے تجزیات اور اصلاح کی تکنیکوں کو یکجا کرتی ہے۔ یہ رکاوٹوں اور مقاصد کی بنیاد پر بہترین عمل کی نشاندہی کرنے کے لیے ریاضیاتی ماڈلز اور الگورتھم کا استعمال کرتا ہے۔

مثال: اعداد و شمار: ڈیلٹا ایئر لائنز نے نسخے کے تجزیات کا استعمال کرتے ہوئے 2022 میں ایندھن کے اخراجات میں $300 ملین سے زیادہ کی بچت کی۔

معلومات: نقلی ماڈلز کا استعمال کرتے ہوئے، کمپنی نے خراب موسم کی وجہ سے ہونے والی تاخیر سے نمٹنے کے لیے متعدد ہنگامی منصوبے بنائے۔

Tools for Prescriptive Analytics:

- IBM Decision Optimization
- KNIME
- AIMMS

نسخے کے تجزیات میں استعمال ہونے والی تین نارمل تکنیکیں (common techniques) یہ ہیں:

1. اصلاح اور نقل Optimization and Simulation

اصلاح کسی مسئلے کا بہترین حل تلاش کرنا ہے جو رکاوٹوں پر غور کرتے ہوئے کسی خاص مقصد کو زیادہ سے زیادہ یا کم کرتا ہے۔ اس کے برعکس،

نقل ایک حقیقی دنیا کی صورت حال کی نقل کرنے اور مختلف حالات میں اس کے رویے کا تجزیہ کرنے کے لیے ایک ماڈل کا استعمال کرتا ہے۔

نسخے کے تجزیات میں، اصلاح اور نقل کی تکنیکیں رکاوٹوں اور مقاصد کے تحت بہترین عمل کی نشاندہی کرتی ہیں۔ مثال کے طور پر، ایک ایئر لائن کمپنی ایندھن کی قیمت، موسم اور مسافروں کی طلب کی بنیاد پر اپنے طیاروں کے لیے بہترین راستے کا تعین کرنے کے لیے اصلاح کی تکنیک استعمال کر سکتی ہے۔

2. گول پروگرامنگ اور جبری اطمینان Goal Programming and Constraint Satisfaction

گول پروگرامنگ ایک ایسی تکنیک ہے جس کا استعمال بہترین طریقہ کار کی نشاندہی کرنے کے لیے کیا جاتا ہے جو بیک وقت متعدد مقاصد کو حاصل کرتا ہے۔ اس میں مقاصد کو ترجیح دینا اور ان کی اہمیت کی بنیاد پر انہیں وزن دینا شامل ہے۔ دوسری طرف، رکاوٹ اطمینان ایک ایسی تکنیک ہے جو عمل کے بہترین طریقہ کی نشاندہی کرنے کے لیے استعمال ہوتی ہے جو رکاوٹوں کے ایک سیٹ کو پورا کرتی ہے۔

نسخہ جات کے تجزیات میں، گول پروگرامنگ اور جبری اطمینان کی تکنیکوں کا استعمال بہترین عمل کی نشاندہی کرنے کے لیے کیا جاتا ہے جو متعدد مقاصد کو حاصل کرتا ہے اور رکاوٹوں کے ایک سیٹ کو پورا کرتا ہے۔ مثال کے طور پر، ایک خوردہ فروش اپنے اسٹورز کے لیے سیلز کے حجم، منافع کے مارجن، اور انوینٹری کی سطحوں کی بنیاد پر بہترین پروڈکٹ مکس کا تعین کرنے کے لیے گول پروگرامنگ کا استعمال کر سکتا ہے۔

3. سفارشی نظام اور تعاونی فلٹرنگ Recommendation Systems and Collaborative Filtering

سفارشی نظام اور اشتراکی فلٹرنگ وہ **تکنیک** ہیں جو صارفین کو ان کے ماضی کے رویے یا ترجیحات کی بنیاد پر ذاتی نوعیت کی سفارشات فراہم کرنے کے لیے استعمال کی جاتی ہیں۔ یہ تکنیک نارمل طور پر ای کامرس، سوشل میڈیا اور تفریحی پلیٹ فارم کے لئے استعمال ہوتی ہیں۔

نسخہ جات کے تجزیات میں، سفارشی نظام اور باہمی تعاون کے ساتھ فلٹرنگ کی تکنیکیں صارفین کے ماضی کے رویے یا ترجیحات کی بنیاد پر ذاتی سفارشات فراہم کرتی ہیں۔ مثال کے طور پر، ایک آن لائن خوردہ فروش اپنے صارفین کو ان کی خریداری کی تاریخ کی بنیاد پر مصنوعات کی سفارش کرنے کے لیے باہمی تعاون کے ساتھ فلٹرنگ کا استعمال کر سکتا ہے۔

D. تشخیصی تجزیات Diagnostic Analytics

تشخیصی تجزیات ماضی کے نتائج کے پیچھے کی وجوہات کو بے نقاب کرنے پر مرکوز ہیں۔ یہ پیٹرن اور بنیادی وجوہات کی شناخت کے لیے ڈیٹا مائننگ، ڈرل ڈاؤن تجزیہ، اور شماریاتی جانچ جیسی تکنیکوں کا استعمال کرتا ہے۔

مثال: ایک SaaS کمپنی صارفین کی قیمتوں کا تجزیہ کرتی ہے تاکہ یہ معلوم کیا جاسکے کہ صارفین کیوں چھوڑ رہے ہیں۔

اعداد و شمار: چرن میں 5% کی منافع میں 25%-95% اضافہ کر سکتی ہے (بائن اینڈ کمپنی)۔

ریئل ٹائم بصیرت: یہ دریافت کر کے کہ منحرف صارفین میں سے 40% کو آن بورڈنگ کے مسائل کا سامنا کرنا پڑا، کمپنی نے اپنے آن بورڈنگ کے عمل کو بہتر بنایا، جس سے چھ مہینوں میں 15% کی کمی واقع ہوئی۔

Tools for Diagnostic Analytics:

- SQL
- Tableau
- Apache Hadoop

1.5 بگ ڈیٹا(Big Data)

بگ ڈیٹا کی بحث میں جانے سے پہلے آیئے ڈیٹا کے معنی کو سمجھیں:

ڈیٹا کیا ہے Data؟

وہ مقداریں، حروف، یا علامتیں جن پر کمپیوٹر کے ذریعے آپریشن کیے جاتے ہیں، جنہیں برقی سگنل کی شکل میں ذخیرہ اور منتقل کیا جاسکتا ہے اور مقناطیسی، نظری، یا مکینیکل ریکارڈنگ میڈیا پر ریکارڈ کیا جاسکتا ہے۔

بگ ڈیٹا کیا ہے؟ What is Big Data

بگ ڈیٹا ڈیٹا کا ایک مجموعہ ہے جو حجم میں بہت بڑا ہے، لیکن وقت کے ساتھ ساتھ اور تیزی سے بڑھ رہا ہے۔ یہ اتنے بڑے سائز اور پیچیدگی کے ساتھ ایک ڈیٹا ہے کہ روایتی ڈیٹا مینجمنٹ ٹولز میں سے کوئی بھی اسے محفوظ یا مؤثر طریقے سے پروسیس نہیں کر سکتا۔ بگ ڈیٹا بھی ایک ڈیٹا ہے لیکن بڑے سائز کے ساتھ۔ بگ ڈیٹا سے مراد انتہائی بڑے اور پیچیدہ ڈیٹا سیٹس ہیں جن پر روایتی ٹولز مؤثر طریقے سے کارروائی نہیں کر سکتے۔ اس کی تعریف بمقابلہ: حجم (بڑے پیمانے پر)، رفتار (تیز رفتار نسل)، اور مختلف قسم (متنوع فارمیٹس جیسے متن، تصاویر اور ویڈیوز) سے کی گئی ہے۔ اضافی پہلوؤں میں سچائی (ڈیٹا کی وشوسنیتا) اور قدر (مفید بصیرت) شامل ہیں۔ بگ ڈیٹا سوشل میڈیا، IoT آلات، اور لین دین کے ریکارڈ جیسے ذرائع سے آتا ہے۔ یہ صحت کی دیکھ بھال، مالیات اور مارکیٹنگ جیسے شعبوں میں AI، پیشین گوئی کرنے والے تجزیات، اور حقیقی وقت میں فیصلہ سازی کو طاقت دیتا ہے۔ Hadoop، Spark، اور NoSQL ڈیٹا بیس جیسی ٹیکنالوجیز اسے ذخیرہ کرنے اور تجزیہ کرنے میں مدد کرتی ہیں۔ بالآخر، بگ ڈیٹا خام معلومات کو کاروبار اور تحقیق کے لیے قابل عمل ذہانت میں تبدیل کرتا ہے۔

بگ ڈیٹا کی مثالیں درج ذیل ہیں Examples of Big Data-

A. **اسٹاک مارکیٹ** Stock Market: نیویارک سٹاک ایکسچینج بگ ڈیٹا کی ایک مثال ہے جو روزانہ تقریباً ایک ٹیرا بائٹ نیا تجارتی ڈیٹا تیار کرتا ہے۔

Source:https://nickledanddimed.com/2024/06/09/holiday-season-and-stock-market-performance-examining-patterns-and-investor-behaviour/

B. **سوشل میڈیا** Social Media: اعداد و شمار سے پتہ چلتا ہے کہ سوشل میڈیا سائٹ Facebook کے ڈیٹا بیس میں ہر روز 500+ ٹیرا بائٹس نیا ڈیٹا داخل ہوتا ہے۔ یہ ڈیٹا بنیادی طور پر تصویر اور ویڈیو اپ لوڈ، پیغامات کے تبادلے، تبصرے ڈالنے وغیرہ کے لحاظ سے تیار کیا جاتا ہے۔

Source: https://medium.com/@f6aiimbz/social-media-is-a-disease-511794a9af3

1.6 بگ ڈیٹا کے اقسام (Types of Big Data)

بگ ڈیٹا کی اقسام درج ذیل ہیں:

I. ساختہ (Structured)

II. غیر ساختہ (Unstructured)

III. نیم ساختہ (Semi-Structured)

1. ساختہ (Structured)

کوئی بھی ڈیٹا جو فکسڈ فارمیٹ کی شکل میں محفوظ، رسائی اور پروسیس کیا جا سکتا ہے اسے 'سٹرکچرڈ' ڈیٹا کہا جاتا ہے۔ وقت گزرنے کے ساتھ ساتھ، کمپیوٹر سائنس میں ٹیلنٹ نے اس قسم کے ڈیٹا کے ساتھ کام کرنے کے لیے تکنیک تیار کرنے میں زیادہ کامیابی حاصل کی ہے (جہاں

فارمیٹ پہلے سے جانا جاتا ہے)اور اس سے قدر بھی حاصل کی ہے۔تاہم، آج کل، ہم ایسے مسائل کی پیشین گوئی کر رہے ہیں جب اس طرح کے ڈیٹا کا سائز بہت زیادہ بڑھ جاتا ہے۔

Examples Of Structured Data

An ‘Employee’ table in a database is an example of Structured Data

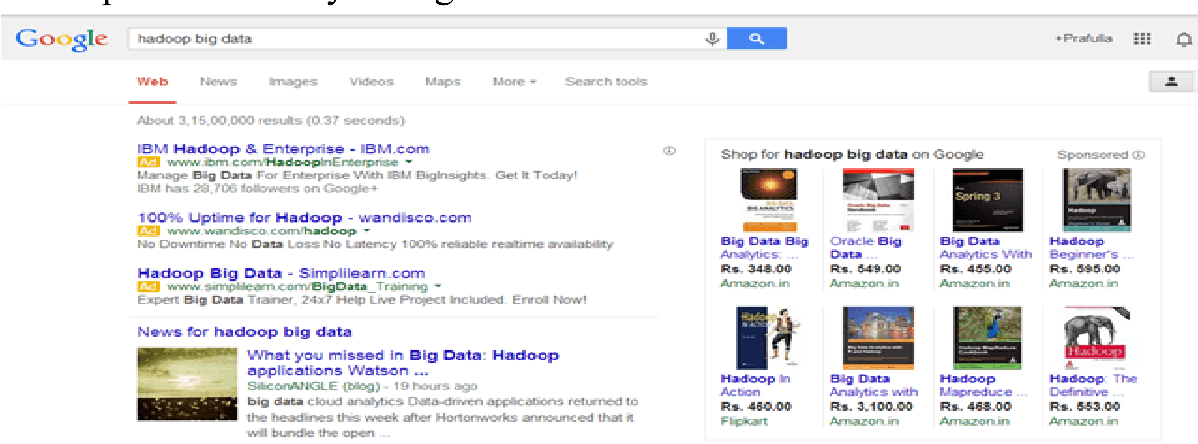
Employee_ID	Employee_Name	Gender	Department	Salary_In_lacs
2365	Rajesh Kulkarni	Male	Finance	650000
3398	Pratibha Joshi	Female	Admin	650000
7465	Shushil Roy	Male	Admin	500000
7500	Shubhojit Das	Male	Finance	500000
7699	Priya Sane	Female	Finance	550000

2. غیر ساختہ Unstructured

نامعلوم شکل یا ساخت کے ساتھ کسی بھی ڈیٹا کو غیر ساختہ ڈیٹا کے طور پر درجہ بندی کیا جاتا ہے۔ سائز بہت بڑا ہونے کے علاوہ، غیر ساختہ ڈیٹا اس سے قیمت حاصل کرنے کے لیے اس کی پروسیسنگ کے لئے متعدد چیلنجز کا سامنا کرتا ہے۔ غیر ساختہ ڈیٹا کی ایک عام مثال ایک متفاوت ڈیٹا سورس ہے جس میں سادہ ٹیکسٹ فائلوں، تصاویر، ویڈیوز وغیرہ کا مجموعہ ہوتا ہے۔ آج کل تنظیموں کے پاس ڈیٹا کا ذخیرہ موجود ہے لیکن بدقسمتی سے، وہ نہیں جانتے کہ اس سے قدر کیسے نکالی جائے کیونکہ یہ ڈیٹا اپنی خام شکل یا غیر ساختہ شکل میں ہے۔

Examples Of Un-Structured Data

The output returned by ‘Google Search’



3. نیم ساختہ (Semi-structured)

نیم ساختہ ڈیٹا ڈیٹا کی دونوں شکلوں پر مشتمل ہو سکتا ہے۔ ہم نیم ساختہ ڈیٹا کو بطور ساختی شکل میں دیکھ سکتے ہیں لیکن اس کی اصل میں

تعریف نہیں کی گئی ہے جیسے کہ RDBMS(Relational Database Management) میں ایک جدول کی تعریف۔ نیم ساختہ ڈیٹا کی مثال XML فائل میں ظاہر کردہ ڈیٹا ہے۔

Examples Of Semi-Structured Data

Personal data stored in an XML file:

<rec><name>Prashant Rao</name><sex>Male</sex><age>35</age></rec>

<rec><name>Seema R.</name><sex>Female</sex><age>41</age></rec>

<rec><name>Satish Mane</name><sex>Male</sex><age>29</age></rec>

<rec><name>Subrato Roy</name><sex>Male</sex><age>26</age></rec>

<rec><name>Jeremiah J.</name><sex>Male</sex><age>35</age></rec>

1.7 بگ ڈیٹا کی خصوصیات(Features of Big Data)

بگ ڈیٹا کی درجہ ذیل خصوصیات ہیں

1. حجم Volume

بگ ڈیٹا کا نام خود ایک سائز سے متعلق ہے جو بہت بڑا ہے۔ ڈیٹا کا سائز ڈیٹا سے باہر کی قدر کا تعین کرنے میں بہت اہم کردار ادا کرتا ہے۔ اس کے علاوہ، آیا کسی خاص ڈیٹا کو اصل میں بگ ڈیٹا کے طور پر سمجھا جاسکتا ہے یا نہیں، ڈیٹا کے حجم پر منحصر ہے۔ لہذا، 'حجم' ایک خصوصیت ہے جسے بگ ڈیٹا سلوشنز سے نمٹنے کے دوران غور کرنے کی ضرورت ہے۔

2. مختلف قسم Variety

بگ ڈیٹا کا اگلا پہلو اس کی ورائٹی ہے۔ مختلف قسم سے مراد متضاد ذرائع اور ڈیٹا کی نوعیت ہے، دونوں ساختہ اور غیر ساختہ۔ پہلے دنوں کے دوران، اسپریڈ شیٹس اور ڈیٹا بیس ہی ڈیٹا کے واحد ذرائع تھے جن پر زیادہ تر ایپلی کیشنز کے ذریعہ غور کیا جاتا تھا۔ آج کل تجزیہ ایپلی کیشنز میں ای میلز، تصاویر، ویڈیوز، مانیٹرنگ ڈیوائسز، پی ڈی ایف، آڈیو وغیرہ کی شکل میں ڈیٹا کو بھی زیر غور لایا جارہا ہے۔ غیر ساختہ ڈیٹا کی یہ قسم ڈیٹا کو ذخیرہ کرنے، کان کنی اور تجزیہ کرنے کے لیے کچھ مسائل پیدا کرتی ہے۔

3. رفتار (Velocity)

اصطلاح 'رفتار' ڈیٹا کی تخلیق کی رفتار سے مراد ہے۔ مطالبات کو پورا کرنے کے لیے ڈیٹا کو کتنی تیزی سے تیار کیا جاتا ہے اور اس پر عملدرآمد کیا جاتا ہے، ڈیٹا میں حقیقی صلاحیت کا تعین کرتا ہے۔

Big Data Velocity اس رفتار سے متعلق ہے جس پر کاروباری عمل، ایپلیکیشن لاگ، نیٹ ورکس، اور سوشل میڈیا سائٹس، سینسرز، موبائل ڈیوائسز وغیرہ جیسے ذرائع سے ڈیٹا کا بہاؤ ہوتا ہے۔ ڈیٹا کا بہاؤ بڑے پیمانے پر اور مسلسل ہوتا ہے۔

4. تغیر پذیری (Variability)

اس سے مراد وہ عدم مطابقت ہے جو بعض اوقات ڈیٹا کے ذریعے دکھائی جا سکتی ہے، اس طرح ڈیٹا کو مؤثر طریقے سے سنبھالنے اور اس کا انتظام کرنے کے عمل میں رکاوٹ پیدا ہوتی ہے۔

1.8 سوشل میڈیا اینالیٹکس (Social Media Analytics)

آج کے ڈیجیٹل دور میں، سوشل میڈیا ہر قسم کے کاروبار کے لیے ایک ناگزیر مواصلاتی چینل بن گیا ہے۔ لیکن صرف سوشل میڈیا کی موجودگی کافی نہیں ہے۔ اس کی صلاحیت کو صحیح معنوں میں فائدہ اٹھانے کے لیے، آپ کو سوشل میڈیا کے تجزیات کی دنیا میں جانے کی ضرورت ہے۔

سوشل میڈیا نے تبدیل کر دیا ہے کہ کاروبار کس طرح گاہکوں کے ساتھ بات چیت کرتے ہیں، برانڈ بیداری پیدا کرتے ہیں، اور سیل کو بڑھاتے ہیں۔ ان پلیٹ فارم سے زیادہ سے زیادہ فائدہ اٹھانے کے لیے، فیصلہ سازی میں سوشل میڈیا کے منظر نامے اور سوشل میڈیا کے تجزیات کے کردار کو سمجھنا ضروری ہے۔

سوشل میڈیا سے مراد وہ آن لائن پلیٹ فارم ہیں جہاں صارف مواد تخلیق، اشتراک کے ساتھ مشغول ہوتے ہیں۔ یہ افراد کاروباری اداروں کو عالمی سطح پر مصنوعات یا خدمات کو مربوط کرنے، بات چیت کرنے اور مارکیٹ کرنے کے قابل بناتے ہیں۔

سوشل میڈیا کی اہم خصوصیات: (Features)

- User-generated content (e.g., posts, images, videos, and blogs).
- Real-time interaction (e.g., comments, likes, shares, and messages).
- Community engagement (e.g., forums, groups, and discussions).
- Targeted marketing (e.g., personalized ads and promotions).

مثال: ایک ٹیک کمپنی صنعت کی معلومات شائع کرنے، مہارت کو راغب کرنے اور پیشہ ور افراد کے ساتھ نیٹ ورک کرنے کے لیے LinkedIn کا استعمال کرتی ہے، جبکہ YouTube کا استعمال مصنوعات کی خصوصیات کو ظاہر کرنے کے لیے کرتی ہے۔

سوشل میڈیا اینالیٹکس کیا ہے؟ (Social Media Analytics)

سوشل میڈیا اینالیٹکس سوشل میڈیا پلیٹ فارم سے ڈیٹا اکٹھا کرنے، تجزیہ کرنے اور اس کی تشریح کرنے کا عمل ہے تاکہ آپ کے سامعین، برانڈ کے تاثرات اور مارکیٹنگ کی مہمات کے بارے میں قیمتی معلومات حاصل کر سکیں۔ یہ آپ کے برانڈ کے ارد گرد آن لائن بات چیت میں ایک ونڈو رکھنے کی طرح ہے، آپ کو اپنی سوشل میڈیا حکمت عملی کو بہتر بنانے کے لیے ڈیٹا پر مبنی فیصلے کرنے کی اجازت دیتا ہے۔

سوشل میڈیا اینالیٹکس (SMA) تمام سائز کے کاروبار میں کارکردگی کو ٹریک کرنے، سامعین کے رویے کو سمجھنے، اور مارکیٹنگ کی حکمت عملیوں کو بہتر بنانے میں ان کی مدد کر کے ایک اہم کردار ادا کرتا ہے۔ تاہم، چھوٹی اور بڑی تنظیموں کا SMA استعمال کرنے کا طریقہ وسائل، مقاصد اور آپریشنز کے پیمانے کی بنیاد پر مختلف ہوتا ہے۔

1.9 مختلف سوشل میڈیا پلیٹ فارم پر سوشل میڈیا اینالیٹکس کا اطلاق

(Application of Social Media Analytics on different Social Media Platforms)

سوشل میڈیا اینالیٹکس میں قابل عمل معلومات کو نکالنے کے لیے سوشل پلیٹ فارم سے ڈیٹا اکٹھا کرنا، پروسیسنگ کرنا اور اس کی ترجمانی کرنا شامل ہے۔ Facebook پر، یہ حصص اور ردعمل جیسے منگنی کے میٹرکس کو ٹریک کرنے میں مدد کرتا ہے، جس سے کاروباری اداروں کو اشتہاراتی ہدف بندی اور مواد کی حکمت عملیوں کو بہتر بنانے میں مدد ملتی ہے۔ ٹویٹر (X) تجزیات ہیش ٹیگ کے رجحانات، جذبات کے تجزیہ، اور اثر انگیز اثرات پر توجہ مرکوز کرتا ہے، جو برانڈ کی نگرانی اور بحران کے انتظام کے لیے مفید ہے۔

انسٹاگرام اور TikTok اثر انگیز مہمات اور وائرل مواد کو بہتر بنانے کے لیے بصری تجزیات، جیسے ویڈیو ویوز اور فالوور ڈیموگرافکس کا فائدہ اٹھاتے ہیں۔ LinkedIn B2B لیڈ جنریشن اور بھرتی کی حکمت عملیوں کے لیے پیشہ ورانہ تنوع کا ڈیٹا استعمال کرتا ہے۔ YouTube تجزیات ویڈیو SEO کو بہتر بنانے کے لیے دیکھنے کے وقت اور سامعین کی برقراری کا اندازہ لگاتا ہے۔ دریں اثنا، Reddit اور Quora اپ ووٹ پیٹرن اور موضوع کی مقبولیت کے ذریعے کمیونٹی سے چلنے والی بصیرت فراہم کرتے ہیں، جو کہ مخصوص مارکیٹنگ میں مدد کرتے ہیں۔ مجموعی طور پر، سوشل میڈیا اینالیٹکس فیصلہ سازی، کسٹمر کی مصروفیت، اور تمام پلیٹ فارم پر ROI کو بڑھاتا ہے۔ ہر سوشل میڈیا پلیٹ فارم میں کارکردگی کو ٹریک کرنے کے لیے منفرد تجزیاتی ٹولز اور خصوصیات ہیں۔ یہاں مگر ہمیں جاننا ضروری ہے کہ مختلف پلیٹ فارم پر SMA کا اطلاق کیسے کیا جاتا ہے:

A. Facebook Analytics

Key Metrics:

- **Page Likes & Follows**: Measures audience growth.
- **Post Engagement**: Tracks reactions, comments, and shares.
- **Reach & Impressions**: Shows how many users saw a post.
- **Click-Through Rate (CTR)**: Measures link clicks.

مثال: کپڑے کاایک برانڈ فیس بک کی اشتہاری مہم چلاتا ہے۔Facebook Analytics کااستعمال کرتے ہوئے، وہ دیکھتے ہیں کہ ویڈیو اشتہارات میں جامد تصویری اشتہارات کے مقابلے میں 30% زیادہ کلک کرنے کی شرح ہوتی ہے، جس کی وجہ سے وہ ویڈیو مارکیٹنگ پر زیادہ توجہ مرکوز کرتے ہیں۔

B. Instagram Insights

Key Metrics:

- **Engagement Rate**: Likes, shares, comments, and saves.
- **Story Views & Replies**: Tracks how many users interact with Instagram Stories.
- **Hashtag Performance**: Measures reach generated by specific hashtags.

مثال: ایک فٹنس متاثر کنندہ انسٹا گرام معلومات اور نوٹس کو ٹریک کرتا ہے کہ ریلز کو بااصول پوسٹس کے مقابلے میں 5 گنا زیادہ مشغولیت ملتی ہے۔ وہ اپنی رسائی کو بڑھانے کے لیے مزید مختصر ورزش کی ویڈیوز بنانا شروع کر دیتے ہیں۔

C. Twitter (X) Analytics

Key Metrics:

- **Tweet Impressions**: Number of times a tweet is seen.
- **Engagement Rate**: Likes, retweets, and replies.
- **Follower Growth**: Tracks increase or decrease in followers.
- **Hashtag Performance**: Identifies trending hashtags.

مثال: ایک ٹیک اسٹارٹ اپ ٹویٹر تجزیات کو ٹریک کرتا ہے اور اسے پتہ چلتا ہے کہ پولز والی ٹویٹ کو نارمل ٹیکسٹ پر مبنی ٹویٹ کے مقابلے میں 40% زیادہ مشغولیت ملتی ہے۔ وہ بات چیت کو فروغ دینے کے لیے پول کا استعمال شروع کر دیتے ہیں۔

Likert Scale

Key Metrics:

- **Profile Views & Followers**: Measures professional reach.
- **Post Engagement**: Tracks like, comments, and shares.
- **Click-Through Rate (CTR):** Shows how many users click on shared links.

مثال: ایک B2B سافٹ ویئر کمپنی نے نوٹس کیا ہے کہ کیس اسٹڈیز پر مشتمل LinkedIn پوسٹس ریگولر پوسٹس کے مقابلے میں 60% زیادہ لیڈز پیدا کرتی ہیں۔ وہ ممکنہ گاہکوں کو اپنی طرف متوجہ کرنے کے لیے مزید کسٹمر کی کامیابی کی کہانیاں بانٹنا شروع کر دیتے ہیں۔

D. YouTube Analytics

Key Metrics:

- **Watch Time & Retention**: Measures how long viewers watch a video.
- **Engagement Rate**: Likes, comments, and shares.
- **Traffic Sources**: Shows where viewers come from (search, social media, external websites).

مثال: ایک گیمنگ YouTuber YouTube کے تجزیات کو چیک کرتا ہے اور اسے پتہ چلتا ہے کہ 10 منٹ سے زیادہ طویل ویڈیوز کو دیکھنے کا وقت زیادہ ہوتا ہے۔ وہ سامعین کی برقراری کو بڑھانے کے لیے طویل، مزید تفصیلی ویڈیوز بنانا شروع کر دیتے ہیں۔

مختصراً، سوشل میڈیا اینالیٹکس (SMA) چھوٹی اور بڑی دونوں تنظیموں کے لیے اہم ہے۔ چھوٹے کاروبار اسے لاگت سے موثر مارکیٹنگ کے لیے استعمال کرتے ہیں، جب کہ بڑی کمپنیاں برانڈ مینجمنٹ اور ڈیٹا پر مبنی فیصلوں کے لیے اس کا فائدہ اٹھاتی ہیں۔ مختلف پلیٹ فارم پر SMA لاگو کر کے، کاروباری مواد کو بہتر بنا سکتے ہیں، کسٹمر کی مصروفیت کو بہتر بنا سکتے ہیں، اور اپنے ROI میں اضافہ کر سکتے ہیں۔

1.10 اکتسابی نتائج (Learning Outcomes)

طلباء نے کاروباری تجزیات کے بنیادی تصورات اور جدید فیصلہ سازی میں اس کی بڑھتی ہوئی اہمیت کے بارے میں واضح سمجھ حاصل کر لی ہے۔ انہوں نے مختلف قسم کے کاروباری تجزیات کے بارے میں سیکھا ہے—وضاحتی، پیشین گوئی، اور نسخہ—اور کس طرح ہر قسم تنظیموں کو ڈیٹا کا تجزیہ کرنے اور کارکردگی کو بہتر بنانے میں مدد کرتی ہے۔ طلباء اب بگ ڈیٹا کے تصور، اس کی اقسام، اور اس کی اہم خصوصیات جیسے حجم، رفتار، اور مختلف قسم کی وضاحت کر سکتے ہیں۔ وہ سوشل میڈیا تجزیات کے کردار کو بھی سمجھتے ہیں اور یہ بھی سمجھتے ہیں کہ کس طرح Facebook، Twitter، اور LinkedIn جیسے پلیٹ فارمز سے ڈیٹا کو صارفین کے رویے اور مارکیٹ کے رجحانات کو سمجھنے کے لیے استعمال کیا جا سکتا ہے۔ مجموعی طور پر، سیکھنے والوں نے تجزیاتی ٹولز اور ڈیٹا پر مبنی بصیرت کو عملی کاروباری حکمت عملیوں کے ساتھ جوڑنے کی صلاحیت تیار کی ہے۔

1.11 کلیدی الفاظ (Keyword)

1) بزنس اینالیٹکس - باخبر کاروباری فیصلوں کی حمایت کے لیے ڈیٹا، شماریاتی ٹولز، اور تکنیک استعمال کرنے کا عمل۔
2) وضاحتی تجزیات - ماضی کی کارکردگی کے رجحانات اور نمونوں کو سمجھنے کے لیے تاریخی ڈیٹا کا تجزیہ کرتا ہے۔

3) پیشین گوئی کے تجزیات - ماضی کے اعداد و شمار کی بنیاد پر مستقبل کے نتائج کی پیشن گوئی کرنے کے لیے شماریاتی ماڈلز اور مشین لرننگ کا استعمال کرتا ہے۔

4) بگ ڈیٹا - انتہائی بڑے اور پیچیدہ ڈیٹا سیٹس جنہیں ذخیرہ کرنے، پروسیس کرنے اور تجزیہ کرنے کے لیے جدید ٹولز کی ضرورت ہوتی ہے۔

5) سوشل میڈیا اینالیٹکس - کاروباری حکمت عملی کی رہنمائی کے لیے سوشل میڈیا پلیٹ فارمز سے ڈیٹا اکٹھا کرنے اور تجزیہ کرنے کی مشق۔

1.12 نمونہ امتحانی سوالات (Model Examination Questions)

درجہ ذیل سوالات کے مختصر ترین جوابات دیں

1) بزنس اینالیٹکس کیا ہے؟
2) بزنس اینالیٹکس کی کسی بھی دو قسموں کے نام بتائیں۔
3) بگ ڈیٹا کو ایک لائن میں بیان کریں۔
4) Predictive Analytics کا مقصد کیا ہے؟
5) سٹرکچرڈ ڈیٹا سورس کی ایک مثال دیں۔

مختصر جوابات کے حامل سوالات (Short Answer Type Questions)

1) جدید کاروباری فیصلہ سازی میں بزنس اینالیٹکس کی اہمیت کی وضاحت کریں۔
2) وضاحتی، پیشین گوئی، اور نسخے کے تجزیات کے درمیان فرق کریں۔
3) بگ ڈیٹا کی خصوصیات کیا ہیں؟
4) کاروباری حکمت عملی میں سوشل میڈیا اینالیٹکس کے کردار کو مختصراً بیان کریں۔
5) وضاحتی تجزیات میں استعمال ہونے والے ٹولز پر مثالوں کے ساتھ بحث کریں۔

طویل جوابات کے حامل سوالات (Long Answer Type Questions)

1) بزنس اینالیٹکس کی وضاحت کریں اور مناسب مثالوں کے ساتھ اس کی اقسام کی وضاحت کریں۔
2) بگ ڈیٹا کیا ہے؟ اس کی اقسام اور حقیقی دنیا کی ایپلی کیشنز پر بحث کریں۔
3) سوشل میڈیا اینالیٹکس کی اہمیت اور مختلف پلیٹ فارمز سے کاروباری بصیرت کے لیے کس طرح لاگو کرتے ہیں اس کی وضاحت کریں۔

اکائی 2۔ ڈیٹا کا فریم ورک

(Framework of Data)

اکائی کے اجزاء		Structure
2.0	تمہید	Introduction
2.1	مقاصد	Objectives
2.2	مواد	Data
2.3	ڈیٹا کی پیشکش	Data Representation
2.4	ڈیٹا کے ذرائع	Data Sources
2.5	ٹیبلیشن اور ڈیٹا کی درجہ بندی	Tabulation and classification of data
2.6	ڈیٹا سائنس	Data Science Statistics
2.7	شماریاتی تجزیہ	Descriptive Data Analytics
2.8	اکتسابی نتائج	
2.9	نمونہ امتحانی سوالات	Model Examination Questions
2.10	تجویز کردہ اکتسابی مواد	Suggested Learning Resources

2.0 تمہید (Introduction)

ڈیٹا کے فریم ورک کی مدد سے یہ سمجھنا کہ ڈیٹا کو کیسے جمع کیا جاتا ہے، ترتیب دیا جاتا ہے، پروسیس کیا جاتا ہے اور استعمال کیا جاتا ہے۔ یہ ان تصورات، فریم ورک اور ماڈل کا خاکہ پیش کرتا ہے جو ڈیٹا کو غیر پروسیس شدہ ان پٹ (Input) سے معلومات بھرے علم میں تبدیل کرنے کی ہدایت کرتے ہیں۔ فریم ورک ڈیٹا مینجمنٹ، تجزیہ اور اسٹوریج کے معیارات کی وضاحت کرکے مستقل مزاجی، انحصار اور افادیت کی ضمانت دیتا ہے۔ اس طرح، یہ عصری معلوماتی نظام اور فیصلہ سازی کا ایک لازمی جزو ہے۔

2.1 مقاصد(Objectives)

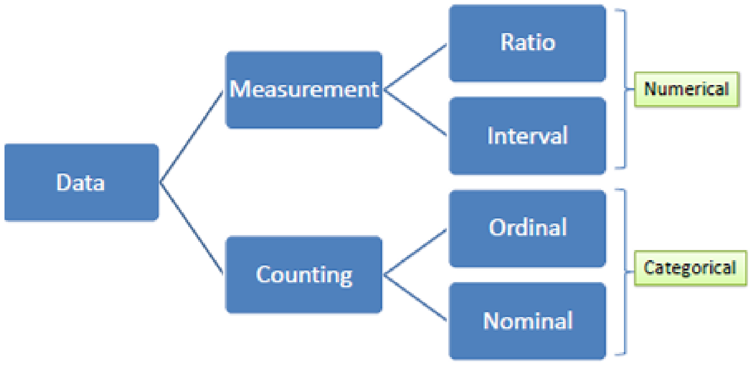
معلوماتی نظام کے اندر ڈیٹا کے بنیادی ڈھانچے، تنظیم اور انتظامیہ کو سمجھنے میں طلباء کی مدد کرنا "ڈیٹا کے فریم ورک" باب کا سیکھنے کا ہدف ہے۔اس باب کا مقصد قارئین کو علمی فیصلے کرنے کے لیے ڈیٹا کو جمع کرنے، درجہ بندی کرنے، ذخیرہ کرنے، عمل کرنے اور جانچنے کے لیے استعمال کیے جانے والے طریقہ کار کے بارے میں مکمل معلومات فراہم کرنا ہے۔اس کا بنیادی مقصد ڈیٹا کی اقسام، ڈیٹا ماڈلز، ڈیٹا کی درجہ بندی، اور کاروبار میں ڈیٹا کے بہاؤ جیسے اہم آئیڈیا متعارف کروانا ہے۔اس باب کے اختتام تک، طلباء کو مختلف ڈیٹا فریم ورک کو پہچاننے کے قابل ہونے کی ضرورت ہے، یہ سمجھنے کی ضرورت ہے کہ وہ تکنیکی اور تجارتی سیاق وسباق میں کیسے کام کرتے ہیں، اور مؤثر ڈیٹا مینجمنٹ اور تجزیہ کے لیے صحیح تکنیکوں کا استعمال کریں۔

2.2 مواد(Data)

مواد نارمل طور پر پیمائش) عددی (یا گنتی) دوٹوک (کے نتائج کی معلومات ہے۔ متغیرات مواد کے لیے پلیس ہولڈر کے طور پر کام کرتے ہیں۔ متغیرات کی دو قسمیں ہیں، عددی اور دوٹوک۔

عددی یا مسلسل متغیر وہ ہے جو کسی بھی قدر کو محدود یا لا محدود وقفہ کے اندر قبول کر سکتا ہے (مثلاً، اونچائی، وزن، درجہ حرارت، خون میں گلوکوز)۔ عددی اعداد و شمار کی دو قسمیں ہیں، وقفہ اور تناسب۔ وقفہ کے پیمانے پر ڈیٹا کو شامل اور گھٹایا جا سکتا ہے لیکن معنی خیز طور پر ضرب یا تقسیم نہیں کیا جا سکتا کیونکہ کوئی حقیقی صفر نہیں ہے۔ مثال کے طور پر، ہم یہ نہیں کہہ سکتے کہ ایک دن دوسرے دن سے دوگنا گرم ہے۔ دوسری طرف، تناسب کے پیمانے پر ڈیٹا میں درست صفر ہوتا ہے اور اسے شامل، گھٹایا، ضرب یا تقسیم کیا جا سکتا ہے (مثال کے طور پر، وزن)۔

ایک زمرہ یا مجرد متغیر وہ ہوتا ہے جو دو یا دو سے زیادہ اقدار (زمرے) کو قبول کر سکتا ہے۔ واضح اعداد و شمار کی دو قسمیں ہیں، برائے نام اور آرڈینل۔ برائے نام ڈیٹا میں زمروں میں کوئی اندرونی ترتیب نہیں ہوتی ہے۔ مثال کے طور پر، " جنس " دو زمروں کے ساتھ، مرد اور عورت۔ اس کے برعکس، آرڈینل ڈیٹا میں زمروں میں ایک اندرونی ترتیب ہوتی ہے۔ مثال کے طور پر، تین منظم زمروں (کم، درمیانے اور اعلی) کے ساتھ " توانائی کی سطح "۔



Source: https://ersoykubraa.medium.com/scale-types-of-data-measurement-4052673ad49

مواد کیا ہے؟

اعداد و شمار کو اعداد کے طور پر بیان کیا جاتا ہے جو حقیقی دنیا سے پیائش کی نمائندگی کرتے ہیں۔ ڈیٹم ایک واحد پیائش ہے۔ ہم اکثر خبریں پڑھتے ہیں جیسے باڑ میں 20 سینٹی میٹر مسلسل بارش یا بانسواڑہ میں 24 گھنٹوں میں 35 سینٹی میٹر بارش یا معلومات جیسے کہ نئی دہلی - ممبئی کا فاصلہ کوٹا اور اٹارسی کے راستے - وڈودرا کے راستے 1385 کلو میٹر ہے - منماڈ ٹرین کے ذریعے 1542 کلو میٹر ہے۔ اس عددی معلومات کو مواد کہا جاتا ہے۔ یہ آسانی سے محسوس کیا جا سکتا ہے کہ آج دنیا بھر میں مواد کی بڑی مقدار دستیاب ہے۔ تاہم، بعض اوقات، ان اعداد و شمار سے منطقی نتائج اخذ کرنا مشکل ہو جاتا ہے اگر وہ خام شکل میں ہوں۔ لہذا، اس بات کو یقینی بنانا ضروری ہے کہ ماپا معلومات الگورتھم سے اخذ کی گئی ہے اور/یا منطقی طور پر اخذ کی گئی ہے اور/یا متعدد مواد سے شماریاتی طور پر شمار کی گئی ہے۔ یہ عام طور پر بہت سے متغیرات سے متاثر ہوتے ہیں جن کی مقداری لحاظ سے بہترین وضاحت کی جا سکتی ہے۔ ان متغیرات کا شماریاتی تجزیہ آج ایک ضرورت بن گیا ہے۔ مثال کے طور پر، کسی علاقے کے کراپنگ پیٹرن کا مطالعہ کرنے کے لیے، فصل کے رقبے، فصل کی پیداوار اور پیداوار، آبپاشی کے علاقے، بارش کی مقدار اور کھاد، کیڑے مار ادویات، کیڑے مار ادویات وغیرہ کے استعمال کے بارے میں اعداد و شمار کی معلومات کا ہونا ضروری ہے۔ شہر کی ترقی کا مطالعہ کرنے کی ضرورت ہے۔ اس طرح، جغرافیائی تجزیہ میں ڈیٹا ایک اہم کردار ادا کرتا ہے۔

مواد کی ضرورت

جغرافیہ کے مطالعہ میں نقشے اہم اوزار ہیں۔ اس کے علاوہ، مظاہر کی تقسیم اور نمو کو بھی اعداد و شمار کے ذریعے ٹیبلر شکل میں بیان کیا گیا ہے۔ ہم جانتے ہیں کہ زمین کی سطح پر بہت سے مظاہر کے درمیان ایک باہمی تعلق موجود ہے۔ یہ تنورملات بہت سے متغیرات سے متاثر ہوتے ہیں جو ہو سکتے ہیں۔

2.3 ڈیٹا کی پیشکش (Data Representation)

آپ نے ایک ایسے شخص کی کہانی سنی ہوگی جو اپنی بیوی اور پانچ سال کے بچے کے ساتھ سفر کر رہا تھا۔ راستے میں اسے ایک دریا عبور کرنا پڑا۔ سب سے پہلے، اس نے دریا کی گہرائی چار پوائنٹس پر 0.6، 0.8، 0.9 اور 1.5 میٹر معلوم کی۔ اس نے اوسط گہرائی کا حساب لگایا 0.95 میٹر۔ اس کے بچے کا قد 1 میٹر تھا۔ چنانچہ وہ انہیں دریا پار کرنے کے لیے لے گیا اور اس کا بچہ دریا میں ڈوب گیا۔ دوسرے کنارے پر بیٹھا سوچ رہا تھا" :لیکہ جو کھا تھا، بچہ دوبہ کیہ؟) "جب اوسط گہرائی ہر ایک کی پہنچ میں تھی تو بچہ کیوں ڈوب گیا؟) اسے شماریاتی غلط فہمی کہا جاتا ہے، جو آپ کو حقیقی صورتحال سے منحرف کر سکتا ہے۔ لہٰذا، حقائق اور اعداد و شمار کو جاننے کے لیے ڈیٹا اکٹھا کرنا ضروری ہے، لیکن اتنا ہی اہم ڈیٹا کی پیش کش ہے۔ آج، تجزیہ پیشکش اور نتیجہ اخذ کرنے میں شماریاتی طریقوں کا استعمال جغرافیہ سمیت تقریباً تمام شعبوں میں اہم کردار ادا کرتا ہے، جو ڈیٹا کا استعمال کرتے ہیں۔ لہذا، اس بات کا اندازہ لگایا جا سکتا ہے کہ کسی رجحان کا ارتکاز، جیسے، آبادی، جنگل یا نقل و حمل یا مواصلات کا نیٹ ورک نہ صرف جگہ اور وقت کے لحاظ سے مختلف ہوتا ہے بلکہ ڈیٹا کا استعمال کرتے ہوئے آسانی سے وضاحت بھی کی جا سکتی ہے۔ دوسرے لفظوں میں، آپ کہہ سکتے ہیں کہ متغیرات کے درمیان تعلق کی وضاحت میں کوالٹیٹیو وضاحت سے مقداری تجزیہ کی طرف تبدیلی آئی ہے۔ لہذا، تجزیاتی ٹولز اور تکنیک ان دنوں مطالعہ کو زیادہ منطقی بنانے اور درست نتیجہ اخذ کرنے کے لیے زیادہ اہم ہو گئے ہیں۔ درست مقداری تکنیک کا استعمال اعداد و شمار کو جمع کرنے اور مرتب کرنے کے آغاز سے لے کر اس کے ٹیبلیشن، تنظیم، ترتیب اور تجزیہ تک نتیجہ اخذ کرنے تک استعمال کیا جاتا ہے۔

2.4 ڈیٹا کے ذرائع (Sources of Data)

مندرجہ ذیل طریقوں سے ڈیٹا اکٹھا کیا جاتا ہے۔

1. بنیادی ذرائع 2. ثانوی ذرائع۔ وہ ڈیٹا جو کسی فرد یا افراد کے گروپ، ادارے/تنظیموں کے ذریعے پہلی بار جمع کیا جاتا ہے، ڈیٹا کے بنیادی ذرائع کہلاتے ہیں۔ دوسری طرف، کسی بھی شائع یا غیر مطبوعہ ذرائع سے جمع کردہ ڈیٹا کو ثانوی ذرائع کہا جاتا ہے۔ تصویر 1.1 ڈیٹا اکٹھا کرنے کے مختلف طریقے دکھاتا ہے۔

پرائمری ڈیٹا کے ذرائع

1. ذاتی مشاہدات

اس سے مراد کسی فرد یا افراد کے گروپ کے ذریعے میدان میں براہ راست مشاہدات کے ذریعے معلومات جمع کرنا ہے۔ فیلڈ سروے کے ذریعے امدادی خصوصیات، نکاسی آب کے نمونوں، مٹی اور قدرتی پودوں کی اقسام کے ساتھ ساتھ آبادی کی ساخت، جنسی تناسب، خواندگی،

نقل و حمل اور مواصلات کے ذرائع، شہری اور دیہی بستیوں وغیرہ کے بارے میں معلومات اکٹھی کی جاتی ہیں۔تاہم، ذاتی مشاہدات کو انجام دینے میں،اس میں شامل افراد کو موضوع کی نظریاتی معلومات اور غیر جانبدارانہ تشخیص کے لیے سائنسی رویہ ہونا چاہیے۔

2۔انٹرویو

اس طریقہ کار میں، محقق مکالموں اور بات چیت کے ذریعے جواب دہندہ سے براہ راست معلومات حاصل کرتا ہے۔تاہم،انٹرویو لینے والے کو علاقے کے لوگوں کے ساتھ انٹرویو کرتے وقت درج ذیل احتیاطی تدابیر اختیار کرنی ہوں گی۔

(i) ان اشیاء کی ایک قطعی فہرست تیار کی جائے جن کے بارے میں انٹرویو لینے والے افراد سے معلومات اکٹھی کی جانی ہیں۔

(ii) انٹرویو کے انعقاد میں شامل افراد کو سروے کے مقصد کے بارے میں واضح ہونا چاہیے۔

(iii) کوئی بھی حساس سوال پوچھنے سے پہلے جواب دہندگان کو اعتماد میں لیا جائے اور اسے یقین دلایا جائے کہ رازداری برقرار رکھی جائے گی۔

(iv) ایک سازگار ماحول بنایا جائے تاکہ مدعی بغیر کسی ہچکچاہٹ کے حقائق بیان کر سکے۔

(v)سوالات کی زبان سادہ اور شائستہ ہونی چاہیے تاکہ جواب دہندگان حوصلہ افزائی کریں اور مانگی گئی معلومات دینے کے لیے آسانی سے راضی ہوں۔

(vi) ایسا کوئی سوال پوچھنے سے گریز کریں جس سے جواب دہندہ کی عزت نفس یا مذہبی جذبات مجروح ہوں۔

(vii)انٹرویو کے اختتام پر، جواب دہندہ سے پوچھیں کہ وہ کون سی اضافی معلومات فراہم کر سکتا ہے،اس کے علاوہ جو اس نے پہلے ہی فراہم کی ہے۔

(viii)آپ کے لیے اپنا قیمتی وقت نکالنے کے لیے آپ کا شکریہ اور شکر گزار ہوں۔

سوالنامہ

اس طریقہ میں، سادہ سوالات اور ان کے ممکنہ جوابات ایک سادہ کاغذ پر لکھے جاتے ہیں اور جواب دہندگان کو دیئے گئے انتخاب میں سے ممکنہ جوابات پر ٹک مارک کرنا ہوتا ہے۔ بعض اوقات، منظم سوالات کا ایک سیٹ لکھا جاتا ہے اور سوالنامے میں کافی جگہ فراہم کی جاتی ہے جہاں جواب دہندہ اپنی رائے لکھتے ہیں۔سوالنامے میں سروے کے مقاصد کا واضح طور پر ذکر ہونا چاہیے۔یہ طریقہ کسی بڑے علاقے کا سروے کرنے میں مفید ہے۔یہاں تک کہ سوالنامہ بھی دور دراز مقامات پر بھیجا جا سکتا ہے۔طریقہ کار کی حد یہ ہے کہ مطلوبہ معلومات فراہم کرنے کے لیے صرف پڑھے لکھے اور پڑھے لکھے لوگوں سے رابطہ کیا جا سکتا ہے۔شیڈول ایک سوالنامے کی طرح ہے جس میں تفتیشی عمل سے متعلق سوالات ہوتے ہیں۔سوالنامے اور شیڈول میں فرق صرف یہ ہے کہ جواب دہندہ خود سوالنامے کو بھرتا ہے،جب کہ ایک مناسب تربیت یافتہ شمار کنندہ

خود جواب دہندگان سے سوال پوچھ کر نظام الاوقات پُر کرتا ہے۔ سوالنامے پر شیڈول کا سب سے بڑا فائدہ یہ ہے کہ پڑھے لکھے اور ناخواندہ دونوں جواب دہندگان سے معلومات اکٹھی کی جاسکتی ہیں۔

دیگر طریقے

مٹی اور پانی کی خصوصیات کے بارے میں ڈیٹا کو مٹی کی کٹ اور پانی کے معیار کی کٹ کا استعمال کرتے ہوئے ان کی خصوصیات کی پیمائش کر کے براہ راست کھیت میں جمع کیا جاتا ہے۔ اسی طرح، فیلڈ سائنس دان ٹرانس ڈوسر کا استعمال کرتے ہوئے فصلوں اور پودوں کی صحت کے بارے میں ڈیٹا اکٹھا کرتے ہیں۔

ڈیٹا کا ثانوی ماخذ

اعداد و شمار کے ثانوی ذرائع شائع شدہ اور غیر مطبوعہ ریکارڈز پر مشتمل ہوتے ہیں جن میں سرکاری اشاعتیں، دستاویزات اور رپورٹیں شامل ہوتی ہیں۔

شائع شدہ ذرائع

مختلف وزارتوں اور حکومت ہند کے محکموں، ریاستی حکومتوں اور ضلعی بلیٹن کی اشاعتیں ثانوی معلومات کے سب سے اہم ذرائع میں سے ایک ہیں۔ ان میں رجسٹرار جنرل آف انڈیا کے دفتر کے ذریعہ شائع کردہ ہندوستان کی مردم شماری، قومی نمونہ سروے کی رپورٹیں، ہندوستانی محکمہ موسمیات کی موسمی رپورٹس اور ریاستی حکومتوں کے ذریعہ شائع کردہ شماریاتی خلاصے، اور مختلف کمیشنوں کے ذریعہ شائع ہونے والی متواتر رپورٹیں شامل ہیں۔ کچھ سرکاری اشاعتیں تصویر میں دکھائی گئی ہیں۔

نیم/نیم سرکاری پبلیکیشنز

مختلف شہروں اور قصبوں کی شہری ترقیاتی اتھارٹیز اور میونسپل کارپوریشنز، ضلع پریشدوں) ضلع کونسلوں) وغیرہ کی اشاعتیں اور رپورٹیں اس زمرے میں آتی ہیں۔

بین الاقوامی پبلیکیشنز

بین الاقوامی پبلیکیشن میں اقوام متحدہ کی مختلف ایجنسیوں کی طرف سے شائع ہونے والی سالانہ کتابیں،

i. رپورٹس اور مونوگراف شامل ہیں، جیسے کہ اقوام متحدہ کی تعلیمی، سائنسی اور ثقافتی تنظیم (UNESCO) ،

ii. اقوام متحدہ کے ترقیاتی پروگرام (UNDP) ، عالمی ادارہ صحت

iii. (WHO)، خوراک اور زرعی تنظیم (FAO) وغیرہ۔ انسانی ترقی کی رپورٹ۔

پرائیویٹ پبلیکیشن

اخبارات اور نجی اداروں کی طرف سے شائع ہونے والی سالانہ کتابیں، سروے، تحقیقی رپورٹس اور مونوگراف اس زمرے میں آتے ہیں۔

اخبارات اور رسائل۔ روزنامہ اخبارات اور ہفتہ وار، پندرہ روزہ اور ماہانہ میگزین ثانوی ڈیٹا کے آسانی سے قابل رسائی ذرائع کے طور پر کام کرتے ہیں۔ الیکٹرانک میڈیا الیکٹرانک میڈیا، خاص طور پر انٹرنیٹ، حالیہ دنوں میں ثانوی ڈیٹا کا ایک بڑا ذریعہ بن کر ابھرا ہے۔

غیر مطبوعہ ذرائع

1. سرکاری دستاویزات

غیر مطبوعہ رپورٹس، مونوگراف اور دستاویزات ثانوی ڈیٹا کا ایک اور ذریعہ ہیں۔ ان دستاویزات کو گورننس کی مختلف سطحوں پر غیر مطبوعہ ریکارڈ کے طور پر تیار اور برقرار رکھا جاتا ہے۔ مثال کے طور پر، متعلقہ گاؤں کے پٹواری کے ذریعہ دیہات کی سطح پر ریونیو کا ریکارڈ رکھا جاتا ہے جو گاؤں کی سطح کی معلومات کا ایک اہم ذریعہ ہے۔

2. نیم سرکاری ریکارڈز

مختلف میونسپل کارپوریشنوں، ضلع کونسلوں اور سول سروسز کے محکموں کی طرف سے تیار کردہ اور ان کی دیکھ بھال کی متواتر رپورٹیں اور ترقیاتی منصوبے نیم سرکاری ریکارڈ میں شامل ہیں۔

3. نجی دستاویزات

ان میں کمپنیوں، ٹریڈ یونینوں، مختلف سیاسی اور غیر سیاسی تنظیموں اور رہائشیوں کی فلاحی انجمنوں کی غیر مطبوعہ رپورٹس اور ریکارڈ شامل ہیں۔

2.5 ٹیبلیشن اور ڈیٹا کی درجہ بندی (Tabulation and classification of data)

ابتدائی یا ثانوی ذرائع سے جمع کردہ اعداد و شمار ابتدائی طور پر کم سے کم فہم کے ساتھ معلومات کے ایک بڑے ہنگامے کے طور پر ظاہر ہوتے ہیں۔ یہ خام ڈیٹا کے طور پر جانا جاتا ہے. بامعنی نتائج نکالنے اور انہیں قابل استعمال بنانے کے لیے خام ڈیٹا ٹیبلیشن اور درجہ بندی کی ضرورت ہے۔ اعداد و شمار کا خلاصہ پیش کرنے کے لیے سب سے آسان آلات میں سے ایک شماریاتی جدول ہے۔ یہ کالموں اور قطاروں میں ڈیٹا کی ایک منظم ترتیب ہے۔ ٹیبل کا مقصد پیشکش کو آسان بنانا اور موازنہ کو آسان بنانا ہے۔ یہ جدول قاری کو مطلوبہ معلومات کو تیزی سے تلاش کرنے کے قابل بناتا ہے۔ اس طرح، جدول تجزیہ کار کے لیے کم سے کم جگہ کے اندر ایک منظم انداز میں ڈیٹا کی ایک بڑی مقدار کو پیش کرنا ممکن بناتی ہیں۔

ڈیٹا کمپلیشن اور پریزنٹیشن

ڈیٹا کو جمع کیا جاتا ہے، ٹیبل کیا جاتا ہے اور اسے جدولی شکل میں یا تو مطلق شرائط، فیصد یا اشاریہ میں پیش کیا جاتا ہے۔

مطلق ڈیٹا:

جب ڈیٹا کو ان کی اصل شکل میں انٹیجرز کے طور پر پیش کیا جاتا ہے، تو اسے مطلق ڈیٹا یا خام ڈیٹا کہا جاتا ہے۔ مثال کے طور پر، کسی ملک یا ریاست کی کل آبادی، فصل کی کل پیداوار یا مینوفیکچرنگ انڈسٹری وغیرہ۔ جدول 1.1 ہندوستان اور کچھ منتخب ریاستوں کی آبادی کا قطعی ڈیٹا دکھاتا ہے۔

Population of India and Selected States/Union Territories, 2011

		Persons	Males	Females
1	2	3	4	5
	INDIA	1,21,05,69,573	62,31,21,843	58,74,47,730
1.	Jammu and Kashmir[2]	1,25,41,302	66,40,662	59,00,640
2.	Himachal Pradesh	68,64,602	34,81,873	33,82,729
3.	Punjab	2,77,43,338	1,46,39,465	1,31,03,873
4.	Chandigarh[3]	10,55,450	5,80,663	4,74,787
5.	Uttarakhand	1,00,86,292	51,37,773	49,48,519
6.	Haryana	2,53,51,462	1,34,94,734	1,18,56,728
7.	National Capital Territory of Delhi	1,67,87,941	89,87,326	78,00,615
8.	Rajasthan	6,85,48,437	3,55,50,997	3,29,97,440
9.	Uttar Pradesh	19,98,12,341	10,44,80,510	9,53,31,831
10.	Bihar	10,40,99,452	5,42,78,157	4,98,21,295

فیصد/تناسب

کچھ وقت کے اعداد و شمار کو تناسب یا فیصد کی شکل میں ٹیبل کیا جاتا ہے جو ایک نارمل پیرامیٹر سے حساب کیا جاتا ہے، جیسے خواندگی کی شرح یا آبادی کی شرح نمو، زرعی مصنوعات یا صنعتی مصنوعات کا فیصد، وغیرہ۔ نیچے دی گئی جدول کئی دہائیوں کے دوران ہندوستان کی شرح خواندگی کو فیصد کی شکل میں پیش کرتی ہے۔ شرح خواندگی کا شمار اس طرح کیا جاتا ہے:

کل لٹریٹس

$$\frac{\text{Total Literates}}{\text{Total Population}} \times 100$$

انڈیکس نمبر

انڈیکس نمبر ایک شماریاتی پیمانہ ہے جو متغیر یا متعلقہ متغیرات کے گروپ میں تبدیلیوں کو دکھانے کے لیے ڈیزائن کیا گیا ہے جو کہ وقت، جغرافیائی لحاظ سے جدول 1.2: خواندگی کی شرح: 1951-2011

مقام یا دیگر خصوصیات۔ واضح رہے کہ انڈیکس نمبر نہ صرف وقت کے ساتھ ہونے والی تبدیلیوں کی پیمائش کرتے ہیں بلکہ مختلف مقامات، صنعتوں، شہروں یا ممالک کے معاشی حالات کا موازنہ بھی کرتے ہیں۔ قیمت اور مقدار میں تبدیلی دیکھنے کے لیے اشاریہ نمبر کو معاشیات اور کاروبار میں بڑے پیمانے پر استعمال کیا جاتا ہے۔ انڈیکس نمبر کے حساب کے لیے مختلف طریقے ہیں۔ تاہم، سادہ مجموعی طریقہ سب سے زیادہ استعمال کیا جاتا ہے. یہ مندرجہ ذیل فارمولے کا استعمال کرتے ہوئے حاصل کیا جاتا ہے :

Year	***Person***	***Male***	***Female***
1951	18.33	27.16	8.86
1961	28.3	40.4	15.35
1971	34.45	45.96	21.97
1981	43.57	56.38	29.76
1991	52.21	64.13	39.29
2001	64.84	75.85	54.16
2011	73.0	80.9	64.6

$\sum q_1$ = Total of the current year production

$\sum q_0$ = Total of the base year production

$$\frac{\sum q_1}{\sum q_0} \times 100$$

نارمل طور پر، بیس سال کی اقدار کو 100 کے طور پر لیا جاتا ہے اور اس کے بعد انڈیکس نمبر کا حساب لگایا جاتا ہے۔ مثال کے طور پر، جدول 1.3 ہندوستان میں خام لوہے کی پیداوار اور 71-1970 سے لے کر 01-2000 تک انڈیکس نمبر میں ہونے والی تبدیلیوں کو ظاہر کرتا ہے۔ بنیادی سال 71-1970 ہے

جدول : 1.3 ہندوستان میں لوہے کی پیداوار

Year	*Production (in million tonnes)*	*Calculation*	*Index Number*
1970-71	32.5	32.5	100
1980-81	42.2	32.5 × 100	130
1990-91	53.7	42.2	165
2000-01	67.4	32.5 × 100 53.7 32.5 × 100 67.4 × 100 32.5	207

ڈیٹاکی پروسیسنگ

خام ڈیٹاکی پروسیسنگ کے لیے منتخب کلاسوں میں ان کی ٹیبلیشن اور درجہ بندی کی ضرورت ہوتی ہے۔ مثال کے طور پر، جدول 1.4 میں دیے گئے ڈیٹا کو یہ سمجھنے کے لیے استعمال کیا جاسکتا ہے کہ ان پر کارروائی کیسے کی جاتی ہے۔ ہم دیکھ سکتے ہیں کہ دیا گیاڈیٹا غیر گروپ شدہ ہے۔ لہذا، پہلا قدم یہ ہے کہ ڈیٹاکا حجم کم کرنے اور اسے سمجھنے میں آسانی پیدا کرنے کے لیے گروپ بندی کی جائے۔

47	*02*	*39*	*64*	*22*	*46*	*28*	*02*	*09*	*10*
89	*96*	*74*	*06*	*26*	*15*	*92*	*84*	*84*	*90*
32	*22*	*53*	*62*	*73*	*57*	*37*	*44*	*67*	*50*
18	*51*	*36*	*58*	*28*	*65*	*63*	*59*	*75*	*70*
56	*58*	*43*	*74*	*64*	*12*	*35*	*42*	*68*	*80*
64	*37*	*17*	*31*	*41*	*71*	*56*	*83*	*59*	*90*

ڈیٹاکی گروپ بندی

خام ڈیٹاکی گروپ بندی کے لیے کلاسوں کی تعداد کا تعین کرنے کی ضرورت ہوتی ہے جن میں خام ڈیٹا کو گروپ کیا جاناہے اور کلاس کے وقفے کیا ہوں گے۔ کلاس وقفہ کا انتخاب اور کلاسز کی تعداد، تاہم، خام ڈیٹاکی حد پر منحصر ہے۔ دے گئے جدول میں دیا گیا خام ڈیٹا 02 سے 96 تک ہے۔ اس لیے ہم آسانی سے ڈیٹا کو دس کلاسوں میں گروپ کرنے کا انتخاب کر سکتے ہیں جس میں ہر گروپ میں دس اکائیوں کے وقفے کے ساتھ، جیسے۔0–10، 10–20، 20–30، وغیرہ۔

Table 1.5: Making Tally Marks to Obtain Frequency

GClaroup	***Numerical of Raw Data***	***Tally Marks***	***Number of Individual***
0-10	02,02,09,06	////	4
10-20	10,15,18,12,17	~~////~~	5
20-30	22,28,26,22,28	~~////~~	5
30-40	39,32,37,36,35,37,31	~~////~~//	7
40-50	47,46,44,43,42,41	~~////~~/	6

50-60	53,57,50,51,58, 59,56,58,56,59	~~////~~ ~~////~~	10
60-70	64,62,67,65, 63,64,68,64	~~////~~ ///	8
70-80	74,73,75,70,74,71	~~////~~ /	6
80-90	89,84,84,80,83	~~////~~	5
90-100	96,92,90,90	////	4
			$f = N = 60$

درجہ بندی کا عمل

گروپوں کی تعداد اور ہر گروپ کے کلاس وقفہ کا تعین ہونے کے بعد، خام ڈیٹا کی درجہ بندی کی جاتی ہے جیسا کہ جدول 1.5 میں دکھایا گیا ہے۔ یہ ایک طریقہ سے کیا جاتا ہے جسے فور اینڈ کراس میتھڈ یا ٹیلی مارکس کے نام سے جانا جاتا ہے۔

سب سے پہلے، جس گروپ میں وہ گر رہا ہے اس کے ہر فرد کو ایک عدد نشان تفویض کیا جاتا ہے۔ مثال کے طور پر، خام ڈیٹا میں پہلا عدد 47 ہے۔ چونکہ، یہ 40-50 کے گروپ میں آتا ہے، ایک عدد نشان جدول 1.5 کے کالم 3 میں درج کیا گیا ہے۔

تعدد کی تقسیم

جدول 1.5 میں ہم نے مقداری متغیر کے خام ڈیٹا کی درجہ بندی کی ہے اور انہیں طبقاتی طور پر گروپ کیا ہے۔ افراد کی تعداد (ٹیبل 1.5 کے چوتھے کالم میں جگہیں) تعدد کے طور پر جانا جاتا ہے اور کالم تعدد کی تقسیم کی نمائندگی کرتا ہے۔ یہ واضح کرتا ہے کہ متغیر کی مختلف اقدار کو مختلف کلاسوں میں کیسے تقسیم کیا جاتا ہے۔ تعدد کو سادہ اور مجموعی تعدد کے طور پر درجہ بندی کیا جاتا ہے۔

سادہ تعدد

اس کا اظہار 'f' سے ہوتا ہے اور ہر گروپ میں آنے والے افراد کی تعداد کی نمائندگی کرتا ہے (ٹیبل 1.6)۔ تمام کلاسوں کو تفویض کردہ تمام تعدد کا مجموعہ، دی گئی سیریز میں انفرادی مشاہدات کی کل تعداد کی نمائندگی کرتا ہے۔ اعداد و شمار میں، یہ علامت N سے ظاہر ہوتا ہے۔

جدول: تعدد کی تقسیم

مجموعی تعدد

اس کا اظہار 'Cf' سے ہوتا ہے اور ہر گروپ میں پچھلی رقم کے ساتھ لگاتار سادہ تعدادات کو شامل کر کے حاصل کیا جا سکتا ہے، جیسا کہ جدول 1.6 کے کالم 3 میں دکھایا گیا ہے۔ مثال کے طور پر، جدول 1.6 میں پہلی سادہ فریکوئنسی 4 ہے۔ 5 کی اگلی فریکوئنسی 4 میں شامل کی جاتی ہے جو کہ اگلی مجموعی فریکوئنسی کے طور پر کل 9 دیتی ہے۔ اسی طرح، ہر اگلے نمبر کو شامل کریں جب تک کہ 60 کی آخری مجموعی تعداد نہ ہو۔

حاصل کیا نوٹ کریں کہ یہ N یا f کے برابر ہے۔

مجموعی تعداد کا فائدہ یہ ہے کہ کوئی آسانی سے یہ بتا سکتا ہے کہ 50 سے کم اسکور کرنے والے 27 افراد ہیں یا 60 میں سے 45 افراد 70 کے اسکور سے نیچے ہیں۔ ہر سادہ فریکوئنسی اس کے گروپ یا کلاس سے وابستہ ہے۔ خصوصی یا گروپوں یا کلاسوں کی تشکیل کے لیے جامع طریقے استعمال کیے جاتے ہیں۔

خصوصی طریقہ

جیسا کہ جدول 1.6 میں دکھایا گیا ہے، اس کے پہلے کالم میں دو نمبر دکھائے گئے ہیں۔ نوٹ کریں کہ ایک گروپ کی اوپری حد اگلے گروپ کی نچلی حد کے برابر ہے۔ مثال کے طور پر، ایک گروپ (20-30) کی اوپری حد 30 ہے، جو اگلے گروپ (30-40) کی نچلی حد ہے، جس سے دونوں گروپوں میں 30 ظاہر ہوتے ہیں۔ لیکن 30 کی قدر رکھنے والا کوئی بھی مشاہدہ اس گروپ میں شامل ہوتا ہے جہاں وہ اپنی کم حد پر ہوتا ہے اور اسے اس گروپ سے خارج کر دیا جاتا ہے جہاں یہ اوپری حد ہوتی ہے جیسا کہ (20-30 گروپوں میں)۔ اسی لیے اس طریقہ کو خصوصی طریقہ کہا جاتا ہے، یعنی ایک گروہ کو اس کی بالائی حدود سے خارج کر دیا جاتا ہے۔ اب آپ اندازہ لگا سکتے ہیں کہ ٹیبل 1.4 کی تمام معمولی قدریں کہاں جائیں گی۔

جدول 1.6 میں گروپوں کی تشریح درج ذیل طریقے سے کی گئی ہے-0 اور 10 سے کم 10 اور 20 سے کم

20 اور 30 سے کم اور 40 سے کم

40 اور 50 سے کم اور 60 سے کم

60 اور 70 سے کم اور 80 سے کم

80 اور 90 سے کم اور 100 سے کم

لہذا، اس قسم کی گروپ بندی میں کلاس دس اکائیوں پر پھیلی ہوئی ہے۔ مثال کے طور پر 20، 21 ،22، 23، 24، 25، 26، 27، 28 اور 29 تیسرے گروپ میں شامل ہیں۔

جامع طریقہ

اس طریقہ کار میں، ایک گروپ کی بالائی حد کے برابر ایک قدر اسی گروپ میں شامل کی جاتی ہے۔ لہذا، یہ جامع طریقہ کے طور پر جانا جاتا ہے.

اس طریقہ کار میں کلاسوں کا ذکر ایک مختلف شکل میں کیا گیا ہے، جیسا کہ جدول 1.7 کے پہلے کالم میں دکھایا گیا ہے۔ نارمل طور پر، ایک گروپ کی اوپری حد اگلے گروپ کی نچلی حدوں کے ساتھ 1 سے مختلف ہوتی ہے۔ یہ نوٹ کرنا ضروری ہے کہ ہر گروپ اس طریقہ کار میں بھی دس

اکائیوں پر پھیلا ہوا ہے۔مثال کے طور پر، 50–59 کے گروپ میں دس قدریں شامل ہیں یعنی 50، 51، 52، 53، 54، 55، 56، 57، 58 اور 59۔اس طریقہ میں، فریکوئنسی کی تقسیم کو تلاش کرنے کے لیے اوپری اور نچلی دونوں حدیں شامل ہیں۔

تعدد کثیر الاضلاع

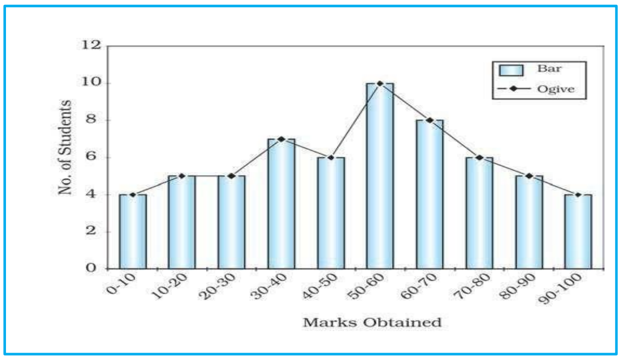
TableFrequency Distribution

Group	*f*	*Cf*
0 – 9	4	4
10 – 19	5	9
20 – 29	5	14
30 – 39	7	21
40 – 49	6	27
50 – 59	10	37
60 – 69	8	45
70 – 79	6	51
80 – 89	5	56
90 – 99	4	60
	f	= *N* = 60

تعدد کی تقسیم کا گراف تعدد کثیر الاضلاع کے نام سے جانا جاتا ہے۔یہ دو یا دو سے زیادہ فریکوئنسی کی تقسیم کا موازنہ کرنے میں مدد کرتا ہے۔ دونوں تعدد کو بالترتیب بار ڈایا گرام اور لائن گراف کا استعمال کرتے ہوئے دکھایا گیا ہے۔

اوگیو

جب تعدد کو شامل کیا جاتا ہے تو انہیں مجموعی تعدد کہا جاتا ہے اور اسے ایک جدول میں درج کیا جاتا ہے



تصویر: فریکوئنسی ڈسٹری بیوشن پولیگون

مجموعی تعدد جدول مجموعی تعدد کی منصوبہ بندی کے ذریعہ حاصل کردہ وکر کو اوگیو (اوجیو کے طور پر تلفظ) کہا جاتا ہے۔ یہ یا تو طریقہ سے کم یا زیادہ طریقہ سے بنایا گیا ہے۔ کم سے کم طریقہ میں، ہم کلاسوں کی اوپری حد سے شروع کرتے ہیں اور تعدد کو شامل کرتے ہیں۔ جب ان تعدد کو پلاٹ کیا جاتا ہے، تو ہمیں ایک بڑھتا ہوا وکر ملتا ہے جیسا کہ ٹیبل اور تصویر میں دکھایا گیا ہے۔ طریقہ سے زیادہ میں، ہم کلاسوں کی نچلی حدوں سے شروع کرتے ہیں اور مجموعی تعدد سے، ہم ہر کلاس کی فریکوئنسی کو گھٹاتے ہیں۔ جب ان تعدد کو پلاٹ کیا جاتا ہے، تو ہمیں ایک زوال پذیر وکر ملتا ہے جیسا کہ جدول 1.9 اور تصویر 1.7 میں دکھایا گیا ہے۔

دونوں انجیر۔ 1.5 اور 1.6 کو ملا کر اوگیو سے کم اور زیادہ کی تقابلی تصویر حاصل کی جاسکتی ہے جیسا کہ جدول اور تصویر میں دکھایا گیا ہے۔

Table: Frequency Distribution less than Method

Less than Method	***Cf***
Less than 10	4
Less than 20	9
Less than 30	14
Less than 40	21
Less than 50	27
Less than 60	37
Less than 70	45
Less than 80	51
Less than 90	56
Less than 100	60

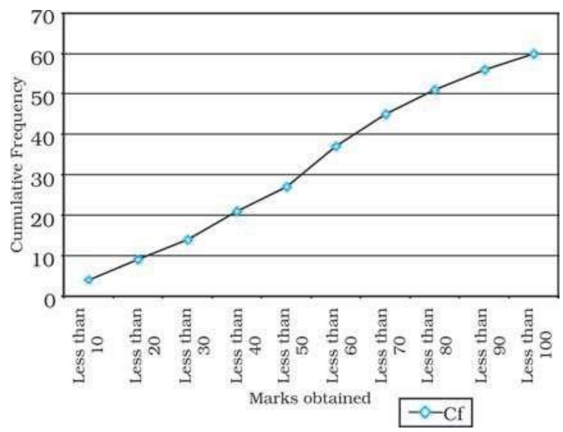


Table: Frequency Distribution more than Method

More than Method	***Cf***
More than 0	60
More than 10	56
More than 20	51
More than 30	44
More than 40	38
More than 50	28
More than 60	20
More than 70	14
More than 80	9
More than 90	4

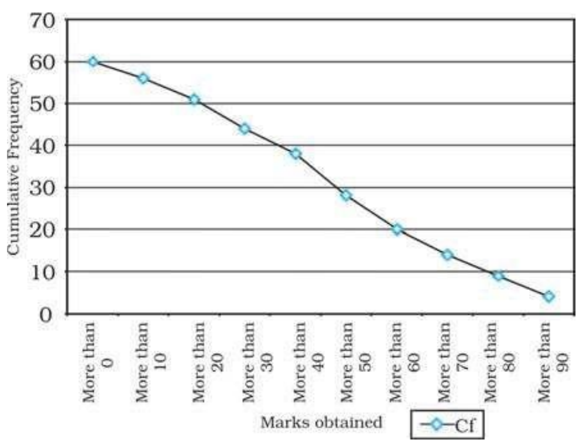
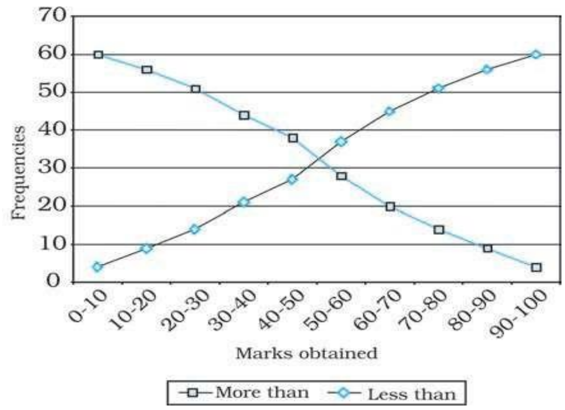


Table: Less than and more than Ogive

Marks obtained	*Less than*	*More than*
0 - 10	4	60
10 - 20	9	56
20 - 30	14	51
30 - 40	21	44
30 - 40	27	38
50 - 60	37	28
60 - 70	45	20
70 - 80	51	14
80 - 90	56	9



2.5 ڈیٹاسائنس (Data Science)

ڈیٹا سائنس جادو سے کم نہیں ہے، اور ڈیٹا سائنسدان ایک جادو گر ہے جو اپنی ٹوپی میں مواد کے ساتھ کرتب دکھاتا ہے۔ اب، جیسا کہ جادو مختلف عناصر پر مشتمل ہے، اسی طرح ڈیٹا سائنس ایک بین الضابطہ میدان ہے۔ ہم اسے مختلف شعبوں جیسا کہ مواد ہیرا پھیری، ڈیٹا ویژولائزیشن، شماریاتی تجزیہ، اور مشین لرننگ کا امتزاج سمجھ سکتے ہیں۔ ان ذیلی ڈومینز میں سے ہر ایک کو یکساں اہمیت حاصل ہے۔

ڈیٹا ہیرا پھیری:

ڈیٹا میں ہیرا پھیری کی تکنیکوں کی مدد سے، آپ کم سے کم کوشش کے ساتھ خام ڈیٹا سے دلچسپ بصیرت حاصل کر سکتے ہیں۔ ڈیٹا ہیرا پھیری معلومات کو پڑھنے کے قابل بنانے کے لیے ترتیب دینے کا عمل ہے۔

قابل فہم انجینئرز ڈیٹا کا استعمال کرتے ہوئے ڈیٹا کی ہیرا پھیری کرتے ہیں۔

ہیرا پھیری کی زبان (DML) ڈیٹا کو شامل کرنے، حذف کرنے یا تبدیل کرنے کے قابل ہے۔ ڈیٹا مختلف ذرائع سے آتا ہے۔ مختلف ڈیٹا کے ساتھ کام کرتے ہوئے، آپ کو اسے اپنے فیصلہ سازی کے عمل میں استعمال کرنے کے لیے اسے منظم، صاف اور تبدیل کرنے کی ضرورت ہے۔ یہ وہ جگہ ہے جہاں ڈیٹا میں ہیرا پھیری فٹ بیٹھتی ہے۔ ڈیٹا کی ہیرا پھیری آپ کو ڈیٹا کو منظم اور مربوط کرنے کی اجازت دیتی ہے جو قابل عمل معلومات کو چلانے میں مدد کرتی ہے۔ ڈیٹا میں ہیرا پھیری، جسے ڈیٹا کی تیاری کے نام سے بھی جانا جاتا ہے، صارفین کو کاروباری ذہانت اور تجزیات کے لیے جامد ڈیٹا کو ایندھن میں تبدیل کرنے کے قابل بناتا ہے۔ بہت سے ڈیٹا سائنسدان ڈیٹا کو ترتیب دینے اور رپورٹیں تیار کرنے کے لیے ڈیٹا کی تیاری کے سافٹ ویئر کا استعمال کرتے ہیں، اس لیے غیر تجزیہ کار اور دیگر اسٹیک ہولڈرز قیمتی معلومات حاصل کرتے ہوئے باخبر فیصلے کر سکتے ہیں۔

ڈیٹاہیر اپھیری کی اہمیت

ڈیٹا میں ہیر اپھیری تنظیموں کے لیے ضرورت کے مطابق ڈیٹا کو منظم اور تجزیہ کرنا آسان بناتی ہے۔ یہ انہیں اہم کاروباری افعال انجام دینے میں مدد کرتا ہے جیسے رجحانات کا تجزیہ کرنا، خریداروں کے رویے، اوران کے مالیاتی ڈیٹا سے آگاہی حاصل کرنا۔

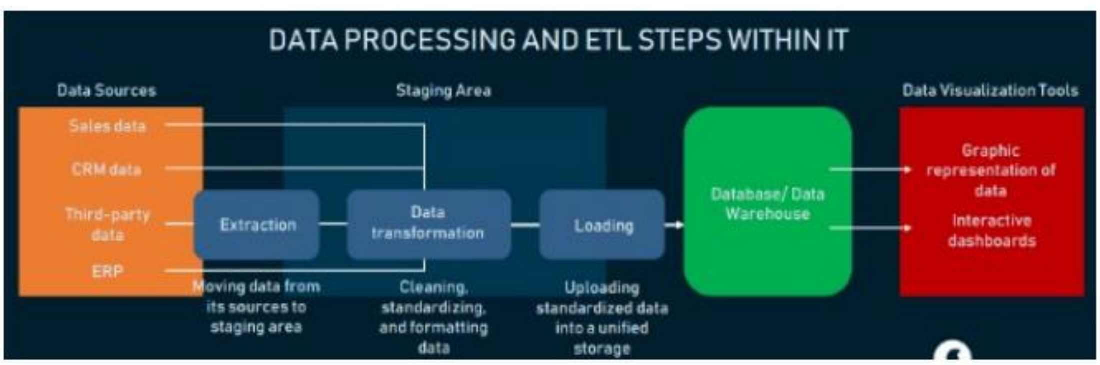
ڈیٹاہیر اپھیر یسے کاروبار کو کئی فوائد ملتے ہیں، بشمول:

- مستقل مزاجی :ڈیٹا کی ہیر اپھیری مختلف ذرائع سے جمع کردہ ڈیٹا میں مستقل مزاجی کو برقرار رکھتی ہے، جس سے کاروبار کو ایک متفقہ نظریہ ملتا ہے جو انہیں بہتر، زیادہ باخبر فیصلے کرنے میں مدد کرتا ہے۔
- قابل استعمال :ڈیٹا میں ہیر اپھیری صارفین کو ڈیٹا کو صاف اور منظم کرنے اور اسے زیادہ مؤثر طریقے سے استعمال کرنے کی اجازت دیتی ہے۔
- پیشن گوئی :ڈیٹا میں ہیر اپھیری کاروبار کو تاریخی ڈیٹا کو سمجھنے کے قابل بناتی ہے اور مستقبل کی پیشین گوئیاں تیار کرنے میں ان کی مدد کرتی ہے، خاص طور پر مالیاتی ڈیٹا کے تجزیہ میں۔
- کلینزنگ :ڈیٹا میں ہیر اپھیری غیر مطلوبہ ڈیٹا کو صاف کرنے اور اہم معلومات رکھنے میں مدد کرتی ہے۔ انٹرپرائزز ریکارڈ کو صاف کر سکتے ہیں، الگ تھلگ کر سکتے ہیں، اور غیر ضروری متغیرات کو بھی کم کر سکتے ہیں، اور اپنی ضرورت کے ڈیٹا پر توجہ مرکوز کر سکتے ہیں۔

ڈیٹاویژولائزیشن:

یہ خام معلومات (متن، نمبر، یا علامت) کو گرافک فارمیٹ میں تبدیل کرنے کی مشق ہے۔ اعداد و شمار کو ایک واضح مقصد کے ساتھ اکائیوں کے درمیان منطقی ارتباط کو ظاہر کرنا رجحانات اور نمونوں کی وضاحت جیسی قسم پر منحصر ہے۔

منطقی کنکشن اور ڈیٹا خود، تصور ایک مناسب شکل میں کیا جا سکتا ہے. لہذا، یہ بالکل آسان ہے، کسی بھی تجزیاتی رپورٹ میں ڈیٹا کی تشریحات کی مثالیں ہوتی ہیں جیسے پائی چارٹس، موازنہ بار، آبادیاتی نقشے اور بہت کچھ۔

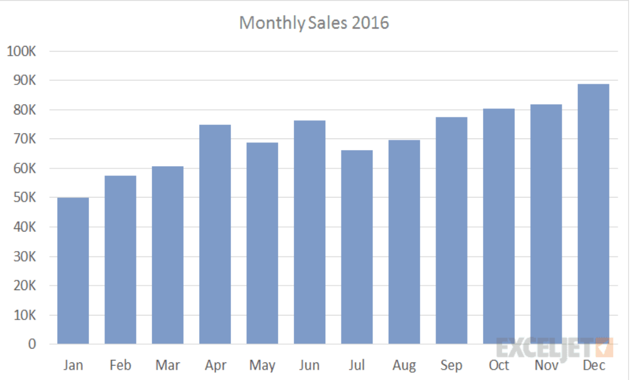


Source: A Complete Guide to Data A Complete Guide to Data Visualization in Business Intelligence

ڈیٹاویژولائزیشن کی نارمل اقسام

i. ہبار چارٹ

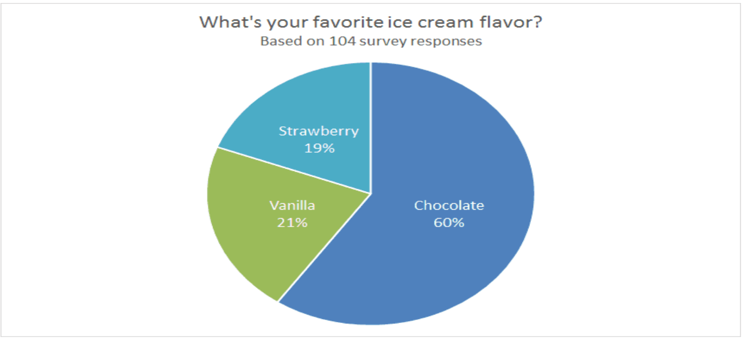
بار چارٹ ڈیٹاکا ئیوں کاایک دوسرے سے موازنہ کرنے کے بنیادی طریقوں میں سے ایک ہے۔اس کی سادہ گرافک شکل کی وجہ سے،ایک بار چارٹ اکثر کاروباری تجزیات میں ایک انٹرایکٹو صفحہ عنصر کے طور پر استعمال ہوتا ہے۔بار چارٹس اس قدر ورسٹائل ہیں کہ اس میں ترمیم کی جائے اور زیادہ پیچیدہ ڈیٹاماڈل دکھائے جائیں۔سلاخوں کو کلسٹرز میں بنایا جاسکتا ہے یااسٹیک کیا جاسکتا ہے،مارکیٹ کے حصوں میں تقسیم کو ظاہر کرنے کے لیے، یااشیاء کے ذیلی زمرہ جات ایساہی افقی بار چارٹس کے لیے بھی ہوتا ہے،جو سلاخوں پر لگائے جانے والے لمبے ڈیٹالیبلز کے لیے زیادہ موزوں ہے۔



کب استعمال کرنا ہے عددی معلومات اشیاء کا موازنہ کرنا، لمبے ڈیٹالیبلز کو فٹ کرنے کے لیے افقی چارٹس کا استعمال کریں۔مزید تفصیلی موازنہ کے لیے ہر چیز کو حصوں میں تقسیم کرنے کے لیے سلاخوں میں اسٹیک رکھیں۔

ii. پائی چارٹ

چارٹ کی ایک اور نارمل قسم جو ہم ہر جگہ دیکھتے ہیں،وہ ہے پائی چارٹ۔اس قسم کا چارٹ کسی بھی مارکیٹنگ یا سیلز ڈیپارٹمنٹ میں استعمال ہوتا ہے، کیونکہ یہ اشیاء کی ساخت یایونٹ سے یونٹ موازنہ کو ظاہر کرنا آسان بناتا ہے۔کب استعمال کرنا ہے : کسی شے کی ترکیب، حصوں کا پوری آبجیکٹ سے موازنہ کرنا۔

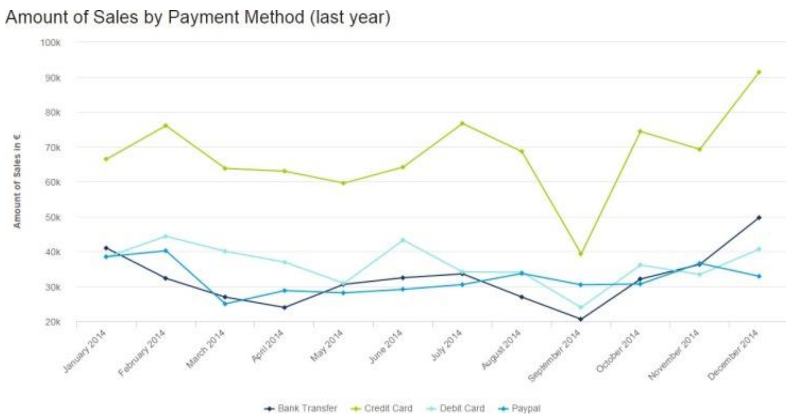


پائی چارٹ آئس کریم کے ذائقہ کی ترجیح کا فیصد ارتباط دکھا رہا ہے۔

iii. لائن گراف

اس قسم کا گراف وقت کے ساتھ ساتھ یونٹ کی قدر کو ظاہر کرنے کے لیے افقی محور اور عمودی محور کا استعمال کرتا ہے۔ متعدد جہتوں سے ڈیٹا کی نمائندگی کرنے کے لیے لائن گراف کو بار چارٹ کے ساتھ بھی ملایا جاسکتا ہے۔

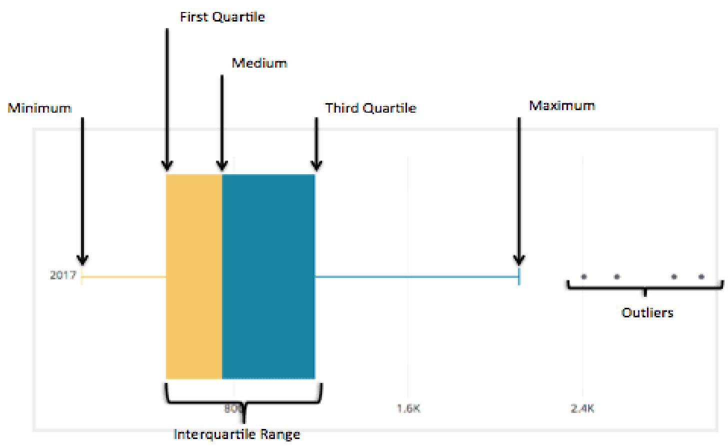
کب استعمال کرنا ہے: ٹائم لائن پر آبجیکٹ ویلیو، وقت کے ساتھ رویے میں رجحانات کو ظاہر کرتی ہے۔



ادائیگی کے طریقوں کے ذریعہ فروخت کا تجزیہ

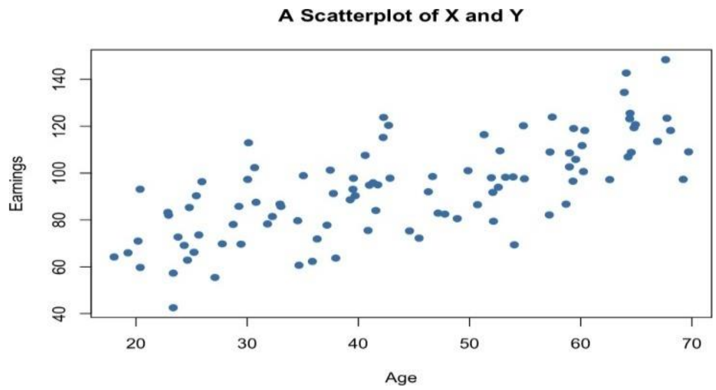
iv. باکس پلاٹ

پہلی نظر میں، ایک باکس پلاٹ کافی پیچیدہ لگتا ہے۔ لیکن اگر ہم مثال کو قریب سے دیکھیں تو یہ واضح ہو جاتا ہے کہ اس میں چوتھائیوں کو افقی انداز میں دکھایا گیا ہے۔ ہمارے یہاں بنیادی عناصر کم از کم، زیادہ سے زیادہ، اور درمیانی ہیں جو پہلے اور تیسرے چوتھائی کے درمیان رکھے گئے ہیں۔ ایک باکس کیا دکھاتا ہے اشیاء کی تقسیم، اور ان کی میڈین سے انحراف۔ کب استعمال کرنا ہے : پیچیدہ چیز کی تقسیم، درمیانی قدر سے انحراف۔ باکس پلاٹ کو 5 چوتھائی حصوں میں تقسیم کیا گیا ہے، جب کہ آؤٹ لیرز کو ایسی چیز کے طور پر دکھایا گیا ہے جو تقسیم کے علاقے سے باہر ہو



Scatter plot .v

اس قسم کا تصور X اور Y محور پر بنایا گیا ہے۔ان کے درمیان، چاروں طرف نقطے رکھے گئے ہیں،جو اشیاء کی وضاحت کرتے ہیں۔ گراف پر نقطے کی پوزیشن یہ بتاتی ہے کہ اس میں کون سی خصوصیات ہیں۔ جیسا کہ لائن گراف کے منور ملے میں، محوروں کے درمیان رکھے ہوئے نقطے ایک تقسیم سیکنڈ میں نظر آتے ہیں۔ اس قسم کے تصور کی واحد حد محوروں کی تعداد ہے۔

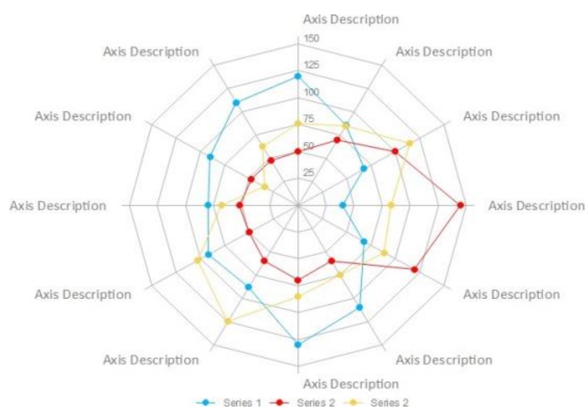


کب استعمال کرنا ہے :اشیاء کی تقسیم دکھانا، گراف پر ہر چیز کے معیار کی وضاحت کرنا۔ ایک اداس بکھرنے والا پلاٹ جو پیسے کمانے میں نوجوانوں کی نااہلی کو ظاہر کرتا ہے۔

vi. ریڈار یا مکڑی کا چارٹ

اس قسم کا چارٹ بنیادی طور پر ریڈیل فیشن میں تیار کردہ لائن چارٹ ہے۔اس میں مکڑی کے جالے کی شکل ہے جو متعدد محوروں اور متغیرات کے ذریعے تخلیق کی گئی ہے۔اس کا مقصد وہی ہے جو لائن چارٹ کا ہے۔ لیکن محوروں کی تعداد کی وجہ سے، آپ مختلف زاویوں سے اکائیوں کا موازنہ کر سکتے ہیں اور مائل کو گرافی طور پر دکھا سکتے ہیں۔

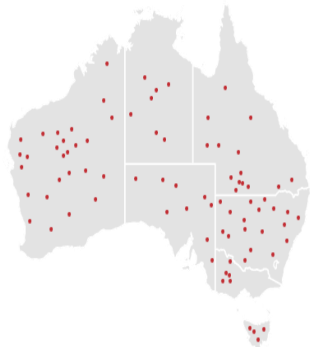
کب استعمال کرنا ہے :ڈیٹا کی خصوصیات کو بیان کرنا، مختلف جہتوں کے ذریعے متعدد اشیاء کا ایک دوسرے سے موازنہ کرنا۔



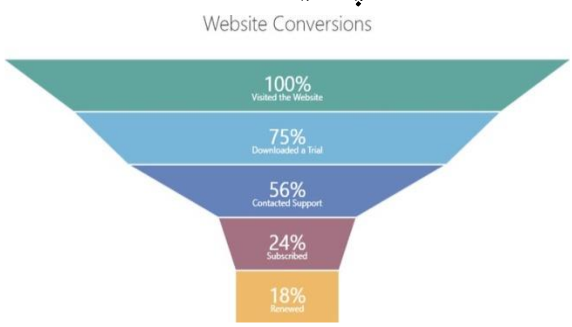
مکڑی کے چارٹ کا ڈھانچہ

vii. ڈاٹ میپ یا کثافت کا نقشہ

نقشے پر تصور کو ظاہر کرنا ڈیٹا کے جغرافیائی ڈومین کے لیے کام اتا ہے۔ کثافت کے نقشے پر رکھے ہوئے نقطوں کی مدد سے بنائے گئے ہیں، جو ہر یونٹ کے مقام کو نشان زد کرتے ہیں۔



