



राजस्थान

राज्य पात्रता परीक्षा (SET)

पेपर - 1

भाग - 3

डेटा इंटरप्रिटेशन, सूचना एवं संचार प्रौद्योगिकी, पर्यावरण एवं प्रदूषण और उच्च शिक्षा संगठन एवं प्रणाली



Index

क्र.सं.	अध्याय	पृ.सं.
UNIT – VII डेटा व्याख्या		
1.	1. डेटा स्रोत, डेटा का अधिग्रहण और वर्गीकरण	1
	2. मात्रात्मक और गुणात्मक डेटा	6
	3. ग्राफिकल प्रतिनिधित्व	8
	4. डेटा और शासन	12
	5. डाटा इंटरप्रिटेशन	13
	6. अभ्यास प्रश्न	32
UNIT – VIII सूचना और संचार प्रौद्योगिकी		
2.	1. आईसीटी	37
	2. इंटरनेट से संबंधित महत्वपूर्ण शब्द	50
	3. उच्च शिक्षा में डिजिटल पहल	62
	4. ई-शासन	84
	5. अभ्यास प्रश्न	88
UNIT – IX लोग, विकास और पर्यावरण		
3.	1. विकास और पर्यावरण	93
	2. मानव और पर्यावरण सहभागिता	96
	3. पर्यावरण के मुद्दें	98
	4. मानव स्वास्थ्य पर प्रदूषकों के प्रभाव	104
	5. प्राकृतिक और ऊर्जा संसाधन	108
	6. प्राकृतिक खतरे और आपदाएँ	111
	7. भारत का पर्यावरण संरक्षण कानून	113
	8. अभ्यास प्रश्न	121
UNIT – X उच्च शिक्षा प्रणाली		
4.	1. प्राचीन भारत में उच्च शिक्षा और शिक्षा के संस्थान	125
	2. भारत में प्राच्य, पारंपरिक और गैर-पारंपरिक शिक्षण कार्यक्रम।	134
	3. व्यावसायिक, तकनीकी और कौशल आधारित शिक्षा	136
	4. मूल्य शिक्षा और पर्यावरण शिक्षा	139
	5. नीतियां, शासन और प्रशासन	154
	6. अभ्यास प्रश्न	164

VII UNIT

डेटा व्याख्या

अध्याय 1

डेटा स्रोत, डेटा का अधिग्रहण और वर्गीकरण

पिछले वर्ष के प्रश्न

प्रश्न 1 सूची I को सूची II से सुमेलित करें

(2022)

सूची I (फाइल एक्सटेंशन)	सूची II (फाइल प्रकार)
(A) mp3	(I) छवि
(B) xls	(II) संगीत
(C) jpeg	(III) डेटाबेस
(D) mdb	(IV) स्प्रेडशीट

नीचे दिए गए विकल्पों में से सही उत्तर चुनें:

- (A) (A)-(I); (B)-(IV); (C)-(II); (D)-(III)
(B) (A)-(II); (B)-(IV); (C)-(I); (D)-(III)
(C) (A)-(II); (B)-(IV); (C)-(III); (D)-(1)
(D) (A)-(III); (B)-(II); (C)-(I); (D)-(IV)

प्रश्न 2 सूची I को सूची II से सुमेलित करें

(2023)

सूची I (प्रश्न)	सूची II (उत्तर)
(A) 512 के दशमलव संख्या के बाइनरी रूप में कितने 1 हैं?	(I) 13
(B) यदि 1 किलोबाइट = 2 बिट्स है, तो x का मान क्या होगा?	(II) 20
(C) जब हेक्साडेसिमल अंक A और B को जोड़ा जाता है, तो प्राप्त हेक्साडेसिमल मान क्या है?	(III) 1
(D) बाइनरी संख्या 00010100 का दशमलव (आधार 10) मान क्या है?	(IV) 15

नीचे दिए गए विकल्पों में से सही उत्तर चुनें:

- (A) (A)-(III), (B)-(IV), (C)-(I), (D)-(II)
(B) (A)-(II), (B)-(1), (C)-(IV), (D)-(III)
(C) (A)-(I), (B)-(III), (C)-(II), (D)-(IV)
(D) (A)-(III), (B)-(I), (C)-(IV), (D)-(II)

विश्लेषण - इस इकाई में आम तौर पर 5 प्रश्न होते हैं, जिनकी मूल अवधारणाएँ इकाई 5 के अध्याय 3 में शामिल की गई थीं। पिछले वर्ष के प्रश्नों (पीवाईक्यू) के लगातार अभ्यास से, इस इकाई में अच्छा स्कोर करना काफी संभव हो जाता है। नियमित अभ्यास से अवधारणाओं को मजबूत करने और आपके प्रदर्शन को बेहतर बनाने में मदद मिलेगी।

अध्ययन के उद्देश्य

- डेटा प्राप्ति और वर्गीकरण
- गुणात्मक और मात्रात्मक डेटा
- फोटोग्राफिक विवरण और डेटा मानचित्र
- डेटा की व्याख्या
- सांख्यिकी और सुशासन

परिचय

- दिए गए सामग्री में मुख्य रूप से कच्चे तथ्य, आंकड़े और सांख्यिकीय जानकारी होती है। डेटा वह प्रक्रिया है जिससे अर्थपूर्ण जानकारी प्राप्त की जाती है। उदाहरण के लिए, किसी संख्या को अक्सर सांख्यिकी कहा जाता है या किसी जानकारी का प्राप्त करना, उसकी गणना करना, प्रक्रिया करना या प्रसंस्करण करना। डेटा को किसी क्रम में व्यवस्थित करके संग्रहित, जांचा और संरक्षित किया जाता है। इसके बाद इसे विभिन्न व्यक्तियों को भेजा जाता है। इस प्रक्रिया में निम्नलिखित चरण शामिल होते हैं:

1. गणना - जोड़, घटाव, गुणा, भाग।
2. तुलना - समान, बड़ा, छोटा, शून्य, सकारात्मक, नकारात्मक।
3. निर्णय लेना - किसी स्थिति के आधार पर विभिन्न चरण।
4. तर्क करना - वांछित परिणाम प्राप्त करने के लिए शब्दों का क्रम।

- केवल संख्याओं की गिनती को प्रक्रिया नहीं कहा जाता। दस्तावेजों में त्रुटियां ढूंढना, कर की व्यवस्था करना आदि भी प्रक्रिया या प्रसंस्करण कहलाते हैं।

सूचना (Information)

- प्रक्रिया के बाद प्राप्त सामग्री को सूचना कहा जाता है। इसका अर्थ एक पूर्ण तथ्य, संख्या या सांख्यिकीय सूचना है। सरल शब्दों में, डेटा को प्रक्रिया करने के बाद जो अर्थपूर्ण डेटा प्राप्त होता है, उसे सूचना कहा जाता है। सूचना एक उपयोगी सामग्री है, जिसके गुण विभिन्न श्रेणियों के अनुसार होते हैं। सूचना निम्नलिखित कारणों से आवश्यक और सहायक होती है:

1. यह सूचना को संक्षेप और अधिक अर्थपूर्ण तरीके से प्रस्तुत करती है।
2. यह वर्तमान और भविष्य के लिए निर्णय लेने में मदद करती है।
3. यह भविष्य का मूल्यांकन करने में सहायक होती है।