ڈاٹ میپ کی ایک سادہ نمائندگی



Source: https://clusterdesign.io/funnel-chart/

viii. فنل چارٹس

یہ اشیاء کے مختلف گروہوں کے درمیان محدود تعلق کو ظاہر کرنے کے لیے بہترین ہیں۔ زیادہ تر منور ملات میں، فنل اشیاء میں فرق کرنے کے لیے جیومیٹرک فارم اور کلر کوڈنگ دونوں کا استعمال کریں گے۔ مثال کل ٹریفک نمبر اور سبسکرائبرز کی تعداد سے شروع ہونے والے تبادلوں کے نتائج دکھاتی ہے۔ اس قسم کا چارٹ اس وقت بھی کارآمد ہوتا ہے جب عمل میں متعدد مراحل ہوتے ہیں۔ اوپر دی گئی مثال پر، ہم دیکھ سکتے ہیں کہ " رابطہ کردہ سپورٹ "کے مرحلے کے بعد، سبسکرائبرز کی تعداد کم ہوگئی ہے۔ کب استعمال کرنا ہے : قدر/آبجیکٹ کے تنگ فیصد کے ساتھ عمل کے مراحل کی عکاسی کرنا۔ تصور کی قسم کا انتخاب کرتے ہوئے، یقینی بنائیں کہ آپ مندرجہ ذیل نکات کو واضح طور پر سمجھتے ہیں:

آپ کے ڈیٹا سیٹ کی تفصیلات : آپ کی کمپنی میں علم یا شعبہ کا ڈومین سامعین : وہ لوگ جن کو آپ معلومات پیش کرنا چاہتے ہیں۔ کنکشن منطق : اشیاء کا موازنہ، تقسیم، تعلق، عمل کی تفصیل، وغیرہ۔ آؤٹ پٹ : بس، کسی کو یہ معلومات دکھانے کی وجہ

2.7 شماریاتی تجزیہ(Descriptive Statistics)

شماریاتی تجزیہ اعداد و شمار کے نمونے جمع کرنے اور ان کا تجزیہ کرنے کا عمل ہے تاکہ نمونوں اور رجحانات کو بے نقاب کیا جاسکے اور یہ اندازہ لگایا جا سکے کہ بہتر اور زیادہ سائنسی فیصلے کرنے کے لیے آگے کیا ہو سکتا ہے۔

ڈیٹا اکٹھا کرنے کے بعد، شماریاتی تجزیہ آپ کے کاروبار میں بہت سی چیزوں کے لیے استعمال کیا جاسکتا ہے۔ کچھ نکات درجہ ذیل ہیں:

- کلیدی نتائج کو پیش کرنے کے لیے گراف یا چارٹ میں ڈیٹا کا خلاصہ پیش کرنا
- ڈیٹا کے اندر اہم اقدامات دریافت کرنا، جیسا کہ اوسط کا حساب لگانا کہ ڈیٹا تھوڑا سا کلسٹر ڈ یا پھیلا ہوا ہے، جو مماثلت کا بھی تعین کرتا ہے۔
- ماضی کے رویے کی بنیاد پر مستقبل کی پیشین گوئیاں کرنا
- کسی تجربے سے مفروضے کی جانچ کرنا

ایسے کئی طریقے ہیں جن سے کاروباری اپنے فائدے کے لیے شماریاتی تجزیہ استعمال کر سکتے ہیں۔ ان میں سے کچھ طریقوں میں یہ شناخت کرنا شامل ہے کہ آپ کے سیلز اسٹاف میں کون خراب کارکردگی کا مظاہرہ کر رہا ہے، کسٹمر ڈیٹا میں رجحانات تلاش کرنا، سرفہرست آپریٹنگ پروڈکٹ لائنوں کو کم کرنا، مالیاتی آڈٹ کرنا، اور ملک کے مختلف خطوں میں سیل کی کارکردگی کس طرح مختلف ہو سکتی ہے اس کی بہتر تفہیم حاصل کرنا شامل ہے۔

کاروبار میں کسی دوسری چیز کی طرح، کاروباری تجزیات میں بھی ایک عمل شامل ہے۔ کاروباری تجزیات میں منظم، اور مرحلہ وار کارروائیاں شامل کرنے کی ضرورت ہے تاکہ کم از کم تضادات کے ساتھ آخر میں سب سے زیادہ بہتر نتیجہ حاصل کیا جاسکے۔

کاروباری تجزیات میں شامل اقدامات:

- فریمنگ بزنس پرابلم: اس مرحلے میں، ہم بنیادی طور پر یہ معلوم کرتے ہیں کہ ہم کس کاروباری مسئلے کو حل کرنے کی کوشش کر رہے ہیں، مثلاً، جب ہم یہ معلوم کرنے کی کوشش کر رہے ہیں کہ سپلائی چین اتنی مؤثر کیوں نہیں ہے جیسا کہ ہونا چاہیے یا ہم فروخت کیوں کھو رہے ہیں۔ یہ بحث نارمل طور پر اسٹیک ہولڈرز کے ساتھ اس وقت ہوتی ہے جب انہیں کاروبار کے کسی بھی حصے میں غیر موثر ہونے کا احساس ہوتا ہے۔
- پرابلم فریمنگ: ایک بار ہمارے پاس مسئلہ بیان ہو جانے کے بعد، ہمیں اس کے بعد سوچنے کی ضرورت ہے کہ اس کاروباری تجزیاتی مسئلے کے لیے تجزیات کیسے کیے جا سکتے ہیں۔ یہاں، ہم میٹرکس اور مخصوص نکات تلاش کرتے ہیں جن کا ہمیں تجزیہ کرنے کی ضرورت ہے۔
- ڈیٹا : جس لمحے ہم اس مسئلے کی نشاندہی کرتے ہیں کہ کس چیز کا تجزیہ کرنے کی ضرورت ہے، اگلی چیز جس کی ہمیں ضرورت ہے وہ ڈیٹا ہے، جس کا تجزیہ کرنے کی ضرورت ہے۔ اس مرحلے میں، ہم نہ صرف ڈیٹا کے مختلف ذرائع سے ڈیٹا حاصل کرتے ہیں بلکہ ڈیٹا

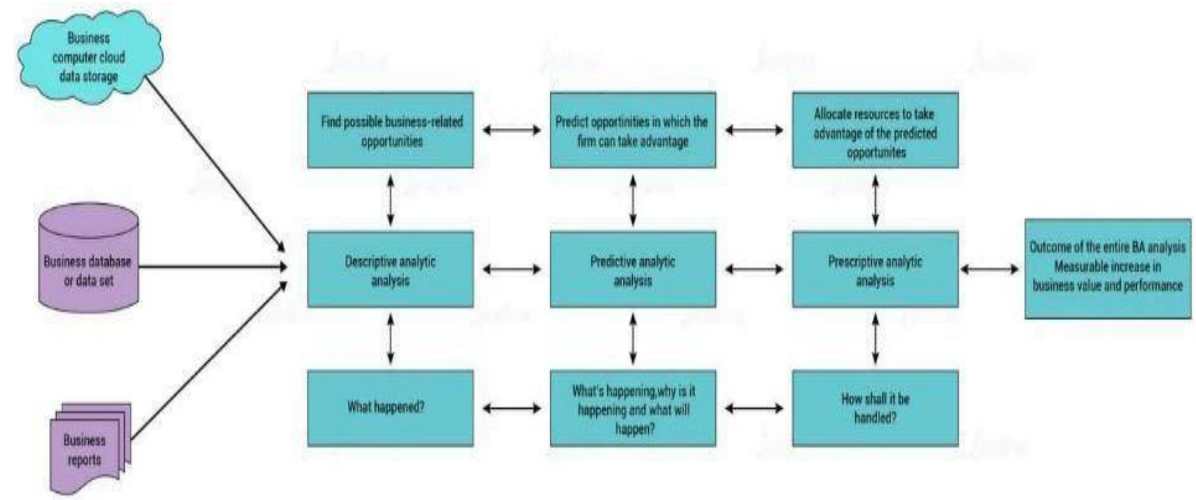
کو صاف بھی کرتے ہیں۔ اگر خام ڈیٹا کرپٹ ہے یا اس میں غلط اقدار ہیں، تو ہم ان مسائل کو دور کرتے ہیں اور ڈیٹا کو قابل استعمال شکل میں تبدیل کر دیتے ہیں۔

- طریقہ کار کا انتخاب اور ماڈل کی ترقی: ڈیٹا تیار ہونے کے بعد، مشکل مرحلہ شروع ہو جاتا ہے۔ اس مرحلے پر، ہمیں یہ طے کرنے کی ضرورت ہے کہ کون سے طریقے استعمال کیے جائیں اور کون سے میٹرکس اہم ہیں۔ اگر ضرورت ہو تو، ٹیم کو مخصوص طریقوں کا پتہ لگانے کے لیے اپنی مرضی کے ماڈل بنانا ہوں گے جو متعلقہ آپریشنز کے لیے موزوں ہوں۔ کئی بار، ہمارے پاس ڈیٹا کی قسم بھی اس طریقہ کار کا حکم دیتی ہے جسے کاروباری تجزیات کرنے کے لیے استعمال کیا جا سکتا ہے۔ زیادہ تر تنظیمیں متعدد ماڈلز بناتی ہیں اور فیصلہ کردہ اہم میٹرکس کی بنیاد پر ان کا موازنہ کرتی ہیں۔
- تعینات : ماڈل کے انتخاب اور حل کے لیے ڈیٹا کا تجزیہ کرنے کے شماریاتی طریقوں کو پوسٹ کریں، اگلا کام جو ہمیں کرنے کی ضرورت ہے وہ ہے حل کو حقیقی وقت کے منظر نامے میں جانچنا۔ اس کے لیے، ہم ڈیٹا پر ماڈلز لگاتے ہیں اور مختلف قسم کی جانکاری تلاش کرتے ہیں۔ میٹرکس اور ڈیٹا ہائی لائٹس کی بنیاد پر، ہمیں اپنے مسئلے کو حل کرنے اور کسی حل کو مؤثر طریقے سے نافذ کرنے کے لیے بہترین حکمت عملی طے کرنے کی ضرورت ہے۔ کاروباری تجزیات کے اس مرحلے میں بھی، ہم متوقع آؤٹ پٹ کا حقیقی وقت کے آؤٹ پٹ سے موازنہ کریں گے۔ بعد میں، اس کی بنیاد پر، ہم فیصلہ کریں گے کہ کیا حل کو دہرانے اور اس میں ترمیم کرنے کی ضرورت ہے یا کیا ہم اس پر عمل درآمد جاری رکھ سکتے ہیں۔

کاروباری تجزیات کا عمل

کاروباری تجزیات کے عمل میں سوالات پوچھنا، ڈیٹا کو دیکھنا، اور مطلوبہ جوابات تلاش کرنے کے لیے اس میں ہیرا پھیری شامل ہوتی ہے۔ اب، ہر تنظیم کے پاس اس عمل کو انجام دینے کے مختلف طریقے ہیں کیونکہ یہ تمام تنظیمیں مختلف شعبوں میں کام کرتی ہیں اور اپنے مخصوص کاروباری ماڈل کی بنیاد پر مختلف میٹرکس کو دوسروں سے زیادہ اہمیت دیتی ہیں۔

چونکہ مختلف تنظیموں کے لیے کاروبار کا نقطہ نظر مختلف ہوتا ہے، اس لیے ان کے حل اور حل تک پہنچنے کے طریقے بھی مختلف ہوتے ہیں۔ بہر حال، وہ تمام اعمال جو وہ کرتے ہیں ان کی درجہ بندی اور ان کے نقطہ نظر کو سمجھنے کے لیے نارمل کیا جا سکتا ہے۔ ذیل میں دی گئی تصویر ایک فرم کے کاروباری تجزیات کے عمل کے مراحل کو ظاہر کرتی ہے:



بزنس اینالیٹکس لائف سائیکل کے چھ مراحل ہیں:

مرحلہ 1: مسئلے کی نشاندہی کرنا

عمل کا پہلا مرحلہ کاروباری مسئلے کی نشاندہی کرنا ہے۔ مسئلہ ایک حقیقی بحران ہو سکتا ہے۔ یہ کاروباری ضروریات کو تسلیم کرنے یا موجودہ عمل کو بہتر بنانے سے متعلق کچھ بھی ہو سکتا ہے۔ کاروباری تجزیات میں یہ ایک اہم مرحلہ ہے کیونکہ یہ واضح طور پر سمجھنا ضروری ہے کہ متوقع نتیجہ کیا ہونا چاہیے۔ جب مطلوبہ نتائج کا تعین کیا جاتا ہے، تو اسے مزید چھوٹے مقاصد میں تقسیم کیا جاتا ہے۔ پھر، کاروباری اسٹیک ہولڈرز مسئلے کو حل کرنے کے لیے درکار متعلقہ ڈیٹا کا فیصلہ کرتے ہیں۔ اس مرحلے میں کچھ اہم سوالات کا جواب دینا ضروری ہے، جیسے: کس قسم کا ڈیٹا دستیاب ہے؟ کیا کافی ڈیٹا ہے؟ وغیرہ وغیرہ۔

مرحلہ 2: ڈیٹا کی تلاش

ایک بار مسئلہ کے بیان کی وضاحت ہو جانے کے بعد، اگلا مرحلہ ڈیٹا اکٹھا کرنا ہے (اگر ضرورت ہو) اور، زیادہ اہم بات یہ ہے کہ ڈیٹا کو صاف کرنا — زیادہ تر تنظیموں کے پاس کافی مقدار میں ڈیٹا ہو گا، لیکن تمام ڈیٹا پوائنٹس درست یا مفید نہیں ہوں گے۔ تنظیمیں مختلف طریقوں سے بہت زیادہ ڈیٹا اکٹھا کرتی ہیں، لیکن بعض اوقات، ڈیٹا سیٹ میں جنک ڈیٹا یا خالی ڈیٹا پوائنٹس موجود ہوتے ہیں۔ ڈیٹا کے یہ ناقص ٹکڑے تجزیہ کو روک سکتے ہیں۔ اس لیے اس ڈیٹا کو صاف کرنا بہت ضروری ہے جس کا تجزیہ کرنا ہے۔

ایسا کرنے کے لیے، آپ کو گمشدہ ڈیٹا کی گنتی کرنی ہو گی، آؤٹ لیرز کو ہٹانا ہو گا، اور دوسرے متغیرات کے مجموعہ کے طور پر نئے متغیرات تلاش کرنا ہوں گے۔ آپ کو ٹائم سیریز کے گراف بھی بنانے کی ضرورت پڑ سکتی ہے کیونکہ وہ نارمل طور پر پیٹرن اور آؤٹ لیرز کی نشاندہی کرتے ہیں۔ آؤٹ لیرز کو ہٹانا بہت ضروری ہے کیونکہ وہ آپ کے بنائے ہوئے ماڈل کی درستگی پر بہت زیادہ اثر ڈال سکتے ہیں۔ مزید یہ کہ، ڈیٹا کو صاف کرنے سے آپ کو ڈیٹا سیٹ کا بہتر احساس حاصل کرنے میں مدد ملتی ہے۔

مرحلہ 3: تجزیہ

ڈیٹا تیار ہونے کے بعد، اگلا کام اس کا تجزیہ کرنا ہے اس مرحلہ کو انجام دینے کے لیے یا جس معلومات کی تلاش کر رہے ہیں ان کو تلاش کرنے کے لیے مختلف قسم کے شماریاتی طریقے (جیسے مفروضے کی جانچ، ارتباط وغیرہ) کی مدد لے سکتے ہیں یا وہ تمام مرحلہ استمال کے جا سکتے ہیں جن کے لیے آپ کے پاس ڈیٹا موجود ہے۔

تجزیہ کرنے کا بنیادی طریقہ ہدف متغیر کے گرد محور ہے، لہذا آپ کو ان عوامل کو مد نظر رکھنا ہو گا جو ہدف متغیر کو متاثر کرتے ہیں۔ اس کے علاوہ بہت سے مفروضوں پر بھی غور کیا جاتا ہے کہ اس کے نتائج کیا ہو سکتے ہیں۔ مرمل طور پر، اس مرحلے پر، ڈیٹا کاٹ دیا جاتا ہے، اور موازنہ کیا جاتا ہے۔ ان طریقوں کے ذریعے، قابل عمل معلومات حاصل کیا جا سکتا ہے۔

مرحلہ 4: پیشن گوئی اور اصلاح

وہ دن گئے جب تجزیات کو رد عمل کے لیے استعمال کیا جاتا تھا۔ آج کے دور میں، کاروباری تجزیات فعال ہونے کے بارے میں ہے۔ اس مرحلے میں، آپ ڈیٹا کو ماڈل کرنے کے لیے پیشین گوئی کی تکنیک، جیسے نیورل نیٹ ورکس یا فیصلے کے درختوں کا استعمال کریں گے۔ پیشین گوئی کی یہ تکنیکیں متغیرات کے درمیان پوشیدہ بصیرت اور تعلقات کو تلاش کرنے میں آپ کی مدد کریں گی، جس سے آپ کو انتہائی اہم میٹرکس پر پیٹرن کو کھولنے میں مزید مدد ملے گی۔ اصولی طور پر، بہت سارے ماڈل بیک وقت استعمال کیے جاتے ہیں، اور سب سے زیادہ درستگی والے ماڈلز کا انتخاب کیا جاتا ہے۔ اس مرحلے میں، بہت ساری شرائط کو بھی پیرامیٹرز کے طور پر چیک کیا جاتا ہے، اور بہت سے سوالات کے جوابات فراہم کیے جاتے ہیں۔

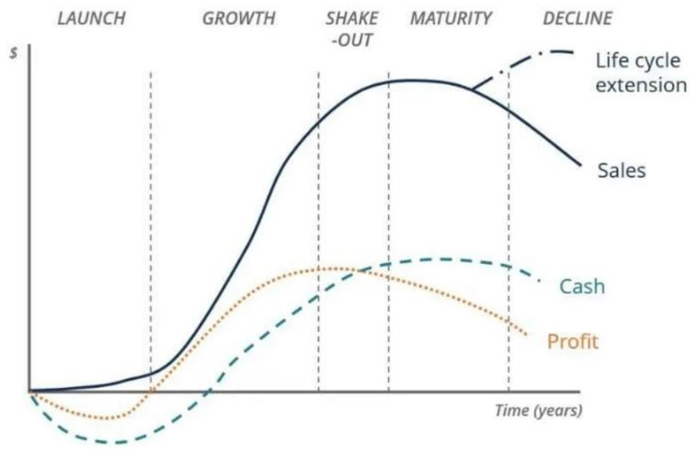
مرحلہ 5: فیصلہ کرنا اور نتیجہ کا اندازہ لگانا

ہدف کے متغیرات پر بنائے گئے اپنے ماڈل سے آپ کو حاصل ہونے والی بصیرتوں سے، تنظیم کے اہداف اور توقعات کو پورا کرنے کے لیے اس مرحلے میں ایک قابل عمل لائحہ عمل قائم کیا جائے گا۔ اس کے بعد کارروائی کا مذکورہ منصوبہ کام پر لایا جاتا ہے، اور انتظار کی مدت شروع ہوتی ہے۔ آپ کو اپنی پیشین گوئیوں کے حقیقی نتائج دیکھنے اور یہ جاننے کے لیے انتظار کرنا پڑے گا کہ آپ اپنی کوششوں میں کتنے کامیاب رہے۔ ایک بار جب آپ نتائج حاصل کر لیتے ہیں، آپ کو ان کی پیائش اور اندازہ کرنا پڑے گا۔

مرحلہ 6: اصلاح اور اپ ڈیٹ کرنا

حل کے نفاذ کے بعد، نتائج کی پیائش اوپر بیان کی گئی ہے۔ اگر آپ کو کچھ ایسے طریقے ملتے ہیں جن کے ذریعے عمل کی منصوبہ بندی کو بہتر بنایا جا سکتا ہے، تو ان پر عمل درآمد کیا جا سکتا ہے۔ اگر ایسا نہیں ہے، تو آپ پورے عمل کے نتائج کو رجسٹر کرنے کے ساتھ آگے بڑھ سکتے ہیں۔ یہ قدم مستقبل میں کسی بھی تجزیات کے لیے بہت اہم ہے کیونکہ آپ کے پاس ہمیشہ بہتر ہونے والا ڈیٹابیس ہو گا۔ اس ڈیٹابیس کے ذریعے، آپ زیادہ سے زیادہ قریب اور قریب جا سکتے ہیں۔

اصلاح اس مرحلے میں،(ROI سرمایہ کاری پر واپسی) کا اندازہ لگانا بھی ضروری ہے۔ کاروباری تجزیات کے لائف سائیکل کے اوپر دیے گئے خاکے پر ایک نظر ڈالیں۔



Business Life Cycle - Understanding the 5 Different StagesSource:

ڈیٹا پرائیویسی کی تعریف اور اہمیت

ڈیٹا پرائیویسی، جسے انفارمیشن پرائیویسی بھی کہا جاتا ہے، ذاتی ڈیٹا کی مناسب ہینڈلنگ، پروسیسنگ اور اسٹوریج سے مراد ہے تاکہ یہ یقینی بنایا جا سکے کہ افراد اپنی معلومات پر کنٹرول برقرار رکھیں۔ ذاتی ڈیٹا میں قابل شناخت معلومات کی ایک وسیع رینج شامل ہو سکتی ہے جیسے کہ نام، پتے، فون نمبر، ای میل ایڈریس، سوشل سیکیورٹی نمبر، اور یہاں تک کہ براؤزنگ رویے یا آن لائن خریداری جیسے ڈیجیٹل فوٹ پرنٹس۔

آج کے ڈیجیٹل دور میں، ڈیٹا بے مثال شرح سے تیار کیا جاتا ہے، اور یہ ڈیٹا اکثر تنظیموں کے ساتھ شیئر یا جمع کیا جاتا ہے۔ درست اقدامات کے بغیر، اس ذاتی معلومات کا غلط استعمال کیا جا سکتا ہے، جس سے شناخت کی چوری، نگرانی، یا اعتماد میں کمی واقع ہو سکتی ہے۔ ڈیٹا پرائیویسی کی اہمیت کو بڑھا چڑھا کر پیش نہیں کیا جا سکتا: یہ ایک بنیادی حق کے طور پر کام کرتا ہے جو افراد کو اپنی ذاتی معلومات پر خود مختاری برقرار رکھنے میں مدد کرتا ہے۔ ڈیٹا کی رازداری کا تحفظ افراد کو اس بات میں محفوظ محسوس کرنے کی اجازت دیتا ہے کہ ان کی ذاتی معلومات کو کس طرح جمع کیا جاتا ہے، ذخیرہ کیا جاتا ہے اور اس کا اشتراک کیا جاتا ہے، اور یہ صارفین اور کاروبار کے درمیان اعتماد کا ماحول پیدا کرتا ہے۔

کاروباری نقطہ نظر سے، ڈیٹا کی رازداری کو برقرار رکھنا بھی مضبوط کسٹمر تعلقات بنانے کے لیے بہت ضروری ہے۔ پرائیویسی کے خطرات کے بارے میں بڑھتی ہوئی بیداری کے ساتھ، صارفین ان کمپنیوں کے ساتھ مشغول ہونے کا زیادہ احتمال رکھتے ہیں جو اپنے ڈیٹا کا احترام کرتی ہیں اور اسے ذمہ داری سے ہینڈل کرتی ہیں۔ ڈیٹا کی خلاف ورزی یا ذاتی معلومات کو غلط طریقے سے ہینڈل کرنے کے نتیجے میں نہ صرف بھاری جرمانے ہو سکتے ہیں بلکہ اس سے شہرت کو بھی اہم نقصان پہنچ سکتا ہے۔

ریگولیٹری لینڈ سکیپ

حالیہ برسوں میں ڈیٹا پرائیویسی کے لیے ریگولیٹری منظر نامے میں نمایاں طور پر ترقی ہوئی ہے، دنیا بھر کی حکومتوں نے افراد کے ذاتی ڈیٹا کی حفاظت کی ضرورت کو تسلیم کیا ہے۔ کاروبار اور صارفین پر یکساں اثرات کے لیے چند نمایاں ضابطے ہیں:

• جنرل ڈیٹا پروٹیکشن ریگولیشن (GDPR): مئی 2018 سے یورپی یونین (EU) میں نافذ، GDPR نے ڈیٹا کی رازداری کے لیے ایک نیا عالمی معیار قائم کیا ہے۔ یہ شفافیت، رضامندی، اور جوابدہی پر زور دیتا ہے، جس کے لیے کاروبار کو واضح ہونا چاہیے کہ وہ کیسے جمع کرتے ہیں،

استعمال کریں، اور ذاتی ڈیٹا کی حفاظت کریں۔ GDPR کے تحت، افراد کو اپنے ڈیٹا تک رسائی حاصل کرنے، تصحیح کی درخواست کرنے، یا اسے حذف کرنے کا مطالبہ کرنے کا حق ہے۔ GDPR کی عدم تعمیل کے نتیجے میں 20€ ملین یا سالانہ عالمی کاروبار کا 4%، جو بھی زیادہ ہو، جرمانہ ہو سکتا ہے۔

• دیگر بین الاقوامی فریم ورک: کینیڈا (PIPEDA کے ساتھ)، برازیل (LGPD کے ساتھ) اور آسٹریلیا (اس کے پرائیویسی ایکٹ کے ساتھ) جیسے ممالک نے بھی ڈیٹا پرائیویسی کے قوانین نافذ کیے ہیں جن پر عمل درآمد کرنے کے لیے کاروبار کی ضرورت ہوتی ہے۔ ذاتی ڈیٹا کی حفاظت کے لیے اقدامات۔ اگرچہ ہر ضابطے کی تفصیلات مختلف ہوتی ہیں، وہ سبھی کاروباروں کے لیے ذاتی ڈیٹا کو شفاف اور ذمہ داری سے سنبھالنے کی ضرورت پر زور دیتے ہیں۔

ان ضوابط نے کمپنیوں کے کام کرنے کے طریقے کو تبدیل کر دیا ہے، انہیں اس بات پر دوبارہ غور کرنے پر مجبور کیا ہے کہ وہ کس طرح ذاتی ڈیٹا اکٹھا، ذخیرہ اور استعمال کرتی ہیں۔ شفافیت، رضامندی، اور ڈیٹا کو کم کرنا اب کسی بھی ذمہ دار ڈیٹا حکمت عملی کے بنیادی اصول ہیں، اور کاروباری اداروں کو تعمیل کو یقینی بنانے کے لیے مسلسل موافقت کرنی چاہیے۔

ڈیٹا پرائیویسی کو لاگو کرنے کے چیلنجز اگرچہ GDPR اور CCPA جیسے ضابطے ذاتی ڈیٹا کی حفاظت کے لیے واضح رہنما خطوط مرتب کرتے ہیں، ڈیٹا پرائیویسی کے طریقوں پر عمل درآمد کاروبار کے لیے اہم چیلنجز کا باعث بنتا ہے۔ قانون کی اسi تعمیل ہمیشہ سیدھی نہیں ہوتی بلکہ تنظیموں کو ایک پیچیدہ منظر نامے سے رو بہ رو کرتی ہے تاکہ یہ یقینی بنایا جاسکے کہ وہ جدت کو دبائے بغیر قانونی تقاضوں کو پورا کرتے ہیں۔

• جدت کے ساتھ رازداری کا توازن: بہت سی تنظیموں کے لیے، ڈیٹا ایک قیمتی اثاثہ ہے جو جدت کو ہوا دیتا ہے۔ کاروبار مصنوعات کو بہتر بنانے، گاہک کے تجربات کو ذاتی بنانے اور ترقی کو بڑھانے کے لیے ڈیٹا کا فائدہ اٹھاتے ہیں۔ تاہم، انہیں جدت کے لیے ڈیٹا کے استعمال اور اس بات کو یقینی بنانے کے درمیان ایک نازک توازن برقرار رکھنا چاہیے کہ وہ افراد کے رازداری کے حقوق کا احترام کرتے ہیں۔ ذاتی ڈیٹا کی حفاظت میں ناکامی مہنگے جرمانے کا باعث بن سکتی ہے، جبکہ حد سے زیادہ پابندی والی ڈیٹا پالیسیاں کمپنی کی اختراع کرنے کی صلاحیت کو روک سکتی ہیں۔

•رضامندی کاانتظام کرنا: ڈیٹاپرائیویسی کے زیادہ تر ضوابط کی بنیاد صارفین کاڈیٹااکٹھا کرنے سے پہلے ان سے رضامندی حاصل کرنا ہے۔ اگرچہ نظریہ میں یہ آسان لگتا ہے، عملی طور پر، رضامندی کاانتظام کرناایک پیچیدہ کام ہے۔ تنظیموں کو واضح طور پر بتاناچاہیے کہ ڈیٹاکااستعمال کیسے کیاجائے گا،اس بات کو یقینی بناناچاہیے کہ رضامندی کو مطلع کیاجائےاور آزادانہ طور پر دیاجائے،اور صارفین کو کسی بھی وقت رضامندی واپس لینے کے اختیارات فراہم کریں۔مزید برآں، کاروباری اکثر مختلف خطوں سے جمع کیے گئے ڈیٹاسے نمٹتے ہیں، ہر ایک کے اپنے رازداری کے قوانین کے ساتھ، رضامندی کے انتظام کے عمل کو مزید پیچیدہ بنا. دیتے ہیں

ڈیٹا ایتھکس(Data Ethics)

ڈیٹااخلاقیات سے ہماراکیامطلب ہے؟

"ٹیکلاش' کی اصطلاح کو تیار کیے ہوئے تقریباً تین سال ہو چکے ہیں: ایک اصطلاح جو ٹیکنالوجی اور ڈیٹاکے وسیع استعمال کے بارے میں بڑھتی ہوئی عالمی تشویش کو بیان کرنے کے لیے استعمال ہوتی ہے۔ مصنوعی ذہانت (AI) اور نئی ابھرتی ہوئی ٹیکنالوجی کے عروج کے ساتھ، یہ کبھی بھی اتنااہم نہیں تھاجتناکہ اب بینکوں اور بیمہ کنندگان کے لیے صارفین کے ساتھ اعتماد کااپناسماجی معاہدہ بنانااور برقرار رکھنا ہے۔ یہ وہ جگہ ہے جہاں ڈیٹاکی اخلاقیات عمل میں آتی ہیں، سوالات پوچھتے ہیں اور ڈیٹاکے استعمال کے بارے میں عملی اقدامات کی رہنمائی کرتے ہیں اور یہ سمجھتے ہیں کہ اسے " غلط فہمی، غلط انتظام، یاغلط استعمال کیاجاسکتا ہے۔

جیساکہ ہم نے کہاہے، ڈیٹااخلاقیات کی کوئی واحد متفقہ تعریف نہیں ہے۔ درحقیقت، گوگل پر اس وقت 'ڈیٹاایتھکس ڈیفینیشن' کے لیے 315 ملین تلاش کے نتائج ہیں۔ان نتائج میں بے شمار تعریفیں ہیں جن میں حد سے زیادہ آسان جملوں سے لے کر فلسفیانہ مضامین تک، فریم ورک، ماڈل، سوچ کی قیادت، اور کیس اسٹڈیز شامل ہیں جو کہ مخصوص تنظیمی باریکیوں کو حل کرنے کے لیے یاتو بہت زیادہ ونیلاہیں یا کسی بھی عملی قدر کے لیے ضرورت سے زیادہ نظریاتی ہیں۔

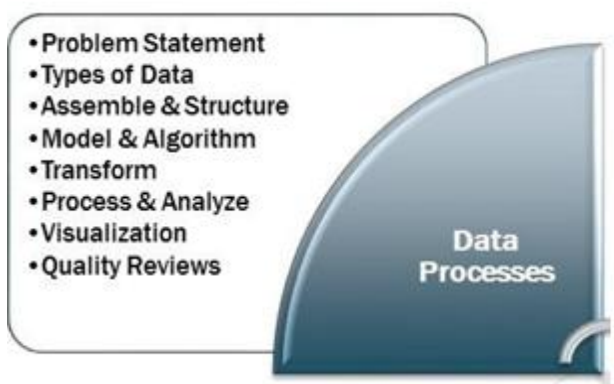
ڈیٹااخلاقیات صرف AI اخلاقیات سے زیادہ ہے؛ یہ اس بارے میں ہے کہ ڈیٹالائف سائیکل میں معلومات کے استعمال کے بارے میں قیمتی فیصلے کس طرح کاروبار کے اخلاقی اسپیکٹرم کے ساتھ تنورملل کرتے ہیں۔ حقیقت یہ ہے کہ ہر کاروبار کا اخلاقی اسپیکٹرم دوسرے کاروبار سے مختلف ہوگا۔اس کے بارے میں اس طرح سوچیں، آپ کاذاتی اخلاقی ضابطہ اخلاق اور اقدار آپ کے ہر دوست، خاندان اور ساتھیوں سے مختلف ہیں؛ کچھ پہلو وسیع پیمانے پر مختلف اور دوسرے ایک جیسے ہو سکتے ہیں، لیکن پھر بھی آپ کی اپنی ذاتی ترجیحات کے ساتھ۔اسی طرح ہر کاروبار پر لاگو ہوتا ہے۔

تاہم، اخلاقیات ہمیشہ اخلاقی مسائل کا صحیح جواب فراہم نہیں کرتی ہیں۔ اخلاقیات - تصورات کے طور پر - ہماری جلد کے نیچے آتے ہیں، اور وہ اکثر جذباتی مسائل سے متعلق ہوتے ہیں، جیسے کہ سیاست اور ذاتی عقائد۔ڈیٹااخلاقیات مختلف نہیں ہے۔

اعداد و شمار کی اخلاقیات کو عملی طور پر کسی تنظیم کے لیے مفید بنانے کے لیے، اسے 'کیا نہیں کرنا چاہیے' سے زیادہ ہونا چاہیے؛ یہ صرف ایک تعمیل ٹک باکس مشق نہیں ہو سکتی ہے جسے برا سمجھا جاتا ہے۔ اور یہ ہاں یا نہیں، صحیح یا غلط سوالات کی بائنری فہرست نہیں ہے۔ ڈیٹا کی اخلاقیات ان اہم سرمئی علاقوں کو تلاش کرنے کے بارے میں ہے جو آسانی سے قواعد وضوابط کی فہرست کے مطابق نہیں ہوتے ہیں، جب کہ یہ سمجھتے ہیں کہ ڈیٹا کا استعمال کرتے وقت تنظیموں کی زندگی کیا اہمیت رکھتی ہے۔ یہ بات چیت کو آپ سے تبدیل کرنے کے بارے میں ہے۔

آپ کو ڈیٹا کے ساتھ کیا کرنا چاہیے۔ یہ اس سے زیادہ اہم ہے کہ آپ ڈیٹا کے ساتھ کیا کر رہے ہیں، یہ اس بارے میں بھی ہے کہ یہ کون کر رہا ہے اور کیسے اور کب۔ ڈیٹا اخلاقیات کی زیادہ تر گفتگو اس بات پر مرکوز ہوتی ہے کہ کون سا ڈیٹا اکٹھا کیا جا رہا ہے اور اسے کیسے استعمال کیا جاتا ہے، اور کس کے ذریعے، لیکن ڈیٹا اخلاقیات کو وسیع تر اخلاقی سوالات (جیسے کہ یہ فیصلے پہلے کس کو کرنا ہیں۔ بینکوں کے لیے ڈیٹا اخلاقیات کا کیا مطلب ہے، ؟ یا اس کا کیا مطلب ہونا چاہیے، ؟ اور انشورنس فراہم کرنے والے صارفین کون ہیں ؟ وغیرہ) اٹھانے کی بھی ضرورت ہے

ڈیٹا اور ڈیجیٹل اختراع کو معاشی ترقی اور سماجی قدر کے ساتھ ساتھ وبائی امراض کے بعد بحالی کے لیے ایک محرک کے طور پر پیش کیا جا رہا ہے۔ بہت سے FIs اس بات کی کھوج کے ساتھ کہ وہ کس طرح ایک زیادہ ذمہ دار ریئل ٹائم آرگنائزیشن ہو سکتی ہے، ڈیٹا اخلاقیات ایک ایسا آلہ اور محافظ ریل ہو سکتا ہے جس کی ضرورت ڈیٹا کے مواقع کو اس قابل غور طریقے سے تیز کرنے کے لیے ہو سکتی ہے جو تنظیم کی اقدار کے ساتھ ہم آہنگ ہو۔



Source: five-different-data-processing-methods

مالیاتی ادارے ہر روز بڑی مقدار میں ڈیٹا وصول کرتی ہیں اور ان کا نظم کرتی ہیں۔ اس طرح کے ڈیٹا کا استعمال صارفین کی مدد کرنے، مصنوعات کی تشکیل، سروس کی ترقی، اور رسک کو منظم کرنے کے لیے کیا جاتا ہے، اور ڈیٹا کی اخلاقیات ان میں سے ہر ایک میں اہم کردار ادا کرتی ہے۔ تعمیل کی سرگرمیاں، مثال کے طور پر، ڈیٹا کے اخلاقیات کے اصولوں کا ایک مشترکہ سیٹ ہے کہ کیا، کہاں، کون، کیسے اور کیوں ڈیٹا کو کسی تنظیم کی اقدار سے ہم آہنگ کرنا کسی بھی شخص کے لیے بنیادی داؤ پر لگے گا۔

FI جو ڈیٹا کو قوت ضرب کے طور پر استعمال کرنا چاہتا ہے۔ اگر ڈیٹا کا درست اندازہ نہیں لگایا جاتا ہے، تو یہ دیکھ سکتا ہے کہ ایک جائز اور بے قصور گاہک قرضوں یا مجرمانہ سرگرمیوں سے انکار کر دیا گیا، اور انسانی اسمگلنگ کے متاثرین کو نظر انداز کر دیا گیا۔ یہ وہ جگہ ہے جہاں یہ سوال پیدا ہوتا ہے کہ کس قسم کے ڈیٹا کا انتظام کون کرتا ہے انتہائی اہم ہو جاتا ہے۔

2.8 اکتسابی نتائج (Learning Outcome)

طلباء نے ڈیٹا کے تصور اور کاروباری فیصلہ سازی میں اس کی اہمیت کے بارے میں مکمل فہم پیدا کیا ہے۔ انہوں نے ڈیٹا کی نمائندگی کے مختلف طریقے سیکھے ہیں جیسے کہ ٹیبل، چارٹ اور گراف، جو معلومات کو واضح اور مؤثر طریقے سے پیش کرنے میں مدد کرتے ہیں۔ طلباء اب ڈیٹا کے مختلف ذرائع، بنیادی اور ثانوی دونوں، اور بامعنی بصیرت پیدا کرنے میں ان کے کردار کو سمجھتے ہیں۔ وہ ڈیٹا کی ٹیبلیشن اور درجہ بندی کے عمل کی وضاحت کر سکتے ہیں، جو خام ڈیٹا کو ایک منظم اور پڑھنے کے قابل شکل میں ترتیب دیتا ہے۔ یونٹ نے ڈیٹا سائنس اور شماریات کے بنیادی اصولوں کو سمجھنے میں بھی ان کی مدد کی ہے، جس سے وہ ڈیٹا کا خلاصہ، تشریح اور تجزیہ کر سکیں۔ آخر میں، وضاحتی ڈیٹا تجزیات کے مطالعہ کے ذریعے، طلباء نے سیکھا ہے کہ کاروباری فیصلوں اور اسٹریٹجک منصوبہ بندی کی حمایت کرنے کے لیے ڈیٹا کے اندر پیٹرن، رجحانات اور تعلقات کی شناخت کیسے کی جائے۔

2.9 نمونہ امتحانی سوالات (Model Examination Questions)

درجہ ذیل سوالات کے مختصر ترین جوابات دیں.

1. ڈیٹا پر مبنی فیصلہ سازی کیا ہے، اور یہ کاروبار میں وجدان پر مبنی فیصلہ سازی سے کیسے مختلف ہے؟
2. ڈیٹا پر مبنی فیصلہ سازی کو جدید کاروباری حکمت عملی کا ایک اہم جز کیوں سمجھا جاتا ہے؟
3. کاروباری تجزیاتی ٹولز مارکیٹنگ، فنانس، اور آپریشنز جیسے شعبوں میں ڈیٹا پر مبنی فیصلہ سازی کی حمایت کیسے کر سکتے ہیں؟
4. ایک ایسے منظر نامے کی وضاحت کریں جہاں ڈیٹا کی ناقص تشریح کی وجہ سے کاروباری فیصلہ غلط ہوا۔ تجزیات سے نتیجہ کیسے بدل سکتا ہے؟
5. مارکیٹ کے غیر یقینی حالات میں پیش گوئی کرنے والے تجزیات اسٹریٹجک فیصلوں کے معیار کو کیسے بڑھا سکتے ہیں؟

مختصر جوابات کے حامل سوالات (Short Answers Type Questions)

1. ڈیٹا پر مبنی فیصلہ سازی کو لاگو کرتے وقت تنظیموں کو کن مشترکہ چیلنجوں کا سامنا کرنا پڑتا ہے، اور وہ ان پر کیسے قابو پا سکتی ہیں؟

2. ڈیٹا یا الگورتھم میں تعصب کس طرح کاروباری فیصلوں کو متاثر کر سکتا ہے، اور اس خطرے کو کم کرنے کے لیے کیا اقدامات کیے جا سکتے ہیں؟

3. فیصلہ سازی کے لیے ڈیٹا اکٹھا کرنے اور استعمال کرنے میں اخلاقی تحفظات پر بحث کریں۔

1. ایک کمپنی کس طرح انتظام کی تمام سطحوں پر ڈیٹا پر مبنی کلچر بنا سکتی ہے؟

2. مؤثر فیصلہ سازی میں ڈیٹا گورننس اور ڈیٹا کا معیار کیا کردار ادا کرتا ہے؟

طویل جوابات کے حامل سوالات (Long Answers Type Questions)

1. ایم بی اے کے طلباء انٹرن شپ یا لائیو پروجیکٹس کے دوران ڈیٹا پر مبنی فیصلہ سازی، اصولوں کا اطلاق کیسے کر سکتے ہیں؟

2. کاروباری تجزیات کے لائف سائیکل کے اہم مراحل کیا ہیں، اور ہر مرحلہ مؤثر فیصلہ سازی میں کس طرح حصہ ڈالتا ہے؟

3. اینالیٹکس لائف سائیکل کے آغاز میں مسئلہ کی تعریف تجزیاتی پروجیکٹ کے نتائج کو کیسے متاثر کر سکتی ہے؟

4. تجزیہ سے پہلے ڈیٹا کی صفائی اور تبدیلی کے لیے کون سی نارمل تکنیکیں استعمال کی جاتی ہیں، اور وہ کیوں ضروری ہیں؟

ایکسل، ایس کیو ایل، یا آر جیسے ٹولز کو ڈیٹا میں ہیرا پھیری کے لیے کیسے استعمال کیا جا سکتا ہے تاکہ بامعنی نمونوں اور رجحانات کو ننگا کیا جا سکے۔

اکائی 3۔ صلاحیت کی ترقی

(Analytics Capability Building)

Structure		اکائی کے اجزاء
Introduction	3.0	تمہید
Objectives	3.1	مقاصد
Capability Building	3.2	صلاحیت کی ترقی
Data Processing	3.3	ڈیٹاپراسیسنگ
A. Capacity Planning	3.4	صلاحیت کی منصوبہ بندی
Steps for Data Processing	3.5	مواد کی صلاحیت کا عمل
Capacity Matrix	3.6	صلاحیت میٹرکس
Data Driven Decision Making	3.7	ڈیٹاپر مبنی فیصلہ سازی
	3.8	اکتسابی نتائج
Model Examination Questions	3.9	نمونہ امتحانی سوالات

3.0 تمہید(Introduction)

صلاحیت کی ترقی ایک تصوراتی نقطہ نظر ہے جوان رکاوٹوں کو سمجھنے پر مرکوز ہے جو کسی تنظیم کو اس کے ترقیاتی اہداف کو حاصل کرنے سے روک سکتی ہیں مگران صلاحیتوں کو بڑھاتی ہیں جو اسے قابل پیمائش اور پائیدار نتائج حاصل کرنے کی اجازت دیتی ہیں۔ لکن شرٹ یہ ہے کہ آپ کی تنظیم کے لیے ایک گائیڈ یا فریم ورک ہو۔ یہ نتائج کو یقینی بنانے کے لیے تجزیاتی اقدامات کو کاروباری حکمت عملی کے ساتھ ہم آہنگ کرنے کی اہمیت پر بھی زور دیتا ہے۔ تجزیاتی صلاحیت کے کلیدی اجزاء کو سمجھ کر، تنظیمیں ڈیٹا کو ایک اسٹریٹجک اثاثہ میں تبدیل کر سکتی ہیں اور ایسی ثقافت کو فروغ دے سکتی ہیں جو ثبوت پر مبنی فیصلوں پر پروان چڑھتی ہے۔

3.1 مقاصد(Objectives)

اس باب کا مقصد طلباء کو یہ سمجھنے میں مدد کرنا ہے کہ فیصلوں کو مطلع کرنے کے لیے ڈیٹا کو کامیابی کے ساتھ استعمال کرنے کے لیے درکار تجزیاتی صلاحیتوں، وسائل، اور تنظیمی فریم ورک کو کیسے بنایا جائے۔ تجزیاتی صلاحیت کو فروغ دینے کے اہم عناصر — ڈیٹا کا بنیادی ڈھانچہ، ٹیکنالوجی کو اپنانا، ٹیلنٹ کی ترقی، اور کاروباری مقاصد کے ساتھ اسٹریٹجک صف بندی — اس باب کا موضوع ہیں۔ کارکردگی کو بہتر بنانے، مسابقتی برتری حاصل کرنے، اور ڈیٹا پر مبنی ثقافت کو فروغ دینے کے لیے کاروباری اداروں کے ذریعے تجزیاتی فریم ورک کی تخلیق اور استعمال طلباء کو سکھایا جائے گا۔ اس باب کے اختتام تک، سیکھنے والوں کو تجزیات کی پختگی کی سطحوں کا اندازہ کرنے، صلاحیت کے فرق کی نشاندہی کرنے، اور تنظیم کے اندر پائیدار تجزیاتی صلاحیتوں کو بنانے کے لیے حکمت عملی تیار کرنے کے قابل ہونا چاہیے۔

3.2 صلاحیت کی ترقی (Capability Building)

کسی فرد ،ٹیم یا ادارے کی صلاحیتوں کو بڑھانا تاکہ وہ اپنا کام بہتر ،تیزی سے اور کامیابی کے ہے کامطلب ہیںاس کا مطلب ہے کمپنی اور اسکے ملازمین کی صلاحیتوں کو مضبوط بنانا تاکہ (Business) ساتھ کر سکیں کاروبار کاروبار بہتر چلے اور زیادہ کامیاب ہو

3.3 ڈیٹاپراسیسنگ(Data Processing)

مندرجہ ذیل صلاحیت سازی کے خاکے میں، منصوبہ بندی، ٹیم کی ترقی، نظم ونسق اور تبدیلی کے انتظام کے چار کواڈرنٹس، اور ڈیٹا کے عمل ایک لازمی باہم مربوط فریم ورک کی نمائندگی کرتے ہیں جو آزادانہ طور پر اور ایک دوسرے سے باہم منحصر طور پر کام کرتا ہے۔ جیسا کہ ایک تنظیم تجزیاتی راستے پر سفر کرنے کی تیاری کرتی ہے، اس صلاحیت کی ترقی کے فریم ورک کے اندر بیان کردہ اہم اجزاء کو سمجھنے اور ان کی شناخت کو مفید اور مؤثر تجزیاتی مصنوعات کی فراہمی میں نمایاں طور پر مدد کرنی چاہیے۔

:Integration of cadastral survey data into building information modelsSource:

3.4 صلاحیت کی منصوبہ بندی

واضح طور پر، ہر وہ شخص جس نے H/HS پالیسیوں اور مشقوں کے شعبوں میں کام کیا ہے وہ اس سیاق وسباق کو سمجھنے کی قدر کی تعریف کرتا ہے جس میں وہ کام کرتے ہیں، موجودہ سرمایہ کاری سے فائدہ اٹھانے کی ضرورت، اور مشترکہ اہداف کے ساتھ اسٹیک ہولڈرز کی طرف سے بیان کردہ حکمت عملیوں پر استوار کرنا۔ اس کے علاوہ، تجزیات کے استعمال کے لیے ایک واضح نقطہ نظر جو اسی طرح اہداف، سامعین، دستیاب وسائل، فنانسنگ، اور پائیداری پر مرکوز ہے۔

:Integration of cadastral survey data into building information modelsSource:

کسی بھی نئے اقدام کی کامیابی یا موجودہ کی توسیع کے لیے اہم۔ ہم نے مندرجہ ذیل حصے میں ایک سادہ جدول بھی فراہم کیا ہے جو اس گائیڈ کے صارفین کو ان کلیدی عناصر میں سے ہر ایک کو واضح کرنے میں مدد کرے گا اور اتنا ہی اہم ہے کہ ان کی ضرورت کیوں ہے۔ اگرچہ پہلی نظر میں یہ اتنا ابتدائی معلوم ہوتا ہے کہ یہ غیر ضروری ہے، بہت سے تجزیاتی پروجیکٹ غلط قدموں پر اتر جاتے ہیں کیونکہ کسی نے بھی واضح طور پر یہ نہیں بتایا کہ قیادت کے لیے ان شرائط کا کیا مطلب ہے۔ اور، نتیجے کے طور پر، ٹیم کے ارکان مختلف مقاصد کا پیچھا کرتے ہیں. جبکہ

Capacity 1 جدول اوپر درج ہر عنصر کے لیے تعریفیں فراہم کرتا ہے، ہم تجویز کرتے ہیں کہ ریاستیں اور علاقے اپنی ٹیموں کے ساتھ کام کرنے کے لیے وقت نکالیں تاکہ ان عناصر میں سے ہر ایک، ان کی تعریفوں، اور ان کی ضرورت کیوں ہو۔ آپ کو وضاحت کی خاطر اضافی شامل کرنا مفید معلوم ہو سکتا ہے۔

صلاحیت کمیونٹی کنکشنز:

اس سے مراد اجتماعی صلاحیت اور کارکردگی کو بڑھانے کے لیے افراد، تنظیموں اور اداروں کے درمیان تعلقات کی تعمیر اور مضبوطی کا عمل ہے۔ یہ مشترکہ اہداف کے حصول کے لیے درکار مہارتوں، وسائل اور بنیادی ڈھانچے کو تیار کرنے کے لیے کمیونٹی کے اندر اشتراک، علم کے اشتراک اور باہمی تعاون پر زور دیتا ہے۔

تجزیات یا تنظیمی ترقی کے تناظر میں، اس کا مطلب ہے ایک ایسا نیٹ ورک بنانا جہاں مہارت، ڈیٹا، اور بہترین طریقوں کو محکموں یا شراکت داروں میں شیئر کیا جائے۔ اس طرح کے رابطے اجتماعی ذہانت سے فائدہ اٹھا کر مسلسل سیکھنے، اختراع اور پائیدار ترقی کو قابل بناتے ہیں۔

.صلاحیت ٹیم کی ترقی

جب کوئی ایجنسی تنظیم میں ڈیٹا اینالیٹکس کے استعمال کو شامل کرنے کے ساتھ آگے بڑھتی ہے تو، جو ٹیم جمع کی جا نی چاہیے وہ مختلف مہارتوں کے حامل افراد پر مشتمل ہونی چاہیے تاکہ صحیح مہارت اور تجربے کو اکٹھا کیا جا سکے۔ تنظیمی ڈھانچے کا کوئی ایک ماڈل معیاری نہیں ہے۔ بلکہ یہ ہر ایجنسی کے لیے منفرد ہے۔ تجزیات پر مرکوز ایک بین الضابطہ ٹیم کا استعمال تعاون کو شناخت کرنے کی اجازت دے کر کامیابی کا بہترین موقع فراہم کرتا ہے۔

:Integration of cadastral survey data into building information modelsSource:

درست سوالات اور ایسے ماڈلز بنانے کے لیے جو کاروبار کو استعمال کر سکیں۔ ٹیم کی ساخت اتنی لچکدار ہونی چاہیے کہ ضرورت کے مطابق مطلوبہ مہارتوں کے اضافے یا گھٹاؤ کی اجازت دی جا سکے۔

ٹیم بلڈنگ ٹیبل جو اس مقالے کے سیکشن 3 میں مندرجہ ذیل ہے ان ملازمتوں کے عنوانات، وضاحتوں، اور مناسب حکمت عملی کی ترقی کے لیے فعال مطابقت کی مزید وضاحت کرتا ہے جس پر غور کیا جانا چاہیے۔ اگرچہ یہ فہرست مکمل نہیں ہے، اس میں ایک ٹیم کے اراکین شامل ہیں

جو کسی بھی تجزیاتی حکمت عملی کی ترقی اور نفاذ کے لیے ضروری ہوں گے: شماریات دان، محققین، مالیاتی تجزیہ کار، کاروباری تجزیہ کار، ڈیٹا آرکیٹیکٹ، تکنیکی عملہ، پروگرام یا موضوع کے ماہرین، ایگزیکٹو یا مینیجر، اور دیگر اسٹیک ہولڈرز۔

صلاحیت گورننس اور تبدیلی کا انتظام

بہت سے ریاستی آئی ٹی سسٹم کی اوسط عمر 20 سال کے لگ بھگ ہے، اس بات کا قوی احتمال ہے کہ تجزیاتی جدید کاری کی کوششوں کے دوران ٹیکنالوجی کی کئی نسلیں چھلانگ لگ جائیں گی۔ آئی ٹی کی تبدیلی کی اس ڈگری کے لیے متاثرہ عملے اور کاروبار کو بااختیار بنانے کے لیے تنظیم اور عمل میں تبدیلی کے انتظام کی ضرورت ہوتی ہے۔

ایجنسیوں کے اندر آپریشن۔ جن اقدامات کی نشاندہی کی گئی ہے ان سے پراجیکٹ کے خطرے کو کم کرنا چاہیے، ایجنسی کو زیادہ سے زیادہ اپنانا چاہیے، اور سرمایہ کاری پر مناسب منافع حاصل کرنا چاہیے۔

:Integration of cadastral survey data into building information modelsSource:

جیسا کہ مندرجہ بالا سیکشن میں صلاحیت 3 جدول میں مزید وضاحت کی گئی ہے، ایک مؤثر تبدیلی کے انتظامی منصوبے کے عناصر تجزیات پر مبنی فیصلہ سازی کو فعال کرنے کے لیے ضروری ایجنسی کے کردار اور ذمہ داریوں کی وضاحت کریں گے۔ اس میں ایک آواز کے قیام کے ذریعے تجزیاتی وسائل کا اشتراک کرنے کے لیے سرپرست ایجنسیوں کو گورننس شامل ہے۔ (ڈیٹا) گورننس کا ڈھانچہ جو واضح فیصلہ سازی اور مواصلاتی منصوبے، پالیسیاں اور طریقہ کار فراہم کرتا ہے۔ اور حوصلہ افزائی

مخصوص مقاصد کے لیے اور قابل اطلاق ریاست اور وفاقی قانون اور ضوابط کے مطابق محفوظ ڈیٹا ایکسچینج کلچر کو فروغ دینا اور انفارمیشن ٹکنالوجی کی جدید کاری سے شروع ہونے والی کاروباری کارروائیوں کی جدت کو شروع کرنے کی ترغیب۔ یہ جانتے ہوئے کہ تبدیلی نارمل طور پر لوگوں کے لیے آسان نہیں ہوتی، تنظیمی تبدیلی کے انتظام (OCM) کے منصوبے کو تجزیاتی تبدیلی کے لیے کاروبار پر مبنی ضرورت کو فعال طور پر شناخت کرنے کی ضرورت ہوتی ہے، یہ منصوبہ ہر شخص کی ذمہ داریوں پر کیسے اثر انداز ہو سکتا ہے اور وہ آگے کی راہ پر گامزن ہونے کے لیے کس طرح مدد کر سکتے ہیں۔

صلاحیت 3 مثال:Commonwealth of Virginia's (CoV) Health and Human Resources

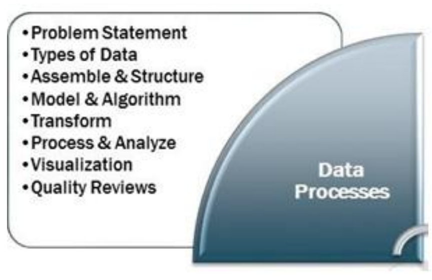
HHRA نے ایک Agency—Enhanced Memorandum of Understanding (eMOU) eMOU تیار کیا ہے جو ایک دوبارہ قابل استعمال ڈیٹا اوور نینس فریم ورک ہے جس کو مختلف مقاصد/پروجیکٹس کے لیے مختلف ڈیٹا اثاثوں کی شناخت کے طور پر بڑھایا جا سکتا ہے۔ ایجنسی نے وفاقی ڈیٹا کے استعمال اور باہمی تعاون کے معاہدے (DURSA) کو فریم ورک کے ماڈل کے طور پر استعمال کیا اور ریاست کے مخصوص، ایجنسی کے وسیع اہداف اور ڈھانچے کو ایڈجسٹ کرنے کے لیے اس میں ترمیم کی۔

HHR 11 ڈویژنوں پر مشتمل ہے، ہر ایک کا اپنا سیکیورٹی آفیسر ہے۔ تمام ڈویژن کے افسران کو ایک ہی صفحہ پر لانے کی پہل سیکریٹریٹ کے دفتر کے اندر ایک مرکزی سرشار عملے کے ذریعے کی گئی تھی تاکہ رفتار اور سمت کو اس بات پر مرکوز رکھا جا سکے کہ ڈیٹا شیئرنگ کو کیسے پورا کیا جائے گا بجائے اس کے کہ ایسا کیوں نہ ہو سکا۔ ای ایم او یو کی ترقی پر، اٹارنی جنرل کو ڈیٹا گورننس کے نئے فریم ورک کے استعمال پر غور کرنا اور اس کی توثیق کرنی پڑی۔

ڈویژنوں کے درمیان پوائنٹ ٹو پوائنٹ ایم او یوز کی تبدیلی کے طور پر۔ اس عمل کا ایک اہم فائدہ eMOU کی حقیقی ترقی کو ختم کرنا تھا کیونکہ اس نے شروع سے ان لوگوں کو شامل کیا اور پھر اسے سرکاری قانونی منظوری حاصل ہوئی۔ 8 جولائی 2014 تک، تمام 11 ڈائریکٹرز نے eMOU استعمال کرنے پر اتفاق کیا تھا۔

3.5 مواد کی صلاحیت کا عمل (Process of Data Capabilities)

صاف، معیاری ڈیٹا کا ہونا ڈیٹا اینالیٹکس کے لیے ایک اہم بنیاد ہے۔ اس طرح کے اعلیٰ معیار کے ڈیٹا کو حاصل کرنے کے لیے عمل کے آغاز سے شروع ہونے کی ضرورت ہے اور پہلے زیر بحث کی گئی صلاحیتوں کو استوار کرنا چاہیے۔ اس بات کا منصوبہ ہونا چاہیے کہ کون سے ڈیٹا کی ضرورت ہے جو لوگ جانتے ہیں۔



:Integration of cadastral survey data into building information modelsSource:

تکنیکی اور کاروباری نقطہ نظر سے ڈیٹا۔ ڈیٹا کے مالکان کے درمیان ڈیٹا کو کس طرح شیئر کیا جائے اس کے لیے ڈیٹا گورننس کا عمل قائم کیا جانا چاہیے۔ ڈیٹا کی سالمیت کو یقینی بنانے کے لیے، صحیح ڈیٹا سیٹ حاصل کرنے کے لیے ایک عمل ہونا چاہیے؛ اس کے ساتھ ساتھ ڈیٹا گودام جیسے مربوط تجزیاتی ماحول میں ڈیٹا کو کیسے نکالا جائے گا، تبدیل کیا جائے گا اور لوڈ کیا جائے گا۔ ڈیٹا کے عمل کو بہتر فیصلہ سازی، پروگرام کے نتائج کو

بہتر بنانے، اور بہتر خدمات کی فراہمی کے بنیادی مقاصد کی حمایت کرنی چاہیے۔ ایجنسیاں اکثر یہ کہتی ہیں کہ وہ ڈیٹا میں ڈوب رہے ہیں، لیکن ان کے پاس قابل استعمال اور قابل عمل معلومات نہیں ہیں۔ درست ڈیٹا کے عمل کو ملازمت دینے سے پالیسی سازوں اور پروگرام کے عملے کی تمام سطحوں کو فیصلے کرنے کے لیے بہتر طور پر مطلع کرنے کے لیے قابل عمل معلومات حاصل کرنے کی سہولت ملتی ہے۔

اہلیت 4 مثال: فلوریڈا اسٹیٹ فیملی نیٹ ورک (FSFN) کے اندر ڈیٹا اینالیٹکس کا استعمال کرتے ہوئے بچوں کی بہبود کے نظام کو بہتر بنانا Mindshare Technologies اور FSFN "فوری نگہداشت کے احتساب اور کیس ایسکلیشن "(ICARE) کا استعمال کر رہے ہیں، ایک تجزیاتی اور کیس مینجمنٹ حل، قابل نگہداشت کے انتظام کی مشق کو دوبارہ بنانے اور قابل عمل کوالٹی مینجمنٹ کو بہتر بنانے کے لیے۔

3.6 صلاحیت میٹرکس (Metrics of Capabilities)

یہ حصہ اس مقالے کے پچھلے حصوں میں بیان کردہ ہر ایک صلاحیت پر مزید تفصیل فراہم کرتا ہے۔ اہلیت 1—4 کے عناصر یا اجزاء کی شناخت اور تعریف کی گئی ہے۔ اس کے علاوہ، ان کی قدر ایک ایجنسی کی وسیع تجزیاتی حکمت عملی یا تجزیاتی منصوبے کے اندر فراہم کی جاتی ہے۔

1 صلاحیت : منصوبہ بندی		
موضوع	تعریف	ہے
مقصد اور سامعین کی شناخت کریں۔	پروجیکٹ کے دائرہ کار، اہم مقاصد، اور نتائج کے مطلوبہ وصول کنندگان کو قائم کرتا ہے۔	اسٹیک ہولڈرز اور دوسروں تک منصوبے کے اسٹریٹجک وژن اور بنیادی مقصد کا خلاصہ اور ان سے رابطہ کرتا ہے، اور زیادہ تر وقت کے ساتھ ساتھ اس میں کوئی تبدیلی نہیں ہوتی۔
مختصر اور طویل مدتی گول سیٹنگ	تکمیل کے لیے ٹائم فریم کے ساتھ کاموں اور سرگرمیوں کی شناخت، ترجیح، اور تفویض کرنے کا عمل۔ یہ آپ کو فعال کرتا ہے۔ مقاصد کے حصول کے لیے منصوبے کی حکمت عملی۔	جوہر میں، یہ مکمل ہونے والی سرگرمیوں کا ایک تفصیلی معاہدہ فراہم کرتا ہے۔ مختصر اور طویل مدتی اہداف کو بیان کرنے سے، نتائج پورے پروجیکٹ کے لائف سائیکل میں حاصل کیے جاتے ہیں اور اس سے اسٹیک ہولڈر کی دلچسپی برقرار رہتی ہے۔

کاروبار اور وسائل کی تشخیص	آپ کی ایجنسی کے اثاثوں اور کاروبار کا ایک جامع جائزہ پراجیکٹ کے آغاز اور دیکھ بھال میں مدد کے لیے فیصلہ سازی اور منصوبہ بندی میں مدد فراہم کرتا ہے۔ اثاثوں اور کاروباری ضروریات کا جائزہ لیا جاتا ہے جن میں نارمل طور پر شامل ہیں : عملہ، RFP/ معاہدے، پروکیورمنٹ، تکنیکی آلات، سافٹ ویئر، فن ترقی، دفتر کی جگہ، اور کاروباری کام جو متاثر ہوتے ہیں۔	کسی تنظیم کو متاثر کرنے والے فیصلے کرنے میں محتاط فیصلے کی منصوبہ بندی شامل ہوتی ہے۔ پراجیکٹس بیرونی عوامل)معاہدے/منظوریوں (پر منحصر ہو سکتے ہیں جن کے لیے تمام اسٹیک ہولڈرز کو بلانے سے پہلے منصوبہ بندی کی جانی چاہیے، بہت جلد بلانے سے اسٹیک ہولڈر کی دلچسپی ختم ہو سکتی ہے۔
بجٹ/لاگت مختص	منصوبہ بند کوششوں کی حمایت کے لیے فنڈنگ کے سلسلے کا تعین کرنا۔	پروجیکٹوں میں اکثر تاخیر ہوتی ہے، یا غیر ارادی اخراجات پروجیکٹ کے آغاز میں پیش نہیں ہوتے۔ ریزرو بجٹ مختص کرنے سے پروجیکٹ کا لائف سائیکل چلتا رہے گا۔
پائیداری / دوبارہ پریوست	دور اندیشی اور منصوبہ بندی کا طریقہ کار، وقت کے ساتھ ساتھ برداشت کرنے یا جاری رکھنے کی صلاحیت۔ خریداریوں کی منصوبہ بندی کریں اور کاروباری طرز عمل بنائیں جن میں کم سے کم دیکھ بھال یا لاگت کے ساتھ توسیع، اضافہ، یا آپریشنل کرنے کی صلاحیت شامل ہو۔	ایجنسیاں اثاثوں میں ڈالر کی بڑی سرمایہ کاری کر سکتی ہیں اور ایجنسیوں کو ایسے اثاثے حاصل کرنے چاہئیں جو اس کی تنظیم کی مستقبل کی ترقی کو پیش کرتے ہیں۔
صلاحیت :2 ٹیم بلڈنگ		
موضوع	**تعریف**	**کیوں ضرورت ہے**
شماریات دان	عملی اعداد و شمار، ریاضی، یا دونوں میں تربیت یافتہ اور تجربہ کار، شماریات دان حقیقی دنیا کے مسائل کو حل کرنے میں مدد کے لیے ڈیٹا اکٹھا کرنے اور تجزیہ کرنے کے لیے شماریاتی طریقے استعمال کرتا ہے۔ یہ قابل اعتماد ڈیٹا تیار کرنے، ماحول کو سمجھنے کے لیے ڈیٹا کا تجزیہ کرنے، اور ڈیٹا سے عملی نتائج اخذ کرنے کے عمل سے	ایک شماریات دان فیصلہ سازی کے عمل میں قابل مقدار اور سائنسی طور پر حمایت یافتہ نقطہ نظر لاتا ہے۔ کمپیوٹنگ طاقت کے بڑھتے ہوئے استعمال کے ساتھ، شماریاتی سرگرمیاں تیز تر اور زیادہ درست ہو گئی ہیں۔ وفاقی مالی اعانت سے چلنے والے منصوبوں کی منصوبہ بندی اور ان پر عمل درآمد کے لیے قابل مقدار جواز اکثر ضروری ہوتا ہے۔ شماریات دانوں

	حاصل ہوتا ہے۔اس کے علاوہ، شماریات دانوں کے پاس اکثر ڈومین مخصوص مہارت ہوتی ہے جو تجزیہ کی سرگرمی کے لیے اضافی بصیرت فراہم کرتی ہے، بشمول متعلقہ سوالات کو سمجھنا اور ممکنہ باہری یا بے ضابطگیوں سے واقفیت۔ شماریات دان ڈیٹا سیٹ کے اوپر الگورتھم لاگو کریں گے تاکہ مزید معلومات حاصل کی جا سکیں، جو کہ سادہ شماریا دیگر آپریشنل رپورٹنگ سے بالاتر ہے۔	کی مہارت تنظیموں کے ڈیٹا کو صاف کر کے اور اسے تجزیہ کے لیے قابل استعمال بنا کر ان کی مدد کر سکتی ہے۔
محقق	محققین دی گئی صنعت یا ڈومین کے علاقے میں جامع تحقیق کرتے ہیں۔ وہ ڈیٹا سیٹس کی شناخت، تجزیہ اور ترکیب کرتے ہیں) کبھی کبھی شماریات دان کے ساتھ مل کر(۔ محقق ایک ڈومین کے وسیع تر منظر نامے کو سمجھتا ہے، اسے اہم تعلیمی تحقیقی نتائج سے آگاہ کیا جاتا ہے، ثبوت پر مبنی طریقوں کے بارے میں علم ہوتا ہے، اور اکثر تعلیمی اداروں سے اچھی طرح سے جڑا ہوتا ہے۔	مالیاتی حکمت عملی تنظیم کے وژن کی حمایت کرتی ہے اور اسٹریٹجک منصوبہ بندی کے لیے کارکردگی کے نتائج کو زیادہ سے زیادہ کرنے میں مدد کرتی ہے۔ مالیاتی تجزیہ کار حصولی کے طریقوں کو نیویگیٹ کر کے اس بات کو یقینی بنانے کے لیے مدد کرتا ہے کہ موثر حل کی تکمیل ہو۔
مالیاتی تجزیہ کار	ایک مالیاتی تجزیہ کار رجحانات کی نشاندہی کرنے، پیشین گوئیاں تیار کرنے، اور لاگت سے فائدہ اور خطرے کا تجزیہ فراہم کرنے کے لیے تنظیم کے مالی ڈیٹا کا استعمال کرتے ہوئے بجٹ سازی اور اکاؤنٹنگ کے افعال فراہم کرتا ہے، جس سے قائدین کو سرمایہ کاری مؤثر حل بنانے میں رہنمائی کی جاتی ہے جو تنظیم کے اسٹریٹجک اہداف کو پورا کرتے ہیں۔ عملہ وفاقی، ریاستی اور مقامی آمدنی کے ذرائع، تبدیلی کے ممکنہ راستے، اثاثہ جات کے انتظام اور وسائل کی کوششوں کے بارے میں علم اور	معروضیت کا معروضی اور باخبر نقطہ نظر فراہم کرنے، میٹرکس کی شناخت اور نتائج کی پیائش، تحقیق اور تجزیے تک پہنچنے کے لیے معیاری طریقے قائم کرنے، سطحی مشاہدات سے بالاتر ہونے والی گہرائی کے ساتھ پیچیدہ حالات کے لیے وضاحت پیش کرنے کے ذریعے کوالٹی اور کوالٹیٹیو تحقیق کا استعمال توجہ کے شعبوں کو بہتر بنانے اور توجہ مرکوز کرنے کے لیے کیا جاسکتا ہے۔

	تفہیم میں حصہ ڈالتا ہے، اور ڈیٹا کو کارکردگی کے کلیدی اشاریوں اور نتائج پر مبنی اقدامات پر لاگو کرتا ہے۔	
کاروباری تجزیہ کا	ایک کاروباری تجزیہ کار تجزیاتی اصولوں اور طریقوں کا استعمال کرتے ہوئے ضروریات کی وضاحت کرنے اور اسٹیک ہولڈرز کو قدر فراہم کرنے والے حل تجویز کرنے کے ذریعے تنظیم میں تبدیلی کو قابل بناتا ہے۔ بہترین طرز عمل یہ حکم دیتے ہیں کہ کاروباری تجزیہ کار کو بھی موجودہ تنظیمی عمل کے مطابق ہونا چاہیے۔ بین الاقوامی انسٹی ٹیوٹ آف بزنس اینالیسس کے مطابق، کاروباری تجزیہ کار" علمی شعبوں "میں کام انجام دیتا ہے جس میں ایلیٹیشن، ضروریات کا انتظام اور مواصلات، انٹرپرائز تجزیہ، ضروریات کا تجزیہ، حل کی تشخیص اور توثیق، اور کاروباری تجزیہ کی منصوبہ بندی اور نگرانی شامل ہیں۔	کاروباری تجزیات کسی تنظیم کی شناخت شدہ پروجیکٹ کو سمجھنے یا فوائد کو تبدیل کرنے، لاگت سے بچنے، آٹومیشن اور بہتری کے نئے مواقع کی نشاندہی کرنے اور مطلوبہ صلاحیتوں کو سمجھنے میں مدد کر سکتے ہیں۔ کاروباری تجزیہ کار صحیح ڈیٹا تلاش کرنے یا استعمال کرنے کے لیے کاروبار اور IT کے درمیان ایک پل بھی بنا سکتا ہے۔
ڈیٹا آرکیٹیکٹ	ڈیٹا آرکیٹیکٹ کاروباری تجزیہ کاروں، سسٹمز ڈیزائنرز، اور پروجیکٹ ٹیموں کے ساتھ مل کر کام کرتا ہے تاکہ کاروباری ضروریات کو پورا کرنے کے لیے ڈیٹا بیس کے حل کو ڈیزائن کیا جا سکے۔ ڈیٹا آرکیٹیکٹ ڈیٹا کی منتقلی، توثیق، کلین اپ، میپنگ، اور ڈیٹا لغات سے وابستہ مسائل کے لیے بھی منصوبہ بندی کرے گا۔ ڈیٹا آرکیٹیکٹ انڈسٹری ڈیٹا ماڈلز، انٹرپرائز پلاننگ، اور ریگولیٹری ڈھانچے جیسے HIPAA، FERPA، Sarbanes Oxley، اور Basel II کے بارے میں علم رکھتا ہے۔	ایک اچھا ڈیٹا آرکیٹیکچر تنظیم کے وژن اور ڈیٹا فن ترقی اور کاروباری فن ترقی کے درمیان قریبی اتحاد کو سمجھتا ہے۔ پیچیدہ کاروباری ماحول اور ٹکنالوجی کے ماحول کی دنیا میں ڈیٹا انضمام یا مختلف ڈیٹا ذرائع کے باہمی تعاون کا استعمال کرتے ہوئے، ڈیٹا آرکیٹیکٹ کی پوزیشن کسی بھی ڈیٹا اینالیٹکس کوشش کی کامیابی کے لیے اہم ہے۔

ٹیکنکل سٹاف	تکنیکی عملہ، جیسا کہ رپورٹ لکھنے والے اور پروگرامرز، تکنیکی حل کی تبدیلیوں یا نفاذ کی حمایت کرتے ہیں جو ڈیٹا اینالیٹکس کی کوششوں کے نتیجے میں ہوتے ہیں۔ تکنیکی عملہ ڈیٹا انٹیگریشن کرنے کے لیے بہت اہم ہے کیونکہ وہ پیچیدہ ڈیٹا سسٹم لینے اور اسے تنظیم کے تجزیہ میں قابل استعمال بنانے میں مدد کر سکتے ہیں۔	ڈیٹا اینالیٹکس کی کوشش کا سب سے کامیاب نتیجہ ڈیٹا سے چلنے والی کارروائیوں اور فیصلوں کا ادارہ جاتی ہونا ہے۔ تکنیکی عملہ نتائج کو ویژولائزیشن ٹولز، ویب ایپلیکیشنز، اور رپورٹس میں شامل کرے گا جو ان معلوماتی ٹولز کو وسیع تر اپنانے اور اسٹیک ہولڈر کمیونٹی کو بااختیار بنانے کی اجازت دیتے ہیں۔ تکنیکی عملہ کاروباری تجزیہ کاروں اور دیگر عملے کو رپورٹ لکھنے اور ڈیٹا کی تفہیم کے علم کو منتقل کر کے قدر میں اضافہ کر سکتا ہے۔ وہ اعداد و شمار کے سیاق وسباق کو بھی لا سکتے ہیں جو تنظیم کے آپریشنل سسٹم سے باہر رہ سکتے ہیں۔
پروگرام کے موضوع کے ماہر	ایک پروگرام کے موضوع کا ماہر تنظیم کے اندر کسی ڈومین یا موضوع کے اندر کام، کام، یا مہارت کو انجام دینے میں ایک اتھارٹی ہے۔ اعداد و شمار کے تجزیاتی اقدام پر منحصر ہے، اسے متعدد مضامین کے ماہرین کی کراس فنکشنل نمائندگی کی ضرورت ہو سکتی ہے۔ اعلیٰ درجے کے ڈومینز کی مثالوں میں انٹیگریٹڈ اہلیت، بچوں کی بہبود، اور چائلڈ سپورٹ شامل ہیں۔ تاہم، ان ڈومینز کے اندر ذیلی زمرے ہو سکتے ہیں جیسے IV-E فنڈنگ، گود لینے، ہاٹ لائن، فرنٹ لائن یا انٹیک، کیس مینجمنٹ، اور اجتماعی سودے بازی) بطور چائلڈ ویلفیئر کا استعمال کرتے ہوئے (جنہیں بھی شامل کیا جانا چاہیے۔	کلیدی پروگرام کے موضوع کے ماہرین شماریات دان اور محقق کے نتائج کو مطلع کریں گے اور ان کی توثیق کریں گے، نیز ان حل یا مداخلتوں کے بارے میں جو تجزیاتی کوششوں کے نتیجے میں نافذ ہو سکتے ہیں۔ پروگرام کے مضامین کے ماہرین کو شامل کرنا شمولیت کی ثقافت کو پروان چڑھانے میں مدد کرتا ہے جو اعداد و شمار سے چلنے والے افعال کو ادارہ جاتی بنانے اور چلانے میں معاون ہو گا۔
ایگزیکٹو مینیجر	ایک ایگزیکٹو مینیجر سٹریٹجک وژن کے ساتھ ڈیٹا اینالیٹکس آپریشن کی نگرانی کا ذمہ دار ہوتا ہے اور بجٹ، وسائل، ٹولز اور آپریشن کی دیگر ضروری ضروریات کا انتظام کرنے کے لیے اختیار کی مطلوبہ سطح کا ہوتا ہے۔	ایک بصیرت والا اور بااثر ایگزیکٹو مینیجر کسی تنظیم میں ڈیٹا اینالیٹکس کے آغاز اور ادارہ سازی کے لیے اہم ہے۔ اعداد و شمار کے تجزیات کی کوششوں کو منظم کرنے کے لیے درکار

		وسائل کی مختلف قسم کے پیش نظر، ایگزیکٹو مینیجر کو تنظیم کی وسعت اور اس سے آگے میٹرکس کے نقطہ نظر میں مقصد کو آگے بڑھانا چاہیے۔
دیگر اسٹیک ہولڈرز	دوسرے اسٹیک ہولڈرز اندرونی یا بیرونی دلچسپی رکھنے والے فریق ہیں جو کسی تنظیم کو متاثر کر سکتے ہیں یا متاثر ہو سکتے ہیں۔ مثالوں میں کسی دوسرے دفتر، ڈویژن، یا محکمہ کا عملہ شامل ہو سکتا ہے۔ ٹھیکیدار وکالت یا صارفین	سٹیک ہولڈرز سمیت کسی خاص عمل یا پروڈکٹ کے استعمال کی خریداری کو حاصل کرنے، عمل کی قدر میں اضافہ کرتا ہے۔

صلاحیت :3 ٹیم بلڈنگ		
موضوع	**تعریف**	**کیوں ضرورت ہے**
گورننس کا ڈھانچہ	ایک ٹیم یا لوگوں کی ٹیمیں جو ایک متعین مقصد کے لیے، یا ایک محدود مدت کے لیے بلائی جاتی ہیں، جو ایک مقررہ مقصد کے حصول کے لیے تنظیم کو منظم کرنے والی ٹیموں کے لیے کردار، ذمہ داریوں، اور فیصلہ سازی کے حقوق کا ایک نظام قائم کرتی ہے۔ ایک سے زیادہ ٹیموں کے ساتھ گورننس کے ڈھانچے نارمل طور پر درجہ بندی کے ہوتے ہیں، جو حتمی مستند فیصلہ سازی کے ساتھ ت نورمل ل کا عمل قائم کرتے ہیں اور اس میں پالیسی، قانونی، پروگرام، مالی اور تکنیکی مفادات کے کراس فنکشنل اسٹیک ہولڈرز کے ایگزیکٹو، مینجمنٹ اور ورکنگ گروپ کی سطحیں شامل ہوتی ہیں۔ حکمرانی قائم کرنے کے لیے درج ذیل نمونے استعمال کیے جا سکتے ہیں۔ پروجیکٹ چارٹر یا ورک گروپ چارٹر	گورننس باڈیز کے اختیارات اور سرگرمیوں کے ذریعے، ادارے باہمی اہداف کے حصول کے لیے اپنی روزمرہ ایجنسی کے کاروبار سے آزادانہ تعاون اور کوششوں کا نظم کر سکتے ہیں۔ ساؤنڈ ڈیٹا گورننس دیگر تمام تجزیاتی سرگرمیوں کی رہنمائی کرے گا، کراس پروگرام تجزیاتی مواقع، مالی بچت، اور پروگرامی استعداد فراہم کرے گا۔ یہ یہ بھی دکھائے گا کہ فیصلہ سازی کی اتھارٹی کسی بھی سرگرمی کے سلسلے میں کہاں ہے اور ساتھ ہی وہ پیرامیٹرز بھی قائم کرے گی جن کے ذریعے مفادات کے حامل اداروں یا گروپوں کو اسٹیک ہولڈرز کے طور پر منتخب کیا جائے۔ ساؤنڈ ڈیٹا گورننس ایک دیئے گئے پروجیکٹ یا سرگرمی میں شامل مختلف ڈیٹا اور سسٹمز کی نسبت حد یا ملکیت کا احساس بھی فراہم کرے گی۔

	اسٹیک ہولڈر تجزیہ) بشمول مالی، سیاسی، اور پروگراماتی اسٹیک ہولڈرز(منصوبے، پالیسیاں، اور طریقہ کار گورننس میچورٹی روڈمیپ یاپلان	
طریقہ کار اور عمل کی دستاویزات	طریقوں، قواعد، رہنما اصولوں، عملوں، اور طریقہ کار کا ایک منتخب نظام جو علم کی ایک شاخ) ڈیٹا اینالیٹکس (سے وابستہ ہے جو صنعت کے معیارات اور بہترین طریقوں پر مبنی ہے جو خاص تحقیق، پروجیکٹ یا مطالعہ کے مقصد کو انجام دینے کے لیے بنائے گئے ہیں۔ لاگو کرنے کے لیے ایک خاص نمونہ، ماڈل یا تکنیک کا انتخاب ایک ایسی کوشش ہے جس میں تمام اسٹیک ہولڈرز سے وقت اور عزم درکار ہوتا ہے۔ اس کوشش کو دستاویز کرنے کے لیے درج ذیل کا استعمال کیا جا سکتا ہے: تحقیق یاپروجیکٹ تجزیات کا طریقہ کار ڈیٹا اکٹھا کرنا، انوینٹری اور ڈیٹا ڈکشنری انفارمیشن مینجمنٹ لائف سائیکل ماسٹر ڈیٹا مینجمنٹ پلان ڈیٹا گورننس کے نفاذ کا منصوبہ کوالٹی مینجمنٹ پلان ڈیٹا چینج مینجمنٹ پلان توثیق اور استعمال کا منصوبہ	ایک ترقی ی فریم ورک یا اصولوں کی بنیاد فراہم کرتا ہے جو فیصلہ سازی کا جواز پیش کرتا ہے۔ ایک متعین طریقہ کار اور معیاری عمل پورے انٹرپرائز میں مستقل مزاجی کو یقینی بناتے ہیں اور دوبارہ قابل استعمال کام کے نمونوں کی تخلیق کے ذریعے کام کے وقت کو تیز کرتے ہیں اور مسلسل تطہیر اور بہتری کی کوششوں کی اجازت دیتے ہیں۔
رازداری، حفاظتی	دستخط شدہ بیانات، دستاویزات یا معاہدوں کا استعمال ڈیٹا کی رسائی، جمع کرنے، ٹرانسمیشن اور اسٹوریج کے ذریعے جسمانی، انتظامی، قانونی اور تکنیکی استعمال کی رازداری اور تحفظ کے لیے	اداروں کو قانون کے ذریعہ ڈیٹا کی حفاظت اور ان پر حکومت کرنے کی ضرورت ہے تاکہ وہ جن لوگوں کی خدمت کرتے ہیں ان کی محفوظ معلومات سے سمجھوتہ

تدابیر، اور حفاظتی ٹولز	کیا جاتا ہے۔ ڈیٹا کی سرگرمیوں کا مقصد مخصوص مقاصد کے لیے واضح طور پر بیان کیا جائے گا اور یہ قابل اطلاق ریاست اور وفاقی قانون اور ضوابط کے مطابق ہونا چاہیے۔ تمام دستاویزات پر دستخط کرنے والے ہوں گے۔ علیحدہ دستاویزات یا بیانات آزادانہ طور پر یا مشترکہ طور پر بنائے جا سکتے ہیں، اور نارمل طور پر کسی پروجیکٹ، مطالعہ یا اقدام کے آغاز میں اس وقت قائم کیے جاتے ہیں جب گورننس قائم ہو رہی ہو: مفاہمت کی یادداشت (MOU) یا معاہدہ (MOA) ڈیٹا شیئرنگ کا معاہدہ سیکیورٹی ایگریمنٹ/SLA ڈیٹا تک رسائی اور رازداری کے بیانات انٹرپرائز کی یادداشت مفاہمت یا معاہدے	کرنے کے خطرے کو کم سے کم کریں۔ زیادہ اہم، تجزیاتی سرگرمیوں کی ساکھ صارفین اور نارمل طور پر عوام کے لیے اس حد تک سنگین خطرے میں پڑ جاتی ہے کہ معلومات تک رسائی کا حق نہ رکھنے والوں کے ذریعے پرائیویسی اور سیکیورٹی پروٹوکول کی خلاف ورزی ہوتی ہے۔
تنظیمی تبدیلی کا انتظام اور ادارہ سازی	واضح اور بروقت منصوبہ بندی، بات چیت، تربیت، تعلیم، اور اس اقدام یا کوشش کے لیے رپورٹنگ کے ساتھ کسی کوشش یا اقدام میں تنظیم کی مصروفیت۔ روزانہ کے کاروباری طریقوں میں ڈیٹا اینالیٹکس کے استعمال کو شامل کرنا ابتدائی کامیابیوں کو نمایاں کرتے ہوئے موجودہ کاروباری کوششوں کے لیے نئے نقطہ نظر کا باعث بن سکتا ہے اور ڈیٹا گورننس انٹرپرائز کے اقدامات کی حمایت کے لیے چیمپیئنز اور ابتدائی اختیار کرنے والوں کو محفوظ بنانے کے لیے ایک حکمت عملی ہے۔ ادارہ جاتی سرگرمیوں کی مثالوں میں شامل ہیں:	یہ تجزیاتی کوشش کو معیاری کاروباری طریقوں میں فعال کرتا ہے۔ باضابطہ منصوبہ بندی اور ادارہ جاتی تحریک کے نفاذ کے ذریعے، ڈیٹا گورننس کی کوششوں کو حاصل کیا جائے گا۔ نتیجہ ٹھوس ڈیٹا گورننس حکمت عملی کے تمام فوائد کا ادراک ہے۔ ڈیٹا گورننس کی حکمت عملی کے فوائد کو محسوس کرنے کے خطرات کو OCM کی کوششوں کے ذریعے فعال طور پر حل کیا جاتا ہے:

	تبدیلی کے انتظام کی تشخیص مواصلات اور تنظیمی تبدیلی کے انتظام کے منصوبے معیاری رپورٹنگ ٹریننگ پلانز، بشمول ڈیٹا کا استعمال، ڈیٹا سیکیورٹی، اور ڈیٹا اینالیٹکس ٹریننگ نصاب	تجزیہ پر مبنی تبدیلی کی وجہ کے بارے میں اسٹیک ہولڈرز کو تعلیم دینے اور مطلع کرنے کے لیے موثر مواصلت۔ مستقبل وژن؟ کیوں؟ یہ مجھ پر کیسے لاگو ہوتا ہے؟ کب؟ کہاں؟ کاروباری عمل کو سمجھنے اور اس کی حمایت کرنے کے لیے کاروباری مالکان کی مشغولیت اور مسلسل شمولیت جو کہ ثقافت کی تبدیلی کے ساتھ شروع ہو سکتی ہے۔ تنظیم کے لیے تعلیم، تربیت یا مہارتوں کو اپ گریڈ کرنا۔ تنظیم کے اندر مخالف مزاحمت کی شناخت، یا تو غیر فعال یا واضح۔ خطرے میں پڑنے والے عملے کو رہنمائی، تعلیم، مواصلات، اور انتظامی خدمات۔

صلاحیت :4 ٹیم بلڈنگ		
موضوع	**تعریف**	**کیوں ضرورت ہے**
مسئلہ بیان اور تقاضے		صلاحیت 1 میں شناخت شدہ مقصد کے حصول کے لیے
ڈیٹا کی اقسام	ڈیٹا ڈکشنری: ڈیٹا ڈکشنری ڈیٹا کے تجزیہ کے ماحول میں تمام ڈیٹا عناصر کا ایک لاگ ہے۔ ہر ڈیٹا عنصر کے لیے لاگ میں معلومات شامل ہونی چاہیے جیسے:	ڈیٹا لغت ڈیٹا اسٹیورڈز کو سمجھنے کے قابل بناتی ہے: ڈیٹا عناصر تجزیہ کے لیے دستیاب ہیں۔ ☐ دیگر ڈیٹا عناصر کو تجزیہ میں لانے کے لیے ممکنہ ☐ ڈیٹا عناصر کی سالمیت ☐ ڈیٹا عناصر کی جامعیت

	نام) ماحول میں معیاری(تفصیل فارمیٹ ذرائع ریکارڈ کا ذریعہ ماخذ) مالک(ماخذ میں نام فعال تاریخ بند تاریخ آبادیاتی ڈیٹا) مثالیں:(نام پتہ — مختلف اقسام فون - مختلف اقسام ای میل جنس ریس	ڈیٹا لغت ٹیکنالوجی کے بنیادی ڈھانچے کا میکانی جزو نہیں ہے بلکہ اس کے مواد کی تفصیل کا ایک لازمی حصہ ہے۔ لغت میں موجود ڈیٹا صرف ماحول میں موجود مواد کا ڈیٹا نہیں ہے، جیسے کہ کیس کا ڈیٹا، فراہم کنندہ کا ڈیٹا، دعوے کا ڈیٹا بلکہ میٹا ڈیٹا، آڈٹ/رسائی ڈیٹا کے ساتھ ساتھ تاریخی ڈیٹا (نیچے تعریفیں دیکھیں)۔ اعداد و شمار کی لغت کو گورننس کے عمل کے حصے کے طور پر مرتب اور برقرار رکھا جانا چاہئے، اوپر Capacity 2 دیکھیں۔ ہیومن سروس کیس مینجمنٹ سسٹمز میں موجود معلومات کی سب سے بنیادی شکل ڈیمو گرافک ڈیٹا ہے جو پیش کیے جانے والے فرد، گھریلو ساخت، کیس ورکرز، فراہم کنندگان اور اسی طرح کی ہے۔ اس قسم کی معلومات تمام پختگی کی سطحوں پر تجزیات کے لیے ضروری ہے اور اکثر بنیادی تجزیاتی کوششوں کا نقطہ آغاز ہوتی ہے۔
	غیر ساختہ ڈیٹا — متن اور تبصرے۔	کیس کے نوٹس اور تبصرے والے فیلڈز میں اکثر قیمتی مواد ہوتا ہے جس کا تجزیہ کرنے کے بعد، خود کیس اور پریکٹس کے حوالے سے ناقابل یقین معلومات حاصل ہوتی ہیں۔ تنظیم کی تمام سطحوں پر فیصلہ سازی کی رہنمائی میں مدد کے لیے یہ معلومات تجزیاتی عمل کے لیے اہم ہے۔
	تاریخی/آڈٹ) ٹرانزیکشنل (ڈیٹا: ڈیٹا تک رسائی اور اپ ڈیٹ ہونے کے ساتھ ہی وہ ڈیٹا جو مستقل طور پر محفوظ کیا جاتا ہے۔ڈیٹا بیس کے آڈٹ فیچر کے	کئی وجوہات کی بنا پر آڈٹ لاگز کی ضرورت ہے۔زیادہ تر نارمل طور پر سیکورٹی فراہم کرنے اور دھوکہ دہی کا پتہ لگانے کے لیے ان کی ضرورت ہوتی ہے۔ ڈیٹا بیس کے مواد پر منحصر ہے،

قانون کے ذریعہ آڈٹ لاگز کی ضرورت ہے، جیسے Sarbanes- Oxley، Payment Card Industry Compliance، HIPAA، CFR، FERPA اور FOIA۔ نوٹ: ڈیٹا بیس میں آڈٹ ڈیٹا کی خصوصیت تمام ڈیٹا بیس پروڈکٹس کی ایک خصوصیت ہے اور نارمل طور پر اس میں کوئی کردار ادا نہیں کرتی ہے جس پر یہ پیپر بحث کر رہا ہے اور اسے صرف معلوماتی سمجھا جانا چاہیے۔ ڈیٹا اینالیٹکس اس ڈیٹا کو صرف اندرونی استعمال کے لیے فراڈ اور سیکیورٹی کے لیے استعمال کرتا ہے۔ ڈیٹا بیس کی آڈٹ خصوصیات کوالٹی کنٹرول اور فرانزک جائزہ لینے کی اجازت دیتی ہیں کہ صارف کے ذریعے، تاریخ اور وقت کے اسٹیمپ کے ساتھ ڈیٹا عنصر کے ذریعے، ڈیٹا بیس کی انتظامیہ میں نگرانی فراہم کرتے ہوئے، کیا لین دین ہوا ہے۔ ایک آڈٹ ڈیٹا میں ہونے والی تبدیلیوں کو ٹریک کرتا ہے۔ تاریخ بن جائے. میٹا ڈیٹا کا بنیادی مقصد صارف یا سبسکرائبر کے ساتھ تبادلہ کیے جانے والے ڈیٹا کے بارے میں متعلقہ معلومات کی دریافت کو آسان بنانا ہے۔ میٹا ڈیٹا سبسکرائبر کو معلومات کے ساتھ مدد کرتا ہے کہ بغیر محنت کی تحقیق کے تبادلہ شدہ ڈیٹا کو کیسے سمجھنا ہے۔ ڈیٹا فائلوں کے مواد اور سیاق وسباق کو بیان کرنے سے، اصل ڈیٹا/فائلوں کی افادیت بہت بڑھ جاتی ہے، انتظامی معاونت کے اخراجات کم ہوتے ہیں اور نظام کی آزادی بہتر ہوتی ہے۔	ذریعے تخلیق کردہ جدولیں وقت کے ساتھ ساتھ ڈیٹا عناصر کے ورژن کا تاریخی منظر پیش کرتی ہیں۔ آڈٹ/رسائی ڈیٹا، جو ڈیٹا کو مسلسل ذخیرہ کیا جاتا ہے کیونکہ ڈیٹا تک رسائی اور اپ ڈیٹ کیا جاتا ہے، جیسے کہ سروس/صارف نام کو اپ ڈیٹ کرنا ڈیٹا، ڈیٹا کو اپ ڈیٹ کرنے کی تاریخ۔ تاریخی ڈیٹا، جو کہ ڈیٹا عناصر کا پچھلا ورژن ہے، جیسے، پہلے اور اب ادا شدہ دعوے میٹا ڈیٹا: دو" میٹا ڈیٹا کی قسمیں "ہیں - ساختی میٹا ڈیٹا، ڈیٹا ڈھانچے کے ڈیزائن اور تفصیلات کے بارے میں یا" ڈیٹا کے کنٹینرز کے بارے میں ڈیٹا -"اور ایپلیکیشن ڈیٹا یا ڈیٹا مواد کی انفرادی مثالوں کے بارے میں وضاحتی میٹا ڈیٹا۔ میٹا ڈیٹا میں موجود نارمل معلومات میں شامل ہیں: واضح اور قطعی کاروباری تفصیل اور اس میں شامل ڈیٹا فیلڈز کا مطلوبہ معنی ڈیٹا عناصر کو ترتیب دینے یا درجہ بندی کرنے کے لیے استعمال کیے گئے مخصوص اصول یا منطق زبان یا مقامی بولی کی تفصیلات) مثال کے طور پر، HTTP یا HL7 ورژن(کوڈ سیٹ مفروضے مخصوص فیلڈز پر لاگو ہوتے ہیں)تاریخ کی شکل، مانیٹری بیس، ICD-10/LOINC/SNOMED کوڈز(	