सूचना के गुण

- हम जानते हैं कि सूचना किसी भी प्रणाली के लिए एक आवश्यक तत्व है, इसलिए सूचना में निम्नलिखित गुण होने चाहिए:

1. सामर्थ्य पूर्णता (Semantic Completeness)
2. शुद्धता (Chastity)
3. सटीकता (Accuracy)
4. पूर्व ज्ञान को सुधारना और निरंतरता बनाए रखना
5. संक्षिप्तता (Conciseness)
6. समय पर उपलब्धता (Timeliness)
7. कार्य संपादन में सहायक (Assistant in Editing the Work)

- डेटा किसी चर (variable) के गुणात्मक या मात्रात्मक मानों को परिभाषित करता है। डेटा "Datum" का बहुवचन है, जिसका शाब्दिक अर्थ है "दिया हुआ" या "कुछ दिया गया।" डेटा को जानकारी की सबसे छोटी इकाई माना जाता है, जिससे अन्य माप और विश्लेषण किए जा सकते हैं।
- डेटा संख्याएं, चित्र, शब्द, तथ्य, आंकड़े या विचार हो सकते हैं। डेटा अपने आप में समझा नहीं जा सकता और जानकारी प्राप्त करने के लिए इसे अर्थपूर्ण जानकारी में परिवर्तित करना आवश्यक होता है। डेटा की व्याख्या के लिए विभिन्न तरीके उपलब्ध हैं।

डेटा स्रोतों को मोटे तौर पर प्राथमिक और माध्यमिक डेटा में वर्गीकृत किया गया है

सांख्यिकीय विश्लेषण में डेटा संग्रह बहुत महत्वपूर्ण भूमिका निभाता है। अनुसंधान में, जानकारी इकट्ठा करने के लिए विभिन्न तरीकों का उपयोग किया जाता है, जो सभी दो श्रेणियों में आते हैं, यानी प्राथमिक और माध्यमिक डेटा (डगलस, 2015)।

जैसा कि नाम से पता चलता है, प्राथमिक डेटा वह होता है जिसे शोधकर्ता द्वारा पहली बार एकत्र किया जाता है जबकि द्वितीयक डेटा वह डेटा होता है जिसे पहले से ही दूसरों द्वारा एकत्र या उत्पादित किया जाता है।

प्राथमिक और द्वितीयक डेटा के बीच कई अंतर हैं, जिनकी चर्चा इस कार्य में की गई है। लेकिन सबसे महत्वपूर्ण अंतर यह है कि—

- प्राथमिक डेटा तथ्यात्मक और मूल होता है जबकि द्वितीयक डेटा केवल प्राथमिक डेटा का विश्लेषण और व्याख्या है।
- जबकि प्राथमिक डेटा को हाथ में समस्या का समाधान पाने के उद्देश्य से एकत्र किया जाता है, द्वितीयक डेटा अन्य उद्देश्यों के लिए एकत्र किया जाता है।
- प्राथमिक और द्वितीयक डेटा के बीच मूलभूत अंतर हैं; प्राथमिक डेटा शब्द का तात्पर्य शोधकर्ता द्वारा पहली बार उत्पन्न किए गए डेटा से है, जबकि द्वितीयक डेटा पहले से ही जांच एजेंसियों और संगठनों द्वारा एकत्र किया गया मौजूदा डेटा है।
- प्राथमिक डेटा स्रोतों में सर्वेक्षण, अवलोकन, प्रयोग, प्रश्नावली, व्यक्तिगत साक्षात्कार आदि शामिल हैं, इसके विपरीत, माध्यमिक डेटा संग्रह स्रोत सरकारी प्रकाशन, वेबसाइट, किताबें, जर्नल लेख, आंतरिक रिकॉर्ड आदि हैं।

प्राथमिक डेटा

प्राथमिक डेटा का अर्थ मूल डेटा है जिसे विशेष रूप से किसी उद्देश्य को ध्यान में रखकर एकत्र किया गया है। इसका मतलब है कि किसी ने पहले ही मूल स्रोत से डेटा एकत्र कर लिया है।

- इस प्रकार एकत्रित किये गये आँकड़े प्राथमिक आँकड़े कहलाते हैं। प्राथमिक डेटा अभी तक प्रकाशित नहीं हुआ है और यह अधिक विश्वसनीय, प्रामाणिक और उद्देश्यपूर्ण है। प्राथमिक डेटा को मनुष्यों द्वारा बदला या परिवर्तित नहीं किया गया है; इसलिए, इसकी वैधता माध्यमिक डेटा से अधिक है।

द्वितीयक डेटा

द्वितीयक डेटा वह डेटा है जो पहले से ही अन्य स्रोतों से एकत्र किया गया है और आसानी से उपलब्ध है। जब हम अपने उद्देश्य के लिए किसी अन्य उद्देश्य से प्राथमिक डेटा के साथ सांख्यिकीय पद्धति का उपयोग करते हैं तो हम इसे द्वितीयक डेटा के रूप में संदर्भित करते हैं। इसका मतलब है कि एक उद्देश्य प्राथमिक डेटा दूसरा उद्देश्य माध्यमिक डेटा है। तो वह द्वितीयक डेटा वह डेटा है जिसका पुनः उपयोग किया जा रहा है। ऐसे डेटा प्राथमिक डेटा की तुलना में अधिक शीघ्रता से प्राप्त किये जा सकते हैं।

- ये द्वितीयक डेटा कई स्रोतों से प्राप्त किया जा सकता है, जिनमें साहित्य, उद्योग सर्वेक्षण, कम्प्यूटरीकृत डेटाबेस और सूचना प्रणालियों से संकलन, और पर्यावरणीय प्रक्रियाओं के कम्प्यूटरीकृत या गणितीय मॉडल शामिल हैं।

डेटा अधिग्रहण

- डेटा किसी भी शोध अध्ययन का सबसे महत्वपूर्ण और महत्वपूर्ण पहलू है। अध्ययन के विभिन्न क्षेत्रों में किए गए शोधकर्ताओं की कार्यप्रणाली अलग-अलग हो सकती है लेकिन प्रत्येक शोध डेटा पर आधारित होता है जिसका जानकारी प्राप्त करने के लिए विश्लेषण और व्याख्या की जाती है। सांख्यिकीय अध्ययन में डेटा मूल इकाई है। जनगणना, जनसंख्या चर, स्वास्थ्य आँकड़े और सड़क दुर्घटनाओं के रिकॉर्ड जैसी सांख्यिकीय जानकारी सभी डेटा से विकसित की जाती हैं।
- डेटा संग्रह तकनीकों के दो स्रोत हैं। प्राथमिक और द्वितीयक डेटा संग्रह तकनीकें, प्राथमिक डेटा संग्रह सर्वेक्षण, प्रयोग या प्रत्यक्ष अवलोकन का उपयोग करता है।
- द्वितीयक डेटा संग्रह दस्तावेजों के विभिन्न स्रोतों या इलेक्ट्रॉनिक रूप से संग्रहीत जानकारी से जानकारी एकत्र करके आयोजित किया जा सकता है, जनगणना और बाजार अध्ययन माध्यमिक डेटा के सामान्य स्रोतों के उदाहरण हैं। इसे "डेटा माइनिंग" के रूप में भी जाना जाता है।

सर्वेक्षण

- सर्वेक्षण कुछ हद तक सामाजिक विज्ञान, प्रबंधन, विपणन और मनोविज्ञान में सबसे अधिक उपयोग की जाने वाली पद्धति है। सर्वेक्षण विभिन्न तरीकों से किया जा सकता है।

प्रश्नावली

- सर्वेक्षण में प्रश्नावली सबसे अधिक उपयोग की जाने वाली विधि है। प्रश्नावली ओपन-एंडेड या क्लोज-एंडेड प्रश्नों की एक सूची है जिसके लिए उत्तरदाता उत्तर देते हैं। एक प्रश्नावली टेलीफोन, मेल के माध्यम से, सार्वजनिक क्षेत्र में या किसी संस्थान में, इलेक्ट्रॉनिक मेल के माध्यम से या फैक्स और अन्य तरीकों से आयोजित की जा सकती है।

साक्षात्कार

- साक्षात्कार प्रतिवादी के साथ आमने-सामने की बातचीत है। यह धीमा है, महंगा है, और वे लोगों को उनकी नियमित नौकरियों से दूर ले जाते हैं, लेकिन वे गहन पूछताछ और अनुवर्ती प्रश्नों की अनुमति देते हैं।

अवलोकन (Observations)

- अवलोकन करने वाले व्यक्ति को यह बताते हुए कि उसका अवलोकन किया जा रहा है, या उसे बताए बिना भी अवलोकन किया जा सकता है। प्राकृतिक परिवेश के साथ-साथ कृत्रिम रूप से निर्मित वातावरण में भी अवलोकन किया जा सकता है।

प्रकाशित मुद्रित स्रोत (Published Printed Sources)

- प्रकाशित मुद्रित स्रोतों की कई प्रकार की श्रेणियां होती हैं। उनकी विश्वसनीयता कई कारकों पर निर्भर करती है, उदाहरण के लिए, लेखक, प्रकाशन कंपनी और प्रकाशन का समय और तारीख। नई तकनीक और शोध नए तथ्यों को उजागर करते हैं, इसलिए नई स्रोतों को प्राथमिकता दी जाती है और पुराने स्रोतों से बचा जाता है।

पुस्तकें (Books)

- आज किसी भी विषय पर किताबें उपलब्ध हैं जिस पर आप शोध करना चाहते हैं। पुस्तकों का उपयोग उस समय से शुरू होता है जब आप अपने शोध के लिए विषय का चयन भी नहीं करते। विषय के चयन के बाद, पुस्तकें यह जानने में मदद करती हैं कि उस विषय पर पहले से कितना कार्य किया जा चुका है, और आप अपना साहित्य समीक्षा (literature review) तैयार कर सकते हैं। पुस्तकें द्वितीयक स्रोत होती हैं, लेकिन द्वितीयक स्रोतों में सबसे प्रामाणिक मानी जाती हैं।

पत्रिकाएँ/आवधिक पत्रिकाएँ

- डेटा संग्रहण के लिए जर्नल और आवधिक पत्रिकाएं तेजी से महत्वपूर्ण हो रही हैं। इसका कारण यह है कि: पत्रिकाएं अद्यतन (up-to-date) जानकारी प्रदान करते हैं, जो कई बार पुस्तकों से संभव नहीं हो पाती तथा पत्रिकाएं प्र शोध के बहुत विशिष्ट विषयों पर जानकारी दी जाती है, जबकि पुस्तकें सामान्य विषयों पर चर्चा करती हैं।

पत्रिकाएं/समाचारपत्र

- पत्रिकाएँ प्रभावशाली भी हैं लेकिन बहुत विश्वसनीय नहीं। इसके विपरीत, समाचारपत्र अधिक विश्वसनीय होते हैं और कई मामलों में जानकारी केवल समाचारपत्रों से ही प्राप्त की जा सकती है, जैसे राजनीतिक अध्ययन के लिए।

प्रकाशित इलेक्ट्रॉनिक स्रोत (Published Electronic Sources)

- इंटरनेट के बढ़ते विकास, तेजी और अधिक पहुंच के कारण यह देखा गया है कि कई जानकारी जो मुद्रित रूप में उपलब्ध नहीं होती, वह इंटरनेट पर आसानी से उपलब्ध होती है। पहले इंटरनेट की विश्वसनीयता पर सवाल उठाए जाते थे, लेकिन आज यह अधिक विश्वसनीय माना जाता है।
- पहले जर्नल और पुस्तकें शायद ही कभी इंटरनेट पर प्रकाशित होती थीं, लेकिन आज लगभग हर जर्नल और पुस्तक ऑनलाइन उपलब्ध है। कुछ सामग्री मुफ्त होती है और कुछ के लिए भुगतान करना पड़ता है।
- ई-पत्रिकाएँ: ई-पत्रिकाएँ मुद्रित पत्रिकाएँ की तुलना में अधिक आसानी से उपलब्ध होते हैं। नवीनतम पत्रिकाएँ को सब्सक्रिप्शन के बिना प्राप्त करना मुश्किल होता है, लेकिन यदि आपके यूनिवर्सिटी में ई-लाइब्रेरी है, तो आप किसी भी पत्रिका को देख सकते हैं, उसे प्रिंट कर सकते हैं, और जो उपलब्ध नहीं हैं, उनके लिए ऑर्डर कर सकते हैं।
- सामान्य वेबसाइटें :- सामान्य वेबसाइटों पर दी गई जानकारी हमेशा विश्वसनीय नहीं होती, इसलिए वेबसाइट के सामग्री को उद्धृत करने से पहले उसकी विश्वसनीयता की जांच करना चाहिए।
- वेबलॉग्स (Weblogs) :- वेबलॉग्स आजकल आम होते जा रहे हैं। ये विभिन्न लोगों द्वारा लिखी गई डायरियां होती हैं। ये डायरियां व्यक्तिगत रूप से लिखी गई डायरियों जितनी ही विश्वसनीय होती हैं।

डेटा का वर्गीकरण

- डेटा वर्गीकरण एक प्रक्रिया है जिसमें डेटा को विभिन्न श्रेणियों में व्यवस्थित किया जाता है ताकि उसका सबसे प्रभावी और कुशल उपयोग किया जा सके।
- वर्गीकरण डेटा को अलग-अलग वर्गों में व्यवस्थित करने का तरीका है ताकि एकत्रित डेटा को एक निश्चित रूप और सुसंगत संरचना दी जा सके, जिससे उनका सबसे व्यवस्थित और प्रभावी तरीके से उपयोग किया जा सके। यह सुविधाजनक व्याख्या के उद्देश्य से सांख्यिकीय डेटा को विभिन्न समझने योग्य समरूप समूहों के अंतर्गत समूहीकृत करने की प्रक्रिया है।
- डेटा वर्गीकरण के लिए उद्योग में तीन मानक दृष्टिकोण अपनाए जाते हैं:
 - सामग्री-आधारित वर्गीकरण (Content-based Classification)
 - संदर्भ-आधारित वर्गीकरण (Context-based Classification)
 - उपयोगकर्ता-आधारित वर्गीकरण (User-based Classification)