	میٹاڈیٹا کے معیارات میں ISO/IEC 11179 ، NIEM،Dublin core، اور NISTs کے سیکورٹی کے انتظامی معیارات شامل ہیں۔	میٹاڈیٹا سیمنٹک انٹر آپریبلٹی کے ذریعے شراکت داروں کے درمیان ڈیٹا کے تبادلے کو فروغ دیتا ہے۔ سیمنٹک انٹر آپریبلٹی کا تعلق صرف ڈیٹا کی پیکیجنگ (نحو) سے نہیں ہے، بلکہ ڈیٹا (Semantics) کے ساتھ معنی کی بیک وقت ترسیل سے ہے۔ یہ وضاحتی معلومات کو شامل کر کے مکمل کیا جاتا ہے یا "ڈیٹا کے بارے میں ڈیٹا" (میٹاڈیٹا) کے طور پر بیان کیا جا سکتا ہے، ہر ڈیٹا عنصر کو کنٹرول شدہ، مشترکہ الفاظ سے جوڑ کر۔ ڈیٹا کا مطلب ڈیٹا کے ساتھ ہی منتقل ہوتا ہے، ایک خود بیان کردہ "معلوماتی پیکیج" میں جو کسی بھی معلوماتی نظام سے آزاد ہے۔ یہ مشترکہ ذخیرہ الفاظ ہے، اور اس کے ایک آنٹولوجی سے منسلک روابط ہیں، جو مشین کی تشریح، اندازہ لگانے اور منطق کی بنیاد اور صلاحیت فراہم کرتا ہے۔
اسمبلی اور ساخت	ڈیٹا کو کمپیوٹر میں ترتیب دیں تاکہ اسے استعمال کیا جا سکے۔	یہ قدم تجزیاتی سرگرمی کے لیے ضروری ہے۔اعداد و شمار کے ڈھانچے انڈیکسنگ، ڈیزائننگ الگورتھم، اور ایپلیکیشن پروسیسنگ جیسے استعمال کے لیے بڑی مقدار میں ڈیٹا کو مؤثر طریقے سے منظم کرنے کا ذریعہ فراہم کرتے ہیں۔
ماڈل اور الگورتھم	صحیح ڈیٹاسیٹ، الگورتھم، اور متغیرات کے انتخاب کا عمل اور کسی خاص کاروباری مسئلے کے لیے ڈیٹا کو فارمیٹ کرنے کے لیے صحیح تکنیک	مناسب تجزیہ کرنے کے لیے ضروری ماڈل اور الگورتھم کی شناخت شناخت شدہ مسئلے کے لیے مطلوبہ نتائج پیدا کرتی ہے۔ یہ صحیح نتیجہ حاصل کرنے کے لیے صحیح ریاضیاتی فارمولہ استعمال کرنے کے مترادف ہے۔
ڈیٹا کو تبدیل کریں۔	متغیرات کو تبدیل کرنا تاکہ ڈیٹا کو مطلوبہ نتیجہ کے لیے استعمال کیا جا سکے۔	اس فنکشن کو نارمل طور پر لاگو کیا جاتا ہے تاکہ ڈیٹا متعین مفروضوں کو زیادہ قریب سے پورا کرے، یا گراف کی تشریح یا ظاہری شکل کو بہتر بنائے۔

عمل اور تجزیہ کریں۔	ڈیٹا کی پروسیسنگ اور تجزیہ بامعنی نتائج پیدا کرنے کے لیے ڈیٹا کو اکٹھا کرنا اور ہیرا پھیری کرنا ہے۔	یہ سرگرمی کاروبار اور کارروائیوں کے بارے میں بصیرت فراہم کرتی ہے تاکہ فیصلہ سازوں کو ایک مطلوبہ نتیجہ یا نتیجہ حاصل کرنے کے لیے ایک دیے گئے منظر نامے میں مداخلت کرنے کے لیے کارروائی کرنے کی اجازت دی جائے (یا نہیں، جیسا کہ یہ ہو سکتا ہے۔
تصور	اعداد و شمار یا معلومات کو بصری اشیاء یا اعداد و شمار کی بصری نمائندگی کے طور پر انکوڈ کر کے بات چیت کرنے کے لیے استعمال ہونے والی تکنیکیں	ڈیٹا ویژولائزیشن کا بنیادی مقصد اعداد و شمار کے گرافکس، پلاٹ، معلوماتی گرافکس، ٹیبلز، اور منتخب کردہ چارٹس کے ذریعے صارفین تک معلومات کو واضح اور مؤثر طریقے سے پہنچانا ہے۔
معیار کے جائزے	ڈیٹا کوالٹی کے جائزے اس بات کا تعین کرنے کے لیے کیے جاتے ہیں کہ ڈیٹا کے استعمال کے لیے موزوں ہے تاکہ کسی مخصوص سیاق وسباق کے مقاصد کو پورا کیا جا سکے۔	اعداد و شمار کے تجزیہ کا نتیجہ تجزیہ کردہ ڈیٹا کے معیار پر بہت زیادہ منحصر ہے۔ کاروباری فیصلوں کو مطلع کرنے کے لیے ڈیٹا اینالیٹکس استعمال کرنے کے مطلوبہ مقصد کے پیش نظر، یہ ضروری اور متوقع ہے کہ ڈیٹا کا معیار اور اس لیے تجزیاتی نتائج درست طور پر نمائندہ ہوں

3.7 ڈیٹا پر مبنی فیصلہ سازی (Data Driven Decision Making)

ڈیٹا پر مبنی فیصلہ سازی: ڈیٹا پر مبنی فیصلہ سازی (DDDM) فیصلوں کو مطلع کرنے اور رہنمائی کے لیے ڈیٹا کے استعمال کا عمل ہے۔ یہ تمام سائز کی تنظیموں کے لیے ایک اہم نقطہ نظر ہے، کیونکہ یہ فیصلہ سازی کے معیار کو بہتر بنانے، لاگت کو کم کرنے اور منافع بڑھانے میں مدد کر سکتا ہے۔

DDDM کے بہت سے فوائد ہیں، بشمول:

بہتر فیصلہ سازی: فیصلوں کو مطلع کرنے کے لیے ڈیٹا کا استعمال کرتے ہوئے، تنظیمیں زیادہ باخبر اور معروضی انتخاب کر سکتی ہیں جن کی کامیابی کے احتمالات زیادہ ہوتے ہیں۔

کم لاگت: ڈی ڈی ڈی ایم ان علاقوں کی نشاندہی کرنے میں مدد کر سکتا ہے جہاں لاگت کو کم کیا جا سکتا ہے جیسے کہ ناکارہ ہونے یا ڈیٹا کو ضائع کرنا۔ منافع میں اضافہ: DDDM منافع میں اضافے کے مواقع کی نشاندہی کرنے میں مدد کر سکتا ہے جیسے کہ نئی منڈیوں کی شناخت یا موجودہ مارکیٹوں میں توسیع۔ بہتر کسٹمر کی اطمینان: DDDM کسٹمر کی ضروریات اور ترجیحات کی شناخت اور ان پر توجہ دے کر کسٹمر کی اطمینان کو بہتر بنانے میں مدد کر سکتا ہے۔

.ڈیٹا پر مبنی بننے کے چیلنجوں کی وضاحت کریں۔

آج بہت سی کمپنیاں ڈیٹا پر مبنی بننے کی کوشش کر رہی ہیں۔ لیکن اس کا بالکل کیا مطلب ہے؟ صرف صحیح ٹولز اور ایپلی کیشنز کو انسٹال کرنے سے زیادہ، ڈیٹا پر مبنی بننے کا مطلب ڈیٹا اور تجزیات کو کاروباری حکمت عملی، اس کے نظام، عمل اور ثقافت کا حصہ بنانا ہے۔ یہ ایک ایسی ذہنیت پیدا کرنے کے بارے میں ہے جس میں تجزیات تمام حقائق پر مبنی کاروباری فیصلوں کی بنیاد بناتے ہیں اور تنظیم کے تمام سطحوں کے ذریعہ اسے قبول کیا جاتا ہے۔

ڈیٹا کوالٹی: ڈیٹا کوالٹی ڈیٹا سے چلنے والی تنظیموں کا ایک اہم پہلو ہے۔ یہ فیصلہ سازی اور تجزیہ کے لیے استعمال ہونے والے ڈیٹا کی مجموعی درستگی، مکمل، مستقل مزاجی اور مطابقت سے مراد ہے۔ اعلی معیار کا ڈیٹا اس بات کو یقینی بناتا ہے کہ بصیرت حاصل کی گئی ہے۔

ڈیٹا قابل اعتماد اور قابل عمل ہے، جو بہتر فیصلہ سازی کا باعث بنتا ہے۔

رسائی، نکالنے، پروسیسنگ اور تجزیہ کے لیے ٹولز: ڈیٹا سے چلنے والی تنظیمیں فیصلہ سازی کے مقاصد کے لیے ڈیٹا تک رسائی، نکالنے، عمل کرنے اور تجزیہ کرنے کے لیے مختلف ٹولز پر انحصار کرتی ہیں۔ یہ ٹولز خام ڈیٹا کو قابل عمل بصیرت میں تبدیل کرنے میں اہم کردار ادا کرتے ہیں جو کاروباری نتائج کو بہتر بنا سکتے ہیں۔

1. ڈیٹا تک رسائی کے اوزار ڈیٹا بیس مینجمنٹ سسٹمز (DBMS)، ڈیٹا ہیں۔
2. گودام اور ڈیٹا لیکس۔
3. ڈیٹا نکالنے کے ٹولز ویب ٹولز، کنیکٹرز اور ڈیٹا مائیگریشن ہیں۔
4. <u>اوزار (Tools)</u>
5. ڈیٹا پروسیسنگ ٹولز ڈیٹا ٹرانسفارمیشن ٹولز اور ڈیٹا انٹیگریشن ٹولز ہیں۔
6. ڈیٹا اینالیسس ٹولز بزنس انٹیلی جنس (BI) ٹولز، شماریاتی ہیں۔
7. تجزیہ کے اوزار اور مشین لرننگ ٹولز

لاگت : یہ ڈیٹا اینالیٹکس ماڈلز کو تیار کرنے اور لاگو کرنے کے اخراجات ہیں۔اس میں ڈیٹا سائنسدانوں کی خدمات حاصل کرنے، تجزیاتی سافٹ ویئر کی خریداری،اور ملازمین کو تجزیاتی ٹولز استعمال کرنے کے بارے میں تربیت دینے کی لاگت شامل ہے۔ تجزیات اور ماڈلنگ کے اخراجات زیادہ ہو سکتے ہیں، لیکن وہ سرمایہ کاری پر نمایاں منافع بھی حاصل کر سکتے ہیں۔

میراثی ڈیٹا : میراثی ڈیٹا ڈیٹا سے چلنے والی تنظیموں کے لیے ایک اثاثہ ہے۔ یہ ماضی کے رجحانات، گاہک کے رویے،اور آپریشنل کارکردگی کے بارے میں تاریخی سیاق و سباق اور بصیرت فراہم کر سکتا ہے جو مستقبل کے فیصلوں سے آگاہ کر سکتا ہے اور کاروباری نتائج کو بہتر بنا سکتا ہے۔ میراثی ڈیٹا مختلف فارمیٹس اور ذرائع میں پایا جاسکتا ہے، بشمول کاغذی ریکارڈ، پرانے سافٹ ویئر سسٹم اور ڈیٹا بیس۔

ڈیٹا سائلوز : ڈیٹا سائلوز الگ تھلگ ڈیٹا سیٹ ہیں جو کسی تنظیم کے اندر دوسرے محکموں یا ٹیموں کے ذریعے آسانی سے قابل رسائی یا قابل استعمال نہیں ہوتے ہیں۔

مضبوط قیادت : ڈیٹا سے چلنے والی تنظیموں میں مضبوط قیادت کامیابی کے لیے ضروری ہے۔ رہنماؤں کو تنظیم کے لیے ایک وژن مرتب کرنے،اس وژن کو حاصل کرنے کے لیے حکمت عملی تیار کرنے اور ڈیٹا پر مبنی فیصلہ سازی کا کلچر بنانے کے قابل ہونے کی ضرورت ہے۔ انہیں تنظیم کے اندر اور باہر دونوں اسٹیک ہولڈرز کے ساتھ مؤثر طریقے سے بات چیت کرنے کے قابل ہونے کی ضرورت ہے۔

ڈیٹا سائنسدان: ڈیٹا سائنسدان ڈیٹا سے قیمتی بصیرت نکالنے کے لیے اعداد و شمار، مشین لرننگ اور ڈیٹا کے تجزیہ میں اپنی مہارت کا استعمال کرتے ہوئے، ڈیٹا سے چلنے والی تنظیموں میں ایک اہم کردار ادا کرتے ہیں۔ وہ تنظیموں کو باخبر فیصلے کرنے، کارکردگی کو بہتر بنانے اور مسابقتی برتری حاصل کرنے میں مدد کرتے ہیں۔

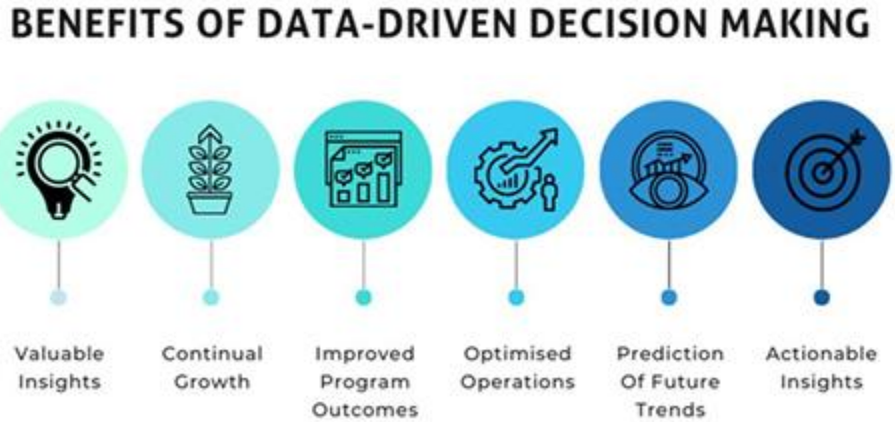
کسی تنظیم میں ڈیٹا پریکٹس کیا ہیں اور اسے ڈیٹا پر مبنی کیسے بنایا جائے؟

کسی تنظیم میں ڈیٹا پریکٹسز: ڈیٹا اینالیٹکس تنظیمی کامیابی کے حصول کے لیے ایک ناقابل تبدیلی ٹول ہے۔ صرف ڈیٹا اکٹھا کرنا، ذخیرہ کرنا اور تجزیہ کرنا کافی نہیں ہے۔ حقیقی ڈیٹا سے چلنے والی، قابل پیمائش ترقی، اور ترقی ڈیٹا سے چلنے والی کمپنی کی ثقافت کے قیام کے ساتھ آتی ہے۔ ڈیٹا سے چلنے والی ہر تنظیم درج ذیل کو استعمال کرتی ہے:

- ڈیٹا مزاحم: کاروبار کسی بھی مقصد کے لیے ڈیٹا کا استعمال نہیں کرتا ہے۔
- ڈیٹا سے متجسس : کاروبار اپنے کاموں کی کامیابی کو بڑھانے کے لیے ڈیٹا کو ضم کرنے اور استعمال کرنے کا ایک موقع دیکھتا ہے۔
- ڈیٹا سے آگاہی : کاروبار اپنے کچھ آپریشنل عملوں کے لیے ڈیٹا اینالیٹکس کا استعمال شروع کرتا ہے۔
- ڈیٹا سیوی : کاروبار آپریشن کے متعدد شعبوں اور مختلف محکموں میں ڈیٹا اینالیٹکس کا استعمال شروع کرتا ہے۔ یہ ڈیٹا کی بصیرت حاصل کرنا بھی شروع کرتا ہے اور فیصلہ سازی کے لیے ان کا استعمال کرتا ہے۔
- ڈیٹا پر مبنی : کاروبار اسٹریٹجک منصوبہ بندی کے قیام، مقاصد کے تعین، اور مستقبل کے خطرات کا اندازہ لگانے کے لیے ڈیٹا کو بنیادی اثاثے کے طور پر استعمال کرتا ہے۔

ڈیٹاپر مبنی فیصلہ سازی کے فوائد

ڈیٹاپر مبنی فیصلہ سازی کے مختلف صنعتوں اور تنظیموں میں بے شمار فوائد ہیں۔ یہاں کچھ اہم فوائد ہیں:



1) بہتر درستگی اور معروضیت :ڈیٹاپر مبنی فیصلہ سازی موضوعی رائے یا تعصب کے بجائے حقائق پر مبنی معلومات اور تجزیہ پر انحصار کرتی ہے۔

2) ڈیٹاکا استعمال کرتے ہوئے، تنظیمیں زیادہ درست اور معروضی فیصلے کر سکتی ہیں، غلطیوں یا گمراہ کن فیصلوں کے خطرے کو کم کر سکتی ہیں۔

3) باخبر فیصلہ سازی :ڈیٹا ماضی کی کارکردگی، موجودہ رجحانات، اور مستقبل کے ممکنہ نتائج کے بارے میں بصیرت فراہم کرتا ہے۔ اعداد و شمار کے تجزیے سے فائدہ اٹھا کر، فیصلہ ساز صورت حال کی ایک جامع تفہیم حاصل کر سکتے ہیں، اور انہیں مفروضوں کی بجائے ثبوت کی بنیاد پر باخبر انتخاب کرنے کے قابل بناتے ہیں۔

4) بہتر مسئلہ حل کرنا :ڈیٹاپر مبنی فیصلہ سازی تنظیموں کو فعال طور پر مسائل کی شناخت اور ان کو حل کرنے کے قابل بناتی ہے۔ متعلقہ اعداد و شمار کا تجزیہ کر کے، تنظیمیں پیٹرن، رجحانات، اور مسائل کی بنیادی وجوہات کی نشاندہی کر سکتی ہیں، جس سے مسئلہ حل کرنے کی مؤثر حکمت عملی اور بہتر نتائج برآمد ہوتے ہیں۔

5) لاگت کی بچت اور وسائل کی اصلاح :ڈیٹا کا تجزیہ تنظیموں کو ناکارہیوں اور بہتری کے لیے شعبوں کی نشاندہی کرنے میں مدد کرتا ہے۔ ڈیٹا پر مبنی بصیرت کی بنیاد پر عمل اور وسائل کی تقسیم کو بہتر بنا کر، کاروبار لاگت کو کم کر سکتے ہیں، فضلہ کو کم کر سکتے ہیں، اور دستیاب وسائل کے زیادہ سے زیادہ استعمال کو بڑھا سکتے ہیں۔

6) بہتر کارکردگی اور پیداواریت :ڈیٹاپر مبنی فیصلہ سازی تنظیموں کو ان عوامل کی نشاندہی کرنے کی اجازت دیتی ہے جو کامیابی یا ناکامی میں کردار ادا کرتے ہیں۔ کارکردگی کی پیمائش کے اعداد و شمار کا تجزیہ کر کے، تنظیمیں پیداواری کو بڑھانے، کاموں کو بہتر بنانے اور بہتر نتائج حاصل کرنے کے لیے حکمت عملیوں کو نافذ کر سکتی ہیں۔

7) کسٹمر کی بہتر تفہیم : ڈیٹا کا تجزیہ کسٹمر کے رویے، ترجیحات اور ضروریات کے بارے میں قیمتی بصیرت فراہم کرتا ہے۔ گاہکوں کو بہتر طور پر سمجھ کر، تنظیمیں اپنی مصنوعات، خدمات، اور مارکیٹنگ کی کوششوں کو کسٹمر کی توقعات کو پورا کرنے کے لیے تیار کر سکتی ہیں، جس سے صارفین کی اطمینان اور وفاداری میں بہتری آتی ہے۔

8) مسابقتی فائدہ : وہ تنظیمیں جو ڈیٹا پر مبنی فیصلہ سازی کو اپناتی ہیں مسابقتی برتری حاصل کرتی ہیں۔ جدت طرازی کو آگے بڑھانے، مارکیٹ کے رجحانات کی نشاندہی کرنے اور گاہک کے مطالبات کا اندازہ لگانے کے لیے ڈیٹا کا فائدہ اٹھا کر، کاروبار مقابلے میں آگے رہ سکتے ہیں اور دوسروں سے پہلے مواقع سے فائدہ اٹھا سکتے ہیں۔

9) رسک مینجمنٹ : ڈیٹا کا تجزیہ تنظیموں کو ممکنہ خطرات کی شناخت اور اندازہ کرنے میں مدد کرتا ہے۔ تاریخی اعداد و شمار، مارکیٹ کے رجحانات، اور دیگر متعلقہ معلومات کا تجزیہ کر کے، فیصلہ ساز خطرے کو کم کرنے کی حکمت عملی تیار کر سکتے ہیں اور باخبر انتخاب کر سکتے ہیں جو ممکنہ خطرات کے اثرات کو کم سے کم کر سکتے ہیں۔

10) چستی اور موافقت : ڈیٹا پر مبنی فیصلہ سازی تنظیموں کو مارکیٹ کے بدلتے ہوئے حالات اور ابھرتے ہوئے رجحانات کا فوری جواب دینے کے قابل بناتی ہے۔ ڈیٹا کی مسلسل نگرانی اور تجزیہ کر کے، تنظیمیں متعلقہ رہنے، کسٹمر کے مطالبات کو پورا کرنے، اور نئے مواقع سے فائدہ اٹھانے کے لیے اپنی حکمت عملیوں اور کارروائیوں کو ڈھال سکتی ہیں۔

11) ثبوت پر مبنی جوابدہی : ڈیٹا پر مبنی فیصلہ سازی اداروں کے اندر جوابدہی کے کلچر کو فروغ دیتی ہے۔ شفافیت، جوابدہی، اور مسلسل بہتری کی سہولت فراہم کرتے ہوئے، فیصلوں اور نتائج کا پتہ دستیاب ڈیٹا سے لگایا جا سکتا ہے۔

ڈیٹا پر مبنی فیصلہ سازی کے چیلنجز

ڈیٹا پر مبنی فیصلہ سازی بہت سے فوائد پیش کرتی ہے، جیسے کہ درستگی میں اضافہ، بہتر کارکردگی، اور بہتر بصیرت۔ تاہم، یہ اپنے چیلنجوں کے اپنے سیٹ کے ساتھ بھی آتا ہے۔ ڈیٹا پر مبنی فیصلہ سازی سے وابستہ کچھ نارمل چیلنجز یہ ہیں:

1. ڈیٹا کا معیار اور وشوسنییتا: بنیادی چیلنجوں میں سے ایک فیصلہ سازی کے لیے استعمال کیے جانے والے ڈیٹا کے معیار اور وشوسنییتا کو یقینی بنانا ہے۔ ڈیٹا نامکمل، غلط، یا متعصب ہو سکتا ہے، جو غلط نتائج اور فیصلوں کا باعث بنتا ہے۔ اس چیلنج سے نمٹنے کے لیے ڈیٹا کی صفائی اور توثیق کے عمل ضروری ہیں۔

2. ڈیٹا انٹیگریشن: تنظیموں کے پاس اکثر ڈیٹا مختلف سسٹمز اور ذرائع میں پھیلا ہوا ہوتا ہے، جس سے ڈیٹا کو مؤثر طریقے سے مربوط اور تجزیہ کرنا مشکل ہو جاتا ہے۔ ڈیٹا انضمام کے لیے محتاط منصوبہ بندی، مطابقت کی جانچ پڑتال اور ڈیٹا گورننس کی ضرورت ہوتی ہے تاکہ ہموار رسائی اور استعمال کو یقینی بنایا جا سکے۔

3. ڈیٹا پرائیویسی اور سیکیورٹی: چونکہ ڈیٹا پر مبنی فیصلہ سازی میں بہت زیادہ ڈیٹا اکٹھا کرنا اور اس کا تجزیہ کرنا شامل ہے، اس لیے رازداری اور سیکیورٹی کے خدشات اہم ہو جاتے ہیں۔ تنظیموں کو ڈیٹا کے تحفظ کے ضوابط پر عمل کرنا، مضبوط حفاظتی اقدامات کو نافذ کرنا، اور حساس ڈیٹا کی حفاظت اور کسٹمر کے اعتماد کو برقرار رکھنے کے لیے اخلاقی رہنما خطوط قائم کرنا چاہیے۔

4. ٹیلنٹ اور ہنر کا فرق: مؤثر ڈیٹا پر مبنی فیصلہ سازی کے لیے ایسے ماہر پیشہ ور افراد کی ضرورت ہوتی ہے جو ڈیٹا کے تجزیہ، شماریات، مشین لرننگ، اور ڈیٹا ویژولائزیشن میں مہارت رکھتے ہوں۔ تاہم، جاب مارکیٹ میں ایسے پیشہ ور افراد کی کمی ہے، جس سے ٹیلنٹ کا خلا پیدا ہوتا ہے جسے بھرنے کے لیے تنظیمیں جدوجہد کرتی ہیں۔

5. تبدیلی کا انتظام: ڈیٹا پر مبنی فیصلہ سازی کی ثقافت کو اپنانے کے لیے اکثر تنظیمی عمل، ورک فلو، اور ذہنیت میں اہم تبدیلیوں کی ضرورت ہوتی ہے۔ تبدیلی کے خلاف مزاحمت اور اسٹیک ہولڈرز کی جانب سے خریداری کا فقدان ڈیٹا پر مبنی طریقوں کے کامیاب نفاذ اور استعمال میں رکاوٹ بن سکتا ہے۔

6. تشریح اور سیاق وسباق: اگرچہ ڈیٹا قیمتی بصیرت فراہم کرتا ہے، لیکن یہ ضروری ہے کہ نتائج کی مناسب سیاق وسباق میں تشریح کی جائے۔ ڈیٹا پر غلط تشریح یا حد سے زیادہ انحصار غلط نتائج اور فیصلوں کا باعث بن سکتا ہے۔ ڈیٹا کی حدود کو سمجھنا اور اسے ڈومین کے علم اور مہارت کے ساتھ جوڑنا بہت ضروری ہے۔

7. تعصب اور منصفانہ: ڈیٹا پر مبنی فیصلہ سازی غیر ارادی طور پر تاریخی ڈیٹا میں موجود تعصبات کو برقرار رکھ سکتی ہے، جس کے نتیجے میں امتیازی نتائج برآمد ہوتے ہیں۔ تعصبات کی شناخت اور تخفیف، انصاف پسندی کو یقینی بنانے، اور اخلاقی فیصلہ سازی کے طریقوں کو فروغ دینے کے لیے احتیاط برتنی چاہیے۔

8. ڈیٹا اوورلوڈ اور فیصلہ کن فالج: دستیاب اعداد وشمار کی کثرت کے ساتھ، معلومات کے زیادہ بوجھ کا خطرہ ہے، جس سے سب سے زیادہ متعلقہ اور قابل عمل بصیرت کی شناخت کرنا مشکل ہو جاتا ہے۔ تنظیموں کو واضح مقاصد کی وضاحت کرنے، کلیدی کارکردگی کے اشارے (KPIs) قائم کرنے، اور فیصلے کے فالج سے بچنے کے لیے مؤثر ڈیٹا ویژولائزیشن تکنیکوں کو استعمال کرنے پر توجہ مرکوز کرنی چاہیے۔

3.8 اکتسابی نتائج (Learning Outcome)

طلباء نے اس بات کی واضح تفہیم تیار کی ہے کہ کس طرح قابلیت کی تعمیر کسی تنظیم کی ڈیٹا کو مؤثر طریقے سے منظم کرنے اور تجزیہ کرنے کی صلاحیت کو مضبوط کرتی ہے۔ انہوں نے ڈیٹا پروسیسنگ کی اہمیت اور خام ڈیٹا کو بامعنی اور مفید معلومات میں تبدیل کرنے میں شامل منظم اقدامات کو جان لیا ہے۔ صلاحیت کی منصوبہ بندی کے تصور کے ذریعے، طلباء یہ سمجھتے ہیں کہ تنظیمی ڈیٹا کے تقاضوں کو پورا کرنے کے لیے وسائل کو مؤثر طریقے سے کیسے مختص اور استعمال کیا جائے۔ انہوں نے کارکردگی کی سطح کا جائزہ لینے، مہارت کے فرق کی نشاندہی کرنے، اور

افرادی قوت کی صلاحیتوں کی منصوبہ بندی کرنے میں صلاحیت کے میٹرکس کے کردار کو بھی دریافت کیا ہے۔ مجموعی طور پر، اس یونٹ نے سیکھنے والوں کو قابلیت کی ترقی کو موثر ڈیٹا مینجمنٹ کے ساتھ مربوط کرنے کے قابل بنایا ہے، اس بات کو یقینی بناتے ہوئے کہ تجزیاتی عمل اسٹریٹجک کاروباری اہداف کے ساتھ ہم آہنگ ہوں اور پائیدار ترقی کی حمایت کریں۔

3.9 نمونہ امتحانی سوالات (Model Examination Questions)

معروضی جوابات کے حامل سوالات (Objective Answer Type Questions)

1. کاروباری تجزیات کے تناظر میں صلاحیت کی ترقی کا کیا مطلب ہے، اور یہ جدید تنظیموں کے لیے کیوں اہم ہے؟
2. کاروباری تجزیات میں ایک مؤثر صلاحیت سازی کی حکمت عملی کے اہم اجزاء کیا ہیں؟
3. ڈیٹا کی کہانی سنانے اور تصور کرنے سے کاروباری ماحول میں فیصلہ سازی کی صلاحیت کیسے بڑھ سکتی ہے؟
4. تنظیمیں اپنے ملازمین کے درمیان تجزیاتی مہارت کے فرق کا اندازہ کیسے لگا سکتی ہیں اور اسے ختم کر سکتی ہیں؟
5. کسی تنظیم کے اندر ڈیٹا سے چلنے والی ثقافت کو فروغ دینے میں قیادت کیا کردار ادا کرتی ہے؟
6. کمپنیاں تجزیات میں صلاحیت سازی کی کوششوں کو اپنے مجموعی اسٹریٹجک اہداف کے ساتھ کیسے ہم آہنگ کر سکتی ہیں؟
7. تنظیموں میں تجزیاتی صلاحیت کی ترقی میں کچھ نارمل رکاوٹیں کیا ہیں، اور ان پر کیسے قابو پایا جا سکتا ہے؟
8. ایک درمیانے سائز کی کمپنی کے لیے صلاحیت بڑھانے کا روڈ میپ تیار کریں جس کا مقصد اگلے تین سالوں میں ڈیٹا پر مبنی بننا ہے۔
9. ایک ایسے کیس کا اندازہ کریں جہاں کمپنی نے کامیابی کے ساتھ تجزیاتی صلاحیت بنائی—کامیابی کے اہم عوامل کیا تھے؟
10. ڈیٹا پر مبنی فیصلہ سازی کیا ہے، اور یہ کاروبار میں وجدان پر مبنی فیصلہ سازی سے کیسے مختلف ہے؟

مختصر جوابات کے حامل سوالات (Short Answers Type Questions)

1. ڈیٹا پر مبنی فیصلہ سازی کو جدید کاروباری حکمت عملی کا ایک اہم جز کیوں سمجھا جاتا ہے؟
2. کاروباری تجزیاتی ٹولز مارکیٹنگ، فنانس، اور آپریشنز جیسے شعبوں میں ڈیٹا پر مبنی فیصلہ سازی کی حمایت کیسے کر سکتے ہیں؟
3. ایک ایسے منظر نامے کی وضاحت کریں جہاں ڈیٹا کی ناقص تشریح کی وجہ سے کاروباری فیصلہ غلط ہوا۔ تجزیات سے نتیجہ کیسے بدل سکتا ہے؟
4. مارکیٹ کے غیر یقینی حالات میں پیش گوئی کرنے والے تجزیات اسٹریٹجک فیصلوں کے معیار کو کیسے بڑھا سکتے ہیں؟

طویل جوابات کے حامل سوالات (Long Answers Type Questions)

1. ڈیٹا پر مبنی فیصلہ سازی کو لاگو کرتے وقت تنظیموں کو کن مشترکہ چیلنجوں کا سامنا کرنا پڑتا ہے، اور وہ ان پر کیسے قابو پا سکتی ہیں؟

2. ڈیٹا یا الگورتھم میں تعصب کس طرح کاروباری فیصلوں کو متاثر کر سکتا ہے، اور اس خطرے کو کم کرنے کے لیے کیا اقدامات کیے جا سکتے ہیں؟
3. فیصلہ سازی کے لیے ڈیٹا اکٹھا کرنے اور استعمال کرنے میں اخلاقی تحفظات پر بحث کریں۔
4. ایک کمپنی کس طرح انتظام کی تمام سطحوں پر ڈیٹا پر مبنی کلچر بنا سکتی ہے؟
5. مؤثر فیصلہ سازی میں ڈیٹا گورننس اور ڈیٹا کا معیار کیا کردار ادا کرتا ہے؟
6. ایم بی اے کے طلباء انٹرن شپ یا لائیو پروجیکٹس کے دوران ڈیٹا پر مبنی فیصلہ سازی، اصولوں کا اطلاق کیسے کر سکتے ہیں؟

اکائی 4۔ شماریات اور فیصلہ سازی

(Statistics and Decision Making)

اکائی کے اجزا

4.0	تمہید	Introduction
4.1	مقاصد	Objective
4.2	علم شماریات کا تصور	StatisticsConcept of
4.3	معنی اور تعریف	Meaning and Definition
4.4	شماریات کی وسعت	Scope of Statistics
4.5	شماریات کی اہمیت	Importance of Statistics
4.6	شماریات کے آلات	Tools of Statistics
4.7	مرکزی رجحان کے معنی اور تعریف	Meaning and Definitions of Measures of Central Tendency)
4.8	اکتسابی نتائج	
4.9	نمونہ امتحانی سوالات	Model Examination Questions

4.0 تمہید(Introduction)

عزیز طلباء! کیا بچپن میں آپ نے کبھی پیسے جمع کیے ہیں۔ ہاں یا نہیں، اگر ہاں تو یاد کیجیے کہ یہ پیسے کہاں سے آتے تھے، کہاں جمع کرتے تھے، پیسوں کا حساب کتاب کس طرح کرتے تھے اور ان پیسوں کو آپ کس لیے استعمال کیے تھے۔ آپ کو یاد ہو گا عیدین، تہوار کے موقع پر والدین بڑے بھائی بہن وغیرہ کبھی کبھی بطور انعام یا عیدی آپ کو پیسے دیے ہوں گے۔ ان تمام پیسوں کو آپ ایک مٹی کے غلہ میں جمع کیے ہونگے۔ ان پیسوں سے آپ عیدین کے موقع پر کچھ کپڑے یا کھلونے خریدے ہوں گے۔ سالانہ امتحان کے بعد اسکول سے ہر سال آپ کو مختلف امتحانات میں حاصل کردہ نشانات کا ایک تعلیمی کارڈ(Progress Card) دیا جاتا ہے۔ جماعت دہم وانٹر میڈیٹ کے حاصل کردہ نشانات پر غور کیجیے، آپ کی کامیابی کا تناسب کا حساب کتاب کس طرح کرتے ہیں۔ کس درجہ میں آپ کامیاب ہو چکے، کتنے فیصد آپ کو

نشانات حاصل ہوئے،آپ کی تعلیمی قابلیت کس موقف میں ہے وغیرہ وغیرہ کااندازہ صرف آپ کے حاصل کردہ نشانات کی بنیاد پر ہی طے کیا جاتاہے۔اس اکائی میں اسی سے متعلق ایک علم ہے جس کو علم شماریات کہتے ہیں اس کے متعلق تفصیل سے مطالعہ کریں گے۔

4.1 مقاصد(Objectives)

اس اکائی کے مطالعہ کے بعد آپ اس قابل ہو جائیں گے کہ :

- آپ علم شماریات کے معنی اور تعریف کو بیان کر سکیں۔
- آپ علم شماریات کی اہمیت وافادیت کو بیان کر سکیں۔
- آپ علم شماریات کے استعمالات کو بیان کر سکیں۔
- آپ علم شماریات کے خصوصیات کو بیان کر سکیں۔
- آپ علم شماریات کی وسعت کو بیان کر سکیں۔

4.2 علم شماریات کا تصور(Concept of Statistics)

علم شماریات ایک اہم علم ہے۔یہ علم کاروبار، تعلیم،آمدنی واخراجات، صنعت و حرفت، رسمی وغیر رسمی تنظیم، پیداوار واخراجات، کھیل کود وغیرہ وغیرہ مختلف میدانوں میں استعمال کرتے ہیں۔اس علم کی بدولت کسی بھی ادارے کی کار کردگی کا اندازہ قائم ہوتا ہے۔ علم شماریات کے بغیر کسی بھی ادارے کی کار کردگی کو محسوب نہیں کر سکتے۔کسی ادارے میں بر سر خدمت سرکاری یا غیر سرکاری ملازم کو خدمت کے سلسلے میں ماہانہ تنخواہ حاصل ہوتی ہے۔ملازم اس تنخواہ کو ماہانہ کے مختلف اخراجات میں اس طرح تقسیم کرتا ہے جس سے تمام اخراجات میں توازن پیدا ہو۔اسی طرح تاجرین بھی ماہانہ گھر کے اخراجات کی منصوبہ بندی کرتے ہیں۔ تجارت کی آمدنی واخراجات کا حساب کتاب رکھتے ہیں اور اخراجات پر کنٹرول کرتے ہوئے اونچا منافع حاصل کرنا چاہتے ہیں۔تاجر کاروبار میں مختلف اقسام کے اخراجات برداشت کرتا ہے۔دوکان کا کرایہ، نوکروں کی تنخواہ، صاف صفائی کے اخراجات، سامان کی خریدی پر سفری اخراجات، بجلی کا خرچہ وغیرہ وغیرہ ان تمام اخراجات کو وہ صرف کاروبار سے حاصل آمدنی سے ہی ادا کرتا ہے۔ بازار میں مسابقت کرتے ہوئے اپنے کاروبار کو بڑھانا چاہتا ہے۔اس لیے کم از کم منافع پر کاروبار کو کرتے ہوئے صارفین کو راغب کرنے کی کوشش کرتا ہے۔اخراجات کا تعین کرنا، قیمت خرید کا حساب کرنا، منافع کی شرح کا تعین کرنا، قیمت فروخت کا تعین کرنا، کاروبار سے حاصل آمدنی کا حساب کرنا وغیرہ وغیرہ شماریات سے تعلق رکھتے ہیں۔یہاں پر ایک صنعت کی مثال کے ذریعہ علم شماریات کا استعمال کس طرح کیا جاتا ہے اس کی وضاحت کی جاتی ہے۔ چپل یا جوتے بنانے والی صنعت پر غور کیجیے۔وہ پیداوار کرنے سے قبل کافی چیزوں پر غور وخوص کرتی ہے۔سماج میں کل افراد کی تعداد کیا ہے، اس میں خواتین اور مرد حضرات کی تعداد کیا ہے، کس عمر کے لوگوں کی تعداد کتنی ہے، بچوں کی تعداد کیا ہے، مدرسہ جانے والے طلباء و طالبات کی تعداد کیا ہے، یہ طلباء و طالبات کی کتنی مقدار شہروں اور کتنی مقدار دیہاتوں میں رہتی ہے، اسی طرح کالج سطح پر اور دفتری سطح پر مرد وعورتوں کی تعداد کیا ہے، ان سب کے پیر کا اوسط

سائز کیا ہوگا، یہ لوگ کس قسم کے جوتے چپل، سینڈل استعمال کرتے ہیں وغیرہ وغیرہ۔ مختلف حصوں میں تجزیہ کرتے ہوئے مختلف سائز کے جوتوں کی تعداد کا تعین کیا جاتا ہے۔ اگر مندرجہ بالا باتوں کو نظر انداز کرتے ہوئے جوتے چپل تیار کریں تو وہ صارفین کے استعمال کے قابل نہیں ہونگے، اور صنعتکار کو نقصان ہوگا۔ صنعتکار کو کامیابی سے ہمکنار ہونے کے لیے علم شماریات کا استعمال کرنا لازمی ہے۔

4.3 معنی اور تعریف (Meaning and Definition)

علم شماریات ایک اہم مضمون ہے۔ زمانۂ قدیم سے لیکر آج تک مختلف طریقوں سے علم شماریات کا استعمال ہوتا رہا ہے۔ ہر دور میں اس کے معنی اور تعریف کی وضاحت کی گئی ہے۔

لفظ Statistics لاطینی لفظ "Status" اور "Statisticum" "Statisticus" اور جرمن لفظ Statistik، اطالوی لفظ Statista اور فرانسیسی لفظ Statistique سے ماخوذ ہے۔ اس کو Political State یا Government سے مراد لیا جاتا ہے۔ قدیم دور (Ancient Period) میں Statistics کو Science of State یا Science of Kings کے طور پر استعمال کیا گیا تھا۔ 1602ء میں شیکسپیئر نے اپنے ڈرامہ Hamlet میں لفظ Statist کا استعمال کیا تھا۔ 1662ء میں John Grount اور William Petty نے مردم شماری کے لیے علم Statistics کو استعمال کیا تھا۔ 1749ء میں جرمنی میں اس لفظ کو استعمال کیا گیا۔ وقت کے بدلنے کے ساتھ ساتھ دنیا بھر میں اس کا استعمال نارمل ہوا اور آج کے دور میں علم شماریات (Statistics) کو ملک کی مردم شماری، سماجی و معاشی مختلف پہلوؤں کے مطالعہ میں استعمال کرتے ہیں۔ Statistics کو اردو زبان میں علم شماریات میں اعداد شمار بھی کہتے ہیں۔

Prof.Prasanta Chandra Mahalamobis کو ہندوستانی علم شماریات کا باوا آدم (Father of Indian Statistics) کہتے ہیں۔ اسی طرح برطانوی ماہر شماریات Sir Ronald Fisher کو جدید شماریات کا باوا آدم (Father of Modern Statistics) کہتے ہیں۔ علم شماریات اطلاقی ریاضی (Applied Mathematics) سے تعلق رکھتی ہے۔ رونالڈ فشر، Florence Nightingale، John Graunt، کارل پیرسن، مہالونوبس، John Tuhey عالمی سطح کے چند اہم ماہرین شماریات ہیں۔

علم شماریات کو مختلف ماہرین نے مختلف تعریفات پیش کیے ہیں۔

Prof. Batingtoo کے مطابق ''علم شماریات کو 'Science of Estimates and Probabilities' کہتے ہیں''۔

Croxton and Cowdon کے مطابق ''علم شماریات عددی مواد کو حاصل کرنے پیشکش اور تجزیہ سے تعلق رکھتا ہے''۔

Selligman کے مطابق ''علم شماریات تحقیق میں مواد کی درجہ بندی اور پیشکش سے تعلق رکھتی ہے''۔

Webster کے مطابق ''علم شماریات کسی ریاست کے حقائق کی وضاحت کرتے ہیں''۔

A.L. Bowley کے مطابق ''شماریات حساب کاری کا علم (Science of Counting) ہے''۔

دوسری تعریف میں شماریات کو(Science of Average)یعنی اوسط کے مطالعہ کو شماریات قرار دیا۔

Youle and Kendal کے مطابق ''شماریات مقداری مواد سے تعلق رکھتا ہے، جس کو مختلف وجوہات کے تحت استعمال کرتے ہیں''۔

4.4 شماریات کی وسعت(Scope of Statistics)

علم شماریات کے معنی اور تعریف سے واقف ہو چکے ہیں۔آپ کو اندازہ ہو چکا ہے کہ علم شماریات کا دائرہ کافی وسیع ہے۔اس کو ہر چھوٹے و بڑے تاجرین صنعتکار، پھیری کے کاروبار کرنے والے تاجرین، دوکاندار، ترکاری فروش وغیرہ وغیرہ بھی استعمال کرتے ہیں۔ تعلیمی اداروں میں اس کا خاص استعمال ہوتا ہے۔کامیابی کے مختلف سطحیں بنائی جاتی ہیں۔اول، دوم، سوم، امتیازی نشانات وغیرہ اور فیصد مقرر کیا جاتا ہے۔ہر سال تعلیمی ادارے جماعت دہم یا اعلیٰ جماعتوں کی کامیابی کی فہرست اور تناسب جاری کرتے ہیں۔تاکہ تعلیمی ادارے کی کارکردگی سے لوگ واقف ہو سکیں۔طبّی ادارے یعنی شفاخانے، دواخانے، دارالمطالعہ، امداد باہمی ادارے و تاجرین، صنعتکار اپنی سالانہ کارکردگی کو رپورٹ کی شکل میں شائع کرتے ہیں۔ذیل میں علم شماریات کی وضاحت کی گئی ہے۔

کاروباری انتظامیہ میں شماریات Statistics in Business Management

علم شماریات ایک اہم علم ہے۔اس کو ہر ایک اپنی ضرورت اور مقصد کے تحت استعمال کرتے ہیں۔کاروباری انتظامیہ کے تمام فیصلے شماریات کے پس منظر میں طے کیے جاتے ہیں۔یہ علم کاروبار کی کامیابی اور ناکامی کے عوامل کی نشاندہی کرتی ہے۔کاروبار کی نوعیت، قسم، مقام، حالات وغیرہ کے اعتبار سے سرمایہ کی موزوں سرمایہ کاری، کاروبار کے مختلف اثاثوں کی خریدی، خام مال ومزدوروں کا تعین، مقدار پیداوار کا تعین، ذخیرہ کی حکمت عملی، تشہیر، صارفین تک سامان کی منتقلی وغیرہ مختلف منورملات میں علم شماریات کے مختلف تکنیک کی مدد لی جاتی ہے۔

معاشی منصوبہ بندی (Economic Planning)

ہر ملک معاشی ترقی واستحکام پر مختلف منصوبوں کی تشکیل دیتی ہے۔استحکام کو قوت بخشنے والے پراجکٹ کی منظوری وغیرہ میں علم شماریات کارآمد ثابت ہوتی ہے۔ موجودہ معاشی ترقی کے پس منظر میں معاشی مسائل کی نشاندہی، مسائل کے حل کے تدابیر، وقت کا تعین، درکار سرمایہ کی مقدار، ممکنہ نتائج، ممکنہ شرح ترقی، معشیت کے مختلف شعبوں پر اثرات وغیرہ وغیرہ اہم امور ہیں۔جن کو معاشی منصوبہ بندی کے دوران غور وخوص کیا جاتا ہے۔ ہر ملک مختلف سماجی و معاشی مسائل سے دوچار ہوتا ہے۔ مسائل کے ممکنہ حل کے لیے حکمت عملی کو قطعیت دینے کے دوران شماری پیمانوں کا استعمال کیا جاتا ہے۔

انتظامیہ میں علم شماریات Statistics in Administration

آپ نے دیکھا کہ علم شماریات ایک وسیع مضمون ہے۔ ہر مضمون اس کو مختلف طریقوں سے استعمال کرتا ہے۔ صنعتی و دفتری انتظامیہ میں بھی شماریات کے مختلف پہلوؤں کا استعمال کیا جاتا ہے۔ صنعتی انتظامیہ، دفتری انتظامیہ، سرمایہ کا انتظامیہ، ملازمین کی بھرتی، ذخیرہ کا انتظامیہ، پیداوار کا انتظامیہ وغیرہ مختلف انتظامی میدان میں شماریات کا استعمال کیا جاتا ہے۔ اس علم کی روشنی میں ضرورت کے تحت وسائل پیداوار کا استعمال کر سکتے ہیں۔ کفایت شعاری اور وسائل کے صحیح استعمال کی ترکیب حاصل ہوتی ہے۔ جس سے ہر شعبہ کی کارکردگی ہمت افزاء ہوتی ہے۔

تعلیم Education

تعلیم ایک اہم شعبہ ہے، جس پر معشیت کے مختلف پلر قائم ہوتے ہیں۔ ہر ملک کی معاشی سرگرمی اس ملک کی تعلیمی سرگرمیوں پر منحصر ہوتی ہے۔ ملک میں شعبہ تعلیم کی کارکردگی کا پتہ لگانے، تعلیم کو بلند کرنے، تعلیمی سرگرمیوں کو رائج کرنے، سائنسی تحقیقاتی اداروں کو قائم کرنے، ملک کی شرح خواندگی میں اضافہ کرنے، تعلیم کو روزگار سے جوڑنے، کامیابی کا تناسب بڑھانے، سائنسی و سماجی علوم کو رائج کرنے، دیہی و شہری تعلیمی تفاوت کو کم کرنے وغیرہ مختلف مسائل کو حل کرنے میں شماریات اہم کردار ادا کرتا ہے۔

صنعت Industry

ہر ملک کی معشیت پر صنعتیں راست طور پر اثرانداز ہوتے ہیں۔ جن ممالک میں قدرتی ذخائر یا وسائل کافی مقدار میں پائے جاتے ہوں، وہاں پر صنعتوں کی کثرت پائی جاتی ہے۔ صنعتوں کے قیام سے قبل قدرتی وسائل کا پتہ لگانے، صنعتوں کے قیام کے مراحل کو تکمیل دینے، صنعتوں کے قیام سے حاصل ممکنہ نتائج کا تعین کرنے وغیرہ اہم امور ہیں جن کو صنعتی پالیسی اختیار کرنے کے دوران غور وخوص کیا جاتا ہے۔ ملک میں چھوٹے پیمانے کی صنعتوں، درمیانی یا بڑے پیمانے کی صنعتوں کے درمیان توازن پیدا کرنے، صنعتوں کی مقام پذیری، صنعتی پیداوار اور فروخت وطلب میں توازن پیدا کرنے وغیرہ مختلف منور مللات میں شماریات کارگرد ثابت ہوتی ہے۔

4.5 شماریات کی اہمیت (Importance of Statistics)

دنیا کا ہر علم اپنی اپنی منفرد اہمیت وافادیت رکھتا ہے۔ نارمل طور پر علم ریاضی کو اہمیت کے اعتبار سے منفرد مقام دیا جاتا ہے۔ اسی طرح سائنسی علوم اور سماجی علوم کو بھی مختلف میدانوں میں اہمیت دی جاتی ہے۔ علم شماریات بھی اہمیت وافادیت کے اعتبار سے کافی منفرد مقام رکھتا ہے۔ سابقہ جماعتوں میں حاصل نشانات کی بنیاد پر اپنی کامیابی کا فیصد معلوم کر چکے ہیں۔ ملک کی گھریلو پیداوار، روزگار، شرح پیدائش و اموات، درآمدات و برآمدات وغیرہ علم شماریات کی مدد سے بہتر طور پر پیش کر سکتے ہیں۔ علم شماریات کی اہمیت کو مختلف علوم کے پس منظر میں آسانی کے ساتھ سمجھ سکتے ہیں۔ ہر علم میں اس کا ایک ہی طریقے سے استعمال نہیں کیا جاتا ہے۔ مضمون کی اہمیت واستعمال کے پس منظر میں علم شماریات کا استعمال کیا جاتا ہے۔ ذیل میں مختلف مضامین کے پس منظر میں علم شماریات کی اہمیت کو بتلایا گیا ہے۔

1۔ معاشیات Economics

علم معاشیات ایک اہم علم ہے۔ کسی ملک کی آمدنی و اخراجات کی کیفیت، شرح پیدائش و اموات، صنعتوں کی تعداد، روزگار کی سطح، بیروزگاری کی سطح، شہری و دیہاتی سہولتیں، صنعتوں کی پیداوار، چھوٹے پیمانے کی صنعتیں، ریاست اور مرکزی حکومت کے موازنہ کی پیشکش سطح غربت، معیار زندگی، درآمدات و برآمدات وغیرہ علم معاشیات سے تعلق رکھتے ہیں۔ انہیں علم شماریات کی مدد سے بہتر طور پر تجزیہ کر سکتے ہیں۔ فیصد یا تناسب وغیرہ علم شماریات کے ہی حصے ہیں۔ اس کو نارمل فہم بنانے کے لیے اعداد کو گراف کی شکل میں بھی پیش کر سکتے ہیں۔ علم شماریات کے زبان میں پیش کرنے سے ہر شخص آسانی کے ساتھ سمجھ سکتا ہے۔ اسی طرح روزگار کو بڑھانے، شرح تعلیم یا خواندگی میں اضافہ، قومی آمدنی میں اضافہ، گھریلو پیداوار میں اضافہ وغیرہ وغیرہ مختلف میدانوں میں ترقی و نمو کے لیے حکمت عملی اختیار کرنے میں بھی علم شماریات کو استعمال کیا جاتا ہے۔ سرکاری اعتبار سے ملازمین میں تنخواہ کی ادائیگی، بونس کی اجرائی، رعایت، شرح سود کا تعین، بنک کاری نظام وغیرہ بھی علم شماریات کا استعمال کرتے ہیں۔ علم شماریات کا استعمال کے بغیر نتائج کا بہتر طور پر تجزیہ کرنا ممکن نہیں۔

2۔ کاروباری انتظامیہ (Business Management)

کاروبار ایک اہم معاشی سرگرمی ہے۔ ہر چھوٹے و بڑے تجارتی اداروں کا اہم مقصد منافع کمانا ہوتا ہے۔ اسی مقصد کے تحت مختلف تاجرین حکمت عملی اختیار کرتے ہیں۔ کاروبار مختلف اقسام کے ہیں۔ پھیری کا کاروبار، موسمی کاروبار، ترکاری و میوہ فروش، کرایہ دوکان، ٹیلر شاپ وغیرہ وغیرہ چند چھوٹے پیمانے کے کاروبار ہیں۔ اسی طرح کپڑے، جوتے، لوہے فولاد، کمپیوٹر ٹکنالوجی وغیرہ وغیرہ چند بڑے پیمانے کے کاروبار ہیں۔ شہروں میں واقع بڑے بڑے مال سنٹر بھی اہم کاروباری ادارے ہیں۔ آپ کو اسکول کے ایام یاد ہونگے۔ اسکول کے اوقات میں اسکول کے باہر آئسکریم اور پھلی، جام، بیر وغیرہ فروخت کرنے کے لیے معمولی تھیلے پر جمائے رکھتے ہیں۔ اسی طرح موسم گرما میں سڑکوں پر شربت کے مختلف اسٹال لگتے ہیں۔ کیا آپ جانتے ہیں یہ سب علم شماریات کا کس طرح استعمال کرتے ہیں۔ اوقات، قیمت خرید، قیمت فروخت کا تعین کرنا، شرح منافع و نقصان کا اندازہ لگانا، صارفین کو راغب کرنا، مقدار پیداوار کا تعین کرنا، لاگتوں پر کنٹرول کرنا، آمدنی و اخراجات میں توازن پیدا کرنا، موسم کے اعتبار سے کاروبار چلانا وغیرہ وغیرہ تمام امور علم شماریات سے تعلق رکھتے ہیں۔ جن کو تاجر مختلف زاویوں سے استعمال کرتا ہے۔

3۔ مالیاتی انتظامیہ (Finance Management)

ہر تاجر بیش ترین منافع حاصل کرنا چاہتا ہے۔ ہر تاجر کے پاس مالیہ کے محدود ذرائع ہوتے ہیں۔ محدود سرمایہ کو اس طرح سرمایہ کاری کرنا چاہتا ہے جس سے اس کو زیادہ سے زیادہ فائدہ حاصل ہو سکے۔ کاروبار کی نوعیت کے اعتبار سے سرمایہ کی رقم مختص کی جاتی ہے۔ چھوٹے کاروبار کے لیے سرمایہ کی کم ضرورت ہوگی۔ کاروبار میں ضرورت سے زائد رقم سرمایہ کاری کی جائے تو زائد سرمایہ سے تاجر کو فائدہ کے بجائے نقصان ہوگا۔ مثلاً گلی کوچوں میں پھر پھر کر آئسکریم فروخت کرنے والا تاجر یا کسی ٹفن سنٹر کا مالک مقدار پیداوار طے کرنے سے قبل وقت، حالات، مقامی صارفین کی تعداد، موسم، گرما، سرما، بارش، آب و ہوا وغیرہ مختلف عوامل کو مد نظر رکھتا ہے۔ ان ہی عوامل کے مد نظر سرمایہ بھی مشغول کرتا ہے۔ کیونکہ تعطیلات، برسات، شدید گرمی وغیرہ ایام میں صارفین کی کم تعداد پائی جاتی ہے۔ بڑے بڑے صنعتکار مالیہ

مشغول کرنے سے قبل ممکنہ فوائد اور اس کے متعلق مختلف حکمت عملیوں کو قطعیت دیتے ہیں۔ پھر کسی موزوں پیداوار پر سرمایہ کاری کو قطعیت دی جاتی ہے۔ کاروبار کی نوعیت یا کاروبار کے اعتبار سے سرمایہ مشغول کرنا کامیاب کاروبار کی نشانی ہے۔ فاضل رقومات کو محفوظ رکھا جاتا ہے تاکہ اس سے کچھ آمدنی حاصل کرسکے۔ غیر ضروری سرمایہ کاری سے بچنے کے لیے علم شماریات کے مختلف پہلوؤں کا استعمال کیا جاتا ہے۔

4۔ علم سماجیات Sociology

علم سماجیات ایک اہم علم ہے۔اس میں سماج سے متعلق مختلف امور کا تجزیہ کیا جاتا ہے۔ طبقات، سماج کے رہن سہن، طور طریقے، رسم و رواج، مرد اور خواتین کی تعداد، مذہبی رسم و رواج، عیدین و تہوار، گھریلو یا خاندانی طور طریقے، زبان، شرح خواندگی، روزگار کی سطح، بیروزگاری کی شرح وغیرہ وغیرہ مختلف سماجی پہلو ہیں جو سماجیات سے تعلق رکھتے ہیں۔ علم شماریات کی مدد سے سماج کے ان مختلف پہلوؤں کا مطالعہ کیا جاتا ہے۔ مختلف امور کو تناسب یا فیصد میں ظاہر کیا جاتا ہے۔ مملکت کے استحطام کے لیے کن کن تناسبوں میں اضافہ کرنا لازمی ہے۔ اس پر حکمت عملی اختیار کی جاتی ہے۔

5۔ علم تاریخ History

علم تاریخ ایک اہم مضمون ہے۔اس میں تاریخی حکمرانوں، بادشاہوں، راجاؤں کی حکومت اور طرز عمل کا مطالعہ کیا جاتا ہے۔ زمانۂ قدیم سے دور حاضر تک دنیا کے مختلف علاقوں کے مختلف حکمرانوں کی زندگی، طرز حکومت، طور طریقے وغیرہ وغیرہ کے متعلق مطالعہ کیا جاتا ہے۔ دور حاضر کی طرح تاریخی دور میں کتنے فیصد افراد تعلیم یافتہ تھے،افراد کا رہن سہن، طور طریقے، روزگار، بیروزگار، آمدنی، معیار زندگی، ترقیات، کاروبار، صنعتوں یا پیداواری اداروں کی نوعیت، مقدار پیداوار، درآمدات و برآمدات، کرنسی نظام، نظم وضبط، بندوبست وغیرہ مختلف تاریخی پس منظر کو شماریات کے پس منظر میں ظاہر کرتے ہیں۔ جس سے تاریخ کو سمجھنے میں مدد ملتی ہے۔

6۔ علم ریاضی Mathematics

علم ریاضی ایک اہم مضمون ہے۔ ہر دور میں اس مضمون کی اہمیت کو قبول کیا جاتا ہے۔ گنتی، اعداد کی درجہ بندی، حساب کتاب، زاویئے فیصد، تناسب وغیرہ علم ریاضی سے تعلق رکھتے ہیں۔ یہ تمام اعداد شمار میں استعمال کیے جاتے ہیں۔ علم ریاضی اور علم اعداد شمار کا آپس میں گہرا تعلق پایا جاتا ہے۔ان ہی کے مختلف پہلو علم ریاضی میں کافی استعمال میں آتے ہیں۔ علم ریاضی کے مسائل کو حل کرنے میں شماریات کی مدد لی جاتی ہے۔

4.6 شماریات کے آلات (Tools of Statistics)

شماریات ایک اہم علم ہے۔ جس میں ہر چھوٹے و بڑے تجارتی، قومی و بین الاقوامی منورملات کا تجزیہ کیا جاتا ہے۔ لین دین کے منورملات، کاروبار کی معاشی سرگرمیوں وغیرہ کے اعتبار سے شماریات کے مختلف ضوابط کا استعمال کیا جاتا ہے۔ اوسط، وسطانیہ، بہتانیہ، کوریلیشن، احتمال وغیرہ نارمل طور پر تجارت میں استعمال مختلف ضوابط ہیں، جس کی مدد سے نتائج اخذ ہوتے ہیں۔ اس کے علاوہ کاروبار یا غیر کاروباری منورملات کی جانچ و پڑتال کے لیے بھی ان کا استعمال کرتے ہیں۔ تحقیق میں مواد کو تجزیہ کرنے، نتائج اخذ کرنے کے لیے

شماریات میں مختلف ضابطہ تکنیک طریقے، ماڈلس کو استعمال کرتے ہیں۔انہیں آلاتِ شماریات کہتے ہیں۔ F.Test، T.Test، Regression Analysis،Chai Square Test، چنداہم شماریات کے آلات ہیں۔ان کی مدد سے حاصل مواد کا سائنسی نقطۂ نظر سے درست انداز میں تجزیہ کرتے ہوئے نتائج کو حاصل کر سکتے ہیں۔

ہتھوڑی، اسکروڈرائیور، پھوڑا، تھرمامیٹر، بادپیا، استھٹسکوپ، مختلف آلات یااوزار ہیں جس کی ایک طبعی شکل ہے۔ جن کو مختلف افعال جانچ وپڑتال میں استعمال کرتے ہیں۔ شماریات میں استعمال تکنیک یاضوابط یاماڈلس کی کوئی طبعی ساخت نہیں ہے۔ بلکہ غیر لمسی آلات ہیں۔ جن کو مطالعہ کرنے، تجزیہ کرنے، نتائج اخذ کرنے، منصوبہ بندی، کنڑول، کارکردگی وغیرہ کی جانچ کے لیے استعمال کرتے ہیں۔ ذیل میں چنداہم شماریات میں استعمال آلات کی وضاحت کی گئی ہے۔

F - Test.1   T - Test.2   Chi Square Test.3

Z - Test.4   Regression Analysis.5   Anova (Analysis of Variance).6

1۔ فشر ٹسٹ (F-Test): ایف ٹسٹ شماریات میں استعمال ہونے والا ایک اہم ٹسٹ ہے۔ یہ Null-Hypothesis میں استعمال ہوتاہے۔Georg W. Snotecor نےSir Ronald Fisher کے اعزاز میں اس ٹسٹ کوF-Test کانام دیاہے۔

اس ٹسٹ کے اہم مفروضات:

اس ٹسٹ کودومتغیرات(Two Variances) کے تقابل کے لیےاستعمال کیاجاتاہے۔ہر وقت اس کی قدر مثبت میں حاصل ہوتی ہے۔ دیے گئے دو اقسام کے نمونوں کے معیاری انحراف (Standard Deviation) کا تقابل کرنے اور ان کی صحت (Variabllity) کی جانچ کے لیےاستعمال کرتے ہیں۔

2۔ ٹی - ٹسٹ (T - Test): ٹی ٹسٹ شماریات میں استعمال ایک اہم ٹسٹ ہے۔ T Distribution کو سب سے پہلے Helmert and Luroth نے 1876ء میں پیش کیا۔ Willian Sealy Gosset نے 1908 میں سب سے پہلے Biometrika میں شائع کیا۔ ٹی ٹسٹ کو Students T. Distribution بھی کہاجاتاہے۔ جہاں پر یہ ملازم تھے وہاں پر اپنے نام کے بجائے محقق یاملازمین اپنے قلمی نام سے مقالے شائع کرنے کی اجازت تھی۔ گوسٹ (Gosset) آئیر لینڈ کے Guinness Brewery میں ملازم تھے۔ چھوٹے نمونہ بندی کے مسائل کو حل کرنے میں دلچسپی رکھتے تھے۔

3۔ کائی اسکوائر ٹسٹ(Chi-Square Test): کائی اسکوائر ٹسٹ شماریات میں استعمال ہونے والاایک اہم ٹسٹ ہے۔اس ٹسٹ کو 1900 ویں صدی میں پیرسن کے کائی اسکوائر پر اپنا تحقیق مقالہ شائع کیا۔ جو جدید شماریات(Modern Statistics) کے بنیاد میں فراہم کیے ہیں۔

4۔ زیڈ ٹسٹ(Z - Test): زیڈ ٹسٹ شماریات میں استعمال ایک اہم آلہ ہے۔ نمونوں کی تعداد زیادہ ہونے پر اس ٹسٹ کواستعمال کیا جاتا ہے۔ نارمل طور پر نمونوں کی تعداد30 سے زیادہ ہونے پر زیڈ ٹسٹ کو استعمال کیا جاتا ہے۔ جب دو اقسام کے نمونوں کے جمیعت (Population) کے اوسط(Mean) اور تغیرات(Variance) معلوم ہوں تب اس کواستعمال کیاجاتاہے۔

تغیرات کا تجزیہ Analysis of Variance (ANOVA)

تغیرات کا تجزیہ علم شماریات میں استعمال ایک اہم آلہ ہے۔اس کو ماہر شماریات رونالڈ فشر (Ronald Fisher) نے پیش کیا۔ یہ Law of Total Variance کے اصول پر مبنی ہے۔ سر رونالڈ فشر نے سب سے پہلے اصطلاح (Variance) کو استعمال کیا۔ 1925 میں اپنی کتاب "Statistical Method for Research Work" میں Analysis of Variance پر تفصیلی روشنی ڈالی۔ ANOVA کو Fisher Analysis of Variance بھی کہتے ہیں۔ دو یا زائد گروپ کے درمیان پائے جانے والے تعلق کو اخذ کرنے کے لیے ANOVA ٹسٹ کا استعمال کیا جاتا ہے۔ تین یا اس سے زائد متغیرات (Variables) کو ٹسٹ کرنے کے لیے اس کا استعمال کیا جاتا ہے۔

ANOVA دو اقسام کے ہیں۔

1. One way ANOVA Test
2. Two way ANOVA Test

4.7 مرکزی رجحان کے معنی اور تعریف

(Meaning and Definitions of Measures of Central Tendency)

علم شماریات ایک اہم علم ہے جو کسی ملک، صنعت، فیکٹری، تعلیمی ادارے، تجارت کے منور مللات کو آسان اعداد میں پیش کرنے کے لیے استعمال کرتے ہیں۔ مرکزی رجحان کے مفہوم کو ذیل مین وضاحت کی گئی ہے۔

1۔ مرکزی رجحان کے پیمانے علم شماریات میں استعمال ہونے والے واحد اعداد ہیں جو متعلقہ اعداد کی ترجمانی کرتے ہیں۔

2۔ مرکزی رجحان کے پیمانے حاصل مواد سے اخذ کردہ واحد اعداد ہوتے ہیں جو مواد کی عکاسی کرتے ہیں۔

3۔ مرکزی رجحان کے پیمانے شماریات (Descriptive Statistic) کی ایک شاخ ہے حاصل مواد کا درمیانی عدد ہوتا ہے جو تمام اعداد کی نمائندگی کرتا ہے۔

4۔ ایسا عدد جو مواد کی خصوصیات اور صفات کو ظاہر کرتا ہے جو مرکزی رجحان کے پیمانے کہلاتے ہیں۔

5۔ ایسا عدد جو اعداد کے اوسط یا درمیانی عدد کو ظاہر کرتا ہے وہ مرکزی رجحان کا پیمانہ کہلاتا ہے۔

4.7.1 نوعیت (Nature)

مرکزی رجحان کے پیمانے علم شماریات کے اہم پہلوؤں میں سے ایک ہے۔ سماجی علوم کے ہر مضمون میں استعمال میں آتے ہیں۔ مرکزی رجحان کے پیمانے دراصل حاصل مواد کا درمیانی یا وسطی عدد ہوتا ہے۔ دیگر تمام اعداد اس ہی اعداد کے اطراف پائے جاتے ہیں۔ یہ سلسلہ وار مواد کا درمیانی عدد ہوتا ہے۔ یہ مواد کا خلاصہ پیش کرتا ہے یہ دئے گئے اعداد کا اوسط ہوتا ہے۔اس اعداد کی مدد سے مجموعی کارکردگی کا

اندازہ قائم ہوتا ہے اکثر ادارے ان ہی اعداد کو بنیاد بناتے ہوئے اپنی پیداوار کی مثبت عملی اختیار کرتے ہیں، تعلیمی ادارے سالانہ کارکردگی کو ظاہر کرنے کے لیے ان اعداد کا استعمال کرتے ہیں۔ یہ انفرادی اور اجتماعی طور پر استعمال کیا جاسکتا ہے۔

چھوٹے پیمانے کے تاجرین زبانی حساب کتاب کے ذریعہ ان کی یومیہ آمدنی واخراجات کا حساب کرلیتے ہیں۔ چھوٹے وبڑے پیمانے کی صنعتوں کے لیے یہ موزوں نہیں ہے۔ صنعتوں کے پیداوار بھاری مقدار میں ہوتے ہیں۔ یہ سالانہ کھاتے تیار کرتے ہیں۔ صنعتیں بازاری طلب اور منصوبہ بندی کو مد نظر رکھتے ہوئے پیداوار جاری کرتے ہیں پیداوروں منصوبہ بندی کے دوران، مقدار پیداوار، لاگت، بازاری طلب ، قیمت، شئے کی نوعیت، ساخت وغیرہ مختلف عوامل کو پیش نظر رکھا جاتا ہے۔ اسکول کے بچوں کے لیے جوتے تیار کرنے والی صنعتیں بغیر حساب کے جوتے تیار نہیں کرتے بلکہ جوتے تیار کرنے پر قبل جوتوں کے مختلف سائز اور مدارس کی تعداد، طلبا کی تعداد، علاقے، شہر، دیہات عوام کی آمدنی کی سطح وغیرہ مختلف عوامل کو مد نظر رکھتے ہوئے جوتوں کے سائز اور ان کی تعداد کا تعین کرتے ہیں۔ جوتوں کے سائز اور ان کی تعداد کے تعین کے دوران طلبا کی اور طالبات کی تعداد کا بھی تجزیہ کرتے ہیں۔ ایک موزوں مقدار کا تعین کرتے ہوئے پیداوار کو جاری کرتے ہیں۔ اس طرح ہر چھوٹے وبڑے صنعتیں حسابات کی مدد سے مقدار پیداوار کا تعین کرتے ہیں اوسط، وسطانیہ وبہتانیہ جامع انداز میں مختلف بھاری پیمانے کے صنعتوں میں کثرت سے استعمال ہوتے ہیں۔

4.7.2 اقسام (Types)

اعداد کا تجزیہ کرتے ہوئے نتائج اخذ کرنے میں علم شماریات اہم کردار ادا کرتا ہے۔ اعداد کی نوعیت کے اعتبار سے ان کا تجزیہ کیا جاتا ہے۔ مرکزی رجحان کے تین اہم اقسام ہیں۔

1۔ اوسط　　2۔ وسطانیہ　　3۔ بہتانیہ

ہر ایک کے متعلق تفصیلی مطالعہ کیجئے۔

1۔ اوسط　　Arithemetic mean or Average

اوسط مرکزی رجحان کی پیائش کی ایک اہم قسم ہے۔ اوسط دئے گئے اعداد کی مرکزی عدد کے حیثیت سے نمائندگی کرتا ہے۔ جس ضابطے کی مدد سے محسوب کیا جاتا ہے اوسط کو حسب ذیل دو حصوں میں تقسیم کیا جاتا ہے۔ اکائی 10 میں اوسط پر تفصیلی روشنی ڈالی گئی ہے۔

(i) ہندسی اوسط (Gementic Mean)　　(ii) ہارمونک اوسط (Harmonic Mean)

(iii) اوزنی اوسط (Weighted Arithmetic Mean)

اوزنی اوسط　　Weighted Arithmetic Mean

اوزنی اوسط مرکزی رجحان کے پیمانوں میں اوسط کی ایک اہم قسم ہے۔ جس میں دیے ہوئے اعداد کی وزن کی بنیاد پر اوسط محسوب کیا جاتا ہے۔ اوزانی اوسط کو محسوب کرنے کا ضابطہ

$$\text{اوزانی اوسط} = \frac{\sum WX}{WX}$$

مثال 1: ذیل میں ایک گھر کے ماہانہ اخراجات کی تفصیل دی گئی ہے۔ جس کی مدد سے اوزانی اوسط محسوب کیجئے۔

متفرق	لباس	تعلیم	ادویات	غذا	اخراجات
10	20	20	20	30	فیصد
1	1	1	1	2	وزن

حل: اوزانی اوسط Weighted Average

WX	وزن	فیصد(W)	تفصیلات(X)	سلسلہ نمبر
60	2	30	غذا	1
20	1	20	ادویات	2
20	1	20	تعلیم	3
20	1	20	لباس	4
10	1	10	متفرق	5
130	6			

$$WX = \frac{\sum WX}{WX} = \frac{130}{6} = 21.6$$

وزنی اوسط $WX = 21.6$

مثال 2: ذیل میں ایک کالم میں طلبا کی تعداد اور ان کی کامیابی کا فیصد دیا گیا ہے۔ جن کی مدد سے وزنی اوسط محسوب کیجئے۔

ایم سی اے	ایم بی اے	ایم کام	ایم اے	کورس
400	300	200	100	طلبا کی تعداد
95	85	90	80	کامیابی کا فیصد

حل: اوزانی اوسط کی تحسیب (Calculation of Weighted Average)

WX	کامیابی کی فیصد	طلبا کی تعداد	کورس	سلسلہ نمبرات
8000	80	100	M.A	1
18000	90	200	M.Co,	2

3	M.B.A	300	85	25500
4	M.C.A	400	95	38000
		1000		89500

$$WX = \frac{\sum wx}{\sum wx}$$

$$= \frac{89500}{1000}$$

وزنی اوسط= 89.5

2۔ وسطانیہ Median

وسطانیہ مرکزی رجحان کا ایک اہم قسم کا پیمانہ ہے جو دیے گئے اعداد کی درمیانی یا وسطی نقطہ کی حیثیت سے نمائندگی کرتا ہے۔یعنی وسطانیہ دیے گئے اعداد کا بالکل درمیانی یا وسطی نقطہ یا عدد ہوتا ہے۔دئے گئے اعداد کو گھٹتی ہوئی یا بڑھتی ہوئی ترتیب میں ترتیب دینے پر ان کے درمیانی عدد کو وسطانیہ کہتے ہیں۔دیے گئے اعداد کی تعداد طاق ہونے پر وسطانیہ کی قدر آسانی سے اخذ کی جاسکتی ہے۔اس کے برخلاف دئے گئے اعداد کی تعداد جفت میں ہوں تب درمیانی دو اعداد کا اوسط وسطانیہ (Median) کہلاتا ہے۔اس میں درمیانی دو اعداد کے حاصل جمع کو'2'سے تقسیم کیا جاتا ہے۔جس پر وسطانیہ حاصل ہوتا ہے۔اکائی 11 میں وسطانیہ پر تفصیلی روشنی ڈالی گئی ہے۔

3۔ بہتانیہ یا کثیریہ Mode

بہتانیہ یا کثیریہ مرکزی رجحان کی پیائش کا ایک اہم قسم ہے۔دیے گئے اعداد ہیں سب سے زیادہ مرتبہ دہرائے جانے والے عدد کو بہتانیہ یا کثیریہ (Mode) کہتے ہیں۔دیے گئے اعداد منفی، مثبت، طاق، جفت کسی بھی نوعیت کے ہو سکتے ہیں۔اعداد میں سب سے زیادہ مرتبہ دہرانے والا عدد یا کثرت سے آیا ہوا عدد بہتانیہ کہلاتا ہے۔اعداد کی تعداد کم ہونے پر مشاہدے کی مدد سے بہتانیہ کو محسوب کیا جاسکتا ہے۔اسی طرح اعداد کی تعداد کثیر مقدار میں ہونے پر ضابطہ کے استعمال کی مدد سے بہتانیہ کو محسوب کیا جاتا ہے۔اکائی 12 میں بہتانیہ پر تفصیلی روشنی ڈالی گئی ہے۔

4.7.3 حسابی اوسط کے معنی اور تعریف (Meaning and Definition of Arithmetic Mean)

حسابی اوسط کو مرکزی رجحان کا ایک اہم پیمانے کے اس کی تفہیم کے لیے ذیل میں چند اہم تعریفات دیے گئے ہیں۔

1۔ حسابی اوسط کو آسان الفاظ mean یا Average یا اوسط کہا جاتا ہے۔

2۔ ایسا عدد جو دیے گئے اعداد کے مجموعے کو تعداد سے تقسیم کرنے پر حاصل ہوتا ہے وہ حسابی اوسط کہلاتا ہے۔

3۔ ایسا عدد جو دیے گئے مواد کی مرکزی حیثیت سے نمائندگی کرتا ہو وہ حسابی اوسط کہلاتا ہے۔

4۔ حسابی اوسط دیے گئے اعداد کے مجموعہ کی نمائندگی کرتا ہے۔

مثال 1: ایک جماعت کے سائنس مضمون میں 10 طلبا کے حاصل نشانات دئے گئے ہیں جن کی عدد سے اوسط محسوب کیجئے۔

72 50 33 71 65 42 45 61 42 89

حل: دس طلبا کے حاصل نشانات کو جمع کر لیا جائے۔

تمام طلبا کی نشانات کا کل یا مجموعہ کو طلبا کی تعداد سے تقسیم کرنے پر اوسط حاصل ہوگا۔

$$= 72 + 50 + 33 + 71 + 65 + 42 + 45 + 61 + 42 + 89 = 570$$

$N = 10$ طلبا کی تعداد $\sum X = 570$ نشانات کا کل $57 =$ اوسط $= 570/10 =$ اوسط

4.7.3.1 اوسط محسوب کرنے کا بالراست طریقہ

(Indirect Method of Calculation of Arthmetic mean)

اوپر کی مثال کو بالراست طریقہ پر اوسط کو محسوب کیا جاتا ہے۔ اس طریقہ کے تحت ذیل مراحل ہیں۔

1۔ دیے گئے اعداد میں سے کسی ایک عدد کو فرض کر لیں۔

2۔ مفروضہ عدد میں سے دیے گئے اعداد کو تفریق کریں جس سے انحراف حاصل ہوتا ہے۔

3۔ انحرافات کو جمع کریں۔

4۔ ذیل کے ضابطہ پر عمل کریں۔

$$X= A+ \sum d/N$$

$$72+50+33+71+65+42+45+61+42+89$$

ان میں سے عدد 65 کو فرض کیا گیا

دیے گئے اعداد کو اس عدد میں سے تفریق کرتے ہوئے مثبت (+) اور منفی (-) علامات کے ساتھ اعداد کو ظاہر کریں۔ جمع اعداد کو جمع کر لیا اور تفریق علامت کے اعداد کو علیحدہ جمع کریں۔ جمع علامت کے اعداد کا مجموعہ +37 اور تفریق علامات کے اعداد کا مجموعہ – 117 ہے ان دونوں کا فرق -80 ہے۔ منفی علامات کے اعداد زیادہ ہیں اس لیے منفی علامت میں ظاہر کیا جاتا ہے۔

ذیل کے ضابطہ میں قیمتیں درج کریں۔

$$X = A + \sum d / \sum N$$

$$= 65 + \frac{-80}{10}$$

$$= 65 - 8$$

$\bar{\sum}$ اوسط $= 57$

مندرجہ ذیل کے راست یا بالراست دونوں طریقوں سے اوسط 57 حاصل ہوتا ہے۔

4.7.3.2 غیر مسلسل سلسلہ میں اوسط کا تخمینہ (Calculation of Mean in case of Discrete Series)

تعددی سلسلہ (Discrete Series) علم شماریات میں مواد کو پیش کرنے کا اہم سلسلہ ہے اس طریقہ میں معطیات (Data) اور ان کی تعداد کو جدول کی شکل میں مواد پیش کیا جاتا ہے اس جدول میں وقفہ جماعت (Class Interval) نہیں پایا جاتا اس میں مواد کو دو کالم میں پیش کیا جاتا ہے۔ پہلا کالم میں عدد اور دوسرا کالم تعداد (Frequency) بتلائی جاتی ہے۔

تعددی سلسلہ (Discrete Series) میں راست طریقے سے اوسط محسوب کرنے کا طریقہ (Direct Method):

راست طریقہ کے تحت اوسط محسوب کرنے کے لیے معطیات اور تعداد کے حاصل ضرب کے مجموعہ کو تعدد کی جملہ تعداد سے تقسیم کرنے پر اوسط حاصل ہوتا ہے۔ اس طریقہ پر اوسط محسوب کرنے کے لیے اس ضابطہ کو استعمال کریں۔

$$\bar{\Sigma} = \sum fx / \sum F$$

مثال 2: ذیل میں شماریات میں حاصل نشانات اور طلبا کی تعداد دی گئی ہے۔ ان کی مدد سے اوسط محسوب کیجئے۔

نشانات	35	55	60	65	70	80	85	90
طلبا کی تعداد	4	5	8	6	9	10	6	2

حل : راست طریقہ کے تحت اوسط محسوب کرنا۔

(X)	(F)	(fx)
35	4	140(35x4)
55	5	275(55x5)
60	8	480(60x8)
65	6	390(65x6)
70	9	630(70x9)
80	10	800(80x10)
85	6	510(85x6)
90	2	180(90x2)
	50=∑x	340=∑fx

$$\bar{\Sigma} = \sum fx / \sum x$$

$= 3405 / 50$

اوسط$\overline{\sum} = 68.1$

بالراست طریقے پراوسط محسوب کرنا(Indirect Method)

اس طریقہ میں دیے گئے معطیات میں سے کسی ایک عدد کو فرض کرتے ہوئے فرق شدہ اعداد کو منفی یا مثبت علامت کے ساتھ ظاہر کریں پھران کو تعدد(F) سے ضرب دیں۔اور ضابطہ میں قیمت درج کریں۔ بالراست طریقہ پر اوسط محسوب کرنے کا ضابطہ۔

$$X = A + \sum fd / \sum f$$

جہاں پر A=فرض کیا ہوا عدد fd=حاصل ضرب کا مجموعہ f=تعدد کا حاصل جمع

مثال 3 :مثال تین کو استعمال کرتے ہوئے بالراست طریقہ پر اوسط محسوب کیجئے۔

حل :بالراست طریقہ پر اوسط محسوب کرنے کا طریقہ:

نشانات(X)	تعدد(F)	d = 65	Fd
35	4	-3-	-120 (4x -30)
55	5	-10	-50 (5x-10)
60	8	-5	-40(8x-5)
65	6	0	0(6x0)
70	9	+5	+45(9x5)
80	10	+15	+150(10x15)
85	6	+20	+120 (6x20)
90	2	+20	+50(2x25)
	50=sf	-0	-210 +365 155=∑fd

$$\bar{\sum}=A+\sum fd/\sum f$$
$$= 65 + 155/50$$
$$=65 + 3.1$$
اوسط= 68.1

مندرجہ بالا پر دونوں طریقوں پر اوسط 68.1 حاصل ہوا ہے۔

4.7.3.3 مسلسل سلسلہ کے تحت اوسط کو محسوب کرنا

(Calculation of Arthmetic Average in case of Continuous Series)

مسلسل سلسلہ (Continuous Series) علم شماریات میں مواد کو ترتیب وار پیش کرنے کا ایک اہم طریقہ ہے اس طریقہ میں مواد کو وقفہ جماعت (Calss Interval) تعدد (Frequency) کے ساتھ بتلایا جاتا ہے۔ معطیات کی تعداد زیادہ ہونے کی صورت میں جدول وقفہ جماعت یا مسلسل سلسلہ کو استعمال کیا جاتا ہے۔ دیے گئے معطیات کی تعداد اور ان کی اونچی اور نچلی سطح کو مد نظر رکھتے ہوئے وقفہ جماعت کی موزوں سطح کا تعین کیا جاتا ہے نارمل طور پر 5 یا 10 وقفہ جماعت کے ساتھ جدول تیار کیا جاتا ہے۔ راست اور بالراست اور اختصاری طریقہ (Step Devition) کے طریقہ پر اوسط کو محسوب کیا جا سکتا ہے۔

مسلسل سلسلہ میں راست طریقہ کے تحت اوسط محسوب کرنا (Calculation of Mode in Direct Method)

مسلسل سلسلہ یا راست طریقہ پر اوسط محسوب کرنے کے لیے ذیل کے مراحل پر عمل کریں۔

1۔ وسطی نقطہ محسوب کریں

2۔ وسطی نقطہ اور تعدد کو ضرب دیں اور اس کا کل محسوب کریں۔

3۔ ضابطہ میں قدریں کو درج کریں۔

$$= \sum mf/\sum f\bar{X}$$

جہاں پر $\bar{X}$= اوسط کی علامت ہے

mf کا حاصل جمع $=\sum mf$

$\sum f$= تعدد کا حاصل جمع

مثال 4: ذیل میں ایک صنعت میں کام کرنے والے مزدوروں کی توسیع تنخواہوں کے تفصیلات دیے گئے ہیں ان کی مدد سے اوسط محسوب کیجیے۔

وقفہ جماعت تنخواہ (f)	1000-1500	1500-2000	2000-2500	2500-3000	3000-3500	3500-4000
ملازمین کی تعداد (x)	5	7	12	10	9	7

حل: اوسط محسوب کرنے کا طریقہ ذیل میں بتلایا گیا ہے۔

وقفہ جماعت تنخواہیں (C-I)	ملازمین کی تعداد (F)	وسطی نقطہ (M)	2 x3 mf
1000-1500	5	1250	(1250x5) 6250
1500-2000	7	1750	(1750x7) 12250
2000-2500	12	2250	(2250x12) 27000
2500-3000	10	2750	(2750x10) 27500
3000-3500	9	3250	(3250x9) 29250
3500-4000	7	3750	(3750x7) 26250
Total	50		∑mf=128500

وسطی نقطہ: وقفہ جماعت کے اوپری اور نچلی سطح کے حاصل جمع کو 2 سے تقسیم کرنے پر وسطی نقطہ حاصل ہوتا ہے۔

مثلاً: 1000 + 1500=2500/2 = 1250

اوسط= Smf/Sf = 128500 / 50 = 2570

اوسط اجرت $\bar{X}$ = 2570 روپے

بالراست طریقہ پر اوسط محسوب کرنے کا طریقہ (Calculation of Average by Indirect Method)

بالراست طریقہ پر بھی اوسط محسوب کیا جا سکتا ہے اس کے لیے ذیل کے مراحل پر عمل کریں:۔

1۔ وقفہ جماعت کا وسطی نقطہ محسوب کریں۔

2۔ وسطی نقطہ میں سے کسی ایک عدد کو فرض کرتے ہوئے اعداد کے فرق کو منفی و مثبت کے ساتھ ظاہر کریں۔

3۔ حاصل عدد کو تعدد سے ضرب دیں۔

4۔ ضابطہ میں قدریں درج کریں۔

وقفہ جماعت اجرت روپے (x)	تعداد (f)	وسطی نقطہ (m)	d=2250	Fd
1000-1500	5	1250	-1000	-5000
1500-2000	7	1750	-500	-3500

2000-2500	12	2250	0	0
2500-3000	10	2750	+500	+5000
3000-3500	9	3250	+1000	+9000
3500-4000	7	3750	+1500	+10500
				-8500 +24500 +16000

طریقہ عمل:(1)وسطی نقطہ 2,250روپے کو فرض کیا گیا ہے۔

(2)فرض کئے ہوئے عدد سے تفریق کرتے ہوئے مثبت و منفی کے ساتھ اعداد

(3)تعدد کے ساتھ حاصل ضرب محسوب کریں۔

$$\bar{X}_{\text{اوسط}} = A + \frac{\Sigma fd}{\Sigma f}$$

$$= 2250 + 16000 / 50$$

$$= 2250 + 320$$

$\bar{X}$ اوسط= 2250روپے

اوسط محسوب کرنے کا اختصار یا انحراف(Step Deviation)کا طریقہ:

اختصار کے طریقہ کے تحت انحراف کے اعداد(d) کو کسی موزوں عدد سے تقسیم کیا جاتا ہے اس تقسیم شدہ عدد کو d1 سے ظاہر کیا جاتا ہے ذیل میں اختصار کے طریقہ کے تحت اوسط کو محسوب کرنے کا طریقہ کی وضاحت کی گئی ہے۔

1۔ وقفہ جماعت سے وسطی نقطہ محسوب کریں

2۔ وسطی نقطہ سے کسی ایک عدد کو فرض کرتے ہوئے اعداد کے فرق کو منفی و مثبت کے ساتھ ظاہر کریں اور ان اعداد کو کسی موزوں عدد سے تقسیم کریں۔

3۔ تقسیم سے حاصل اعداد کو تعدد سے ضرب دیں۔

4۔ ضابطہ میں قدریں درج کریں۔ بالراست طریقہ پر اوسط کی تحسیب:

وقفہ جماعت اجرت	تعداد(f)	وسطی نقطہ(M)	d=2250	$d^1/100$	fd^1
1000-1500	5	1250	-1000	-10	-50

1500-2000	7	1750	-500	-5	-35
2000-2500	12	2250	0	0	0
2500-3000	10	2750	+500	+5	+50
3000-3500	9	3250	+1000	+10	+90
3500-4000	7	3750	+1500	+15	+105
					-85 +245 $\sum fd^1$+160

1۔ اس مثال میں وسطی نقطہ 2.250 کو فرض کرتے ہوئے تفریق کرنے کے بعد حاصل امداد کو 100 سے تقسیم کیا گیا ہے اس کالم کو d^1 سے ظاہر کیا گیا۔

2۔ تقسیم کے بعد حاصل اعداد کو تعدد سے ضرب دیں اور ضابطہ میں قدریں درج کریں۔

جہاں پر $\sum fd1 = 160$

$\sum R = 50$

$C =$

$$= A + \frac{\sum fd1}{\sum f} X C \bar{X}$$

$$= 2250 + \frac{160}{50} \text{ x } 100$$

$$= 2250 + 3.2 \text{ x } 100$$

$$= 2250 + 320$$

$$\bar{X}_{\text{اوسط}} = 2570$$

4.7.3.4 حسابی اوسط کے خصوصیات (Properties or Features of Arthmetic Mean)

حسابی اوسط کے چند اہم خصوصیات کو ذیل میں بتلایا گیا ہے۔

1۔ حسابی اوسط میں دئے گئے معطیات یا اعداد کے انحرافات کا مجموعہ ہمیشہ صفر ہوتا ہے۔ یعنی $\Sigma(X-X)=0$

2۔ حسابی اوسط کے انحرافات کے مربعوں (Square Deviation) کا حاصل جمع اقل ترین ہوتا ہے یعنی $\Sigma(X-X)^2$

3۔ حسابی اوسط کے ماول کی عدد یا مدات کا جمع کرنے یا تفریق کرنے سے حسابی اوسط کی قدر متاثر نہیں ہوتی۔

4۔ حسابی اوسط دئے گئے معطیات کے درمیانی رشتہ یا مقام رکھتا ہے۔

5۔ حسابی اوسط کی قدر منفی، مثبت یا تفریق ہو سکتی ہے۔

حسابی اوسط کے خوبیاں Merits of Arthmetic Mean

حسابی اوسط کے اہم خوبیوں کو ذیل میں بتلایا گیا ہے۔

1۔ حسابی اوسط کو حسابی ضابطوں کی مدد سے آسانی سے محسوب کیا جا سکتا ہے۔

2۔ حسابی اوسط کو ہر تعلیم یافتہ اور غیر تعلیم یافتہ شخص آسانی سے سمجھ سکتا ہے۔

3۔ حسابی اوسط کو تعلیمی، غیر تعلیمی، کھیل کود، ذخائر ملکی منور ملات، درآمدت برآمدات منصوبہ بندی، محصول سازی، قیمت سازی، اخراجات میں تخفیف، نفع ونقصان وغیرہ ہر میدان میں استعمال کیا جا سکتا ہے۔

4۔ حسابی اوسط قابل اعتبار ہوتا ہے۔

5۔ حسابی اوسط دیے گئے تمام مدات کا نمائندہ ہوتا ہے۔

6۔ حسابی اوسط کو مستقبل منصوبہ بندی وغیرہ کے لیے بھی استعمال کیا جاتا ہے۔

7۔ ہر قسم کے مواد کا حسابی اوسط محسوب کیا جا سکتا ہے۔ مختلف طریقوں سے حسابی اوسط کو محسوب کرنے پر حسابی اوسط کی مساوی قدر حاصل ہوتی ہے۔

خامیاں (Demerits)

حسابی اوسط کے چند اہم خامیوں کو ذیل میں بتلایا گیا ہے۔

1۔ حسابی اوسط کو محسوب کرنے کے دوران معمولی سے غلطی بھی غلط نتائج حاصل ہو سکتے ہیں۔

2۔ حسابی اوسط کو صرف مقداروں (Quantitive) صورت میں اخذ کیا جا سکتا ہے لیکن یہ معیاروں (Qualitative) منور ملات جیسے ایمانداری عادت وغیرہ میں حسابی اوسط کو محسوب کرنا ممکن نہیں۔

3۔ دیے گئے مواد میں کسی ایک مواد کی عدم دستیابی سے صحت کے ساتھ حسابی اوسط کو محسوب کرنا ممکن نہیں۔

4۔ اوسط کا ترسیمی اظہار ممکن ہیں۔

4.7.4 وسطانیہ کے معنی وتعریف (Meaning and Definitions of Median)

وسطانیہ علم شماریات میں استعمال ہونے والا اہم پیمانہ ہے

☆ وسطانیہ کی تفہیم کے لیے اس کے چند تعریفات کو ذیل میں دیے گئے ہیں۔

☆ وسطانیہ دیے گئے اعداد یا معطیات کا وسطی نقطہ ہوتا ہے۔

☆ وسطانیہ مرکزی رجحان کا ایک اہم پیمانہ ہے دیے گئے اعداد کو گھٹتی ہوئی یا بڑھتی ہوئی ترتیب میں ترتیب دینے پر حاصل درمیانی عدد کو وسطانیہ کہتے ہیں۔

☆ وسطانیہ اعداد کا درمیانی عدد ہوتا ہے جس کہ دائیں اور بائیں جانب مساوی مقدار میں اعداد پائے جاتے ہیں۔

☆ وسطانیہ اعداد کا درمیانی عدد ہوتا ہے جو دیے گئے مورد کو دو حصوں میں تقسیم کرتا ہے۔

4.7.4.1 وسطانیہ کی نوعیت(Nature of Median)

وسطانیہ مرکزی رجحان کے پیمانوں میں سے ایک اہم پیمانہ ہے وسطانیہ ایک ایسا وسطی نقطہ ہے جو دیے گئے اعداد کے بالکل وسطی مقام پر ایک نصف حصہ اس سے اوپر اور باقی نصف حصہ اس کے نچلے پایا جاتا ہے وسطانیہ دیے گئے معطیات کو دو برابر حصوں میں تقسیم کرتا ہے وسطانیہ سے اوپری اعداد اور نچلے اعداد کے تجزیہ سے مختلف پہلو وُجود میں آتے ہیں۔

Al-Biruni نے Mid-Range کی اصطلاح استعمال کی جو وسطانیہ کو ظاہر کرتی ہے۔وسطانیہ کا تصور 13ویں صدی عیسوی میں زیر بحث رہی لیکن وسیع پیمانے پر سائنٹفک انداز میں اس کو وسعت حاصل نہیں ہوئی۔ برطانوی ماہر ریاضی Edward Wright نے اپنی کتاب "Certaine Errors in Navigation" میں 1599 میں پہلی مرتبہ وسطانیہ کو پیش کیا۔ 1843 میں Antoine Augustin Courrot نے پہلی مرتبہ اصطلاح وسطانیہ (Valewr Median) median کو احتمالی تقسیم میں دو مساوی حصوں میں تقسیم کے لیے استعمال کیا تھا۔

4.7.4.2 وسطانیہ کی خصوصیات(Properties or Features of Median)

وسطانیہ کے خصوصیات کو ذیل میں بتلایا گیا ہے۔

1۔ وسطانیہ دیے گئے اعداد کا درمیانی عدد ہوتا ہے۔

2۔ وسطانیہ کو محسوب کرنے کے لیے دیے گئے اعداد کو بڑھتے ہونے یا گھٹتی ہوئی ترتیب میں ترتیب دیا جاتا ہے۔

3۔ وسطانیہ اعداد یا مواد (Data) کی تعداد سے متاثر ہوتا ہے۔

4۔ وسطانیہ کو ترسیم کی مدد سے ظاہر کیا جا سکتا ہے۔

5۔ معیاری صفت کو ظاہر کرنے کے لیے وسطانیہ ایک اہم پیمانہ ہے۔

4.7.4.3 وسطانیہ کے فوائد و خامیاں(Merits and Demerits of median)

وسطانیہ علم شماریات کا ایک اہم پیمانہ ہے اس کے چند اہم خوبیوں یا فوائد کو ذیل میں بتلایا گیا ہے۔

فوائد Merits

1۔ وسطانیہ کو محسوب کرنا کافی آسان ہے اس کی تحسیب میں کوئی مشکلات نہیں ہے۔

2۔ وسطانیہ کی تفہیم کافی آسان ہیں۔

3۔ وسطانیہ کو گراف میں آسانی سے پیش کر سکتے ہیں۔

4۔ وسطانیہ میں دیے ہوئے اعداد کا وسطی یا درمیانی عدد ہوتا ہے یہ اوسط سے غیر متاثر ہوتا ہے۔

5۔ وسطانیہ کو گروہ اور غیر گروہ بند مواد دونوں صورتوں میں آسانی سے محسوب کر سکتے ہیں۔

6۔ وسطانیہ کی معیاری پیمائش میں بھی استعمال کیا جاتا ہے۔

7۔ اوسط تعلیم یافتہ شخص بھی وسطانیہ کو آسانی محسوب کر سکتا ہے۔

خامیاں Demerits

وسطانیہ کے چند اہم خامیوں کو ذیل میں بتلایا گیا ہے۔

1۔ وسطانیہ کو محسوب کرنے کے لیے دئے ہوئے اعداد کو گھٹتی ہوئی یا بڑھتی ہوئی ترتیب میں ترتیب دینا لازمی ہے بصورت دیگر اس کو محسوب کرنا کافی مشکل ہے۔

2۔ جفت اعداد کی صورت میں وسطانیہ کو محسوب کرنا کافی آسان ہے لیکن طاق اعداد کی صورت میں مشکل پیش آتی ہے۔

3۔ وسطانیہ کو الجبرائی انداز میں استعمال ممکن نہیں۔

4۔ وسطانیہ کو کسری اور فیصد سے ظاہر نہیں کیا جا سکتا۔

5۔ اعداد میں تبدیلی کے ساتھ وسطانیہ کی قدر میں تبدیلی واقع ہوتا ہے۔

4.7.4.4 وسطانیہ کی تحسیب(Calculation of Median)

حسابی اوسط کی طرح وسطانیہ کو دیے گئے اعداد یا مواد کی بنیاد پر محسوب کیا جاتا ہے ذیل میں دو طرح کے مواد پائے جاتے ہیں۔

(i) غیر گروہ بند مواد (ii) گروہ بند مواد

1۔ غیر گروہ بند مواد کی صورت میں وسطانیہ کی تحسیب:

دیا گیا مواد کی جدول یا گروپ کی شکل میں نہ پایا جاتا ہو تو اس کو غیر گروہ بند مواد کہتے ہیں ایسی صورت میں مواد بے ترتیب پایا جاتا ہے اعداد ترتیب میں نہیں پائے جاتے' دیے گئے اعداد کو بڑھتی ہوئی یا گھٹتی ہوئی ترتیب میں ترتیب دیا جاتا ہے ان اعداد کا درمیانی عدد وسطانیہ کہلاتا ہے اس کے لیے ذیل کا ضابطہ استعمال کیا جاتا ہے۔

Md =

جہاں پر Md=میڈین یا وسطانیہ کی علامت

N=مواد کی تعداد

I=ضابطہ کے مطابق ایک جمع کرنا

مثال(1): دیے گئے اعداد کا وسطانیہ محسوب کیجئے۔ 15, 4,18,2,9,7,5,16

حل: دیے گئے اعداد کو بڑھتی ہوئی یا گھٹتی ہوئی ترتیب میں ترتیب دیں یہاں پر بڑھتی ہوئی ترتیب میں اعداد کو ترتیب دیا گیا ہے۔

2, 4, 5, 7, 8, 9, 15, 16, 18

$$\text{Md} = \left(\frac{N+1}{2}\right) th\ Item \qquad 9 = \text{N}$$ جہاں پر

$$= \left(\frac{9+1}{2}\right)$$

$$= \left(\frac{10}{2}\right)$$

$$Md = 5^{th} \text{ Item}$$

دیے گئے اعداد کا پانچواں عدد وسطانیہ ہوگا اوپر ترتیب دیے گئے مواد میں پانچواں عدد 8 ہے اس لیے ان اعداد میں وسطانیہ 8 ہے۔

طاق اعداد کی صورت میں وسطانیہ راست طور پر محسوب کیا جاتا ہے۔

مثال(2): ذیل میں دیے گئے اعداد کا وسطانیہ محسوب کیجئے۔ 65, 70, 55, 49, 38, 32, 72, 80, 30, 35, 51, 51

حل: دئے گئے اعداد کو بڑھتی ہوئی یا گھٹتی ہوئی ترتیب میں ترتیب دیں یہاں پر بڑھتی ہوئی ترتیب میں ترتیب دیا گیا

30, 32, 35, 38, 37, 49, 51, 65, 70, 72, 80

وسطانیہ $= \left(\frac{N+1}{2}\right) th\ Item$

دیے گئے اعداد کی تعداد 12 یعنی جفت ہے

وسطانیہ $= \left(\frac{N+1}{2}\right) th\ Item$

$= \frac{13}{2}$th item

وسطانیہ = 6.5 واں عدد

وسطانیہ 6.5 واں عدد سے مراد چھٹواں اور ساتواں عدد کا اوسط وسطانیہ ہوگا۔

ترتیب دیے گئے اعداد میں چھٹواں عدد 49 اور ساتواں عدد 51 ہے اس لیے ان دونوں کا اوسط وسطانیہ ہوگا۔

وسطانیہ = 7 واں عدد + 6 واں عدد

2

$= \frac{49+51}{2}$

$= \frac{100}{2}$

Md وسطانیہ = 50

گروہ بند مواد میں وسطانیہ کی تحسیب Calculation of Median in Group of Data

ایسا مواد جو جدول یا گروپ کی شکل میں پیش کیا گیا ہو اس کو گروہ بند مواد کہتے ہیں اس کے دو اقسام ہیں۔

(a) غیر مسلسل سلسلہ (Discrete Series) (b) مسلسل سلسلہ (Continuous Series)

(a) غیر مسلسل سلسلہ میں وسطانیہ کو محسوب کرنا:

غیر مسلسل سلسلہ گروہ بند مواد کی ایک قسم ہے اس میں مواد سلسلہ وار نہیں پایا جاتا لیکن مواد کو ترتیب وار دو جدول میں پیش کیا جاتا ہو اس کو غیر مسلسل سلسلہ کہتے ہیں اس صورت میں وسطانیہ محسوب کرنے کے لیے دیے گئے جدول کے ساتھ یکجائی تعدد کا تیسرا کالم بنائیں تعدد میں اگلے تعدد کو جمع کرتے ہوئے یکجائی تعدد جمع کریں۔ ضابطہ میں قدریں درج کرتے ہوئے وسطانیہ کو جو محسوب کیا جاتا ہے۔

$$\text{وسطانیہ}=\left(\frac{N+1}{2}\right) th\ Item$$

مثال(3): ذیل میں ایک جماعت میں حاصل نشانات اور طلبا کی تعداد دی گئی ہے جب کہ مدد سے وسطانیہ محسوب کیجئے۔

نشانات(X)	15	20	25	30	35	40	45	50
تعدد(F)	4	6	8	9	3	5	4	1

حل : دیے گئے مواد کو ذیل میں ترتیب دیں۔

نشانات	تعداد(F)	یکجائی تعدد CF
15	4	4
20	6	10 (4+6)
25	8	18(10+8)
30	9	27 (18+9)
35	3	30 (27+3)
40	5	35(30+5)
45	4	39(35+4)
50	1	40(39+1)
	40	

$$\text{وسطانیہ}=\left(\frac{N+1}{2}\right) th\ Item$$

$$=\left(\frac{40+1}{2}\right)$$

$$=\left(\frac{41}{2}\right)\text{th item}$$

20.5 واں عدد = وسطانیہ

20.5 واں عدد یکجائی تعدد کے 27 میں آتا ہے اس کے مقابلے میں موجود عدد 30 وسطانیہ ہوتا ہے۔

(b) مسلسل سلسلہ میں وسطانیہ کو محسوب کرنا:

مسلسل سلسلہ مواد کو پیش کرنے ایک اہم طریقہ اس میں مواد کو مسلسل وقفہ جماعت کے ساتھ جدول میں پیش کیا جاتا ہے۔اس طریقہ میں دیے گئے سوال کے ساتھ تیسرا کالم یکجائی تعدد کا بنایا جاتا ہے۔تعدد کے مجموعہ کو 2 سے تقسیم کرنے پر حاصل عدد کو یکجائی تعدد میں نشاندہی کی جاتی ہے اس کے مقابل وقفہ جماعت میں وسطانیہ پایا جاتا ہے ذیل کے ضابطہ کی مدد سے وسطانیہ محسوب کیا جاتا ہے۔

$$\text{Md. } L_1 + \frac{\frac{N}{4} - 4}{F1} \text{xi}$$

جہاں پر:

Md: وسطانیہ کی ملامت

N: تعدد کی جملہ تعداد

L_1: وسطانیہ لائن کے وقفہ جماعت کی نچلی سطح

F_1: وسطانیہ لائن سے پہلے کی لائن میں موجود یکجائی تعدد

i: وقفہ جماعت کی لمبائی یا سعت

مثال 4: ذیل میں ایک اسٹور میں کام کرنے والے مزدور اور ان کی یومیہ اجرت کے تفصیلات دیے گئے ہیں جن کی مدد سے وسطانیہ محسوب کیجئے۔

ملازمین کی تعداد (F)	وقفہ جماعت تنخواہ (X)
2	400-500
3	500-600
5	600-700
8	700-800
4	800-900
10	900-1000
6	1000-1100
2	1100-1200

حل :دیے گئے جدول کو اس طرح تیار کیجئے۔

وقفہ جماعت تنخواہ(C.I)	ملازمین کی تعداد(F)	یکجائی تعدد(CF)
400-500	2	2
500-600	3	5(2+3)
600-700	5	10(5+5)
700-800	8	18(10+8) C
800-900	4 F1	22(18+4)
900-1000	10	32(22+10)
1000-1100	6	38(32+6)
1100-1200	2	40
	40	

وسطانیہ کی قدر کی قطار کو معلوم کرنے کے لیے

$$=\left(\frac{N+1}{2}\right) th\ Item$$

$$=\left(\frac{40}{2}\right)$$

$$= 20 \text{ th item}$$

20 کو یکجائی تعدد کی قطار میں نشاندہی کرتے ہوئے ایک لکیر کھینچ لیجیے اس قطار میں آئے وقفہ جماعت میں نچلی حد کو L_1 اور اوپری حد کو L_2 کہتے ہیں اور اس کے سامنے تعدد کو F اور اس قطار کے اوپری یکجائی تعدد کو C سے تعبیر کہتے ہیں ضابطہ میں قدریں درج کریں 20 واں عدد یکجائی تعدد 22 کی قطار میں آتا ہے اس کے سامنے وقفہ جماعت 800-900 ہے جو وسطانیہ کی لائین ہے۔

$$L_1 + \frac{\frac{N}{2}-C}{F1} Md = X\ i$$

جہاں

$18 = C$ $\quad$ $4 = F_1$ $\quad$ $20 = N/2$ $\quad$ $800 = L_1$

C=18 $\quad$ F1=4 $\quad$ N/2 =20 $\quad$ L1=800 جہاں

i: وقفہ جماعت کی لمبائی 10=

قیمتیں درج کرنے پر

$$Md = 800 + \frac{20-18}{2} \times 100$$
$$= 800 + \frac{2}{4} 100$$
$$= 800 + \frac{200}{4}$$
$$= 800 + 50$$

وسطانیہ Median = 850

وسطانیہ کو محسوب کرنے کا دوسرا ضابطہ اس طرح ہے۔

$$Md = L_1 + \frac{L^2 - L^1}{F} (m-c)$$
$$= 800 + \frac{900-800}{4} (20-18)$$
$$= 800 + \frac{100}{4} (2)$$
$$800 + \frac{200}{4}$$
$$= 800 + 50$$

وسطانیہ= 850

4.7.4.5 تقسیمی قدریں (Partition Values)

وسطانیہ دیے گئے معطیات کی وسطی قدریں ہوتی ہے جو دیے گئے اعداد کو دو مساوی حصوں میں تقسیم کرتی ہے تجزیہ کے لیے اعداد کو مزید کئی حصوں میں تقسیم کر سکتے ہیں جو اس طرح ہیں۔

1۔چوتھائیہ 2۔ عشریہ 3۔صدویہ

1۔ چوتھائیہ Quartiles

دیے گئے اعداد یا قدروں کو چار حصوں میں تقسیم کرنے پر حاصل ہر ایک حصہ کو چوتھائیہ کہتے ہیں ہر چوتھائیہ کو Q_3, Q_2, Q_1 سے ظاہر کیا جاتا ہے۔ پہلا چوتھائیہ کو زیریں چوتھائیہ کہتے ہیں یہ 25% کو ظاہر کرتا ہے Q_2 کو 50% سے ظاہر کرتے ہیں۔ تیسرا چوتھائیہ Q_3، 75% کو ظاہر کرتا ہے Q_3 کو بالائی چوتھائیہ کہتے ہیں Q_3 اور Q_1 کے درمیانی وسعت کو (Inter quartile range) بین چوتھائی سعت کہتے ہیں۔

2۔ چوتھائیوں کو محسوب کرنا Calculation Quartiles

غیر گروہ بندی غیر مسلسل اور مسلسل سلسلہ میں چوتھائیوں کو محسوب کیا جاتا ہے۔

غیر گروہ بندی میں چوتھائیہ کو محسوب کرنا:

غیر گروہ بندی میں مواد ترتیب یا جدول کی شکل میں نہیں پایا جاتا بلکہ بے ترتیب مواد پایا جاتا ہے غیر گروہ بند مواد کی صورت میں دیے گئے اعداد کو بڑھتی ہوئی یا گھٹتی ہوئی ترتیب میں ترتیب دیا جائے ذیل کے ضابطہ کی مدد چوتھائیہ محسوب کیا جاتا ہے۔

$$Q1 = \quad (\frac{N+1}{4})th\ term$$

$$Q1 = \quad 2(\frac{20+1}{4})th\ term$$

$$Q1 = \quad 3(\frac{21+1}{4})th\ term$$

مثال5: ذیل میں ایک جماعت کے 20 طلبہ کے 100 نشانات میں سے حاصل نشانات دیے گئے ہیں Q1،Q2 اور Q3 محسوب کیجئے

35, 65, 70, 90, 95, 60, 80, 82, 74, 95, 55, 61, 67, 78, 91, 40, 53, 41, 52, 63

$$Q_1 = \left(\frac{N+1}{4}\right)^{\text{th item}}$$

حل: دیے گئے اعداد کو اور ترتیب میں ترتیب دیں اور ضابطہ میں قدریں درج کریں

$$Q1 = \quad (\frac{N+1}{4})th\ term$$

$$Q1 = \quad (\frac{20+1}{4})th\ term$$

$$Q1 = \quad (\frac{21}{4})th\ term$$

Q1=5.25

35, 40, 41, 52, 53, 55, 60, 61, 63, 65, 67, 70, 74, 78, 80, 82, 90, 91, 95, 97

$$Q_1 = 5\text{th item} + \frac{1}{4}\ (6\text{th item} - 5\text{th item})$$

اوپر اعداد کی بڑھتی ہوئی ترتیب میں اعداد کو ترتیب دیا گیا ہے اس میں پانچواں عدد 53 اور چھٹواں عدد 55 ہے۔

$$=53 + \frac{1}{4} (55-53)$$

$$= 53 + \frac{1}{4} X 2$$

$$53 + \frac{2}{4}$$

$$53 + 0.5$$

$$Q_1 = 53.5$$

$$Q_2 = 2\left(\frac{20+1}{4}\right)\text{th item}$$

$$= 2\left(\frac{20+1}{4}\right)\text{th item}$$

$$= 2\left(\frac{21}{4}\right)\text{th item}$$

$$= 2X5.25$$

$$= 10.5 \text{ th item}$$

اوپر بڑھتی ہوئی ترتیب میں اعداد کو ترتیب دیا گیا ہے اس میں دسواں عدد 65 اور گیارواں عدد 67 ہے۔

$$Q_2 = 10\text{th item} + \frac{1}{4} (11\text{th item}-10\text{th item})$$

$$= 65 + \frac{1}{4} (67-65)$$

$$= 65 + \frac{1}{4} (2)$$

$$= 65 + \frac{1}{4} X_2 = 65 + \frac{2}{4}$$

$$= 65+0.5$$

$$Q_2 = 65.5$$

$$Q_3 = 3\left(\frac{N+1}{4}\right) \text{th item}$$

$$= 3\left(\frac{20+1}{4}\right) \text{th item}$$

$$= 3\left(\frac{21}{4}\right) \text{th item}$$

$$= 3\,(5.25)$$

$$Q_3 = 15.75 \text{ th item}$$

اوپر ترتیب دے گئے اعداد میں پندرہواں اور سولہواں عدد کے درمیان وسطانیہ ہوتا ہے پندرہواں عدد 80 اور سولہواں عدد 82 ہے۔

$$Q_3 = 15\text{th item} + \frac{1}{4}\,(16\text{th item} - 15\text{th item})$$

$$= 80 + + \frac{1}{4}\,(82\text{-}80) \text{ th item}$$

$$= 80 + \frac{1}{4}\,(2)$$

$$= 80 + \frac{2}{4}$$

$$= 80 + 0.5$$

$$= 80.5$$

مسلسل سلسلہ میں چوتھائیوں کی تحسیب (Calculation of Quartiles in Continuous Series)

مسلسل سلسلہ مواد کو پیش کرنے کا ایک اہم طریقہ ہے اس طریقہ میں مواد وقفہ جماعت تعدد کے ساتھ ترتیب میں پایا جاتا ہے۔

مثال 6: ذیل میں دیے گئے تفصیلات کی مدد سے وسطانیہ اور چوتھائیوں کو محسوب کیجیے۔

نشانات وقفہ جماعت (CI)	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
طلبا کی تعداد (F)	2	4	6	8	11	7	12	13	9	8

حل

نشانات	طلبا کی تعداد (F)	یکجائی تعدد (CF)
0-10	2	2
10-20	4	6 (2+4)
20-30	6	12 (2+8)
30-40	8	20 (12+8)
40-50	11	31 (20+11)
50-60	7	38 (31+7) C
60-70	12F1	50 (38+12)
70-80	13	63 (50+13)
80-90	9	72 (63+9)
90-100	8	80 (72+8)
	N = 80	

Median line N = $\left(\frac{N}{4}\right)$ th item $\left(\frac{80}{4}\right)$ = 40th item

یکجائی تعدد میں 40واں عدد شمار ہونے والے قطار کو منتخب کریں یہ وقفہ جماعت 60-70 کی قطار میں واقع ہے۔

$$Md = L_1 + \frac{\frac{N}{2} - C}{F1} Xi$$

جہاں پر $60 = L_1$، $38 = C \quad 12 = F_1 \quad 40 = N/2$

قیمتیں درج کرنے پر $= 60 + \left(\frac{40-38}{12}\right) x\ 10$

$$= 60 + \frac{2}{12} X10$$

$$= 60 + \frac{20}{12}$$

$$60 + 1.67$$

Md وسطانیہ $= 61.67$

پہلے چوتھائیہ Q_1 کی تحسیب = کل تعدد کو چار سے تقسیم کریں۔

$$\text{Quarties Line} = \left(\frac{N}{4}\right) \text{th item}$$

$$= \left(\frac{80}{4}\right) = 20 \text{ th item}$$

یکجائی تعداد میں 20عدد والے قطار کو منتخب کریں اور ذیل کے ضابطہ میں قیمتیں درج کریں یہ وقفہ جماعت 30-40 کے قطار میں واقع ہے۔

$$Q_1 = L_1 + \frac{3\left(\frac{N}{4} - C\right)}{F1} Xi$$

جہاں پر $30 = L_1$ $\quad 8 = F_1$ $\quad 12 = C$ $\quad 10 = i$

قیمتیں درج کرتے ہیں

$$= 30 + \frac{20-12}{8} \text{ X } 10$$

$$= 30 + \frac{8}{8} \text{ X } 10$$

$$= 30 + \frac{80}{8}$$

$$= 30 + 10$$

$$Q_1 = 40$$

$$Q_3 = 3 \left(\frac{N}{4}\right) \text{th item}$$

$$= 3 \text{ x } \left(\frac{80}{4}\right) \text{th item}$$

$$= 3 \text{ x } 20$$

$$Q_3 \text{ Line} = 60 \text{ thitem}$$

یکجائی تعداد میں 60عدد والے قطار کو منتخب کریں اور ذیل کے ضابطہ میں قیمتیں درج کریں یہ وقفہ جماعت 70-80 کے قطار میں واقع ہے۔

$$Q_3 = L_1 + \frac{3\left(\frac{N}{4} - C\right)}{F1} Xi$$

جہاں پر $70 = L_1$ $50 = C$ $13 = F_1$ $10 = i$

قیمتیں درج کرتے ہیں

$$= 70 + \frac{60-50}{13} X10$$

$$= 70 + \frac{10}{13} x\ 10$$

$$= 70 + \frac{100}{13}$$

$$= 70 + 7.692$$

$$Q_1 = 77.692$$

3۔ عشریے Deciles

عشریہ ایک عربی لفظ ہے دیے گئے اعداد کو 10 مساوی حصوں میں تقسیم کرنے پر حاصل ہر ایک حصہ عشرہ کہلاتا ہے۔ ہر حصہ میں تمام عدد کا 10% پر مشتمل ہوتا ہے اس کو علامت D_1 سے ظاہر کیا جاتا ہے پہلا عشریہ D_1' دوسرا عشریہ D_2 کہلاتا ہے۔

4۔ صدویہ Percentile

دیے گئے اعداد یا معطیات کو 100 مساوی حصوں میں تقسیم کرنے پر حاصل ہر ایک حصہ 1% ہوتا ہے جن کو صدویہ کہتے ہیں یہ کل تعداد کا 1% ہوتا ہے اس کو P_1 سے ظاہر کیا جاتا ہے۔ دسویں صدویہ کو P_{10}' پچاسویں صدویہ کو P_{50} وغیرہ سے ظاہر کیا جاتا ہے۔

4.7.5 بہتانیہ کے معنی اور تعریف (Meaning and Definitions of Mode)

بہتانیہ یا کثیرہ کی آسان تفہیم کے لیے ذیل کے تفریقات کا مطالعہ کیجئے۔

1۔ ایسا عدد جو سب سے زیادہ مرتبہ دہرایا جاتا ہو وہ بہتانیہ کہلاتا ہے۔

2۔ ایسا عدد جو کثرت سے واقع ہوتا ہو وہ بہتانیہ یا کثریہ کہلاتا ہے۔

3۔ دیے گئے معطیات (data) میں سب سے زیادہ مرتبہ دہرائے جانے والے عدد کو بہتانیہ یا کثریہ کہتے ہیں۔

4۔ Zizek کے مطابق بہتانیہ دیے گئے اعداد میں سب سے زیادہ مرتبہ دہرانے والا عدد ہوتا ہے اس کے اطراف اعداد پھیلے ہوئے ہوتے ہیں۔

5۔ A.M Tuttle کے مطابق دیے ہوئے اعداد کے اطراف میں سب سے زیادہ مرتبہ دہرانے والا عدد ہوتا ہے۔

6۔ ماہر شماریات کارل پیرسن(Karl Pearson)نے 1895میں اصطلاح Modeاستعمال کیا ۔لفظ Modeدراصل فرانسی لفظ "ala mode"سے ماخوذ ہے جس کے لفظی معنی "Fashimable"کے ہیں۔ala modeکے انگریزی معنیStylish

Fashonableسے مراد لیاجاتاہے کثرت سے دہرائے جانے والے رواج کو فیشن سے مراد لیاجاتاہے۔

4.7.5.1 نوعیت اور دائرہ کار(Nature and Scope)

بہتانیہ مرکزی رجحان کی پیمائش کاایک اہم آلہ ہے یہ ہر چھوٹے و بڑے پیداوار اداروں، صنعتوں، منصوبوں سرکاری وغیرہ و سرکاری منور ملات وغیرہ میں کثرت سے استعمال ہوتاہے۔دیے گئے اعداد میں سب سے زیادہ مرتبہ دہرانے والا عدد بہتانیہ یا کثیریہ کہلاتاہے۔باقی اعداداِسی عدد کے اطراف پائے جاتے ہیں۔علم شماریات ایک اہم ریاضیاتی مضمون ہے سماجی و ملکی مسائل کی نشاندہی اور ان کے حل کے لیے اس کااستعمال کیاجاتاہے۔ہندوستان میں ہر 10سال میں مردم شماری کی جاتی ہے۔عمر، جنس، تعلیم، علاقہ، شہر، دیہات، ذات، مذہب، زبان وغیرہ وغیرہ مختلف انداز میں تجزیہ کیاجاتاہے حاصل نتائج کو مسائل حل کرنے اور وسائل کا بہتر استعمال کے لیے حکمت عمل اختیار کرنے میں استعمال کرتے ہیں۔اس طرح قومی پیداوار، قومی درآمدات و برآمدات، شرح پیدائیش واموات، روزگار، بے روزگاری، چھوٹے صنعتیں، بھاری پیمانے کی صنعتیں وغیرہ کے مطالعہ کے ذریعہ معشیت کے بنیادی مسائل کو حل کر سکتے ہیں۔

کثیریہ کے مطالعہ سے مندرجہ ذیل نکات اخذ کیے جاسکتے ہیں۔

1۔ سب سے زیادہ آبادی کس عمر میں موجود ہے۔

2۔ سب سے زیادہ بے روزگاری کس عمر کے افراد یں شامل ہے۔

3۔ سب سے زیادہ تعلیم یافتہ طبقہ کہاں موجود ہے۔

4۔ سب سے زیادہ غیر تعلیم یافتہ افراد کہاں پر موجود ہیں۔

5۔ سب سے زیادہ مسائل کونسے حصہ میں حل کرنا ضروری ہے۔

6۔ سب سے زیادہ آبادی کی گنجائیت کہاں پائی جاتی ہے۔

4.7.5.2 خصوصیات(Features)

بہتانیہ کی اہم خصوصیات کو ذیل میں بتلایا گیاہے۔

1۔ بہتانیہ سب سے زیادہ مرتبہ دہرائے جانے والا عدد ہوتاہے یہ مرکزی رجحان کے اہم پیمانوں میں سے ایک ہے۔

2۔ معطیات یااعداد میں تبدیلی کے ساتھ بہتانیہ میں بھی تبدیلی واقع ہوتی ہے۔

3۔ معمولی مواد کی صورت میں بھی بہتانیہ کو محسوب کیاجاسکتاہے۔

4۔ بہتانیہ دیے گئے اعداد و شمار کی ترجمانی کرتاہے۔

4.7.5.3 فوائد وخوبیاں(Advantages or Merits)

بہتانیہ کے فوائد کو ذیل میں ظاہر کیا گیا:

1۔ بہتانیہ ایک آسان اور قابل فہم اصطلاح ہے جس کو ہر نارمل و خاص فرد آسان سے سمجھ سکتا اور اس کی تحسیب کافی آسان ہے۔

2۔ بہتانیہ کا معیشت کے ہر شعبے میں استعمال کیا جاتا ہے۔

3۔ تجارتی و صنعتی میدان میں بہتانیہ کو کثرت سے استعمال ہوتا ہے۔

4۔ صنعتیں، چھوٹے و بڑے پیداواری ادارے منصوبہ بندی میں بہتانیہ کا کثرت سے استعمال کرتے ہیں اور اسی کی بنیاد پر حکمت عملی طے کرتے ہیں۔

5۔ ہر چھوٹے و بڑے پیداواری اداروں سرکاروں وغیرہ سرکاری تنظیموں کو کامیابی سے ہمکنار کرنے کے لیے بہتانیہ کا استعمال کرتے ہیں۔

6۔ بہتانیہ معیاری مواد(Qualitative Data)میں کافی استعمال ہوتا ہے۔

7۔ بہتانیہ کوOpen-endedجدول میں بھی آسانی سے اخذ کیا جاسکتا ہے۔

8۔ بہتانیہ کو ترسیمی شکل میں آسانی سے ظاہر کر سکتے ہیں۔

9۔ بہتانیہ معطیات کا بہتر نمائندگی کرتا ہے۔

10۔ بہتانیہExtreme Valueسے متاثر نہیں ہوتاExtreme Valueکی عدم موجودگی ہیں بھی بہتانیہ کو محسوب کیا جاتا ہے۔

11۔ صارفین کی اکثریت کی خواہش کو معلوم کرنے کے لیے بہتانیہ کا استعمال کیا جاتا ہے۔

خامیاں Disadvantages

بہتانیہ کے خامیوں کو ذیل میں بتلا گیا ہے۔

1۔ بہتانیہ کی تحسیب آسان ہونے کے باوجود بعض اوقات تحسیب یا مشکلات درپیش ہوتے ہیں بہتانیہ کی واضح انداز میں جامع تعریف نہیں ہے۔

2۔ الجبرائی منورملات میں بہتانیہ کی تحسیب ممکن نہیں۔

3۔ مواد کی تبدیلی کے ساتھ ساتھ بہتانیہ میں بھی تبدیلی واقع ہوتی ہے اس طرح بہتانیہ کو مستقل تصور کرنا ممکن نہیں۔

4۔ بہتانیہ الجبرائی طرز میں ظاہر نہیں کر سکتے۔

5۔ Bimodelاورmultimodelمیں مواد کی صورت میں بہتانیہ کی تحسیب ممکن نہیں۔

6۔ نمونوں یا مواد کی تبدیلی کے ساتھ بہتانیہ میں بھی تبدیلی واقع ہوتی ہے۔

7۔ بہتانیہ کوئی مستقل عدد نہیں ہے جو مسائل کو حل کرنے میں سازگار ہو۔

8۔ بہتانیہ دیے گئے مواد یا معطیات کی نمائندگی نہیں کرتا۔

9۔ بہتانیہ کو محسوب کرنے کے مختلف ضوابطہ ہیں ضوابطہ میں تبدیلی سے نتائج میں بھی تبدیلی واقع ہوتی ہے۔

10۔ بہتانیہ کی واضح تعریف موجود نہیں ہے۔

4.7.5.4 استعمالات یا اطلاق (Uses or Application)

بہتانیہ کے اہم استعمالات کو ذیل میں بتلایا گیا ہے۔

1۔ تعلیمی میدان 2۔ صنعتی میدان

3۔ منصوبہ بندی 4۔ تجارتی میدان

1۔ تعلیمی میدان

تعلیمی میدان میں بہتانیہ کا جامع و وسیع معنوں میں استعمال کیا جاتا ہے تعلیمی پالیسی بھی بہتانیہ کی بنیاد پر قائم کی جاتی ہے۔ سماجی ضرورتوں کو مدنظر رکھتے ہوئے کثرت کے ساتھ طلب شدہ مضامین کو نصاب میں شامل کرنے میں بہتانیہ کا سہارا لیا جاتا ہے۔ تعلیم بالغان، تعلیم اطفال، اسکولی تعلیم، روزگار پر مبنی تعلیم، تکنیکل تعلیم، سائنسی تعلیم وغیرہ کی فراہمی مختلف منورملات میں بہتانیہ کا استعمال کیا جاتا ہے۔

2۔ صنعتی میدان

صنعتی میدان میں بہتانیہ کا کثرت سے استعمال کیا جاتا ہے ہر چھوٹے و بڑے پیداوری و صنعتی اداروں میں اس کا استعمال کیا جاتا ہے ریڈی میڈ ملبوسات تیار کرنے والی صنعتیں، جوتے چپل کی صنعتیں ضروریات زندگی کے سامان (ظروف) تیار کرنے والی صنعتیں وغیرہ۔ بازار میں سب سے زیادہ مقدار میں جس سائز یا ڈیزائن کی طلب ہو اسی کی سائز میں چپل، جوتے ملبوسات وغیرہ تیار کیے جاتے ہیں۔ بہتانیہ صنعتی پیداوار یا مقدار اور سائز کے تعین میں مدد ملتی ہے۔ جس سے صنعتی وسائل یا پیداواروں کے ضائع ہونے سے محفوظ رہتے ہیں۔

3۔ منصوبہ بندی

منصوبے مستقبل کے لیے مدون کیے جاتے ہیں۔ منصوبوں کی تدوین کے دوران بہتانیہ کا کثرت سے استعمال کیا جاتا ہے۔ صنعتی و پیداواری ادارے اپنے پیداواری کا تعین کرنا، پیداواری کی تقسیم، صارفین تک رسائی کی حکمت عملی وغیرہ میں کثیریہ کا استعمال کیا جاتا ہے۔ اسی کے سبب کاروبار کو فروغ اور استحکام حاصل ہوتا ہے۔ کثیریہ کو نظرانداز کرنے کی صورت میں اس کے فوائد سے محروم ہونے کا خدشہ بڑھتا ہے۔ اور ایک کامیابی تجارتی حکمت عملی کو اختیار کرنا مشکل اور دشوار ہوتا ہے۔

4۔ تجارتی میدان

ہر چھوٹے و بڑے تاجرین کثیریہ کو مدنظر رکھتے ہوئے تجارتی حکمت عملی اختیار کرتے ہیں۔ ذخیرہ اندازی، قیمت سازی، صارفین سے تعلقات، مال کی نکاسی، تجارتی اوقات کا تعین وغیرہ وغیرہ منورملات میں کثیریہ کا استعمال کرتے ہیں۔

4.7.5.5 بہتانیہ کی تحسیب(Computation of Mode)

مواد کی فراہمی کی بنیاد پر بہتانیہ کو مختلف طریقوں سے محسوب کیا جاتا ہے آپ جانتے ہیں کہ مواد یا معطیات غیر گروہ بند(Un grouped) گروہ بند(Grouped)کی شکل میں ہوتا ہے۔ذیل میں بہتانیہ کی تحسیب کے طریقوں کی وضاحت کی گئی ہے۔

غیر گروہ بند مواد Ungrouped Data

ایسا مواد بے ترتیب یا گروہ کی شکل میں نہ پایا جاتا ہواس کو غیر گروہ بند مواد کہتے ہیں۔اس میں مواد بے ترتیب پھیلا ہوا ہوتا ہے۔یعنی غیر گروہ بند مواد ایک ایسا مواد ہے جس میں معطیات(Data)بے ترتیب ہوتے ہیں یعنی گروہ یا جماعت یا ترتیب سے نہیں پایا جاتا ہے۔معطیات کی تعداد کم ہونے پر مشاہدہ کے ذریعہ آسانی سے محسوب کیا جاسکتا ہے۔

ذیل میں جماعت پنجم کے 10 طلبا کی عمریں دی گئی ہیں۔

عمریں سال میں:10, 10, 11, 10, 11, 12, 10, 12, 11, 10

مندرجہ بالا مثال میں ظاہر ہوتا ہے کہ جماعت پنجم میں 10 سال کی عمر والے طلبا کی تعداد زیادہ ہے اس طرح بہتانیہ 10 (mode)سال ہوگا۔

انفرادی سلسلہ(Calculation of mode in individual Series)میں بہتانیہ کی تحسیب

ایسے معطیات جو بے ترتیب یا گروہ بندی میں نہ ہو اعداد یا معطیات انفرادی طور پر دئے گئے ہوں انہیں انفرادی سلسلہ کہتے ہیں۔

مثلاً: 2, 8, 2, 6, 9, 5, 4, 7, 3, 2, 6, 8اس مثال میں اعداد کی کوئی ترتیب نہیں ہے بلکہ یہ بے ترتیب پھیلے ہوئے ہیں۔

انفرادی سلسلہ میں بہتانیہ کی تحسیب کے مراحل:

انفرادی سلسلہ میں بہتانیہ(mode)کو محسوب کرنے کے لیے ذیل کے مراحل پر عمل کریں۔

1۔ دیے گئے اعداد کو صعودی(بڑھتی ہوئی) یا نزولی(گھٹتی ہوئی) ترتیب میں ترتیب دیں۔

2۔ دیے گئے اعداد میں سب سے زیادہ مرتبہ دہرائے جانے والا عدد بہتانیہ کہلاتا ہے اس کی نشاندہی کیجئے۔

3۔ بہتانیہ کو علامت 'Z' سے ظاہر کیا جاتا ہے۔

نوٹ: دیے گئے اعداد میں کوئی بھی عدد سب سے زیادہ مرتبہ نہ دہرانے پر بہتانیہ کو محسوب نہیں کیا جاسکتا ہے۔

مثال: ذیل میں دیے گئے اعداد کا بہتانیہ محسوب کیجئے۔

17, 15, 10, 8, 12, 15, 21, 24, 15, 8, 6, 9, 15, 18, 21

حل: دیے گئے اعداد یا معطیات کو صعودی(بڑھتی ہوئی) ترتیب میں ترتیب دیں(چھوٹے عدد سے بڑے عدد کی جانب رخ ہونا)

Z=6, 8, 8, 9, 10, 12, 15, 15, 15,15, 17, 18, 21, 21, 24

مندرجہ بالا اعداد میں عدد 15 چار مرتبہ دہرایا گیا ہے اس لیے بہتانیہ 15 ہوگا۔یعنی Z=15

نوٹ: صعودی (بڑھتی) ہوئی ترتیب چھوٹے عدد سے شروع کر کے ترتیب وار بڑے اعداد ترتیب دیا جاتا ہے اور آخر میں سب سے بڑا عدد حاصل ہوتا ہے۔ نزولی ترتیب میں بڑے عدد سے شروع کیا جاتا ہے اور آخر میں سب سے چھوٹے عدد حاصل ہوتا ہے۔ مشاہدے کے ذریعہ سب سے زیادہ مرتبہ دہرائے جانے والا عدد بہتانیہ کو اخذ کر سکتے ہیں۔

بعض اوقات بہتانیہ کی تحسیب ممکن نہ ہونے کی صورت میں ذیل کے ضابطہ کی مدد سے بہتانیہ کو محسوب کیا جا سکتا ہے۔

Mode= 3 Median- 2 Mean

تعددی یا مجرد سلسلہ میں بہتانیہ کی تحسیب Calculation of Mode in Discrete series

دیے گئے معطیات ترتیب وار دو قطار میں ترتیب ہوں اس کو Discrete Series کہتے ہیں۔ پہلے قطار میں اعداد اور دوسرے قطار میں ان کی تعداد کو بتلایا جاتا ہے ذیل میں Discrete Series کی مثال دیا گیا ہے۔

اعداد	9	11	15	19	22	23	28	29	32
تعدد	2	6	12	17	8	9	7	5	4

مندرجہ بالا دو قطار میں سے ہیں اعداد اور دوسرے قطاریں ان کی تعداد کو بتلایا گیا ہے۔

مثال 2: ذیل میں دیے گئے مواد کی مدد سے بہتانیہ محسوب کیجئے۔

عمر (X)	12	13	18	21	25	30	35	40	42	45
تعدد (F)	3	5	6	9	10	15	11	8	6	6

حل: سب سے زیادہ دو مرتبہ دہرائے جانے والے مواد کو بہتانیہ کہتے ہیں مشاہدہ کی مدد سے 30 عدد 15 مرتبہ دہرایا ہے اس لیے موڈ (mode) 30 عدد ہو گا۔ حسابی اعتبار سے اس کو حسب ذیل طریقوں سے محسوب کیا جاتا ہے۔

گروپ بندی جدول (Grouping Table)

عمر(x)	تعدد(F)	II	III	IV	V	VI
12	3	-	-	-	-	-
13	5	8(3+5)	-	14(3+5+6)	-	-
18	6	-	11 (5+6)	-	20 (5+6+9)	-
21	-	15	-	-	-	-

25	10	-	19 (9+10)	34 (9+19+15)	-	25 (6+9+10)
30	15	25	-	-	36 (10+15+11)	-
35	11	-	26 (15+11)	-	-	34 (15+11+8)
40	8	19	-	25	-	-
42	6	-	14 (8+6)	-	20(8+6+6)	-
45	6	12	-	-	-	-

تجزیاتی جدول(Analysis Table)

Colum	1	2	3	4	5	6	Total
12							
13							
18							
21				✔			1
25		✔		✔	✔		3
30	✔	✔	✔	✔	✔	✔	6
35			✔		✔	✔	3
40						✔	1
42							
45							

مندرجہ بالا تجزیاتی جدول میں 30 عدد 6 مرتبہ دہرایا ہے اس لیے بہتانیہ 30 ہوگا۔

بہتانیہ کی تحسیب کے لیے ذیل کے مراحل پر عمل کریں۔

1۔ دیے گئے جدول میں مزید 5 کالم قطار ترتیب دیں۔

2۔ قطار (کالم) ایک تعدد کو ظاہر کریگا۔

3۔ قطار II میں پہلے دو تعدد کے جوڑ کا حاصل جمع کو بتلائیں۔ مثلاً (3+5=8, 6+9=15, 10+15=25, 8+6=14)

4۔ قطار III میں پہلے تعدد کو چھوڑ کر دو تعدد کے جوڑ کے حاصل جمع کو بتلائیں۔

(5+6=11, 9+10=19, 15+11=26, 8+6=14)

(14=6+8,26=11+15,19=10+9,11=6+5)

5۔ چوتھے قطار میں پہلے تین تعدد کے حاصل جمع کو بتلائیں۔ (3+5+6=14, 9+10+15=34, 14=6+3 11+8+6=25)

6۔ پانچویں قطار میں پہلے تعدد کو چھوڑ کر تین تعدد کے حاصل جمع کو بتلائیں۔

(5+6+9=20, 10+15+11=36, 8+6+6=20)

7۔ چھٹویں قطار میں تعدد کے پہلے دو اعداد کو چھوڑ کر تین تعدد کے حاصل جمع کو بتلائیں۔

(6+9+10=25, 15+11+8=34)

8۔ تجزیاتی جدول تیار کریں اس میں 8 قطار ہونگے۔

9۔ گروپ بندی جدول میں ہر کالم میں سب سے بڑے عدد کی نشاندہی کرتے ہوئے تجزیاتی جدول ہیں جن اعداد کو جمع کرنے سے بڑا اعدد حاصل ہوا ہے ان کے سامنے (✔) صحیح کا نشان لگائیں۔

10۔ سب سے زیادہ مرتبہ دہرائے ہوا عدد بہتانیہ ہوگا۔

مسلسل سلسلہ میں بہتانیہ کی تحسیب Calculation of mode in Continuous Series

معطیات یا اعداد کو وقفہ جماعت کے ساتھ جدول میں پیش کیا جائے۔ وہ Continuous Series کہلاتا ہے اس جدول میں سوال کی موزونیت کے اعتبار سے وقفہ جماعت اور تعدد (F) کی تعداد دی جاتی ہے۔ ذیل کے ضابطہ کی عدد سے بہتانیہ محسوب کیا جاتا ہے۔

$$Mo = L_1 + \frac{F_1 - F_0}{2F_1 - F_0 - F_2} X Cl$$

جہاں پر Mo= بہتانیہ کی علامت

L_1= وقفہ جماعت کی نچلی سطح

F_1= بہتانیہ کے وقفہ جماعت کی تعدد

F_0= بہتانیہ کے وقفہ جماعت سے پہلے کی تعدد

F_2= بہتانیہ کے وقفہ جماعت کے بعد کا تعدد

$C_1=$ وقفہ جماعت (Class Interval)

مثال 3: ذیل میں ایک صنعت میں کام کرنے والے مزدوری کی یومیہ اجرت دی گئی ہے۔ جن سے بہتانیہ محسوب کیجئے۔

وقفہ جماعت روپے	500-600	600-700	700-800	800-900	900-1000	1000-1100
مزدوروں کی تعداد (F)	4	7	9	5	4	1

حل:

وقفہ جماعت روپے	مزدوروں کی تعدد
500-600	4
600-700	7 F_0
700-800	9 F_1
800-900	5 F_2
900-1000	4
1000-1100	1
	30

1۔ مشاہدہ کے ذریعہ اخذ ہوتا ہے کہ وقفہ جماعت 700-800 میں سب سے زیادہ مزدوروں کی تعدد ہے تب بہتانیہ بھی اسی وقفہ جماعت میں ہوگا۔

2۔ علامتوں کے اعتبار سے اعداد کو تعین کریں۔5=L1=700, 9=F1, 7=F0, F2ضابطوں میں قدروں کو درج کریں۔

$$= 30 + \frac{17-7}{2(17)-7-8} X10$$

$$= 30 + \frac{10}{34-15} X10$$

$$= 30 + \frac{100}{19}$$

$$= 30 + 4.2$$

4.2= 3

آپ نے حسابی یا ضابطہ کی مدد سے بہتانیہ کے تحسیب کے طریقہ سے واقف ہو چکے ہیں بہتانیہ کو ترسیم کی مدد سے بھی ظاہر کیا جاتا ہے۔ ترسیم کے ذریعہ بہتانیہ کو اخز کرنے کے لیے ذیل کے طریقہ کار کو اختیار کریں۔

1۔ ترسیمی کاغذ پر X اور Y محسور کے خطوط ڈالیں

2۔ افقی خط X محور پر وقفہ جماعت اور محودی خط Y محور پر تعدد (F) کو بتلائیں۔

3۔ تعدد (Frequency) کی مدد سے ہسٹو گرام تیار کریں۔

4۔ سب سے زیادہ تعدد والے ہٹو گرام سے پہلے اور بعد کے متصل کونوں سے خطوط کھینچے (مثال ترسیمی کاغذ پر دیکھے)۔

5۔ کھینچے ہوئے خطوط سے حاصل نقطہ تقاطع بہتانیہ ہو گا۔

6۔ نقطہ تقاطع سے X محور تک لکیر کھینچے جو بہتانیہ پر قائم ہو گا۔

7۔ مثال کی مدد سے بتلائے گئے ترمیمی اظہار کا مشاہدہ کیجئے۔

ذیل کے مثال میں ترسیم کی مدد سے بہتانیہ کی تحسیب کو بتلایا گیا:

مثال 4: ذیل میں ایک مضمون میں حاصل کردہ نشانات اور طلبا کی تعداد دی گئی جن کی مدد سے ترسیم پر بہتانیہ تحسیب کیجیے۔

70-80	60-70	50-60	40-50	30-40	20-30	10-20	0-10	نشانات وقفہ جماعت
2	6	5	8	17	7	3	2	طلبا کی تعداد

حل: ذیل میں ترسیمی کاغذ پر X محور پر نشانات اور Y محور پر طلبا کی تعداد کو بتلایا گیا ہے۔

حسابی اعتبار سے بہتانیہ کی تحسیب

$$Mo = L1 + \frac{F1 - Fo}{2F1 - Fo - F2} XCI$$

10=CI ' 8=F2 ' 7=Fo ' 17=F1 ' 30=L1

قیمتیں درج کرنے پر

$$= 30 + \frac{17 - 7}{2(17) - 7 - 8} X10$$

$$= 30 + \frac{10}{34 - 15} X10$$

$$= 30 + \frac{100}{19}$$

$$= 30 + 4.2$$

$$Mo = 34.2$$

ترسیمی کاغذ پر بہتانیہ وقفہ جماعت 30-40 کے درمیان واقع ہے ترسیمی کاغذ پر ٹھیک طور پر اعداد کو ظاہر کرنے سے بہتانیہ کی صحیح قدر اخز ہوتی ہے ضابطہ کی مدد سے اخز کردہ بہتانیہ کی قدر 34.2 ہے جو وقفہ جماعت 30-40 کے درمیانی واقع ہے۔

بہتانیہ (کثیریہ) کا ترسیمی تعین Graphical Determination

آپ نے حسابی یا ضابطہ کی مدد سے بہتانیہ کے تحسیب کے طریقہ سے واقف ہو چکے ہیں۔ مگر بہتانیہ کو ترسیم کی مدد سے بھی ظاہر کیا جاتا ہے۔ ترسیم کے ذریعہ بہتانیہ کو اخذ کرنے کے لیے ذیل کے طریقہ کار کو اختیار کریں۔

1۔ ترسیمی کاغذ پر x اور y محسور کے خطوط ڈالیں۔

2۔ افقی خط x محور پر وقفہ جماعت اور عمودی خط y محور پر تعدد (F) کو بتلائیں۔

3۔ تعدد (Frequency) کی مدد سے ہسٹو گرام تیار کریں۔

4۔ سب سے زیادہ تعدد والے ہسٹو گرام سے پہلے اور بعد کے مستقل کونوں سے خطوط کھینچیے۔ (مثال ترسیمی کاغذ پر دیکھیے)۔

5۔ کھینچے ہوئے خطوط سے حاصل نقطہ تقاطع بہتانیہ ہوگا۔

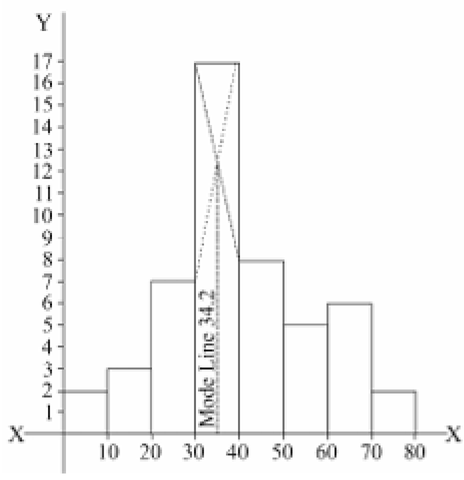
6۔ نقطہ تقاطع سے x محسور تک لکیر کھینچے جو بہتانیہ پر قائم ہوگا۔

7۔ مثال کی مدد سے بتلائے گیے ترسیمی اظہار کا مشاہدہ کیجیے۔

ذیل کے مثال میں ترسیم کی مدد سے بہتانیہ کی تحسیب کو بتلایا گیا ہے۔

مثال 5: ذیل میں ایک مضمون میں حاصل کردہ نشانات اور طلبا کی تعداد دی گئی ہے جن کی مدد سے ترسیم پر بہتانیہ محسوب کیجیے۔

نشانات وقفہ جماعت	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
طلباء کی تعداد	2	3	7	17	8	5	6	2



حل: ذیل میں ترسیمی کاغذ تعدد کرتے ہوئے سفید کاغذ پر نشانات اور طلباء کی تعداد کو بتلایا گیا ہے۔

حسابی اعتبار سے بہتانیہ کی تحسیب

$$M_o = L_1 + \frac{F_1 - F_o}{2F_1 - F_o - F_2} \times CI$$

$$10 = CI \ ' \ 8 = F_2 \ ' \ 7 = F_o \ ' \ 17 = F_1 \ ' \ 30 = L_1$$

قیمتیں درج کرنے پر

$$= 30 + \frac{17 - 7}{2(17) - 7 - 8} \times 10$$

$$= 30 + \frac{10}{34 - 15} \times 10$$

$$= 30 + \frac{100}{19} \times 10$$

$$= 30 + 4.2$$

بہتانیہ $M_O = 34.2$

ترسیمی کاغذ پر بہتانیہ وقفہ جماعت 30-40 کے درمیان واقع ترسیمی کاغذ پر ٹھیک طور پر اعداد کو ظاہر کرنے سے بہتانیہ کی صحیح قدر اخذ ہوتی ہے۔ ضابطہ کی مدد سے اخذ کردہ بہتانیہ کی قدر 34.2 ہے جو وقفہ جماعت 30-40 کے درمیان واقع ہے۔

4.8 اکتسابی نتائج (Learning Outcome)

طلباء نے شماریات کے تصور، معنی اور تعریف کے ساتھ ساتھ کاروبار اور تحقیق میں اس کی گنجائش اور اہمیت کے بارے میں ٹھوس سمجھ حاصل کی ہے۔ انہوں نے سیکھا ہے کہ کس طرح اعداد و شمار فیصلہ سازی میں معاونت کے لیے ڈیٹا اکٹھا کرنے، ترتیب دینے، خلاصہ کرنے اور تشریح کرنے کے لیے ایک طاقتور ٹول کے طور پر کام کرتے ہیں۔ طلباء اب اعداد و شمار کے مختلف شماریاتی ٹولز کو مؤثر طریقے سے شناخت اور استعمال کر سکتے ہیں۔ مزید برآں، انہوں نے مرکزی رجحان کے اقدامات کے بارے میں علم تیار کیا ہے، بشمول ان کے معنی، تعریفیں، اور اطلاقات، جو انہیں بڑے ڈیٹا سیٹوں کا خلاصہ کرنے اور عام اقدار کی شناخت کرنے کے قابل بناتے ہیں۔ مجموعی طور پر، سیکھنے والے اعداد و شمار کو منظم طریقے سے تجزیہ کرنے اور باخبر کاروباری فیصلے کرنے کے لیے شماریاتی تصورات کو لاگو کرنے کے لیے لیس ہیں۔

4.8 نمونہ امتحانی سوالات (Model Examination Questions)

خالی جگہوں کو پر کیجیے۔

1۔ لفظ Mode.........زبان سے ماخوذ ہے۔

2۔ بہتانیہ کی تعریف...........ہے۔

3۔ مسلسل سلسلہ میں بہتانیہ کو محسوب کرنے کا ضابطہ...........ہے۔

4۔ بہتانیہ کو........علامت سے ظاہر کیا جاتا ہے۔

5۔ 700-600 کا وسطی نقطہ.........ہے۔

6۔ طاق اعداد کی مثالیں.............ہیں۔

7۔ جفت اعدا کی مثالیں..............ہیں۔

8۔ Q_1 کو محسوب کرنے کا ضابطہ..........ہے۔

9۔ مسلسل سلسلہ میں وسطانیہ کو محسوب کرنے کا ضابطہ...........ہے۔

10۔ 12 ’5 ’15 ’8 ’2 ’7 کی بڑھتی ہوئی ترتیب...........ہے۔

مختصر جوابات کے حامل سوالات

1۔ ذیل کے معطیات کی عدد سے بہتانیہ محسوب کیجئے۔

20, 18, 22, 16, 15, 20, 26, 22, 23, 22, 15, 18, 22

2۔ ذیل کے مواد کی عدد سے بہتانیہ محسوب کیجئے۔

20	19	18	14	12	9	تعدد
10	6	2	3	4	3	F

طویل جوابات کے حامل سوالت

6۔ بہتانیہ کی وضاحت کرتے ہوئے اس کے فوائد اور خامیوں کو بیان کیجئے۔

7۔ بہتانیہ کے استعمالات پر تفصیلی نوٹ لکھئے۔

8۔ ذیل کے مواد کی مدد سے بہتانیہ محسوب کیجئے۔

60-70	50-60	40-50	30-40	20-30	10-20	0-10	وقفہ جماعت(I)
4	4	8	5	4	2	3	تعدد(F)

9۔ ذیل میں ایک جماعت کے مضمون شماریات میں حاصل کردہ نشانات دیے گئے ہیں۔ گروپ جدول کی مدد سے بہتانیہ محسوب کیجئے۔

وقفہ جماعت نشانات	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
طلبا کی تعداد	3	1	5	7	8	6	7	8	3	2

1۔ وسطانیہ سے کیا مراد ہے۔

2۔ وسطانیہ کے فوائد اور خامیوں کو بیان کیجئے۔

3۔ ذیل کے اعداد کے مدد سے وسطانیہ محسوب کیجئے 3’ 8’ 12’ 9’ 7’ 15’ 18’ 22

4۔ وسطانیہ کے خصوصیات بیان کیجئے۔

5۔ زیریں چوتھائیہ اور بالائی چوتھائیہ سے کیا مراد ہے۔

6۔ ذیل کے تفصیلات کی عدد سے وسطانیہ محسوب کیجئے۔

نشانات X	35	45	55	65	75	85	95
تعوز F	4	3	9	5	4	2	3

7۔ ذیل کے تفصیلات کی عدد سے وسطانیہ محسوب کیجئے۔

نشانات X	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
طلباء کی تعداد F	3	2	7	9	11	13	6	5	3	1

8۔ ذیل میں ایک صنعت میں کام کرنے والے مزدوروں کی یومیہ اجرت اور تعدادی گرامس کی مدد سے وسطانیہ محسوب کیجئے۔

اجرت رہے X	300-400	400-500	500-600	600-700	700-800	800-900	900-1000
مزدوروں کی تعداد F	3	4	8	6	4	3	2

9۔ ذیل میں دیئے گئے تفصیلات کی مدد سے Q_1, Q_2 اور Q_3 محسوب کیجئے۔

وقفہ جماعت نشانات	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
تعداد F	1	3	5	7	11	9	2	2

10۔ وسطانیہ محسوب کرنے کے مختلف طریقوں میں وضاحت کیجئے۔

اکائی 5۔ انتشار کی تحسیب

(Measurement of Dispersion)

5.0	تمہید	Introduction
5.1	مقاصد	Objectives
5.2	معنی اور تعریف	Meaning and Definition
5.3	انتشار کی اہمیت	Importance of Dispersion
5.4	اچھے انتشار کے اہم خصوصیات	Features of Good Dispersion
5.5	انتشار کی نوعیت	Nature of Dispersion
5.6	معیاری انحراف کے معنی اور تعریف	Meaning and Definition of Standard Deviation
5.7	اکتسابی نتائج	
5.8	نمونہ امتحانی سوالات	Model Examination Questions

5.0 تمہید (Introduction)

ہر تاجر صنعت کار یا پیدا کنندے کاروبار کے موقف کو اخذ کرنے کے لیے ریاضی و شماریات کے مختلف پیمانوں کا استعمال کرتے ہیں۔ نارمل منافع، خالص، منافع، اخراجات کا تناسب، اوسط، وسطانیہ اور بہتانیہ وغیرہ چند اہم مقبول طریقے ہیں جو کاروبار کے ساتھ ساتھ کھیل کود تعلیمی اداروں میں استعمال کرتے ہیں۔ تعلیمی کارکردگی کو ظاہر کرنے کے لیے اوسط کا زیادہ استعمال کیا جاتا ہے۔ کرکٹ کے مقابلوں کے دوران میدان میں آنے والے کھلاڑیوں کا تعارف پیش کیا جاتا ہے۔ کھلاڑی کے اعظم ترین اسکور، رن، وکٹ اور چھکے کے ساتھ ساتھ اقل ترین مظاہرہ کا بھی ذکر کیا جاتا ہے۔ علم ریاضی میں قدرت کا حاصل طالب علم ریاضی میں زیادہ نشانات حاصل کر سکتا ہے۔ لیکن کسی دوسرے مضمون میں اس کے برعکس نشانات حاصل ہو سکتے ہیں روشنی یا شعاع لا متناہی حد تک پھیل جاتی ہے۔ شعاعوں کا دور تک پھیل جانا یا منتشر ہونا ایک بنیادی صفت ہے۔ اعداد یا معطیات کے درمیان شدید فرق انتشار کہلاتا ہے۔ اس اکائی میں آپ انتشار کے متعلق معلومات حاصل کریں گے۔

5.1 مقاصد (Objectives)

اس اکائی کے مطالعہ کے بعد آپ اس قابل ہو جائیں گے کہ

- انتشار کے معنی و مفہوم کو بیان کر سکیں گے۔
- انتشار کی اہمیت کی وضاحت کر سکیں گے۔
- انتشار کے استعمالات کی وضاحت کر سکیں گے۔
- انتشار کو اخذ کرنے کے طریقوں کی وضاحت کر سکیں گے۔

5.2 معنی اور تعریف (Meaning and Definition)

انتشار کے معنی و مفہوم کی تشریح کے لیے ذیل کے تعریفات کا مطالعہ کیجئے۔

1۔ دیے ہوئے معطیات کا پھیلاؤ انتشار کہلاتا ہے۔

2۔ دیے ہوئے معطیات یا مواد مرکزی رجحان کے مساوی نہ ہو انتشار کہلاتا ہے۔

3۔ معطیات یا مواد کا مرکزی رجحان سے مختلف ہونا انتشار کہلاتا ہے۔

4۔ حسابی اوسط کے انحراف انتشار کہلاتا ہے۔

5۔ دیے ہوئے اعداد یا معطیات میں تسلسل یا یکسانیت کا نہ ہونا انتشار کہلاتا ہے۔

Simpson and Kafka کے مطابق اوسط (Average) بذات خود تمام مواد کے متعلق معلومات فراہم نہیں کرتا یا مواد کی خصوصیات کو ظاہر نہیں کرتا تب تک اس کے اطراف پھیلے ہوئے اعداد کا مطالعہ نہ کریں یعنی اوسط سے اسی وقت موزوں مطلب اخذ کر سکتے ہیں جب کہ اس کے اطراف کے مواد کے متعلق معلومات حاصل کریں۔

علم شماریات میں انتشار کو Variability یا Scatler یا Speed بھی کہا جاتا ہے۔ مواد کا پھیلاؤ، انتشار کو ظاہر کرتا ہے۔ وسیع یا بڑے علاقے پر محیط یا پھیلاؤ کو ظاہر کرتا ہے۔ علم شماریات میں یہ مواد کا پھیلاؤ کو انتشار سے مراد لیا جاتا ہے۔ نقطہ Dispersion یونانی (Greek) لفظ 'Disapora' سے ماخوذ ہے۔ یہ لاطینی لفظ Dispersine سے ماخوذ ہے۔ جس کے معنی پھیلاؤ یا دور تک پھیلنا سے ہے۔ انتشار علم ریاضی کے ساتھ ساتھ دیگر علوم میں اکثر استعمال ہوتا ہے مواد یا دیے ہوئے اعداد کا پھیلاؤ انتشار کہلاتا ہے۔ انتشار کو ہم اپنی روز مرہ زندگی میں استعمال کرتے رہتے ہیں ایک نارمل مثال کے ذریعہ انتشار کے مفہوم کی وضاحت ہوتی ہے۔ ایک کاغذ کے چند چھوٹے چھوٹے ٹکڑوں کو فرش پر پھینک دیں جس جگہ یا جہاں تک کاغذ کے ٹکڑے پھیلے ہو وہاں تک انتشار کہلائے گا۔ کھیل کود کے دوران پانی کے پائپ سے پچکاری کے پھیلاؤ پر غور کیجئے۔ خوشی کے موقع پر پھولوں کو ہوا میں پھیلاتے ہیں۔ برقی بلب سے نکلنے والی شعاعوں پر غور کیجئے۔ روشنی کے پھیلاؤ کا علاقہ اس کا انتشار کہلاتا ہے۔

5.3 انتشار کی اہمیت(Importance of Dispersion)

انتشار علم شماریات کا ایک اہم حصہ ہے۔ جس کو مختلف مقاصد کے تحت استعمال کیا جاتا ہے۔ شعبہ تجارت، درآمدات و برآمدات، پیداور معیار، اہداف کا تعین وغیرہ مختلف میدانوں میں اس کا استعمال کیا جا سکتا ہے۔ ذیل میں انتشار کی اہمیت کو بیان کیا گیا ہے۔

1۔ تغیرات کے اسباب کا پتہ لگانا: انتشار کی مدد سے مواد یا معطیات پر مختلف عوامل کے اثرات کا آسانی سے پتہ چلایا جا سکتا ہے۔

2۔ نمونہ بندی کی صحت کی جانچ: مختلف تحقیقی میدان میں آبادی سے حاصل نمونوں کی صحت مندی کے ساتھ جانچنے میں انتشار کا استعمال کیا جاتا ہے۔

3۔ اوسط کی جانچ یا اعتبار کرنا:

اوسط علم شماریات کا نارمل و مقبول پہلو ہے۔ اوسط کی قدر کی سچائی کی جانچ اور قابل اعتبار کو محسوب کرنے کے لیے انتشار کا استعمال کیا جاتا ہے۔ حاصل مواد یا نتائج اوسط کے قریب ہو یا انتشار قابل نظر انداز مواد میں ہو تب حاصل مواد پر اعتبار میں اضافہ ہوتا ہے۔ اور نتائج کو قبول کیا جاتا ہے۔ اس طرح حاصل نتائج اوسط سے دور یا اوسط اور حاصل نتائج میں کافی فرق ہو تو اس پر کم اعتبار قبول کیا جاتا ہے۔

شماریات کے دوسرے پیمانوں کی پیمائش:

انتشار کی مدد سے علم شماریات کے مختلف قدروں کو محسوب کرنے کے لیے استعمال کیا جاتا ہے۔ Fvalue کوریلیشن کی قدر Regression،Degree of Correction وغیرہ کو محسوب کرنے کے لیے استعمال کیا جاتا ہے۔

تحقیق میں استعمال: انتشار کو محققین اپنی تحقیقی نتائج کو اخذ کرنے میں استعمال کرتے ہیں خاص کر سماجی علوم یا سماجی مسائل یا سماجیات کے موضوع پر اس کا خاص استعمال ہوتا ہے۔ سماجی، آمدنی، تعلیم، سماجی و معاشی موقف معیار زندگی، آمدنی کی سطح وغیرہ کے تخمینہ میں انتشار کا استعمال کیا جاتا ہے۔

ماحولیات کے تبدیلی کے اثر کا مطالعہ: انتشار کی مدد سے موسم و ماحول میں تبدیل کے اثرات کا مطالعہ کے لیے ہی استعمال کیا جاتا ہے۔ رزو بروز موسمی تبدیلی، ماحولیات میں تبدیلی، صنعتوں میں تبدیلی، رسم و رواج میں تبدیلی سے سماج پر اثر انداز مختلف عوامل کا مطالعہ کرنے میں انتشار کا استعمال کیا جاتا ہے۔ انتشار کو متغیرات کی تبدیلی اور اس کے معکوس اثرات کو مطالعہ کرنے کے ساتھ ساتھ بااصول کنٹرول میں مدد ملتی ہے۔

5.4 اچھے انتشار کی اہم خصوصیات(Features of Good Dispersion)

انتشار کے چند اہم مثالی خصوصیات کو ذیل میں بتلایا گیا ہے۔

1۔ انتشار قابل فہم اور ہر عام آدمی کے لیے قابل قبول ہوتا ہے۔

2۔ مختلف طریقوں سے انتشار کی قدر یا قیمت یکساں حاصل ہونی چاہیے۔ مختلف طریقوں کو اختیار کرنے کے باوجود حاصل نتائج میں مشترک ہوں۔

3۔ انتشار تمام مواد کی نمائندگی کرتا ہے۔

4۔ انتشار کو الجبرائی انداز میں پیش کیا جاتا ہے۔

5.5 اوسط انحراف (Mean Deviation)

اوسط انحراف کے معنی اور تعریف کی تفہیم کے لیے حسب ذیل تعریفات کا مطالعہ کیجئے۔

1۔ دیے ہوئے اعداد کا اوسط (Average) سے تقابل کرتے ہوئے فرق کو محسوب کرنا اوسط انحراف کہلاتا ہے۔

2۔ اوسط انحراف ایک اہم پیمانہ ہے جس کی مدد سے فراہم کیے گئے اعداد کا اوسط کی مدد سے انحراف محسوب کرنا اوسط انحراف کہلاتا ہے۔

3۔ دیے گئے اعداد اور اوسط کے درمیان انحراف کو اوسط انحراف کہلاتا ہے۔

4۔ دیے ہوئے اعداد سے حاصل اوسط اور حقیقی اعداد کے درمیان فرق کو اوسط انحراف کہتے ہیں۔

5.5.1 اوسط انحراف کی خصوصیات (Features of Mean Deviation)

اوسط انحراف کے چند اہم خصوصیات کو ذیل میں بتلایا گیا ہے۔

1۔ اوسط انحراف کا تعلق دیے ہوئے اعداد پر منحصر ہوتے ہیں۔

2۔ یہ دیے ہوئے اعداد اور اوسط کے درمیان فرق کو ظاہر کرتا ہے۔

3۔ اوسط انحراف کی تفہیم کافی آسان ہے۔

4۔ اوسط انحراف کی معاشی میدان، کاروبار، صنعتی پیداوار، کامرس، تجارت منصوبہ بندی، موازنہ سازی وغیرہ مختلف میدانوں میں استعمال ہوتا ہے۔

5.5.2 اوسط انحراف کی خامیاں (Limitations of Mean Deviation)

اوسط انحراف کے چند اہم خامیوں کو ذیل میں بتلایا گیا ہے۔

1۔ اوسط انحراف کا الجبرائی انداز پیش کرنا مشکل ہے۔

2۔ اوسط انحراف کا انحصار دیے ہوئے اعداد پر منحصر ہے۔

3۔ اوسط انحراف کا سائنسی انداز میں تخمینہ ممکن نہیں ہے۔

4۔ اوسط انحراف میں حسابی علامتوں کو نظر انداز کیا گیا ہے۔

5۔ اوسط انحراف کے نتائج صد فیصد درست ہونا لازمی نہیں ہے۔

6۔ علم سماجیات میں اوسط انحراف کا بہت ہی کم استعمال ہوتا ہے۔

5.5.3 انفرادی سلسلے میں انحراف کا تعین کرنا

(Determination of Deviation in case of Individual Series)

انفرادی سلسلے میں اوسط انحراف کا تعین کے لیے ذیل کے مراحل پر عمل کریں۔

1۔ دیے ہوئے اعداد کا اوسط محسوب کریں۔

2۔ دیے ہوئے اعداد کا اوسط سے تقابل کرتے ہیں انحراف محسوب کریں۔

3۔ ضابطہ میں قیمت درج کریں۔

$$MD = \frac{\Sigma|D|}{N}$$

نوٹ: دو عمودی خط کے درمیان D ہے انہیں ماڈ D پڑھا جائے۔

مثال 1: دیے ہوئے اعداد کی مدد سے اوسط انحراف محسوب کیجئے۔

3,8,6,7,5,2,4,5

حل: دیے ہوئے اعداد کا اوسط محسوب کریں۔

40=3+8+6+7+5+2+4+5 اوسط

$$\overline{X} = \frac{\Sigma X}{4} = \frac{40}{8} = 5$$ اوسط

اوسط انحراف کو ذیل کے شکل میں وضاحت کی گئی ہے۔

اعداد (X)	-X-X= $\overline{X}$=5
5	0
4	1
2	3
5	0
7	2
6	1
8	3
3	2

	Ed=12

$$MD = \frac{\Sigma|D|}{N}$$
$$= \frac{12}{8}$$

اوسط انحراف $MD=1.5$

5.5.4 اوسط انحراف کی قدر ضربی (Coefficient of Mean Deviation)

اوسط انحراف کو اوسط سے تقسیم کرنے پر حاصل عدد اوسط انحراف کی قدر ضربی کہلاتا ہے۔ یہ دراصل اوسط انحراف اور اوسط کے درمیان تعلق کو ظاہر کرتی ہے۔ ذیل کے ضابطے کی مدد سے اوسط انحراف کی قدر ضربی محسوب کیا جاتا ہے۔

Coefficient of Mean Deviation = Mean Deviation / Mean

مثال:2 ذیل کے اعداد کی مدد سے اوسط انحراف کی قدر ضربی محسوب کی جاتی ہے۔

اوسط انحراف 1.5=، اوسط 5=

Coefficient of Mean Deviation = Mean Deviation (with mean) / Mean

5/1.5=

Coefficient of Mean Deviation = 0.3

5.5.5 تعددی سلسلہ میں اوسط انحراف کی تحسیب

(Computation of Mean Deviation in Discrete Series)

تعددی سلسلہ میں دو قطار میں مواد فراہم کیا جاتا ہے۔ پہلے قطار میں اعداد کی کیفیت (X) اور دوسرے قطار میں ان کی تعداد (F) کو ظاہر کیا جاتا ہے۔

مثال 3: ذیل میں دیے گئے مواد کی مدد سے اوسط انحراف اور اس کی قدر ضربی محسوب کیجئے۔

نشانات(X)	10	20	30	40	50
طلباء کی تعداد(F)	1	3	5	4	2

حل: اوسط کی مدد سے اوسط انحراف کی تحسیب:

نشانات(X)	طلبا کی تعداد(F)	fx (3)	$X-\bar{X}$ (4)	2x4 $FX-(X-\bar{X})$ (5)
10	1	10	22	22

20	3	60	12	36
30	5	150	2	10
40	4	160	8	32
50	2	100	18	36
	Sf=15	Sfd=480		Sfd=136

15=Sf, 480=Sfx قیمتیں درج کرنے پر

$$\text{اوسط}\ \overline{X} = \frac{\Sigma fx}{\Sigma N}$$

$$= \frac{480}{15}$$

$$\text{اوسط}\ \overline{X} = 32$$

$$= \frac{\Sigma fdy}{\Sigma F}$$

$$\text{Mean Deviation (with mean)} = \frac{136}{15}$$

Mean Deviation (with mean) = 9.06

Coefficient of Mean Deviation = Mean Deviation / Mean

32 = (Mean)اوسط   9.06 = (Mean Deviation) اوسط انحراف

Coefficient of Mean Deviation = 9.06 / 32

5.5.6 مسلسل سلسلہ میں اوسط کی مدد سے اوسط انحراف کی تحسیب

(Calculation of Mean Deviation with the help of Mean)

مسلسل سلسلہ میں وقفہ جماعت کے ساتھ تعداد دی جاتی ہے۔اس صورت میں حسب ذیل مراحل کے ذریعہ اوسط انحراف محسوب کیا جاسکتا ہے۔

مثال:4 ذیل میں مضمون شماریات میں حاصل نشانات اور طلبا کی تعداد دی گئی ہے اس کے مدد سے اوسط کے ذریعہ اوسط انحراف اور اس کی قدر ضربی محسوب کیجئے۔

وقفہ جماعت(CI)	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
تعدد(F)	2	2	2	5	4	2	2	1

حل: اوسط کے ذریعہ اوسط انحراف کی تحسیب

نشانات وقفہ جماعت (1)	طلبا کی تعداد (F) (2)	وسطی نقطہ (X) MP (3)	mf (4)	X-38 (5)	$F(X-\overline{X})$ 2x5 (6)
0-10	2	5	10	33 (5-38)	66 (33x2)
10-20	2	15	30	23 (15-38)	46(23x2)
20-30	2	25	50	13 (25-36)	26 (13x2)
30-40	5	35	175	3(35-38)	15 (5x3)
40-50	4	45	180	7 (45-38)	38 (4x7)
50-60	2	55	110	17 (55-38)	34 (17x2)
60-70	2	65	130	27(65-38)	54 (27x2)
70-80	1	75	75	37 (75-38)	37 (1x37)
	Sf = 20		Smf = 760		Sfd = 306

$$\text{اوسط } \overline{X} = \frac{\Sigma mf}{\Sigma F}$$

$$= \frac{760}{20}$$

$$\text{اوسط } \overline{X} = 38$$

$$\text{Mean Deviation with mean} = \frac{\Sigma f(X-\overline{X})}{\Sigma F}$$

$$= \frac{306}{20}$$

Mean Deviation = 20.4

Coefficient of Mean Deviation = Mean Deviation / Mean

Coefficient of Mean Deviation = 20.4 / 38

اوسط انحراف کی قدر ضربی = 0.5368

5.6 معیاری انحراف کے معنی اور تعریف

(Meaning and Definition of Standard Deviation)

معیاری انحراف کے معنی اور تعریف کی تفہیم کے لیے ذیل کے تعریفات کا مطالعہ کیجئے۔

1۔ حسابی اوسط کے مشاہدات کے اخذ کردہ انحرافات کے مربعوں کے اوسط کے جذرالمربع کو معیاری انحراف کہتے ہیں۔

2۔ معیاری انحراف انتشار کو محسوب کرنے کا ایک طریقہ ہے زیادہ انتشار سے معیاری انحراف کی اونچی قدر حاصل ہوتی ہے اور کم انتشار کی صورت میں معیاری انحراف کی کم قدر حاصل ہوتی ہے۔

معیاری انحراف مواد کے انتشار یا پھیلاؤ کی جانچ کرتی ہیں۔ معیاری انحراف کی کم قدر اوسط (mean) سے قربت، نزدیک یا قریب کو ظاہر کرتی ہے۔ اسی طرح بلند یا اونچا معیاری انحراف اعداد کا اوسط سے دور یا پھیلاؤ کو ظاہر کرتا ہے۔ معیاری انحراف کو انگریزی حراف SD سے اور یونانی علامت (S) سگما سے ظاہر کیا جاتا ہے۔

لفظ معیاری انحراف (Standrad Deviation) کو سب سے پہلے 1893 کو کارل پیرسن (Karl Pearson) نے برطانوی ماہر ریاضی داں نے استعمال کیا۔ انتشار کے مطالعہ کے لیے معیاری انحراف کا استعمال کیا جاتا ہے۔

5.6.1 معیاری انحراف کی نوعیت (Nature of Standard Deviation)

انحراف علم ریاضی میں استعمال ایک اہم پیمانہ ہے یہ حاصل اوسط انحراف کے مربع کے مجموعہ ہوتا ہے۔ اس کی مدد سے اوسط سے انتشار کیفیت کو محسوس کیا جاتا ہے۔ یہ اوسط سے نزدیک اور دوری یا پھیلاؤ کو ظاہر کرتا ہے۔ معیاری انحراف کی کم قدر اوسط سے قربت اور دوری اوسط سے پھیلاؤ کو ظاہر کرتا ہے۔ یعنی معیاری انحراف اوسط سے قربت کی سطح یا قربت کے معیار کو ظاہر کرتی ہے۔ تاجرین صنعت کار، تعلیمی ادارے وغیرہ اس کو استعمال کرتے ہیں۔ مکانات کی تجارت کی تاجرین مقام یا علاقہ کے حساب سے مکان کی قیمت کے تعین میں معیاری انحراف کو استعمال کرتے ہیں۔ اس کی قدر صفر تا ایک کے درمیان ہوتی ہے۔ لیکن اس کی قدر منفی نہیں ہوتی۔

5.6.2 معیاری انحراف کا استعمال واطلاق (Uses and Application of Standard Deviation)

معیاری انحراف کے چند اہم استعمالات کو ذیل میں بتلایا گیا ہے۔

1۔ انتشار کی کیفیت کو محسوب کرنا۔

2۔ اوسط اور انتشار کے درمیان موجود فرق یا پھیلاؤ کو محسوب کرنا۔

3۔ Volatility کو محسوب کرنے کے لیے استعمال کیا جاتا ہے۔

4۔ مواد یا معطیات کا تقابل کرنے کے لیے استعمال کیا جاتا ہے۔

5۔ کاروباری جوکھم کی پیمائش کے لیے استعمال کیا جاتا ہے۔

6۔ صارفین یا گاہکوں کی تشفی، رد عمل، ذخیرہ، قیمتوں کی سطح وغیرہ کو محسوب کرنے کے لیے استعمال کیا جاتا ہے۔

7۔ صنعتی پیداوار کے معیار کی جانچ اور غلطیوں یا نقائص کی پیمائش کے لیے استعمال کیا جاتا ہے۔

8۔ متوقع اور حقیقی مواد کے فرق کو محسوب کرنے کے لیے استعمال کیا جاتا ہے۔

9۔ تحقیقی و تجزیہ کے ذریعہ نتائج کو محسوب کرنے کے لیے استعمال کیا جاتا ہے۔

10۔ تجارتی و تعلیمی کا کردگی کو ظاہر کرنے کے لیے استعمال کیا جاتا ہے۔

5.6.3 معیاری انحراف کی خوبیاں(Merits of Standard Deviations)

معیاری انحراف کی اہم خوبیوں کو ذیل میں بتلایا گیا ہے۔

1۔ انتشار کے تمام پیمانوں میں معیاری انحراف کو بہتر اور برتر سمجھا جاتا ہے۔ کیوں کہ معیاری انحراف میں انتشار کے اچھے لوازمات پائے جاتے ہیں۔

2۔ معیاری انحراف معطیات کے تمام مشاہدوں پر مشتمل ہوتا ہے۔

3۔ معیاری انحراف نمونہ کاری کی تبدیلی یا اتار چڑاؤ سے متاثر نہیں ہوتا۔

4۔ معیاری انحراف انتشار کے دوسرے پیمانوں کے مقابلے بہتر ہے۔

5۔ معیاری انحراف الجبرائی وضاحت کے لیے موزوں پیمانہ ہے۔

6۔ مخلوط معیاری انحراف کو آسانی سے محسوب کیا جاسکتا ہے۔

7۔ انتشار کی صحت کے ساتھ تحسیب میں مدد ملتی ہے۔

5.6.4 معیاری انحراف کی خامیاں(Demerits of Standard Deviations)

معیاری انحراف کی چند اہم خامیوں کو ذیل میں بتلایا گیا ہے۔

1۔ دو یا زیادہ سلسلوں کے انتشار کے تقابل میں معیاری انحراف کا استعمال نہیں کیا جاسکتا۔

2۔ معیاری انحراف کی تحسیب پیچیدہ عمل ہے۔

3۔ کھلے کنارے کی صورت میں معیاری انحراف کی تحسیب مشکل ہوتی ہے۔

4۔ معیاری انحراف کی تحسیب کے لیے علم ریاضی کے اصولوں پر کامل قدرت کا ہونا ضروری ہے۔

5.6.5 انفرادی سلسلہ میں معیاری انحراف کی تحسیب

(Calculation of Standard Deviation in case of Individual Series)

انفرادی سلسلہ اعداد کو پیش کرنے کا ایک اہم طریقہ ہے۔اس طریقہ ہے اعداد کو غیر گروہ بند انداز میں پیش کیا جاتا ہے۔اعداد بے ترتیب پھیلے ہوئے ہوتے ہیں۔معیاری انحراف کو راست اور بالراست دونوں طریقوں پر محسوب کر سکتے ہیں۔

معیاری انحراف کی تحسیب کا راست طریقہ(Direct Method of Standard Deviation)

مثال 1: ذیل کے مواد کی مدد سے معیاری انحراف(SD) محسوب کیجئے۔

1, 3, 5, 4, 6, 11

حل: 1۔دیے ہوئے اعداد کا اوسط محسوب کریں۔

2۔اوسط سے اعداد کے انحراف کو محسوب کریں۔

3۔حاصل انحرافی اعداد کا مربع محسوب کریں۔

ذیل میں معیاری انحراف کو محسوب کرنے کے طریقہ کو بتلایا گیا ہے۔

اعداد(X)	$(dx=X-\bar{X})$ $\bar{X}=5$	dx^2
1	-4	16
3	-2	4
4	-1	1
5	0	0
1	+1	1
11	+6	36
X=30	- 7 + 7 Σdx 0	$\Sigma dX^2=58$

اوسط= $\bar{X} = \frac{\Sigma X}{\Sigma N}$

جہاں $\Sigma X =$ اعداد کا مجموعہ $= 30$، ΣN اعداد کی تعداد $= 6$

$$\overline{X} = \frac{30}{6}$$

$$\overline{X} = 5$$

$$\Sigma D \text{ معیاری انحراف } = \sqrt{\frac{\Sigma dX^2}{N}}$$

جہاں پر $\Sigma dX^2 = 58$، $N = 6$، قیمتیں درج کرنے پر

$$= \sqrt{\frac{58}{6}}$$

$$= \sqrt{9.67}$$

$$SD\ (s) \text{ معیاری انحراف } = 3.1096$$

$$\text{Coefficient of Standard Deviation} = \frac{\sigma}{\overline{X}} = \frac{SD}{Mean}$$

$$= \frac{3.1096}{5}$$

$$\text{Coefficient of Standard Deviation} = 0.62192$$

5.6.6 معیاری انحراف کی تحسیب کا بالواسطہ طریقہ

(Indirect Method for Calculation of Standard Deviation)

بالراست طریقہ پر معیاری انحراف کی تحسیب کے لیے ذیل کے مراحل پر عمل کریں۔

1۔ فرضی اوسط عدد کی مدد سے انحراف محسوب کریں۔

2۔ انحراف کا مربع محسوب کریں۔

3۔ ضابطہ میں قدریں درج کریں۔

اعداد(X)	dx = (X-AM) A=5	Σdx^2
1	-4	16 (-4 x -4)
3	-2	4 (-2 x -2)
4	-1	1 (-1 x -1)
5	0	0
6	+1	1 (1 x 1)
11	+6	36 (6 x 6)
	- 7 + 7 Σdx= 0	$\Sigma dX^2=58$

بالراست طور پر معیاری انحراف کی تحسیب کا ضابط

جہاں پر $58 = \Sigma dX^2$ $\quad$ $0 = \Sigma dx$ $\quad$ $6 = N$

$$\sigma = \sqrt{\frac{\Sigma dX^2}{N} - \left(\frac{\Sigma dX}{N}\right)^2}$$

$$= \sqrt{\frac{58}{6} - \frac{0}{6}}$$

$$= \sqrt{9.67 - 0}$$

SD معیاری انحراف = 3.1096

راست اور بالراست دونوں طریقوں سے معیاری انحراف کی قدر 3.1096 ہے۔ جو دونوں طریقوں میں مساوی قدر اخذ کی جاتی ہے۔

$$\text{Variance} = \sigma^2$$

$$= 3.1097 \times 3.1097$$

$$\text{Variance} = 9.670$$

5.6.7 تعددی سلسلہ میں معیاری انحراف کی تحسیب

(Calculation of Standard Deviation in Discrete Series)

تعددی سلسلہ میں درکار مواد کو دو قطار میں فراہم کیا جاتا ہے۔ تعددی سلسلہ میں راست اور بالراست دونوں طریقوں پر معیاری انحراف کو محسوب کیا جاتا ہے۔

1۔ راست طریقہ (حقیقی اوسط) Direct Method:

معیاری انحراف کو حقیقی اوسط کی مدد سے محسوب کرنا راست طریقہ کہلاتا ہے۔ اس طریقہ کے تحت حسب ذیل مراحل پر عمل کریں۔

1۔ حقیقی اوسط محسوب کریں۔

2۔ حقیقی اوسط کی مدد سے انحراف (dx) کو محسوب کریں۔

3۔ انحراف کا مربع محسوب کریں۔

4۔ انحراف کے مربع کو متعلقہ تعدد (F) سے ضرب دیں۔

5۔ ضابطہ میں قدریں درج کریں۔

5.6.8 بالواسطہ طریقہ (مفروضہ اوسط) کا طریقہ Indirect Method (Assumed Mean)

بالواسطہ یا مفروضہ اوسط کے مدد سے معیاری انحراف کو محسوب کیا جاتا ہے۔ اس طریقہ کے تحت دیے ہوئے اعداد میں سے کسی ایک عدد کو اوسط فرض کریں اور اس سے انحراف کو محسوب کریں۔ اس طریقہ کے تحت حسب ذیل مراحل پر عمل کریں۔

1۔ دیے ہوئے اعداد سے کسی ایک عدد کو بطور اوسط فرض کریں۔

2۔ فرض کیے ہوئے عدد سے اعداد کا انحراف محسوب کریں جس کو dx سے ظاہر کیا جاتا ہے۔

3۔ حاصل انحرافی اعداد (dx) کا مربع محسوب کریں جس کو (dx^2) سے ظاہر کیا جاتا ہے۔

4۔ انحراف کے مربع کو متعلقہ تعدد (f) سے ضرب دیں جس کو $\sum fdx^2$ سے ظاہر کیا جاتا ہے۔

5۔ ضابطہ میں قدریں درج کریں۔

مثال2: ذیل کے تفصیلات کی مدد سے راست اور بالراست دونوں طریقوں پر معیاری انحراف اور تفاوت/تغیر (variation) محسوب کیجئے۔

X	8	10	12	14	16	18	20
تعدد(F)	1	2	4	5	3	3	2

حل: راست اور حقیقی اوسط کے طریقہ پر معیاری انحراف کی تحسیب۔

X	F	fx	d(x-x)	fd	fd2(fdxd)
6	1	6	-8	-8	64
8	2	16	-6	-12	72
12	5	60	-2	-10	20
14	4	56	0	0	0
16	3	48	+2	6	12
18	3	54	+4	12	48
20	2	40	+6	12	72
	N=20	$\sum Fx=280$			$\sum Fd^2=288$

$$\overline{X} = \frac{280}{20}$$

$$\overline{X} = 4$$

$$\sigma = \sqrt{\frac{\Sigma dX^2}{N}}$$

$$= \sqrt{\frac{288}{20}}$$

$$= \sqrt{14.4}$$

معیاری انحراف $= 3.7947$

$$\text{Variance} = \sigma^2$$

$$= 3.7947 \times 3.7947$$

$$\text{Variance} = 14.4$$

بالراست (مفروضہ اوسط) کے طریقہ کے تحت معیاری انحراف کی تحسیب

X	F	d A = 16	fd	Fdxd fd^2
6	1	-10	-10	100
8	2	-8	-16	128
12	5	-4	-20	80
14	4	-2	-8	16
16	3	0	0	0
18	3	+2	6	12
20	2	+4	8	32
	N = 20		-54 +14 ∑Fd= - 40	$\sum Fd^2=368$

$$\text{معیاری انحراف } \sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$

$$= \sqrt{\frac{368}{20} - \left(\frac{40}{20}\right)^2}$$

$$= \sqrt{18.4 - 4}$$

$$= \sqrt{14.4}$$

$$\text{معیاری انحراف } \sigma = 3.7947$$

5.6.9 مسلسل سلسلہ میں معیاری انحراف کی تحسیب

(Calculation of Standard Deviation in Continuous Series)

مسلسل سلسلہ گروہی مواد ہے جس میں وقفہ جماعت کے ساتھ تعدد دیا جاتا ہے۔ مسلسل سلسلہ میں معیاری انحراف کے لیے ذیل کے مراحل پر عمل کریں۔

1۔ مسلسل سلسلہ کے ہر وقفہ جماعت کا وسطی نقطہ محسوب کریں۔

2۔ راست طریقہ پر اوسط محسوب کریں۔

3۔ حقیقی اوسط کی مدد سے وسطی نقاط کا تقابل کرتے ہوئے انحراف کی قدر کو محسوب کریں۔

مثال 3: ذیل میں دیے ہوئے مواد کی مدد سے راست طریقہ پر معیاری انحراف محسوب کریں۔

نشانات(X)	0-10	10-20	20-30	30-40	40-50	50-60	60-70
تعدد(F)	1	2	3	4	5	4	1

حل: حقیقی اوسط کی مدد سے معیاری انحراف کی تحسیب

وقفہ جماعت	F	وسطی نقطہ Mp	Mpf	d= x̄=38 Mp - x̄	fd(1x5) 6	fd^2 fdxd
0-10	1	5	5	33-	33-	1089
10-20	2	15	30	-23	-46	1058
20-30	3	25	75	-13	-39	507
30-40	4	35	140	-13	-12	36
40-50	5	45	225	+7	35	245

50-60	4	55	220	+17	68	1156
60-70	1	65	65	+27	27	729Type equation here.
	N = 20		Mf = 760		-130 +130 $\sum Fd=0$	$\sum Fd^2=4820$

$$\bar{x} = \frac{\sum mf}{\sum f} \quad \bar{x}=\frac{760}{20} \quad \bar{x}=38$$

$$= \sqrt{\frac{\sum fd^2}{\sum f}} = \sqrt{\frac{4820}{20}} = \sqrt{241}$$

$$\text{(SD)معیاری انحراف} = 15.524\sigma$$

5.6.10 فرضی اوسط کی مدد سے معیاری انحراف کی تحسیب

(Calculation of Standard Deviation with the help of Assumed Mean)

1۔ وسطی نقطہ محسوب کریں۔

2۔ وسطی نقطہ میں سے کسی ایک عدد کو اوسط فرض کریں۔

وسطی نقطہ 35 کو اوسط عدد فرض کیا گیا۔

فرضی اوسط (35) کی مدد سے معیاری انحراف کی تحسیب

وقفہ جماعت	تعدد F	وسطی نقطہ mp	X-A A=35	Fd	fdxd fd^2
0-10	1	5	-30	-30	900
10-20	2	15	-20	-40	800
20-30	3	25	-10	-30	300
30-40	4	35	0	0	0
40-50	5	45	+10	+50	500
50-60	4	55	+20	+80	1600
60-70	1	65	+30	+30	900
	N = 20			-100	$Fd^2=5000$

				+100 ∑fd=+60	

$$=\sqrt{\frac{\sum fd^2}{N}-\left[\frac{\sum fd}{N}\right]2\sigma}$$

جہاں پر $5000 = \sum fd^2$ $\quad 20 = N \quad 6 = \sum fd,$

$$=\sqrt{\frac{5000}{20}-\left[\frac{60}{20}\right]2}=\sqrt{250-9}$$

$$=\sqrt{241} = 15.524$$

اقدامی انحراف/اختصار کے طریقہ پر معیاری انحراف کے تحسیب

Calculation of Standard Deviation by Step Deviation Method

وقفہ جماعت	F تعدد	Mp وسطی نقطہ	A=35	d/10	fd	fd^2 fdxd
0-10	1	5	-30	-3	-3	9
10-20	2	15	-20	-2	-4	8
20-30	3	25	-10	-1	-3	3
30-40	4	35	0	0	0	0
40-50	5	45	+10	+1	5	5
50-60	4	55	+20	+2	8	16
60-70	1	65	+30	+3	3	9
					-10 +16 ∑fd +6	fd-50

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left[\frac{\Sigma fd}{N}\right]^2} \times i$$

جہاں پر $50 = \Sigma fd^2$ $\quad 20 = N \quad 6 = \Sigma fd, \quad 10 = i,$

$$= \sqrt{2.5 - 0.09} \times 10$$

$$= \sqrt{2.41} \times 10$$

$$= 1.5524 \times 10$$

$$= 15.524$$

مندرجہ بالا تینوں طریقوں پر معیاری انحراف کی قدر 15.524 ہو گی۔ طریقہ تبدیل ہے لیکن قیمت میں کوئی تبدیلی واقع نہ ہو گی

5.7 اکتسابی نتائج (Learning Outcome)

طلباء نے بازی کے معنی اور تعریف اور ڈیٹا کا خلاصہ اور تشریح کرنے میں اس کی اہمیت کے بارے میں واضح سمجھ پیدا کر لی ہے۔ انہوں نے اچھی بازی کی اہم خصوصیات اور ڈیٹا سیٹس میں تغیر یا پھیلاؤ کا اندازہ کرنے کا طریقہ سیکھا ہے۔ طلباء بازی کی نوعیت کی وضاحت کر سکتے ہیں اور تغیر کے مختلف پیمانوں کے درمیان فرق کر سکتے ہیں۔ انہوں نے معیاری انحراف کا بھی علم حاصل کر لیا ہے، بشمول اس کی تعریف، حساب اور اطلاق، جس سے وہ اس بات کا اندازہ لگا سکتے ہیں کہ ڈیٹا پوائنٹس وسط سے کیسے ہٹتے ہیں۔ مجموعی طور پر، سیکھنے والے ڈیٹا کی تغیرات کا مؤثر طریقے سے تجزیہ کرنے کے قابل ہوتے ہیں اور اسے زیادہ درست اور باخبر کاروباری فیصلے کرنے کے لیے استعمال کرتے ہیں۔

5.8 نمونہ امتحانی سوالات (Model Examination Questions)

خالی جگہوں کو پر کیجئے۔

1۔ معیاری انحراف کو سب سے پہلے ....نے استعمال کیا۔

2۔ راست طریقہ پر معیاری انحراف کو محسوب کرنے کا ضابطہ...............ہے۔

3۔ کار پیرسن کا تعلق...............ملک سے ہے۔

4۔ معیاری انحراف کو...............علامت سے ظاہر کیا جاتا ہے۔

5۔ جذر المربع کو...............علامت سے ظاہر کیا جاتا ہے۔

مختصر جوابات کے حامل سوالات

1۔ معیاری انحراف سے کیا مراد ہے۔ وضاحت کیجئے۔

2۔ معیاری انحراف کے استعمالات کو بیان کیجئے۔

3۔ دیے ہوئے اعداد کا معیاری انحراف محسوب کیجئے۔10,9,8,6,5,3

4۔ دیے ہوئے اعداد کی مدد سے حقیقی اوسط کے طریقہ پر معیاری انحراف محسوب کیجئے۔4,6,8,10,12,18,20,22

5۔ ذیل میں دیے ہوئے تفصیلات کی مدد سے معیاری انحراف اور اس کی قدر ضربی محسوب کیجئے۔

X	10	20	30	40	50	60	70	80
F	1	2	2	3	4	5	2	1

طویل جوابات کے حامل سوالات

6۔ دیے ہوئے اعداد کی مدد سے راست اور بالراست طریقہ پر معیاری انحراف اور تفاوت محسوب کیجئے۔

X	10	20	30	11	14	17	20	23
F	1	2	2	3	4	5	2	1

7۔ ذیل میں ایک مضمون میں حاصل نشانات اور طلبا کی تعداد دی گئی ہے۔ جن کی مدد سے راست اور بالراست طریقہ پر معیاری انحراف محسوب کیجئے۔

X	0-10	10-20	20-30	30-40	40-50	50-60	60-70
F	1	2	2	3	5	4	3

8۔ ذیل کے تفصیلات کی مدد سے راست بالراست اور اختصار کے طریقہ کے تحت معیاری انحراف محسوب کیجئے

X	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
F	1	1	3	4	5	7	6	3

9۔ ایک صنعت میں ملازم مزدوروں کی یومیہ اجرت دی گئی ہے۔ جس کی مدد سے راست باالرست اور اختصار کے طریقہ پر معیاری انحراف محسوب کیجئے۔

X	100-200	200-300	300-400	400-500	500-600	600-700
F	4	3	5	8	6	4

اکائی6۔ حتمال اور احتمال کی تقسیم

(Probability and Probability Distribution)

Structure	اکائی کے اجزاء
Introduction	6.0 تمہید
Objectives	6.1 مقاصد
Meaning and Definition of Probability	6.2 احتمال کا معنی اور تعریف
Important Terminology of Probability	6.3 احتمال کی اہم اصطلاحات
Probability Rules	6.4 احتمال کے اصول
Types of Probabilities Distributions	6.5 احتمالی تقسیم کے اقسام
Learning Outcomes	6.6 اکتسابی نتائج
Model Examination Questions	6.7 نمونہ امتحانی سوالات

6.0 تمہید (Introduction)

احتمال ریاضی کی ایک شاخ ہے جو دی گئی شرائط کے تحت رونما ہونے والے ایونٹس کے احتمالات کی پیائش سے متعلق ہے۔ یہ کاروبار، اقتصادیات، سائنس، انجینئرنگ، اور روزمرہ کی زندگی جیسے شعبوں میں بہتر فیصلہ سازی کو قابل بنا کر، غیر یقینی صورتحال کو درست کرنے کا ایک منظم طریقہ فراہم کرتا ہے۔ احتمالی نظریہ ریاضی کے اصولوں کو استعمال کرتے ہوئے نتائج کی پیشین گوئی، خطرات کا اندازہ لگانے اور حکمت عملی بنانے میں مدد کرتا ہے۔

دوسری طرف، ایک احتمالی تقسیم بیان کرتی ہے کہ بے ترتیب متغیر کے تمام ممکنہ نتائج کے لیے احتمالات کیسے تفویض کیے جاتے ہیں۔ یہ مجرد (discrete) ہو سکتا ہے، جہاں نتائج قابل شمار ہوتے ہیں (مثال کے طور پر، سکے کو اچھالنا، ڈائیس کو رول کرنا)، یا مسلسل، جہاں نتائج ایک حد میں آتے ہیں (جیسے، اونچائی، وقت، یا وزن کی پیائش)۔ یہ تقسیمیں، جیسے بائنومیل، پوائسن، اور نارمل تقسیم، شماریاتی تجزیہ اور مفروضے کی جانچ کی بنیاد بناتے ہیں۔

ڈیٹا کا تجزیہ کرنے، حقیقی دنیا کے حالات کی ماڈلنگ، اور باخبر پیشین گوئیاں کرنے کے لیے احتمال اور احتمالی تقسیم کو سمجھنا بہت ضروری ہے۔ یہ یونٹ احتمالی تقسیم کی کلیدی اقسام کے ساتھ احتمال کے بنیادی تصورات، قواعد اور اطلاقات کا احاطہ کرے گا، جس سے سیکھنے

والوں کو تعلیمی، پیشہ ورانہ اور تحقیقی سیاق وسباق میں ان تصورات کو مؤثر طریقے سے لاگو کرنے کے قابل بنایا جائے گا۔

6.1 مقاصد(Objectives)

اس اکائی کا مطالعہ کرنے کے بعد آپ:

- احتمال کے بنیادی تصورات اور تعریفوں ،بشمول اقسام اور نقطہ نظر کو سمجھ سکیں گے۔
- مجرد اور مسلسل احتمالی تقسیم کے درمیان فرق کر سکیں گے۔
- کلیدی احتمالی تقسیم کو لاگو کرنے کی صلاحیت حاصل کر سکیں گے جیسےBinomial،Poisson،اورNormal۔
- احتمالی نظریات اور تقسیم کے ماڈلز کا استعمال کرتے ہوئے حقیقی زندگی کے مسائل کا تجزیہ کر سکیں گے۔

6.2 احتمال کا معنی اور تعریف(Meaning and Definition of Probability)

ہماری روزمرہ کی زندگی میں "احتمال" یا "موقع" کی اصطلاح بہت عام استعمال ہوتی ہے۔ کبھی کبھی، ہم کہتے ہیں "شاید کل بارش ہو سکتی ہے"، "شاید مسٹرX آج اپنی کلاس لینے آئیں"، "شاید آپ ٹھیک کہہ رہے ہیں"۔ یہ تمام اصطلاحات، احتمال اور احتمال کا ایک ہی مطلب ہے۔ لیکن اعداد و شمار میں احتمال کا نارمل آٹیلی کے خیال کے برعکس کچھ خاص مفہوم ہوتا ہے۔

احتمال کا نظریہ 17ویں صدی میں تیار کیا گیا تھا۔ اس کی ابتدا کھیلوں، سکے پھینکنے، نرد پھینکنے، پیکٹ سے کارڈ کھینچنے سے ہوئی ہے۔ 1954 میں Antoine Gornbandنے اس کے لیے ایک پہل اور دلچسپی لی تھی۔

ان کے بعد، شماریات میں بہت سے مصنفین نے سابق کے ذریعہ دیے گئے خیال کو دوبارہ تشکیل دینے کی کوشش کی ہے۔ "احتمال" شماریات کے بنیادی ٹولز میں سے ایک بن گیا ہے۔ بعض اوقات، شماریاتی تجزیہ احتمال کے نظریہ کے بغیر مفلوج ہو جاتا ہے۔

"Probability of a given event is defined as the expected frequency of occurrence of the event among events of a like sort." (Garrett)

ریاضی میں احتمال کی تعریف Probability Definition in Math

احتمال ایونٹ(likelihood of an event)کے پیش آنے کے احتمال کا ایک پیمانہ ہے۔ بہت سے ایونٹس کی پوری یقین کے ساتھ پیشین گوئی نہیں کی جا سکتی۔ ہم صرف اس ایونٹ کے پیش آنے کے احتمالات کا اندازہ لگا سکتے ہیں، یعنی اس کا استعمال کرتے ہوئے اس کے ہونے کا کتنا احتمال ہے۔ احتمال 0 سے 1 تک ہو سکتا ہے، جہاں 0 کا مطلب ہے کہ ایونٹ ناممکن ہے اور 1 کسی خاص واقعے کی نشاندہی کرتا ہے۔ کلاس 10 کے لیے احتمالات طلباء کے لیے ایک اہم موضوع ہے، جو اس موضوع کے تمام بنیادی تصورات کی وضاحت کرتا ہے۔ نمونے (sample)میں تمام ایونٹس کا احتمال 1 تک بڑھ جاتا ہے۔

مثال کے طور پر، جب ہم ایک سکے کو ٹاس کرتے ہیں، یا تو ہمیں ہیڈ یا ٹیل مل جاتا ہے، صرف دو ممکنہ نتائج ممکن ہیں (H, T)۔ لیکن جب دو سکے پھینکے جائیں تو چار ممکنہ نتائج نکلیں گے، یعنی {(H, H), (H, T), (T, H), (T, T)}۔

تناسب کے طور پر احتمال (Probability as Ratio)

ریاضیاتی طور پر بیان کردہ یا اظہار کردہ ایونٹ کا احتمال تناسب کہلاتا ہے۔ غیر جانبدار سکے کے گرنے کا احتمال 2/1 ہے، اور نرد کا احتمال 6/1 ہے۔ یہ تناسب، جنہیں احتمالی تناسب کہا جاتا ہے، کی تعریف اس حصے سے کی جاتی ہے، جس کا عدد مطلوبہ نتائج یا نتائج کے برابر ہوتا ہے، اور جس کا اعشاریہ کل ممکنہ نتائج کے برابر ہوتا ہے۔

مزید آسان الفاظ میں، نرد پر کسی بھی عدد کے ظاہر ہونے کا احتمال (مثلاً، 4) 6/1 ہے۔

ایونٹ ہونے کا احتمال P(E)=سازگار نتائج کی تعداد/نتائج کی کل تعداد

اس طرح، ایک احتمال ایک عدد یا تناسب ہے جو 0 سے 1 تک ہوتا ہے۔ کسی واقعے کے لیے صفر جو واقع نہیں ہو سکتا اور 1 کسی واقعے کے لیے، جس کا واقع ہونا یقینی ہے۔

کلاسیکل احتمال (Classical Probability)

احتمالات کا کلاسیکل نقطہ نظر قدیم ترین اور سادہ ترین مکتب فکر میں سے ایک ہے۔ اس کی ابتدا 18 ویں صدی میں ہوئی ہے جو احتمالات کے کھیلوں جیسے سکے پھینکنا، نرد، کارڈ ڈرائنگ وغیرہ سے متعلق احتمال کی وضاحت کرتی ہے۔

احتمال کی تعریف ایک فرانسیسی ریاضی دان نے دی ہے جس کا نام "Laplace" ہے۔ ان کے مطابق احتمال مساوی طور پر ممکنہ کیس (favourable cases) کی تعداد کے درمیان سازگار کیس کی تعداد کا تناسب ہے۔ یا دوسرے الفاظ میں، کلاسیکل نقطہ نظر کے ذریعہ تجویز کردہ تناسب یہ ہے:

احتمال=سازگار کیس کی تعداد/ممکنہ صورتوں کی تعداد

Probability = Number of favourable cases/Number of equally likely cases

مثال کے طور پر، اگر ایک سکہ اچھالا جاتا ہے، اور اگر اس سے پوچھا جاتا ہے کہ ھیڈ کے ہونے کا کیا احتمال ہے، تو موافق صورت کی تعداد=1، مساوی طور پر ممکنہ صورتوں کی تعداد=2۔ ھیڈ کا احتمال=½

Symbolically it can be expressed as:

P = Probability (A) = a/n, q = Probability (B) or (not A) = b/n

1 – a/n = b/n = (or) a + b = 1 and also p + q = 1

p = 1 – q, and q = 1 – p and if a + b = 1 then so also a/n + b/n = 1

اس نقطہ نظر میں احتمال 0 سے 1 تک مختلف ہوتا ہے۔ جب احتمال صفر ہوتا ہے تو یہ ظاہر کرتا ہے کہ اس کا ہونا ناممکن ہے۔ اگر احتمال 1 ہے تو

وقوع پذیر ہونے کا یقین ہے، یعنی ایونٹ واقع ہونے کا پابند ہے۔

Example:

From a bag containing 20 black and 25 white balls, a ball is drawn randomly. What is the probability that it is black.