डेटा वर्गीकरण के उद्देश्य

- असमान डेटा को समान विशेषताओं वाले समरूप समूहों में विभाजित करना।
- विभिन्न समूहों के बीच समानता स्थापित करना।
- प्रभावी तुलना को आसान बनाना।
- जटिल, बेतरतीब और बिखरे हुए डेटा को एक संक्षिप्त, तार्किक, समरूप और समझने योग्य रूप में प्रस्तुत करना।
- जटिल डेटा की स्पष्टता और सादगी बनाए रखना।
- स्वतंत्र और आश्रित चर (variables) की पहचान करना और उनके संबंध को स्थापित करना।
- प्रभावी और तार्किक विश्लेषण के लिए विविध डेटा के लिए एक सुसंगत प्रकृति स्थापित करना;
- तार्किक और प्रभावी मात्रा निर्धारण करना
- एक अच्छे वर्गीकरण में स्पष्टता, एकरूपता, पैमाने की समानता, उद्देश्यपूर्णता, सटीकता, स्थिरता, लचीलापन और अस्पष्टता जैसी विशेषताएं होनी चाहिए।

वर्गीकरण की विशेषताएं

एक अच्छे वर्गीकरण में निम्नलिखित गुण होने चाहिए:

- स्पष्टता (Clarity)
- समरूपता (Homogeneity)
- पैमाने की समानता (Equality of Scale)
- उद्देश्यपूर्णता (Purposefulness)
- सटीकता (Accuracy)
- स्थिरता (Stability)
- लचीलापन (Flexibility)
- अस्पष्टता की अनुपस्थिति (Unambiguity)
- वर्गीकरण दो प्रकार का होता है, अर्थात् मात्रात्मक वर्गीकरण, जो चर या मात्रा के आधार पर होता है; और गुणात्मक वर्गीकरण (विशेषताओं के अनुसार वर्गीकरण)। पहला चरों को समूहीकृत करने का तरीका है, जैसे कि सामंजस्यपूर्ण समूहों में चरों की मात्रा निर्धारित करना, जबकि बाद वाला डेटा को विशेषताओं या गुणों के आधार पर समूहित करता है। फिर, यह एकाधिक वर्गीकरण या द्विभाजित वर्गीकरण हो सकता है।
- पहला किसी गुणवत्ता या विशेषताओं के आधार पर कई (दो से अधिक) समूह बनाने का तरीका है, जबकि बाद वाला एक निश्चित गुणवत्ता की उपस्थिति या अनुपस्थिति के आधार पर दो समूहों में वर्गीकरण है।
- डेटा वर्गीकरण, सूचना सुरक्षा के संदर्भ में, संवेदनशीलता के स्तर के आधार पर डेटा का वर्गीकरण है और विश्वविद्यालय पर प्रभाव यह होना चाहिए कि डेटा को प्राधिकरण के बिना प्रकट किया जाए, बदला जाए या नष्ट किया जाए। डेटा का वर्गीकरण यह निर्धारित करने में मदद करता है कि उस डेटा को सुरक्षित रखने के लिए कौन से आधारभूत सुरक्षा नियंत्रण उपयुक्त हैं। सभी संस्थागत आंकड़ों को तीन संवेदनशीलता स्तरों या वर्गीकरणों में से एक में वर्गीकृत किया जाना चाहिए:

A.	प्रतिबंधित डाटा
	डेटा को प्रतिबंधित के रूप में वर्गीकृत किया जाना चाहिए जब उसका अनधिकृत प्रकटीकरण, संशोधन या विनाश विश्वविद्यालय या इसके सहयोगियों के लिए उच्च स्तर का जोखिम उत्पन्न कर सकता है। प्रतिबंधित डेटा के उदाहरणों में राज्य या संघीय गोपनीयता नियमों द्वारा संरक्षित डेटा और गोपनीयता समझौतों द्वारा संरक्षित डेटा शामिल हैं। प्रतिबंधित डेटा पर उच्चतम स्तर का सुरक्षा नियंत्रण लागू किया जाना चाहिए।
B.	निजी डाटा
	डेटा को निजी के रूप में वर्गीकृत किया जाना चाहिए जब उसका अनधिकृत प्रकटीकरण, संशोधन या विनाश विश्वविद्यालय या उसके सहयोगियों के लिए मध्यम स्तर का जोखिम उत्पन्न कर सकता है। डिफॉल्ट रूप से, ऐसा संस्थागत डेटा जो स्पष्ट रूप से प्रतिबंधित या

	सार्वजनिक के रूप में वर्गीकृत नहीं है, उसे निजी डेटा माना जाना चाहिए। निजी डेटा की सुरक्षा के लिए उचित स्तर के सुरक्षा नियंत्रण लागू किए जाने चाहिए।
C.	सार्वजनिक डाटा
	डेटा को सार्वजनिक के रूप में वर्गीकृत किया जाना चाहिए जब उस डेटा के अनधिकृत प्रकटीकरण, परिवर्तन या विनाश से विश्वविद्यालय और उसके सहयोगियों को बहुत कम या कोई जोखिम न हो। सार्वजनिक डेटा के उदाहरणों में प्रेस विज्ञप्ति, पाठ्यक्रम जानकारी और शोध प्रकाशन शामिल हैं। हालाँकि सार्वजनिक डेटा की गोपनीयता की सुरक्षा के लिए बहुत कम या कोई नियंत्रण की आवश्यकता नहीं है, लेकिन सार्वजनिक डेटा के अनधिकृत संशोधन या विनाश को रोकने के लिए कुछ स्तर के नियंत्रण की आवश्यकता है।



अध्याय – 2

मात्रात्मक और गुणात्मक डेटा

पिछले वर्ष के प्रश्न

प्रश्न 1 निम्नलिखित शब्दों को अमूर्तता के बढ़ते क्रम में व्यवस्थित करें: (2023)	
(a) मोबाइल टेलीफोन	(b) एंड्रॉइड फोन
(c) संचार उपकरण	(d) टेलीफोन
नीचे दिए गए विकल्पों में से सही उत्तर चुनें:	
(A) (b), (a), (d), (c)	(B) (b), (d), (a), (c)
(C) (c), (d), (a), (b)	(D) (c), (a), (b), (d)

गुणात्मक डेटा (Qualitative Data)

गुणात्मक डेटा विश्लेषण को एक शब्द में समझाया जा सकता है – **श्रेणीबद्ध**। इसमें डेटा को संख्यात्मक मानों या पैटर्न के माध्यम से वर्णित नहीं किया जाता है, बल्कि वर्णनात्मक संदर्भ (जैसे, पाठ) के माध्यम से परिभाषित किया जाता है। गुणात्मक डेटा आमतौर पर व्यक्ति-से-व्यक्ति तकनीकों का उपयोग करके एकत्र किया जाता है। इन तकनीकों में शामिल हैं:

- **अवलोकन (Observations):** अवलोकन समूह में होने वाले व्यवहारिक पैटर्न का विवरण किया जाता है। इस पैटर्न में किसी गतिविधि में बिताया गया समय, गतिविधि का प्रकार और उपयोग की गई संचार विधि शामिल हो सकता है।
- **फोकस समूह (Focus Groups):** लोगों को समूहों में संगठित कर उनके साथ किसी शोध के विषय पर सहयोगात्मक विचार-विमर्श उत्पन्न करके प्रासंगिक प्रश्नों के माध्यम से चर्चा की जाती है।
- **द्वितीयक शोध (Secondary Research):** जिस तरह से व्यवहार के पैटर्न का अवलोकन किया जाता है, उसी तरह विभिन्न प्रकार के दस्तावेज़ीकरण स्रोतों को उनके सामग्री प्रकार के आधार पर कोड और वर्गीकृत किया जा सकता है।
- **साक्षात्कार (Interviews):** यह वर्णनात्मक डेटा एकत्र करने के लिए सबसे प्रभावी विधियों में से एक है। साक्षात्कार के उत्तरों को विषय, विषयवस्तु या श्रेणियों के आधार पर वर्गीकृत किया जा सकता है। यह दृष्टिकोण अत्यधिक केंद्रित डेटा खंडन (segmentation) की सुविधा प्रदान करता है।