Solution: Probability of a black ball = 20/45 = 4/9 = p, Probability of a white ball = 25/45 = 5/9 = q

p = 4/9 and q = 5/9

(p + q= 4/9 + 5/9= 1)

کلاسیکل نقطہ نظر کے نقصانات Demerits of the classical approach:

1) کلاسیکل نقطہ نظر صرف سکے، نرد، تاش وغیرہ تک محدود ہے۔
2) یہ بعض صورتوں میں اصل نتیجہ کی وضاحت نہیں کر سکتا۔
3) اگر مساوی احتمالی صورتوں کی تعداد زیادہ ہے، تو احتمالی تناسب کی قدروں کا پتہ لگانا مشکل ہے، اور
4) اگر مساوی طور پر ممکنہ کیس کی تعداد 00 ہے، تو یہ نقطہ نظر ناکافی ہے۔

Solved Examples

1) There are 6 pillows in a bed, 3 are red, 2 are yellow and 1 is blue. What is the probability of picking a yellow pillow?

Ans: The probability is equal to the number of yellow pillows in the bed divided by the total number of pillows, i.e. 2/6 = 1/3.

2) There is a container full of coloured bottles, red, blue, green and orange. Some of the bottles are picked out and displaced. Sumit did this 1000 times and got the following results:

- No. of blue bottles picked out : 300
- No. of red bottles : 200
- No. of green bottles : 450
- No. of orange bottles : 50

a) What is the probability that Sumit will pick a green bottle?

Ans: For every 1000 bottles picked out, 450 are green.

Therefore, P(green) = 450/1000 = 0.45

b) If there are 100 bottles in the container, how many of them are likely to be green?

Ans: The experiment implies that 450 out of 1000 bottles are green.

Therefore, out of 100 bottles, 45 are green.

6.3 احتمال کی اہم اصطلاحات(Important Terminology of Probability)

احتمال کی اہم اصطلاحات:

1) باہمی خصوصی ایونٹس:Mutually Exclusive Events

ایونٹس(events) کو باہمی طور پر خصوصی کہا جاتا ہے جب وہ ایک ساتھ نہیں ہوتے ہیں۔ایونٹس میں، اگر ایک ایونٹ ٹرائل میں موجود رہتا ہے، تو دوسرے ایونٹس ظاہر نہیں ہوں گے۔دوسرے الفاظ میں، ایک کا وقوع باقی تمام کے وقوع کو روکتا ہے۔

مثال کے طور پر:

اگر لڑکی خوبصورت ہے تو وہ بدصورت نہیں ہو سکتی۔اگر گیند سفید ہے تو وہ سرخ نہیں ہو سکتی۔اگر ہم مردہ اور زندہ جیسے دیگر ایونٹس کو لے لیں تو یہ کہا جا سکتا ہے کہ ایک شخص یا تو زندہ یا مردہ ہو سکتا ہے۔

لیکن جھوٹ بیک وقت زندہ اور مردہ نہیں ہو سکتا۔اگر کوئی سکہ اچھالا جائے تو یا تو ہیڈ نظر آئے گا یا ٹیل نظر آئے گی۔لیکن دونوں ایک ہی وقت میں ظاہر نہیں ہو سکتے۔اس کا مطلب یہ ہے کہ ایک سکے کو اچھالنے میں، ھیڈ اور ٹیل کا ہونا باہمی طور پر خصوصی ایونٹس کے تحت آتا ہے۔

علامتی طور پر، اگر 'A' اور 'B' ایونٹس باہمی طور پر مخصوص ہیں، تو ایونٹس کے احتمال کا اندازہ P(A) یا P(B) میں لگایا جا سکتا ہے۔باہمی خصوصی ایونٹس میں، P (AB) = 0

2) آزاد اور منحصر ایونٹس:Independent and Dependent Events

دو یا زیادہ ایونٹس کو آزاد کہا جاتا ہے جب ایک ایونٹ دوسرے پر اثر انداز نہ ہو۔یہ اس حقیقت کی نشاندہی کرتا ہے کہ اگر ایک ایک کر کے ٹرائل کیے جائیں تو ایک ٹرائل دوسرے ٹرائل سے متاثر نہیں ہوتا ہے۔اور یہ بھی کہ، ایک ٹرائل کبھی بھی دوسرے ٹرائلز کے بارے میں کچھ بیان نہیں کرتا۔

مثال کے طور پر:

ایک سکے کو اچھالنے کے ایونٹس آزادانہ ایونٹس ہیں۔ اگر ایک سکے کو ایک ایک کر کے اچھالا جائے تو ایک ٹرائل دوسرے سے متاثر نہیں ہوتی۔ ٹرائل کی ساعت میں، ہیڈ یا ٹیل مخروطی ہو سکتا ہے، جو کبھی بھی یہ بیان نہیں کرتا کہ دوسرے ٹرائل میں کون سا ایونٹ پیش آئے گا۔ لہٰذا دو ہیڈی ٹرائل پہلی ٹرائل سے مکمل طور پر آزاد ہے۔

منحصر ایونٹس وہ ہوتے ہیں جن میں ٹرائل میں ایک ایونٹ کا وقوع پذیر ہونا اور نہ ہونا ایک ہی ٹرائل کی موجودگی کو متاثر کر سکتا ہے۔ یہاں، ایونٹس ایک دوسرے پر منحصر ہیں.

مثال:

اگر پلے کارڈز کے ایک پیکٹ سے کارڈ نکالا جاتا ہے اور اسے تبدیل نہیں کیا جاتا ہے، تو دوسرے ٹرائل میں احتمال تبدیل ہو جائے گا۔

3) یکساں طور پر ممکنہ ایونٹس Equally Likely Events :

ایونٹس کو یکساں احتمال کہا جاتا ہے جب وقوع پذیر ہونے کا مساوی احتمال ہو۔ اگر ایک ایونٹ دوسرے ایونٹس کی طرح واقع نہیں ہوتا ہے، تو ایونٹس کو یکساں احتمال نہیں سمجھا جاتا ہے۔ یا دوسرے لفظوں میں، ایونٹس کو یکساں احتمال کہا جاتا ہے جب ایک ایونٹ دوسرے سے زیادہ کثرت سے پیش نہیں آتا ہے۔

مثال کے طور پر:

اگر ایک غیر جانبدار سکہ یا نرد پھینکا جاتا ہے، تو توقع کی جا سکتی ہے کہ طویل عرصے میں ہر ایک کا چہرہ برابر تعداد میں ہو گا۔ ایک اور مثال میں، تاش کھیلنے کے ایک پیکٹ میں، ہم توقع کرتے ہیں کہ ہر کارڈ یکساں طور پر ظاہر ہو گا۔ اگر کوئی سکہ یا نرد متعصب ہے، تو ہر چہرے کے یکساں طور پر ظاہر ہونے کی توقع نہیں ہے۔

4) سادہ اور جامع ایونٹس Simple and Compound Events :

سادہ ایونٹس میں، ہم سادہ ایونٹس کے ہونے یا نہ ہونے کے احتمال کے بارے میں سوچتے ہیں۔ جب بھی ہم سکہ اچھالتے ہیں تو ہم ھیڈ اور ٹیل کے ایونٹس کی موجودگی پر غور کر رہے ہوتے ہیں۔ ایک اور مثال میں، اگر ایک تھیلے میں 10 سفید گیندیں اور 6 لال گیندیں ہیں، جب بھی ہم لال گیند کے ڈرائنگ کا احتمال معلوم کرنے کی کوشش کر رہے ہیں، تو یہ سادہ ایونٹس میں شامل ہے۔

مرکب ایونٹس:Compound events

لیکن دوسری طرف، جب ہم دو یا دو سے زیادہ ایونٹس کے مشترکہ وقوع پر غور کرتے ہیں، تو یہ ایک مرکب ایونٹ بن جاتا ہے۔ سادہ ایونٹس کے برعکس، یہاں ایک سے زیادہ ایونٹس کو مد نظر رکھا گیا ہے۔

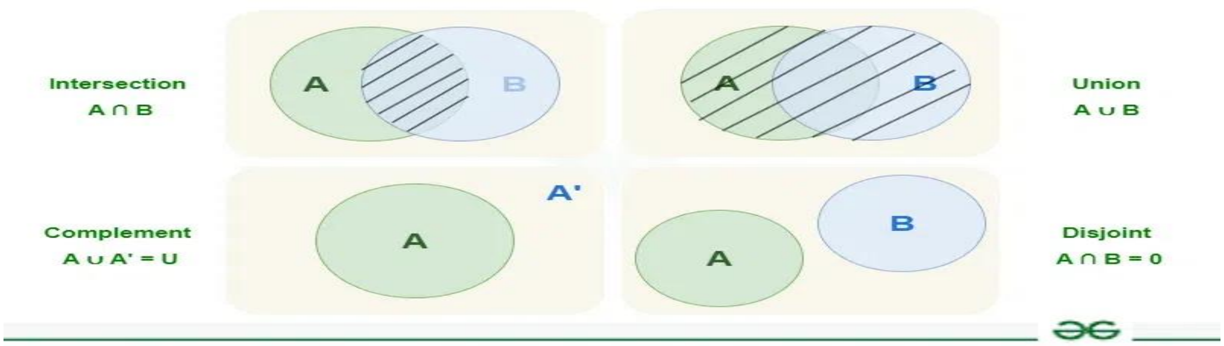
مثال کے طور پر:

اگر ایک تھیلے میں 10 سفید اور 6 لال گیندیں ہیں، اور اگر ہم 3 سفید گیندوں کا احتمال معلوم کرنے کی کوشش کر رہے ہیں تو لگاتار 3 گیندوں کی قرعہ اندازی کی جائے۔ یہ مثال اس حقیقت کو بیان کرتی ہے کہ ایونٹس کو دو سے زیادہ ایونٹسی صورتوں میں سمجھا جاتا ہے۔

6.4 احتمال کے اصول(Probability Rules)

احتمالی اصول بے ترتیب مظاہر(random phenomena) کا تجزیہ اور سمجھنے کا ایک منظم طریقہ فراہم کرتے ہیں، چاہے وہ کسی تجربے(experiment) کے نتائج کی پیشین گوئی کر رہا ہو یا مالیات میں خطرے کا اندازہ لگا رہا ہو۔ یہ بنیادی اصول، بشمول اضافہ، ضرب، اور تکمیلی اصول، ایونٹس کے وقوع پذیر ہونے کے احتمالات کا حساب لگانے کے لیے استعمال ہوتے ہیں اور بے ترتیب تجربات میں مختلف نتائج کے احتمالات کا تعین کرنے میں مدد کرتے ہیں۔

مسائل کو حل کرنے کے لیے استعمال ہونے والے احتمالات کے مختلف اصول ہیں:



Source: https://www.geeksforgeeks.org/

1) اضافے کا اصول Addition Rule

احتمال کا اضافی اصول دو ایونٹس کے اتحاد کے احتمال سے متعلق ہے۔ دو ایونٹس کے ملاپ کا احتمال انفرادی احتمالات کو شامل کرکے اور ان کے تقاطع کے احتمال کو گھٹا کر حاصل کیا جا سکتا ہے۔

اگر A اور B دو ایونٹس ہیں، تو ایونٹ A یا ایونٹ B کے رونما ہونے کا احتمال بذریعہ دیا جاتا ہے:

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- $P(A \cap B)$ represents the probability of both events A and B occurring simultaneously.
- یہ اصول اس وقت لاگو ہوتا ہے جب ایونٹس A اور B ایک دوسرے سے الگ نہیں ہوتے ہیں۔

Example: Suppose we toss a fair coin. Let A be the event of getting a head, and let the multiplication be the event of getting a tail. The probability of either getting a head or a tail is given by:

P(AUB) = P(A) + P(B)

= 1/2 + 1/2 = 1

2) ضرب کا اصول Multiplication Rule

احتمال کے ضرب کا اصول اس وقت لاگو ہوتا ہے جب ہم دو آزاد ایونٹس کے انقطاع کے احتمال کو تلاش کرنا چاہتے ہیں۔ اگر A اور B آزاد ایونٹس ہیں، تو A اور B دونوں ایونٹس کے ہونے کا احتمال بذریعہ دیا جاتا ہے:

P(A∩B) = P(A) × P(B)

یہ اصول اس وقت برقرار رہتا ہے جب ایک ایونٹ کا رونما ہونا دوسرے ایونٹ کے احتمال کو متاثر نہیں کرتا ہے۔

Example: Consider rolling a fair six-sided die. Let A be the event of rolling an even number, and let B be the event of rolling a number less than 4. Since these events are independent, the probability of rolling an even number and a number less than 4 is given by:

P(A∩B) = P(A) × P(B)

P(A) = 3/6 = 1/2

P(B) = 3/6 = 1/2

P(A∩B) = 1/2×1/2 = ¼

3) مشروط احتمال Conditional Probability

مشروط احتمال ایک ایونٹ کے پیش آنے کا احتمال ہے کہ ایک اور ایونٹ پہلے ہی واقع ہو چکا ہے۔ اگر A اور B ایونٹس ہیں، تو A دیئے گئے B کے مشروط احتمال کو P(A|B) سے ظاہر کیا جاتا ہے اور اس طرح شمار کیا جاتا ہے:

P(A|B) = P(A∩B)/P(B)

یہ اصول اس بات کی مقدار بتاتا ہے کہ کسی دوسرے واقعے کے وقوع پذیر ہونے کی روشنی میں ایک ایونٹ کا احتمال کس طرح تبدیل ہوتا ہے۔

Example: Let's continue with the example of drawing a card from a standard deck of 52 playing cards. Suppose B is the event of drawing a face card (jack, queen, or king), and A is the event of drawing a heart. The probability of drawing a heart given that the card drawn is a face card is given by:

P(A) = 13/52 = 1/4

P(B) = 12/52

P(A∩B) = 3/52

$P(A|B) = P(A \cap B)/P(B)$

$P(A|B) = (3/52)/ (12/52) = 1/4$

This means that given the card drawn is a face card, there is a 1/4 chance that it is also a heart.

4) فیصلہ کن اصول Decision Rule

اعداد و شمار میں فیصلہ کرنے کا اصول ایک رہنما خطوط ہے جو اس بات کا تعین کرتا ہے کہ مخصوص معیارات، جیسے ٹیسٹ کے اعداد و شمار یا p-values کی بنیاد پر کسی کالعدم مفروضے (Hypothesis) کو مسترد کرنا ہے یا قبول کرنا ہے۔

فیصلہ اصول ایک رسمی طریقہ کار ہے جو شماریاتی مفروضے کی جانچ میں استعمال کیا جاتا ہے یہ فیصلہ کرنے کے لیے کہ آیا کالعدم مفروضے (Null Hypothesis) کو مسترد کرنا ہے۔ کالعٹیل مفروضہ نارمل طور پر پہلے سے طے شدہ پوزیشن یا کوئی اثر یا کوئی فرق نہ ہونے کے بیان کی نمائندگی کرتا ہے۔ فیصلے کا اصول ان شرائط کا خاکہ پیش کرتا ہے جن کے تحت ایک متبادل مفروضے کے حق میں کالعٹیل مفروضے کو مسترد کر دیا جائے گا، جو بتاتا ہے کہ کوئی اثر یا فرق ہے۔

یہ کیسے کام کرتا ہے:

i. معیار طے کرنا Setting Criteria: فیصلہ کا اصول اہمیت کی سطح پر مبنی ہے (اکثر الفا، α کے طور پر بیان کیا جاتا ہے)، جو کہ کالعٹیل مفروضے کو مسترد کرنے کی حد کی وضاحت کرتا ہے۔ نارمل اہمیت کی سطحیں 0.05، 0.01، یا 0.10 ہیں۔

ii. ٹیسٹ کے اعداد و شمار Test Statistics: فیصلہ کا اصول ٹیسٹ کے اعداد و شمار کی اہم اقدار کو بیان کرتا ہے۔ مثال کے طور پر، ایک ٹیل والے ٹیسٹ میں، اگر ٹیسٹ کے اعداد و شمار ایک خاص اہم قدر سے زیادہ ہو تو null hypothesis کو مسترد کر دیا جاتا ہے۔ دو ٹیل والے ٹیسٹ میں، اگر ٹیسٹ کے اعداد و شمار دو اہم اقدار میں سے کسی ایک سے باہر ہو تو کالعٹیل مفروضے کو مسترد کر دیا جاتا ہے۔

iii. P-Values: بہت سے نارمل ڈسٹریبیوشن میں، فیصلے کے اصول میں p-value (اعداد و شمار کا مشاہدہ کرنے کا احتمال جو کہ null hypothesis سچ ہے) کا اہمیت کی سطح سے موازنہ کرنا شامل ہے۔ اگر p-value α سے کم ہے، تو null hypothesis کو مسترد کر دیا جاتا ہے۔

فیصلہ سازی کے اصول کی اقسام Types of Decision Rules

a) ڈیٹرمنسٹک رولز Deterministic Rules: یہ مخصوص ان پٹ (input) کی بنیاد پر ایک واضح نتیجہ فراہم کرتے ہیں، جیسے کہ اگر ٹیسٹ کے اعداد و شمار ایک اہم قدر سے زیادہ ہو تو کالعٹیل مفروضے کو مسترد کرنا۔

b) احتمالی اصول Probabilistic Rules: یہ غیر یقینی صورتحال کو شامل کرتا ہے اور احتمالات کی بنیاد پر نتائج فراہم کرتا ہے، جو اکثر زیادہ پیچیدہ شماریاتی ماڈلز میں استعمال ہوتے ہیں۔

فیصلے کے اصول کا استعمال Applications of Decision Rule

فیصلہ سازی کے اصول مختلف شعبوں میں ضروری ہیں، بشمول طب، مالیات اور سماجی علوم، جہاں وہ اعداد و شمار کے ثبوت کی بنیاد پر باخبر فیصلے کرنے میں محققین کی رہنمائی کرتے ہیں۔ وہ مفروضوں کی درستگی اور اعداد و شمار کے تجزیہ سے اخذ کردہ نتائج کی وشوسنیتا کا تعین کرنے میں مدد کرتے ہیں۔

خلاصہ Summary

احتمالی اصول غیر یقینی صورتحال کی مقدار درست کرنے اور بے ترتیب مظاہر کا تجزیہ کرنے کے لیے ایک منظم فریم ورک فراہم کرتے ہیں۔ ان اصولوں کو سمجھ کر اور انہیں مختلف منظر ناموں میں لاگو کرنے سے، ہم سائنس اور انجینئرنگ سے لے کر مالیات اور روزمرہ کی زندگی تک کے شعبوں میں باخبر فیصلے اور پیشین گوئیاں کر سکتے ہیں۔ فیصلے کے اصول شماریاتی تجزیہ کے لیے بنیادی ہیں، جو تجرباتی اعداد و شمار کی بنیاد پر مفروضے کی جانچ اور فیصلہ سازی کے لیے ایک منظم انداز فراہم کرتے ہیں۔

Sample Questions on Probability Rules

Question 1: You flip a fair coin twice. What is the probability of getting exactly one head?

Solution:

To solve this question, we can use the complement rule.

Probability of getting exactly one head is the complement of getting either two heads or two tails.

Probability of getting two heads: $P(HH)=1/2\times1/2=1/4$

Probability of getting two tails: $P(TT)=1/2\times1/2=1/4$

So, the probability of getting exactly one head is:

$1-P(HH)-P(TT)=1-1/4-1/4=1/2$

Question 2: A bag contains 5 red balls and 3 blue balls. You randomly select two balls from the bag without replacement. What is the probability of selecting one red ball and one blue ball?

Solution:

We can solve this question using the multiplication rule.

First, we find the probability of selecting a red ball, then the probability of selecting a blue ball after selecting a red ball.

P(red)=5/8

P (blue after red) =3/7

So, the probability of selecting one red ball and one blue ball is:

P (red and blue) = P(red)×P (blue after red)

$$= 5/8 \times 3/7 = 15/56$$

Question 3: You roll a fair six-sided die. What is the probability of rolling a number divisible by 2 or 3?

Solution:

We can solve this question using the addition rule.

We find the probability of rolling a number divisible by 2, the probability of rolling a number divisible by 3, and subtract the probability of rolling a number divisible by both 2 and 3 (i.e., 6).

Probability of rolling a number divisible by 2: P (divisible by 2) =3/6=1/2

Probability of rolling a number divisible by 3: P (divisible by 3) =2/6=1/3

Probability of rolling a number divisible by both 2 and 3: P (divisible by 6) =1/6

So, the probability of rolling a number divisible by 2 or 3 is:

P (divisible by 2 or 3) = P (divisible by 2) +P (divisible by 3) −P (divisible by 6) = 1/2+1/3−1/6 = 5/6

Question 4: A card is drawn from a standard deck of 52 playing cards. What is the probability of drawing a face card or a heart?

Solution:

We can solve this question using the addition rule.

We find the probability of drawing a face card, the probability of drawing a heart, and subtract the probability of drawing a card that is both a face card and a heart.

Probability of drawing a face card: P (face card) =12/52 (There are 3 face cards in each suit)

Probability of drawing a heart: P(heart)=13/52

Probability of drawing a card that is both a face card and a heart: P (face card and heart) =3/52

(There are 3 face cards that are hearts) So, the probability of drawing a face card or a heart is:

P (face card or heart) =P (face card) +P(heart)−P (face card and heart)

=12/52+13/52−3/52 = 22/52 =11/26

Question 5: Two dice are rolled. What is the probability of getting a sum greater than 9?

Solution:

We can solve this question by listing all possible outcomes of rolling two dice and counting the outcomes where the sum of the numbers is greater than 9.

There are 6×6 = 36 possible outcomes when rolling two dice. Outcomes with a sum greater than 9: (4, 6), (5, 5), (5, 6), (6, 4), (6, 5), (6, 6)

So, there are 6 outcomes where the sum is greater than 9. Therefore, the probability of getting a sum greater than 9 is:

P (sum > 9) =6/36=1/6

6.5 احتمالی تقسیم کے اقسام(Types of Probabilities Distributions)

احتمالی تقسیم کی اقسام Types of Probabilities Distributions

Binomialاور Poisson کی تقسیم اعداد و شمار اور ڈیٹا کے تجزیہ میں استعمال ہونے والی مجرد احتمالی تقسیم کی دو اہم اقسام ہیں۔ binomial distribution independent trials کی ایک مقررہ تعداد میں کامیابیوں کی تعداد کو ماڈل کرتی ہے، ہر ایک کی کامیابی کے یکساں احتمال کے ساتھ۔ binomial distribution بڑے پیمانے پر فنانس، بیالوجی، اور کوالٹی کنٹرول جیسے شعبوں میں استعمال ہوتی ہے۔

دوسری طرف Poisson تقسیم ماڈل وقت یا جگہ کے ایک مقررہ وقفہ میں وقوع پذیر ہونے والے ایونٹس کی تعداد کو پیش کرتا ہے جس کی موجودگی کی ایک مستقل اوسط شرح ہوتی ہے۔اس کی خصوصیت ایک واحد پیرامیٹر سے ہوتی ہے—فی یونٹ وقت کے ایونٹس کی اوسط تعداد۔ Poisson کی تقسیم اکثر شعبوں جیسے کہ وبائی امراض، فنانس اور ٹیلی کمیونیکیشن میں استعمال ہوتی ہے۔

نارمل ڈسٹری بیوشن (normal distribution)، جسے Gaussian distribution بھی کہا جاتا ہے، ایک مسلسل احتمالی تقسیم ہے جو اکثر قدرتی مظاہر جیسے کہ اونچائی اور وزن کو بیان کرنے کے لیے استعمال ہوتی ہے۔ یہ اس کی گھنٹی کے سائز کا منحنی خطوط ہے، جو متوازی ہے اور وسط کے گرد مرکز ہے۔ معیاری انحراف کروکی چوڑائی کا تعین کرتا ہے اور اعداد و شمار کے تغیر کو بیان کرتا ہے۔

1) بائنومیل ڈسٹری بیوشن Binomial Distribution

بائنومیل ڈسٹری بیوشن ایک مجرد (discrete) احتمالی تقسیم ہے جو بائنری (ہاں/نہیں) تجربے کے آزاد ٹرائلز کی ایک مقررہ تعداد میں کامیابیوں کی تعداد کو بیان کرتی ہے۔ یہ شماریات میں سب سے زیادہ استعمال ہونے والی احتمالی تقسیم میں سے ایک ہے۔

binomial distribution تجربات کی ایک سیریز سے پیدا ہوتی ہے جسے Bernoulli trials کہا جاتا ہے۔ ہر ٹرائل کے نتیجے میں کامیابی یا ناکامی ہوتی ہے، اور ہر ٹرائل میں کامیابی کا احتمال یکساں ہوتا ہے۔ تقسیم کی وضاحت دو پیرامیٹرز سے ہوتی ہے:

- n: ٹرائلوں کی تعداد۔
- p: ایک ہی ٹرائل میں کامیابی کا احتمال۔

ایک binomial random variable کا احتمالی ماس فنکشن بذریعہ دیا جاتا ہے:

$$P(X=k) = {}_{n}C_{k} * p^{k} * (1-p)^{n-k}$$

Where:

- n is the number of trials,
- k is the number of successes,
- p is the probability of success,

- nCk is the binomial coefficient, representing the number of ways to choose k successes out of n trials.
- $q = 1-p$

Example 1. suppose we flip a coin 3 times. We can use the formula above to determine the probability of obtaining 0 heads during these 3 flips:

Solution: P(X=0) = 3C0 * .50 * (1-.5)3-0 = 1 * 1 * (.5)3 = 0.125

Example 2. If a coin is tossed 3 times, then find the probability of getting exactly two heads.

X be the random variable for number of heads

The formula for the required probability is given by:

$P(X=k) = {}_{n}C_{k} * p^{k} * (1-p)^{n-k}$

Here, n = 3, k = 2, p =1/2, q = 1/2

⇒ P (X = 2) = 3C2 (1/2)2 (1/2)3-2

⇒ P (X = 2) = 3 (1/2)3

⇒ P (X = 2) = 3/8

بائنومیئل ڈسٹری بیوشن کی مثالیں Examples of Binomial Distribution

بائنومیئل ڈسٹری بیوشن کی کچھ مثالیں ذیل میں زیر بحث ہیں:

- مینوفیکچرنگ میں کوالٹی کنڑول: ایک فیکٹری لائٹ بلب تیار کرتی ہے، اور ہر بلب کے خراب ہونے کا 2% احتمال ہوتا ہے۔ اگر 100 روشنی کے بلبوں کا بے ترتیب نمونہ لیا جائے تو، ہم عیب دار بلبوں کی ایک خاص تعداد کے احتمال کو تلاش کرنے کے لیے binomial distribution کا استعمال کر سکتے ہیں۔
- کلینکل ٹرائلز: ایک نئی دوا کا تجربہ کیا جا رہا ہے، اور یہ 70% منورملات میں مؤثر معلوم ہوتی ہے۔ کلینیکل ٹرائل میں، اس دوا سے 10 مریضوں کا علاج کیا جاتا ہے۔
- کھیلوں کے اعداد و شمار: باسکٹ بال کے کھیل میں، ایک کھلاڑی کی کامیابی کی شرح 80% ہے۔ ایک کھیل کے دوران، کھلاڑی 15 فری تھرو لیتا ہے۔

2) پوائسن ڈسٹری بیوشن (Poisson Distribution)

پوائسن ڈسٹری بیوشن ایک مجرد احتمالی تقسیم ہے جو وقت یا جگہ کے ایک مقررہ وقفہ میں وقوع پذیر ہونے والے ایونٹس کی دی گئی تعداد کے

احتمال کو ظاہر کرتی ہے، بشر طیکہ یہ ایونٹس معلوم مستقل اوسط شرح کے ساتھ اور آخری ایونٹ کے بعد سے آزادانہ طور پر رونما ہوں۔ اس کا نام فرانسیسی ریاضی دان Siméon Denis Poisson کے نام پر رکھا گیا ہے۔

Poisson تجربہ درج ذیل صفات کے ساتھ شماریاتی تجربہ ہے:

- تجربے کے نتائج کو کامیابیوں یا ناکامیوں کے طور پر درجہ بندی کیا جاتا ہے۔
- کسی خاص علاقے میں کامیابیوں کی اوسط تعداد کو μ سے ظاہر کیا جاتا ہے، اور یہ جانا جاتا ہے۔
- کسی تقریب میں کامیابی کا احتمال خطے کے سائز کے متناسب ہے۔
- ایک انتہائی چھوٹے خطے میں کسی تقریب میں کامیابی کا احتمال تقریباً صفر ہے۔
- ایک مخصوص خطہ کئی شکلیں لے سکتا ہے، جیسے لمبائی، رقبہ، حجم، وقت کی مدت وغیرہ۔

پوسن کے احتمال کو شمار کرنے کے لیے استعمال ہونے والا فارمولا، ایک بار جب کسی مخصوص علاقے میں کامیابیوں کی اوسط تعداد (μ) دستیاب ہو جاتی ہے، مندرجہ ذیل دیا جاتا ہے:

$$P(x; \mu) = \frac{(e^{-\mu})(\mu^x)}{x!}$$

- e: It is a constant that is approximately equal to 2.71828. In fact, e is the base of the natural logarithm system.
- μ: It is the average number of successes occurring in a specified region.
- x: It is the exact number of successes occurring in a specified region.
- P (x; μ): It is the Poisson probability of exact x successes occurring in a Poisson experiment. Here the mean number of successes is μ

Example 1: A call center receives an average of 3 calls per minute. What is the probability that exactly 4 calls will be received in a given minute?

Solution: Given:

$\mu = 3$, $x = 4$

P (4; 3) = (e^-3 * 3^4) / 4!

= (0.0498 * 81) / 24

= 4.0338 / 24

≈ 0.1681 or 16.81%.

Example 2: A website gets an average of 2 user sign-ups per hour. What is the probability that in one hour, there will be no sign-ups?

Solution: Given:

μ = 2,   x = 0

P (0; 2) = (e^-2 * 2^0) / 0!

= (0.1353 * 1) / 1

= 0.1353 or 13.53%.

پوائسن ڈسٹری بیوشن کی مثالیں Examples of Poisson Distribution

Poissonڈسٹری بیوشن کا استعمال کرتے ہوئے ماڈل کی جاسکتی ہے کہ کچھ سب سے زیادہ نارمل مثالیں ہیں:

- کال سینٹر Call Center: ایک کال سینٹر اوسطاً 5 کالیں فی منٹ وصول کرتا ہے۔ ہم ایک منٹ میں ایک مخصوص تعداد میں کال موصول ہونے کا احتمال تلاش کرنے کے لیے Poisson کی تقسیم کا استعمال کر سکتے ہیں۔
- ٹریفک کا بہاؤ Traffic Flow: اوسطاً، ہر 10 منٹ میں 2 کاریں ایک چیک پوائنٹ سے گزرتی ہیں۔ ہم 10 منٹ میں چوکی سے گزرنے والی کاروں کی ایک خاص تعداد کا احتمال تلاش کرنے کے لیے پوسن کی تقسیم کا استعمال کر سکتے ہیں۔
- ویب ٹریفک Web Traffic: ایک ویب سائٹ فی گھنٹہ اوسطاً 10 وزٹ حاصل کرتی ہے۔ ہم ایک گھنٹے میں مخصوص تعداد میں وزٹ حاصل کرنے کا احتمال تلاش کرنے کے لیے Poisson کی تقسیم کا استعمال کر سکتے ہیں۔

3) نارمل ڈسٹری بیوشن یا گاسی ڈسٹری بیوشن یا بیل کریو Normal Distribution or Gaussian Distribution or Bell Curve

احتمالی نظریہ میں، نارمل ڈسٹری بیوشن یا گاوسی تقسیم ایک بہت ہی نارمل مسلسل احتمالی تقسیم ہے۔ نارمل ڈسٹری بیوشن کو بعض اوقات غیر رسمی طور پر گھنٹی کرو کہا جاتا ہے۔ ایک نارمل مساوات مسلسل احتمالی تقسیم کے خاندان کو بیان کرتی ہے۔ اس کا استعمال حقیقی قدر والے بے ترتیب متغیرات کو پیش کرنے کے لیے کیا جاتا ہے جن کی تقسیم معلوم نہیں ہے۔ نارمل ڈسٹری بیوشن میں استعمال ہونے والا فارمولا درج ذیل ہے:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$f(x)$ = probability density function

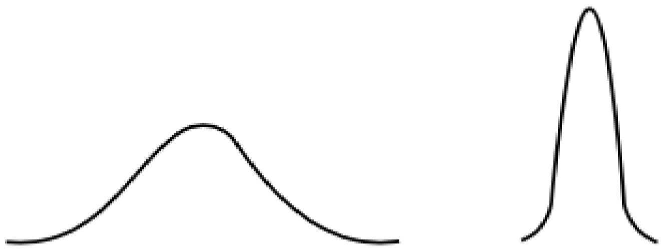
σ = standard deviation

μ = mean

Here, μ is the mean of a normal distribution, x is a normal random variable, σ is the standard deviation of normal distribution, π = 3.14159, and e = 2.71828.

ایک نارمل ڈسٹری بیوشن کو گرافی طور پر پیش کیا جا سکتا ہے اور گراف کو متاثر کرنے والے دو عوامل اوسط(average) اور معیاری انحراف(standard deviation) ہیں۔ گراف کے مرکز کے مقام کا تعین تقسیم کے وسط سے ہوتا ہے، اور گراف کی چوڑائی اور اونچائی کا تعین معیاری انحراف سے ہوتا ہے۔

نارمل ڈسٹری بیوشن کا کرو بڑے معیاری انحراف کی صورت میں چوڑا اور چھوٹا ہوتا ہے۔ تاہم، ایک چھوٹے معیاری انحراف کی صورت میں، نارمل ڈسٹری بیوشن کا کرو تنگ اور لمبا ہوتا ہے۔ نارمل ڈسٹری بیوشن کا کرو گھنٹی کی شکل کا ہوتا ہے اور تمام نارمل ڈسٹری بیوشن ہم آہنگ نظر آتی ہیں۔



اوپر کی تصویر میں، دیکھیں کہ بائیں طرف کی تصویر دائیں طرف کی تنگ اور لمبے گھماؤ کے مقابلے میں چھوٹی اور چوڑی ہے۔ ایسا اس لیے ہے کہ دائیں جانب کی کریو میں ایک چھوٹا معیاری انحراف ہے

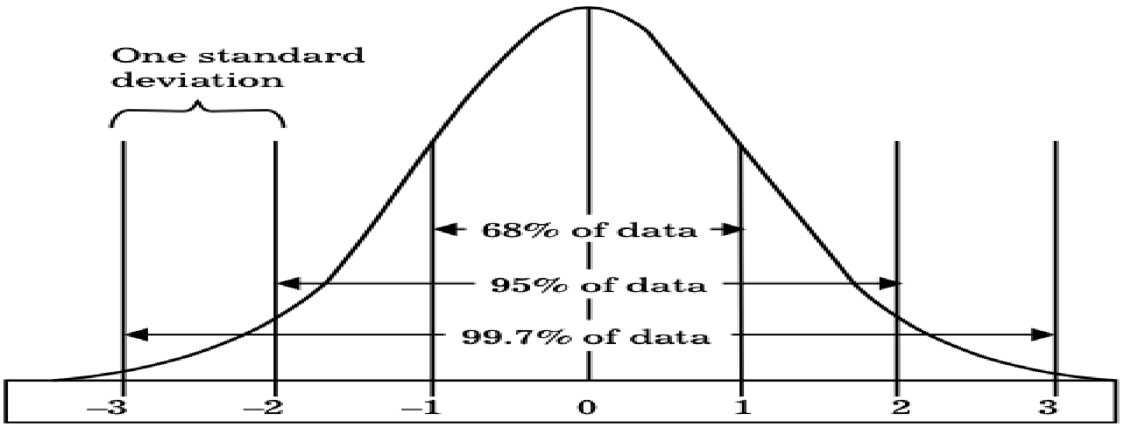
نارمل ڈسٹری بیوشن کرو کی خصوصیات درج ذیل ہیں The Properties of a normal distribution :curves are as follows

- نارمل کرو کے تحت کل رقبہ 1 ہے۔
- اوسط، موڈ، اور میڈین سب برابر ہیں۔
- نارمل کرو وسط کے ارد گرد سڈول ہے، μ۔
- تمام نارمل منحنی خطوط اوسط μ کے بارے میں ہم آہنگ ہیں، σ ± μ پر موڑنے کا ایک نقطہ ہے، اور گھنٹی کے سائز کے ہوتے ہیں۔
- x کی تمام اقدار کے لیے، تمام نارمل منحنی خطوط مثبت ہیں۔ یعنی 0 <f(x) تمام x کے لیے۔
- f (x) کی حد جب x لا محدودیت کے قریب پہنچتی ہے تو 0 ہے، اور f (x) کی حد جب x منفی لا محدودیت کے قریب پہنچتی ہے بھی 0 ہے۔
- کسی بھی نارمل کرو کی زیادہ سے زیادہ اونچائی μ = x ہے۔
- ایک نلارمل منحنی شکل کا انحصار اس کے وسط، یعنی μ، اور اس کے معیاری انحراف σ پر ہوتا ہے۔

پہلے ذکر کی گئی خصوصیات کے علاوہ، ہر نارمل کرو (مطلب یا معیاری انحراف سے قطع نظر) 68-95-99.7 اصول کے مطابق ہے۔ یہ ایک اصول ہے جو نارمل ڈسٹری بیوشن میں استعمال کیا جاتا ہے فیصد کی قدروں کو یاد رکھنے کے لیے جو اوسط کے ارد گرد دو، چار، اور چھ معیاری انحراف کی حد میں ہوتی ہیں۔ اصول سے متعلق درست قدریں 68.27%، 95.45%، اور 99.73% ہیں۔ اس اصول کے مطابق:

- نارمل منحنی خطوط کے تحت رقبہ کا 68% اوسط $(x \pm 1\alpha)$ کے دو معیاری انحراف کے اندر آتا ہے۔
- نارمل منحنی خطوط کے تحت تقریباً 95% رقبہ اوسط کے چار معیاری انحراف $(x \pm 2\alpha)$ کے اندر آتا ہے۔
- نارمل منحنی خطوط کے تحت تقریباً 99.7% رقبہ اوسط کے چھ معیاری انحراف کے اندر آتا ہے $(x \pm 3\alpha)$۔

معیاری نارمل ڈسٹری بیوشن کا اوسط 1 اور معیاری انحراف 1 ہوتا ہے۔ 68-95-99.7 رول کے ساتھ ایک معیاری نارمل ڈسٹری بیوشن کو دکھایا گیا ہے۔



Example: Find the value of the probability density function at x = 10, given that the distribution has a mean μ = 8 and a standard deviation σ = 2.

Solution: (1

$$f(10) = (1 / (2\sqrt{2\pi})) * e^{\wedge} (-1/2 * ((10 - 8)/2) ^{\wedge}2)$$
$$= (1 / (2\sqrt{2\pi})) * e^{\wedge} (-1/2 * (1) ^{\wedge}2)$$
$$= (1 / (2 * 2.5066)) * e^{\wedge} (-0.5)$$
$$= (1 / 5.0132) * 0.6065$$
$$\approx 0.1210$$

Answer: $f(10) \approx 0.1210$

6.7 اکتسابی نتائج (Learning Outcome)

طلباء نے احتمال کے معنی اور تعریف اور کاروبار اور شماریاتی تجزیہ میں اس کے کردار کی ٹھوس سمجھ حاصل کر لی ہے۔ انہوں نے احتمال کی اہم

اصطلاحات کو سیکھ لیا ہے، جس سے وہ تصورات کو واضح اور درست طریقے سے بات چیت کرنے کے قابل بناتے ہیں۔ طلباء اب بنیادی مسائل کو حل کرنے اور واقعات کے امکانات کا حساب لگانے کے لیے احتمالی اصولوں کا اطلاق کر سکتے ہیں۔ انہوں نے امکانات کی تقسیم کی مختلف اقسام کو بھی دریافت کیا ہے، ان کی خصوصیات اور حقیقی دنیا کے ڈیٹا کی ماڈلنگ میں ایپلی کیشنز کو سمجھنا ہے۔ مجموعی طور پر، سیکھنے والے غیر یقینی صورتحال کا تجزیہ کرنے، پیشین گوئیاں کرنے، اور ڈیٹا پر مبنی فیصلہ سازی کی حمایت کرنے کے لیے احتمالی تصورات اور تقسیم کو استعمال کرنے کے لیے لیس ہیں۔

6.8 نمونہ امتحانی سوالات((Model Examination Questions)

درج ذیل بیان کو صحیح یا غلط بتائیں:

1) احتمالی تقسیم میں تمام احتمالات کا مجموعہ ہمیشہ 1 کے برابر ہوتا ہے۔ صحیح
2) بائنومیئل ڈسٹری بیوشن میں، ٹرائلوں کی تعداد لا محدود ہونی چاہیے۔ غلط
3) ایک نارمل ڈسٹری بیوشن اس کے اوسط کے بارے میں ہم آہنگ ہے۔ صحیح
4) Poisson ڈسٹری بیوشن کا استعمال اس وقت کیا جاتا ہے جب ٹرائلوں کی تعداد مقرر ہو اور ہر ٹرائل کے دو نتائج ہوتے ہیں۔ غلط
5) ایک ناممکن ایونٹ کا احتمال صفر ہے۔۔ صحیح

مختصر جوابات کے حامل سوالات(Short Answer Type Questions)

1) احتمال کی وضاحت کریں اور اس کی پیائش کے لیے کسی بھی دو طریقوں کی وضاحت کریں۔
2) نارمل ڈسٹری بیوشن کی کوئی بھی تین خصوصیات بیان کریں۔
3) بائنومیئل ڈسٹری بیوشن کے مفروضے کیا ہیں؟
4) احتمالی تقسیم میں متوقع قدر کے تصور کی وضاحت کریں۔

طویل جوابات کے حامل سوالات(Long Answer Type Questions)

1) Poisson ڈسٹری بیوشن کو اس کے فارمولے، خصوصیات اور حل شدہ دو مثالوں کے ساتھ سمجھائیں۔
2) نارمل ڈسٹری بیوشن کو تفصیل سے بیان کریں۔ اس کے علاوہ، شماریات میں اس کی اہمیت پر بحث کریں۔

احتمالات پر مبنی سوالات(Questions based on Probability):

1) ایک بیگ میں 5 لال گیندیں، 3 سبز گیندیں، اور 2 نیلی گیندیں ہیں۔ بیگ سے لال گیند یا سبز گیند نکالنے کا کیا احتمال ہے؟
2) ایک منصفانہ سکہ دو بار پھینکا جاتا ہے۔ لگاتار دو ھیڈ ملنے کا کیا احتمال ہے؟

3) تاش کے ایک ڈیک میں 52 کارڈ ہوتے ہیں، جن میں سے 13 دل ہوتے ہیں۔ڈیک سے دل نہ نکالنے کا کیا احتمال ہے؟

4) 52 تاشوں کا ایک معیاری ڈیک بدلا ہوا ہے۔اگر تیار کردہ کارڈ ایک چہرہ کارڈ ہے (جیک، ملکہ، یا بادشاہ)، تو اس بات کا کیا احتمال ہے کہ کارڈ بھی ایک سپیڈ ہے؟

Answer: - *1. 0.8,* *2. 0.25,* *3. 0.75,* *4. 0.25*

اکائی 7۔ کوریلیشن کا تجزیہ

(Correlation Analysis)

اکائی کے اجزاء

7.0 تمہید — Introduction

7.1 مقاصد — Objectives

7.2 کوریلیشن کے معنی، تعریف اور نوعیت — Meaning, Definition and Nature of Correlation

7.3 کوریلیشن اور عناد تحریک — Correlation and Causation

7.4 کوریلیشن کی ضرورت — Need of Correlation

7.5 کوریلیشن کی اقسام — Types of Correlation

7.6 کارل پیرسن کا کوفیشنٹ آف کوریلیشن کی نوعیت

Nature of Karl Pearson of Coefficient of Correlation

7.7 پیرسن کے کوریلیشن کے مفروضات

Assumptions of Pearson's Coefficient of Correlation

7.8 کوریلیشن کے کوفیشنٹ کی تشریح — Interpretation of Coefficient of Correlation

7.9 کوریلیشن کا کوفیشنٹ اور احتمالی خطا — Coefficient of Correlation and Probable Error

7.10 درجاتی کوریلیشن کے معنی و تعریف — Meaning and Definition of Rank Correlation

7.11 اسپیرمین کوریلیشن کے کوفیشنٹ کے خدوخال

Features of Spearman's Coefficient of Correlation

7.12 رینک کوریلیشن کے فائدے اور نقصانات

Advantages and Disadvantages of Rank Correlation

7.13 رینک کوریلیشن کے کوفیشنٹ کے استعمالات — Uses of Coefficient of Correlation

7.14 اکتسابی نتائج

7.15 نمونہ امتحانی سوالات — Model Examination Questions

7.0 تمہید(Introduction)

شماریات میں اب تک ہم نے صرف ایک متغیر سے متعلق سوالات کا مطالعہ کیا ہے لیکن عملی طور پر ہم ایسے کئی بے شمار مسائل کا سامنا کرتے ہیں جس میں دو یا زائد متغیرات کا استعمال ہوتا ہے۔ اگر دو مقداریں کچھ اس انداز میں حرکت کر رہی ہوں کہ ایک مقدار کی حرکت دوسری مقدار کے برابر، مساوی ہوں، کچھ فرق نہ ہو تب یہ مقداریں ہم رشتہ، ہم آہنگ ہوں گی۔ مثال کے طور پر اگر شوہر اور بیوی کی عمر کے درمیان، کسی شئے کی قیمت اور طلب ایک نکتہ تک بارش میں اضافہ اور چاول کی پیداوار ٹیلی ویژن لائسنس اور سینما جانے والوں کی تعداد وغیرہ کے درمیان تعلق پایا جائے، زیر غور متغیرات کے درمیان تعلق کے دائرہ کی پیائش کوریلیشن کے تجزیہ کے ذریعہ ہوگی۔ کوریلیشن کی پیائش کو کوریلیشن کا کوفیشنٹ یا کوریلیشن اعشاریہ کہتے ہیں جس کو ایک عدد کی صورت میں رخ اور زاویہ کو ظاہر کیا جاتا ہے۔ کوریلیشن تجزیہ متغیرات کے درمیان تعلق قربت کی پیائش میں مستعمل ٹیکنکس سے متعلق ہے۔

7.1 مقاصد(Objectives)

اس اکائی کے مطالعہ کے بعد آپ اس قابل ہو جائیں گے کہ :

- کوریلیشن کے مطالعہ کی اہمیت سے واقف ہو سکیں گے۔
- ہم رشتگی کے مفہوم سے واقف ہو سکیں گے۔
- کوریلیشن کے مطالعہ کے اسباب و علل سے واقف ہو سکیں گے۔
- کوریلیشن کے اقسام پر بحث کر سکیں گے۔
- کوریلیشن کے مطالعہ کے طریقوں سے آگاہ ہو سکیں گے۔

7.2 کوریلیشن کے معنی، تعریف اور نوعیت

(Meaning, Definition and Nature of Correlation)

متغیرات کے درمیان کوریلیشن کی نشاندہی کی جاتی ہے کہ ایک متغیر کی قدر میں تبدیلی ہو رہی ہو تب دوسرا متغیر بھی اسی خصوصی سمت میں تبدیل ہوتا ہے۔ شماریات میں دو متغیرہ تقسیم میں ہم یہ دریافت کرنے میں دلچسپی رکھتے ہیں کہ زیر مطالعہ دو متغیروں کے درمیان کوئی تعلق ہے۔ کوریلیشن ایک شماریاتی آلہ ہے جو دو متغیروں کے درمیان تعلق کا مطالعہ کرتا ہے اور کوریلیشن تجزیہ میں دو متغیروں کے درمیان کس حد تک تعلق ہے اس مطالعہ کے لئے مختلف طریقہ اور ٹیکنکس کا استعمال شامل ہے دو متغیروں کو اسی وقت ہم رشتہ سمجھا جاتا ہے جب کہ ایک متغیرے میں تبدیلی کے نتیجے میں دوسرے متغیرے میں تبدیلی ہو۔ اس طرح کوریلیشن شماریاتی آلہ ہے جو دو یا زائد متغیروں کے درمیان تجزیہ میں ہماری مدد کرتا ہے۔

بعض اوقات ایسا محسوس ہوتا ہے کہ محصلہ مختلف متغیرات کی قدریں باہم تعلق رکھتے ہیں۔ایسا ممکن ہے کہ ایسا تعلق ہم دو سلسلوں میں حاصل کر سکتے ہیں جیسے ایک گروہ کا قد اور وزن۔ایسا مشاہدہ کیا جا سکتا ہے قد میں اضافہ کے ساتھ وزن میں اضافہ ہوتا ہے۔ نتیجتاً بلند قامت افراد کم قامت افراد سے زیادہ وزنی ہوتے ہیں۔اسی طرح اگر ایک شئے کی قیمت اور مختلف قیمتوں پر فروخت کردہ مقدار ان دونوں کی معلومات کو جمع کریں تو دو سلسلے حاصل ہوں گے۔ایک متغیرہ شئے کی مختلف قیمتوں اور دوسرا متغیرہ ان قیمتوں پر فروخت کردہ مقدار ہوگی۔ نتیجتاً ہم ایسے دو سلسلوں میں کچھ تعلق ضرور پائیں گے۔ شئے کی قیمتوں میں اضافہ کے ساتھ شئے کی فروخت کردہ مقدار گھٹتی جاتی ہے۔اس طرح ہم یہ نتیجہ اخذ کر سکتے ہیں کہ قیمت اور طلب میں کچھ تعلق ہے ایسا تعلق کئی سلسلوں میں پایا جا سکتا ہے مثال کے طور پر قیمت اور رسد۔ افراد کا قد اور وزن۔ نیشکر اور شکر کی قیمتیں۔شوہر اور بیویوں کی عمریں وغیرہ۔

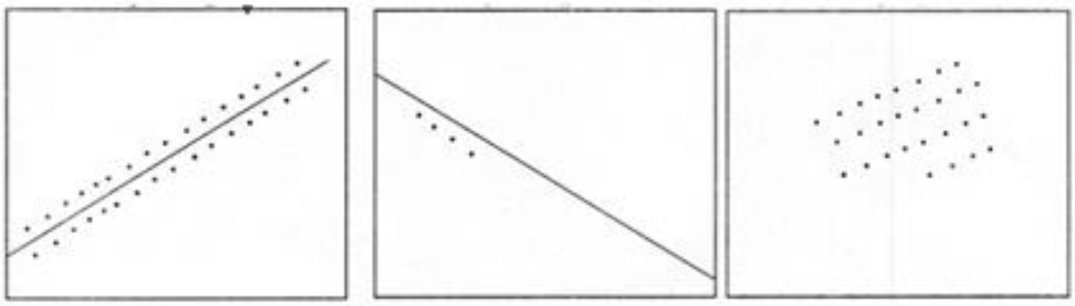
7.2.1 کوریلیشن کا کوفیشنٹ(Coefficient of Correlation)

کوریلیشن کے کوفیشنٹ کی تحسیب دو متغیروں کے درمیان تعلق کے حدود کے مطالعہ کے لئے کی جاتی ہے۔ جیسا کہ پہلے کہا جا چکا ہے یہ حقیقت ہے کہ دو متغیروں کے درمیان کوریلیشن ہے اس کا یہ مطلب نہیں ہوتا کہ یہ تعلق مستقل یا فعال ہے۔اگر متغیر کی قدر معلوم ہو یہ ممکن نہ ہوگا کہ دوسرے متغیر کی قدر ہو بہو حاصل ہو۔ ایسا صرف اسی وقت کیا جا سکتا ہے جب کہ دو متغیروں کے درمیان خطی تعلق(Linear relationship) پایا جائے۔کچھ سلسلے ہیں جن میں خطی تعلق(Linear relationship) موجود ہوتا ہے جیسے حقیقی اعداد اور ان کے مربعے یا جذر المربع ہمیشہ خطی تعلق پیش کرتے ہیں۔اسی طرح خطی تعلق دو سلسلوں میں پایا جا سکتا ہے۔ایک مختلف دائروں کے قطر سے متعلق دوسرا ان کے رقبوں سے متعلق۔ معاشی اعداد و شمار میں ایسے تعلقات کبھی کبھی پائے جاتے ہیں۔اس میں کوئی شک نہیں کہ قیمتوں میں اضافہ سے طلب میں کمی واقع ہوگی لیکن تعلق فعال نہیں ہے۔ قیمت اور طلب سے متعلق دو سلسلوں میں تغیر کے درمیان کوئی مستقل تناسب ،نسبت نہیں پائی جاتی۔

مکمل کوریلیشن(Perfect correlation):اگر دو متغیروں کے درمیان تعلق ایسا ہو کہ ایک کی قدر میں اضافہ سے دوسرے کی قدر میں اضافہ یا کمی قائم تناسب میں ہو تب ان کے درمیان کوریلیشن مکمل(Perfect correlation) ہوگا۔اگر دونوں سلسلے ایک ہی سمت میں حرکت پذیر ہوں اور تغیرات ،تبدیلی متناسب ہو ان کے درمیان مکمل مثبت کوریلیشن (Perfect positive correlation) ہوگا۔اگر اس کے برخلاف دو سلسلے معکوس ،مخالف سمت میں حرکت پذیر ہوں اور ان کی قدروں کے درمیان تغیر ہمیشہ متناسب ہو یہ مکمل منفی کوریلیشن(Perfect negative correlation) کی مثال ہوگی۔ایسا بھی ہو سکتا ہے کہ دو سلسلوں کے درمیان تغیر میں کوئی تعلق نہ ہو ایسی صورت میں ان کے درمیان کوئی کوریلیشن نہ ہوگا۔

جیسا کہ پہلے کہا جا چکا ہے معاشی اعداد و شمار میں مکمل مثبت یا منفی کوریلیشن عموماً نہیں پایا جاتا چونکہ معاشی سلسلوں کے درمیان تعلق کبھی کبھار فعال ہوتا ہے ایسے اعداد و شمار میں چونکہ متعلقہ سلسلے مکمل طور پر ایک دوسرے پر انحصار نہیں کرتے اس لئے کوریلیشن مکمل نہیں ہوتا۔ پرفیکٹ کوریلیشن اس وقت حاصل ہوتا ہے جب کہ دو سلسلوں کے درمیان مکمل باہمی انحصار پایا جاتا ہو۔

Perfect Negative Correlation　　Perfect Positive Correlation



(i)　(ii)　(iii)

Source:https://www.simplypsychology.org/correlation.html

مذکورہ بالا گراف سے یہ مشاہدہ کیا جاسکتا ہے کہ x اور y کی قیمتوں کے مقابل قدریں، قیمتیں خط مستقیم میں پائی گئی ہیں۔ شکل نمبر(i) میں x اور y کے درمیان پرفیکٹ پازیٹیو کوریلیشن کی نشاندہی کی گئی ہے چونکہ دو سلسلوں کی قیمتوں میں تغیر ہمیشہ قائم متناسب ہے اور وہ ایک ہی سمت میں حرکت پذیر ہیں۔ اس کے برخلاف شکل نمبر(ii) x اور y کے درمیان پرفیکٹ نیگیٹیو کوریلیشن کی نشاندہی کی گئی ہے چونکہ ان کی قیمتوں کے درمیان تغیرات قائم متناسب ہیں اور دونوں معکوس، مخالف سمت میں حرکت پذیر ہیں۔

یہ معلوم ہونے کے بعد یہ ضروری ہے کہ کوریلیشن کی ایسی پیمائش حاصل کی جائے جو مقداری اصطلاح میں کوریلیشن کی ڈگری کی صحیح نشاندہی کر سکے۔ اس کی پیمائش ایسی ہو کہ اس کی انتہائی قیمتیں پرفیکٹ پازیٹیو یا پرفیکٹ نیگیٹیو کوریلیشن کی نمائندگی کرے اور درمیان میں کوئی قیمت ہو تو کوریلیشن غیر موجود ہو۔ ایسی پیمائش کوریلیشن کے کوفیشنٹ نے دی ہے۔

اگلے صفحات میں ہم کوریلیشن کے کوفیشنٹ پر بحث کرنے جارہے ہیں۔ یہ بحث ہمیشہ مثبت کوریلیشن اور منفی کوریلیشن یعنی 1+ اور -1 کے درمیان پھیلی ہوئی ہوگی۔ جب پرفیکٹ کوریلیشن ہوگا اس کی قیمت 1+ اور جب پرفیکٹ نیگیٹیو کوریلیشن ہوگا تب اس کی قیمت 1- ہوگی۔ اس کا درمیانی نکتہ '0' ہے جو کوریلیشن کی غیر موجودگی کو ظاہر کردے گا۔ جیسا کوریلیشن کے کوفیشنٹ کی قیمت گھٹتی جائے گی دو متغیروں کے درمیان پازیٹیو کوریلیشن کی بھی قیمت گھٹتی جائے گی جب یہ قیمت صفر تک پہنچتی ہے یہ اس بات کی نشاندہی کرتی ہے کہ ان کے درمیان کوریلیشن نہیں ہے۔ اور اگر مزید گھٹتی ہے تو منفی قیمتیں حاصل ہوتی ہیں یعنی صفر سے کم۔ یہ نیگیٹیو کوریلیشن کی نشاندہی کرتی ہیں جب یہ قیمت -1 تک پہنچتی ہیں تب یہ دو متغیروں کے درمیان پرفیکٹ نیگیٹیو کوریلیشن ہونے کا ثبوت ہے۔

مذکورہ بالا بیان کردہ نکات کا مطالعہ قبل ازیں دیے گئے گراف سے کیا جاسکتا ہے۔ جب متغیرہ کی قیمتیں شکل نمبر(i) میں دی گئی قیمتوں کے مماثل ہوں تب پرفیکٹ پازیٹیو کوریلیشن ہوگا یا کوریلیشن کا کوفیشنٹ 1+ ہے لیکن جب وہ شکل نمبر 1 میں دئے گئے قیمتوں کے مماثل ہوں تب بھی وہ پازیٹیو کوریلیشن ہے لیکن مکمل (پرفیکٹ) نہ ہوگا یا کوریلیشن کی قیمت 1+ سے کم ہوگی لیکن 0 سے زیادہ ہوگی۔ جب قیمتیں شکل نمبر 3 کے مماثل ہو تب اعداد و شمار کے درمیان کوئی کوریلیشن نہ ہوگا یا کوریلیشن کے کوفیشنٹ کی قیمت 0 ہوگی جب قیمتیں شکل نمبر(ii) کے مماثل ہو تب نیگیٹیو کوریلیشن ہوگا اگرچہ کہ پرفیکٹ نہ ہو۔ اس کا مطلب یہ ہوگا کہ کوریلیشن کی قیمت 0 سے زیادہ ہوگی (منفی

جانب) لیکن 1- سے کم ہو گی۔ تاہم اگر قیمتیں متغیر ہوں جیسے شکل نمبر 6 میں دیا گیا ہے یہاں پر فیکٹ نیگیٹیو کوریلیشن ہے یا دوسرے الفاظ میں کوریلیشن کے کوفیشنٹ کی قیمت 1- ہو گی۔

7.2.2 کوریلیشن کے دائرہ کار (Scope of Correlation)

کوریلیشن (کوریلیشن) کے مطالعہ کی افادیت طبیعیات کے ساتھ ساتھ سوشل سائنس اور دونوں میں بے پایاں افادیت کا حامل ہے تاہم، ہم اپنے آپ کو کوریلیشن کے مطالعہ کی افادیت کو صرف سوشل سائنس تک محدود رکھیں گے۔

1. کوریلیشن کا مطالعہ فیصلہ سازی سے جڑی غیر یقینی کیفیت کی حد کو کم کرتا ہے۔ س سوشل سائنس میں خصوصاً کاروباری دنیا میں پیش بینی ایک اہم مظہر ہے اور کوریلیشن کا مطالعہ ہمیں پیش بینی پر زیادہ بھروسہ کرنے میں مددگار ہے۔
2. معاشی برتاؤ کو سمجھنے میں کوریلیشن تجزیہ بہت ہی مددگار ہے۔ یہ ہمیں ایسے متغیرات کی نشاندہی میں مدد کرتا ہے جن پر دوسرے متغیرات کا انحصار ہوتا ہے۔ یہ ان عناصر کے مطالعہ میں مددگار ہے جن سے معاشی ایونٹس متاثر ہوتے ہیں مثال کے طور پر ہم ان عناصر کو معلوم کر سکتے ہیں جو قیمتوں میں اضافہ یا کم پیداواریت کے ذمہ دار ہیں۔
3. کوریلیشن کا مطالعہ ہمیں ایسے عناصر کی نشاندہی میں مددگار ہے جو متاثرہ معاشی حالات کو مستحکم کر سکے۔
4. کوریلیشن کا مطالعہ ہمیں ایک متغیرہ میں متوقع تبدیلی کا تخمینہ کرنے میں مددگار ہے جو اس سے متعلقہ متغیرہ میں یک مخصوص مقدار میں تبدیلی کے ساتھ ہوتا ہے۔ مثال کے طور پر کوریلیشن کا مطالعہ ہمیں قیمتوں میں تبدیلی سے طلب میں تبدیلی کو معلوم کرنے میں مددگار ہے۔
5. مختلف متغیرات کے درمیان بین تعلقاتی مطالعہ تحقیق کے فروغ اور علم کے نئے ابواب کے آغاز میں ایک بہت ہی مددگار آلہ ہے۔

اس طرح کوریلیشن کے مطالعات مختلف النوع مقاصد کے لئے بہت زیادہ مستعمل اور تشریح کے لئے ایک بنیادی آلہ تصور کیا جاتا ہے۔

7.3 کوریلیشن اور عناد تحریک (Correlation and Causation)

اگرچہ کہ لفظ کوریلیشن دو یا زیادہ متغیرات کے باہمی انحصار کے پس منظر میں استعمال کیا جاتا ہے اس کے باوجود یہ ہر گز ضروری نہیں کہ وجہ اور اثر کے تعلق کی نشاندہی ہو جائے۔ درج ذیل کسی ایک یا زیادہ وجوہات کی بنا دو متغیروں کے درمیان کوریلیشن پائی جا سکتی ہے۔

1. دو متغیروں کے درمیان وجہ اور اثر کا تعلق ہوتا ہے مثال کے طور پر گرمی اور درجہ حرارت میں کوریلیشن ہو سکتی ہو گی کیونکہ گرمی درجہ حرارت پر اثرانداز ہوتی ہے۔ ان کے درمیان وجہ اور اثر کا تعلق ہو سکتا ہے تاہم س سوشل سائنس میں ایسا تعلق شاذ و نادر ہی پایا جاتا ہے' اسکی وجہ یہ ہے کہ سوشل سائنس میں بکثرت عناد تحریک ہوتی ہے ساتھ ہی ساتھ ایک متغیر کئی عناصر سے متاثر ہو سکتا ہے جس کے انفرادی اثرات کو الگ نہیں کیا جا سکتا۔ مثال کے طور پر طلب افراط زر' برآمدی پالیسی اور دیگر کئی ایک دوسرے عناصر ہیں تبدیلی کی وجہ سے قیمتوں

میں اضافہ ہو سکتا ہے۔ یہ معلوم کرنا تقریباً ناممکن ہوتا ہے کہ قیمتوں میں اضافہ کونسے عنصر کی وجہ سے ہوتا ہے۔اس طرح سوشل سائنسیس میں کوریلیشن اتفاقاً ہی وجہ اور اثر تعلق کی نشاندہی کرتا ہے۔

2. دونوں ہم رشتہ Correlated متغیرات تیسرے متغیرہ یا ایک سے زیادہ متغیرات سے متاثر ہوتے ہیں مثال کے طور پر چاول کی قیمتوں اور جوٹ (پٹ سن) کی قیمتوں کے درمیان زیادہ کوریلیشن پائیں گے۔ حقیقت میں یہ معلوم ہوگا کہ دونوں اشیاء کی قیمتیں ان کی پیداوار سے متاثر ہوئی ہو اور پیداوار بارش سے متاثر ہوئی ہو۔ ہو سکتا ہے کہ چاول کی قیمتوں اور جوٹ کی قیمتوں کے درمیان کوئی حقیقی تعلق نہ ہو ایسی صورتوں میں کوریلیشن گمراہ کن نتائج دے سکتا ہے۔

3. متعلقہ متغیرات ہو سکتا ہے کہ باہمی طور پر ایک دوسرے کو متاثر کر رہے ہوں تاکہ ان میں سے کوئی وجہ اور اثر کا باعث نہ بنا ہو۔ایسی صورتحال خصوصاً معاشیات اور تجارت کے میدان میں درست، صحیح ہو سکتی ہے مثال کے طور پر ایک شئے کے لئے طلب قیمتوں میں اضافہ کے نتیجہ میں کمی واقع ہو سکتی ہے نارمل طور پر کوئی یہ فرض کر سکتا ہے کہ قیمت وجہ اور طلب اثر ہے تاہم ایسا ہو سکتا ہے کہ شئے کی طلب میں اضافہ مستقبل میں متوقع قلت کے خوف سے ہوا ہو اور جس کے نتیجے میں قیمتوں میں اضافہ ہوا اب طلب وجہ ہو سکتی ہے اور قیمت اثر ہو سکتی ہے۔

4. کئی مرتبہ کوریلیشن کو دو متغیرات کے درمیان بغیر کسی حقیقی تعلق کے نوٹس کیا گیا۔ایسا اتفاقی طور پر ہو سکتا ہے نارمل طور پر ایسا اس وقت ہوتا ہے جب کہ موجودات سے ایک بہت چھوٹا سا نمونہ (Sample) چنا گیا ہو۔ مثال کے طور پر ایک چھوٹا سا نمونہ اجرتوں اور پیداواریت میں بہت زیادہ کوریلیشن (کوریلیشن) پائی جائے گی تاکہ ہم یہ یقین کر لیں کہ زیادہ اجرتیں زیادہ پیداواریت کا باعث ہوئی ہیں۔ موجودات میں حقیقی صورتحال اس کے برعکس ہو سکتی ہے یہ کہ زیادہ اجرتیں کم پیداواریت کا باعث ہوئی ہوں یا ان میں کوئی کوریلیشن (کوریلیشن) نہ پائی جائے۔ لہٰذا یہ ضروری ہے کہ دو متغیرات کے درمیان کوریلیشن کے تجزیہ کے دوران ہمیں جلد بازی میں کسی نتیجہ پر پہنچ جائیں۔

ایک ایسی صورتحال بھی ہو سکتی ہے کہ زیر مطالعہ دو متغیرات کے دوران بے معنی یا جعلی کوریلیشن پایا جائے۔ایسا بھی ہو سکتا ہے کہ کوئی ہر سال بیواؤں کی تعداد اور ٹیلی ویژن سیٹس کے برآمدات کے درمیان کوئی تعلق نہیں ہو سکتا۔ یہی وجہ ہے کہ اس کو بآسانی سمجھا جا سکتا ہے کہ کوریلیشن صرف متعلقہ متغیرات کے درمیان پایا جاتا ہے۔

مذکورہ بالا نکات سے ظاہر ہے کہ کوریلیشن صرف حسابی تعلق ہے اور یہ ضروری نہیں کہ متغیرات کے درمیان تعلق وجہ اور اثر کی دلالت کرتا ہے۔

مزید برآں یہ بھی یاد رکھا جائے گا کہ کوریلیشن کی موجودگی اس بات کی ضمانت نہیں کہ دو متغیرات کے درمیان تعلق ہمیشہ ایک ہی قسم کا ہوگا۔ نارمل طور پر بلند قد و قامت شوہر کی بیویاں بلند قد و قامت والی ہوتی ہیں اور پستہ قد شوہر پستہ قد و قامت والی بیویاں رکھ سکتے ہیں۔ لیکن اسکا مطلب ہر گز یہ نہیں کہ ہر بلند قد و قامت شوہر بلند قد و قامت والی بیوی رکھے گا۔ کئی بلند قد و قامت شوہر پستہ قد و قامت والی بیویاں رکھ سکتے ہیں اور پستہ قد و قامت والے شوہر بلند قد و قامت والی بیویاں رکھ سکتے ہیں حتیٰ کہ ایک شوہر دو بیویاں رکھ سکتا ہے ایک بلند قد و قامت اور دوسری پستہ قد و قامت والی۔ اسی طرح قیمت اور طلب میں کوریلیشن ہو سکتا ہے تاکہ عمومی طور پر جب کبھی قیمتوں میں اضافہ ہوگا تب

طلب کم ہو گی اور قیمتیں کم ہوں گی تو طلب میں اضافہ ہو گا۔ لیکن اس کا یہ مطلب نہیں کہ جب قیمت میں اضافہ ہو گا طلب میں کمی ہو گی۔ یہ ممکن ہے کہ قیمت میں اضافہ کے ساتھ طلب میں بھی اضافہ ہو گا۔ بر بنائے حقیقت معاشی اور سماجی مظہر مختلف عناصر اعداد و شمار پر اثر انداز ہوتے ہیں ساتھ ہی ساتھ یہ نہایت مشکل بلکہ ان عوامل کے اثرات کا الگ الگ مطالعہ کرنا ناممکن ہے۔

تعریف:

کوریلیشن کی مختلف ماہرین نے اپنے اپنے منفرد انداز میں تعریف کی ہے چند ماہرین کی تعریفات کو ذیل میں پیش کیا گیا ہے۔

1. W.I. King کے مطابق "کوریلیشن کا مطلب اعداد و شمار کا دو سلسلوں یا گروہوں کے درمیان کچھ عارضی تعلق کا وجود ہے"۔

2. Ya Lun Chow نے کوریلیشن کی تعریف یوں کی ہے کہ "کوریلیشن تجزیہ دو متغیرات کے درمیان تعلقات کی ڈگری کے تعین کی کوشش ہے"۔

3. Croxton and Cowder کے مطابق "جب تعلق مقداری نوعیت کا ہو، تعلق کی دریافت کاری اور پیمائش کے لئے موزوں شماریاتی آلہ اور اس کو مختصر اور جامع ضابطہ کوریلیشن کے طور پر جانا جاتا ہے"۔

4. Simpson and Kafka کے مطابق "کوریلیشن تجزیہ دو یا زائد متغیرات کے درمیان ایک بندشِ خیالات جوڑ سے متعلق ہے"۔

5. A.M. Tuttle نے کوریلیشن کی تعریف یوں کی ہے "کوریلیشن دو یا زائد متغیرات کے درمیان حسابی تعلق کا ایک تجزیہ ہے"۔

6. Tippet کے مطابق "ہماری پیشن گوئی میں غیر یقینی کے حدود میں کمی کوریلیشن کا اثر ہے"۔ دو متغیرات کو، کوریلیٹیڈ اُو وقت کہا جائے گا جبکہ ایک متغیر میں تبدیلی کے نتیجہ میں دوسرے متغیر میں بھی مشابہ / مماثل تبدیلی ہو۔ اس طرح کوریلیشن ایک شماریاتی آلہ ہے جو دو یا دو سے زیادہ متغیرات میں Correlation کے تجزیہ میں ہماری مدد کرتا ہے۔

مختلف سلسلوں کے درمیان تعلق سے متعلق تجزیہ کے مسئلہ کو تین اقدامات کی صورت میں دیکھا جائے گا۔

1. تعین کرنا کہ آیا تعلق موجود ہے اور اگر ایسا ہے تو پیمائش کرنا۔
2. ٹیسٹنگ کرنا کہ آیا یہ بامعنی ہے۔
3. کیا کوئی وجہ اور اثر کا تعلق قائم بھی ہے۔

اس باب میں صرف اول الذکر پہلو پر بحث ہو گی۔ دوسرے پہلو کے لئے باب "Test of Significance" کا حوالہ دیا جا سکتا ہے۔ تیسرے پہلو کے لئے تجزیہ ہو گا جو وجہ اور اثر تعلق قائم کرے گا اور اس کے ساتھ شماریاتی برتاؤ مشکل ہو گا۔ سگریٹ نوشی میں اضافہ اور پھیپھڑوں کے درمیان زیادہ اور نمایاں کوریلیشن ثابت نہیں کیا جا سکتا کہ سگریٹ نوشی کی وجہ سے پھیپھڑوں کا کینسر ہوتا ہے وجہ اور اثر تعلق کا ثبوت صرف اسی وقت سامنے آئے گا۔ جب کہ موثر عناصر کا جامع مطالعہ کیا جائے۔ یہ بھی نوٹ کیا جائے گا کہ دو شماریاتی متغیرات کے درمیان کچھ تعلق درکار ہو گا جو جوڑ میں مشاہدے سے جڑا ہو گا ہر جوڑ سے دو متغیرات کی قدر کے منجملہ ہر ایک کی قدر ہو گی نارمل طور پر

جوڑ میں تعلق کسی بھی نوعیت کا ہو سکتا ہے جیسے وقت پر مشاہدات یا ایک مدت کے بعد یا مختلف جگہوں' مقامات یا ایک عرصہ کے بعد یا مختلف مقامات پر۔

قربت کے زاویہ کو محسوب کرنے کی بنیاد ریگریشن مساوات پر ہے۔تاہم ریگریشن ایکویشن کے بغیر بھی کوریلیشن تجزیہ پیش کرنا ممکن ہے۔

7.4 کوریلیشن کی ضرورت(Need of Correlation)

کسی بھی ترتیب کو کوریلیشن ایک معنی عطا کرتی ہے۔کوریلیشن کا تجزیہ بنیادی تعلیمی نفسیاتی تحقیق کے لیے ضروری ہے۔نارمل طور پر تمام نفسیاتی تحقیقات پر ہمرشتگی کا اطلاق ہوتا ہے۔اس میں نفسیاتی اور تعلیمی جانچ کا انحصار، بھروسہ اور ہر ایٹم کے تجزیہ پر ہوتا ہے۔ بعض اوقات اعداد و شمار Hypothesis کے مطابق نہیں ہوتے۔اس کی پیمائش اور جانچ بھی کوریلیشن کی مناسبت سے کی جاتی ہے۔ایک معلوم متغیر کی بنیاد پر دوسرے متغیر کی پیش قیاسی کی جاسکتی ہے۔کوریلیشن کی بنیاد پر نفسیاتی اور تعلیمی ماڈلس اور نظریات کو آگے بڑھایا جاسکتا ہے۔اعداد و شمار کی یا متغیروں کی باریک بینی سے گروہ بندی کے لیے بھی کوریلیشن کو استعمال کیا جاسکتا ہے۔ہمہ اختلافی شماریاتی ٹیسٹس جیسے فیکٹر اینالائسیس، ممیز تجزیہ(Discriminant Analysis) کے لیے استعمال کیا جاسکتا ہے۔متغیرات کے ایکدوسرے پر اثرات سے الگ رکھنے کے لیے بھی کوریلیشن کی ضرورت ہوتی ہے۔

7.5 کوریلیشن کے اقسام(Types of Correlation)

کوریلیشن کی مختلف انداز سے درجہ بندی کی جاسکتی ہے۔کوریلیشن کی درجہ بندی تین اہم طریقوں سے کی جاتی ہے۔(1) مثبت یا منفی (2) سادہ' جزوی اور بکثرت (3) خطی اور غیر خطی۔

1۔ مثبت یا منفی Positive and Negative

کوریلیشن مثبت یا منفی ہو سکتا ہے جب دو متغیرات کی قیمتیں ایک ہی سمت حرکت کر رہی ہوں نتیجتاً ایک متغیرہ کی قیمت میں اضافہ دوسرے متغیرہ کی قیمت میں اضافہ سے منسلک ہو یا ایک متغیرہ کی قیمت میں کمی دوسرے متغیرہ کی قیمت میں کمی کی وجہ سے ہو رہی ہو تو یہ مثبت کوریلیشن ہے اسکے بر خلاف متغیرات کی قیمت میں تبدیلی' ایک دوسرے کے متضاد ہو یعنی ایک متغیرہ کی قیمت میں اضافہ ہو رہا ہو تو دوسرے متغیرہ کی قیمت میں کمی واقع ہو رہی ہو تب منفی کوریلیشن تصور کیا جائے گا۔کچھ اعداد و شمار ایسے ہیں جن کا کوریلیشن مثبت ہوتا ہے وہیں دوسرے اعداد و شمار کے لئے منفی بھی ہوتا ہے اس طرح قیمت اور رسد مثبت انداز میں ہم رشتہ ہیں جب قیمتوں میں اضافہ ہوتا ہے تو رسد میں بھی اضافہ ہوتا ہے اور جب قیمتوں میں کمی ہوتی ہے تو رسد میں بھی کمی ہوتی ہے۔ عموماً قیمت اور طلب میں کوریلیشن منفی ہوتا ہے۔ قیمتوں میں اضافہ طلب میں کمی واقع ہوتی ہے اور قیمتوں میں کمی سے طلب میں بھی اضافہ ہوتا ہے۔

مثبت کوریلیشن:

قیمت	10	15	20	25	40	30	20	10
رسد	100	110	115	130	150	140	115	100

منفی کوریلیشن:

قیمت	10	15	20	25
رسد	60	80	90	100

قیمت	25	20	15	10
رسد	60	80	90	100

مثبت کوریلیشن سلسلے کی چند مثالیں نیچے دی گئی ہیں:

1. قداور وزن
2. خاندان کی آمدنی اور اشیائے تعیش پر خرچ
3. بارش کی مقدار اور فصل کی پیداوار
4. ایک شئے کی قیمت اور رسد

منفی کوریلیشن سلسلے کی چند مثالیں ذیل میں دی گئی ہیں:

1. ایک شئے کی طلب اور قیمت
2. ایک پرفیکٹ گیس کا حجم اور دباؤ
3. اون کے کپڑوں کی فروخت اور دن کا ٹمپریچر وغیرہ۔

2۔ سادہ مرکب اور جزوی کوریلیشن:

سادہ کوریلیشن میں ہم صرف دو متغیرات کا مطالعہ کرتے ہیں جیسے قیمت اور طلب مرکب کوریلیشن میں ہم تین یا زیادہ عناصر کے درمیان تعلق کا مطالعہ کرتے ہیں۔ جیسے پیداوار' بارش اور کھاد کا استعمال۔ جزوی کوریلیشن میں اگرچیکہ دو عناصر سے زیادہ عناصر ہوتے ہیں لیکن کوریلیشن کا مطالعہ دو عناصر کے درمیان ہوتا ہے دیگر عناصر کو مستقل طور پر فرض کیا جائے گا۔

3۔ خطی اور غیر خطی (کروی لائنر) کوریلیشن:

جب دو متغیرات کی قدروں میں تغیر برابر تناسب میں ہوں تب کوریلیشن کو خطی کہا جائے گا اگر ہر دفعہ قیمت میں10%اضافہ ہوتب رسد میں بھی20%اضافہ ہوگا۔ان دو متغیروں کے درمیان خطی تعلق پایا جاتا ہے۔ان میں تعلق کی قسم $Y = a + bx$ ہے۔ یہ خط مستقیم کی ایک مساوات ہے۔ اگر مقابل دو ایسے سلسلوں کی قدروں کو گراف پیپر پر مرتسم کریں تو خط مستقیم حاصل کیا

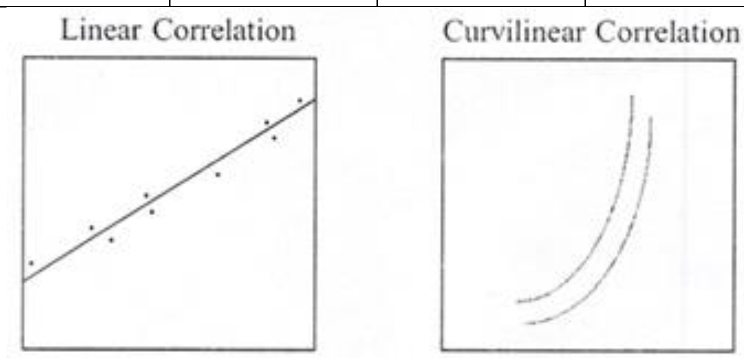
جاسکے۔ تاہم معاشی اور سماجی ڈاٹا میں ایسا تعلق بہت نایاب ہوتا ہے۔ ایسا نارمل طور پر سائنس میں پایا جاتا ہے۔ معاشی اعداد و شمار میں دو متغیرات میں تبدیلی عموماً یکساں یعنی Constant نہیں ہوتی۔ ایسی صورتوں میں دو متغیرات کے مقابل اعداد و شمار سے خط مستقیم حاصل نہ ہوگا کوریلیشن منحنی یا غیر خطی ہوگا۔ اس طرح خطی کوریلیشن وہ ہے جہاں متعلقہ متغیرات میں تغیرات کا تناسب مستقل (Constant) ہوگا۔ اور غیر خطی وہ ہے جہاں یہ تناسب لچکدار ہوگا۔

خطی کوریلیشن:

X	2	4	6	8	10
Y	5	10	15	20	25

غیر خطی کوریلیشن:

X	2	4	6	8	10
Y	5	8	12	15	25



Source: http://simplypsychology.org/correlation.html

7.6 کارل پیرسن کا کوفیشنٹ آف کورلیشن کی نوعیت

(Nature of Karl Pearson of Coefficient of Correlation)

کارل پیرسن (1867-1936) برطانیہ کا رہنے والا تھا۔ وہ ایک ماہر حیاتیات اور ماہر شماریات تھا۔ اس نے کوریلیشن کے کوفیشنٹ کا ضابطہ دیا۔ کوریلیشن کی پیمائش کے کئی حسابی طریقے ہیں۔ ان میں سے ایک کو کارل پیرسن کا طریقہ بھی ہے۔ یہ طریقہ Pearson's Coefficient of Correlation کے نام سے مشہور ہے اس ضابطہ کو بہت زیادہ استعمال کیا جاتا ہے۔ پیرسن کے کوریلیشن کے کوفیشنٹ کو "r" علامت سے ظاہر کیا جاتا ہے۔ کارل پیرسن کے مطابق دو متغیرات کے کوریلیشن کے کوفیشنٹ کو دو سلسلوں کے مختلف ایٹمس کے مقابل اعداد کے اوسط حسابیہ سے لئے گئے انحراف کے حاصل ضرب کے مجموعہ کو معیاری انحراف اور زیر مشاہدہ جوڑ کی تعداد سے تقسیم

کر کے حاصل کیا جاتا ہے۔ اس طرح اگر $x_1, x_2, x_3 .... x_n$ پہلے متغیرہ کے مختلف ایٹمس کے اوسط حسابیہ کے انحراف ہوں اور y_1, y_2, y_3

y_n.......................دوسرے مقابل متغیرہ کے اوسط حسابیہ کے انحراف ہوں ان مقابل انحراف کے حاصل ضرب کا مجموعہ $\sum xy$ ہوگا۔ اگر دو متغیروں کے معیاری انحراف بالترتیب $\sigma_1 \sigma_2$ ہوں اور اگر 'n' زیر مشاہدہ جوڑ کی تعداد ہو تب r (کوریلیشن کا کوفیشنٹ) یوں ہوگا۔

$$r = \frac{\sum xy}{n\sigma_1\sigma_2}$$

مذکورہ بالا ضابطے سے یہ بات واضح ہے کہ اگر $\sum xy$ مثبت (پازیٹیو) ہو تب کوریلیشن کا کوفیشنٹ بھی پازیٹیو ہوگا جو دونوں سلسلوں کے کوریلیشن کے مثبت ہونے کی نشاندہی کرے گا۔ اس کے برخلاف اگر $\sum xy$ نیگیٹیو (منفی) ہو تب کوریلیشن کا کوفیشنٹ بھی منفی ہوگا دونوں سلسلوں کے درمیان کوریلیشن کے نیگیٹیو ہونے کی نشاندہی کرے گا۔ نارمل طور پر اگر ایک سلسلے کے پازیٹیو اور نیگیٹیو انحراف دوسرے سلسلے کے پازیٹیو اور نیگیٹیو سے بھی جڑے ہوں تب $\sum xy$ پازیٹیو ہوگا۔ نارمل طور پر اگر ایک سلسلے کے پازیٹیو انحراف دوسرے متغیرہ کے نیگیٹیو انحراف سے جڑے ہوں یا برعکس ہوں تب $\sum xy$ کی قدر نگیٹیو ہوگی۔ اگر ایک متغیرہ کے پازیٹیو اور نیگیٹیو دوسرے متغیرہ کے انحراف سے بے نیازی/طور پر جڑے ہوں تب $\sum xy$ صفر ہوگا یا اس کے قریب ہوگا اور اس بات کی نشاندہی کرے گا کہ دو سلسلوں کے درمیان کوئی کوریلیشن نہیں ہے۔ کوریلیشن کے کوفیشنٹ کی قدر ہمیشہ 1+ اور 1- کے درمیان ہوگی۔ یہ ایک سے زیادہ بھی ہو سکتا۔

کارل پیرسن کا مذکورہ بالا ضابطہ دونوں سلسلوں کے درمیان Covariance کے مطالعہ کی بنیاد پر وضع کیا گیا ہے۔ دونوں سلسلوں کے درمیان Covariance کو یوں لکھا جاتا ہے۔

$$\text{Covariance} = \frac{\sum xy}{x}$$

جہاں x اور y دونوں سلسلوں کے اوسط حسابیہ سے انحراف سے متعلق ہیں کوریلیشن کا مطالعہ کرنے کے لئے دو سلسلوں کے Covariance کو ان کے معیاری انحراف کے حاصل ضرب سے تقسیم کیا گیا ہے لہذا

$$r = \frac{\text{Covariance of the two Series}}{\sqrt{(\text{Variance of Series 1})\ (\text{Variance of Series 2})}}$$

$$= \frac{\text{Covariance}}{\sigma_1 \times \sigma_2}$$

$$= \frac{\sum xy}{n\sigma_1\sigma_2}$$

اس فارمولے کو Product Moment formula of Coefficient of Correlation کے طور پر جانا جاتا ہے۔