मात्रात्मक डेटा (Quantitative Data)

यदि मात्रात्मक डेटा विश्लेषण को एक शब्द में समझाया जाए (जो वास्तव में संभव नहीं है), तो वह शब्द होगा – **संख्यात्मक**। डेटा विश्लेषण में कुछ निश्चितताएं होती हैं, लेकिन यह निश्चित है कि यदि आपके शोध में संख्याएं शामिल नहीं हैं, तो वह मात्रात्मक

शोध नहीं है। जब डेटा विश्लेषण की बात आती है तो कुछ निश्चितताएं होती हैं, लेकिन आप निश्चित हो सकते हैं कि यदि आप जिस शोध में संलग्न हैं, उसमें कोई संख्या शामिल नहीं है, तो यह मात्रात्मक शोध नहीं है। मात्रात्मक विश्लेषण उन प्रक्रियाओं के एक समूह को संदर्भित करता है जिसके द्वारा संख्यात्मक डेटा का विश्लेषण किया जाता है। अधिकांशतः, इसमें मानक विचलन, माध्य और माधिका जैसे सांख्यिकीय मॉडलिंग का उपयोग शामिल होता है।

- **माध्य:** माध्य प्रतिक्रियाओं के एक सेट के लिए एक संख्यात्मक औसत का प्रतिनिधित्व करता है। डेटा सेट (या एकाधिक डेटा सेट) के साथ काम करते समय, माध्य संख्याओं के एक विशिष्ट सेट के केंद्रीय मान का प्रतिनिधित्व करेगा। यह डेटा सेट के भीतर मूल्यों की संख्या से विभाजित मूल्यों का योग है। अवधारणा का वर्णन करने के लिए जिन अन्य शब्दों का उपयोग किया जा सकता है वे हैं अंकगणितीय माध्य, औसत और गणितीय अपेक्षा।
- **मानक विचलन (Standard Deviation):** यह मात्रात्मक विश्लेषण में अक्सर उपयोग किया जाने वाला सांख्यिकीय शब्द है। मानक विचलन माध्य के आसपास प्रतिक्रियाओं के वितरण को दिखाता है। यह उत्तरों में सुसंगतता की डिग्री को परिभाषित करता है। औसत के साथ मिलकर यह डेटा सेट की गहरी जानकारी प्रदान करता है।
- **आवृत्ति वितरण (Frequency Distribution):** यह एक माप है जो डेटा सेट में उत्तर की उपस्थिति की दर को मापता है। उदाहरण के लिए, सर्वेक्षण का उपयोग करते समय, आवृत्ति वितरण में एक विशिष्ट क्रमिक पैमाने पर प्रतिक्रिया प्रकट होने की संख्या निर्धारित करने की क्षमता होती है (यानी, सहमत, दृढ़ता से सहमत, असहमत, आदि)। आवृत्ति वितरण डेटा बिंदुओं के बीच सहमति की डिग्री का निर्धारण करने में महत्वपूर्ण भूमिका निभाता है।

गुणात्मक और मात्रात्मक के बीच का अंतर नीचे दिया गया है।

गुणात्मक (Qualitative)	मात्रात्मक (Quantitative)
मानव और सामाजिक विज्ञान को समझने के लिए उपयोगी।	वैज्ञानिक और प्रायोगिक शोध विधियों से संख्यात्मक डेटा उत्पन्न करता है।
यह विधि समग्रता (Holistic) पर आधारित है।	मात्रात्मक शोध विशेष और केंद्रित होता है।
इसमें डेटा का संश्लेषण करने के लिए तर्कशक्ति आगमनात्मक (Inductive) होती है।	इसमें तर्कशक्ति निष्कर्षात्मक (Deductive) होती है।
यह सृजनात्मक (Inventor) होती है।	यह महत्वपूर्ण (Crucial) होती है।
यह उद्देश्यपूर्ण सैंपलिंग पर आधारित होती है, जिसमें छोटे समूह या नमूना आकार चुना जाता है।	यह रैंडम सैंपलिंग पर आधारित होती है, जिसमें बड़ी प्रतिनिधिक सैंपलिंग पूरी जनसंख्या के लिए उपयोग की जाती है।
इसमें मौखिक डेटा (Oral Data) एकत्र किया जाता है।	इसमें मापने योग्य डेटा एकत्र किया जाता है।
यह प्रक्रिया-उन्मुख (Process-Oriented) होती है।	मात्रात्मक विधि में विशिष्ट और मापने योग्य परिणामों पर ध्यान केंद्रित किया जाता है।



अध्याय – 3 ग्राफिकल प्रतिनिधित्व

डेटा का दृश्यात्मक प्रतिनिधित्व

उदाहरण के लिए, डेटा को कई तरीकों से क्रमबद्ध और प्रस्तुत किया जा सकता है

1. तालिका
2. चित्रलेख
3. बार चार्ट
4. हिस्टोग्राम
5. पाइ चार्ट
6. रेखा ग्राफ

कभी-कभी डेटा एक से अधिक टेबल, पाइ चार्ट आदि में हो सकता है। उनका उद्देश्य केवल मात्रात्मक सेल परीक्षण नहीं है। बल्कि तुलनात्मक और विश्लेषणात्मक कौशल का पता लगाना भी है।

टेबल के मुख्य भाग

एक तालिका में निम्नलिखित मुख्य भाग होने चाहिए

1. **तालिका का शीर्षक (Title of the Table):**
प्रत्येक तालिका का एक स्पष्ट और उचित शीर्षक होना आवश्यक है। यह बताता है कि तालिका में कौन-से आंकड़े हैं, वे किस समय के हैं और किस स्थान से संबंधित हैं। शीर्षक को संक्षिप्त, स्पष्ट और कुछ स्थानों पर व्याख्यात्मक होना चाहिए। शीर्षक में आकर्षकता भी होनी चाहिए ताकि पाठक का ध्यान तुरंत तालिका की ओर जाए और उसे समझने में समय और प्रयास न लगे।
2. **तालिका संख्या (Table No.):**
प्रत्येक तालिका की शुरुआत में उसे एक संख्या दी जानी चाहिए। तालिका संख्या से किसी विशेष तालिका को ढूँढना आसान हो जाता है। यदि तालिकाओं की संख्या अधिक हो, तो उन्हें उचित तरीके से क्रमांकित किया जाना चाहिए। आमतौर पर, तालिका संख्या को शीर्षक के ऊपर और केंद्र में रखा जाता है।

तालिका संख्या "शीर्षक"

रेखा उपशीर्षक	मुख्य स्तंभ शीर्षक			कुल
	उपशीर्षक	उपशीर्षक	उपशीर्षक	
उप-प्रविष्टियां		मध्य भाग		
कुल				