راست طریقہ Direct Method

اس طریقہ میں شامل اقدامات درج ذیل ہوں گے :

1. دونوں سلسلوں کے اوسط حسابیہ (Arithmatic Mean) کو دریافت کریں۔ $(\overline{x}_1\overline{y}_2)$
2. سلسلے کے ہر ایک ایٹم کے لئے اس کے اوسط حسابیہ سے انحراف دریافت کریں (x اور y)۔
   یہاں $x = (x - \overline{x})$ اور $y = (y - \overline{y})$
3. ان انحراف کا مربع معلوم کر کے جمع کریں۔ $(\sum x^2, \sum y^2)$
4. دونوں سلسلوں کے متعلقہ انحرافی اعداد کو ضرب دے کر جمع کریں۔ $(\sum xy)$
5. مذکورہ بالا قیمتوں کو درج ذیل فارمولے کے مطابق لکھ کر مختصر کریں۔

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \times \sum y^2}} \quad \text{یا} \quad r = \frac{\sum xy}{\sqrt[n]{\frac{\sum x^2}{n}} \times \sqrt[n]{\frac{\sum y^2}{n}}} = \frac{\sum xy}{n\sigma_1\sigma_2}$$

تمثیل 1: درج ذیل دس فرم کے فروخت اور اشتہار کاری کے مصارف کے درمیان کوریلیشن کا کوفیشنٹ دریافت کیجئے (اعداد 000 میں)

فرمس	1	2	3	4	5	6	7	8	9	10
فروخت	50	50	55	60	65	65	65	60	60	50
اشتہار کاری اخراجات	11	13	14	16	16	15	15	14	13	13

Calculation of Coefficient of Correlation

Sales	Deviation from Mean 58	Deviation Square	اشتہاری اخراجات	Deviation from Mean 14	Devaition Square	Product of deviation, xy
50	-8	64	11	-3	9	24
50	-8	64	13	-1	1	8
55	-3	9	14	0	0	0

60	+2	4	16	+2	4	4
65	+7	49	16	+2	4	14
65	+7	49	15	+1	1	7
65	+7	49	15	+1	1	7
60	+2	4	14	0	0	0
60	+2	4	13	-1	1	-2
50	-8	64	13	-1	1	8
∑x 580 $x=\frac{580}{10}=58$ ∑	-27 +27 = 0 ∑x	x^2 = ∑360	∑y 140 $=1\bar{y}\frac{140}{12}$ 4	+6 0 -6 ∑y	$\sum y^2 = 22$	xy = ∑70

$$\text{Mean of } x = \frac{\sum x}{n} = \frac{580}{10} = 58$$

$$y = \frac{\sum y}{n} = \frac{140}{10} = 14$$

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \times \sum y^2}} = \frac{+70}{\sqrt{360 \times 22}} = \frac{70}{\sqrt{7920}}$$

$$= \frac{+70}{88.9943818}$$

$$= +0.7865$$

تمثیل 2: نیچے پیکنگ کی لاگت اور فروخت کے دیے گئے اعداد کے مطابق کارل پیرسن کے کوریلیشن کے کوفیشنٹ کی تحسیب کیجئے۔

پیکنگ کی لاگت (000روپیوں میں)	39	65	62	90	82	75	25	98	36	78
فروخت(لاکھوں میں)	47	53	58	86	62	68	60	91	51	84

Calculation of Coefficient of Correlation

پیکنگ کی لاگت '000) (روپیوں میں	Deviation from Mean (65)	Square of Deviation	Sales in) (lakh	Deviation from Mean (65)	Square of Deviation	Product of Deviations

39	-26	676	47	-19	361	+494
65	+0	0	53	-13	169	0
62	-3	9	58	-8	64	+24
90	+25	625	86	+20	400	500
82	+17	289	62	-4	16	-68
75	+10	100	68	+2	4	+20
25	-40	1600	60	-6	36	+240
98	+33	1089	91	+25	625	+825
36	-29	841	51	-15	225	+435
78	+13	169	84	+18	224	+234
$\sum x = 70$ $x = 10$	$\sum x = 0$	$x^2 =$ $\sum 5398$	$y =$ $\sum 660$ $x \sum 10$	$\sum y = 0$	$y^2 =$ $\sum 2224$	$xy =$ $\sum 2704$

$$\text{Mean of Series } (x) = \frac{\sum x}{n} = \frac{650}{10} = 65$$

$$(y) = \frac{\sum y}{n} = \frac{660}{10} = 66$$

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \times \sum y^2}} = \frac{2704}{\sqrt{(65)^2(66)^2}} \frac{2704}{\sqrt{5398 \times 2224}} = \frac{2704}{\sqrt{123005152}} = \frac{2704}{3462.84516}$$

$$= 0.78086$$

تمثیل 3: درج ذیل نتائج دو سلسلوں کے درمیان ان کے اوسط حسابیہ سے حاصل کئے گئے ہیں کوریلیشن کا کوفیشنٹ تحسیب کیجئے۔

	سلسلہ 'y'	سلسلہ 'y'
ایٹمس کی تعداد	7	7
اوسط حسابیہ	4	8
اوسط حسابیہ سے انحرافی اعداد کے مربعوں کا مجموعہ	28	76
متعلقہ اوسط حسابیہ سلسلہ x اور y کے حاصل ضرب کا مجموعہ	46	

حل: دیا گیا ہے کہ:

$$x = 7,\ \bar{x} = 4,\ \bar{y} = 8,\ \sum x^2 = 28\ \sum x^2 = 76\ \sum xy = 46$$

کارل پیرسن کی کوریلیشن کا کوفیشنٹ یا

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \times \sum y^2}} = \frac{46}{\sqrt{28 \times 76}}$$

$$= \frac{46}{\sqrt{2128}} = \frac{46}{46.1302504}$$

$$= 0.997$$

تمثیل 4: درج ذیل نتائج سے کوریلیشن کا کوفیشنٹ تحسیب کیجئے

$$x = 10\ \ \sum x = 100\ \ \sum y = 150$$

$$\sum (x-10)^2 = 180\ \ \sum (y-15)^2 = 215$$

$$\sum (x-10)(y-15) = 60$$

$$\sum (x-10)^2, 180,\ \sum (y-15)^2 = 215$$

$$\sum (x-10)(y-15) = 60$$

Mean of the first Series or $\bar{x} = \frac{\sum x}{n} = \frac{100}{10} = 10$

Second source or $\bar{y} = \frac{\sum x}{n} = \frac{150}{10} = 15$

$$\sum (x-10) = Sum$$

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \times \sum y^2}} = \frac{60}{\sqrt{180 \times 215}} = \frac{60}{196,7213156}$$

$$= +\ 0.305$$

تمثیل 5: نیچے دیئے گئے اعداد و شمار سے ایٹمس کی تعداد 'n' دریافت کیجئے۔

$$r = +0.5\ \sum xy = 120,\ \sigma_y = 8, \sum x^2 = 90$$

جہاں x اور y انحراف اور حسابیہ سے لئے گئے ہیں۔

حل: دیا گیا ہے $r = +0.5\ \sum xy = 120,\ \sigma_y = 8, \sum x^2 = 90$

$$Now\ \sigma_x^2 = \frac{\sum x^2}{n} = \frac{90}{n}$$

$$r = \frac{\sum xy}{n\sigma_x \times \sigma_y} or\ r^2 = (0.5)^2 = 0.25$$

$$0.25 = \frac{(120)^2}{n^2 \times \left(\frac{90}{n}\right) \times 64} = \frac{14400}{n \times 90 \times 64} = \frac{14400}{5760n} = \frac{2.5}{n}$$

$$\frac{0.25}{100} \times \frac{2.5}{n} = 25n = 2.5 + 100$$

$$25n = 250 = n = \frac{250}{25} = 10$$

2. راست طریقہ(Direct method):

کوریلیشن کے کوفیشنٹ کی تحسیب سلسلوں کے مختلف اعداد، قیمتوں کے اوسط حسابیہ، مفروضہ حسابیہ کے انحراف کے بغیر کی جاسکتی ہے۔اسکے لئے درج ذیل فارمولہ استعمال کیاجاتاہے۔

$$r = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{N\sum x^2 - (\sum x)^2}\sqrt{N\sum y^2 - (\sum y)^2}}\ or$$

تمثیل6: درج ذیل اعداد وشمار سے کارل پیرسن کے کوریلیشن کاکوفیشنٹ تحسیب کیجئےاوراس کی قیمت کی تشریح کیجیے۔

طلباء کارول نمبر	1	2	3	4	5	
اکنومکس میں نشانات	48	35	17	23	47	= 170/5=34
شماریات میں نشانات	45	20	40	25	45	=175/5=35

حل :

x	X^2	y	Y2	xy
48	2304	45	2025	2160
35	1225	20	400	700
17	289	40	1600	680
23	529	25	625	575
47	2209	45	2025	2115
170	6556	175	6675	6230

$$N = 5, \quad \sum xy = 6230, \sum x = 170$$
$$\sum xy = 175, \sum x^2 = 6556, \sum y^2 = 6230$$

$$r = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{N\sum x^2 - (\sum x)^2}\sqrt{N\sum y^2 - (\sum y)^2}}$$

$$= \frac{5\times 6230 - (170*175)}{\sqrt{5\times 6556 - (170)^2}\sqrt{5\times 6675 - (175)^2}}$$

$$= \frac{5\times 6230 - (170*175)}{\sqrt{5\times 6556 - (170)^2}\sqrt{5\times 6675 - (175)^2}} = \frac{31{,}150 - 29{,}750}{\sqrt{32{,}780 - 28{,}900}\sqrt{33{,}375 - 30{,}625}}$$

$$= \frac{1400}{\sqrt{3880}\sqrt{2750}}$$

$$= \frac{1400}{62.289646\times 52.4404424}$$

$$= \frac{1400}{3266.49659}$$

$$= 0.42859$$

تمثیل 7: نیچے دیئے گئے اعداد سے راست طریقہ سے کوریلیشن کا ضرب تحسیب کیجئے۔

x:	12	9	8	10	11	13	7
y:	14	8	6	9	11	12	3

حل:

x	x^2	y^1	y^2	Xy
12	144	14	196	168
9	81	8	64	72
8	64	6	36	48
10	100	9	81	90
11	121	11	121	121
13	169	12	144	156
7	49	3	9	21
70	728	63	651	676

$$r = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{N\sum x^2 - (\sum x)^2}\sqrt{N\sum y^2 - (\sum y)^2}}$$

$$= \frac{7\times 676 - 70\times 63}{\sqrt{7428 - (70)^2}\sqrt{7\times 651 - (63)^2}} = \frac{4732 - 4410}{\sqrt{5096 - 4900}\sqrt{4557 - 3969}}$$

$$= \frac{4732 - 4410}{\sqrt{5096 - 4900}\sqrt{4557 - 3969}}$$

$$= \frac{322}{\sqrt{196}\sqrt{588}}$$

$$= \frac{322}{14\times 24.2487113}$$

$$= \frac{322}{339.481958}$$

$$= 0.9485$$

تمثیل 8: x اور y کی قیمتیں نیچے دی گئی ہیں ان کے درمیان کوریلیشن کا کوفیشنٹ تحسیب کیجیے۔

x :	65	66	67	67	68	69	70	72
y :	67	68	65	68	72	72	69	71

x	y	$x2$	$y2$	(xy)
65	67	4225	4489	4355
66	68	4356	4624	4488
67	65	4489	4225	4355
67	68	4489	4624	4556
68	72	4624	5184	4896
69	72	4761	5184	4968
70	69	4900	4761	4830
72	71	5184	5041	5112
544	552	3,7028	3,8132	3,7560

$$r = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{N\sum x^2 - (\sum x)^2}\sqrt{N\sum y^2 - (\sum y)^2}}$$

$$= \frac{37560 - 544\times 552}{\sqrt{8\times 37560 - (544)^2}\sqrt{8\times 38132 - (552)^2}}$$

$$= \frac{3,00,480 - 300288}{\sqrt{296224 - 295936}\sqrt{305056 - 304704}}$$

$$= \frac{192}{\sqrt{288}\sqrt{352}}$$

$$= \frac{192}{16.9705627\times 18.761663}$$

$$= \frac{192}{318.395978} = 0.6030$$

شارٹ کٹ طریقے Shortcut Methods(مفروضہ اوسط حسابیہ یامبدے کی تبدیلی کے استعمال سے)

راست طریقے میں کوریلیشن کے کوفیشنٹ کی تحسیب کاری میں دونوں سلسلوں کے حقیقی اوسط حسابیہ کی تحسیب کرناہوتاہے اور ان اوسط حسابیہ سے دو سلسلوں کے مختلف ایٹمس سے انحرافی اعداد کو معلوم کیاجاتاہے۔ اگر سلسلہ کا اوسط حسابیہ اعشاریہ میں ہو جیسے 80.5 یا 94.08 میں ہو تب انحرافی اعداد کی تحسیب مشقت طلب اور تکلیف دہ ہو جاتی ہے اور یہ کام ذہنی تھکان کا باعث ہوتاہے کیونکہ جب اعداد اعشاریہ میں ہوں تب انحراف کی تحسیب اور x^2، y^2اور xy کی تحسیب اکتانے والا کام ہو جاتاہے۔ ایسی صورت میں یہ مشورہ دیاجاتاہے کہ اوسط حسابیہ کو فرض کریں اور ضابطہ (فارمولے) میں تبدیلی لائی جس کی وجہ سے انحراف کے معلوم کرنے کا عمل نہایت آسان ہو جاتاہے۔ فارمولے میں ضروری تبدیلی کے بعد اس کی اصلیت تبدیل ہوتی ہے اور نہ ہی کوریلیشن کے کوفیشنٹ کی قیمت پر کوئی فرق واقع ہوتاہے۔ r کی قدر بھی اتنی ہی ہوگی جتنی کہ حقیقی اوسط حسابیہ کے استعمال سے ہوا کرتی تھی۔

حقیقی اوسط حسابیہ کے عدم استعمال کی صورت میں فارمولہ (ضابطہ) درج ذیل ہوگا۔

$$r = \frac{\sum dxdy - \frac{(\sum dx)(\sum dy)}{n}}{\sqrt{\sum dx^2 - \frac{(\sum dx)^2}{n}}\sqrt{\sum dy^2 - \frac{(\sum dy)^2}{n}}} \text{........(1)}$$

جہاں $\sum dx$ اور $\sum dy$ بالترتیب مفروضہ اوسط حسابیہ سے سلسلہ x اور y کے لئے انحرافی اعداد معلوم کئے جائیں گے۔

$\sum dx^2$ اور $\sum dy^2$ بالترتیب x اور y مربعوں کا مجموعہ ہے۔

$\sum dxdy$ مفروضہ اوسط حسابیہ سے x اور y کے مقابل انحراف کے حاصل ضرب کا مجموعہ ہے۔ مذکورہ بالا فارمولہ یوں بھی لکھا جاسکتاہے۔

$$r = \frac{\sum dxdy - n\left(\frac{\sum dx}{n}\right)\left(\frac{\sum dy}{n}\right)}{n\sqrt{\frac{\sum dx^2}{n} - \frac{(\sum dx)^2}{n}}\sqrt{\sum dy^2 - \left(\frac{\sum dy^2}{n}\right)}} \text{........(2)}$$

یایوں بھی لکھا جاسکتا ہے۔

$$r = \frac{\sum dxdy - (\sum dx)(\sum dy)}{\sqrt{(n\sum dx^2 - (\sum dx)^2}\sqrt{n\sum dy^2 - (\sum dy)^2}} \text{........(3)}$$

مذکورہ بالا تمام ضابطوں کے استعمال کی صورت میں r کی قیمت میں کوئی فرق واقع نہ ہوگا۔

اقدامات

(i) x اور y سلسلوں میں کسی ایک عدد کو اوسط حسابیہ فرض کریں۔

(ii) مفروضہ اوسط حسابیہ سے x اور سلسلہ سے انحرافی اعداد معلوم کریں تب ہمیں dx اور dy حاصل ہوگا۔

(iii) dx اور dy کا مربع بنائیں اور ان مربع اعداد کا مجموعہ $\sum dx^2$ اور $\sum dy^2$ معلوم کریں۔

(iv) x اور y سلسلوں کے مقابل انحرافی اعداد کو ایک دوسرے سے ضرب دیں۔(dy اور dx) اور حاصل ضرب کا مجموعہ $\sum dxdy$ کے حصول کا باعث ہوگا۔

مذکورہ بالا فارمولے میں قیمتیں درج کر کے کوریلیشن کا کوفیشنٹ یا r معلوم کر سکتے ہیں۔

تمثیل 9: نیچے دیے گئے اعداد سے x اور y کی قیمتوں کے درمیان کوریلیشن کا کوفیشنٹ تحسیب کیجئے۔

x:	15	16	17	18	19	20
y:	80	75	60	40	30	20

حل: Calculation of Coefficient of Correlation

x	Dev. of x from AM 17 dx	dx^2	y	Dev. of y from AM	Step Dev. 1-5	Dy^2	Dxdy
15	-2	4	80	+40	+8	64	16
16	-1	1	75	+35	+7	49	-7
17	0	0	60	+20	+4	16	0
18	+1	1	40	0	0	0	0
19	+2	4	30	-10	-2	4	-4
20	+3	9	20	-20	-4	16	12
n = 6	$\sum Dx+3$	$\sum dx^2 19$	x=6	+65	+13	$\sum dy^2 149$	-39

Step Diviation لینے پر فارمولہ میں کسی تطبیق کی ضرورت نہیں ہوتی کیونکہ Scale کی تبدیلی سے کوریلیشن پر کوئی اثر نہیں پڑتا۔

$$r=\frac{\sum dxdy-\frac{(\sum dx)(\sum dy)}{n}}{\sqrt{\sum dx^2-\frac{(\sum dx)^2}{n}}\sqrt{\sum dy^2-\frac{(\sum dy)^2}{n}}}$$

$$=\frac{-39-(+3)(+13)}{\sqrt{19-\frac{(3)^2}{6}}\sqrt{149-\frac{(3)^2}{6}}}=\frac{-45.5}{\sqrt{17.5}\sqrt{120.83}}$$

$$=\frac{-45.5}{4.18330013\times 10.99227}=\frac{-45.5}{45.9839645}=-0.98947$$

تمثیل10: درج ذیل اعداد و شمار سے کوریلیشن کا کوفیشنٹ تحسیب کیجئے اور نتائج پر تبصرہ کیجئے۔

x:	16	12	18	4	3	10	5	12
Y:	23	22	24	17	19	20	18	21

حل:

x	dx	dx^2	y	dy	dy^2	Dxdy
16	+12	144	23	+6	36	72
12	+8	64	22	+5	25	40
18	+14	196	24	+7	49	98
4	0	0	17	0	0	0
3	-1	1	19	+2	4	-2
10	+6	36	20	+3	9	18
5	+1	1	18	+1	1	1
12	+8	64	21	+4	16	32
80	+48	506	164	+28	140	259

$$r=\frac{\sum dxdy-\frac{(\sum dx)(\sum dy)}{n}}{\sqrt{\sum dx^2-\frac{(\sum dx)^2}{n}}\sqrt{\sum dy^2-\frac{(\sum dy)^2}{n}}}$$

$$=\frac{259-\frac{48\times 28}{8}}{\sqrt{506-\frac{(48)^2}{8}}\sqrt{140-\frac{(28)^2}{8}}}=\frac{259-168}{\sqrt{506-288}\sqrt{140-98}}$$

$$=\frac{91}{\sqrt{218}\sqrt{42}}=\frac{91}{14.765\times 6.4807}$$

$$=\frac{91}{95.688}$$

$$=0.951$$

لوگارتھم کا استعمال

$$r=\frac{91}{\sqrt{218}\sqrt{42}}=\text{AL}\left[\log 91-\frac{1}{2}(\log 218+\log 42)\right]$$

$$=\text{AL}\left[1.9590-\frac{1}{2}(2.3385+1.6232)\right]$$

$$=\text{AL}\left[1.9590-\frac{1}{2}\times 3.96177\right]$$

$$=\text{AL}\left[1.9590-1.980885\right]$$

$$\text{AL } 1.9782$$

$$=0.9506$$

تمثیل 11: ایک مسابقتی امتحان کے نتائج درج ذیل ہیں۔

امیدوار کی عمر	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23
شریک امتحان	200	300	100	50	150	400	250	150	25	75
کامیاب عمر	124	180	65	34	99	252	145	81	12	33

مسابقتی امتحان میں کامیاب امیدوار اور ان کی عمر کے درمیان کوریلیشن کا کوفیشنٹ تحسیب کیجئے۔

حل: اس سوال کے حل کے دوران دو نکات زیر غور ہوں گے۔

(i) عمر کی وقفہ جماعت کی نمائندگی ان کی درمیانی قدر کرے گی۔

(ii) کامیاب طلباء کی تعداد نارمل بنیادی اصولوں پر کی جائے گی یعنی فیصد یا فی ہزار یا کوئی دوسری ایسی قدر/قیمت۔

لہٰذا عمر پر مبنی گروہ 13-14 کی نمائندگی وقفہ جماعت کی درمیانی قدر پر ہوگی۔

یعنی 13.5-اور کامیاب امیدواروں کا فیصد $\frac{124}{200}$x100 یا 62 ہوگا۔

اسی طرح 14-15 عمر کے گروہ کی نمائندگی 14.5 سے ہوگی اور کامیاب امیدواروں کا فیصد $\frac{180}{300}$x100 یا 68 ہوگا۔

مذکورہ بالا بنیاد پر کوریلیشن کا کوفیشنٹ بآسانی تحسیب کیا جاسکے گا۔

عمر کا گروہ	درمیانی قدر	اوسط پر مبنی انحراف dx	مفروضہ حسابیہ dx^2	کامیاب امیدواروں کا فیصد	مفروضہ اوسط حسابیہ 59 سے انحراف	dy^2	dxdy
13-14	13.5	-4	16	62	+3	9	-12
14-15	14.5	-3	9	60	+1	1	-13
15-16	15.5	-2	4	+65	+6	36	-12
16-17	16.5	-1	1	68	+9	81	-9
17-18	17.5	0	0	66	+7	49	0
18-19	18.5	+1	1	63	+4	16	+4
19-20	19.5	+2	4	58	-1	1	-2
20-21	2.5	+3	9	54	-5	25	-15
21-22	21.5	+4	16	48	-11	121	-44
22-23	22.5	+5	25	44	-15	225	-75
n=10		∑dx+5	$\sum dx^2$+85	x-10	∑dy-2	$dy^2$56 ∑4	dxdy=- ∑168

$$r=\frac{\sum dxdy-\frac{(\sum dx)(\sum dy)}{n}}{\sqrt{\sum dx^2-\frac{(\sum dx)^2}{n}}\sqrt{\sum dy^2-\frac{(\sum dy)^2}{n}}}$$

$$=\frac{-168-\frac{+5\times -2}{2}}{\sqrt{85-\frac{(5)^2}{2}}\sqrt{564-\frac{(-2)^2}{10}}}=\frac{-168-0}{\sqrt{85-2.5}\sqrt{564-0.4}}$$

$$=\frac{-168}{\sqrt{82.5}\sqrt{563.6}}=\frac{-168}{9.08295\times 23.7902612}$$

$$=\frac{-168}{215.631605}=-0.779$$

7.7 پیرسن کے کوریلیشن کے مفروضات

(Assumptions of Pearson's Coefficient of Correlation)

پیرسن کے کوریلیشن کا کوفیشنٹ درج ذیل مفروضات پر مشتمل ہے۔

I. پہلا یہ کہ متغیرات کے درمیان خطی تعلق ہوتا ہے یعنی منتشر ڈیا گرام پر دو متغیرات کو مرتسم کیا جائے تو ایک خط مستقیم تشکیل پائے گا اس کو آپ نقاط میں مرتسم کیا ہوا پائیں گے۔

II. دوسرا یہ کہ حسب معمول (نارمل) تقسیم کی تشکیل کے لئے زیر مطالعہ دو متغیرات کئی خود سر وجوہات سے متاثر ہوں گے۔ متغیرات جیسے قد، وزن، قیمت، طلب، رسد وغیرہ متاثر ہوں گے اور ایسی قوتیں حسب معمول تقسیم کی تشکیل ہو گی۔

دو سلسلوں میں موجود ایٹمس کی تقسیم پر اثر انداز ہونے والی قوتوں کے درمیان وجہ اور اثر کا تعلق ہوتا ہے۔ اگر متغیرات کے درمیان ایسے تعلق کی تشکیل نہ ہو یعنی اگر متغیرات آزاد ہوں تب کوئی کوریلیشن نہ ہو گا۔ مثال کے طور پر آمدنی اور قد کے درمیان کوئی تعلق نہیں ہو گا کیونکہ قوتیں جو ان متغیرات پر اثر انداز ہوتی ہیں وہ نارمل نہیں ہیں۔

پیرسن کے کوفیشنٹ کی خوبیاں اور خرابیاں

خوبیاں:

تعلقات کی ڈگری کی پیمائش کے لئے مستعملہ حسابی طریقوں کے منجملہ کارل پیرسن کا طریقہ حد درجہ مشہور ہے۔ کوریلیشن کا کوفیشنٹ نہ صرف کوریلیشن کی ڈگری کو صرف ایک عدد میں ملخص کرتا ہے بلکہ قیمت کا تعین کرتا ہے کہ کوریلیشن مثبت ہے یا منفی ہے۔

تاہم کوفیشنٹ کی افادیت کا انحصار اس کے پیمانے کے علم و مفہوم کے ساتھ ساتھ خرابیوں سے بھی واقف ہو۔

خرابیاں:

1. کوریلیشن کے کوفیشنٹ کا ہمیشہ یہ مفروضہ رہا ہے کہ تعلق خطی ہوتا ہے اس حقیقت سے قطع نظر کہ مفروضہ درست بھی ہے یا نہیں۔

2. Coefficient کی قدر کی تشریح کے دوران غیر معمولی چوکس ہونا ضروری ہے کیونکہ اکثر Coeffient کی قدر کی غلط تشریح کی جاتی ہے۔

3. Coefficient کی قدر پر انتہائی ایٹمس سے حد سے زیادہ اثرانداز ہوتے ہیں۔

4. دوسرے طریقوں کے مقابلے میں یہ طریقہ کوریلیشن کے کوفیشنٹ کی قدر کو معلوم کرنے میں بہت زیادہ وقت لگتا ہے۔

7.8 کوریلیشن کے کوفیشنٹ کی تشریح (Interpretation of Coefficient of Correlation)

کوریلیشن کا کوفیشنٹ دو اعداد کے سیٹس کے درمیان تعلق کے زاویہ کی پیمائش کرتا ہے۔ چونکہ تخمینہ جات پر اعتباریت کا انحصار تعلق کی قربت پر ہوتا ہے۔ کوریلیشن کے کوفیشنٹ کی قیمت پر تشریح کرنے کے دوران انتہائی احتیاط برتنا لازمی ہے ورنہ گمراہ کن نتائج نکالے جا سکتے ہیں۔ بد قسمتی سے کوریلیشن کے کوفیشنٹ کی تشریح کا زیادہ تر انحصار تجربہ پر ہوتا ہے۔ r کی اہمیت صرف اسی وقت سمجھ میں آسکتی ہے جب کہ کوریلیشن کے کئی سوالات حل کئے جائیں اور r کی مختلف قیمتوں میں اضافہ اعداد و شمار کی نوعیت کو دیکھنے سے پتہ چلتا ہے۔ محقق کے لئے لازمی ہے کہ وہ اپنے اعداد و شمار سے اچھی طرح واقف رہے تاکہ تشریح و تائید میں کسی بھی غلطی سے پرہیز کیا جا سکے۔ اس کے لئے ضروری ہے کہ وہ اعداد و شمار کے تعلقات اور نظریہ سے واقف ہو یا واقف ہو جانا چاہئے تاکہ اہم متعلقہ امور کو منطقی استدلال اور ذہانتی تحقیق کی بنیاد پر نتائج پر پہنچا جا سکے۔ تاہم درج ذیل میں نارمل اصول دئے گئے ہیں جو r کی قیمت کی تشریح میں مددگار ہوں گے۔

1۔ جب r کی قیمت 1+ ہو اس کا مطلب متغیرات کے درمیان پرفیکٹ پازیٹو تعلق ہے۔

2۔ جب r کی قیمت 1- ہو اس کا مطلب متغیرات کے درمیان پرفیکٹ نیگیٹو تعلق ہے۔

3۔ جب r کی قیمت 0 ہو اس کا مطلب متغیرات کے درمیان کوئی تعلق نہیں ہے یعنی متغیرات میں کوریلیشن نہیں پایا جاتا ہے۔

4۔ جب 1+ 'r' اور 1- سے قریب ہو متغیرات کے درمیان قربت ہوگی اور اگر r کی قیمت 0 سے قریب ہو تو یہ سمجھا جائے گا کہ تعلق میں قربت نہیں ہے۔ اس قیمت کے پرے جانا پر خطر ہوگا یعنی محفوظ نہ ہوگا۔ r کی مکمل تشریح کا انحصار صورتحال پر ہوگا ان میں سے ایک سیمپل کا سائز ہوگا۔ ایسا کہا جا سکتا ہے کہ جب ایک متغیر کی قیمت دے کر دوسرے متغیر کی قیمت کا تخمینہ کرنا ہو r کی قیمت جتنی زیادہ ہوگی اتنا ہی بہتر ہوگا۔ تعلق کی قربت r کے متناسب نہ ہو۔ اگر r کی قیمت 0.8 ہو تو اس کا مطلب یہ نہ ہوگا کہ اس کا تعلق 0.4 کے دوگنا ہے حقیقت میں یہ ایک دوسرے کے قریب ہیں۔

7.9 کوریلیشن کا کوفیشنٹ اوراحتمالی خطا (Coefficient of Correlation and Probable Error)

کوریلیشن کے کوفیشنٹ کی احتمالی خطا اس کی قیمت کی تشریح میں مددگار ہے۔احتمالی خطا کی مدد سے کوفیشنٹ(Coefficient) کی قیمت کی اعتباریت کا تعین ممکن ہے۔ جیسا کہ اب تک یہ مشاہدہ کیا گیا کہ اس کا انحصار Random sampling کی شرائط حالات پر تھا۔ کوریلیشن کے کوفیشنٹ کی احتمالی خطایوں حاصل کی جائے گی۔

$$\text{P.E}_r = 0.6745\frac{1-r^2}{\sqrt{N}}$$

جہاں

r = Correlaton Coefficient

N = Number of pairs of Observation

1۔ اگر 'r' کی قیمت احتمالی خطا سے کم ہو تب کوریلیشن کا کوئی ثبوت نہ ہوگا یعنی 'r' کی قیمت کوئی اہمیت کی حامل نہیں ہے۔

2۔ اگر 'r' کی قیمت احتمالی خطا کا چھ گنا زیادہ ہو تب کوریلیشن کا کوفیشنٹ عملی طور پر یقینی ہے یعنی 'r' کی قیمت اہمیت کی حامل ہے۔

3۔ کوریلیشن کے کوفیشنٹ کی قیمت سے احتمالی خطا کی قیمت جمع اور تفریق کرنے سے ہم اندرون پاپولیشن متوقع کوریلیشن کے کوفیشنٹ میں بالترتیب اوپری اور نچلے حدود معلوم کر سکتے ہیں جہاں کہ کوریلیشن کا کوفیشنٹ واقع ہوگا۔

$$\text{P} = r \pm \text{P.E}$$

جہاں P(rho) ہے جو پاپولیشن میں کوریلیشن کو ظاہر کرتا ہے۔اب ہم احتمالی خطا معلوم کریں گے یہ فرض کرتے ہوئے کہ کوریلیشن کا کوفیشنٹ 0.8 اور ایٹمس کے سیمپل جوڑ 16 ہے۔ ہمیں دستیاب ہے۔

$$P = 0.6745\frac{1-(0.8)^2}{\sqrt{16}}$$

$$= 0.6745\frac{1-0.64}{4}$$

$$= 0.6745\frac{0.36}{4}$$

$$= 0.6745 \times 0.09$$

$$= 0.060705$$

پاپولیشن میں کوریلیشن کے حدود یوں ہوں گے $r \pm P.E$ یعنی $r \pm P.E$ یا 0.74-0.86 ایسی مثالیں بالکل نارمل ہیں جہاں کوریلیشن کا کوفیشنٹ 0.5 یا حتیٰ کہ 0.4 صریحاً ایک مصنف یا محقق کے لئے بھی کوریلیشن کو موزوں طور پر زیادہ تصور کیا جائے گا۔ باوجود یہ کہ کوریلیشن کا کوفیشنٹ 0.5 کا مطلب صرف 25% تغیر کو واضح کیا گیا۔ کوریلیشن کا کوفیشنٹ 0.4 کا مطلب صرف 16% تغیر کی توضیح کی گئی۔

احتمالی خطا کے استعمال کے لئے شرائط

احتمالی خطا کو موزوں طور پر صرف اسی وقت استعمال کیا جاسکتا ہے جب کہ درج ذیل تین شرائط کی تکمیل ہو۔

1. اعداد و شمار ایک حسب معمول (نارمل) تعددی قوس کے برابر ہو (گھنٹی کے شکل کی قوس)۔
2. شماریاتی پیمائش جس کے لئے PE محسوب کیا گیا ہو لازمی طور پر سیمپل سے تحسیب کیا گیا ہو۔
3. سیمپل کا انتخاب بے لاگ انداز میں ہوا ہو اور انفرادی ایٹمس لازمی طور پر آزاد ہوں۔

تاہم عمومی طور پر ان شرائط کی تکمیل ہوتی ہے جیسا کہ کوریلیشن کے کوفیشنٹ کی اعتباریت کا تعین بڑی حد تک اس کی موزونیت کی خارجی ٹیسٹس کی بنیاد پر ہوتا ہے جن کی اکثر شماریاتی حیثیت ہوتی ہے۔

تمثیل 15: اگر $r = 0.6$ اور $N = 64$ ہو، کوریلیشن کے کوفیشنٹ کی احتمالی خطا دریافت کیجئے اور آبادی 'r' کے لئے حدود کا تعین کیجئے۔

$$P.E_r = 0.6745\frac{1-r^2}{\sqrt{N}}$$
$$= 0.6745\frac{1-(0.6)^2}{\sqrt{64}}$$
$$= 0.6745\frac{1-0.36}{8}$$
$$= 0.6745\frac{0.08}{8}$$
$$= 0.6745 \times 0.08$$
$$= 0.05396$$

Limits of Population Correlation = 0.6 ± 0.054

= 0.546 to 0.654

7.10 درجاتی کوریلیشن کے معنی و تعریف (Meaning and Definition of Rank Correlation)

اسپیرمین درجاتی کوریلیشن دو متغیرات کے درمیان قوت اور سمت (مثبت / منفی) کی جانچ کاری ایک طریقہ ہے۔ شماریات میں درجاتی کوریلیشن کا کوفیشنٹ دراصل چارلس ایڈورڈ اسپیرمین کے نام پر لکھا گیا ہے۔ اس کو لاطینی زبان کے حرف (p (rho سے نشاندہی کی جاتی ہے۔ درجاتی کوریلیشن کا انحصار دو متغیرات کے درمیان درجات / درجہ کی پیمائش پر ہے۔ دو متغیرات کے درمیان اسپیرمین کی کوریلیشن ان دو متغیرات کے درمیان درجاتی قدروں کے برابر ہوتی ہے۔ اسپیرمین کی کوریلیشن ایک جیسے تعلق کا جائیزہ لیتی ہے (خطی یا غیر خطی)۔ اگر سلسلوں کے درمیان ایک ہی عدد کی بار بار تکرار نہ ہو تو اسپیرمین کی مکمل کوریلیشن (1+) پائی جاتی ہے یا جب متغیرات ایک دوسرے سے مکمل طور پر ایک جیسے ہوں تب (1-) مکمل کوریلیشن پائی جاتی ہے۔

تعریف (Definition)

اسپیرمین کی کوریلیشن کے کوفیشنٹ کی تعریف یوں کی گئی کہ: ''درجاتی متغیرات کے درمیان کارل پیرسن کی کوریلیشن کہتے ہیں''۔

ویکیپیڈیا(Wikipedia)میں یوں تعریف کی گئی ہے کہ :''شماریات کے کئی مختلف متغیرات کے درجات کے درمیان تعلق کی پیائش یاایک ہی متغیر کے مختلف درجات کی پیائش کودرجاتی کوریلیشن کہتے ہیں''۔

بعض اوقات ایسے مسائل سے سامنا ہوتا ہے جہاں ایک سلسلے کے مختلف ایٹمس کو سلسلہ وار ترتیب دیا جاسکتا ہے لیکن ان کی قدر مقداری پیائش مشکل ہوتی ہے۔ مثال کے طور پر ایک کلاس میں ٹیچر کے لئے یہ ممکن ہوتا ہے کہ وہ اپنے طلباء کو ذہانت کے اعتبار سے بڑھتی ہوئی اور گھٹتی ہوئی قدر میں ترتیب دے۔ اگرچیکہ ذہانت کی مقداری پیائش نہیں کی جاسکتی۔ بلاشبہ طلبا کی ذہانت کے بارے میں مقداری مطالعہ امتحان کے انعقاد اور نشانات(Infallible)دے کر کیا جاسکتا ہے لیکن اس طریقہ کو بھی بے خطا نہیں کہا جاسکتا۔ کئی وصف ایسے ہیں جن کی مقداری پیائش کی جاسکتی ہے مثال کے طور پر ایمانداری، کردار، اخلاق وغیرہ۔ اگر ایسے دو اوصاف کے مطالعہ کی ضرورت محسوس کی جائے جیسے ذہانت اور خوبصورتی کارل پیرسن کے کوریلیشن کے کوفیشنٹ کی تحسیب نہیں کی جاسکتی کیونکہ ان اوصاف کو ایک قطعی قدر سے متصف نہیں کیا جاسکتا۔ تاہم ایک ایسا طریقہ پایا جاتا ہے جس کی مدد سے ہم ایسے اوصاف کے درمیان کوریلیشن کا مطالعہ کر سکتے ہیں۔ اس طریقہ کو برطانیہ کے ایک ماہر نفسیات چارلس ایڈورڈ اسپیرمین نے 1904 میں ترتیب دیا۔

اس طریقہ میں 'x' اور 'y' متغیروں کو رینک (درجات) A اور B کی صورت میں دئے جاتے ہیں۔ ذہانت اور خوبصورتی کے درمیان کوریلیشن کے مطالعہ کی صورت میں ہم 10 یا 20 افراد کا انتخاب کر کے ان کی ذہانت کے اعتبار سے رینک دے دیں۔ سب سے زیادہ ذہین فرد کو 1 کا رینک دیں ایسے ہی ہر ایک کو رینک ایک کے بعد دیگرے دیتے ہوئے آخری فرد تک پہنچیں۔ اسی طرح ہم خوبصورتی کو بھی رینک دے سکتے ہیں بہت خوبصورت کو 1 رینک دیں ایسے ہی ہر ایک خوبصورتی کی مناسبت سے رینک دیتے ہوئے آخر تک پہنچیں جو کم سے کم خوبصورت ہے۔

اس طرح ہمارے پاس دو اوصاف کے لئے رینکس کے دو سیٹ ہوں گے اس طریقہ پر 10 افراد کو ترتیب دیا جائے تو ہمارے پاس 1-10 رینک ہوں گے۔ ایک مرتبہ رینک دینے کا کام ہو جائے تب اسپیرمینس رینک کوریلیشن کے ضابطہ سے ان دو سلسلوں کے لئے کوریلیشن کا کوفیشنٹ معلوم کر سکتے ہیں۔ یہ ضابطہ کچھ یوں ہے

$$r_s = 1 - \frac{6\Sigma d^2}{\mathrm{n}(\mathrm{n}^2 - 1)} or 1 - \frac{6\Sigma d^2}{n^3 - n}$$

r_s اسپیرمینس رینک کوریلیشن اور 'd' ایک فرد کے دو اوصاف کے لئے رینک کے فرق کو ظاہر کرتا ہے اور 'n' جوڑ کی تعداد کو ظاہر کرتا ہے۔ کوفیشنٹ کی قدر کی تشریح ایسے ہی ہو گی جیسے کارل پیرس کے کوریلیشن کے کوفیشنٹ ہوں گے۔ اس کی قدر کے حدود بھی 1+ اور 1- ہوں گے۔ جب 'r_s+' ہو اس کا مطلب یہ ہو گا دو اوصاف ذہانت اور خوبصورتی کے درمیان مکمل یکسانیت ہے اگر کوریلیشن 1- ہو اس کا مطلب یہ ہو گا کہ ذہانت اور خوبصورتی کے درمیان یکسانیت نہیں ہے۔

اس کوفیشنٹ کی تحسیب کاری کے سوالات دو قسم کے ہیں:

(A) جب کہ حقیقی رینکس دئے جائیں۔ (B) جب کہ حقیقی رینکس نہ دئے جائیں۔

ان دو قسم کے سوالات میں ہر ایک کے لئے مشکلات اس وقت ہوتی ہیں جب کہ دو افراد کے رینکس ایک ہی ہوں۔ ایسے سوالات کے لئے ضابطہ میں تبدیلی کی ضرورت ہوتی ہے۔

A. جب کہ رینکس دئے جائیں

جہاں ایٹمس کے حقیقی رینکس دئے جائیں تب کوفیشنٹ کو معلوم کرنے کے لئے درج ذیل اقدامات کئے جائیں۔

(i) رینکس کے درمیان فرق معلوم کریں(R_1-R_2)اور ان کی نشاندہی 'd' سے کیجئے۔

(ii) d_2 کی تحسیب کیجئے اور ان کا ٹوٹل کیجئے $\sum d^2$ حاصل ہوگا۔

(iii) کوفیشنٹ کو معلوم کرنے کے لئے ضابطہ استعمال کیجئے

$$r_s = 1 - \frac{6\Sigma d^2}{n^3 - n}$$

تمثیل 1: 15 طلباء کے ریاضی اور طبیعیات میں نشانات ذیل میں دئے گئے ہیں:

ریاضی میں رینکس	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
طبیعیات میں رینکس	10	7	2	6	4	8	3	1	11	15	9	5	14	12	13

رینک کوریلیشن کے کوفیشنٹ کو اسپیر مین کے ضابطے کو استعمال کرتے ہوئے کیجئے۔

ریاضی میں رینکس R_1	طبیعیات میں رینکس R_2	(R_1-R_1) d	d^2
1	10	-9	81
2	7	-5	25
3	2	+1	1
4	6	-2	4
5	4	+1	1
6	8	-2	4
7	3	+4	16
8	1	+7	49
9	11	-2	4
10	15	-5	25
11	9	+2	4
12	5	+7	49
13	14	-1	1
14	12	+2	4
15	13	+2	4
$n = 15$		$\Sigma d = 0$	$\Sigma d^2 = 272$

$$r_s = 1 - \frac{6\Sigma d^2}{n^3 - n}$$
$$= 1 - \frac{6 \times 272}{(15)^3 - 15}$$
$$= 1 - \frac{1632}{1632}$$
$$= 1 - \frac{1632}{3375 - 15}$$
$$= 1 - \frac{1632}{3360}$$
$$= 1 - 0.4857$$
$$= +0.514$$

تمثیل 2: ایک مقابلہ حسن میں تین ججوں نے درجے یوں دیئے:

پہلا جج	1	6	5	10	3	2	4	9	7	8
دوسرا جج	3	5	8	4	7	10	2	1	6	9
تیسرا جج	6	4	9	8	1	2	3	10	5	7

اسپیر مین کے ضابطہ سے رینک کوریلیشن محسوب کیجئے۔

حل:

R_1	R_2	R_3	$R_1 - R_2$ $d12$	$R_1 - R_3$ $d13$	$R_2 - R_3$ $d23$	d^2 12	d^2 13	d^2 23
1	3	6	-2	-5	-3	4	25	9
6	5	4	+1	+2	+1	1	4	1
5	8	9	-3	-4	-1	9	16	1
10	4	8	+6	+2	-4	36	4	16
3	7	1	-4	+2	+6	16	4	36
2	10	2	-8	0	+8	64	0	64
4	2	3	+2	+1	-1	4	1	1
9	1	10	+8	-1	-9	64	1	81
7	6	5	+1	+2	+1	1	4	1
8	9	7	-1	+1	+2	1	1	4
						200	60	214

$$r_S = 1-\frac{6\Sigma d^2}{n^3-n}$$
$$=1-\frac{6\times 14}{10^3-10}=1-\frac{1284}{100-10}$$
$$=1-\frac{284}{990}$$
$$=1-1.2969$$
$$=0.2896$$

مقابلہ حسن میں پہلے اور دوسرے جج کے درمیان کوریلیشن منفی ہے یعنی ان کی رائے ایک دوسرے سے مختلف ہے۔اسی طرح دوسرےاور تیسرے جج کی رائے بھی ایک دوسرے سے مختلف ہے۔ لیکن پہلے اور تیسرے جج کی رائے یکساں ہےان کا کوریلیشن مثبت ہے۔

تمثیل3: ایک کمپنی کے ڈبنچر کی قیمتوں اور حصص کی قیمتوں کے درمیان میں رینکس کوریلیشن+0.8 ہے اگر رینک میں فرق کے مربعوں کا مجموعہ 33 ہوتب 'n' کی قدر دریافت کیجئے۔

$$r_S = 1-\frac{6\Sigma d^2}{n^3-n} \qquad r_s=0.8 \qquad \Sigma d^2=33$$
$$0.8=1-\frac{6\times 33}{n^3-n} \ or\ 1-\frac{198}{n^3-n}$$
$$0.8=1-\frac{198}{n^3-n}=0.2(n^3-n)=198$$
$$n^3-n=\frac{198}{0.2}$$
$$n^3-n=990 \text{ یا } n^3-1000-n+10=0 \text{ یا}$$
$$\text{یا } n^3-10^3-(n-10)=0 \text{ یا } (n-10)(n^2+10n+100)-(n-10)=10$$
$$(n-10)(n^2+10n+99)=0\ \ i.e\, n\times 10 (\because a^3+b^3)=(a+b)(a^2+ab+b^2)$$
$$n^2+10n+99=0$$
$$n^2+10n=-99$$
$$n^2+n\ =$$
$$(n+0)[n^2+10n+100-1]=0$$
$$(n-10)[n^2+10n+99]=0$$
$$n-10=0\ (or)\ n^2+10n+99=0$$
$$n=10\ \ (or)\ n^2+10n+99=0$$

33 ہو تب 'n' کی قدر دریافت کیجئے۔

$$r_s = 0.8$$
$$\Sigma d^2 = 33$$

$$r_S = 1-\frac{6\Sigma d^2}{n^3-n}$$

$$0.8 = 1-\frac{6\times 33}{n^3-n} \ or\ 1-\frac{198}{n^3-n}$$

$$0.8 = 1-\frac{198}{n^3-n} = 0.2(n^3-n) = 198$$

$$n^3-n = \frac{198}{0.2}$$

$n^3 - n = 990$ یا $n^3 - 1000 - n + 10 = 0$ یا

$n^3 - 10^3 - (n-10) = 0$ یا $(n-10)(n^2+10n+100) - (n-10) = 10$ یا

$$(n-10)(n^2+10n+99) = 0 \ \ i.e\, n\times 10(\because a^3+b^3) = (a+b)(a^2+ab+b^2)$$

$$n^2+10n+99 = 0$$

$$n^2+10n = -99$$

$$n^2+n \ \ =$$

$$(n+0)[n^2+10n+100-1] = 0$$

$$(n-10)[n^2+10n+99] = 0$$

$$n-10 = 0\ (or)\ n^2+10n+99 = 0$$

$$n = 10\ \ (or)\ n^2+10n+99 = 0$$

اس گروہ میں ایٹمس کی تعداد دس ہے۔

تمثیل 4: شماریات اور اکاؤنٹنسی میں 10 طلباء کے حاصل کردہ نشانات کے رینک کوریلیشن کا کوفیشنٹ 0.2 پایا گیا۔ بعد میں پایا گیا کہ ایک طالب علم کے دو مضامین کے رینکس میں غلطی سے 7 کے بجائے 9 شمار کر لیا گیا۔ درست رینک کوریلیشن کا کوفیشنٹ دریافت کیجئے۔

اعداد دئے گئے ہیں:

$$n = 10 \quad r_s = 0.2$$

$$r_s = 1 - \frac{6\Sigma d^2}{n^3 - n}$$

$$0.2 = 1 - \frac{6\Sigma d^2}{10^3 - 10}$$

$$0.2 = 1 - \frac{6\Sigma d^2}{100 - 10} =$$

$$0.2 = 1 - \frac{6\Sigma d^2}{990}$$

$$\frac{6\Sigma d^2}{990} = 1 - 0.2$$

$$\frac{6\Sigma d^2}{990} = 0.8$$

$$6\Sigma d^2 = 0.8 \times 990$$

$$6\Sigma d^2 = 792$$

$$\Sigma d^2 = \frac{792}{6}$$

$$\Sigma d^2 = 132$$

یوں ہوگا Σd^2 صحیح $= 132 - 9^2 - 7^2$

$$= 132 - 81 + 49$$

$$= 100$$

$$correct\ r_s = 1 - \frac{6\Sigma d^2}{n^3 - n} = 1 - \frac{6 \times 100}{100 - 10} = 1 - \frac{600}{990}$$

$$= 1 - 0.60606$$

$$= 0.39$$

(b) جب کہ رینکس نہ دئے گئے ہوں

درجاتی کوریلیشن کا ضابطہ اس وقت استعمال کیا جاسکتا ہے جبکہ ہم متغیرات سے سامنا کر رہے ہوں جن کی مقداری پیمائش کی جاسکتی ہے۔ اگر ہمیں حقیقی اعداد و شمار دئے (رینکس نہ دیں) جائیں تب ہم ان اعداد کو رینکس میں تبدیل کر سکتے ہیں۔ اعلیٰ سے ادنیٰ تک رینکس دے دیں یعنی پہلے 1 ایک اور اس کے بعد دوسرے اعداد تک رینک دیا جائے۔ یہاں رینک بڑھتی ہوئی ترتیب یا پھر گھٹتی ہوئی ترتیب میں دیا

جائے۔دونوں متغیرات کے لئے یہی طریقہ کار اپنایا جائے۔

تمثیل 5: درج ذیل اعداد سے رینک کوریلیشن کا کوفیشنٹ تحسیب کریں۔

x:	60	34	40	50	45	41	22	43	42	66	64	46
y:	75	32	34	40	45	33	12	30	36	72	41	57

حل:

X	Rank 1	y	Rank 2	Diff. In Ranks (d)	d^2
60	3	75	1	+2	4
34	11	32	10	+1	1
40	10	34	8	+2	4
50	4	40	6	-2	4
45	6	45	4	+2	4
41	9	33	9	0	0
22	12	12	12	0	0
43	7	30	11	-4	16
42	8	36	7	+1	1
66	1	72	2	-1	0
64	2	41	5	-3	9
46	5	57	3	+2	4
n=12				0	$\sum d^2 48$

$$r_s = 1 - \frac{6\Sigma d^2}{n^3 - n}$$

$$= 1 - \frac{6 \times 48}{12(12^2 - 1)}$$

$$= 1 - \frac{288}{12(144 - 1)}$$

$$= 1 - \frac{288}{1728 - 12}$$

$$=1-\frac{288}{1716}$$

$$=1-0.16667$$

$$0.88$$

(c)یکساں رینکس Equal Ranks

بعض اوقات جہاں ایک سے زائد ایٹمس ہوں جن کی قیمت ایک ہو ایسے ایٹمس کو مشترکہ رینکس دئے جائیں گے۔ یہ رینک۔ رینکس کا اوسط ہوگا۔ جو ایک دوسرے سے کسی قدر مختلف ہوگا۔ جب یہ رینکس دے دئے جاتے ہیں رینک کو ریلیشن کے کوفیشنٹ کو مزید کو ریلیشن کی ضرورت ہوگی کیونکہ مذکورہ بالا ضابطہ مفروضہ کی بنیاد پر ہے مختلف ایٹمس کے رینکس مختلف ہیں اور رینک ایک سے زائد مرتبہ نہیں دیا جاسکتا۔

اگر ایک سلسلہ میں 'm' ایٹمس ہیں جن کا رینک مشترکہ ہے تب رینک کوریلیشن کے کوفیشنٹ میں تصیح ہوگی $\frac{1}{2}(m^3-m)$ کو Σd^2 کی قدر میں جوڑا جائے گا۔ اگر ایسے گروپس جن کا رینک مشترکہ ہو اور ایک سے زائد ہوں تب اس قدر کو جوڑتے جانا ہوگا۔ لہٰذا رینک کوریلیشن کے ضابطہ کو درج ذیل طریقہ سے تبدیل کرتے جانا ہوگا۔

$$r=1-\left[\frac{\Sigma d^2+\frac{1}{12}(m^3-m)+\frac{1}{12}(m^3-m).....}{n^3-n}\right]$$

تمثیل 6: درج ذیل اعداد کے لئے رینک کوریلیشن کی تحسیب کیجئے۔

x:	48	33	40	9	16	16	65	24	16	57
y:	13	13	24	6	13	4	20	9	6	19

حل:

X	Rank 1	y	Rank 2	Diff. In Ranks (d)	d^2
48	3	13	5.5	-2.5	6.25
33	5	13	5.5	-0.5	0.25
40	4	24	1	+3.0	9.00
9	10	6	8.5	+1.5	2.25

16	8	15	4	+4.0	16.00
16	8	4	10	-2.0	4.00
65	1	20	2	-1.0	1.00
24	6	9	7	-1.0	1.00
16	8	6	8.5	0.5	0.25
57	2	19	3	1.0	1.00
n=12				0	$\sum d^2=41.00$

مذکورہ بالا جدول میں x سلسلے میں عدد 16 تین مرتبہ آیا ہے ان تمام ایٹمس کا رینک 8 جو 8,7 اور 9 کا اوسط ہے اعداد جن کو یہ رینک دیا گیا ہے اس کا مطلب یہ ہوگا کہ ان قیمتوں میں کچھ فرق پایا گیا ہے۔ y سلسلے میں عدد 13 اور 6 دونوں دو مرتبہ واقع ہوئے ہیں۔ ان کے رینکس بالترتیب 5.5 اور 8.5 ہیں۔ ان مشترکہ رینکس کی وجہ سے رینک کوریلیشن کو درست کرنا ہوگا۔

ان کی تصحیح کے لئے ہمیں $\left[\frac{1}{2}(m^3-m)\right]$ کو Σd^2 کی قدر میں جوڑنا ہوگا۔ x سلسلے میں اس کی قدر $\frac{1}{2}(2^3-2)$ کے برابر ہوگی کیونکہ عدد 16 اس سلسلے میں تین مرتبہ دہرایا گیا ہے۔ y سلسلوں میں مشترکہ رینک رکھنے والے دو ایسے گروپس ہیں۔ پہلے گروپ میں یہ تصحیح $\left[\frac{1}{2}(m^3-m)\right]$ کیونکہ 13 دو مرتبہ دہرایا گیا ہے اور دوسرے گروپ میں تصحیح کی قدر $\frac{1}{2}(2^3-2)$ کے برابر ہوگی کیونکہ اس سلسلے میں عدد 6 دو مرتبہ دہرایا گیا ہے۔

$$r_s = 1 - \frac{6\left[\Sigma d^2 + \frac{1}{12}(m^3 - m) + \frac{1}{2}(m^3 - m) + \frac{1}{12}(m^3 - m)\right]}{n^3 - n}$$

$$= 1 - \frac{6\left[41 + \frac{1}{12}(3^3 - 3) + \frac{1}{2}(2^3 - 2) + \frac{1}{12}(2^3 - 2)\right]}{10^3 - 10}$$

$$= 1 - \frac{6\left(41 + \frac{1}{12}(27 - 3) + \frac{1}{2}(6 - 2) + \frac{1}{12}(6 - 2)\right)}{100 - 10}$$

$$= 1 - \frac{6\left(41 + \frac{1}{12_1}\overset{2}{24} + \frac{1}{12_1} \times \overset{1}{6} + \frac{1}{12_2}\overset{1}{6}\right)}{100 - 10}$$

$$= 1 - \frac{6\left(41 + 2 + \frac{1}{2} + \frac{1}{2}\right)}{990}$$

$$= 1 - \frac{6(44)}{990}$$

$$= 1 - \frac{264}{990}$$

$$= 1 - 0.2667$$

$$= 0.738$$

7.11 اسپیرمین کوریلیشن کے کوفیشنٹ کے خدوخال

(Features of Spearman's Coefficient of Correlation)

1. دو متغیروں کے درمیان رینکس کے فرق کا مجموعہ صفر ہو گا علامتی طور پر $\sum d=0$ ۔

2. اسپیرمین کا کوریلیشن کا کوفیشنٹ نان پیرامیٹرک ( ) ہے کیونکہ کسی زیر مشاہدہ نمونہ سے آبادی کی شکل کے بارے میں مشکل مفروضات نہیں لئے جاتے۔

3. اسپیرمین کوریلیشن کے کوفیشنٹ کچھ اور نہیں بلکہ کارل پیرسن کے کوریلیشن کا کوفیشنٹ ہے رینکس کے درمیان ہے لہٰذا اس کی تشریح بھی کارل پیرسن کی تشریح کی طرح کی جائے گی۔

7.12 رینک کوریلیشن کے فائدے اور نقصانات

(Advantages and Disadvantages of Rank Correlation)

فائدے

1. چونکہ اس طریقے میں $\sum d$ یا R_1 اور R_2 میں فرق ہمیشہ صفر کے برابر ہوتا ہے اس لئے تحسیب پر روک فراہم کرتا ہے۔

2. چونکہ اسپیرمین کے رینک کوریلیشن وہی ہے جو رینکس کے درمیان کارل پیرسن کے کوریلیشن کے کوفیشنٹ کے برابر ہے اس لئے اس کی تشریح بھی کارل پیرسن کے کوریلیشن کے کوفیشنٹ کی طرح کی جاتی ہے۔

3. رینک کوریلیشن کارل پیرسن کے کوریلیشن کے کوفیشنٹ کے برخلاف یونیورس میں نارمل حالات کے متقاضی نہیں ہوتے جس میں سے نمونہ سیمپل لیا گیا تھا۔

4. رینک کوریلیشن سمجھنے اور اطلاق کے لئے نہایت آسان ہے۔ تاہم پیرسن کا کوفیشنٹ مکمل معلومات کی بنیاد پر ہوتا ہے جب کہ اسپیرمین کا کوفیشنٹ صرف رینک کی بنیاد پر ہوتا ہے۔ r کی قیمتیں عموماً دونوں طریقوں میں الگ الگ ہوتی ہیں۔

5. اسپیرمین کے رینک کا طریقہ ہی کوریلیشن کے مطالعہ کا ایک راستہ ہے جو معیاری اعداد و شمار کے درمیان کیا جاتا ہے جس کی پیمائش اعداد میں نہیں کی جاسکتی لیکن سلسلہ نشان میں ترتیب دیا جاسکتا ہے۔

نقصانات

1. یہ طریقہ دوہرے تعددی جدول میں استعمال نہیں کیا جاسکتا۔ یا Bivariate frequency distribution میں کی صورت میں استعمال نہیں کیا جاسکتا۔

2. اس کا استعمال سہولت بخش طریقہ پر صرف اسی وقت کیا جاسکتا ہے جب کہ n چھوٹا ہو ورنہ حساب دہی تکلیف دہ یا مشقت طلب ہو جاتی ہے۔

7.13 رینک کوریلیشن کے کوفیشنٹ کے استعمالات (Uses of Coefficient of Correlation)

1. ابتدائی اعداد و شمار رینکس کی صورت میں ہوں۔

2. اگر تعداد کم ہو (جیسے 25 یا 30) بعض اوقات داخلی اعداد و شمار پر Apply کیا جاتا ہے جہاں مشابہت پیدا کرنا ہو جب کہ r کی قیمت کی تحسیب زیادہ وقت طلب ہو۔ اس کے لئے Interval data کو دونوں متغیرات کو رینکس کی ترتیب میں منتقل کیا جاتا ہے۔ اگر 'N'30 سے زیادہ ہو تو اس کو دس کی رینکنگ مشقت طلب اور اس فارمولے کے ذریعہ متوقع وقت کی بچت کے خلاف ہو جاتی ہے۔

7.14 اکتسابی نتائج(Learning Outcome)

طلباء نے باہمی تعلق کے معنی، تعریف، اور نوعیت کی مکمل تفہیم تیار کی ہے اور یہ کہ یہ متغیر کے درمیان تعلق کو کیسے ماپتا ہے۔ وہ ارتباط اور سبب کے درمیان فرق کر سکتے ہیں اور ڈیٹا پیٹرن کا تجزیہ کرنے میں ارتباط کی ضرورت کو سمجھ سکتے ہیں۔ طلباء نے باہمی ربط کی اقسام اور کارل پیئرسن کے ارتباط کے گتانک کی نوعیت اور مفروضات سیکھ لیے ہیں، ساتھ ہی یہ بھی سیکھ لیا ہے کہ ارتباط کے گتانک کی تشریح کیسے کی جائے اور ممکنہ غلطی کے ساتھ اس کے تعلق کو کیسے سمجھا جائے۔ مزید برآں، انہوں نے درجہ کے ارتباط کے معنی اور تعریف، اسپیئرمین کے گتانک کی خصوصیات، اور اس کے فوائد اور نقصانات کو دریافت کیا ہے۔ مجموعی طور پر، سیکھنے والے ڈیٹا میں تعلقات کا تجزیہ کرنے اور اعداد و شمار کے ثبوت کی بنیاد پر باخبر فیصلے کرنے کے لیے پیئرسن اور اسپیئرمین کے باہمی تعلق کی تکنیکوں کو لاگو کرنے کے لیے لیس ہیں۔

7.15 نمونہ امتحانی سوالات(Model Exam Questions)

معروضی جوابات کے حامل سوالات(Objective Answer Type Questions- True of False)

1. Correlation کا مطلب ہے کہ دو یا دو سے زیادہ متغیرات کے درمیان تعلق یا رشتہ موجود ہے۔
2. شماریات میں ارتباط کی کوئی اہمیت نہیں ہے۔
3. Positive اور Negative correlation ارتباط کی دو اقسام ہیں۔
4. Pearson اور Spearman کے طریقے ارتباط کو ماپنے کے دو طریقے ہیں۔
5. Correlation ہمیشہ causation یعنی سبب و اثر ظاہر کرتا ہے۔

مختصر جوابات کے حامل سوالات(Short Answers Type Questions)

1. Correlation کا کیا مطلب ہے؟
2. شماریات میں ارتباط کیوں اہم ہے؟
3. کسی بھی دو قسم کے ارتباط کا ذکر کریں۔
4. ارتباط کی پیمائش کے کوئی دو طریقے لکھیں۔

طویل جوابات کے حامل سوالات(Long Answers Type Questions)

1. شماریاتی تجزیہ میں ارتباط کے معنی اور اہمیت کی وضاحت کریں۔
2. مثالوں کے ساتھ ارتباط کی مختلف اقسام بیان کریں۔
3. ارتباط کی پیمائش کے لیے استعمال ہونے والے کسی بھی دو طریقوں کی وضاحت کریں۔

اکائی 8۔ ریگریشن کا تجزیہ

(Regression Analysis)

اکائی کے اجزاء

8.0 تمہید Introduction

8.1 مقاصد Objectives

8.2 ریگریشن کا مفہوم، تعریف اور نوعیت Meaning, Definition and Nature of Regression

8.3 ریگریشن کے تجزیہ کے استعمالات Uses of Regression Analysis

8.4 کوریلیشن اور ریگریشن کے تجزیہ میں فرق Difference between correlation and Regression analysis

8.5 ریگریشن کے مطالعہ کے طریقے Methods of Studying Regression

8.6 حسابی اوسط سے انحراف کا طریقہ Methods of Deviations from Mean

8.7 اکتسابی نتائج

8.8 نمونہ امتحانی سوالات Model Examination Questions

8.0 تمہید (Introduction)

دو متغیروں کے درمیان اس حقیقت کا انکشاف ہونے کے بعد کہ دونوں میں قریبی تعلق پایا جاتا ہے، ہم اس بات میں دلچسپی رکھیں گے کہ ایک معلوم قیمت دینے کے بعد دوسری نامعلوم قیمت کی پیش بینی کریں۔ مثال کے طور پر اگر ہم جانتے ہیں کہ اشتہار کاری اور فروخت دونوں کوریلیٹیڈ ہیں تب ہم کو دی گئی تشہیر پر خرچ شدہ رقم کی مدد سے فروخت کی رقم معلوم کر سکتے ہیں۔ اسی طرح اگر ہم یہ جانتے ہیں کہ چاول کی پیداوار، بارش دونوں کوریلیٹیڈ ہیں تب ہم معینہ چاول کی پیداوار کے لئے بارش کی مقدار کتنی ہونی چاہئے اس کا اندازہ لگا سکتے ہیں۔ تجزیہ ریگریشن دو متغیرات کے درمیان اوسط تعلق کا انکشاف کرتا ہے یہی وہ عنصر ہے جو تخمینہ یا پیش بینی کو ممکن بناتا ہے۔

8.1 مقاصد (Objectives)

اس اکائی کے مطالعہ کے بعد آپ اس قابل ہو جائیں گے کہ:

- ریگریشن کا تجزیہ کی تعریف کر سکیں۔
- ریگریشن کا تجزیہ کے استعمالات سے واقف ہو سکیں۔
- کوریلیشن اور ریگریشن کا تجزیہ میں فرق کر سکیں۔
- ریگریشنی خطوط اور مساوات سے واقف ہو سکیں۔
- تخمینہ کی معیاری خطا کی تحسیب کر سکیں۔
- تجزیہ ریگریشن کی خامیوں سے واقف ہو سکیں۔

8.2 ریگریشن کے تجزیہ کا مفہوم، تعریف اور نوعیت

(Meaning, Definition and Nature of Regression)

اصطلاح Regression کے معنی ڈکشنری میں "عمل ریگریشن" (Act of Returning) واپسی (Going back) کے ہیں۔ اصطلاح Regression کو سب سے پہلے سر فرانسس گالٹن نے 1877 میں باپ اور بیٹوں کے قد کے درمیان تعلق کے مطالعہ کے دوران استعمال کیا۔ اس نے اس لفظ کو اپنے ایک مضمون "Regression towards Mediocrity in Herditary Stature" میں استعمال کیا۔ "قد" کے بارے میں اس کے مطالعہ سے ایک ہزار "باپ اور بیٹوں" سے ایک دلچسپ تعلق کا انکشاف ہوا یعنی لانبے باپ لانبے بیٹے اور پستہ قد باپ کے پستہ قد بیٹے ہوتے ہیں لیکن لانبے قد کے والدین کے گروپ کے بیٹوں کا اوسط قد اپنے والد سے کم نکلا اور پستہ قد والد کے بیٹوں کا قد اپنے والد سے زیادہ تھا۔ وہ خط جو عمل ریگریشن کے رجحان یا واپسی کو گالٹن نے "Regression line" کہا ہے۔ یہ اصطلاح آج بھی اس خط کو بیان کرنے کے لئے استعمال کی جاتی ہے جو نقاط کے گروہ کے رواں رجحان کی نمائندگی کے لئے کھینچا گیا ہو لیکن یہ تصور زیادہ دنوں تک باقی نہیں رہا کیونکہ یہ گالٹن کے Stepping back تصور کردہ اصل اثرات کا حامل نہیں تھا۔ ان دنوں جدید مصنفین میں Regression line کے بجائے Estimating line کے استعمال کرنے کے رجحان میں اضافہ ہوا ہے کیونکہ اصطلاح Estimating line اپنے اندر جامعیت رکھتی ہے۔

تعریف:

اصطلاح Regression کا مفہوم اچھی طرح سمجھنے کے بعد اب اس کی تعریف پر ایک اچٹتی ہوئی نظر ڈالیں گے۔ مختلف مصنفین نے اس کی تعریف الگ الگ انداز سے کی ہے۔

1. Taro Yamane کے ریگریشن انالائیسنس کی تعریف ان الفاظ میں کی ہے "دو یا دو سے زیادہ متغیرات کے درمیان موجبانہ تعلق کو دریافت کرنے کے لئے معاشیات اور کاروباری تحقیق میں اکثر استعمال کی جانے والی ٹیکنکس میں سے یہ ایک ہے۔"

2. Ya-Lun Chou کے مطابق Regression Analysis متغیرات کے درمیان تعلقات قائم کرنے کی کوشش ہے جیسے متغیرات کے درمیان تفاعلی تعلق کا مطالعہ ہے اسی وجہ سے پیشن گوئی میکانزم فراہم کرتا ہے۔

مذکورہ بالا تعریفات سے یہ بالکل صاف ہے کہ ریگریشن انالیسیس ایک آلہ ہے جس کی مدد سے ہم معلوم قدر سے دوسری نامعلوم قدر معلوم کرنے کے موقف میں ہوتے ہیں۔ وہ متغیر جس کو نامعلوم متغیر کو پیش بینی کرنے کے لئے استعمال کیا جاتا ہے اس کو Independant variable یا Explanatory variable کہا جاتاہے اور جس متغیر کو ہم پیش بینی کرنے کی کوشش کررہے ہیں اس کو Dependant variable explained variable کہا جاتاہے۔ Independant variable کو x اور Depended variable کو y سے تعبیر، موسوم کیا جاتاہے۔ استعمال کردہ تجزیہ سادہ خطی ریگریشنی تجزیہ (Simple linear regression analysis) کہا جاتا ہے۔ یہ سادہ یاآسان اس لئے ہے کہ یہاں صرف ایک ہی پیش گو یا آزاد متغیر ہوتاہے اور خطی (Linear) اس لئے کہ خود منحصر (Depended) اور آزاد متغیرات کے درمیان مفروضہ خطی تعلق ہوتاہے۔ اصطلاح Linear کا مطلب خط مستقیم کی وہ مساوات ہے جو $y=a+bx$ کو ترتیب دیتی ہے۔ جہاں a اور b اٹل ہیں جن کو دو متغیرات کے درمیان موجود اوسط تعلق کو بیان کرنے کے لئے استعمال کیا جاتاہے۔

یہ بات نوٹ کی جائے کہ اصطلاحات Dependent اور Independent دونوں Dependence کے ریاضیاتی اور تفاعلی معنی مفہوم سے متعلق ہیں۔ اس کا مطلب ہر گز یہ نہیں کہ متغیرات کے درمیان کوئی لازمی طور پر وجہ اور اثر کا تعلق ہے اس کا مطلب سادہ طور پر یہ ہے کہ Dependent متغیر y کی قیمتوں کا تخمینہ معلوم کیا جاسکتاہے بشرطیکہ x Independent variable کی قیمتیں تفاعل ریاضی میں x اور y کو شامل کرکے دی جائیں ایسی صورت میں y کی قیمتوں کا انحصار x کی قیمتوں پر ہوگا۔ چاہے متغیر x، متغیر y کی تبدیلی باعث ہو یا نہ ہو۔ مثال کے طور پر تشہیر پر خرچ کے اعداد و شمار پر ایک شئے کی فروخت کے تخمینہ کرنے کے دوران فروخت کو عموماً Dependent متغیر کے طور پر لیا جاتاہے تاہم تشہیر پر خرچ میں تبدیلی کی وجہ سے فروخت میں تبدیلی کے تناظر میں ان دو عناصر میں موقعتی بندش ہو بھی سکتی ہے اور نہیں بھی۔ حقیقت میں بعض صورتوں میں وجہ۔ اثر تعلق بظاہر نظر آنے والے تعلق کے برعکس ہوتاہے۔ سادہ ریگریشنی تجزیہ میں دو متغیرات ہوتے ہیں ایک کو Independent variable یا Regressor یا Predector یا Explanator کہتے ہیں اس متغیر کی قیمتوں کی بنیاد پر دوسرے متغیر کی پیش بینی کی جاتی ہے۔ یہ دوسرا متغیر جس کی قیمتوں کی پیش بینی کی جاتی ہے اس کو Dependent یا Regressed یا Explained متغیر کہا جاتاہے۔

ریگریشن کا مطالعہ صرف دو متغیرات کی حدتک محدود ہو تو اس کو Simple regression اور اگر ریگریشنی تجزیہ کا مطالعہ وقت واحد میں دو سے زاید متغیرات پر ہو تو اس کو Multiple regression کہا جاتاہے۔

8.3 ریگریشن کے تجزیہ کے استعمالات (Uses of Regression Analysis)

ریگریشنی تجزیہ شماریاتی نظریہ کی ایک شاخ ہے جس کا استعمال تمام عصری شعبہ حیات میں زیادہ سے زیادہ استعمال کیا جاتاہے۔ معاشیات میں معاشی نظریہ اور معاشی زندگی کو ترتیب دینے والے معاشی متغیرات کے درمیان تعلق کی پیائش کاری یا تخمینہ کاری کے لئے یہ ایک بنیادی ٹیکنک ہے مثال کے طور پر اگر ہم دو متغیرات قیمت x اور طلب y کا ایک دوسرے سے قریبی تعلق ہے تب ہم x کی قیمت دریافت

کر سکتے ہیں جب کہ y کی قیمت دی جائے یاy کی قیمت معلوم کر سکتے ہیں جب کہ x کی قیمت دی جائے۔ ٹھیک اسی طرح اگر ہم ٹیکس اور شئے کی قیمت میں اضافہ کا قریبی تعلق ہو تب ہم ٹیکس کی رقم کو مقدار یا شرح کی مناسبت سے شئے کی متوقع قیمت معلوم کر سکتے ہیں۔ چنانچہ ریگریشن کے مطالعہ سے ہمیں پتہ چلتا ہے کہ یہ ماہرین معاشیات اور تاجرین کے لئے کافی مددگار ہے۔ ریگریشن کے استعمالات صرف ماہرین معاشیات اور تاجرین کے لئے ہی محدود نہیں ہیں بلکہ اس کا اطلاق قدرتی، طبعی اور سماجی سائنس تک ہوتا ہے اس کے فائدے مختصراً ذیل میں دیکھے جا سکتے ہیں۔

1. ریگریشنی تجزیہ Dependent متغیر کی تخمینی قیمت Independent متغیر سے فراہم کی جاتی ہے اس تخمینی قیمت کے حصول کے لئے مستعملہ آلہ ریگریشنی خط (Regression line) ہے ریگریشنی خط x اور y متغیرات کے درمیان موجود وسطی تعلق کو بیان کرتا ہے یعنی یہ y کی دئی گئی قیمتوں کے لئے x کی اوسط قیمتیں نمایاں کرتا ہے۔اس خط کی مساوات کو ریگریشنی مساوات (Regression equation) کہتے ہیں۔ یہ مساوات Dependent متغیر کا تخمینہ فراہم کرتی ہے جب کہ Independent متغیر کو مساوات میں درج کیا جائے۔

2. ریگریشن تجزیہ کا دوسرا استعمال ریگریشن لائین کے بطور اساس استعمال میں درآئیں غلطیوں کا سد باب کرنا۔اس مقصد کے لئے تخمینہ کی معیاری خطا (Standard error) کی تحسیب کرنا۔

3. ریگریشن کوایفیشینٹس کی مدد سے ہم کوریلیشن کوایفیشینٹ کی تحسیب کر سکتے ہیں کوریلیشن کوایفیشینٹ (r) کا مربع Coefficient of determination کہلاتا ہے۔ یہ دو متغیرات کے درمیان کوریلیشن کے اتحاد کے زاویہ کی پیائش کرتا ہے۔ ریگریشنی مساوات کے پیش کردہ Dependent variable کے ویرینس تناسب کا جائزہ لیتی ہے۔ نارمل طور پر... کی قدر زیادہ ہو تو ارتسام کے لئے بہتر اور Predictive device کے طور پر ریگریشن مساوات اتنی زیادہ بہتر ہو گی۔

8.4 کوریلیشن اور ریگریشن کا تجزیہ میں فرق

(Difference between correlation and Regression analysis)

کوریلیشن اور ریگریشن تجزیہ دونوں مختلف مفروضات کے تحت ترتیب دئے گئے ہیں۔ یہ مختلف قسم کی معلومات بہم پہنچاتے ہیں لیکن ہمیشہ یہ واضح نہیں ہوتا کہ کونسا پیمانہ دئے گئے حالات کے لئے کب اور کیوں استعمال کیا جائے۔ان دونوں کے درمیان درج ذیل فرق پایا جاتا ہے۔

1. چونکہ کوفیشنٹ 'Coefficient 'x اور 'y' کے درمیان Covariability کے زاویہ کی پیائش ہے ریگریشن تجزیہ کا مقصد متغیرات کے درمیان تعلق کی نوعیت کا مطالعہ ہے تاکہ ہم ایک کی قیمت کی پیشن گوئی دوسری قیمت کی بنیاد پر کر سکیں۔ دو متغیرات کے درمیان جتنا مضبوط تعلق ہو گا اتنا ہی تخمینہ جات پر اعتماد کیا جا سکتا ہے۔

.2 کوریلیشن دو متغیرات کے درمیان تعلق کے زاویہ کی دریافت کے لئے صرف ایک آلہ ہے اور اس لئے ہم یہ نہیں کہہ سکتے کہ ایک متغیر وجہ(Cause)اور دوسرا اثر(Effect) ہے۔ مثال کے طور پر ایک مخصوص مدت یا ایک شئے کے لئے قیمت اور طلب کے درمیان کوریلیشن زیادہ ہو تب یہ نہیں کہا جاسکتا کہ قیمت وجہ اور طلب اثر ہے۔ تاہم ریگریشن تجزیہ میں ایک متغیر کوDependentجب کہ دوسرے کو Independent کے طور پر لیا جاسکتا ہے۔ اس طرح یہاں وجہ اوراثر تعلقات کا مطالعہ ممکن ہے۔ یہاں یہ نوٹ کیا جائے گا کہ اتحاد(Association)کی موجودگی Causationتسبیب کا باعث نہیں لیکن جہاں تسبیب(Causation)ہو وہاں اتحاد لازمی طور پر ہوتا ہے۔ متغیرات کے درمیان اتحاد کی موجودگی یا غیر موجودگی کا پتہ شماریاتی ثبوت سے لگایا جاسکتا ہے چاہے تسبیب پائی جائے یا نہ پائی جائے۔ اس کا انحصار خالصتاً استدلال(Reasoning)پر ہوتا ہے۔

.3 کوریلیشن تجزیہ میں xy سمت کی پیمائش اور دو متغیرات xاور yکے درمیان خطی تعلق(Linear relationship)کا زاویہ ہوتا ہے۔ rxyاور ryx متناسب، متشاکل(Symmetric)ہیں۔($r_{xy} = r_{yx}$)یعنی xاور yمیں کون Dependent متغیر اور کون Independent متغیر ہے کوئی معنی نہیں رکھتا۔ ریگریشن تجزیہ میں ریگریشن کوفیشنٹ bxyاور byx متناسب Symmetric نہیں ہیں یعنی bxy ≠ byxاور اس لئے یہاں کونسا متغیر Dependent ہے اور کونسا متغیر Independent ہے کافی فرق پڑ جاتا ہے۔

.4 کبھی کبھار دو متغیرات کے درمیان بے معنی، مہمل کوریلیشن بھی پایا جاتا ہے جس کا کوئی عملی تعلق نہیں ہوتا۔ جیسے گروپ کی آمدنی میں اضافہ یا وزن میں اضافہ۔ تاہم بے معنی/ مہمل ریگریشن نہیں ہوتا۔

.5 کوریلیشن کا کوفیشنٹ پیمانے اور مبدے میں تبدیلی سے آزاد ہوتا ہے۔

ریگریشن اور کوریلیشن تجزیہ میں کچھ مشترکہ ہے تو یہ ہے کہ کوریلیشن کا کوفیشنٹ(r)اور ریگریشن کوفیشنٹ(bxy)اور(byx) دونوں کی علامت(r)ایک ہی مشترک ہے۔ اگر دیئے گئے Level of significanceپرbکی قدر نمایاں ہو تو اسی سطح پرr بھی نمایاں ہوگا۔

متغیرات میں تعلق خطی(Linear)اور غیر خطی(Non linear)ہو سکتا ہے۔ دو متغیرات میں خطی تعلق اس وقت ہوگا جب کہ Independent variableیعنی xمیں ایک اکائی کی تبدیلی ہو تو 'Dependent variable 'yمیں بھی اتنی ہی تبدیلی ہوگی۔ جب دو متغیرات میں خطی تعلق ہو تب ریگریشن خطوط کو Dependent variableکی قیمتوں کو دریافت کرنے کے لئے استعمال کیا جاسکتا ہے جب ہم دو متغیرات (xاور y) کو ایک منتشر ڈیاگرام پر مرتسم کرتے ہیں اور دو خطوط کھینچتے ہیں جو مرتسم نقاط سے گذرتے ہیں تو ان خطوط کو Regression linesکہا جاتا ہے۔ Linear Regressionمیں یہ خط سیدھے ہوتے ہیں یہ خطوط دو مساوات کی بنیاد پر ہوتے ہیں انہیں Regression equationsکہتے ہیں۔ یہ مساوات ایک متغیر کی معلوم قیمت سے نامعلوم قیمت کو دریافت کرنے پر صحیح تخمینہ کا باعث ہوتی ہے۔

8.5 ریگریشن کے مطالعہ کے طریقے (Methods of Studying Regression)

ریگریشن کے مطالعہ کے دو طریقے ہیں:

(i) ترسیمی طریقہ ((ii Graphically)) الجبرائی طریقہ (Algebraically)

(i) ریگریشن کے مطالعہ کے ترسیمی طریقہ Graphic Study of Regression

جب ریگریشن کا مطالعہ ترسیمی طریقہ سے کیا جاتا ہے تو ہمیں منتشر ڈیا گرام (Scatter diagram) کھینچنا ہوگا۔ ایک منتشر ڈیا گرام x اور y متغیرات کے ایک جوڑ مقداروں کے لئے ایک پوائنٹ پر مشتمل ہوتا ہے۔ عموماً x متغیر کو افقی پیمانے (Horizontal scale) اور y متغیر کو عمودی پیمانے (Vertical scale) پر ظاہر کیا جاتا ہے۔ جب تمام متعلقہ قیمتوں، مقداروں کو منتشر ڈیا گرام پر مرتسم کر دیا جاتا ہے تو ہمیں دو ریگریشن خطوط کھینچنے پڑتے ہیں تا کہ متغیرات x اور y کی مقداروں کی پیش گوئی کی جا سکے۔ وہ ریگریشن خط جس کو x کی قیمت، مقدار پر y کی مقداروں کی پیش گوئی کے لئے استعمال کیا جاتا ہے x پر y کا ریگریشن خط (Regression line of y on x) کہلاتا ہے۔ اسی طرح ریگریشن جس کو y کی قیمت، مقدار پر x کی مقداروں کی پیش گوئی کے لئے استعمال کیا جاتا ہے y پر x کا ریگریشن (Regression line of x on y) کہلاتا ہے۔ x اور y کے درمیان کوریلیشن کا کوفیشنٹ مکمل، کامل ہو یعنی اس کی قیمت 1+ یا 1- ہو تب صرف ایک ریگریشن لائین ہو گی کیونکہ ایسی صورتوں میں دو سلسلوں میں تغیرات ہمیشہ ایک اٹل عدد سے اضافہ یا کم کی صورت میں پائے جاتے ہیں۔

جب ہم Independent variable میں تبدیلیوں کے تناسب Dependent variable میں تبدیلیوں کو ظاہر کر رہے ہوتے ہیں تو سیدھے خطوط کھینچیں گے تب اس کو Linear Regression کہا جاتا ہے۔ تاہم اگر دو متغیرات کے درمیان تعلق خط مستقیم کی طرح نہ ہو بلکہ کوئی دوسرے تفاعلی تعلق ہو جیسے $y = x^2$ ایسی صورت میں متغیرات کے درمیان Non-linear Regression کا ہوتا ہے۔ زیر مطالعہ باب میں ہم صرف Linear Regression کا مطالعہ کر رہے ہیں۔

Linear Regression Lines کو کھینچیں۔ ان خطوط کو یوں کھینچا جا سکتا ہے۔

(a) حسب منشا قوس کا طریقہ (Free hand curve method)

(b) اقل ترین مربعوں کا طریقہ (Method of least squares)

(a) حسب منشا قوس کا طریقہ Free hand Curve Method

اس طریقہ میں ہم x اور y کی قیمتوں کے جوڑوں کو منتشر ڈیا گرام کی شکل میں مرتسم کرتے ہیں۔ قیمتوں کے ایک جوڑ کے لئے ایک پوائنٹ مقرر ہوتا ہے۔ ان خطوط کے منجملہ ایک خط کو کچھ اس انداز سے کھینچا جاتا ہے کہ y سلسلے کے پازیٹیو انحراف جو اس کے اوسط حسابیہ سے لئے جاتے ہیں منفی (Negative) انحراف سے تنسیخ ہو جائے خط کے ایک جانب کے انحرافی اعداد کا مجموعہ دوسری جانب کے انحرافی اعداد کے مجموعے کے برابر، مساوی ہو۔ یہ x پر y کا ریگریشن خط ہو گا۔

دوسری ریگریشن لائین کچھ اس انداز سے کھینچی جائے گی x سلسلے کے پازیٹیوانحرافی اعداد جو اس کے اوسط حسابیہ سے لئے جاتے ہیں منفی (Negative) انحرافی اعداد سے تنسیخ، قطع ہو جائے اس ریگریشن لائین کو y پر x ریگریشن لائین کہا جائے گا۔ یہ دو ریگریشن لائین ایک دوسرے کو اوسط حسابیہ کے نکتہ پر قطع کریں گے۔ دونوں متغیرات کے درمیان کامل پازیٹیو یا نیگیٹیو کوریلیشن ہو تب صرف ایک ہی ریگریشن لائین ہو گی۔

تاہم حسب منشا قوس کے طریقہ سے ریگریشن لائین کھینچنا بہت مشکل ہوتا ہے۔ عموماً منتشر ڈیا گرام میں دھاگہ کے ایک ٹکڑے کو کچھ اس انداز سے مسلسل ایڈجسٹ کیا جاتا ہے کہ پازیٹیو اور نیگیٹو انحرافی اعداد ایک دوسرے کو قطع کریں۔ ایک مرتبہ جب قوسین کھینچ لی جاتی ہیں تب ہم ریگریشن لائین x پر y کی قیمتوں کی پیش گوئی کی جاسکتی ہے۔ اسی طرح x کی قیمتوں کو y پر x ریگریشن لائین سے پیش گوئی کی جاسکتی ہے۔

Methods of least squares

حسب منشاء قوس کے طریقہ میں ریگریشن لکیر کھینچنے میں حائل دشواریوں سے بچنے کے لئے x اور y سیریز کے حرکات کے درمیان ایک ریاضیاتی تعلق قائم کیا جاتا ہے اور x اور y سلسلوں کے اضافی حرکات کی نمائندگی کے لئے الجبریائی مساوات حاصل کی جاتی ہیں۔ ایسے طریقوں میں ایک طریقہ Method of least squares بھی ہے۔ اس طریقہ میں ہم ایک متغیر کی دی گئی قیمتوں کے درمیان انحراف کے مربعوں کے مجموعہ اور لائین آف بیسٹ فٹ کی تخمینی قیمتوں کو کم سے کم کیا جاتا ہے۔ x پر y ریگریشن لائین مصرحہ کی قدر، قیمت کے لئے بہترین تخمینہ دیتی ہے اور اسی طرح y پر x ریگریشن لائین مصرحہ y کی قدر x کے لئے بہترین تخمینہ کا باعث ہے۔

اگر y کی قیمتیں y محور پر مرتسم کی جائیں تب x پر y کی ریگریشن لائین کچھ ایسی ہو گی جو Vertical انحراف کے مربعوں کے مجموعہ کم کرے گی۔ اسی طرح اگر x کی قیمتیں x محور پر مرتسم کی جائیں y پر x یگریشن لائین کچھ ایسی ہو گی جو Horizontal انحراف کے مربعوں کے مجموعہ کو کم کرے گی۔

ہم نے میتھڈ آف لیسٹ اسکوائرز پر کوریلیشن کے باب میں مختصراً بحث کی ہے اور Line of best fit خط مستقیم y=a+bx حاصل ہو گی اور یہ کہ میتھڈ آف لیسٹ اسکوائرز میں اس خط کو درج ذیل حسب معمول مساوات کی مدد سے حاصل کیا جاسکتا ہے۔

اگر x اور y کی قیمتیں مذکورہ بالا مساوات میں درج کی جائیں تو ہمیں a اور b کی قیمتیں حاصل ہوں گی اس کے علاوہ ہمیں x پر y ریگریشن لائین بھی حاصل ہو گی یہاں y Dependent متغیر اور 'x' Independent' متغیر ہے۔ y پر x ریگریشن لائین حاصل کرنے کے لئے ہمیں x کو Dependent متغیر اور y کو Independent متغیر تصور کرنا ہو گا اس طرح ہمیں دو ریگریشن لائینس کے لئے دو مساوات حاصل ہوں گی۔

درج ذیل مثالیں مذکورہ بالا نکات کی توضیح کریں گی۔

تمثیل 1: درج ذیل اعداد و شمار کے لئے ریگریشن لائینس کو مرتسم کیجئے۔

x:	1	2	3	4	5
y:	166	184	142	180	338

خط مستقیم کی مساوات کو حاصل کرنے کے لئے دی گئی قیمتوں کو ہم حسب معمول درج ذیل مساواتوں میں استعمال کریں گے۔

$$\Sigma(y) = na + b(\Sigma x)$$

$$\Sigma(xy) = a\Sigma(x) + b(\Sigma x^2)$$

یہ مساواتیں ہمیں a اور b کی قیمتیں دیں گی ہم خط مستقیم کی مساوات میں درج کریں گے۔

$$y = a + bx$$

یہ ہمیں x پر y ریگریشن لائین دے گی اس سوال میں a = 100 اور b = 34 ہو گی۔ اس طرح مساوات یوں ہو گی:

x	x^2	y	xy	y^2
1	1	166	166	27556
2	4	184	368	33856
3	9	142	246	20164
4	16	180	720	32400
5	25	338	1690	114244
5 - 15	55	1010	3370	228220

$$\Sigma y = na + b(\Sigma x) \quad .... \ (1)$$

$$1010 = 5a + 15b$$

$$202 = a + 3b \quad .... \ (1)$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \quad .... \ (2)$$

$$3370 = 15a + 55b$$

$$674 = 3a + 11b \quad .... \ (2)$$

مساوات نمبر (1) کو (3) سے ضرب دینے پر

$$3(202) = a + 3b \quad .... \ (1)$$

$$606 = 3a + 96$$

تفریق کرنے پر

$$674 = 3a + 116$$

$$+68 = +2b$$

$$b = \frac{68}{2}$$

$$b = 34$$

b کی قیمت مساوات نمبر (1) میں درج کرنے پر

$$202 = a + 3 \times 34$$

$$a = 202 - 202$$

$$a = 100$$

اسی طرح x اور y کا تبادل کرنے پر خط مستقیم کی مساوات ہمیں x اور y کی قیمتیں فراہم کریں گے

$$x = a + by$$

اور دو حسب معمول مساواتیں یوں ہوں گی:

$$\Sigma x = na + b(\Sigma y)$$

$$\Sigma(xy) = a\Sigma y + b(\Sigma y^2)$$

ان مساوات کے ساتھ y پر x خط مستقیم کی مساوات

$$\Sigma x = 5a + b1010$$

$$15 = 5a + b1010 \text{ ....(1)}$$

$$3370 = 1010a + b\ 228220 \text{ ....(2)}$$

مساوات نمبر 1 کو 202 سے ضرب دینے پر

$$202(15 = 5a + b\ 1010)$$

$$3030 = 1010a + 204020b \text{ ....(1)}$$

$$3370 = 1010a + 228220b$$

$$+340 = +24200b$$

$$b = \frac{340}{24200} = 0.0140$$

b کی قیمت مساوات نمبر 1 میں درج کرنے پر

$$15 = 5a + 0.0140495868 \times 1010$$

$$-5a = 14.190 = 15$$

$$-5a + = 14.190 = 15$$

$$+5a = +0.8099173$$

$$a = \frac{0.8099173}{5} = 0.162$$

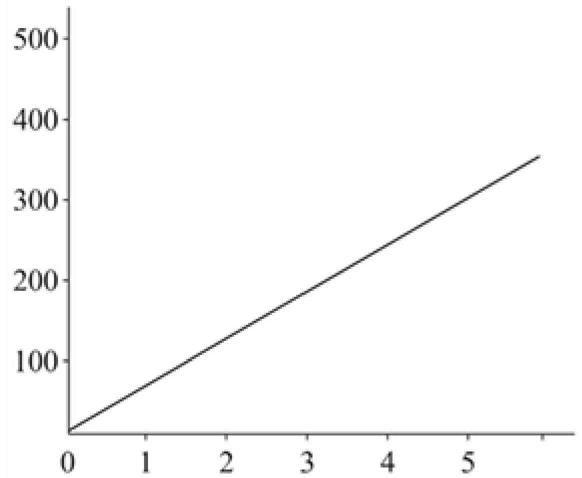
y پر x کی مساوات... $x = 0.162 + 0.014y$

چونکہ 'a' کی قیمت حسب معمول مساوات 0.162 اور 'b' کی 0.014 ہو گی۔
اس طرح خط مستقیم کی مساواتیں یالائین آف بیسٹ فٹ یوں ہوں گی۔

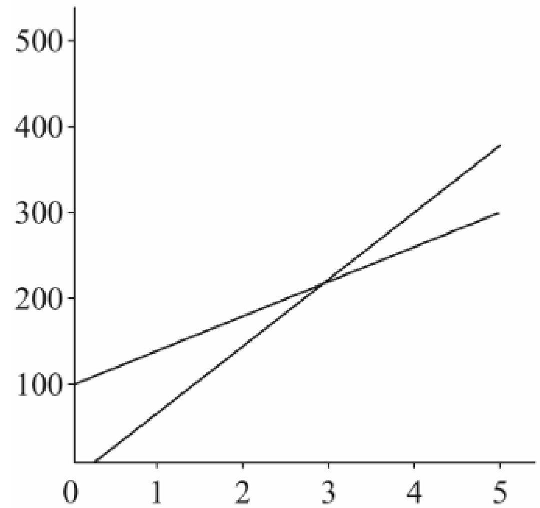
$$y = 100 + 34x \quad .... \quad (1)$$

$$y = 0.162 + 0.014y \quad .... \quad (2)$$

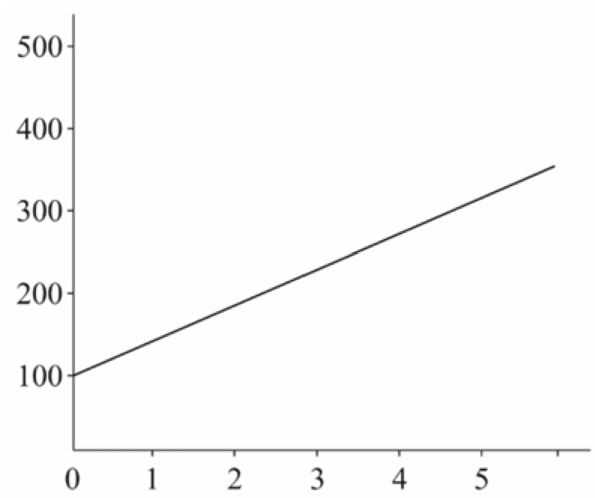
پہلی مساوات سے ہم y کی کوئی دو قیمتیں دریافت کریں گے اور چند x کی قیمتوں کو ترسیمی کاغذ پر y پر x کی ریگریشن لائین حاصل کرنے کے لیے مرتسم کریں گے۔ یہ دو خطوط نکتہ اواسط پر ایک دوسرے کو قطع کرے گے ان خطوط سے درج ذیل ترسیم حاصل ہو گی۔