बार चार्ट

बार चार्ट एक ग्राफ है जो डेटा की श्रेणियों के बीच तुलना दिखाने के लिए बार का उपयोग करता है। बार क्षैतिज या लंबवत हो सकते हैं। ग्राफ में 2 बिंदु हैं। एक अक्ष (axis) डेटा की तुलना की जा रही श्रेणियों का वर्णन करता है और दूसरा अक्ष डेटा के संख्यात्मक मानों को दर्शाता है। जो डेटा के मूल्यों का प्रतिनिधित्व करते हैं, इससे कोई फर्क नहीं पड़ता कि कौन सी अक्ष या अक्ष हैं। लेकिन यह निर्धारित करेगा कि बार ग्राफ में क्या दिखाया गया है। यदि विवरण क्षैतिज अक्ष पर है, तो बाहरी भाग ऊर्ध्वाधर रूप से उन्मुख होगा। यदि मान क्षैतिज अक्ष के साथ हैं, तो बार क्षैतिज रूप से उन्मुख होंगे।

3. उपशीर्षक प्रत्येक (Subtitle):

एक तालिका में कई स्तंभ होते हैं। इन स्तंभों के शीर्ष को उपशीर्षक या कैप्शन कहा जाता है। उपशीर्षक को स्तंभ के केंद्र में रखा जाना चाहिए। एक उपशीर्षक के तहत कई शीर्षक हो सकते हैं। जब विभिन्न स्तंभों में शब्दों को माप की विभिन्न इकाइयों में मापा जाता है, तो संबंधित इकाई को उपशीर्षक से हटा दिया जाता है।

4. पंक्ति शीर्षक (Row Title):

पंक्तियों के शीर्षक को पंक्ति शीर्षक या लाइन शीर्षक कहा जाता है। इसे तालिका में दो बार दिया जाता है।

5. तालिका का मुख्य भाग (Table Cover):

यह तालिका का सबसे मुख्य और महत्वपूर्ण भाग है। इसका आकार और स्वरूप डेटा के आधार पर पहले से तय किया जाना चाहिए। इस भाग में डेटा को उपशीर्षकों और पंक्ति शीर्षकों के अनुसार व्यवस्थित किया जाता है।

6. रेखांकन और स्थान (Underline and Spacing):

तालिका में स्थान छोड़ना और रेखाएं बनाना एक महत्वपूर्ण हिस्सा है। उचित स्थान और रेखाओं के उपयोग से तालिका अधिक आकर्षक और प्रभावी बनती है। यह तालिका को स्पष्ट और समझने में आसान बनाता है।

7. पाद टिप्पणियां (Foot Note):

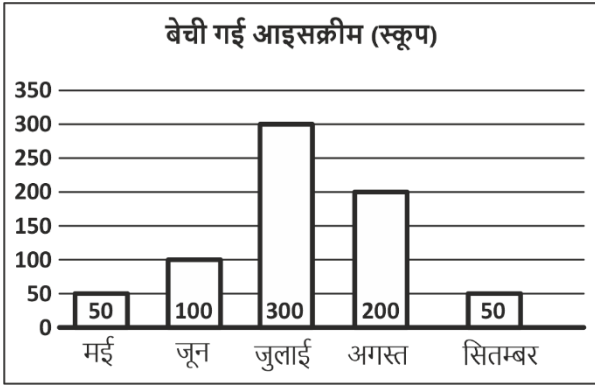
कभी-कभी तालिका में दिए गए आंकड़ों या शब्दों की व्याख्या के लिए टिप्पणियां आवश्यक हो सकती हैं। इन्हें तालिका के नीचे दिया जाता है। हालांकि, कोशिश की जानी चाहिए कि टिप्पणियों का उपयोग कम से कम हो।

बार चार्ट के प्रकार

बार चार्ट या बार ग्राफ के कई प्रकार होते हैं, और वे हमेशा एक-दूसरे के स्थान पर उपयोग नहीं किए जा सकते। किस प्रकार के बार ग्राफ का उपयोग करना है, यह इस बात पर निर्भर करता है कि हम किस चीज़ की तुलना करना चाहते हैं। पहले हम कुछ सरल बार ग्राफ पर चर्चा करेंगे।

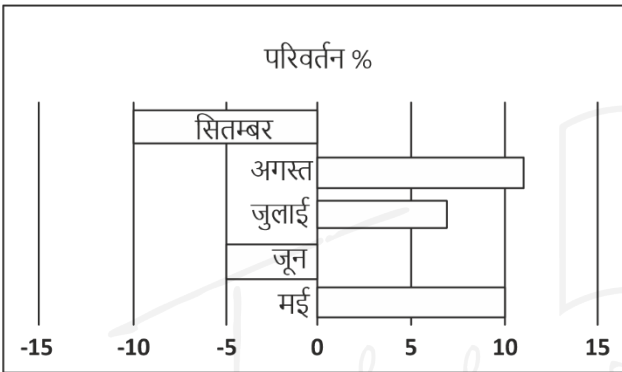
वर्टिकल या लंबवत बार चार्ट

सरल वर्टिकल बार ग्राफ तब सबसे अच्छा होता है, जब आपको दो या दो से अधिक स्वतंत्र चर (variables) के बीच तुलना नहीं करनी हो। प्रत्येक चर एक निश्चित मान से संबंधित होगा और इसलिए इसे क्षैतिज मान के लिए निर्धारित किया जा सकता है।

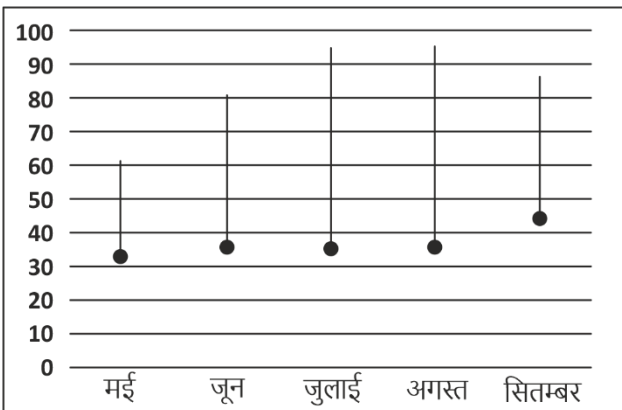


क्षैतिज बार चार्ट

यदि आपके डेटा में नकारात्मक और सकारात्मक मान हैं, लेकिन फिर भी दो या दो से अधिक निश्चित स्वतंत्र चर के बीच तुलना है, तो क्षैतिज ग्राफ़ रखना उचित है। ऊर्ध्वाधर अक्ष क्षैतिज अक्ष के मध्य में उन्मुख हो सकता है। किसका ऋणात्मक एवं धनात्मक मान प्रदर्शित किया जा सकता है?