مذکورہ بالا گراف سے یہ مشاہدہ کیا جاسکتا ہے کہ دو سلسلوں کے اوسط حسابیہ پر دو ریگریشن خطوط ایک دوسرے کو قطع کرتے ہیں۔
اگر ہم دی گئی x کی قیمت پر y کی کوئی قیمت دریافت کرتے ہیں تب ہم x سلسلہ سے عمودی خط (x کی دی گئی قیمت پر) کھینچیں گے تب x پر y ریگریشن لائین جس نکتہ پر قطع کرے گی وہ نکتہ y کی محسوب کردہ قیمت کی نشاندہی کرے گا جس کو y محور پر پڑھا جاسکے گا۔ (x پر y ریگریشن لائین سے x محور پر متوازی خط کو عمودی طور پر اس نکتہ سے کھینچا جائے گا جہاں یہ دونوں ملتے ہوں) تب x = 2 اور 168 'y' ہو گا۔ اسی طرح x سلسلے کی قیمتیں y پر x ریگریشن لائین کو استعمال کرتے ہوئے دریافت کی جاسکتی ہیں۔
جیسا کہ اس سے قبل نشاندہی کی گئی کہ ریگریشن لائین یالائین آف بیسٹ فٹ وہ ہے جس کو متغیر کی دی گئی قیمتوں اور تخمینی قیمتوں کے درمیان انحرافی مربع کا فرق کم ترین ہو۔ ہماری اس تمثیل میں y سلسلہ عمودی پیمانے پر ہے اس لئے اصل اعداد اور ریگریشن لائین کے درمیان (عمودی پیمانے پر ناپے گئے ہیں) کم ترین ہوں گے۔ x سلسلوں کے لئے افقی پیمانے کے مربعے کم ترین ہوں گے۔ درج ذیل دو گراف اس نکتہ کی وضاحت کریں گے۔



مذکورہ بالا گراف میں x پر y ریگریشن لائین کے ساتھ ساتھ ہم y کے اصل اعداد اور عمود اعداد' اصل اور محسوبہ اعداد کے درمیان انحراف کو ظاہر کرتے ہیں۔ ان انحرافی اعداد کا مجموعہ کم ترین ہوگا جب کہ انحراف ریگریشن لائین کی قیمتوں سے لیا جائے گا۔
درج ذیل گراف بھی y پر x ریگریشن مساوات میں x سلسلے کی قیمتوں اور اصل قیمتوں کے درمیان انحراف کو اسی طرح ظاہر کرتا ہے۔



مذکورہ بالا گراف سے یہ واضح ہے کہ x سلسلوں کے ریگریشن لائین کے انحرافی اعداد کم ترین ہیں اور اسلئے ان کے مربعوں کا مجموعہ بھی کم ترین ہوگا۔ حقیقت میں انحراف کا مجموعہ ہمیشہ صفر ہوتا ہے اس لئے انحرافی اعداد کے مربعے بھی کم ترین ہوتے ہیں۔

ریگریشن لائنیس دو ہی کیوں؟

اکثر یہ سوال بہت زیادہ پوچھا جاتا ہے کہ x اور y کی قیمتوں کو حاصل کرنے کے لئے ریگریشن لائنیس دو ہی کیوں؟ اور ایک ہی ریگریشن لائین مقصد کی تکمیل کیوں نہیں کرتی؟ جواب نہایت آسان ہے اور وہ یہ ہے کہ جب تک کہ ان کے درمیان پرفیکٹ پازیٹیو یا نیگیٹو کوریلیشن کی نشاندہی نہیں ہو جاتی دونوں x اور y سیریز کے لئے انحرافی اعداد کے مربعوں کا مجموعہ کو کم ترین نہیں کیا جاسکتا۔ پرفیکٹ کوریلیشن کی صورت میں ایک ریگریشن لائین کا ہونا کافی ہوتا ہے کیونکہ x اور y سلسلوں کے ایک ہی قسم کے انحرافی اعداد ہوتے ہیں۔ عموماً سوشیل

سائنس میں شاذ و نادر ہی پرفیکٹ کوریلیشن پایا جاتا ہے۔ اس وجہ سے x سلسلوں کے انحرافی اعداد کے مربعوں کے مجموعہ کو کم ترین کرنے کے لئے ایک ہی لائین کافی ہوتی ہے اور دوسری ریگریشن لائین y سلسلے کے انحرافی اعداد کا لحاظ کرتی ہے مذکورہ بالا کھینچے گئے گراف A اور B سے یہ نکتہ واضح ہو جاتا ہے۔ گراف نے y سلسلوں اور گراف B نے x سلسلوں کے انحرافی اعداد کے مربعوں کو کم ترین کر دیا ہے۔ پہلی صورت انحرافی اعداد کی پیائش عمودی محور اور دوسرے کو افقی محور پر پیائش کی جائے گی۔

ریگریشن مساوات (Regression equations): ریگریشن مساواتیں ریگریشن لائنیں کا الجبریائی اظہار ہیں۔ چونکہ دو ریگریشن لائنیں ہیں اور دو ریگریشن مساواتیں ہیں۔ y پر x ریگریشن لائین دی گئیں کی y قیمتوں کے لئے x کی ممکنہ حد تک اوسط قیمتیں دیتی ہے اور اسی طرح x پر y ریگریشن لائین دی گئی x کی قیمتوں کے لئے y کی ممکنہ حد تک اوسط قیمتیں دیتی ہے۔ جیسا کہ ریگریشن مساوات y پر x کو y کی دی گئی تبدیل ہوتی رہی قیمتوں کے لئے x کی قیمتوں میں تغیر کو بیان کرنے کے لئے استعمال ہوگی اور اسی طرح x پر y ریگریشن مساوات x کی دی گئی تبدیل ہوری رہی قیمتوں کے لئے y کی قیمتوں میں تغیر کو بیان کرنے کے لئے استعمال ہوگی۔

x پر y کی ریگریشن مساوات $x = a + by$ اور x پر y کی ریگریشن مساوات $y = a + bx$ ہے۔

یہ خط مستقیم کی مساواتیں ہیں۔ ان مساواتوں میں a اور b کی قیمتیں یکساں ہیں جو ریگریشن لائنیں کی جگہ یا مقام کا تعین کریں گی۔ پیرامیٹر a ریگریشن لائین کی سطح کی نشاندہی کرے گا (مبدے کے نیچے یا اوپر کے خط کا فاصلہ)۔ پیرامیٹر 'b' لائین کے ڈھلان کا تعین کرے گا یعنی y میں فی اکائی تبدیلی کے مقابل x میں تبدیلی یا اس کے برعکس۔

مذکورہ بالا مساوات میں a اور b کی قیمتیں Method of lease squares سے دریافت کی جا سکتی ہیں اس کا حوالہ پہلے دیا جا چکا ہے۔ a اور b کی قیمتیں نیچے دی گئیں سادہ مساوات کی مدد سے معلوم کی جا سکتی ہیں۔

$$\text{(i)}\ \Sigma y = na + b\Sigma x \qquad \text{(ii)}\ \Sigma x = na + b\Sigma y$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \qquad \Sigma xy = a\Sigma y + b\Sigma y^2$$

ہم ایک تمثیل کے ذریعہ a اور b کی قیمتیں معلوم کرنے کی ٹیکنک معلوم کریں گے اور اس کی مدد سے ریگریشن مساوات بھی حاصل کریں گے۔

تمثیل 2: درج ذیل اعداد و شمار کی مدد سے Method of least squares کو استعمال کرتے ہوئے دو ریگریشن مساوات حاصل کریں گے۔

x:	2	4	6	8	10
y:	5	7	9	8	11

حل:

x	x^2	y	xy	y^2
2	5	10	4	25
4	7	28	16	49
6	9	54	36	81
8	8	64	64	64
10	11	110	100	121
$\sum y=30$	$\sum y=40$	$\sum xy=266$	$\sum y^2=220$	$\sum y^2=340$

Regrassion equation y on x = y = a + bx

aاور bکی قیمتیں معلوم کرنے کے لئے درج ذیل دو مساواتیں استعمال کی جائیں گی۔

$$\sum x = na + b\sum y \quad ......(1)$$

$$\sum xy = a\sum x + b\sum x^2 \; + \quad ......(2)$$

جدول کی قیمتوں کو مساوات میں درج کرنے پر

$$40 = 5a + 30b \quad ......(1)$$

$$266 = 30a + 220b \quad ......(2)$$

مساوات نمبر(1) کو 6 سے ضرب دینے پر

$$240 = 30a + 180b$$

$$\underline{266 = 30a + 220b}$$

$$+26 = \; +40b$$

$$b = 26 \div 40 = 0.65$$

bکی قیمت کو مساوات نمبر 1 میں درج کرنے پر

$$40 = 5a + 30 \text{ x } 0.65$$

$$40 = 5a + 19.5$$

$$-5a = 19.5 - 40$$

$$+a = \frac{20.5}{5}$$

xپر yر یگریشن مساوات میں aاور bکی قیمتیں درج کرنے پر یہ xپر yر یگریشن مساوات ہے

$$y = 4.1 + 0.65x$$

Regression equation y on x = a + by

دوسادہ مساوات ہیں

$\Sigma x = na + b\Sigma y$ .... (1)

$\Sigma xy = a\Sigma x + b\Sigma x^2$ .... (2)

ان مساوات میں قیمتیں درج کرنے پر

$30 = 5a + 40b$ .... (1)

$266 = 40a + 340b$ .... (2)

مساوات نمبر (1) کو 8 سے ضرب دینے پر

$240 = 40a + 320b$ .... (1)

$266 = 40a + 340b$ .... (2)

$+26 = +20b$

$b = \frac{26}{20} = 1.3$

b کی قیمت کو مساوات نمبر 1 میں درج کرنے پر

$30 = 5a + 40 \times 1.3$ ...(1)

$30 = 5a + 52$

$-5a + 52$

$+5a + 52$

$a = \frac{-22}{5} = -4.4$

y پر x ریگریشن مساوات میں a اور b کی قیمت درج کرنے پر

$x = -4.4 + 1.3y$

یہ y پر x ریگریشن ہے۔

8.6 حسابی اوسط سے انحراف کا طریقہ (Method of Deviations from Mean)

Method of least squares سے ریگریشن مساوات حاصل کرنے کا طریقہ اکتانے والا اور بہت زیادہ حسابات پر مبنی ہے۔ ریگریشن مساوات کو حاصل کرنے کا آسان طریقہ وہ ہے جس میں x اور y کی حقیقی قیمتیں لینے کے بجائے ان کے حسابی اوسط سے انحراف لیا جائے۔ یہ عمل بہت زیادہ حسابات سے بچائے گا اور وہی رزلٹ حاصل ہوگا جو کہ ہمیں Method of least square سے حاصل ہوتا رہا ہے۔ اس طریقہ کے تحت ریگریشن مساوات کی تحسیب کا ضابطہ، فارمولہ یہ ہے۔

y پر x ریگریشن مساوات ہے۔

$$(x-\overline{x}) = r\frac{\sigma x}{\sigma y} = (y-\overline{y})$$

جہاں $\overline{x}$ ‘x سلسلوں کے حقیقی حسابی اوسط کو $\overline{y}$ ‘y سلسلوں کے حقیقی حسابی اوسط کو بتاتا ہے۔ ‘x r اور y کے درمیان کوریلیشن کے کوفیشنٹ کو اور σx اور σy، x اور y سلسلوں کے معیاری انحراف کو بتاتا ہے۔

$r\frac{\sigma x}{\sigma y}$ ‘y پر x کا ریگریشن کا کوفیشنٹ کہلاتا ہے اور اس کی bxy سے نشاندہی کی جاتی ہے۔

اس طرح

$$bxy = r\frac{\sigma x}{\sigma y} = \frac{\Sigma xy}{n\sigma x\sigma y}\times\frac{\Sigma x}{n\sigma x\sigma y}\times\frac{\sigma x}{\sigma y} = \frac{\Sigma xy}{n\sigma y^2}$$

$$= \frac{\Sigma xy}{n\times\frac{\Sigma y^2}{n}} = \frac{\Sigma xy}{\Sigma y^2}$$

جہاں x اور y حسابی اوسط سے انحراف ہے۔

اس طرح r‘ σx اور σy کی قیمتوں کو معلوم کرنے کے بجائے ہم راست طور پر bxy کی قیمت x اور y سلسلوں کو ان کے حسابی اوسط سے انحرافی اعداد کے حاصل ضرب کو y سلسلے کے حسابی اوسط سے انحرافی اعداد کے مربعوں کے مجموعہ سے تقسیم کرکے حاصل کر سکتے ہیں۔

x پر y ریگریشن مساوات

$$(y-\overline{y}) = r\frac{\sigma y}{\sigma x}(x-\overline{x})$$

$r\frac{\sigma x}{\sigma y}$ کو x پر کا ریگریشن کا کوفیشنٹ کہلاتا ہے۔ اس کی قیمت $\frac{\Sigma xy}{\Sigma x^2}$ ہے جیسا کہ y پر x ریگریشن مساوات کی صورت میں ثابت کیا گیا تھا۔

اس طرح دو ریگریشن مساوات کی صورت یوں ہو گی۔

i) Regression equation x on y

$$(x-\overline{x}) = \frac{\Sigma xy}{\Sigma y^2}(y-\overline{y})$$

ii) Regression equation y on x

$$(y-\overline{y}) = \frac{\Sigma xy}{\Sigma x^2}(x-\overline{x})$$

Regression coefficientکی خصوصیات

ریگریشن کے کوفیشنٹ bxyاور byxکی چند خصوصیات ذیل میں بیان کی گئی ہیں۔

1. bxyاور byx کے جیومٹریائی اوسط کو ریلیشن کے کوفیشنٹ کی قیمت معلوم کرنے کا باعث ہے۔ اس کا مطلب یہ ہے کہ $\sqrt{bxy\ x\ byx}$=اس کو یوں ثابت کیا جاسکتا ہے۔

$$bxy = r\frac{\sigma x}{\sigma y} \qquad byx = r\frac{\sigma y}{\sigma x}$$

$$bxy \times byx = r\frac{\sigma x}{\sigma y} \times r\frac{\sigma y}{\sigma x} = r^2$$

$byx = r\frac{\sigma y}{\sigma x}$ اور $bxy = r\frac{\sigma x}{\sigma y}$، $\qquad r = \sqrt{bxy \times byx}$

2. دونوں ریگریشن کے کوفیشنٹ کی الجبرائی علامتیں ایک ہی ہوتی ہیں۔دونوں پازیٹیو ہوں گی یا پھر دونوں نیگیٹو ہوں گی۔

یہ ممکن نہیں کہ ایک کی علامت '+'ہواور دوسرے کی علامت '-'ہو۔وجہ یہ ہے کہ اب تک جہاں تک معیاری انحراف کا تعلق ہے یہ ہمیشہ پازیٹیو ہوتے ہیں۔صرف کوریلیشن کے کوفیشنٹ پازیٹیو یا نیگیٹیو ہو سکتے ہیں۔چونکہ اوران کی الجبرائی علامتوں کا تعین علامت rسے کیا جاتاہے اگر یہ دونوں پازیٹیں ہوں توbxyاور byxبھی پازیٹیو ہوں گے اگر دونوں نیگیٹیو ہوں توbxyاور byxبھی نیگیٹیو ہوں گے۔

3. bxyاور byxدونوں کی قیمتیں انفرادی طور پر ایک سے زیادہ نہیں ہو سکتیں۔دونوں میں سے کسی ایک کی قیمت ایک سے زیادہ ہو سکتی ہے لیکن دونوں کی نہیں ہو سکتیں۔اس کی وجہ یہ ہے کہ $\sqrt{byx - byx}$ اور اگر دونوں bxyاور byxانفرادی طور پر قیمتیں ایک سے زیادہ ہوں تو کوریلیشن کا کوفیشنٹ یا rبھی ایک سے زیادہ ہوگا جو کہ ممکن نہیں ہے۔

درج ذیل تمثیل مذکورہ بالا ریگریشن مساواتوں کو حاصل کرنے کے طریقہ کے استعمال کو ظاہر کرے گی۔

تمثیل3: درج ذیل اعداد و شمار کی مدد سے Method of least squaresکواستعمال کرتے ہوئے دو ریگریشن مساوات حاصل کریں گے۔

x:	2	4	6	8	10
y:	5	7	9	8	11

حل:

x	x-$\bar{x}$ x	x^2	y	y-$\bar{y}$ y	x^2	xy
2	-4	16	5	-3	9	12
4	-2	4	7	-1	1	2

6 8 10	0 +2 +4	0 4 16	9 8 11	+1 0 +3	1 0 9	0 0 12
Sx 30 Mean 6	Sx 0	Sx^2 40	Sy = 40 Mean 8	Sy 0	Sy 20	Sxy

Regression equation *y* on *x* on *y*

$$x-\bar{x}=r\frac{\sigma y}{\sigma x}(y-\bar{y})$$

$$x-\bar{x}=\frac{\Sigma xy}{\Sigma y^2}(y-\bar{y})$$

$$x-6=\frac{26}{20}(y-8)$$

$$x-6=1.3(y-8)$$

$$x-6=1.3y-10.4$$

$$x=1.3y-10.4+6$$

$$x=1.3y-4.4$$

Regression equation *y* on *x* on *y*

$$(y-\bar{y})=r\frac{\sigma y}{\sigma x}(x-\bar{x})$$

$$(y-\bar{y})=\frac{\Sigma xy}{\Sigma x^2}(x-\bar{x})$$

$$y-8=\frac{26}{40}(x-6)$$

$$y-8=0.65(x-6)$$

$$y-8=0.65x-3.90$$

$$y=0.65x-3.9+8$$

$$y=0.65x-4.1$$

یہ مشاہدہ کیا جاسکتا ہے کہ یہ ریگریشن مساوات ویسی ہی ہیں جیسے تمثیل نمبر 2 میں Method of least squares سے حاصل کی گئی تھیں۔

مفروضہ حسابی اوسط سے انحراف Deviations from Assumed Mean

جہاں دو x اور y سلسلوں کا حسابی اوسط اعشاریہ میں ہو ریگریشن مساوات حاصل کرنے کے لئے مذکورہ بالا طریقہ تکلیف دہ مشقت طلب ہو جاتا ہے ایسی صورت میں انحراف مفروضہ حسابی اوسط سے لیا جائے گا۔ ریگریشن مساوات میں تحسیب کردہ فارمولہ وہی رہے گا تاہم کچھ آسانی ممکن ہے y پر x ریگریشن مساوات یوں ہوگی۔

$$x-\bar{x}=r\frac{\sigma y}{\sigma x}(y-\bar{y})$$

$$bxy\ \ r=r\frac{\sigma x}{\sigma y}\left[\frac{\Sigma dxdy-\frac{(\Sigma dx\times\Sigma dy)}{N}}{\sigma x\ \ \sigma y}\right]\times\frac{\sigma x}{\sigma y}$$

$$bxy=\frac{\Sigma dxdy-\frac{\Sigma dx\times\Sigma dy}{N}}{\sigma y^2}$$

$$bxy=\frac{\Sigma dxdy-\frac{\Sigma dx\times\Sigma dy}{N}}{\Sigma dy^2-\frac{(\Sigma dy)^2}{N}}$$

اسی طرح x پر y ریگریشن مساوات

$$y-\bar{y}=r\frac{\sigma y}{\sigma x}(x-\bar{x})$$

$$byx\ or\ r\frac{\sigma y}{\sigma x}=\left[\frac{\Sigma dxdy-\frac{(\Sigma dx\times\Sigma dy)}{N}}{\sigma x\times\sigma y}\right]\times\frac{\sigma y}{\sigma x}$$

$$byx=\frac{\Sigma dxdy-\frac{\Sigma dx\times\Sigma dy}{N}}{\sigma x^2}$$

$$byx=\frac{\Sigma dxdy-\frac{\Sigma dx\times\Sigma dy}{N}}{\Sigma dx^2-\frac{(\Sigma dx)^2}{N}}$$

ایک بار bxy اور byx کی قیمتیں راست σ(y اور σ r_1x کی تحسیب کے بغیر) معلوم کر لئے جائیں تب انہیں ریگریشن مساوات کے فارمولے میں درج کیا جا سکتا ہے۔

گروہی اعداد و شمار کی صورت میں دورخی تعددی جدول کو تھوڑا تبدیل کیا جاتا ہے چونکہ کوریلیشن کے ٹیبل میں ہم کو تعدد دی جاتی ہیں اس لئے ان قیمتوں کو تعدد سے ضرب دیا جاتا ہے۔ دوسری اہم تبدیلی یہ کہ اگر Step deviation لئے گئے ہوں تب y پر x ریگریشن لائین کی صورت میں فارمولے کے ذریعے محصلہ قیمتوں کو $\frac{ix}{iy}$ سے ضرب دیا جائے اور y پر x ریگریشن لائین کی صورت میں $\frac{iy}{ix}$ سے ضرب دیا جائے۔ (ix اور iy اور y اور x سلسلوں کے وقفہ جماعت ہیں) یہ ضروری ہے کیونکہ جب تک کہ کوریلیشن کا کوفیشنٹ غیر مشابہ نہ ہو ریگریشن کے کوفیشنٹ پیمانے میں تبدیلی سے متاثر ہوں گے اگرچہ کہ یہ مبداء کی تبدیلی سے متاثر نہیں ہوتے۔

3 اس طرح گروہی اعداد کی صورت میں

$$bxy \text{ or } r\frac{\sigma x}{\sigma y} = \frac{\Sigma dxdy - \frac{(\Sigma fdx \times \Sigma fdy)}{N}}{\Sigma fdy^2 - \frac{(\Sigma fdy)^2}{N}} \times \frac{ix}{iy}$$

$$byx = r\frac{\sigma y}{\sigma x} = \frac{\Sigma dxdy - \frac{(\Sigma fdx \times \Sigma fdy)}{N}}{\Sigma fdy^2 - \frac{(\Sigma fdy)^2}{N}} \times \frac{iy}{ix}$$

اس کو درج ذیل تمثیل سے سمجھا جاسکتا ہے۔

تمثیل 4: تمثیل نمبر 2 میں دیئے گئے اعداد و شمار سے x سلسلے سے 5 اور y سلسلے سے 6 کو مفروضہ حسابی اوسط لے کر ریگریشن مساوات ترتیب دیجئے۔

x	Deviation from AM 5 dx	dx^2	y	Deviation from AM 6 dy	dy^2	$dxdy$
2	-3	9	5	-1	1	+3
4	-1	1	7	+1	1	1
6	+1	1	9	+3	9	+3
8	+3	9	8	+2	4	+6
10	+5	25	11	+5	25	+25
Sx 30	Sdx 5	Sd^2 45	Sy=40	Sdy +10	Sy^2 40	S$dxdy$ +36

$$bxy = \frac{\Sigma dxdy - \frac{\Sigma dx \times \Sigma dy}{N}}{\Sigma dy^2 - \frac{(\Sigma fdy)^2}{N}} = \frac{36 - \frac{5\times 10}{5}}{40 - \frac{(10)^5}{5}} = \frac{26}{20} = +1.3$$

$$byx = \frac{\Sigma dxdy - \frac{\Sigma dx \times \Sigma dy}{N}}{\Sigma dx^2 - \frac{(\Sigma dx)^2}{N}} = \frac{36 - \frac{5\times 10}{5}}{40 - \frac{(5)^2}{5}} = \frac{26}{40} = +0.65$$

bxy اور byx کی قیمتیں وہی ہیں جن کو ہم نے تمثیل نمبر 2 اور 3 میں پایا تھا۔

تمثیل 5: نیچے دیئے گئے اعداد و شمار سے Method of least squares سے ریگریشن مساوات کی تحسیب کیجئے۔ دو ریگریشن لکیروں کو مرتسم کیجئے اور گراف سے x کی قیمت معلوم کیجئے جب کہ y=110 ہو اور y کی قیمت معلوم کیجئے جب کہ x =23 ہو۔ نتیجہ کی جانچ الجبریائی طریقہ سے کیجئے۔ کوریلیشن کے کوفیشنٹ کی قیمت بھی معلوم کیجئے۔

x :	10	15	35	40	50
y :	100	30	110	80	120

حل:

x	x^2	y	xy	y^2
10	100	1000	100	10,000
15	90	1350	225	8,100
35	110	3850	1225	12,100
40	80	3200	1600	6,400
50	120	6000	2500	14,400
Sx 150	Sy 500	Sxy 15400	Sx^2 5652	Sy^2 51000

Regression equation y on $x = y = a + by$

Two normal equations are

$$\Sigma y = na + b\Sigma x \qquad 150 = 5a + 500b \qquad ....(1)$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \qquad 15400 = 500a + 51000b \; .....(2)$$

ہم ایک تمثیل کے ذریعہ a اور b کی قیمتیں معلوم کرنے کی ٹیکنک معلوم کریں گے اور اس کی مدد سے ریگریشن مساوات بھی

مساوات نمبر (1) کو 100 سے ضرب دینے پر

$$15400 = 500a + 50000b \; .....(1)$$

$$15400 = 500a + 51000b \; .....(1)$$

$$+400 = 1000b$$

$$b = +\frac{400}{1000} = 0.4$$

b کی قیمت مساوات نمبر 1 میں درج کرنے پر

$$150 = 5a + 500 \times 0.4$$

$$50 = 5a + 200$$

$$-5a = 200 - 150$$

$$-a = \frac{50}{5}$$

$$a = -10$$

مساوات میں درج کرنے پر

$$x = -10 + 0.4y$$

Regression equation y on $x = y = a + by$

دوحسب معمول مساوات

$$\Sigma y = na + b\Sigma x \qquad 500 = 5a + 150b$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \qquad 15400 = 150a + 5650b$$

$$500 = 5a + 150b \qquad .....(1)$$

$$15400 = 150a + 5650b \quad .....(2)$$

مساوات نمبر (1) کو 30 سے ضرب دینے پر

$$15000 = 150a + 4500b \quad .....(1)$$

$$15400 = 150a + 5650b \quad .....(2)$$

$$+ \ 400 = \ +1150b$$

$$b = \frac{400}{450} = 0.3478 - 0.347826087$$

b کی قیمت مساوات نمبر (1) میں درج کرنے پر

$$500 = 5a + 150b \quad ....(1)$$

$$500 = 5a + 150 \times 0.347826087$$

$$500 = 5a + 52.1739131$$

$$-5a = 52.1739131 - 500$$

$$+5a = +447.826087$$

$$+a = \frac{447.826087}{5} = 89.5652134$$

$$y = 89.5652175 + 0.347826087x$$

۔ ہونگی اس طرح دو ریگریشن مساوات یوں

ان ریگریشن مساوات کو مرتسم کرنے کے لئے ہمیں x کی دو قیمتیں اور کچھ قیمتیں y کی درکار ہیں تاکہ x پر y کی ریگریشن لائین کھینچ سکیں۔اسی طرح ہمیں y کی دو قیمتیں اور کچھ x کی قیمتیں درکار ہیں تاکہ y پر x کی ریگریشن لائین کھینچ سکیں۔

ہم جانتے ہیں کہ دو ریگریشن لائین x اور y سلسلے کے حسابی اوسط کے نکتہ پر ایک دوسرے کو قطع کرتے ہیں۔ لہٰذا ہر صورت میں دو نکتوں میں ایک نکتہ سلسلے کی حسابی اوسط قیمت ہو گی۔

Mean x = 30, Mean y = 100

لائینس کھینچنے کے لئے ہمیں x کی ایک قیمت اور y کی ایک قیمت کی ضرورت ہو گی۔ چلیں x کی قیمت دریافت کرتے ہیں جب کہ y = 70 ہو۔اس مقصد براری کے لئے ہم y پر x ریگریشن مساوات کا استعمال کریں گے۔ وہ مساوات ہے۔

$$x = -10 + 0.4y$$

اگر y کی قیمت 70 ہو تب (x = -10 + (0.4y x 70 = 18

اس طرح ہمیں x کی دو قیمتیں حاصل ہوئیں جب کہ 30 = x' 100 = y جب کہ 70 = y اور 18 = x ہیں۔
اسی طرح اب ہم y کی قیمت معلوم کریں گے جب کہ x=23 ہو ہم اس مقصد کے لئے x پر y کی ریگریشن مساوات استعمال کریں گے۔

$$y = 89.562174 + 0.3467826087\ x$$

اگر x = 23 ہو تو

$$y = 89.562174 + 0.3467826097\ x\ 23$$

$$y = 89.562174 + 8 = 97.56$$

اس طرح ہمیں y کی دو قیمتیں حاصل ہوئیں x = 30' y = 100 جب کہ x = 23، y = 97.56
ان قیمتوں کے ساتھ ہم دو ریگریشن لائنس کھینچ سکتے ہیں گراف یوں ہو گا۔(صفحہ نمبر 126)
مذکورہ بالا نتائج کی الجبریائی طور پر یوں جانچ کر سکتے ہیں۔

(i) y کی دریافت جب کہ x = 23 ریگریشن مساوات x پر y استعمال کریں۔

$$y = \frac{10,300}{115} + \frac{40}{115}$$ جبکہ x = 23

(ii) x کی دریافت جب کہ y = 110 ریگریشن مساوات y پر x استعمال کریں۔ جبکہ y = 10 ہو

$$x = -10 + 0.4y$$

$$= 10 + 0.4\ x\ 110$$

$$= -10 + 44 = 34$$

الجبریائی مساوات سے گرافی نتائج درست پائے گئے۔

Coefficient of Correlation is $\sqrt{bxy\ x\ byx}$

چونکہ دونوں bxy اور byx پازیٹیو ہیں اس لئے کوریلیشن بھی پازیٹیو ہو گا۔

تمثیل 6: درج ذیل اعداد و شمار سے دو ریگریشن مساوات حاصل کیجئے۔

فروخت	91	97	108	121	67	124	51	73	111	57
خرید	71	75	69	97	70	91	39	61	80	47

Compulation of Regression Euqation

Sales x	Dev. from Mean 90x	x	x^2	Dev. from Mean 70	y^2	xy
91	+1	1	71	+1	1	1
97	+7	49	75	+5	25	35
108	+18	324	69	-1	1	-18
121	+31	961	97	+27	729	+837
67	-23	529	70	0	0	0
124	+34	1156	91	+21	441	714
51	-39	152	39	+21	961	1209
73	-17	289	61	-9	81	153
111	+21	441	80	+10	100	210
57	-33	1089	47	-23	529	759
Sx 900	Sx 0	Sx^2 6360	Sy 700	Sy 0	Sy^2 2868	Sxy 3900

$$\text{Mean } \bar{x} = \frac{900}{10} = 90 \quad \text{Mean } \bar{y} = \frac{700}{100} = 70$$

$$bxy = \frac{\Sigma xy}{\Sigma x^2} = \frac{3900}{63900} = 0.6132$$

$$bxy = \frac{\Sigma xy}{\Sigma y^2} = \frac{3900}{28686} = 1.361$$

Regresrew equation x on $y = x - \bar{x} = bxy\,(y - \bar{y})$

$$= (x-90) = +1.361(y-70)$$

$$= (x-90) = +1.361y - 95.27$$

$$= x = 1.361y - 95.27 + 90$$

$$x = 1.361y - 5.27 \qquad \text{.... (1)}$$

Regression equation y on x - $(y-\bar{y}) = bxy(x-\bar{x})$

$$(y-70) = 0.6132(x-90)$$

$$y - 70 = 0.6132x - 55.188$$

$$y = 0.6132x - 55.188 + 70$$

$$y = 0.6132x + 14.812 \qquad \text{.... (2)}$$

تمثیل 7: دو ججوں کے ایک پیانل C اور D نے ایک خوبصورتی کے مقابلے میں آزادانہ طور پر سات امیدواروں کو درج ذیل نشانات دیئے۔

:مظاہرہ	1	2	3	4	5	6	7
C: کے نشانات	46	42	44	40	43	41	45
D: کے نشانات	40	38	36	35	39	37	41

آٹھویں مظاہرے میں جج D شریک نہ ہوسکے اس لئے جج C نے 37 نشانات دیئے اگر جج D بھی شریک ہوئے ہوتے تو آٹھویں مظاہرے میں ان سے کتنے نشانات کی توقع کی جاسکتی تھی۔

حل: C کی جانب سے دیئے گئے نشانات کی نشاندہی x سے اور جج D کی جانب سے دیئے گئے نشانات y سے کی جائے گی ہمیں y کی قیمت معلوم کرنی پڑے گی جب کہ $x = 37$ ہو۔ایسا ریگریشن مساوات x پر y معلوم کرنے پر ہوگا۔

Compulation of Regression Euqation y on x

Sales x	Dev. from Mean 90x	x	y	Dev. from Mean 28 y	y^2	xy
46	+3	9	40	+2	4	+6
42	-1	1	38	0	0	0
44	+1	1	36	-2	4	-2
40	-3	9	35	-3	9	+9
43	0	0	39	+1	1	0
41	-2	4	37	-1	1	+2
45	+2	4	41	+3	9	+6
Sx 301	0	Sx^2 28	Sy 266	0	Sy^2 = 28	Sxy = +21

$$\text{Mean } x \text{ or } \overline{x} = \frac{301}{7} = 43 \quad \text{Mean } y \text{ or } \overline{y} = \frac{266}{7} = 38$$

$$bxy = \frac{\Sigma xy}{\Sigma x^2} = \frac{+21}{28} = +0.75$$

Regresrew equation x on $y = (y - \overline{y}) = \text{by}x\,(x - \overline{x})$

$$y - 38 = 0.75(x - 43)$$

$$y - 38 = 0.75x - 32.25$$

$$y = 0.75x - 32.25 + 38$$

$$y = 0.75x + 5.75$$

اگر $x = 37$ ہو تو $y = 0.75 \times 37 + 5.75$

$$y = 27.75 + 5.75 = 33.5$$

اس طرح اگر جج D موجود ہوتے تو وہ 33.5 نشانات دئے ہوتے۔

8.9 اکتسابی نتائج (Learning Outcome)

طلباء نے ریگریشن کے معنی، تعریف، اور نوعیت اور متغیر کے درمیان تعلقات کا تجزیہ کرنے میں اس کے کردار کے بارے میں ایک جامع سمجھ حاصل کی ہے۔ انہوں نے نتائج کی پیشن گوئی کرنے اور ڈیٹا پر مبنی فیصلہ سازی کی حمایت میں ریگریشن کے تجزیہ کے استعمال کو سیکھا ہے۔ طلباء اب کوریلیشن اور رجعت کے تجزیے کے درمیان فرق کر سکتے ہیں، یہ سمجھتے ہوئے کہ ہر طریقہ شماریاتی تجزیہ میں مختلف مقاصد کو کیسے پورا کرتا ہے۔ انہوں نے ریگریشن کا مطالعہ کرنے کے طریقوں اور وسط سے انحراف کے طریقوں کی کھوج کی ہے، جس سے وہ رجعت

کے گتانکوں کی درستگی سے حساب اور تشریح کر سکتے ہیں۔ مجموعی طور پر، سیکھنے والے ریگریشن کی تکنیکوں کو ماڈل تعلقات میں لاگو کرنے، پیشین گوئیاں کرنے،اور کاروباری تناظر میں قابل عمل بصیرت فراہم کرنے کے لیے لیس ہیں۔

8.8 نمونہ امتحانی سوالات(Model Examination Questions)

معروضی جوابات کے حامل سوالات Objective type questions

1. __________ متغیروں میں قریبی تعلق پایا جاتا ہے۔ (دو)
2. اصطلاح(Regression)کے معنی __________ ہے۔ (واپسی)
3. __________ نےRegressionکو پہلے استعمال کیا۔ (سر فرانسیس گالٹن)
4. ریگریشنی تجزیہ __________ کی ایک شاخ ہے۔ (شماریاتی نظریہ)
5. ریگریشن کے مطالعہ کے دو طریقے ایک ترسیمی دوسرا __________ ہے۔ (الجبرائی)
6. اصطلاح ریگریشن کو گالٹن نے 1877 میں استعمال کیا۔ (صحیح/غلط)
7. انحراف کا مجموعہ ہمیشہ صفر ہوتا ہے۔ (صحیح/غلط)
8. ریگریشنی خطوط الجبریائی اظہار نہیں ہیں۔ (صحیح/غلط)
9. معیاری انحراف ہمیشہ پازیٹیو ہوتا ہے۔ (صحیح/غلط)
10. تخمینی قیمت کے حصول کے لیے مستعملہ آلہ ریگریشنی خط ہے۔ (صحیح/غلط)

مختصر جوابات کے حامل سوالات Short answer type questions

1. ریگریشن کے کیا معنی ہیں؟
2. ریگریشنی تجزیہ سے کیا مراد ہے؟
3. ترسیمی طریقہ کا کیا مطلب ہے؟
4. الجبریائی طریقہ کیا ہے؟
5. جسب منشا قوس کے طریقہ کے بارے میں کیا جانتے ہو؟

طویل جوابات کے حامل سوالات Long answer type questions

1. تجزیہ ریگریشن کی تعریف کیجیے اور اس کے استعمالات واضح کیجیے۔
2. کوریلیشن اور ریگریشن تجزیہ کے درمیان فرق کیجیے۔

3. ریگریشن کے مطالعہ کے طریقہ کیا ہیں؟ تفصیل سے بحث کیجیے۔

4. ریگریشن کے کوفیشنٹ کی خصوصیات بیان کیجیے۔

5. ریگریشن مساوات ترتیب دیجیے۔

سال	:	2001	2002	2003	2004	2005	2006	2007	2008
انتظمی اخراجات (ہزاروں میں)	Q:	12	13	15	23	24	38	42	48
فروخت (لاکھوں میں)	y:	5.0	5.6	5.8	7.0	7.2	8.8	9.2	9.5

جواب: $y=3.87+0.125x$

6. ریگریشن مساوات ترتیب دیجیے۔

A:	6	2	10	4	8
B:	9	11	5	8	7

7. ریگریشن مساوات ترتیب دیجیے۔

M:	15	27	27	30	34	38	46
N:	120	140	150	170	180	200	250

جواب: y on x $y=40.62 + 4.2657\ x$

x on y $k=0.2189\ y-6.8383$

اکائی 9۔ وقتی سلسلہ کا تجزیہ

(Time Series Analysis)

اکائی کے اجزا

9.0 تمہید — Introduction

9.1 مقاصد — Objectives

9.2 وقتی سلسلہ کا تجزیہ۔ معنی اور تعریف — Meaning and Definition of Time Series Analysis

9.3 وقتی سلسلہ کی نوعیت اور دائرہ کار — Nature and Scope of Time Series

9.4 وقتی سلسلوں کی اہمیت، مقاصد اور افادیت — Importance and Utility of Time Series Analysis

9.5 وقتی سلسلہ کے اجزاء — Components of Time Series

9.6 ابتدائی ترمیمات — Introductory Modification

9.7 رجحان کی پیمائش — Measurement of Trend

9.8 اکتسابی نتائج

9.9 نمونہ امتحانی سوالات — Model Examination Questions

9.0 تمہید (Introduction)

وقتی سلسلے دراصل ایسے سلسلوں سے متعلق ہے جن کا ایک عنصر وقت بھی ہے۔ یعنی اگر ہم متواتر ادوار کے لیے تاریخ وار قدروں یا قیمتوں کو ترتیب دیں یہ بھی ایک سلسلہ ہوگا۔ جیسے پیداوار، طلب، رسد، لاگتیں، قیمتیں، قومی آمدنی وغیرہ کا سلسلہ جو دس یا بیس یا زاید برسوں پر مشتمل ہو۔ اسی طرح ہم ہمارے ملک کی آبادی کا سات دہوں کا جائزہ لیں۔ یہ بھی ایک سلسلہ کی مثال ہوگی۔ ایسی ہی مثالیں دی جاسکتی ہیں۔ ایسے سلسلوں کا تجزیہ وقت سلسلوں کا تجزیہ Analysis of Time Series کہلاتا ہے۔ وقتی سلسلوں کا تجزیہ پیش قیاسی اور ماضی کے مظاہرہ کا جائزہ لینے کے لیے کیا جاتا ہے۔ تاجر، ماہر شماریات اور ماہر معاشیات، آبادی، قومی آمدنی، اُجرتیں، تنخواہیں اور اشیاء کی قیمتوں وغیرہ کی تخمینہ کاری کے لیے فطرتاً دلچسپی رکھتے ہیں۔ حقیقت میں ماہر معاشیات اور تجّار کی کامیابی اور ناکامی کا انحصار بڑی حد تک درست پیش بینی پر ہوتا ہے۔ مستقبل کے تعلق سے پیش بینی زیر مشاہدہ متغیر کے برتاؤ کے تجزیہ کی بنیاد پر کی جاتی ہے۔ اس طرح کسی شئے کی مستقبل میں طلب یا کسی

فرم یا کمپنی کے منافع کے حصول کی پیش بینی گذشتہ برسوں میں کمائے گئے نفع یا طلب کے تجزیہ پر کی جاسکتی ہے۔لہٰذا کسی ملک کے تمام معاشی مسائل کے مطالعہ کی اہمیت وقتی سلسلوں کے تجزیہ پر ہوتا ہے۔ یہاں یہ بات ذہن نشین کرلیں کہ وقتی سلسلوں کا تجزیہ صرف تجار، ماہرین معاشیات یا ماہرین شماریات کے لیے ہی فایدہ بخش ہے۔ بلکہ دیگر شعبہ حیات میں بھی کافی اہمیت کا حامل ہے۔ جیسے ماہر سماجیات اگر کسی ملک جرائم اور حادثات سے متعلق صورتحال کا مطالعہ کرنا اور ان کے تدارک کے لیے سفارشات پیش کرنا چاہتا ہو تو اس کے ضروری ہے کہ وہ گذشتہ برسوں میں اس ملک کے امن وضبط اور لا اینڈ آرڈر سے متعلق اعداد و شمار کا جائیزہ لے کر ہی ایسا کر سکتا ہے۔

9.1 مقاصد (Objectives)

اس اکائی کے مطالعہ کے بعد آپ اس قابل ہو جائیں گے کہ

- وقتی سلسلوں کے تجزیہ کے پس منظر معنی اور تعریف کی توضیح کر سکیں۔
- وقتی سلسلوں کی تعریف بیان کر سکیں۔
- وقتی سلسلوں کی افادیت کو واضح کر سکیں۔
- وقتی سلسلوں کے اجزائے ترکیبی کو سمجھا سکیں۔
- وقتی سلسلوں کے تجزیہ کے اقدامات کا خلاصہ کر سکیں۔
- رجحان کی پیائش کے طریقے شمار کر سکیں۔
- رجحان کی پیائش پر سوالات حل کر سکیں۔

9.2 وقتی سلسلے کا تجزیہ-معنی اور تعریف

(Meaning and Definitionof Analysis of Time Series)

وقتی سلسلے کا تجزیہ ایک شماریاتی ٹیکنک ہے جو مختلف ادوار سے متعلق اعداد و شمار یا رجحان کا تجزیہ ہے۔وقتی سلسلوں میں ایک مخصوص مدت کے اعداد و شمار کا سلسلہ ہوتا ہے۔اس کا اطلاق اسٹاک مارکٹ کے تجزیہ، زلزلے کی پیش قیاسی، معاشی پیش بینی، مردم شماری وغیرہ بہت زیادہ ہوتا ہے۔ مشاہدات کا یہ سلسلہ جاری رہتا ہے، کبھی تھمتا نہیں۔ مشاہدات کی لگاتار پیائش ہوتی رہتی ہے۔آج کے عصری دور میں جبکہ دنیا ٹیکنک رُخی ہو گئی ہے۔ یہی وجہ ہے اعداد و شمار کی جمع بندی بھی ٹیکنالوجی کی مدد سے بڑی آسان ہو گئی ہے۔ اگر اس جمع شدہ اعداد و شمار کا تجزیہ کیا جائے تو بیش قیمت نتائج ہاتھ آتے ہیں۔ حالیہ برسوں میں ڈاٹا سائنس اور مشینوں کے ذریعہ اکتساب کا تصور زندگی کے تمام شعبہ جات کا مرکزی رجحان کا مرکز بن چکا ہے۔ پیش بینی کے طریق عمل میں انقلابی تبدیلیاں رونما ہو چکی ہیں۔ شماریاتی ماڈلنگ بھی کم خرچ پر مبنی تشریحی ٹیکنک میں تبدیل ہو چکی ہے۔ وقتی سلسلوں کا تجزیہ کا آغاز خود رجعتی Auto Regression اور حرکتی اوسط Moving Average سے ہوتا

ہے۔اس لیے وقتی سلسلوں کا گہرائی سے مطالعہ آج کے دور میں ضروری ہو گیا ہے۔یہی وجہ ہے کہ مختلف مفکرین اور مصنفین نے بڑی گہرائی و گیرائی سے جائزہ لے کر اپنے خیالات کا اظہار کیا ہے۔ان کو ذیل میں مختصر اور جامع انداز میں پیش کیا جاتا ہے۔

1) Kenny and Keeping کے مطابق ''وقت کی بنیاد پر اعداد و شمار کا ایک جوڑ وقتی سلسلہ کہلاتا ہے''۔

2) Croxton and Cowden نے یوں تعریف کی ہے ''وقتی سلسلہ تاریخ وارانہ اعداد و شمار کی ترتیب پر مشتمل ہوتا ہے''۔

3) Patterson کے الفاظ میں ''وقتی سلسلہ شماریاتی اعداد و شمار پر مشتمل ہوتا ہے جو بغور متواتر افزائش جمع بندی، اندراج کا نتیجہ ہے''۔

4) Maslow نے وقتی سلسلوں کی تعریف یوں کی ہے ''وقت سلسلے کو ایک مدت کے دوران ایک مظاہرے میں تبدیلیوں کو ظاہر کرتے ہیں، یعنی تغیرات کی وجہ سے تاریخ وار مقدار میں تبدیلیوں کی شدت''۔

5) Morris Hamburg کے خیال میں ''وقتی سلسلے شماریاتی مشاہدات کا تاریخ وار ترتیب کا ایک جوڑ ہے''۔

9.3 وقتی سلسلوں کی نوعیت اور دائرہ کار (Nature and Scope of Time Series)

وقتی سلسلہ کے اعداد و شمار کی قدر و خصوصیت ایک جیسی ہونی چاہیے، یعنی اعداد و شمار یکساں ہونے چاہیں۔ ساتھ ہی ساتھ وقتی سلسلوں کی تمام اصطلاحات ایک مظاہر سے متعلق ہونی چاہیے۔ جیسے فی کس آمدنی، قومی آمدنی وغیرہ۔ یہ اعداد و شمار بیس سال، پچیس سال تک درج شدہ حالت میں دستیاب ہونا چاہیے۔ سلسلہ میں شامل مختلف قیمتوں کے درمیان وقفہ مساوی اور یکساں ہو۔ وقتی سلسلے معاشیات اور کاروباری میدان میں زیادہ مددگار ہوتے ہیں، اور ساتھ ساتھ آج کل محققین بھی زیادہ استعمال کر رہے ہیں۔

9.4 وقتی سلسلوں کی اہمیت، مقاصد اور افادیت

(Importance & Utility Analysis of Time Series)

وقتی سلسلوں کی اہمیت، مقاصد اور افادیت درجہ ذیل نکات سے ظاہر ہوتی ہے۔

1۔ پیش قیاسی میں مددگار (Support in Forecasting)

کسی بھی زیر مطالعہ متغیر کے مستقبل میں رویہ کی پیش قیاسی سال گذشتہ کے حالات کے تجزیہ کی بنیاد پر کی جاتی ہے۔ مثال کے طور پر فروخت سے متعلق وقتی سلسلہ کے تجزیہ سے مستقبل میں فروخت کی مقدار کی پیش قیاسی کی جاسکتی ہے۔ یہ اعداد و شمار مستقبل کے لیے لائحہ عمل کی ترتیب میں مددگار ثابت ہوتے ہیں۔ مستحکم معیشت میں منصوبہ کاری کا انحصار زیر مطالعہ مدت کے دوران مختلف متغیرات کے تجزیہ پر ہوتا ہے، تاکہ ملک کی ترقی کی شرح میں اضافہ کیا جاسکے۔

2۔ حصول کامیابی کے جائزہ میں مددگار(Support in Evaluation of Achievement)

منصوبہ کی ترقی کا جائزہ اور نظرثانی وقتی سلسلوں کے اعداد و شمار کی بنیاد پر کی جاتی ہے۔ جیسے پنج سالہ منصوبہ کی ترقی کا فیصلہ خام قومی اشیاء میں سالانہ نشوونما کی شرح کی روشنی میں کیا جاتا ہے۔اسی طرح ہمارے ملک میں افراطِ زر اور بڑھتی قیمتوں پر قابو پانے کے طریقہ کار کا جائزہ وقتی سلسلوں کے تجزیہ کی بنیاد پر مختلف اشیاء کی قیمتوں کا مطالعہ کرکے لیا جاتا ہے۔اس طرح وقتی سلسلے حصول کامیابی میں مددگار ہوتے ہیں۔

3۔ متغیر کے گذشتہ برسوں میں طرزِ عمل کے تجزیہ میں مددگار

(Support in the Analysis of Past Behaviour of a Variable)

کسی زیرِ مشاہدہ متغیر پر مختلف عناصر کے اثرات کا جائزہ بھی گذشتہ برسوں کے اعداد و شمار کے تجزیہ کے بعد لیا جاسکتا ہے۔ یہ مشاہدات دراصل زیرِ مطالعہ مسئلہ پر مختلف ایک جیسے عناصر کے جوڑ کے اثرات کا تجزیہ کرکے انہیں الگ الگ کرتے ہیں۔ جیسے کسی متغیر پر مختلف موسم کے اثرات کا انکشاف کریں گے اور ٹیکنالوجیائی ترقی یا آبادی میں اضافہ ، قومی آمدنی میں تبدیلی کو وقتی سلسلوں کے تجزیہ کے ذریعہ الگ کیا جاسکتا ہے، جبکہ یہ مذکورہ عناصر طویل مدتی ہوتے ہیں۔

4۔ تقابلی مطالعہ میں مددگار(Support in Comparative Studies)

آپ نے وقتی سلسلوں کی تعریف پڑھی ہوگی۔ یہ تاریخ وار اور متواتر ہوتے ہیں۔ جس کی وجہ سے مختلف ادوار اور عناصر کا تقابلی مطالعہ اور تجزیہ بہتر انداز میں کیا جاسکتا ہے۔وقتی سلسلے تقابل کے لیے عصری بنیاد فراہم کرتے ہیں اور ان اجزاء کے اثرات کو الگ الگ کرتے ہیں۔اسی بنیاد پر جغرافیائی تقابل بھی کیا جاسکتا ہے۔

9.5 وقتی سلسلوں کے اجزاء(Components of Time Series)

کسی مدت سے متعلق متغیر کی قیمت میں تبدیلیوں کے عناصر کے اثرات کا نتیجہ ہوتی ہے۔ جیسے صارفین کے ذوق میں تبدیلی، عادات میں تبدیلی، پیداواری لاگت میں تبدیلی، قومی آمدنی میں تبدیلی وغیرہ ایک متغیر کی قیمت میں تبدیلی ایسی وقتوں کے آپسی ٹکراؤ کا نتیجہ ہوتی ہے۔اگر ان قوتوں کی ترکیب میں تبدیلی نہ ہو تو یہ متغیرات کی قدر و قیمت میں تبدیلی نہ ہوتی، اگر توازن میں ذرا سا بھی خلل پڑ جائے تو ان کی قیمتوں میں بھی خلل واقع ہوگا، یعنی قدر و قیمت میں تبدیلی ہوگی اور پھر توازن میں آجاتے ہیں۔ لیکن حقیقت میں یہ منورملہ اس کے برعکس ہوتا ہے۔یہ قوتیں توازن میں نہیں ہوتیں بلکہ ان قوتوں کے درمیان مسلسل خلل ہونے پر ایک عرصہ کے بعد متغیرات کی قدروں میں تبدیلی ہوتی ہے۔ان تبدیلیوں کی شدت کا اندازہ نہیں لگایا جاسکتا۔ہاں ان کے درمیان تبدیلی کی شدت کا اندازہ اس وقت ہوتا ہے جب ہم متعلقہ متغیرات کی قیمتوں میں تبدیلیوں کا مشاہدہ کرتے ہیں۔ان قوتوں کے اثرات کی درجہ بندی کی جاسکتی ہے۔اس زمرہ بندی کو Components of Time Series کہا جاتا ہے۔ان کی تفصیل درجہ ذیل ہے۔

1.	Secular Trend or Long Term Movement (T)	طویل مدتی حرکت/سیکولر رجحان
2.	Seasonal Variations (S)	موسمی تغیرات
3.	Cyclical Variations (C)	چکری تغیرات
4.	Irregular Variations (I)	غیر مسلسل تغیرات

ایک مظاہر کی قیمت (y) کا کسی وقت مشاہدہ (t) مذکورہ بالا وقتی سلسلوں کے اجزاء کے خالص اثرات ہیں۔ ان اجزاء کی تفصیل ذیل میں دی گئی ہے۔

1۔ سیکولر رجحان/طویل مدتی حرکت (Secular Trend / Long Term Movement)

ایک طویل عرصہ کے دوران وقتی سلسلوں کے اعداد و شمار کے نارمل رجحان میں کبھی اضافہ، کبھی کمی یا الگ الگ بھی ہو سکتے ہیں۔ اس رجحان کو سیکولر رجحان کہتے ہیں۔ لفظ Secular لاطینی زبان کے لفظ "Speculum" سے اخذ کیا گیا ہے۔ اس کے معنی مہد یا نسل کے ہیں۔ اس طرح سیکولر رجحان کا مطلب کسی مظاہر کا طویل مدتی رجحان ہے۔ سمپسن اور کافکا کے الفاظ میں رجحان کو سیکولر یا طویل مدتی رجحان بھی کہتے ہیں۔ یہ دراصل ایک سلسلہ کا بنیادی رجحان ہے۔ ایک مدت کے دوران رجحان میں اضافہ یا کمی بھی ہو سکتی ہے۔ رجحان کے تصور میں مختصر مدت، اہتزاز شامل نہیں، لیکن ایک طویل مدت کے دوران بتدریج مسلسل حرکت ہے۔

a) سیکولر رجحان اوپر کی طرف یا نیچے کی طرف ہوتا ہے۔ مثال کے طور پر ہمارے ملک کی آبادی کا اضافی رجحان ہے۔ اسی طرح بچوں کی شرح اموات میں کمی کا رجحان ہے۔ تاہم اس کا ہرگز یہ مطلب نہیں کہ رجحان ہمیشہ اضافی یا کمی کا ہوتا ہے۔ کبھی رجحان اضافی ہو سکتا ہے اور کبھی کمی کا ہو سکتا ہے یا پھر یکساں اور مستحکم ہو سکتا ہے۔ لیکن عمومی طور پر طویل مدت میں رجحان اضافی، کمی اور مستحکم ہو سکتا ہے۔

b) نارمل طور پر رجحان ایسی قوتوں کا نتیجہ ہوتا ہے جن میں کم و بیش طویل مدت میں بتدریج یا پھر آہستہ تبدیلی عمل میں آتی ہے۔ ان عناصر میں عوام کے ذوق و شوق آبادی اور عادات میں تبدیلیاں، ٹیکنالوجیائی تبدیلیاں، نئے قدرتی وسائل کی دریافت وغیرہ شامل ہیں۔ یہ عناصر بتدریج آہستہ اثرانداز ہوتے ہیں۔

c) اصطلاح طویل مدت ایک اضافی اصطلاح ہے۔ یہ طویل مدت کسی پسِ منظر میں استعمال ہوتی ہے۔ اس کا جواب آسان نہیں ہے۔ لیکن اس جواب کا انحصار زیر مشاہدہ مسئلہ کی نوعیت پر ہوتا ہے، اگر ہم قومی آمدنی، زرعی پیداوار، درآمدات و برآمدات کے چار، پانچ، چھ برس پر مشتمل اعداد و شمار کی ضرورت ہو گی۔ مختصراً ایک عمومی اصول یہ ہے کہ کسی سلسلہ کی مدت طویل ہو تو نتائج بہترین حاصل ہوتے ہیں۔

d) خطی اور غیر خطی رجحانات کا انحصار بھی اعداد و شمار پر ہوتا ہے۔ اگر ہم ان اعداد و شمار کو گراف پر مترسم کریں تو خط سیدھا ہو تو خطی رجحان ہو گا۔ اسی طرح کسی مظاہر کے اعداد و شمار کو مرتسم کرنے پر خط سیدھا نہ تو یہ غیر خطی رجحان ہو گا۔

e) یہ ضروری نہیں کہ وقتی سلسلے ہمیشہ بڑھتے اور گھٹتے رجحان کے حامل ہوں۔ بعض اوقات قیمتوں میں لچک ہو بڑھتے گھٹتے ہوں تو وقت کے گذرنے کے ساتھ اس میں تبدیلی نہیں آتی۔ جیسے درجہ حرارت ہمیشہ تبدیل ہوتا رہتا ہے۔ موسمی حالات ہمیشہ ایک جیسے نہیں رہتے۔ ایسی صورت میں تقابل مشکل ہوتا ہے۔ اس لیے اگر مدت زیادہ ہو تو اعداد و شمار کا ایک سلسلہ ہوتا ہے۔ یہی وجہ ہے کہ ہم دوادوار پر مبنی سلسلوں کا تقابل آسانی سے کر سکتے ہیں۔ اگر ہم رجحانی قدروں کو مظاہر کی قدروں سے الگ کر دیں تو لچک رہ جاتی ہے۔ جیسے موسمی چکری اور غیر مسلسل۔ اب ہمارے لیے یہ ممکن ہو جاتا ہے کہ ہم مختصر مدتی اس لچک کے سلسلوں کا مشاہدہ کریں۔

2۔ موسمی تغیرات (S) (Seasonal Variations)

موسمی تغیرات وقتی سلسلوں کی ان حرکات سے متعلق ہے جو قوتوں کی وجہ سے جن کی نوعیت متوازن ہم آہنگ اور بالترتیب ہوتی ہے اور ہر موسم میں ایک مدت کے بعد دہرائی جاتی ہیں۔ ان کے تغیرات کے درمیان مدت ایک سال سے بھی کم ہوتی ہے۔ انSeasons کی مدت ہفتہ اور ماہوار، سہ ماہی، ششماہی ہو سکتی ہے۔ اس کا انحصار مظاہر کی نوعیت پر ہوتی ہے۔ یہ ایسے عناصر کا نتیجہ ہوتی ہے جن کی مقدار میں یکساں طور پر یا مسلسل کمی و بیشی ہوتی رہتی ہے۔ نارمل طور پر یہ سلسلہ معاشی اعداد و شمار سے متعلق ہوتا ہے جو تبدیل ہوتے رہتے ہیں۔ جیسے اشیاء کی پیداوار، ان کی قیمتیں، صرف کی مقدار، طلب و رسد کی مقدار وغیرہ۔ ان میں تبدیلی موسمیSeasonal ہوتی ہے۔ ان تبدیلیوں کو مسلسل اور یکسانیت کے ساتھ دیکھا جا سکتا ہے۔ مثال کے طور پر کسی موسم میں پھل یا غذائی اجناس کی پیداوار کاٹی جاتی ہے۔ اس موسم میں اس شئے کی قیمتیں کم ہوتی ہیں۔ اس کے برخلاف دوسرے موسم میں قیمتیں زیادہ ہوتی ہیں۔ قیمتوں میں، پیداوار میں، طلب اور رسد میں تبدیلی ہمیشہ ہوتی رہتی ہے۔ اس کی وجوہات کو ذیل میں دیکھا جا سکتا ہے۔

a) موسمی اثرات جیسے موسمی حالات اور آب و ہوا: یہ عناصر ادواری اور موسمی تغیرات میں بہت بڑا کردار ادا کرتے ہیں۔ جیسے ٹھنڈے مشروبات اور آئیس کریم کی طلب موسم گرما میں بڑھ جاتی ہے۔ موسم برسات میں چھتریوں کی فروخت میں اضافہ ہوتا ہے۔ موسم سرما میں چائے اور گرم ملبوسات کی فروخت میں اضافہ ہوتا ہے۔

b) مقامی رہن سہن: عادات رسم و رواج کا مختلف موسم میں الگ الگ اثر پڑتا ہے۔ اسی طرح شادیوں کے سیزن میں نارمل طور پر سونے کے بھاؤ میں اضافہ ہو جاتا ہے۔ عیدین، رمضان، دیوالی، دسہرہ، کرسمس وغیرہ کے موقع پر کئی اشیاء کی قیمتوں میں اضافہ ہو جاتا ہے۔ سال بہ سال ان کی مقدار طلب و رسد، اشیاء خوردنی وغیرہ کی قیمتوں میں تبدیلی کی پیش قیاسی آسانی سے کی جا سکتی ہے۔ کیونکہ ان عوامل کا اثر مدتی اور مسلسل ہوتا ہے۔

سلسلوں پر مبنی تغیرات کا مشاہدہ و مطالعہ تاجر، صارفین پیدا کنندے، فیصلہ سازی میں مصروف عمل افراد کے لیے مفید ثابت ہوتا ہے۔

3۔ چکری تغیرات (C) (Cyclical Variations)

کسی سلسلے میں چکری تغیرات متوالی تغیرات ہوتے ہیں۔ جن کا دورانیہ ایک سال سے زاید ہو سکتا ہے۔ چکری تغیرات اگرچیکہ مسلسل ہوتے ہیں لیکن ان کے دورانیہ میں یکسانیت نہیں ہوتی۔ ایسا اس لیے ہوتا ہے کہ دورانیہ ہمیشہ یکساں نہیں ہوتا۔ ان تغیرات کا بغور

مشاہدہ کیا جائے تو معلوم ہوگا کہ معاشیات اور کاروبار میں چکری تغیرات دراصل کاروباری چکر کا نتیجہ ہوتے ہیں۔کاروباری چکر کا آغاز خوشحالی Boom سے ہوتا ہے۔اس کے نیچے کی طرف حرکت کرتا ہے۔یعنی ریگریشن Recession کے دور میں آتا ہے۔مزید نیچے کی طرف حرکت کرنے پر سرد بازاری Depression کی طرف آجاتا ہے۔یوں کاروباری چکر کے چار مراحل ہوتے ہیں۔خوشحالی، ریگریشن، سرد بازاری اور بحالی۔ان مراحل کو یکساں طور پر دیکھا جاسکتا ہے۔لیکن دورانیہ مستقل ودائمی نہیں ہوتا۔ماہرین معاشیات کا کہنا ہے کہ بعض اوقات یہ دورانیہ تین سال، چار سال، پانچ سال کبھی سات یا دس سال میں مکمل ہوتا ہے۔

4۔ غیر مسلسل تغیرات (Irregular Variations)

بعض اوقات غیر مسلسل تغیرات مختلف عناصر کی وجہ سے وجود میں آتے ہیں۔یہ عناصر کبھی کبھار اثرانداز ہوتے ہیں۔ان کی نوعیت حادثاتی ہوتی ہے۔جیسے قحط، ہڑتال، تالہ بندی، سیلاب، جنگیں، زلزلے وغیرہ یہ کسی خاص مدت کے دوران وقوع پذیر نہیں ہوتے، بلکہ ہنگامی طور پر سامنے آجاتے ہیں۔اس لیے انہیں Random Fluctuations یا Chance Fluctuations کہتے ہیں۔بعض اوقات ان کی شدت اتنی ہوتی ہے کہ چکری لچک میں اضافہ ہو جاتا ہے۔چونکہ غیر متوالی ہوتے ہیں۔اس لیے اس نقصانات سے نمٹنے یا تلافی کے بچاؤ اقدامات نہیں کیے جاسکتے۔ان کے اثرات بھی غیر متوقع اور غیر یقینی ہوتے ہیں۔لیکن ان کے اثرات کو چکری اور موسمی اثرات سے الگ نہیں کیا جاسکتا۔

9.6 ابتدائی ترمیمات (Introductory Modifications)

وقتی سلسلوں کے تجزیہ سے قبل دستیاب اعداد و شمار میں چند ضروری تبدیلیاں کرنا ہوتا ہے۔تاکہ اعداد و شمار میں یکسانیت لائی جاسکے۔ نارمل طور پر درجہ ذیل نکات میں یکسانیت لائی جاتی ہے۔

1۔ معیادی کمی وبیشی (Time Disparity)

اگر اعداد ماہوار دستیاب ہیں تو معیادی کمی وبیشی کی ضرورت ہوتی ہے۔کیونکہ سال کے تمام مہینوں میں دنوں کی تعداد یکساں نہیں ہوتی۔جیسے جنوری میں 31، فبروری میں 28 یا 29، مارچ میں 31 وغیرہ۔اگر تاجر کے پاس ماہانہ پیداواری سلسلہ ہو تو اور ہر دن پیداوار میں تبدیلی نہ ہو یکسانیت ہو تو ایسی صورت میں ہر مہینہ میں دنوں کی تعداد یکساں نہ ہونے کی وجہ سے ہر مہینہ پیداوار مختلف ہوگی۔ تبدیلی میں بھی یکسانیت نہیں ہوگی۔مثال کے طور پر جنوری، فبروری، مارچ، جون، جولائی، اگست وغیرہ میں پیداوار میں تبدیلی ہوگی۔اگر ان اعداد و شمار میں یکسانیت نہ ہو تو تجزیہ کے نتائج صحیح ودرست نہ ہوں گے۔اس لیے اعداد و شمار کی اس کمی وبیشی کو ہر ماہ کی پیداوار کو اس مہینہ کی دنوں کی تعداد سے بھی تقسیم کر کے اعداد و شمار میں یکسانیت لائی جاتی ہے۔ہر دن کی اوسط پیداوار کو 12/365 سے ضرب دے کر تقسیم کیا جاتا ہے۔سال کبیسہ Leap Year ہونے کی صورت میں 12/366 سے ضرب دے کر تقسیم کیا جائے گا۔

2۔ آبادی میں تبدیلیاں(Population Changes)

آبادی میں تبدیلی اس وقت ضروری ہو جاتی ہے جبکہ آبادی میں تبدیلی کی وجہ سے متغیر پر اثرات مرتب ہوتے ہوں۔ جیسے اگر ہم قومی آمدنی کے اعداد و شمار کا تجزیہ کر رہے ہوں یا پیداوار کے اعداد و شمار کا تجزیہ کر رہے ہوں تب آبادی میں بھی تبدیلیاں کرنا ضروری ہوتا ہے۔ قومی آمدنی میں اس وقت اضافہ ہوتا ہے جبکہ کام سے وابستہ آبادی میں تبدیلی ہو۔ کسی شئے کی پیداوار میں بھی اس وقت تبدیلی ہوتی ہے جب کام پر لگائے گئے مزدوروں کی تعداد میں کمی ہوتی ہے۔ دونوں صورتوں میں اعداد و شمار میں یکسانیت متعلقہ کام سے وابستہ افراد کی تعداد سے تقسیم کر کے لائی جاسکتی ہے۔ ایسا کرنے سے فی کس آمدنی اور فی مزدور پیداوار کو بآسانی معلوم کیا جاسکتا ہے۔ اعداد و شمار میں یکسانیت ہونے کی وجہ سے تجزیہ بھی ثمر آور ہوتا ہے۔ ان نتائج پر اعتماد کیا جاسکتا ہے۔

3۔ قیمتوں میں تبدیلیاں(Price Changes)

جہاں کہیں قیمتوں میں تبدیلی ناگزیر ہو جائے اور جہاں کہیں ہمارے پاس رواں قیمتوں کا سلسلہ ہو اور ہم حقیقی قدر میں تبدیلیاں چاہتے ہوں تو تطبیقات لازماً کرنا پڑتا ہے۔ جبکہ ہمارے پاس قیمت سے متعلق سلسلے موجود ہوں۔ اور ہمارے دلچسپی کا مرکز مقداری تبدیلیوں کا مشاہدہ مطالعہ ہو۔ پہلی صورت میں رواں قدروں کو رواں قیمتوں کے تناسب کو اساسی قیمتوں پر محمول کریں، جیسے

$$\text{Current Value x}\frac{\textit{Base Price Index}}{\textit{Current Price Index}}$$

4۔ اعداد و شمار کی تقابل پذیری(Comparability of the Data)

ٹھوس اور مضبوط نتائج حاصل کرنے کے لیے وہ اعداد و شمار جن کا تجزیہ کیا جارہا ہو تقابلی جائزے کے لائق ہو۔ تاہم جب ہم وقتی سلسلوں کے تعلق سے نتائج اخذ کر رہے ہوں تو ضروری ہے کہ سال گذشتہ کے اعداد و شمار بھی ہوں۔ اعداد و شمار میں یکسانیت ایک مسئلہ ہوتا ہے۔ بعض اوقات اصطلاح تبدیلی کی تعریف اور بعض اوقات نمونہ Sample کے سائیز میں تبدیلی یا پھر کبھی جغرافیائی حدود میں تبدیلی کی تعریف مختلف ہونے کی وجہ سے اعداد و شمار تقابل پذیری نہیں ہوتے۔ اس لیے اعداد و شمار کی نوعیت میں یکسانیت لائی جا کر تقابل پذیر بنایا جاتا ہے۔

وقتی سلسلوں کا تجزیہ(Analysis of Time Series)

وقتی سلسلوں کا تجزیہ دو اہم اقدامات پر مشتمل ہوتا ہے۔ ان کو ذیل میں بیان کیا جاتا ہے۔

1) ان قوتوں کی نشاندہی جن کے تنور ملل سے وقتی سلسلوں میں تغیرات پیدا ہوتے ہیں۔

2) دیگر عناصر کو برقرار رکھتے ہوئے ان قوتوں کے اثرات کو الگ کر کے تجزیہ اور پیمائش آزادانہ کریں۔

وقتی سلسلوں کے تجزیہ کے لیے ریاضیاتی ماڈلس(Mathematical Models for Analysis of Time Series)

وقتی سلسلوں کے تجزیہ کے لیے کئی ماڈلس ہیں۔ لیکن وقتی سلسلوں سے مختلف اجزاء کے انحلال کے لیے عمومی طور پر دو ماڈلس استعمال کیے جاتے ہیں۔ ان کو ذیل میں بیان کیا جاتا ہے۔

1۔ جمعی ماڈل/پیمانہ(Addition Model)

اس طریقہ کے مطابق وقتی سلسلوں کا انحلال اس مفروضہ پر کیا جاتا ہے کہ مختلف اجزاء کے اثرات کی نوعیت جمعی ہوتی ہے، یعنی

$$y = T + S + C + I$$

جہاں

y = Time Series

T = Trend

S = Seasonal

C = Cyclical

I = Irregular Variation

اس ماڈل کے مفروضہ کے مطابق وقتی سلسلوں کے چار اجزاء ایک دوسرے سے آزاد ہوتے ہیں۔ کسی ایک عنصر کا باقی تین عناصر پر کوئی اثر نہیں پڑتا۔ یعنی رجحان Trend کا موسمی، چکری، غیر مسلسل تغیر پر اثر نہیں پڑتا۔ اسی طرح یہ بھی فرض کیا جائے کہ موسمی کا چکری پر اور چکری کا غیر مسلسل تغیر پر کوئی اثر نہیں پڑتا۔ لیکن اس کے برعکس عملی نقطۂ نظر سے دیکھا جائے تو یہ مفروضہ حقیقت سے بعید ہے۔ یہ عناصر ایکدوسرے پر اثرانداز ہوتے ہیں۔

2۔ ضربی ماڈل(Multiplication Model)

ضربی ماڈل کے مطابق وقتی سلسلوں کا انحلال اس مفروضہ پر کیا جاتا ہے کہ وقتی سلسلوں کے چار اجزاء(T, S, C & I) کے اثرات ایکدوسرے سے الگ نہیں ہوتے۔ حقیقت میں اس ماڈل کا یہ مفروضہ ہے کہ ان کے اثرات باہمی طور پر مربوط ہیں، یعنی

$$y = T \times S \times C \times I$$

C,S اور I اضافی تغیرات ہیں، اور انہیں شرح یا اعشاری اعداد کی صورت میں بتایا جاتا ہے اور یہ عدد ایک سے اوپر اور نیچے قیمت بدلتے رہتے ہیں۔ یہ فرض کیا جاتا ہے کہ C,S اور I کا جیومیٹریائی اوسط ایک ہوگا۔ اس ضربی ماڈل کو Logs کی اصطلاح میں ظاہر کیا جاسکتا ہے۔

$$\log y = \log T + \log S + \log C + \log I$$

مذکورہ بالا کا یہ مطلب ہوا کہ دیے گئے وقتی سلسلوں کی قدروں کو logs کی صورت میں ضربی ماڈل جمعی ماڈل ہے۔ کاروبار اور معاشیات میں ضربی ماڈل کو اپناتے ہیں۔ اسی طرح کئی مخلوط ماڈل بھی ہوتے ہیں۔ جیسے

$$y = TCS + I$$

$$y = TC + SI$$

$$y = T + SCI$$

$$y = TS + CI$$

جمعی اور ضربی ماڈل تفریق اور تقسیم کے طریق عمل میں وقتی سلسلوں کے مختلف اجزاء کی قدروں کو حاصل کرنے کے لیے استعمال کرتے ہیں۔ اگر ہمیں رجحانی قدر یں معلوم ہوں تو تب جمعی ماڈل کے مطابق

$$y - T = S - C + I$$

ضربی ماڈل میں

$$SxCxl=\frac{y}{T}=\frac{Original\ Value}{Trend\ Values}$$

اگر اعداد و شمار سالانہ ہوں تب اس میں موسمی اتار چڑھاؤ نہ ہوں گے۔

$$y = T \times C \times I$$

یا

$$C \times I = \frac{y}{T}$$

اس طرح وقتی سلسلوں کے تجزیہ کے لیے ہمیں پہلے رجحانی قدریں معلوم کرنا ہوگا۔

9.7 رجحان کی پیمائش (Measurement of Trend)

رجحانی قدروں کے تعین کے لیے درجہ ذیل طریقے پائے جاتے ہیں۔

1) ترسیمی/اختیاری طریقہ (Freehand / Graphic Method)
2) نصف اوسط کا طریقہ (Method of Semi - Averages)
3) حرکتی اواسط کا طریقہ (Method of Moving Averages)
4) اقل ترین مربعوں کا طریقہ (Method of Least Squares)

1۔ ترسیمی/اختیاری طریقہ (Freehand / Graphic Method)

رجحانی قدروں کی دریافت کا سب سے سادہ اور آسان طریقہ ہے۔ سب سے پہلے وقتی سلسلے کی قیمتوں کو ترسیمی کاغذ پر ہسٹو گرام (مستطیلی ترسیم) کی شکل میں مرتسم کر دیا جاتا ہے۔ اس کے بعد ان ہسٹو گرام کے کناروں کو ملاتے ہوئے ہاتھ سے ایک قوس کھینچی جاتی ہے۔ یہ قوس اعداد و شمار کے عمومی رجحان کی نمایندگی کرتی ہے۔ ایک ہموار قوس کھینچنے کے طریقِ عمل کے دوران موسمی، چکری اور غیر مسلسل تغیرات نکال دیے جاتے ہیں۔ تاہم اختیاری طریقہ کے تحت ہموار قوس کھینچنا ایک مشکل کام ہے۔ کیونکہ ہر شخص اپنے اپنے اسٹائیل سے قوس کھینچتا ہے، جبکہ اعداد و شمار وہی ہوتے ہیں۔

فائدے (Advantages)

ترسیمی طریقہ کے فایدے درجہ ذیل ہیں۔

1) اختیاری طریقہ کے تحت رجحانی ترسیم وقتی سلسلوں کی خصوصیت کو سمجھنے میں مددگار ہے،ان خصوصیات میں اتار چڑھاؤ اور سادگی ہے۔

2) یہ طریقہ خطی اور غیر خطی تمام قسم کے رجحان کو بیان کرنے میں استعمال کیا جاتا ہے۔

نقصانات(Disadvantages)

ترسیمی طریقہ کے نقصانات درجہ ذیل ہیں۔

1) انفرادی اختیاری تمیزی اور فیصلہ پر حد سے زیادہ انحصار کرنا ہے جو کہ غلط بھی ہو سکتا ہے۔

2) اس میں پیچیدہ ریاضیاتی ٹیکنکس شامل نہیں ہیں، جس کی وجہ سے اس پر اعتماد کرنا مشکل ہوتا ہے۔

3) یہ طریقہ مقداری تناظر میں رجحان کی پیائش کرنے میں مددگار نہیں ہے۔

4) ایک ہی قسم کے اعداد و شمار کے لیے مختلف افراد کو مختلف رجحانی قوسیں حاصل ہوتی ہیں۔ جس کی وجہ سے نتائج اخذ کرنے میں مشکل ہوتی ہے۔

طریقہ کار

اس طریقہ کے تحت گراف پر اعداد و شمار اور قوس کھینچنے کا طریقہ ذیل میں دیا گیا ہے۔

1) ترسیم Graph پر وقتی سلسلہ کے اعداد و شمار کو مرتسم کریں۔

2) مرتسم کردہ معلومات یعنی نقاط کے رجحان کی سمت کا جائیزہ لیں۔

3) اپنی اختیار تمیزی کے مطابق سیدھا خط کھینچیں جو اعداد و شمار کی بہترین ترجمانی کرتا ہو، یہی خط رجحان کی سمت کو ظاہر کرے گا۔

تمثیل 1 : درجہ ذیل اعداد و شمار کے لیے اختیاری طریقہ سے رجحانی خط مرتسم کیجیے۔

سال	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
آمدنی روپیوں میں	30	40	50	40	55	60	65	60	70	80

2۔ نصف اواسط کا طریقہ (Method of Semi Averages)

اس طریقہ Semi Averages تحسیب کرکے رجحانی قیمتیں معلوم کی جاتی ہیں۔ نصف اوسط کا مطلب سلسلہ کو دو حصوں میں تقسیم کرکے اوسط نکالا جاتا ہے۔ جیسے اگر اعداد و شمار 2011-2020 تک دیے جائیں۔ ہمیں اندازہ ہے کہ یہ اعداد و شمار دس سال پر مشتمل ہیں۔ پہلا نصف 2011-2015 اور دوسرا نصف 2016-2020 ہوگا۔ اوسط کا مطلب حسابی اوسط ہے۔ ان دونوں کا اوسط معلوم کیا جاتا

ہے۔ان اوسط کی دونوں قیمتوں کو ہر نصف کے درمیان میں مرتسم کیا جاتا ہے اور پھر ان دو نقاط کو ایک سیدھی لکیر کے ذریعہ ملادیا جاتا ہے۔یہی رجحانی خط ہو گا۔

	دوسرا نصف	پہلا نصف	
	2016-2020	2011-2015	مدت
	$\bar{x}_2 = \frac{y_6+y_7+y_8+y_9+y_{10}}{5}$	$\bar{x}_1 = \frac{y_1+y_2+y_3+y_4+y_5}{5}$	نیم/نصف اوسط
	$\bar{x}_2 = \frac{2016+2017+2018+2019+2020}{5}$	$\bar{x}_1 = \frac{2011+2012+2013+2014+2015}{5}$	
	$= \frac{10,090}{5}$	$= \frac{10,065}{5}$	
	$= \underline{2018}$	$= \underline{2013}$	

تمثیل 2 : نصف/نیم اوسط کو استعمال کرتے ہوئے درجہ ذیل اعداد و شمار سے رجحان کا تعین کیجیے۔

سال	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
نفع	50	47	60	55	40	65	70	65	70	75

اب ترسیمی کاغذ پر 51 کو پہلے پانچ سال کے درمیان میں 50.4 کو اور دوسرے پانچ سال کے درمیان 69 کو مرتسم کیا جائے گا اور ان نقاط کو سیدھی لکیر کے ذریعہ جوڑا جائے گا۔

1۔ نیم اوسط کے طریقہ کے فایدے (Advantages of Semi - Average Method)

نیم اوسط کے طریقہ کے فایدے درجہ ذیل ہیں۔

1) یہ طریقہ نہایت سادہ اور سمجھنے میں آسان ہے۔

2) اس کو پیمائش اقل ترین مربعوں کے طریقہ کے حرکتی اوسط سے تقابل کیا جاسکتا ہے۔

3) اس کا سب سے بڑا فایدہ شخصی فیصلہ پر منحصر نہیں ہوتا۔

4) اس کا اطلاق خطی رجحان یا تقریباً خطی رجحان پر بھی ہوتا ہے۔

2۔ نیم اوسط کے طریقہ کے نقصانات (Disadvantages of Semi - Average Method)

نیم اوسط کے طریقہ کے نقصانات درجہ ذیل ہیں۔

1) یہ طریقہ اس مفروضہ پر ہے کہ خطی رجحان ہوتا ہے، جبکہ یہ صحیح نہیں ہے۔

2) یہ اس وقت موزوں نہیں ہے جبکہ اس کی پیشکشی کا اوسط چھوٹا ہو۔

3) نیم اوسط کے حصول کے لیے اوسط حسابیہ کا استعمال بھی مناسب نہیں، کیونکہ اس کے اسعمال کے نقصانات ہیں۔

3۔ حرکتی اوسط کا طریقہ(Method of Moving Average)

حرکتی اوسط اتار چڑھاؤ کو کم کرنے کا ایک سادہ آلہ ہے۔ ساتھ ہی اس کے نتائج صحیح اور درست ہوتے ہیں۔ اس طریقہ میں حرکتی اوسط کے درمیانی نقطہ کے لیے اوسط نکالا جاتا ہے۔ اس طریقہ کار میں قوس سطح اور ہموار ہوتی ہے اور اتار چڑھاؤ میں بھی کمی کا باعث ہوتی ہے۔ اس طریقہ میں حرکتی اوسط معلوم کرنے کے لیے مدت کے تعلق سے فیصلہ لیا جاتا ہے۔ مثال کے طور پر اگر حرکتی اوسط کے لیے مدت پانچ سال مقرر کی جاتی ہے تب پہلے سال کے ایٹمس کا اوسط نکالا جاتا ہے اور اس کے سامنے لکھ دیا جاتا ہے۔ پھر 2، 3، 4، 5، 6 کے ایٹمس کا اوسط نکالکر اس کے سامنے لکھ دیا جاتا ہے۔ پھر 3، 4، 5، 6، 7 کے ایٹمس کا اوسط نکال کر اس کے سامنے لکھ دیا جاتا ہے۔ اس عمل کو اس وقت تک جاری رکھا جاتا ہے جب تک کہ آخری پانچ ایٹمس شامل نہ ہو جائیں۔ اس کو ایسا بھی بتایا جاسکتا ہے۔

پہلی قدر کا حرکتی اوسط $= \frac{I}{t}(x_1 + x_2 + x_3..........)$

دوسری قدر کا حرکتی اوسط $= \frac{I}{t}(x_2 + x_3 + x_4..........)$

تیسری قدر کا حرکتی اوسط $= \frac{I}{t}(x_3 + x_4 + x_5..........)$

4۔ حرکتی اوسط کے طریقہ کے فائدے(Advantages of Method of Moving Average)

حرکتی اوسط کے طریقہ کے فایدے درجہ ذیل ہیں۔

1) دوسرے طریقوں کے مقابلے میں یہ طریقہ نہایت سادہ اور تحسیب کے لیے آسان ہے۔

2) وقتی سلسلوں کے تجزیہ میں استعمال کیا جاسکتا ہے۔ چاہے وہ رجحانات موسمی ہوں، چکری یا غیر مسلسل اجزا ہوں۔

3) اس کی تحسیب ریاضیاتی تحسیب کی طرح پیچیدہ نہیں ہے۔

4) اس طریقہ میں اوسط کے تعین کے لیے سلسلے کی تمام قیمتوں کو حساب میں لیا جاتا ہے۔

5) یہ ایک معروضی طریقہ ہے، شخصی قیاس اور فیصلوں کی گنجائش نہیں ہوتی۔

5۔ حرکتی اوسط کے طریقہ کے نقصانات(Disadvantages of Method of Moving Average)

حرکتی اوسط کے طریقہ کے نقصانات درجہ ذیل ہیں۔

1) اس طریقہ میں ابتدائی اور اختتامی اعداد کو چھوڑ دینا ہوتا ہے۔

2) اس طریقہ میں چکر یا دیگر حرکات سامنے آسکتے ہیں، جبکہ ان وجود اصل اعداد و شمار میں نہیں ہوتا۔

3) اس طریقہ کے تحت مدت اور متغیرات کی قدر کے درمیان کو تفاعلی تعلق قایم نہیں کیا جاسکتا۔

4) اس طریقہ کے تحت غیر مسلسل تغیرات کا مکمل تعین نہیں کیا جاسکتا۔

تمثیل 3: تین سالہ حرکتی اوسط کے طریقہ سے رجحانی قدریں دریافت کیجیے۔

سال	2014	2015	2016	2017	2018	2019	2020
نفع	56	69	73	77	85	92	95

حل:

سال	نفع	تین سالہ حرکتی میزان	تین سالہ حرکتی اوسط	مرکزی قیمت Central Value
2014 2015 2016 2017 2018 2019 2020	56 69 73 77 85 92 95	198 219 235 254 272	=66.0$\frac{198}{3}$ =73.0$\frac{219}{3}$ 378.33$\frac{235}{3}$ =84.67$\frac{254}{3}$ =90.67$\frac{272}{3}$	69.5 75.5 81.5 87.67 2020

تمثیل 4: درجہ ذیل اعداد و شمار کے لیے چار سالہ حرکتی اوسط تحسیب کیجیے۔

سال	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
پیداوار (ٹنوں میں)						270	280	285	295	305

حل:

سال	پیداوار	چار سالہ میزان	چار سالہ اوسط
2011	250		
2012	260	250+260+275+235=1120	$\frac{1120}{4}=280.0$
2013	275	260+275+235+255=1025	$\frac{1025}{4}=256.25$
2014	235	275+235+255+270=1035	$\frac{1040}{4}=260.0$
2015	255	235+255+270+280=1040	$\frac{1090}{4}=272.5$
2016	270	255+270+280+285=1090	$\frac{1130}{4}=282.5$
2017	280	270+280+285+295=1130	$\frac{1165}{4}=291.25$
2018	285	280+285+295+305=1165	
2019	295		
2020	305		

4۔ اقل ترین مربعوں کا طریقہ (Least Squares Methods)

رجحانی قدروں کو حاصل کرنے کا سب سے اچھا طریقہ اقل ترین مربعوں کا طریقہ ہے۔ اس طریقہ کو اپنا کر سیدھا رجحانی خط حاصل کیا جاسکتا ہے۔ ان خطوط کو Line of best Fit کہتے ہیں۔ یہ وہ خط ہے جہاں مختلف نکات سے انحراف کا مجموعہ صفر کے برابر ہوتا ہے۔ یعنی خط ایک جانب مختلف نقاط کا عمودی فاصلہ کی پیمائش اور میزان کیا جائے تو حاصل عدد دوسری جانب مختلف نقاط کے فاصلہ کی پیمائش اور میزانی کے برابر ہوگا۔ ایسا اس لیے ہوتا ہے کیونکہ ان انحراف کے مربعوں کا مجموعہ دوسرے خطوط کے استعمال سے حاصل انحراف کے مربعوں کے مجموعہ سے کم ہوگا۔ اوسط حسابیہ سے انحراف کا مجموعہ صفر ہوتا ہے اور یہاں لائین آف بیسٹ فٹ سے انحراف کے مربعوں سے کم ہوتا ہے۔ اسی حقیقی تصور کی بنیاد پر اس طریقہ کو Method of Least Square کہتے ہیں۔

فائدے (Advantages)

اقل ترین مربعوں کے طریقہ کے فائدے درجہ ذیل ہیں۔

(1) یہ ایک مقصدی طریقہ ہے جس کے نتائج میں اتار چڑھاؤ نہیں ہوتا۔

(2) اس کی تحسیب نہایت آسان ہے۔

(3) یہ طریقہ رجحانی قدروں کا تعین کرتا ہے اور موسمی، چکری اور غیر مسلسل تغیرات پر روشنی ڈالنے کا باعث ہے۔

(4) چونکہ اس طریقہ میں رجحان کی تحسیب کی بنیاد الجبراء پر ہے، اس لیے یہاں قیاس کے لیے کوئی جگہ نہیں۔

(5) یہ ایک لچکدار طریقہ ہے۔ کسی بھی دور کے رجحانی قدریں تحسیب کیجیے جواب بالکل درست ہوتا ہے۔

نقصانات (Disadvantages)

اقل ترین مربعوں کے طریقہ کے نقصانات درجہ ذیل ہیں۔

(1) یہ طریقہ مختصر سلسلوں کے لیے موزوں ہے اور ایسے سلسلوں کے لیے غیر ضروری ہے۔

(2) مدت/دور (x) اور ایک متغیر (y) کی قیمتوں کے درمیان تفاعلی تعلق قائم نہیں کرتا۔

(3) یہ ایک مشقت طلب طریقہ ہے۔ دوسرے طریقوں کی بہ نسبت حسابی عمل زیادہ ہوتا ہے۔

(4) یہ صرف مستقبل قریب کی قیمت کا تخمینہ کر سکتا ہے۔ لیکن مستقبل متغائیر کا تخمینہ نہیں کر سکتا۔

'a' اور 'b' کی قدروں کا تخمینہ ('Computation of Values of 'a' and 'b)

'a' اور 'b' کی قدروں کو درجہ ذیل دو مساوات کو حل کر کے حاصل کیا جاسکتا ہے۔

$$\sum y = Na + b\sum x \text{..........1}$$

$$\sum xy = a\sum x + b\sum x^2 \text{.....2}$$

یہاں 'N' رجحانی قدروں کی مدت ہے اور 'x' وقت کی اکائی ہے۔

اس طرح y کا مطلب حقیقی قدریں ہیں۔ y_c - na

مساوات کو بتایا جاسکتا ہے۔

$$y = a + b\,x$$

رجحانی قدروں کے لیے اقل ترین مربعوں کا طریقہ

$$y_c = a + bx$$

$$a = \frac{y}{n}, \quad b = \frac{xy}{x^2}$$

مساوات کی مزید وضاحت اچھی طرح سمجھ لیں۔

(انحراف) آزادانہ متغیر	$x =$
جدار کھنا y intercept	$a =$
ڈھلان کی نشاندہی اور 'x' میں تبدیلیوں کو نمایاں کرنا	$b =$
مشاہدہ کی تعداد	$n =$
xاور y کا حاصل ضرب	$xy =$
x کا مربع	$x^2 =$
y کی قدروں کا مجموعہ	$\sum y =$
x کی قدروں کا مربع	$\sum y^2 =$

تمثیل4: اقل ترین مربعوں کے طریقہ سے رجحانی خطِ مستقیم مرتسم کیجیے۔اور 2021 میں نفع کی پیش بینی کیجیے۔

سال	2013	2014	2015	2016	2017	2018	2019
نفع(روپیوں میں)	80	90	92	83	94	99	92

حل:

رجحانی قدروں کی تخمین

سال	نفع x	x	x^2	xy	رجحانی قدریں Trend Values (y)
2013	80	-3	9	-240	84
2014	90	-2	4	-180	86
2015	92	-1	1	-92	88
2016	83	0	0	0	90
2017	94	+1	1	+94	92
2018	99	+2	4	+198	94
2019	92	+3	9	+276	96
$n = 7$	30	$\sum x = 0$	$\sum x^2 28$	56	

خطِ مستقیم کی مساوات $y_c = a + bx$

$$a = \frac{\sum y}{n} = \frac{630}{7} = 90, \quad b = \frac{\sum xy}{\sum x^2} = \frac{56}{28} = 2$$

$$y_c = a + bx$$

2013 = 90+2(-3) = 90-6=84

2014 = 90+2(-2) = 90-4=86

2015 = 90+2(-1) = 90-2=88

2016 = 90+2(0) = 90-0=90

2017 = 90+2(+1) = 90+2=92

2018 = 90+2(+2) = 90+4=94

2019 = 90+2(+3) = 90+6=96

2021 = 90+2(5) = 90+10=100 پیش بینی

9.8 اکتسابی نتائج (Learning Outcome)

طلباء نے ٹائم سیریز کے تجزیہ کے معنی اور تعریف اور وقت کے ساتھ جمع کیے گئے ڈیٹا پوائنٹس کی جانچ میں اس کے کردار کے بارے میں واضح سمجھ پیدا کرلی ہے۔ انہوں نے ٹائم سیریز کی نوعیت اور دائرہ کار سیکھ لیا ہے اور یہ کہ ڈیٹا میں پیٹرن، رجحانات اور تغیرات کی شناخت کے لیے اسے کس طرح استعمال کیا جاتا ہے۔ طلباء پیشن گوئی اور حکمت عملی کی منصوبہ بندی میں ٹائم سیریز کے تجزیہ کی اہمیت اور افادیت کی وضاحت کر سکتے ہیں۔ انہوں نے ٹائم سیریز کے اجزاء کا بھی مطالعہ کیا ہے، بشمول رجحان، موسمی، چکری، اور فاسد تغیرات، اور ڈیٹا کا مؤثر طریقے سے تجزیہ کرنے کے لیے تعارفی ترمیم کو سمجھتے ہیں۔ آخر میں، سیکھنے والے رجحان کی پیمائش کے لیے طریقوں کو لاگو کر سکتے ہیں، جو انہیں تاریخی ڈیٹا کی تشریح کرنے اور کاروباری فیصلہ سازی کے لیے باخبر پیشین گوئیاں کرنے کے قابل بنا سکتے ہیں۔

9.9 نمونہ امتحانی سوالات (Model Examination Questions)

معروضی جوابات کے حامل سوالات

1) __________ کا تجزیہ ایک شماریاتی ٹیکنک ہے۔ (وقتی سلسلے)

2) لفظ __________ لاطینی زبان سے اخذ کیا گیا ہے۔ (سیکولر)

3) __________ متوالی تغیرات ہوتے ہیں۔ (چکری تغیرات)

4) غیر مسلسل تغیرات کی مثالیں __________ ہیں۔ (قحط، تالہ بندی، سیلاب، جنگیں)

5) سال کبیسہ میں __________ دن ہوتے ہیں۔ (366)

6) آج کی دنیا ٹیکنک رُخی ہے۔ (صحیح/غلط)

7) وقتی سلسلوں کے اعداد و شمار کی خصوصیات یکساں ہونی چاہیے۔ (صحیح/غلط)

8) رجحان کو سیکولر یا طویل مدتی رجحان نہیں کہتے۔ (صحیح/غلط)

9) ہنگامی حالات غیر متوالی ہوتے ہیں۔ (صحیح/غلط)

10) جمعی ماڈل میں اجزا کے اثرات کی نوعیت جمعی نہیں ہوتی۔ (صحیح/غلط)

مختصر جوابات کے حامل سوالات

1) وقتی سلسلوں کا تجزیہ کیا ہے؟

2) وقتی سلسلوں کے تجزیہ کی تعریف کیجیے۔

3) درجہ ذیل پر مختصر نوٹ لکھیے۔

i. سیکولر رجحان ii. موسمی تغیرات iii. چکری تغیرات iv. غیر مسلسل تغیرات

4) جمعی ماڈل کسے کہتے ہیں؟

5) اختیاری طریقہ کیا ہے؟

طویل جوابات کے حامل سوالات

1) وقتی سلسلوں کے تجزیہ کی تعریف کیجیے اور ان کی اہمیت کو واضح کیجیے۔

2) وقتی سلسلوں کے اجزاء کو بیان کیجیے۔

3) ابتدائی تطبیقات کو تفصیل سے بیان کیجیے۔

4) رجحان کی پیمائش کے مختلف طریقوں کو تفصیل سے بیان کیجیے۔

5) درجہ ذیل اعداد و شمار کے رجحان کو اختیاری طریقہ سے گراف کیجیے۔

سال	2014	2015	2016	2017	2018	2019
نفع (روپیوں میں)	100	120	95	105	108	102

6) سادہ نیم اواسط کے طریقہ کے تحت درجہ ذیل اعداد و شمار کے لیے رجحان کا تعین کیجیے۔

سال	2001	2002	2003	2004	2005	2006	2007	2008	2009
پیداوار (ٹنوں میں)	16	14	20	18	19	12	23	28	25

7) نیم اواسط کے طریقہ سے درجہ ذیل اعداد و شمار کے لیے رجحانی خط مرتسم کیجیے۔

سال	2016	2017	2018	2019	2020	2021
فروخت	100	120	140	150	130	200

8) درجہ ذیل اعداد سے تین سالہ حرکتی اواسط کے طریقہ سے رجحان کا تعین کیجیے۔

سال	2012	2013	2014	2015	2016	2017	2018
خرید (ہزاروں میں)	4	12	2	10	6	14	4

9) درجہ ذیل اعداد و شمار کو استعمال کرتے ہوئے تین سالہ حرکتی اواسط کے طریقہ سے رجحان کو مرتسم کیجیے۔

سال	2010	2011	2013	2013	2014	2015	2016	2017	2018	2019	2020	2021
نفع	85	80	74	88	90	83	85	80	74	58	70	53