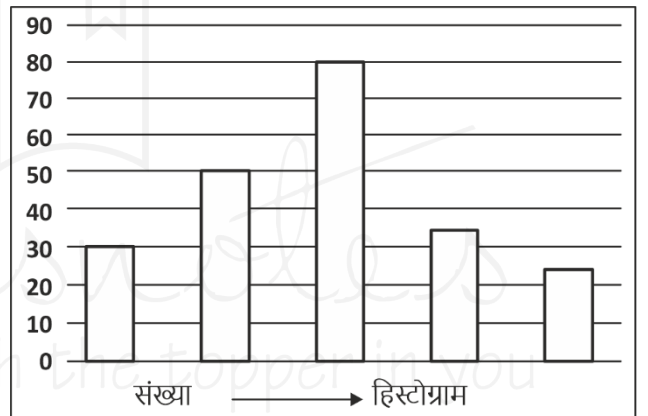
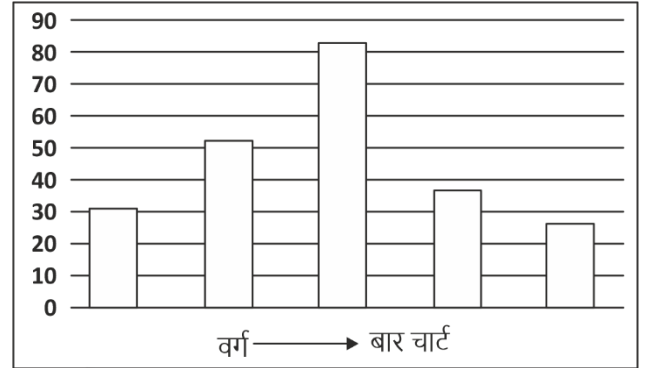
एक रेंज बार चार्ट प्रत्येक स्वतंत्र चर के लिए डेटा की एक श्रेणी का प्रतिनिधित्व करता है। तापमान सीमाएँ या मान सीमाएँ श्रेणी ग्राफ़ के लिए डेटा के सामान्य सेट होते हैं। उपरोक्त ग्राफ़ के विपरीत, डेटा एक सामान्य श्रवण बिंदु से प्रारंभ नहीं होता है। लेकिन वह विशेष बिंदु डेटा की सीमा के लिए कम संख्या से शुरू होता है। एक रेंज बार ग्राफ़ क्षैतिज या ऊर्ध्वाधर हो सकता है।



हिस्टोग्राम डेटा

- हिस्टोग्राम सांख्यिकीय जानकारी को दिखाने का एक तरीका है। यह दो चर के बीच औसत का वर्णन करता है। यह संख्यात्मक या डेटा वितरण को समान भागों में ग्राफ़ के रूप में प्रदर्शित करता है। यह एक सतत चर की प्राथमिक

वितरण (फ्रिक्वेंसी डिस्ट्रिब्यूशन) को दर्शाता है। इसे सबसे पहले कार्ल प्रियसन ने प्रस्तावित किया था। हिस्टोग्राम तालिका में दिखाए गए आवृत्तियों (फ्रिक्वेंसी) को दर्शाता है। इसे आयतों के रूप में प्रदर्शित किया जाता है, जो क्रमबद्ध अंतराल (इंटरवल) में बंटा होता है। प्रत्येक आयत का क्षेत्रफल संबंधित अंतराल की आवृत्ति के समानुपाती होता है। आयत की ऊंचाई भी संबंधित अंतराल की आवृत्ति घनत्व (फ्रिक्वेंसी डेंसिटी) के समानुपाती होती है, अर्थात् आकार और अंतराल चौड़ाई के अनुपात को दिखाती है।



- हिस्टोग्राम में कुल क्षेत्रफल आंकड़ों की कुल संख्या के बराबर होता है। एक हिस्टोग्राम को संबंधित आवृत्तियों के साथ सामान्यीकृत (ओवरनॉर्मलाइज) करके भी दिखाया जा सकता है। यदि हम सामान्य रूप के हिस्टोग्राम की बात करें, तो इसमें क्षैतिज रेखा (horizontal line) पर स्वतंत्र या स्वतंत्र चर (independent variable) और लंबवत रेखा (vertical line) पर आश्रित या आश्रित चर (dependent variable) दिखाए जाते हैं। इन आंकड़ों को अलग-अलग रंगों में या किसी कवर किए गए क्षेत्र के रूप में दर्शाया जाता है।
- हिस्टोग्राम बार चार्ट (bar chart) से बहुत मिलता-जुलता होता है। लेकिन इसमें चर (variable) निरंतर (continuous) होते हैं। हिस्टोग्राम और साधारण बार ग्राफ़ के बीच का मुख्य अंतर यह है कि हिस्टोग्राम में प्रत्येक बार (bar) एक श्रृंखला (series) या श्रेणी (category) के आश्रित चर को दर्शाता है, न कि केवल एक बिंदु (single dotted point) को।

डेटा मैपिंग

- डेटा मैपिंग वह प्रक्रिया है जिसमें किसी स्रोत फाइल (source file) के डेटा फील्ड्स को उनके संबंधित लक्षित फील्ड्स (target fields) से जोड़ा जाता है। आवश्यक डेटा तक पहुंच किसी संगठन को अधिक सफल बना सकती है। किसी तरह, डेटा को समझना आसान हो जाता है, जब इसे चित्रात्मक (विजुअल) रूप में दिखाया जाए।
- चित्रात्मक डेटा (visual data) लोगों को यह समझने में मदद करता है कि विभिन्न अवधारणाएं (concepts) कैसे उत्पन्न होती हैं और वे एक-दूसरे से कैसे संबंधित हैं। डेटा मैपिंग इसमें मदद करती है।
उदाहरण के लिए, Excel में 'Name,' 'Email,' और 'Phone' जैसे फील्ड्स को एक डिलीमिटेड फाइल (delimited file) के संबंधित फील्ड्स से मैप किया जाता है, जो हमारा गंतव्य (destination) है।
- डेटा मैपिंग संगठनों को प्रक्रियाओं और कार्यों के बीच लिंक दिखाने में मदद करती है कि कौन-सा कार्य कैसे उपयोग में लाया जाए। हमारा 40% तंत्रिका तंतुओं (nerve fibres) का कनेक्शन मस्तिष्क से केवल रेटिना में होता है। डेटा मैपिंग यह दिखाने में मदद करती है कि डेटा के विभिन्न हिस्से उपयोगी और सहायक क्यों हैं।
- डेटा मैपिंग से ग्राहक रुझानों (customer trends) का रियल टाइम में पता लगाया जा सकता है। यह रुझानों के कारणों का विश्लेषण करने, पुराने डेटा और अन्य गणनाओं को समझने में मदद करती है। इसके अलावा, हम डेटा मैपिंग सॉफ्टवेयर का उपयोग करके अपने डेटा की तुलना प्रतिस्पर्धियों के डेटा से भी कर सकते हैं। सही डेटा मैपिंग का उपयोग करने पर आपका व्यवसाय तेजी से बढ़ सकता है।
- डेटा मैपिंग बड़े पैमाने पर मैप्स बनाने के लिए भी काम करती है। किसी संगठन के Salesforce में एक शक्तिशाली डेटा मैपिंग सॉफ्टवेयर प्रोग्राम होता है, जिसे उपयोग में लाया जा सकता है। यह रियल टाइम (real-time) में भी मदद करता है। हम क्लाउड नेटवर्क से कनेक्ट होकर रियल टाइम में जानकारी प्राप्त कर सकते हैं।
- डेटा मैपिंग सभी व्यवसायों के लिए उपयोगी होती है। उदाहरण के लिए, यदि हम खुदरा क्षेत्र (retail sector) में हैं, तो हम डेटा मैपिंग का उपयोग यह गणना करने के लिए कर सकते हैं कि छूट (discount sales) हमारी कुल बिक्री (sales totals) को कैसे प्रभावित करती है। इसी तरह, वित्तीय निर्णय (financial decisions) और निवेश (investment) से जुड़े फैसले भी डेटा मैपिंग की सहायता से किए जा सकते हैं।
- डेटा चाहे आंतरिक (internal) हो या बाहरी (external), यह लगातार अधिक बिखरा हुआ (dispersed) और विशाल (voluminous) हो रहा है। ऐसे में डेटा का प्रभावी उपयोग और उससे जुड़े उपयोगी निष्कर्ष (actionable insights) निकालना बहुत महत्वपूर्ण है।