10) درجہ ذیل اعداد و شمار کی مدد سے سیدھا خطی رجحان اقل ترین مربعوں کے طریقہ سے ترتیب دیجیے۔

سال	2015	2016	2017	2018	2019
مدات	45	46	35	46	85

11) اقل ترین مربعوں کے طریقہ سے رجحانی قدروں کی تحسیب کیجیے اور 2012 کے لیے پیداوار کا تخمینہ کیا ہوگا۔

سال	2016	2017	2018	2019	2020
پیداوار (ٹنوں میں)	18	24	28	32	40

12) ذیل میں سالانہ فروخت کی مقدار دی گئی۔

سال	2013	2014	2015	2016	2017	2018	2019
فروخت	30	33	40	41	43	44	45

a) اقل ترین مربعوں کے طریقہ سے رجحانی خط مستقیم مرتسم کیجیے۔

b) 2022 کے لیے فروخت کا تخمینہ کیا ہوگا۔

13) درجہ ذیل پیداوار سے متعلق اعداد و شمار کو استعمال کرتے ہوئے اقل ترین مربعوں کے طریقہ سے رجحانی خط مستقیم مرتسم کیجیے اور رجحانی قدروں کی تحسیب کیجیے اور 2022 کے لیے پیداوار کا تخمینہ کیجیے۔

سال	2014	2015	2016	2017	2018	2019	2020
پیداوار (ٹنوں میں)	64	60	68	62	66	70	72

اکائی 10۔ نمونے بندی کا نظریہ اور ٹیسٹ آف سگنیفیکنس

(Sampling Theory and Test of Significance)

Structure		اکائی کے اجزاء
Introduction	10.0	تمہید
Objectives	10.1	مقاصد
Sampling Theory and Tests of Significance	10.2	نمونہ بندی کا نظریہ اور ٹیسٹس آف سگنیفیکنس
The Basics of Sampling	10.3	نمونہ بندی کا بنیادی جایزہ
Sampling Procedures	10.4	نمونہ بندی کے طریقہ کار
Sample Size Determination	10.5	نمونے کے حجم کا تعین
Sampling Distribution	10.6	نمونہ جاتی تقسیم
Central Limit Theorem	10.7	مرکزی حد کا نظریہ
Hypothesis Testing	10.8	مفروضے کی جانچ :
Large Sample Test and Small Sample Test	10.9	لارج سیمپل ٹیسٹ اور اسمال سیمپل ٹیسٹ
Keywords	10.10	کلیدی الفاظ
Model Exam Questions	10.11	نمونہ امتحانی سوالات
Suggested Learning Resources	10.12	مزید مطالعے کے لیے تجویز کردہ کتابیں

10.0 تمہید(Introduction)

پچھلے یونٹ میں ہم نے ٹائم سیریز تجزیہ (Time Series Analysis) پر بات کی تھی جس میں بااصول و قفوں پر جمع کیے گئے ڈیٹا کی مدد سے رجحانات(trends) موسمی تبدیلیوں(seasonal fluctuations) چکراتی تغیرات (cyclic variations)اور بےاصول اتار چڑھاؤ (irregular movements) کا جائزہ لیا گیا تھا۔اس ضمن میں فری ہینڈ کرو-free) hand curve)سیمی ایوریجز(semi averages) موونگ ایوریجز (moving averages) اور لیسٹ اسکوائر میتھڈ (least squares method)جیسے مختلف طریقوں پر تفصیلی گفتگو کی گئی تھی تاکہ مستقبل کے رجحانات کی بہتر پیش گوئی کی جاسکے۔

اسی تجزیاتی سفر کو جاری رکھتے ہوئے یونٹ 10 میں ہم نمونے بندی کا نظریہ (Sampling Theory) اور ٹیسٹ آف سگنیفیکنس (Tests of Significance) کا تعارف کروا رہے ہیں۔مکمل آبادی (population) کا تجزیہ وقت لاگت اور دستیابی کی وجوہات کی بنا پر اکثر مشکل یا ناممکن ہوتا ہے ایسے میں نمونہ بندی (sampling) ایک اہم حل فراہم کرتی ہے۔اس یونٹ کے آغاز میں ہم نمونے بندی کے بنیادی تصورات اور طریقہ کار پر روشنی ڈالیں گے جس میں رینڈم (Random) اور نان رینڈم (Non-Random)طریقے اور نمونے کے مناسب حجم (sample size) کے تعین کے اہم پہلوؤں کو واضح کیا جائے گا۔اس کے بعد ہم نمونے کے تقسیمی طریقہ کار (sampling distributions) اور مرکزی حد نظریہ (Central Limit Theorem) پر تفصیل سے گفتگو کریں گے جو نمونوں کی اوسط (sample mean) کی نارمل تقسیم (normal approximation) کی وضاحت کرنے والا اہم اصول ہے چاہے اصل تقسیم (original distribution) کوئی بھی ہو۔

مزید اس تفہیم کی بنیاد پر یونٹ کے آخر میں ہم مختصراً لیکن نہایت اہم موضوع مفروضے کی جانچ (Hypothesis Testing)کا تعارف پیش کریں گے۔یہاں بڑے نمونہ (large sample) اور چھوٹے نمونہ (small sample) کے لئے ٹیسٹ آف سگنیفیکنس (Tests of Significance) کو بیان کیا جائے گا تاکہ نمونے کے ڈیٹا کی مدد سے آبادی کے بارے میں دعووں یا مفروضات کا جائزہ لیا جاسکے۔یہ معلومات کاروباری تجزیہ کاری (business analytics) میں فیصلہ سازی کے عمل کو مزید مستحکم اور مؤثر بنانے کے لئے نہایت مفید ثابت ہوں گی۔

10.1 مقاصد (Objectives)

اس اکائی کا مطالعہ کرنے کے بعد آپ مندرجہ ذیل نکات کو سمجھ سکیں گے:

- نمونہ بندی (Sampling) کے بنیادی تصورات اور طریقہ کار کی وضاحت کر سکیں۔
- نمونے کے حجم (Sample Size) کا درست تعین کرنے کے اصول اور طریقے سمجھ سکیں۔
- مرکزی حد کے نظریے (Central Limit Theorem) کی اہمیت اور اس کے عملی استعمال کو واضح کر سکیں۔

- مفروضے کی جانچ (Hypothesis Testing) کے بنیادی طریقے اور ٹیسٹس آف سگنیفیکنس (Tests of Significance) کی اہمیت بیان کر سکیں۔

10.2 نمونہ بندی کا نظریہ اور ٹیسٹس آف سگنیفیکنس

(Sampling Theory and Tests of Significance)

نمونہ بندی کا نظریہ (Sampling Theory) شماریات کا ایک بنیادی اور انتہائی اہم شعبہ ہے جو کسی آبادی (population) کے منتخب کردہ نمائندہ حصے یا نمونہ کی مدد سے پوری آبادی کی خصوصیات کے بارے میں نتائج اخذ کرنے میں ہماری مدد کرتا ہے۔ مکمل آبادی کے تمام عناصر (elements) کا مطالعہ کرنا اکثر وقت پیسہ اور دیگر وسائل کے لحاظ سے عملی طور پر ممکن نہیں ہوتا۔ ایسی صورتحال میں نمونہ بندی کے ذریعے آبادی کی نمائندگی کرنے والے منتخب عناصر کی مدد سے شماریاتی نتائج اخذ کیے جاتے ہیں۔

نمونہ بندی کا بنیادی مقصد یہ ہے کہ نمونے کی خصوصیات آبادی کی خصوصیات سے قریب تر ہوں تاکہ حاصل ہونے والے نتائج قابل اعتماد ہوں۔ اس ضمن میں اہم پہلو نمونے کی نوعیت اور حجم (sample size) کا درست تعین ہے۔ غلط یا غیر مناسب نمونہ منتخب کرنے سے نتائج غلط یا غیر معتبر ہو سکتے ہیں۔

دوسری جانب ٹیسٹس آف سگنیفیکنس (Tests of Significance) شماریاتی تجزیے کا وہ اہم جزو ہیں جن کے ذریعے نمونے کے ڈیٹا سے آبادی کے متعلق کیے گئے مفروضات (hypotheses) کی جانچ کی جاتی ہے۔ ان ٹیسٹس کے ذریعے معلوم کیا جاتا ہے کہ آبادی کے بارے میں بیان کردہ مفروضہ یا دعویٰ نمونے کے نتائج کی روشنی میں قابل قبول ہے یا اسے مسترد کیا جانا چاہیے۔

ٹیسٹس آف سگنیفیکنس کو عمومی طور پر دو اقسام میں تقسیم کیا جاتا ہے: بڑے نمونہ کے ٹیسٹس (Large Sample Tests) اور چھوٹے نمونہ کے ٹیسٹس (Small Sample Tests) ۔ یہ دونوں اقسام ڈیٹا کی نوعیت اور نمونے کے حجم کے مطابق استعمال کی جاتی ہیں۔

مختصراً نمونہ بندی کا نظریہ اور ٹیسٹس آف سگنیفیکنس ہمیں محدود ڈیٹا سے درست نتائج حاصل کرنے کے لیے بنیادی اصول اور طریقے فراہم کرتے ہیں جن سے کاروباری فیصلوں اور شماریاتی تجزیوں کو بہتر اور مؤثر بنایا جا سکتا ہے۔

10.3 نمونہ بندی کا بنیادی جائزہ (The Basics of Sampling)

نمونہ بندی (Sampling) شماریاتی تجزیہ کا ایک ایسا طریقہ ہے جس میں کسی بڑی آبادی (population) میں سے کچھ مخصوص عناصر (elements) کو منتخب کیا جاتا ہے تاکہ آبادی کی خصوصیات کا اندازہ لگایا جا سکے۔ دوسرے الفاظ میں نمونہ بندی کا عمل پوری آبادی کا مکمل مطالعہ کرنے کے بجائے اس کے کچھ منتخب نمائندہ عناصر پر مشتمل چھوٹے گروپ (sample) کا مطالعہ کرنے پر مشتمل ہوتا ہے۔

نمونہ بندی کی تعریف یوں کی جا سکتی ہے:

"نمونہ بندی ایک ایسا شماریاتی عمل ہے جس کے ذریعے بڑی آبادی میں سے چند عناصر منتخب کر کے ان عناصر کی خصوصیات اور نتائج کو بنیاد بنا کر پوری آبادی کے بارے میں اندازے اور تجزیے کیے جاتے ہیں۔"

نمونہ بندی کی ضرورت اس لیے پیش آتی ہے کیونکہ مکمل آبادی کا براہِ راست مطالعہ کرنا اکثر اوقات عملی طور پر مشکل یا ناممکن ہوتا ہے۔ یہ طریقہ کار اخراجات وقت وسائل اور محنت کو نمایاں طور پر کم کرنے میں مدد فراہم کرتا ہے۔ تاہم اس طریقے کی کامیابی کا انحصار منتخب کردہ نمونے کی درستگی اور نمائندگی (representativeness) پر ہوتا ہے۔ ایک اچھا نمونہ آبادی کی خصوصیات کا درست اور حقیقی عکاس ہوتا ہے۔

مختصراً نمونہ بندی آبادی کے وسیع تجزیے کی بنیادی رکاوٹوں کو عبور کرنے میں مدد دیتی ہے اور قابلِ اعتماد نتائج فراہم کرنے میں کلیدی کردار ادا کرتی ہے بشرطیکہ نمونے کا انتخاب درست اور نمائندہ ہو۔

10.4 نمونہ بندی کے طریقہ کار (Sampling Procedures)

نمونہ بندی کے طریقہ کار (Sampling Procedures) شماریاتی تجزیے کا اہم جزو ہیں جن کے ذریعے آبادی (population) میں سے نمائندہ نمونے (representative samples) منتخب کیے جاتے ہیں۔ نمونہ بندی کے طریقہ کار کو بنیادی طور پر دو اہم اقسام میں تقسیم کیا جاسکتا ہے:

رینڈم طریقہ کار (Random Sampling Methods)

نان رینڈم طریقہ کار (Non-Random Sampling Methods)

1. رینڈم طریقہ کار :(Random Sampling Methods) رینڈم طریقہ کار سے مراد وہ طریقے ہیں جن میں آبادی کے ہر عنصر (element) کو منتخب ہونے کا مساوی موقع (equal chance) حاصل ہوتا ہے۔ اس سے نمونے کی نمائندگی میں انصاف اور غیر جانبداری برقرار رہتی ہے۔ رینڈم طریقہ کار کی ذیلی اقسام درج ذیل ہیں:

سادہ رینڈم نمونہ بندی :(Simple Random Sampling) اس میں آبادی کے ہر عنصر کے منتخب ہونے کے مساوی مواقع ہوتے ہیں۔ یہ طریقہ نارمل طور پر قرعہ اندازی (lottery) کے ذریعے انجام دیا جاتا ہے۔

منظم نمونہ بندی :(Systematic Sampling) اس میں آبادی کو ایک ترتیب (systematic order) میں رکھا جاتا ہے اور ہر مخصوص وقفے پر عناصر منتخب کیے جاتے ہیں۔

طبقاتی نمونہ بندی :(Stratified Sampling) اس طریقے میں آبادی کو مختلف ذیلی گروہوں یا طبقات (strata) میں تقسیم کر کے ہر طبقے سے رینڈم نمونے منتخب کیے جاتے ہیں۔

کلسٹر نمونہ بندی :(Cluster Sampling) اس میں پوری آبادی کو مختلف گروہوں یا کلسٹرز (clusters) میں تقسیم کر کے چند کلسٹرز کا انتخاب کیا جاتا ہے اور پھر ان منتخب کلسٹرز میں موجود تمام عناصر یا چند عناصر کا مطالعہ کیا جاتا ہے۔

2. نان رینڈم طریقہ کار :(Non-Random Sampling Methods) یہ طریقے ایسے ہیں جن میں آبادی کے ہر عنصر کو منتخب ہونے کا مساوی موقع حاصل نہیں ہوتا۔ان میں جانبداری (bias) کا احتمال موجود ہوتا ہے۔اس کی ذیلی اقسام مندرجہ ذیل ہیں:

سہولتی نمونہ بندی :(Convenience Sampling) اس میں آسانی اور سہولت کی بنیاد پر عناصر کو منتخب کیا جاتا ہے۔

مقصدی نمونہ بندی :(Purposive Sampling) اس میں تحقیق کے مقصد کے مطابق عناصر کو جان بوجھ کر منتخب کیا جاتا ہے۔

کوٹہ نمونہ بندی :(Quota Sampling) اس میں آبادی کے مخصوص کوٹے مقرر کر کے نمونے منتخب کیے جاتے ہیں۔

سنوبال نمونہ بندی :(Snowball Sampling) اس طریقے میں ابتدائی طور پر چند عناصر منتخب کیے جاتے ہیں جو مزید عناصر کی شناخت اور انتخاب میں مدد کرتے ہیں۔

مختصراً نمونہ بندی کے طریقہ کار (Sampling Procedures) کا درست انتخاب تحقیق کے مقصد اور حالات پر منحصر ہوتا ہے۔ درست طریقہ کار کا انتخاب نتائج کی درستگی اور قابل اعتماد ہونے کو یقینی بناتا ہے۔

10.5 نمونے کے حجم کا تعین (Sample Size Determination)

نمونے کے حجم (Sample Size) سے مراد وہ منتخب شدہ عناصر (Elements) کی تعداد ہے جنہیں آبادی (Population) کی نمائندگی کے لیے نمونے میں شامل کیا جاتا ہے۔ تحقیق کے نتائج کی درستگی اور قابل اعتماد ہونے کا براہِ راست تعلق نمونے کے حجم کے درست تعین کے ساتھ ہے۔اگر نمونے کا حجم ضرورت سے زیادہ کم ہو تو نتائج غیر معتبر (Unreliable) ہو سکتے ہیں اور اگر حجم بہت زیادہ ہو تو اضافی وقت محنت اور مالی وسائل ضائع ہونے کا خطرہ ہوتا ہے۔

نمونے کے حجم کے تعین میں مندرجہ ذیل بنیادی عوامل اہمیت رکھتے ہیں:

نمونے کے حجم کو متاثر کرنے والے عوامل:

a) آبادی کا حجم:(Population Size)

جس آبادی سے نمونہ منتخب کرنا ہو اس کا کل حجم اہمیت رکھتا ہے۔ عموماً بڑی آبادیوں کے لیے نسبتاً بڑے نمونے درکار ہوتے ہیں۔

b) مطلوبہ درستگی کی سطح:(Desired Level of Precision)

نتائج کی درستگی یا ٹھیک ٹھیک (Accuracy) ہونا جتنی زیادہ ضروری ہو اتنا ہی نمونے کا حجم بڑا ہونا چاہیے۔مثال کے طور پر اگر ہم 95% یقین (Confidence) کے ساتھ نتیجہ حاصل کرنا چاہتے ہیں تو نمونے کا حجم اس حساب سے مقرر کرنا ہوگا۔

c) اعتماد کی سطح:(Confidence Level)

تحقیقی مطالعے میں عمومی طور پر اعتماد کی سطح 90% 95% یا 99% مقرر کی جاتی ہے۔اعتماد کی سطح جتنی زیادہ ہو اتنا ہی بڑا نمونے کا حجم درکار ہوگا۔

d) آبادی میں تغیر کی شرح:(Population Variability)

اگر آبادی میں تغیر یا تنوع (Variability) کی شرح زیادہ ہو تو نمائندہ نمونہ حاصل کرنے کے لیے نسبتاً بڑے حجم کا نمونہ منتخب کرنا پڑتا ہے۔

نمونے کے حجم کے تعین کے طریقے:

نمونے کے حجم کو تعین کرنے کے لیے مندرجہ ذیل اہم طریقے اختیار کیے جاتے ہیں:

a) شماریاتی فارمولے:(Statistical Formulas)

نمونے کے حجم کا تعین کرنے کے لیے درج ذیل بنیادی شماریاتی فارمولہ استعمال کیا جاتا ہے:

$$n = \frac{Z^2 \cdot p \cdot q}{e^2}$$

جہاں:

- n نمونے کا مطلوبہ حجم
- Z اعتماد کی سطح کے لحاظ سے Z ویلیو (مثلاً 95% اعتماد کے لیے 1.96)
- p آبادی میں متوقع تناسب یا اوسط (Expected proportion)
- q تناسب کا بقیہ حصہ (p-1)
- e قابل قبول غلطی کا حاشیہ (Margin of error)

مثال کے طور پر اگر ہم 95 % اعتماد کی سطح کے ساتھ 5% غلطی کے مارجن کے ساتھ نمونہ منتخب کرنا چاہتے ہیں اور متوقع تناسب 50% (0.5) فرض کر لیں تو نمونے کے حجم کا فارمولا کچھ اس طرح ہوگا:

$$n \approx \frac{(1.96)^2 \times 0.5 \times 0.5}{(0.05)^2} = 384.16$$

یعنی تقریباً 384 عناصر کا نمونہ اس صورتحال میں ضروری ہوگا۔

b) ماہرین کی رائے:(Expert Judgment)

کبھی کبھی تحقیقی حالات میں شماریاتی فارمولے کا استعمال ممکن نہ ہو تو ماہرین کی رائے کی بنیاد پر بھی نمونے کا حجم متعین کیا جاسکتا ہے۔ یہ طریقہ عموماً ابتدائی (exploratory) تحقیقات میں استعمال ہوتا ہے۔

c) سابقہ مطالعات:(Previous Studies)

بعض اوقات گزشتہ تحقیقی مطالعات کی بنیاد پر بھی نمونے کے حجم کا تعین کیا جاتا ہے۔ اس طریقے میں سابقہ مطالعات میں استعمال ہونے والے نمونوں کو بنیاد بنا کر مناسب حجم کا انتخاب کیا جاتا ہے۔

نمونے کے حجم کا درست تعین تحقیق کے معیار اور نتائج کی درستگی کو یقینی بنانے کے لیے انتہائی اہم ہے۔ نمونے کا حجم جتنا درست طریقے سے طے کیا جائے گا تحقیق کے نتائج اتنے ہی قابل بھروسہ اور مؤثر ہوں گے۔

10.6 نمونہ جاتی تقسیم(Sampling Distribution)

نمونہ جاتی تقسیم (Sampling Distribution) شماریات کا وہ اہم تصور ہے جو نمونے (Sample) کے ذریعے حاصل کیے گئے شماریاتی اشاریوں (جیسے اوسط تناسب وغیرہ) کی ممکنہ اقدار کی تقسیم کو ظاہر کرتا ہے۔ دوسرے لفظوں میں جب ایک ہی آبادی (Population) سے مختلف نمونے بار بار منتخب کیے جائیں تو ہر نمونے سے حاصل کردہ نتائج میں کچھ فرق ہوگا۔ انہی نتائج کی تقسیم کو نمونہ جاتی تقسیم (Sampling Distribution) کہتے ہیں۔

مثال کے طور پر اگر کسی یونیورسٹی میں 5000 طلبہ کی آبادی میں سے 50 طلبا کا ایک نمونہ منتخب کر کے ان کی اوسط عمر (Mean age) معلوم کی جائے اور یہی عمل بار بار دہرایا جائے تو ہر مرتبہ اوسط کی قدر تھوڑی بہت مختلف ہوگی۔ ان تمام ممکنہ اوسطوں کی تقسیم کو ہی "نمونہ جاتی تقسیم" یا Sampling Distribution کہتے ہیں۔

اہم نکات:

- نمونہ جاتی تقسیم کی شکل نارمل طور پر نارمل تقسیم (Normal Distribution) کے قریب ہوتی ہے۔
- اس تقسیم کی خصوصیات مثلاً اس کی اوسط تغیر (Variance) اور معیاری انحراف (Standard Deviation) کو آبادی کی خصوصیات کا اندازہ لگانے کے لیے استعمال کیا جاتا ہے۔
- اگر نمونے کا حجم کافی بڑا ہو تو نمونہ جاتی تقسیم ہمیشہ ایک نارمل ڈسٹری بیوشن (Normal distribution) کے قریب ہوتی ہے جیسا کہ مرکزی حد کا نظریہ (Central Limit Theorem) بتاتا ہے۔

نمونہ جاتی تقسیم (Sampling Distribution) تحقیقی نتائج کے شماریاتی تجزیے کے لیے بنیادی اہمیت رکھتی ہے۔ اس کی مدد سے محققین کو آبادی کے متعلق درست اندازے لگانے اور شماریاتی امتحانات (Tests of Significance) کے لیے بنیاد فراہم کرنے میں آسانی ہوتی ہے۔

10.7 مرکزی حد کا نظریہ :(Central Limit Theorem)

مرکزی حد کا نظریہ (Central Limit Theorem) شماریات کا ایک بنیادی اصول ہے جو بیان کرتا ہے کہ اگر کسی بھی آبادی (Population) سے مختلف نمونے (Samples) بار بار منتخب کیے جائیں اور ہر نمونے کے اوسط (Mean) کا حساب لگایا جائے تو یہ اوسطیں ایک خاص نمونہ جاتی تقسیم (Sampling Distribution) تشکیل دیتی ہیں۔ یہ تقسیم نارمل طور پر نارمل ڈسٹری بیوشن (Normal Distribution) کے قریب ہوتی ہے خاص طور پر جب نمونے کا حجم (Sample Size) بڑا ہو۔ مرکزی حد کے نظریے کے مطابق اگر نمونے کا حجم عموماً 30 یا اس سے زائد ہو تو چاہے آبادی کی اصل تقسیم کی نوعیت کوئی بھی ہو (چاہے نارمل ہو یا نہ ہو) نمونہ جاتی اوسطوں کی تقسیم ہمیشہ Normal تقسیم کے قریب ہوگی۔

مثال کے طور پر اگر بھارت کے مختلف شہروں سے اشیاء کی قیمتوں کا مطالعہ کرنے کے لیے متعدد نمونے منتخب کیے جائیں تو ہر نمونے کی اوسط قیمت میں کچھ نہ کچھ فرق آئے گا۔ لیکن اگر ایسے متعدد نمونوں کی اوسطیں جمع کر کے دیکھا جائے تو ان کی تقسیم ہمیشہ ایک نارمل ڈسٹری بیوشن (Normal Distribution) کے قریب ہو گی۔ اس اصول کی وجہ سے محققین نسبتاً چھوٹے نمونوں سے بھی پورے ملک کی آبادی کے متعلق قابلِ اعتماد نتائج اخذ کر سکتے ہیں۔

مرکزی حد کا نظریہ شماریاتی تجزیے میں بنیادی اہمیت رکھتا ہے کیونکہ یہ محققین کو یقین دلاتا ہے کہ بڑے نمونے سے حاصل کردہ نتائج نارمل ڈسٹری بیوشن کی صورت اختیار کرتے ہیں۔ یوں شماریاتی تجزیے میں درستگی اور آسانی پیدا ہوتی ہے۔

10.8 مفروضے کی جانچ: (Hypothesis Testing)

ٹیسٹس آف سگنیفیکنس (Tests of Significance) کا ایک اہم اور بنیادی جزو مفروضے کی جانچ (Hypothesis Testing) ہے۔ یہ شماریاتی طریقہ کار ہے جس کے ذریعے آبادی (Population) کے بارے میں بنائے گئے کسی دعوے یا مفروضے (Hypothesis) کو نمونے (Sample) کے ڈیٹا کے ذریعے پرکھا جاتا ہے۔

مفروضے کی جانچ کے بنیادی تصورات:

مفروضے کی جانچ میں دو اہم مفروضے قائم کیے جاتے ہیں:

a) بنیادی مفروضہ: (Null Hypothesis - H_0)

یہ وہ ابتدائی مفروضہ ہے جسے تحقیق کے آغاز میں درست مان کر جانچ کا آغاز کیا جاتا ہے۔ مثال کے طور پر "کالج کے طلباء کی اوسط عمر 22 سال ہے۔" یہ بنیادی مفروضہ کہلاتا ہے۔

b) متبادل مفروضہ: (Alternative Hypothesis - H_1)

یہ وہ مفروضہ ہے جو بنیادی مفروضے کی نفی یا مخالفت میں ہوتا ہے۔ مثال کے طور پر "کالج کے طلباء کی اوسط عمر 22 سال نہیں ہے۔" اس مفروضے کو نمونے کے شواہد سے ثابت کرنے کی کوشش کی جاتی ہے۔

ٹیسٹس آف سگنیفیکنس (Tests of Significance) کے ذریعے فیصلہ کیا جاتا ہے کہ بنیادی مفروضہ (H_0) قبول کیا جائے یا رد کر کے متبادل مفروضے (H_1) کو قبول کیا جائے۔ اس فیصلے کا انحصار نمونے سے حاصل شدہ شماریاتی شہادتوں (Statistical evidence) پر ہوتا ہے۔

10.9 لارج سیمپل ٹیسٹ اوراسمال سیمپل ٹیسٹ (Large Sample Test and Small Sample)

a) لارج سیمپل ٹیسٹ:(Large Sample Test)

لارج سیمپل ٹیسٹ وہ ٹیسٹس آف سگنیفیکنس (Tests of Significance) ہیں جنہیں نمونے کے حجم (Sample Size) کے بڑے ہونے کی صورت میں استعمال کیا جاتا ہے۔ نارمل طور پر اگر نمونے کا حجم 30 یا اس سے زائد ہو تو اسے "لارج سیمپل" تصور کیا جاتا ہے۔ بڑے نمونے کے ٹیسٹس کا استعمال عموماً Z- ٹیسٹ کی مدد سے کیا جاتا ہے۔ **مثال:** اگر ہم بھارت میں صارفین کی ایک بڑی تعداد (مثلاً 100 صارفین) کی ماہانہ اوسط آمدنی کو جانچ کر یہ جاننا چاہتے ہیں کہ آیا یہ اوسط آمدنی ₹25000 سے نمایاں طور پر مختلف ہے یا نہیں تو ہم لارج سیمپل ٹیسٹ کا استعمال کرتے ہوئے Z- ٹیسٹ کی مدد سے اس مفروضے کو جانچیں گے۔

b) اسمال سیمپل ٹیسٹ:(Small Sample Test)

اسمال سیمپل ٹیسٹ وہ ٹیسٹ ہیں جو عموماً تب استعمال کیے جاتے ہیں جب نمونے کا حجم 30 سے کم ہو۔ چھوٹے نمونوں میں نارمل طور پر T- ٹیسٹ (t-Test) کا استعمال کیا جاتا ہے۔ اس ٹیسٹ کے لیے یہ فرض کرنا ضروری ہے کہ آبادی کی تقسیم نارمل (Normal) یا اس کے قریب ہو۔ مثلاً اگر کسی کمپنی میں صرف 15 ملازمین موجود ہیں اور کمپنی یہ جانچنا چاہتی ہے کہ ان ملازمین کی اوسط ماہانہ تنخواہ ₹30,000 سے مختلف ہے یا نہیں تو اسمال سیمپل ٹیسٹ (T-Test) کا استعمال کر کے اس مفروضے کو پرکھا جائے گا۔

لارج سیمپل ٹیسٹ (Large Sample Test) اور اسمال سیمپل ٹیسٹ (Small Sample Test) دونوں اہم شماریاتی تکنیک ہیں جو مختلف حالات میں نمونے کے حجم اور ڈیٹا کی نوعیت کے مطابق استعمال کی جاتی ہیں۔ یہ ٹیسٹس کاروباری اور انتظامی فیصلوں میں مددگار ثابت ہوتے ہیں۔

10.10 اکتسابی نتائج

طلباء نے نمونے لینے کی بنیادی باتوں اور شماریاتی تجزیہ کے لیے نمائندہ نمونوں کے انتخاب کی اہمیت کے بارے میں ٹھوس سمجھ حاصل کر لی ہے۔ انہوں نے نمونے لینے کے مختلف طریقہ کار اور قابل اعتماد نتائج کو یقینی بنانے کے لیے نمونے کے مناسب سائز کا تعین کرنے کا طریقہ سیکھا ہے۔ طلباء نمونے لینے کی تقسیم کے تصور اور مرکزی حد نظریہ کی وضاحت کر سکتے ہیں، جو تخمینہ شماریات کی بنیاد بناتے ہیں۔ انہوں نے مفروضے کی جانچ کرنے، بڑے نمونے کے ٹیسٹ اور چھوٹے نمونے کے ٹیسٹ کے درمیان فرق کرنے، اور نمونے کے سائز اور ڈیٹا کی خصوصیات کی بنیاد پر درست طریقے استعمال کرنے کی صلاحیت تیار کی ہے۔ مجموعی طور پر، سیکھنے والے نمونے لینے کی تکنیکوں اور شماریاتی ٹیسٹوں کا استعمال کرنے سے لیس ہوتے ہیں تاکہ درست اندازہ لگایا جا سکے اور ڈیٹا پر مبنی کاروباری فیصلوں کی حمایت کی جا سکے۔

10.11 نمونہ امتحانی سوالات(Model Exam Questions)

معروضی جوابات کے حامل سوالات(Objective Answer Type Questions)

1. نمونہ بندی (Sampling) سے کیا مراد ہے؟
2. بے ترتیب نمونہ بندی (Random Sampling) کا ایک طریقہ لکھیں۔
3. مرکزی حد کے نظریے (Central Limit Theorem) کے مطابق نمونے کا کم سے کم حجم کتنا ہونا چاہیے؟
4. نمونے کے حجم (Sample Size) کا تعین کیوں ضروری ہے؟
5. اسمال سیمپل ٹیسٹ (Small Sample Test) کب استعمال کیا جاتا ہے؟
6. بنیادی مفروضہ (Null Hypothesis) کو کس علامت سے ظاہر کرتے ہیں؟
7. نمونہ جاتی تقسیم (Sampling Distribution) کی اہمیت کا کوئی ایک نکتہ بیان کریں۔
8. مرکزی حد کے نظریے (Central Limit Theorem) کا بنیادی مقصد کیا ہے؟
9. دانستہ نمونہ بندی (Judgmental Sampling) کیا ہے؟
10. ٹیسٹس آف سگنیفیکنس (Tests of Significance) کس لیے استعمال کیے جاتے ہیں؟

مختصر جوابات کے حامل سوالات (Short Answers Type Questions)

1. نمونہ بندی (Sampling) کی اہمیت کو مختصراً وضاحت کریں۔
2. بے ترتیب (Random) اور غیر بے ترتیب (Non-Random) نمونہ بندی میں فرق واضح کریں۔
3. نمونے کے حجم (Sample Size) کو متاثر کرنے والے کسی تین عوامل کی وضاحت کریں۔
4. نمونہ جاتی تقسیم (Sampling Distribution) کو مختصراً بیان کریں۔
5. مرکزی حد کے نظریے (Central Limit Theorem) کا عملی استعمال مختصر مثال کے ذریعے واضح کریں۔

طویل جوابات کے حامل سوالات (Long Answers Type Questions)

1. نمونہ بندی کے طریقہ کار (Sampling Procedures) تفصیل کے ساتھ بیان کریں اور ہر طریقے کی ایک مناسب مثال دیں۔
2. مفروضے کی جانچ (Hypothesis Testing) کیا ہے؟ بنیادی اور متبادل مفروضے کو مثالوں کے ذریعے واضح کریں۔
3. نمونے کے حجم (Sample Size Determination) کا تعین شماریاتی فارمولے (Statistical Formula) کے ذریعے وضاحت کریں ایک فرضی مثال بھی دیں۔

4. لارج سیمپل ٹیسٹ (Large Sample Test) اوراسمال سیمپل ٹیسٹ (Small Sample Test) میں فرق کومثالوں کے ساتھ واضح کریں۔

5. ٹیسٹس آف سگنیفیکنس (Tests of Significance) کیاہیں؟کاروباری فیصلوں میں ان کی اہمیت کووضاحت کریں۔

اکائی 11۔کائی اسکوائر کا تجزیہ

(Chi Square Analysis)

Structure	اکائی کے اجزاء
Introduction	11.0 تمہید
Objectives	11.1 مقاصد
Meaning and Definition of Chi-Square Analysis	11.2 کائی اسکوائر تجزیہ کے معنی اور تصور
Importance of Chi Square Analysis in Business Analytics	11.3 کاروباری تجزیات میں کائی اسکوائر تجزیہ کی اہمیت
Test for Goodness of Fit	11.4 ٹیسٹ فار گڈ نیس آف فٹ
Test for Independence of Attributes	11.5 ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس
Limitations or Conditions of Chi-Square Test	11.6 کائی اسکوائر ٹیسٹ کی حدود یا شرائط
Learning Outcomes	11.7 اکتسابی نتائج
Model Exam Questions	11.8 نمونہ امتحانی سوالات

11.0 تمہید (Introduction)

گزشتہ یونٹ(یونٹ10)میں آپ نے نمونہ بندی کے نظریے (Sampling Theory) اور ٹیسٹس آف سگنیفیکنس (Tests of Significance) کے بنیادی تصورات کو تفصیل کے ساتھ سمجھا۔آپ نے دیکھا کہ آبادی (Population) کے متعلق نمونے کے ذریعے حاصل کردہ معلومات کی مدد سے مفروضات (Hypotheses) کو کس طرح جانچا جاتا ہے۔اس سلسلے میں آپ نے بڑے اور چھوٹے نمونوں سے متعلق شماریاتی ٹیسٹس Z- ٹیسٹ اور t- ٹیسٹ) بھی پڑھے۔

اب اسی موضوع کو آگے بڑھاتے ہوئے موجودہ یونٹ (یونٹ 11) میں آپ ایک اہم غیر پیرامیٹرک ٹیسٹ یعنی **کائی اسکوائر تجزیہ (Chi-Square Analysis)** کے بارے میں جانیں گے۔ کائی اسکوائر ٹیسٹ ایک ایسا شماریاتی طریقہ کار ہے جس کے ذریعے مختلف کیٹیگریز (Categories) یا خصوصیات (Attributes) کے مابین تعلق یا وابستگی کا تجزیہ کیا جاتا ہے۔ اس یونٹ میں آپ خاص طور پر دو اہم ٹیسٹس، ٹیسٹ فار گڈنیس آف فٹ (Test for Goodness of Fit) اور ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس (Test for Independence of Attributes) کو تفصیل سے پڑھیں گے۔

اس یونٹ کا مقصد آپ کو عملی مثالوں کے ذریعے کائی اسکوائر ٹیسٹ کا مؤثر استعمال سکھانا ہے تاکہ آپ کاروباری مسائل اور تحقیقی تجزیات میں بہتر نتائج حاصل کر سکیں۔

11.1 مقاصد (Objectives)

اس اکائی کا مطالعہ کرنے کے بعد، آپ مندرجہ ذیل نکات کو سمجھ سکیں گے:

- کائی اسکوائر تجزیہ (Chi-Square Analysis) کے بنیادی تصور اور استعمال کو واضح کر سکیں۔
- ٹیسٹ فار گڈنیس آف فٹ (Test for Goodness of Fit) کا عملی استعمال سمجھ کر اس کی وضاحت کر سکیں۔
- ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس (Test for Independence of Attributes) کے ذریعے خصوصیات کے درمیان تعلق کو جانچ سکیں۔
- کائی اسکوائر ٹیسٹ کے عملی اطلاق میں موجود حدود یا شرائط (Limitations or Conditions) کو واضح کر سکیں۔

11.2 کائی اسکوائر تجزیہ کے معنی اور تعریف

(Meaning and Definition of Chi-Square Analysis)

کائی اسکوائر تجزیہ (Chi-Square Analysis) شماریات میں استعمال ہونے والا ایک اہم غیر پیرامیٹرک (Non-parametric) ٹیسٹ ہے، جس کے ذریعے دو یا زیادہ خصوصیات (Attributes) یا زمرہ جات (Categories) کے درمیان موجود تعلق یا فرق کو جانچا جاتا ہے۔ دوسرے لفظوں میں، یہ ٹیسٹ کسی آبادی (Population) میں خصوصیات کے حقیقی مشاہدہ شدہ تعدد (Observed Frequency) اور متوقع تعدد (Expected Frequency) کے مابین فرق کو واضح کرتا ہے۔

کائی اسکوائر تجزیے کا تصور:(Concept)

کائی اسکوائر تجزیہ بنیادی طور پر مشاہداتی ڈیٹا (Observed Data) اور متوقع ڈیٹا (Expected Data) کے درمیان فرق (Difference) کو ظاہر کرتا ہے۔اس کا بنیادی فارمولا یہ ہے:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

یہاں:

O = مشاہدہ کردہ قدر(Observed Value)

E = متوقع قدر(Expected Value)

کائی اسکوائر تجزیے کا نتیجہ جتنا بڑا ہو گا، اتنا ہی واضح طور پر دونوں ڈیٹا(مشاہدہ شدہ اور متوقع) کے درمیان فرق نمایاں ہو گا۔

مثال کے طور پر، فرض کریں کہ بھارت میں ایک موبائل کمپنی کا دعویٰ ہے کہ اس کی مصنوعات ملک کے چار مختلف علاقوں (شمال، جنوب، مشرق اور مغرب) میں یکساں مقبول ہیں۔ اب اس دعوے کو جانچنے کے لیے ہر علاقے سے مصنوعات کی فروخت کا ڈیٹا حاصل کیا جائے گا۔اس کے بعد کائی اسکوائر ٹیسٹ کی مدد سے دیکھا جائے گا کہ کیا حقیقی فروخت (Observed) اور کمپنی کی توقع کردہ فروخت (Expected)کے درمیان فرق نمایاں (Significant) ہے یا نہیں۔

اگر حساب کردہ χ^2 کی قدر مخصوص معیار (Critical value) سے زیادہ ہو تو کمپنی کا دعویٰ رد کر دیا جائے گا، بصورتِ دیگر قبول کر لیا جائے گا۔

کائی اسکوائر تجزیہ (Chi-Square Analysis) آبادی کے زمرہ جاتی ڈیٹا کے تجزیے کے لیے ایک اہم اور مؤثر شماریاتی ٹیسٹ ہے۔ اس کی مدد سے کاروباری ادارے اور محققین مختلف خصوصیات یا زمرہ جات کے درمیان تعلق اور اختلاف کو جانچ کر بہتر اور قابل اعتماد فیصلے کر سکتے ہیں۔

11.3 کاروباری تجزیات میں کائی اسکوائر تجزیہ کی اہمیت

(Importance of Chi Square Analysis in Business Analytics)

کائی اسکوائر تجزیہ (Chi-Square Analysis) کاروباری تجزیات (Business Analytics) کا ایک اہم شماریاتی طریقہ ہے جو کاروباری فیصلوں کے لیے کلیدی کردار ادا کرتا ہے۔اس ٹیسٹ کے ذریعے کاروباری ادارے مختلف خصوصیات یا عوامل (Attributes or Factors)کے مابین تعلقات کا تعین کر کے بہتر فیصلہ سازی کر سکتے ہیں۔

کاروباری تجزیات میں کائی اسکوائر تجزیہ کی اہمیت درج ذیل پہلوؤں سے واضح ہوتی ہے:

1) گاہک کے رویے کا تجزیہ:(Customer Behavior Analysis)

کاروباری ادارے اکثر جاننا چاہتے ہیں کہ صارفین کی مخصوص خصوصیات (مثلاً صنف، عمر، علاقہ) کا خریداری کے رویے کے ساتھ کیا تعلق ہے۔ کائی اسکوائر ٹیسٹ کے ذریعے ان تعلقات کو جانچا جا سکتا ہے۔ **مثال** :اگر بھارت میں کوئی ای-کامرس کمپنی یہ جاننا چاہتی ہے کہ صارفین کی عمر (Age) اور مصنوعات کے انتخاب (Product Choice) کے درمیان کوئی تعلق ہے یا نہیں تو کائی اسکوائر ٹیسٹ سے یہ تعلق واضح طور پر معلوم کیا جا سکتا ہے۔

2) صارفین کے رویے کی تحقیق:(Consumer Behavior Analysis)

کائی اسکوائر تجزیہ کاروبار میں صارفین کے رویے اور رجحانات کی شناخت کے لیے وسیع پیمانے پر استعمال ہوتا ہے۔ اس کے ذریعے صارفین کی ترجیحات، خریداری کے نمونے(Buying Patterns) ، اور مصنوعات کی مقبولیت جانچی جاتی ہے۔ **مثال:** ایک ریٹیل اسٹور (Retail Store) یہ تجزیہ کر سکتا ہے کہ کیا گاہک کی آمدنی اور مخصوص برانڈ کے انتخاب کے درمیان تعلق موجود ہے یا نہیں۔

3) مارکیٹ سیگمنٹیشن:(Market Segmentation)

کائی اسکوائر تجزیہ مارکیٹ کو مختلف سیگمنٹس یا حصوں میں تقسیم کرنے کے لیے مؤثر ٹول ہے۔ اس کے ذریعے یہ معلوم کیا جا سکتا ہے کہ مختلف صارفین کے گروہوں (مثلاً مرد، خواتین، طلبا، پیشہ ور افراد) کی ترجیحات کس طرح مختلف ہیں۔ **مثال:** ایک موبائل کمپنی بھارت میں مختلف عمر یا پیشوں سے تعلق رکھنے والے صارفین کی ترجیحات کو جانچ کر مارکیٹ کو مؤثر انداز سے تقسیم کر سکتی ہے۔

4) مصنوعات اور خدمات کی منصوبہ بندی:(Product Planning)

کائی اسکوائر تجزیہ اداروں کو مصنوعات کی طلب یا مقبولیت کے بارے میں فیصلہ سازی میں مدد فراہم کرتا ہے۔ ادارے مصنوعات کی فروخت کے رجحانات کا تجزیہ کر کے مستقبل کے کاروباری فیصلے لے سکتے ہیں۔ **مثال:** ایک بھارتی FMCG کمپنی کائی اسکوائر تجزیہ کے ذریعے یہ فیصلہ کر سکتی ہے کہ مختلف شہروں میں کون سی مصنوعات زیادہ مقبول ہیں، اور اسی کے مطابق پروڈکشن یا سپلائی میں تبدیلیاں کر سکتی ہے۔

کائی اسکوائر تجزیہ (Chi-Square Analysis) کاروباری تجزیات (Business Analytics) میں فیصلہ سازی کو مستحکم بنانے کے لیے اہم ہے۔ اس ٹیسٹ کی مدد سے کاروبار میں شماریاتی شواہد کی بنیاد پر مؤثر فیصلے کیے جاتے ہیں، جس سے نہ صرف درست منصوبہ بندی ممکن ہوتی ہے بلکہ کاروباری کارکردگی بھی بہتر ہوتی ہے۔

11.4 ٹیسٹ فار گڈ نیس آف فٹ(Test for Goodness of Fit)

ٹیسٹ فار گڈ نیس آف فٹ (Test for Goodness of Fit) شماریاتی تجزیے کا ایک اہم طریقہ ہے جس کی مدد سے معلوم کیا جاتا ہے کہ کیا نمونے (Sample) سے حاصل شدہ ڈیٹا کسی مخصوص شماریاتی تقسیم(Statistical Distribution) ، مثلاً نارمل تقسیم(Normal Distribution) ، بائنومیئل تقسیم (Binomial Distribution) یا دیگر نظریاتی تقسیموں سے مطابقت رکھتا ہے یا نہیں۔ دوسرے الفاظ میں، یہ ٹیسٹ جانچتا ہے کہ نمونے سے حاصل کردہ مشاہدہ شدہ اقدار (Observed Values) نظریاتی طور پر متوقع اقدار (Expected Values) کے ساتھ کس حد تک مطابقت رکھتی ہیں۔

یہ ٹیسٹ کائی اسکوائر تجزیے (Chi-Square Analysis) کے ذریعے انجام دیا جاتا ہے، جس کے تحت حقیقی طور پر مشاہدہ کردہ ڈیٹا (Observed Data) کو متوقع ڈیٹا (Expected Data) سے موازنہ کرکے ان کے فرق کی اہمیت کو جانچا جاتا ہے۔

کائی اسکوائر کی قدروں کا حساب لگا کر اسے ٹیبل ویلیو سے موازنہ کیا جاتا ہے۔ اگر حساب شدہ کائی اسکوائر کی قدر ٹیبل میں دی گئی اہم قدر (Critical value) سے کم یا برابر ہو تو نمونہ متوقع تقسیم (Expected distribution) سے مطابقت رکھتا ہے، بصورت دیگر اسے رد کر دیا جاتا ہے۔

مثلاً، فرض کریں کہ ایک بھارتی کنفیکشنری کمپنی دعویٰ کرتی ہے کہ اس کے تیار کردہ ٹافیوں کے پانچ مختلف ذائقے صارفین میں یکساں مقبول ہیں۔ اس دعوے کو پرکھنے کے لیے مارکیٹ سے 500 ٹافیوں کا نمونہ لیا گیا۔ اب کائی اسکوائر ٹیسٹ فار گڈ نیس آف فٹ کے ذریعے دیکھا جائے گا کہ کمپنی کا دعویٰ صارفین کی اصل ترجیحات کے مطابق ہے یا نہیں۔ اگر مشاہدہ شدہ فروخت متوقع فروخت سے نمایاں طور پر مختلف ہو تو کمپنی کا دعویٰ مسترد کر دیا جائے گا۔ ٹیسٹ فار گڈ نیس آف فٹ (Test for Goodness of Fit) کاروباری تجزیات (Business Analytics) میں وسیع پیمانے پر استعمال ہونے والی تکنیک ہے، جس کے ذریعے ڈیٹا کے عملی اور نظریاتی ماڈل کے درمیان مطابقت جانچی جاتی ہے۔ اس ٹیسٹ کے نتائج کاروباری فیصلہ سازی، منصوبہ بندی اور کارکردگی کی بہتری کے لیے اہم ہوتے ہیں۔

سوال:

ایک سکہ 100 بار اُچھالا گیا۔ اس میں 58 بار ہیڈ (Heads) اور 42 بار ٹیل (Tails) آیا۔ کیا ہم کہہ سکتے ہیں کہ یہ سکہ منصفانہ (Fair) ہے یا نہیں؟ اس بات کو جانچنے کے لیے "ٹیسٹ فار گڈ نیس آف فٹ (Test for Goodness of Fit)" کو 5% سطحِ اہمیت (level of significance) پر استعمال کریں۔

جواب:

ہم یہ جانچنے کے لیے کہ سکہ منصفانہ ہے یا نہیں، کائی اسکوائر آزمائش (Chi-square Test for Goodness of Fit) استعمال کریں گے۔ اگر سکہ منصفانہ ہو تو ہم توقع کرتے ہیں کہ 100 بار اُچھالنے پر 50 بار ہیڈ اور 50 بار ٹیل آئے گا۔

ہم درج ذیل فارمولا استعمال کرتے ہیں:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

For **Heads**: Observed (O) = 58, Expected (E) = 50

$$\frac{(58 - 50)^2}{50} = \frac{64}{50} = 1.28$$

For **Tails**: Observed (O) = 42, Expected (E) = 50

$$\frac{(42 - 50)^2}{50} = \frac{64}{50} = 1.28$$

Now, we add both values:

$$\chi^2 = 1.28 + 1.28 = 2.56$$

اب ہم اس کو تنقیدی قیمت (critical value) سے موازنہ کریں گے۔ 5% سطحِ اہمیت اور 1 درجہ آزادی (degree of freedom) کے لیے جدول سے تنقیدی قیمت=3.841 حاصل ہوتی ہے۔

چونکہ $2.56 < 3.841$

لہٰذا ہم صفر مفروضہ (Null Hypothesis) کو رد نہیں کرتے۔ اس کا مطلب ہے کہ ہمارے پاس یہ کہنے کے لیے کافی ثبوت نہیں کہ سکہ غیر منصفانہ ہے۔ لہٰذا ہم نتیجہ نکالتے ہیں کہ یہ سکہ منصفانہ ہے۔

11.5 ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس (Test for Independence of Attributes)

ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس (Test for Independence of Attributes) کائی اسکوائر تجزیے (Chi-Square Analysis) کا اہم اور کثرت سے استعمال ہونے والا شماریاتی ٹیسٹ ہے۔ اس ٹیسٹ کے ذریعے دو یا زائد خصوصیات یا اوصاف (Attributes) کے درمیان تعلق (Relationship) یا آزادی (Independence) کو جانچا جاتا ہے۔ دوسرے الفاظ میں، یہ معلوم کیا جاتا ہے کہ دو صفات آپس میں آزاد (Independent) ہیں یا ان کے درمیان کوئی واضح تعلق موجود ہے۔

اس ٹیسٹ کا بنیادی مقصد یہ جانچنا ہوتا ہے کہ دو مختلف خصوصیات کے درمیان موجود تعلق حقیقی اور شماریاتی لحاظ سے نمایاں (Significant) ہے یا محض اتفاقیہ (by chance) ہے۔

اگر کیلکولیٹ کردہ χ^2 ویلیو جدول میں دی گئی اہمیت کی سطح (Significance Level) کی کریٹیکل ویلیو سے زیادہ ہو تو ہم کہتے ہیں کہ خصوصیات کے درمیان تعلق موجود ہے اور یہ اتفاقی نہیں بلکہ شماریاتی طور پر نمایاں ہے۔ دوسری صورت میں ہم ان خصوصیات کو ایک دوسرے سے آزاد (Independent) تصور کرتے ہیں۔

مثلاً، فرض کریں کہ ایک بھارتی کمپنی یہ جاننا چاہتی ہے کہ اس کی مصنوعات کو پسند کرنے میں صارفین کی جنس (مرد/خواتین) اور ان کی عمر کا گروپ (نوجوان/بالغ/بزرگ) کے درمیان کوئی تعلق ہے یا نہیں۔ اس مقصد کے لیے کمپنی ایک سروے کرتی ہے اور صارفین کے جواب کی مدد سے ڈیٹا تیار کرتی ہے۔ پھر کمپنی اس ڈیٹا پر "ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس" لاگو کرتی ہے۔ اگر ٹیسٹ کے نتائج نمایاں (Significant) ہوں تو کمپنی یہ تسلیم کرتی ہے کہ جنس اور عمر صارفین کی پسند پر اثرانداز ہوتے ہیں۔

مثال

	بزرگ	بالغ	نوجوان
مرد	80	120	100
خواتین	70	130	150

اوپر دی گئی جدول (Table) کے ذریعے کائی اسکوائر ٹیسٹ کا اطلاق کیا جائے گا اور دیکھا جائے گا کہ آیا جنس اور عمر کے گروپ کے درمیان صارفین کی پسندیدگی کا تعلق ہے یا دونوں خصوصیات آزاد ہیں۔ اگر تعلق نمایاں (Significant) پایا گیا تو کمپنی اپنی مارکیٹنگ پالیسی کو اس مناسبت سے ترتیب دے گی۔

ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس (Test for Independence of Attributes) کاروباری اور تحقیقی میدانوں میں فیصلہ سازی کے لیے اہم ہے۔ اس ٹیسٹ سے دو خصوصیات یا عوامل کے تعلق کو درست طریقے سے سمجھ کر کاروباری حکمت عملیوں اور انتظامی فیصلوں میں بہتری لائی جا سکتی ہے۔

سوال:

ایک سروے میں 100 افراد سے پوچھا گیا کہ وہ چائے پسند کرتے ہیں یا کافی، اور ان کی جنس (male/female) بھی نوٹ کی گئی۔

مندرجہ ذیل جدول میں ان کے جوابات دیے گئے ہیں:

	Tea	**Coffee**	**Total**
Male	30	20	50
Female	10	40	50
Total	40	60	100

کیا جنس (Gender) **اور مشروب کی پسند** (Drink Preference) **ایک دوسرے سے آزاد** (Independent) **ہیں؟ 5% سطحِ اہمیت (level of significance) پر ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس** (Test for Independence of Attributes) **کریں۔**

جواب:

اس مسئلے کو حل کرنے کے لیے ہم کائی اسکوائر آزمائش (Chi-Square Test) کا استعمال کریں گے تاکہ یہ جان سکیں کہ جنس اور مشروب کی پسند کے درمیان کوئی تعلق ہے یا نہیں۔

Step 1: State the Hypotheses

H0: Gender and drink preference are **independent** (no association).

H1: Gender and drink preference are **not independent** (there is an association).

Step 2: Calculate Expected Frequencies

We use the formula:

$$E = \frac{(\text{Row Total}) \times (\text{Column Total})}{\text{Grand Total}}$$

	Tea (Expected)	Coffee (Expected)
Male	50×40/100=20	50×60/100=30
Female	50×40/100=20	50×60/100=30

Step 3: Apply the Chi-square Formula

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Category	**Observed (O)**	**Expected (E)**	**$(O-E)^2 / E$**
Male - Tea	30	20	5.00
Male - Coffee	20	30	3.33
Female - Tea	10	20	5.00
Female - Coffee	40	30	3.33
Total			**16.66**

$$\chi^2 = 5.00 + 3.33 + 5.00 + 3.33 = 16.66$$

Step 4: Find the Critical Value

Degrees of freedom: $(df) = (rows - 1) \times (columns - 1) = (2 - 1)(2 - 1) = 1$

At 5% level of significance and 1 degree of freedom, the **critical value** from the Chi-square table is **3.841**.

Step 5: Decision

$$\chi^2_{calculated} = 16.66 > \chi^2_{critical} = 3.841$$

Since

We **reject the null hypothesis**.

11.6 کائی اسکوائر ٹیسٹ کے تحدیدات و یا شرائط

(Limitations or Conditions of Chi-Square Test)

کائی اسکوائر ٹیسٹ (Chi-Square Test) شماریاتی تجزیہ کا ایک مقبول طریقہ کار ہے لیکن اس ٹیسٹ کے درست استعمال کے لیے کچھ شرائط (Conditions) اور حدود (Limitations) ہوتی ہیں، جنہیں ذہن میں رکھنا ضروری ہے۔ ان حدود کو نظرانداز کرنے سے ٹیسٹ کے نتائج غلط یا غیر معتبر ہو سکتے ہیں۔

کائی اسکوائر ٹیسٹ کی اہم شرائط و حدود درج ذیل ہیں:

ڈیٹا کا آزاد ہونا:(Independence of Observations)

کائی اسکوائر ٹیسٹ استعمال کرنے کے لیے ضروری ہے کہ نمونے کے مشاہدات ایک دوسرے سے آزاد ہوں۔ یعنی ایک مشاہدہ دوسرے مشاہدے پر اثرانداز نہ ہو۔

1) نمونے کا حجم:(Sample Size)

نمونے کا حجم کافی بڑا ہونا چاہیے۔ عام طور پر یہ شرط ہے کہ ہر خانے (Cell) میں متوقع تعدد (Expected frequency) کم از کم 5 ہونی چاہیے۔ اگر متوقع تعدد 5 سے کم ہو تو نتائج قابلِ اعتماد نہیں ہوتے۔

2) زمرہ جاتی ڈیٹا:(Categorical Data)

کائی اسکوائر ٹیسٹ صرف زمرہ جاتی (Categorical) یا معیاری (Qualitative) ڈیٹا پر لاگو ہوتا ہے۔ مقداری ڈیٹا (Quantitative Data) پر اس کا اطلاق درست نہیں ہوتا۔

3) فری کوئنسی ڈیٹا:(Frequency Data)

یہ ٹیسٹ صرف تعدد کی صورت میں حاصل کردہ ڈیٹا (Frequency Data) کے لیے موزوں ہے۔ فیصد یا تناسب کی صورت میں براہِ راست استعمال کرنا درست نہیں۔

4) آزاد مشاہدات:(Independent Observations)

مشاہدات آزاد (Independent) ہونے چاہئیں۔ایک مشاہدہ دوسرے کو متاثر نہ کرے، یعنی ایک جواب دوسرے جواب پر منحصر نہیں ہونا چاہیے۔

5) مخصوص شماریاتی تقسیم پر انحصار نہ ہونا:

کائی اسکوائر ٹیسٹ ایک غیر پیرامیٹرک ٹیسٹ ہے۔اس لیے اس کی اطلاق کے لیے آبادی کی کوئی خاص شماریاتی تقسیم فرض کرنا ضروری نہیں ہوتا۔

کائی اسکوائر ٹیسٹ کے درست استعمال کے لیے مذکورہ بالا حدود اور شرائط کا خیال رکھنا ضروری ہے۔ان شرائط کو مد نظر رکھتے ہوئے نتائج قابل اعتماد اور درست ثابت ہوتے ہیں۔

11.7 اکتسابی نتائج(Learning Outcome)

طلباء نے Chi-Square Analysis کے معنی اور تعریف کے بارے میں واضح فہم پیدا کر لیا ہے اور واضح ڈیٹا کا تجزیہ کرنے میں اس کی مطابقت ہے۔انہوں نے مفروضوں کی جانچ کرنے اور ڈیٹا پر مبنی فیصلے کرنے کے لیے کاروباری تجزیات میں Chi-Square Analysis کی اہمیت کو جان لیا ہے۔طلباء اب یہ تعین کرنے کے لیے فٹ کی خوبی کے لیے ٹیسٹ کر سکتے ہیں کہ مشاہدہ کیا گیا ڈیٹا متوقع تقسیم سے کس حد تک میل کھاتا ہے، اور صفات کی آزادی کے لیے ٹیسٹ متغیرات کے درمیان تعلقات کو جانچنے کے لیے۔مزید برآں، انہوں نے Chi-Square ٹیسٹ کی حدود اور شرائط کو سمجھ لیا ہے، بشمول وہ مفروضے جنہیں درست نتائج کے لیے پورا کرنا ضروری ہے۔ مجموعی طور پر، سیکھنے والے دوٹوک ڈیٹا میں پیٹرن کا اندازہ لگانے اور اسٹریٹجک کاروباری فیصلوں کی حمایت کرنے کے لیے Chi-Square تکنیکوں کو لاگو کرنے کے لیے لیس ہیں۔

11.8 نمونہ امتحانی سوالات(Model Exam Questions)

معروضی جوابات کے حامل سوالات (Objective Answer Type Questions)

1. کائی اسکوائر ٹیسٹ (Chi-Square Test) کا بنیادی مقصد کیا ہے؟
2. کائی اسکوائر ٹیسٹ کا فارمولا کیا ہے؟
3. کائی اسکوائر ٹیسٹ کا اطلاق کس قسم کے ڈیٹا پر کیا جاتا ہے؟
4. کائی اسکوائر ٹیسٹ میں "E" سے کیا مراد ہے؟
5. ٹیسٹ فار گڈنیس آف فٹ (Test for Goodness of Fit) کس لیے استعمال ہوتا ہے؟

6. ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس (Test for Independence of Attributes) کا بنیادی مقصد کیا ہے؟
7. کائی اسکوائر ٹیسٹ کے لیے ڈیٹا کا آزاد (Independent) ہونا کیوں ضروری ہے؟
8. کائی اسکوائر ٹیسٹ کی متوقع تعدد (Expected Frequency) کم از کم کتنی ہونی چاہیے؟
9. کیا کائی اسکوائر ٹیسٹ مقداری (Quantitative) ڈیٹا پر لاگو ہو سکتا ہے؟
10. کاروبار میں کائی اسکوائر ٹیسٹ کا کوئی ایک استعمال لکھیں۔

مختصر جوابات کے حامل سوالات (Short Answers Type Questions)

1. کائی اسکوائر تجزیہ (Chi-Square Analysis) کے معنی اور تصور کو مختصر وضاحت کے ساتھ بیان کریں۔
2. ٹیسٹ فار گڈنیس آف فٹ (Test for Goodness of Fit) کو مثال کے ساتھ واضح کریں۔
3. کائی اسکوائر ٹیسٹ میں متوقع (Expected) اور مشاہدہ شدہ (Observed) تعدد سے کیا مراد ہے؟ مثال دیں۔
4. کائی اسکوائر ٹیسٹ کی کوئی تین اہم حدود (Limitations) بیان کریں۔
5. کاروباری تجزیات میں کائی اسکوائر ٹیسٹ (Chi-Square Test) کی اہمیت مختصراً بیان کریں۔

طویل جوابات کے حامل سوالات (Long Answers Type Questions)

1) کائی اسکوائر تجزیہ (Chi-Square Analysis) کیا ہے؟ اس کا فارمولا، بنیادی تصورات اور کاروبار میں استعمال واضح کریں۔
2) ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس (Test for Independence of Attributes) کیا ہے؟ تفصیلی مثال دے کر اس کی وضاحت کریں۔
3) ٹیسٹ فار گڈنیس آف فٹ (Test for Goodness of Fit) کی اہمیت واضح کریں، نیز اسے کاروباری میدان میں مثال کے ساتھ بیان کریں۔
4) کائی اسکوائر ٹیسٹ کی حدود اور شرائط (Limitations or Conditions) کون سی ہیں؟ ان کی تفصیلی وضاحت کریں۔
5) کاروباری تجزیات (Business Analytics) میں کائی اسکوائر تجزیے کی اہمیت واضح کریں اور عملی زندگی سے متعلق کم از کم دو مثالیں دیں۔

اکائی 12۔ تغیر کا تجزیہ

(Analysis of Variance)

اکائی کے اجزاء	Structure
12.0 تمہید	Introduction
12.1 مقاصد	Objectives
12.2 انووا(ANOVA) کی تعریف	Meaning of ANOVA
12.3 انووا (ANOVA) کی ضرورت	Need for ANOVA
12.4 انووا (ANOVA) کا مقصد	Objective of ANOVA
12.5 ایک طرفہ انووا(ANOVA)ر	One-Way ANOVA
12.6 دوطرفہ انووا (ANOVA)	Two-Way ANOVA
12.7 دوطرفہ انووا(ANOVA) کے اقسام	Types of Two-Way ANOVA
12.8 کاروباری تجزیات میں انووا (ANOVA) کی اہمیت	Importance of ANOVA in Business Analytics
12.9 اکتسابی نتائج	Learning Outcomes
12.9 نمونہ امتحانی سوالات	Model Exam Questions

12.0 تمہید(Introduction)

گزشتہ یونٹ (یونٹ 11) میں آپ نے کائی اسکوائر تجزیہ (Chi-Square Analysis) کے بنیادی تصورات سیکھے، جس میں ٹیسٹ فار گڈنیس آف فٹ (Test for Goodness of Fit) اور ٹیسٹ فار انڈیپینڈنس آف ایٹریبیوٹس (Test for Independence of Attributes) شامل تھے۔ ان ٹیسٹس کے ذریعے مختلف زمرہ جاتی (Categorical) متغیرات کے مابین تعلق یا آزادی کو پرکھنے کا طریقہ سمجھا گیا۔

اب یونٹ 12 میں ہم ایک اور اہم شماریاتی تکنیک تجزیہ تغیر (Analysis of Variance - ANOVA) پر گفتگو کریں گے، جو دو یا زیادہ گروپوں کے درمیان اوسط (Mean) کے فرق کو جانچنے کے لیے استعمال کی جاتی ہے۔ جب ہمیں یہ معلوم کرنا ہو کہ مختلف گروپس کے اوسط میں کوئی معنی خیز فرق موجود ہے یا نہیں، تو ANOVA ایک مؤثر تکنیک ثابت ہوتی ہے۔

اس یونٹ میں آپ ایک طرفہ تجزیہ تغیر (One-Way ANOVA) اور دو طرفہ تجزیہ تغیر (Two-Way ANOVA) کا تجزیہ کریں گے۔ ساتھ ہی، دو طرفہ ANOVA کے وہ کیسز بھی سیکھیں گے جہاں عوامل کے درمیان تنور ملل (Interaction) موجود ہوتا ہے اور وہ کیسز بھی جہاں تنور ملل موجود نہیں ہوتا۔

یہ یونٹ آپ کو کاروباری اور تحقیقی تجزیات میں ANOVA کے عملی استعمال کو سمجھنے میں مدد دے گا، خاص طور پر جب مختلف گروپوں کے ڈیٹا کا موازنہ درکار ہو۔

12.1 مقاصد (Objectives)

اس اکائی کا مطالعہ کرنے کے بعد، آپ مندرجہ ذیل نکات کو سمجھ سکیں گے:

- تجزیہ تغیر (ANOVA) کے بنیادی تصورات کو سمجھ سکیں اور اس کی ضرورت و مقصد کو واضح کر سکیں۔
- ایک طرفہ تجزیہ تغیر (One-Way ANOVA) اور دو طرفہ تجزیہ تغیر (Two-Way ANOVA) کے درمیان فرق کو مثالوں کے ذریعے بیان کر سکیں۔
- Two-Way ANOVA میں تنور ملل (Interaction) کی موجودگی اور عدم موجودگی کے اثرات کو جانچ سکیں۔
- ANOVA کی کاروباری تجزیات، تحقیق اور دیگر عملی میدانوں میں اہمیت کو واضح کر سکیں۔

12.2 انووا (ANOVA) کی تعریف (Definition of Analysis of ANOVA)

نووا (Analysis of Variance - ANOVA) ایک شماریاتی تکنیک ہے جو دو یا زیادہ گروپوں کے درمیان اوسط (Mean) کے فرق کو جانچنے کے لیے استعمال کی جاتی ہے۔ اگر ہم یہ جاننا چاہتے ہیں کہ مختلف گروپس کے درمیان کوئی معنی خیز فرق موجود ہے یا نہیں، تو ANOVA اس کا بہترین طریقہ ہے۔

یہ طریقہ اس مفروضے (Hypothesis) کو پرکھنے کے لیے استعمال ہوتا ہے کہ تمام گروپوں کے اوسط برابر ہیں یا نہیں۔ اگر فرق پایا جائے تو ANOVA یہ طے کرنے میں مدد دیتا ہے کہ آیا یہ فرق محض اتفاقیہ (By Chance) ہے یا شماریاتی طور پر نمایاں (Statistically Significant) ہے۔

12.3 انووا(ANOVA) کی ضرورت (Need for ANOVA)

جب ہمیں دو سے زیادہ گروپوں کے اوسط (Mean) کا موازنہ کرنا ہو، تو ANOVA ایک مؤثر شماریاتی تکنیک ثابت ہوتی ہے۔ اگر صرف دو گروپس کا موازنہ کرنا ہو تو t-Test استعمال کیا جا سکتا ہے، لیکن اگر گروپوں کی تعداد تین یا زیادہ ہو تو t-Test بار بار کرنے سے شماریاتی غلطیوں (Errors) کا احتمال بڑھ جاتا ہے۔ اس مسئلے کو حل کرنے کے لیے ANOVA استعمال کیا جاتا ہے، جو تمام گروپوں کا بیک وقت موازنہ کرنے میں مدد دیتا ہے۔

ANOVA کی ضرورت درج ذیل وجوہات کی بنا پر محسوس کی جاتی ہے:

1. متعدد گروپس کے موازنے میں آسانی:

اگر ہمارے پاس تین یا اس سے زیادہ گروپس ہوں تو ANOVA ہمیں ایک ہی ٹیسٹ کے ذریعے ان کا موازنہ کرنے میں مدد دیتا ہے، بجائے اس کے کہ ہم متعدد t-Tests استعمال کریں۔

2. شماریاتی غلطیوں سے بچاؤ:

بار بار t-Test استعمال کرنے سے غلطی (Type I Error) کے احتمالات بڑھ جاتے ہیں، جبکہ ANOVA ایک ہی بار میں تمام گروپس کا موازنہ کر کے اس خطرے کو کم کرتا ہے۔

3. گروپس کے اندر اور گروپس کے درمیان تغیر کا موازنہ:

ANOVA صرف اوسط کا موازنہ نہیں کرتا، بلکہ یہ گروپوں کے اندرونی (Within-Group) اور باہری (Between-Group) تغیر (Variance) کا بھی جائزہ لیتا ہے۔

4. کاروباری اور تحقیقی تجزیہ میں مدد:

مارکیٹ ریسرچ، ہیلتھ کیئر، تعلیمی میدان، اور فنانس میں مختلف گروپوں کے ڈیٹا کا تجزیہ کرنے کے لیے ANOVA کا استعمال ناگزیر ہے۔

12.4 انووا ANOVA() کا مقصد (Objective of ANOVA)

ANOVA کا بنیادی مقصد یہ طے کرنا ہے کہ آیا دو یا زیادہ گروپوں کے درمیان اوسط میں کوئی معنی خیز فرق موجود ہے یا نہیں۔ اگر فرق پایا جاتا ہے تو ANOVA یہ بھی بتاتا ہے کہ آیا یہ فرق محض اتفاقیہ (Random) ہے یا شماریاتی لحاظ سے اہم (Significant) ہے۔

ANOVA کے مقاصد درج ذیل ہیں:

متعدد گروپس کے اوسط کا شماریاتی تجزیہ کرنا ANOVA: اس بات کو پرکھنے میں مدد دیتا ہے کہ کیا مختلف گروپوں کے درمیان کوئی حقیقی فرق موجود ہے یا نہیں۔

شماریاتی اہمیت (Statistical Significance) کی جانچ ANOVA: اس فرق کی اہمیت کو جانچ کر یہ طے کرتا ہے کہ کیا گروپس کے درمیان پایا جانے والا فرق صرف اتفاقیہ (By Chance) ہے یا واقعی کسی وجہ سے ہے۔

کاروباری اور سائنسی فیصلوں میں معاونت: کاروباری ادارے اور تحقیقاتی ادارے ANOVA کی مدد سے بہتر فیصلے کر سکتے ہیں، مثلاً مختلف مارکیٹنگ اسٹریٹیجیز، پروڈکٹ کی کارکردگی، یا کسی طبّی علاج کے اثرات کا تجزیہ۔

گروپس کے اندر اور گروپس کے درمیان تغیر کی شناخت ANOVA: یہ بھی دیکھتا ہے کہ کسی تجربے میں گروپس کے اندرونی فرق (Within-group variance) اور گروپس کے درمیان فرق (Between-group variance) کتنا نمایاں ہے۔

12.5 ایک طرفہ انووا (ANOVA) (One-Way ANOVA)

ایک طرفہ تجزیہ تغیر (One-Way ANOVA) ایک شماریاتی تکنیک ہے جو ایک آزاد متغیر (Independent Variable) کی بنیاد پر تین یا زیادہ گروپوں کے اوسط (Mean) کا موازنہ کرنے کے لیے استعمال کی جاتی ہے۔ اس کا بنیادی مقصد یہ جانچنا ہوتا ہے کہ کیا مختلف گروپوں کے اوسط میں کوئی معنی خیز فرق (Significant Difference) موجود ہے یا نہیں۔

One-Way ANOVA کا مطلب یہ ہے کہ ہم صرف ایک آزاد متغیر (Factor) کو تبدیل کر کے گروپوں کے اوسط میں فرق کا تجزیہ کر رہے ہیں۔

One-Way ANOVA کے مفروضے:(Assumptions)

One-Way ANOVA کو درست طریقے سے لاگو کرنے کے لیے درج ذیل شرائط پوری ہونی چاہیے:

1. نمونے آزاد ہونے چاہیے:(Independence of Samples)

ہر گروپ میں موجود مشاہدات ایک دوسرے سے آزاد (Independent) ہونے-چاہیے

2. گروپس میں ڈیٹا نارمل ڈسٹری بیوشن (Normally Distributed) ہونا چاہیے:

گروپس کے اندر موجود ڈیٹا نارمل تقسیم (Normal Distribution) کے قریب ہونا چاہیے۔

3. واریانس (Variance) یکساں ہونا چاہیے:(Homogeneity of Variance)

تمام گروپس میں تغیر (Variance) تقریباً برابر ہونا چاہیے تاکہ موازنہ درست ہو۔

One-Way ANOVA کا شماریاتی فارمولا:

$$F = \frac{\text{Between Group Variance}}{\text{Within Group Variance}}$$

جہاں:

- Between Group Variance: مختلف گروپس کے اوسط (Mean) میں پائے جانے والا فرق۔
- Within Group Variance: ہر گروپ کے اندر موجود ڈیٹا میں پائے جانے والا فرق۔
- F-Statistic: یہ شماریاتی قدر (Statistic) ہمیں بتاتی ہے کہ آیا گروپس کے درمیان پایا جانے والا فرق معنی خیز (Significant) ہے یا نہیں۔

12.6 دوطرفہ 12.2 انووا (Two-Way ANOVA)

دوطرفہ تجزیہ تغیر (Two-Way ANOVA) ایک شماریاتی تکنیک ہے جو دو آزاد متغیرات (Independent Variables) کے اثر کو جانچنے کے لیے استعمال کی جاتی ہے۔ One-Way ANOVA کے برعکس، جہاں صرف ایک فیکٹر (Factor) کا اثر جانچا جاتا ہے، Two-Way ANOVA میں دو مختلف عوامل (Factors) کے مشترکہ اثرات کو بھی پرکھا جاتا ہے۔

یہ تکنیک خاص طور پر اس وقت استعمال کی جاتی ہے جب یہ جانچنا ہو کہ دو مختلف متغیرات کسی ایک تابع متغیر (Dependent Variable) پر کس طرح اثرانداز ہو رہے ہیں اور آیا ان دونوں کے درمیان کوئی تنورملل (Interaction Effect) موجود ہے یا نہیں۔

Two-Way ANOVA کے مفروضے: (Assumptions)

Two-Way ANOVA کے درست اطلاق کے لیے درج ذیل مفروضے پورے ہونا ضروری ہیں:

1. ڈیٹا نارمل تقسیم (Normally Distributed) ہونا چاہیے۔
2. تمام گروپوں میں واریانس (Variance) یکساں ہونا چاہیے۔
3. مشاہدات آزاد ہونے چاہئیں (Independent Observations)۔
4. دو مختلف آزاد متغیرات (Factors) کا اثر جانچنے کے لیے مناسب نمونے ہونے چاہئیں۔

Two-Way ANOVA کا شماریاتی فارمولا:

$$F = \frac{\text{Variation Due to Factor A, Factor B, and Interaction}}{\text{Within-Group Variation}}$$

جہاں:

- Factor A پہلا آزاد متغیر
- Factor B دوسرا آزاد متغیر
- Interaction دونوں عوامل کے مشترکہ اثرات
- Within-Group Variation ہر گروپ کے اندر موجود تغیر

12.7 دوطرفہ انووا کی اقسام (Types of Two-Way ANOVA)

Two-Way ANOVA کو نارمل طور پر **دو اقسام** میں تقسیم کیا جاتا ہے:

1) Two-Way ANOVA بغیر تنور ملل (Two-Way ANOVA Without Interaction)

یہ وہ صورت ہے جہاں دو آزاد متغیرات (Independent Variables) کا اثر ایک تابع متغیر (Dependent Variable) پر جانچا جاتا ہے، لیکن یہ فرض کیا جاتا ہے کہ ان دونوں متغیرات کے درمیان کوئی تنور ملل (Interaction) موجود نہیں ہے۔ اس تجزیے میں دونوں عوامل (Factors) کا الگ الگ اثر دیکھا جاتا ہے، لیکن یہ پرکھا نہیں جاتا کہ آیا یہ دونوں عوامل ایک دوسرے پر اثرانداز ہوتے ہیں یا نہیں۔

فرض کریں کہ ایک تعلیمی تحقیق میں یہ جانچنا ہو کہ طلبہ کی جنس (Male/Female) اور تدریسی طریقہ (Lecture-based/Practical-based) کس حد تک طلبہ کی تعلیمی کارکردگی پر اثرانداز ہوتا ہے۔

- آزاد متغیر 1 (Factor A): طلبہ کی جنس (Male, Female)
- آزاد متغیر 2 (Factor B): تدریسی طریقہ (Lecture-based, Practical-based)
- تابع متغیر: امتحانی نمبرز

Two-Way ANOVA Without Interaction **یہ بتائے گا کہ:**

- کیا تدریسی طریقہ کارکردگی پر اثر ڈالتا ہے؟
- کیا جنس کے لحاظ سے کارکردگی میں فرق پایا جاتا ہے؟
- لیکن یہ نہیں بتائے گا کہ آیا تدریسی طریقہ اور جنس ایک ساتھ مل کر کارکردگی پر اثر ڈالتے ہیں یا نہیں۔

2) Two-Way ANOVA ساتھ تنور ملل (Two-Way ANOVA With Interaction)

یہ وہ صورت ہے جہاں دو آزاد متغیرات کے انفرادی اثرات کے ساتھ ساتھ ان کے باہمی تنور ملل (Interaction Effect) کا بھی تجزیہ کیا جاتا ہے۔ اس کا مطلب یہ ہے کہ ہم یہ بھی جانچ سکتے ہیں کہ آیا ایک متغیر کا اثر دوسرے متغیر کی موجودگی میں مختلف ہوتا ہے یا

نہیں۔اس تجزیے میں دونوں عوامل کے انفرادی اثرات کے علاوہ یہ بھی دیکھا جاتا ہے کہ کیا یہ دونوں عوامل ایک دوسرے پر اثر انداز ہوتے ہیں؟

فرض کریں کہ ایک بھارتی موبائل کمپنی یہ جانچنا چاہتی ہے کہ مختلف شہروں (ممبئی، دہلی، حیدرآباد) اور تشہیری مہم (TV Ads, Social Media Ads) کس حد تک موبائل فون کی فروخت پر اثر انداز ہوتے ہیں۔

- آزاد متغیر1 (Factor A): شہر (ممبئی، دہلی، حیدرآباد)
- آزاد متغیر2 (Factor B): اشتہاری طریقہ (TV Ads, Social Media Ads)
- تابع متغیر: موبائل فون کی فروخت

Two-Way ANOVA With Interaction کے ذریعے درج ذیل سوالات کے جوابات حاصل ہوں گے:

- کیا مختلف شہروں میں فروخت کا فرق معنی خیز ہے؟
- کیا TV اور Social Media اشتہارات کی تاثیر میں فرق پایا جاتا ہے؟
- کیا کسی مخصوص شہر میں کوئی خاص اشتہاری طریقہ زیادہ مؤثر ہے؟ یہ تنور ملل کا اثر ہے

Illustration 1: Two random samples were drawn from two normal populations and their values are:

A	**65**	**66**	**73**	**80**	**82**	**84**	**88**	**90**	**92**
B	64	66	74	78	82	85	87	92	93

Test whether the two populations have the same variance at the 5% level of significance.

(Given: $F = 3.36$ at 5% level for $v_1 = 10$ and $v_2 = 8$)

Solution: Let us take the null hypothesis that the two populations have not the same variance.

Applying F-test:

$$F = \frac{S_1^2}{S_2^2}$$

A (X_1)	$(X_1 - \bar{X}_1) = x_1$	x_1^2	B (X_2)	$(X_2 - \bar{X}_2) = x_2$	x_2^2

65	-15	225	64	-19	361
66	-14	196	66	-17	289
73	-7	49	74	-9	81
80	0	0	78	-5	25
82	2	4	82	-1	1
84	4	16	85	2	4
88	8	64	87	4	16
90	10	100	92	9	81
92	12	144	93	10	100
			95	12	144
			97	14	196

$$\Sigma X_1 = 720, \quad \Sigma x_1 = 0, \quad \Sigma x_1^2 = 798$$

$$\Sigma X_2 = 913, \quad \Sigma x_2 = 0, \quad \Sigma x_2^2 = 1298$$

$$\bar{X}_1 = \frac{\Sigma X_1}{n_1} = \frac{720}{9} = 80$$

$$\bar{X}_2 = \frac{\Sigma X_2}{n_2} = \frac{913}{11} = 83$$

$$S_1^2 = \frac{\Sigma x_1^2}{n_1 - 1} = \frac{798}{9 - 1} = 99.75$$

$$S_2^2 = \sum x_2^2 / n_2 - 1 = \frac{734}{11 - 1} = 129.8$$

$$F = \frac{S_1^2}{S_2^2} = \frac{99.75}{129.8} = 0.768$$

At 5 percent level of significance, for $v_1 = 10$ and $v_2 = 8$, the table value of $F_{0.05} = 3.36$.

The calculated value of F is less than the table value. The hypothesis is accepted. Hence the two populations have not the same variance. ✓

12.8 کاروباری تجزیات میں ANOVA کی اہمیت

(Importance of ANOVA in Business Analytics)

تجزیہ تغیر (ANOVA - Analysis of Variance) کاروباری تجزیات میں ایک انتہائی مؤثر شماریاتی تکنیک ہے، جو مختلف عوامل (Factors) کے کاروباری کارکردگی پر اثرات کو جانچنے کے لیے استعمال کی جاتی ہے۔ ANOVA کے ذریعے ادارے یہ معلوم کر سکتے ہیں کہ آیا مختلف گروپس کے درمیان کارکردگی یا نتائج میں کوئی شماریاتی لحاظ سے نمایاں فرق موجود ہے یا نہیں۔

ANOVA کا بنیادی مقصد یہ ہوتا ہے کہ کاروباری ادارے اپنے فیصلوں کو ڈیٹا پر مبنی شواہد (Data-Driven Evidence) کے ذریعے بہتر بنا سکیں، اور کسی خاص پالیسی یا حکمت عملی کی مؤثریت (Effectiveness) کو جانچ سکیں۔

کاروباری تجزیات میں ANOVA کی اہمیت:

1) مارکیٹ ریسرچ:(Market Research)

مارکیٹ ریسرچ میں ANOVA ایک طاقتور ٹول کے طور پر استعمال ہوتا ہے۔ کمپنیاں مختلف مارکیٹ سیگمنٹس میں مصنوعات کی فروخت کا موازنہ کرنے کے لیے ANOVA استعمال کرتی ہیں۔ مثال: ایک بھارتی فاسٹ فوڈ چین یہ جاننا چاہتی ہے کہ کیا ممبئی، دہلی، اور حیدرآباد میں اس کی مصنوعات کی اوسط فروخت میں کوئی نمایاں فرق موجود ہے یا نہیں۔ ANOVA کے ذریعے کمپنی فیصلہ کر سکتی ہے کہ کون سے شہر میں سب سے زیادہ مانگ (Demand) ہے اور اپنی کاروباری حکمت عملی کو اس کے مطابق ترتیب دے سکتی ہے۔

2) قیمتوں کی پالیسی:(Pricing Strategy)

ANOVA قیمتوں کے تعین (Pricing Strategy) میں بھی مددگار ثابت ہوتا ہے۔ اگر کسی کمپنی کو مختلف قیمتوں کے ماڈلز کا تجزیہ کرنا ہو تو ANOVA کے ذریعے وہ یہ معلوم کر سکتی ہے کہ کون سی قیمت زیادہ مؤثر ہے۔ مثال: اگر ایک موبائل کمپنی اپنے فون کی ₹15,000، ₹20,000، اور ₹25,000 کی قیمتوں پر فروخت کے اثرات کا موازنہ کرنا چاہتی ہے تو One-Way ANOVA کی مدد سے وہ دیکھ سکتی ہے کہ کون سی قیمت صارفین کے لیے زیادہ پرکشش ہے۔

3) اشتہارات اور مارکیٹنگ:(Advertising & Marketing)

ANOVA مختلف اشتہاری مہمات (Ad Campaigns) کی مؤثریت کو جانچنے کے لیے استعمال کیا جاتا ہے تاکہ کاروباری ادارے یہ معلوم کر سکیں کہ کون سی حکمت عملی زیادہ کامیاب ہے۔ مثال: فرض کریں کہ ایک ای-کامرس کمپنی سوشل میڈیا، ٹی وی، اور بل بورڈز کے ذریعے اشتہار چلا رہی ہے۔ Two-Way ANOVA کے ذریعے کمپنی یہ جانچ سکتی ہے کہ کون سا پلیٹ فارم سب سے زیادہ مؤثر ہے اور آیا مخصوص اشتہاری پلیٹ فارم کسی خاص عمر کے گروپ (مثلاً نوجوان، بالغ، یا بزرگ) پر مختلف اثر ڈال رہا ہے یا نہیں۔

4) انسانی وسائل اور کارکردگی کا تجزیہ:(HR & Performance Analysis)

ANOVA مختلف ملازمین کی کارکردگی کے تجزیے کے لیے استعمال کیا جا سکتا ہے، جس سے ادارے اپنی ہیومن ریسورس پالیسی کو بہتر بنا سکتے ہیں۔ مثال: ایک کمپنی اپنے ملازمین کی کارکردگی کو تین مختلف تربیتی پروگرامز (Training Programs) میں تقسیم کر کے جانچنا چاہتی ہے۔ ANOVA کی مدد سے معلوم کیا جا سکتا ہے کہ کون سا تربیتی پروگرام سب سے زیادہ مؤثر ہے اور کس سے ملازمین کی کارکردگی میں سب سے زیادہ بہتری آئی ہے۔

5) مصنوعات کی کوالٹی کنٹرول:(Product Quality Control)

مینوفیکچرنگ انڈسٹری میں ANOVA مصنوعات کے معیار (Quality Control) کو جانچنے کے لیے استعمال ہوتا ہے۔ مختلف پروڈکشن لائنز کے درمیان معیار کے فرق کو جانچنے کے لیے ANOVA کا اطلاق کیا جاتا ہے۔ مثال: ایک آٹوموبائل کمپنی اپنے تین مختلف مینوفیکچرنگ یونٹس میں بننے والی گاڑیوں کی اوسط کارکردگی کا موازنہ کرنے کے لیے One-Way ANOVA کا استعمال کر سکتی ہے تاکہ یہ معلوم ہو سکے کہ کون سی پروڈکشن لائن سب سے زیادہ معیاری نتائج دے رہی ہے۔

ANOVA کاروباری تجزیات (Business Analytics) میں فیصلہ سازی کے لیے ایک طاقتور شماریاتی ٹول ہے، جو مختلف عوامل (Factors) کے اثرات کو جانچنے اور مختلف گروپس کے درمیان نمایاں فرق کو پرکھنے میں مدد دیتا ہے۔ اس کی مدد سے مارکیٹ ریسرچ، مصنوعات کی قیمتوں، اشتہارات، ملازمین کی کارکردگی، اور کوالٹی کنٹرول جیسے اہم کاروباری فیصلوں کو بہتر بنایا جا سکتا ہے۔

ANOVA کا درست استعمال کاروباری اداروں کو ڈیٹا پر مبنی شواہد کے ساتھ بہتر اور مؤثر فیصلے کرنے میں مدد فراہم کرتا ہے، جس سے مسابقتی برتری (Competitive Advantage) حاصل کی جا سکتی ہے۔

12.9 نمونہ امتحانی سوالات (Model Exam Questions)

معروضی جوابات کے حامل سوالات(Objective Answer Type Questions)

1) تجزیہ تغیر (ANOVA) کا بنیادی مقصد کیا ہے؟
2) ANOVA میں **F-Statistic** کس چیز کا تجزیہ کرتا ہے؟
3) One-Way ANOVA میں کتنے آزاد متغیرات (Factors) کا تجزیہ کیا جاتا ہے؟
4) Two-Way ANOVA میں کتنے آزاد متغیرات (Factors) ہوتے ہیں؟
5) One-Way ANOVA اور Two-Way ANOVA میں بنیادی فرق کیا ہے؟
6) ANOVA میں تنور ملل (Interaction Effect) سے کیا مراد ہے؟
7) ANOVA میں صفر مفروضہ (Null Hypothesis - H_0) کیا ظاہر کرتا ہے؟
8) اگر ANOVA کی $p\text{-value} < 0.05$ ہو تو اس کا کیا مطلب ہے؟

9) ANOVA میں Between Group Variance اور Within Group Variance کا کیا مطلب ہے؟

10) Two-Way ANOVA کس قسم کے ڈیٹا پر استعمال کیا جاتا ہے؟

مختصر جوابات کے حامل سوالات (Short Answers Type Questions)

1) ANOVA کی ضرورت اور اس کے عملی استعمال کی وضاحت کریں۔

2) One-Way ANOVA کے مفروضے (Assumptions) کیا ہیں؟ تفصیل سے بیان کریں۔

3) Two-Way ANOVA کے دو مختلف اقسام (With & Without Interaction) میں فرق واضح کریں۔

4) کاروباری تجزیات (Business Analytics) میں ANOVA کا کردار بیان کریں۔

5) ANOVA اور t-Test میں کیا فرق ہے؟ تقابلی جائزہ پیش کریں۔

طویل جوابات کے حامل سوالات (Long Answers Type Questions)

1) تجزیہ تغیر (ANOVA) کی بنیادی تعریف، ضرورت اور کاروباری و تحقیقی میدان میں اس کی اہمیت پر تفصیلی نوٹ لکھیں۔

2) One-Way ANOVA کا تفصیل سے تجزیہ کریں، اس کا فارمولا، مفروضے اور عملی مثال کے ساتھ وضاحت کریں۔

3) Two-Way ANOVA کے مفہوم کو واضح کریں، اس کے استعمال اور مختلف اقسام (With & Without Interaction) کو مثالوں کے ساتھ بیان کریں۔

4) ANOVA میں F-Statistic کی تشریح کریں، اس کے حساب کے طریقہ کار اور عملی اطلاق کو تفصیل سے بیان کریں۔

5) کسی کاروباری یا تعلیمی تحقیق کے لیے ANOVA کے استعمال کی وضاحت کریں اور ایک حقیقی مثال کے ساتھ اس کے اطلاق کا طریقہ کار بیان کریں۔

MAULANA AZAD NATIONAL URDU UNIVERSITY

CENTRE FOR DISTANCE AND ONLINE EDUCATION

MBA I SEMESTER

Semester Examination, Nov/Dec 2025

Course Code: MMBA123CCT Subject: Business Analytics & Statistics

Total Marks: 70 Times: 3 Hours

یہ پرچہ تین حصوں پر مشتمل ہے: حصہ اول، حصہ دوم، حصہ سوم۔ ہر جواب کے لیے الفاظ کی تعداد اشارہ ہے۔ تمام حصوں سے سوالوں کا جواب دینا لازمی ہے۔

1. حصہ اول میں دس لازمی سوالات ہیں جو کہ معروضی سوالات / خالی جگہوں پر کرنا / مختصر جواب والے سوالات ہیں۔ ہر سوال کا جواب لازمی ہے۔ ہر سوال کے لیے 1 نمبر ہے۔ (1x10 = 10 marks)
2. حصہ دوم میں 8 سوالات ہیں۔ اس میں سے طالب علم کو کوئی پانچ سوالوں کے جواب دینے ہیں۔ ہر سوال کا جواب تقریباً (100) الفاظ پر مشتمل ہو۔ ہر سوال کے لیے 6 نمبرات مختص ہیں۔ (5x6 = 30 marks)
3. حصہ سوم میں 5 سوالات ہیں۔ اس میں سے طالب علم کو کوئی تین سوالوں کے جواب دینے ہیں۔ ہر سوال کا جواب تقریباً (250) الفاظ پر مشتمل ہو۔ ہر سوال کے لیے 10 نمبرات مختص ہیں۔ (3x10 = 30 marks)

حصہ اول

درج ذیل تمام سوالوں کا جواب دیں۔ **(10 x 01 = 10 Marks)**

1) بزنس اینالیٹکس کیا ہے؟
2) بزنس اینالیٹکس کی کسی بھی دو قسموں کے نام بتائیں۔
3) مارکیٹ کے غیر یقینی حالات میں پیش گوئی کرنے والے تجزیات اسٹریٹجک فیصلوں کے معیار کو کیسے بڑھا سکتے ہیں؟
4) کسی تنظیم کے اندر ڈیٹا سے چلنے والی ثقافت کو فروغ دینے میں قیادت کیا کردار ادا کرتی ہے؟
5) طاق اعداد کی مثالیں...............ہیں۔
6) کارپیرسن کا تعلق...............ملک سے ہے۔
7) ایک ناممکن ایونٹ کا احتمال صفر ہے۔ صحیح
8) جمعی ماڈل میں اجزا کے اثرات کی نوعیت جمعی نہیں ہوتی۔ (صحیح/غلط)
9) تجزیہ تغیر (ANOVA) کا بنیادی مقصد کیا ہے؟
10) ANOVA میں **F-Statistic** کس چیز کا تجزیہ کرتا ہے؟

حصہ دوم

درج ذیل سوالات میں سے کسی پانچ کی وضاحت کریں۔ **(5 x 6 = 30 Marks)**

1. وضاحتی تجزیات میں استعمال ہونے والے ٹولز پر مثالوں کے ساتھ بحث کریں۔
2. ایک کمپنی کس طرح انتظام کی تمام سطحوں پر ڈیٹا پر مبنی کلچر بنا سکتی ہے؟
3. مؤثر فیصلہ سازی میں ڈیٹا گورننس اور ڈیٹا کا معیار کیا کردار ادا کرتا ہے؟
4. ایک ایسے منظر نامے کی وضاحت کریں جہاں ڈیٹا کی ناقص تشریح کی وجہ سے کاروباری فیصلہ غلط ہوا۔ تجزیات سے نتیجہ کیسے بدل سکتا ہے؟
5. ذیل کے مواد کی عدد سے بہتانیہ محسوب کیجئے۔

تعدد	9	12	14	18	19	20
F	3	4	3	2	6	10

6. دیے ہوئے اعداد کی مدد سے راست اور بالراست طریقہ پر معیاری انحراف اور تفاوت محسوب کیجئے۔

X	10	20	30	11	14	17	20	23
F	1	2	2	3	4	5	2	1

7. احتمالی تقسیم میں متوقع قدر کے تصور کی وضاحت کریں۔

سادہ نیم اواسط کے طریقہ کے تحت درجہ ذیل اعداد و شمار کے لیے رجحان کا تعین کیجیے۔

سال	2001	2002	2003	2004	2005	2006	2007	2008	2009
پیداوار (ٹنوں میں)	16	14	20	18	19	12	23	28	25

8. کاروباری تجزیات (Business Analytics) میں ANOVA کا کردار بیان کریں۔

حصہ سوم

درج ذیل سوالات میں سے کسی تین کی وضاحت کریں۔ **(3 x 10 = 30 Marks)**

1) سوشل میڈیا اینالیٹکس کی اہمیت اور مختلف پلیٹ فارمز اسے کاروباری بصیرت کے لیے کس طرح لاگو کرتے ہیں اس کی وضاحت کریں۔

2) اینالیٹکس لائف سائیکل کے آغاز میں مسئلہ کی تعریف تجزیاتی پروجیکٹ کے نتائج کو کیسے متاثر کر سکتی ہے ؟

3) ایم بی اے کے طلباء انٹرن شپ یا لائیو پروجیکٹس کے دوران ڈیٹا پر مبنی فیصلہ سازی، اصولوں کا اطلاق کیسے کر سکتے ہیں ؟

4) ذیل میں دیئے گئے تفصیلات کی مدد سے Q_1, Q_2 اور Q_3 محسوب کیجئے۔

وقفہ جماعت نشانات	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
تعداد F	1	3	5	7	11	9	2	2

5) کسی کاروباری یا تعلیمی تحقیق کے لیے ANOVA کے استعمال کی وضاحت کریں اور ایک حقیقی مثال کے ساتھ اس کے اطلاق کا طریقہ کار بیان کریں۔