- सामान्यतः, डेटा मैपिंग निम्नलिखित गतिविधियों में मदद करती है:
 - **डेटा इंटीग्रेशन (Data Integration):**
डेटा स्रोत और लक्ष्य (source and destination) की स्कीमा में अंतर को कवर करने के लिए डेटा मैपिंग उपकरण उपयोग किए जाते हैं। यह व्यवसायों को अलग-अलग डेटा बिंदुओं से जानकारी को आसानी से समेकित (consolidate) करने में मदद करता है।
 - **डेटा माइग्रेशन (Data Migration):**
डेटा को एक डेटाबेस से दूसरे डेटाबेस में ले जाना। इस प्रक्रिया को सफलतापूर्वक पूरा करने के लिए कोड-फ्री डेटा मैपिंग समाधान का उपयोग करना महत्वपूर्ण है, जो प्रक्रिया को स्वचालित (automate) कर सके।
 - **डेटा वेयरहाउसिंग (Data Warehousing):**
डेटा वेयरहाउस में डेटा मैपिंग का मतलब स्रोत (source) और लक्ष्य (target) टेबल या गुणों (attributes) के बीच कनेक्शन बनाना है।
 - **डेटा ट्रांसफॉर्मेशन (Data Transformation):**
डेटा ट्रांसफॉर्मेशन में डेटा मैपिंग पहला कदम है। यह जानकारी के साइलो (silos) को तोड़ने और उपयोगी निष्कर्ष निकालने के लिए आवश्यक है।
 - **डेटा मैपिंग की तकनीकें (Data Mapping Techniques):**
डेटा प्रबंधन प्रक्रिया में डेटा मैपिंग एक महत्वपूर्ण कदम है, लेकिन यह जटिल और समय लेने वाला हो सकता है। स्वचालन (automation) के स्तर के आधार पर, डेटा मैपिंग तकनीकों को दो प्रकारों में विभाजित किया जा सकता है:
 1. **मैन्युअल डेटा मैपिंग (Manual Data Mapping):**
यह प्रक्रिया हाथ से कोडिंग (hand-coded) द्वारा की जाती है और असीमित लचीलापन (unlimited flexibility) प्रदान करती है।
 2. **अर्ध-स्वचालित डेटा मैपिंग (Semi-Automated Data Mapping):**
स्कीमा मैपिंग को सेमी-ऑटोमेटेड डेटा मैपिंग तकनीक के रूप में वर्गीकृत किया जाता है। इसमें दो डेटा ऑब्जेक्ट्स की पहचान की जाती है, जो अर्थपूर्ण रूप से संबंधित (semantically related) होते हैं, और फिर उनके बीच मैपिंग बनाई जाती है।
- ## आंकड़ों या डेटा की व्याख्या
- आंकड़ों की व्याख्या (Data Interpretation) का मतलब है जानकारी को समझने के लिए उसका विश्लेषण करना और उसके निहितार्थ (implications) और महत्व को निर्धारित करना। यह उन प्रक्रियाओं को लागू करने को संदर्भित करता है, जिनके माध्यम से आंकड़ों की समीक्षा की जाती है ताकि एक सूचित (informed) निष्कर्ष तक पहुंचा जा सके। इसके तुलनात्मक माप के लिए एक अच्छा निर्णय लिया जाना चाहिए। इसमें निम्नलिखित पैमाने शामिल हैं।

-
- 1. नाममात्र पैमाना** यह एक गैर-संख्यात्मक (non-numerical) पैमाना है। इसे न तो क्रमबद्ध किया जा सकता है और न ही मात्रात्मक रूप से तुलना की जा सकती है। इसमें चर (variable) विशिष्ट और संपूर्ण (specific and exhaustive) होते हैं।
 - 2. कर्मसूचक संख्या या क्रमिक पैमाना :** इसमें ऐसी श्रेणियां (categories) होती हैं, जो संपूर्ण (wholes) होती हैं और तार्किक मांग (logical demand) के साथ व्यवस्थित होती हैं। उदाहरण: गुणवत्ता रेटिंग (Good, Very Good, Fair) या सहमति रेटिंग (Agree, Strongly Agree)।

- 3. अंतराल पैमाना :** यहां आंकड़ों को समान दूरी (equal distance) के आधार पर श्रेणियों में विभाजित और समूहित किया जाता है। इसमें हमेशा एक मनमाना (arbitrary) 0 बिंदु होता है।
- 4. अनुपात** इसमें उपरोक्त तीनों विशेषताएं शामिल हैं डेटा की व्याख्या करते समय एक विश्लेषण को कुछ अन्य कारकों के अलावा सहसंबंध कारक और संयोग आदि के बीच अंतर को समझने की कोशिश करनी चाहिए।



अध्याय-4 डेटा और शासन



- आज के तेज़ी से बदलते और प्रतिस्पर्धात्मक व्यावसायिक वातावरण में डेटा गवर्नेंस की आवश्यकता है। अब, जब संगठनों के पास आंतरिक और बाहरी डेटा के बड़े और विविध भंडार को संग्रहित करने का अवसर है, तो उन्हें इसे अधिकतम मूल्य देने, जोखिमों को प्रबंधित करने और लागत को कम करने के लिए एक अनुशासन की जरूरत होती है।
- डेटा गवर्नेंस प्रक्रियाओं, भूमिकाओं, नीतियों, मानकों और मापदंडों (metrics) का एक समूह है, जो यह सुनिश्चित करता है कि जानकारी का उपयोग प्रभावी और कुशल तरीके से हो, ताकि संगठन अपने लक्ष्यों को प्राप्त कर सके। डेटा गवर्नेंस परिभाषित करता है कि कौन, किस डेटा पर, किस परिस्थिति में, किस विधि का उपयोग करके, कौन सा कार्य कर सकता है।
- डेटा गवर्नेंस सुनिश्चित करता है कि डेटा से संबंधित भूमिकाएं स्पष्ट रूप से परिभाषित हैं और संगठन में जिम्मेदारी और जवाबदेही (accountability) पर सहमति है। एक अच्छी तरह से नियोजित डेटा गवर्नेंस ढांचा (framework) रणनीतिक, व्यावहारिक (tactical) और संचालनात्मक (operational) भूमिकाओं और जिम्मेदारियों को कवर करता है।
- डेटा और गवर्नेंस रणनीति तैयार करते समय, हमें सावधान रहने की आवश्यकता है।
- **डेटा गवर्नेंस डेटा प्रबंधन नहीं है:** डेटा प्रबंधन से तात्पर्य किसी संगठन की पूर्ण डेटा जीवनचक्र आवश्यकताओं के प्रबंधन से है। **डेटा गवर्नेंस** डेटा प्रबंधन का एक प्रमुख घटक (core component) है, जैसे डेटा वेयरहाउसिंग।
 1. **डेटा गवर्नेंस मास्टर डेटा प्रबंधन नहीं है:** मास्टर डेटा प्रबंधन किसी संगठन की प्रमुख संस्थाओं की पहचान करने और फिर इस डेटा की गुणवत्ता में सुधार करने पर ध्यान केंद्रित करता है।
 2. **डेटा गवर्नेंस डेटा प्रबंधन नहीं है:** डेटा प्रबंधक डेटा संपत्तियों की देखभाल करते हैं, यह सुनिश्चित करते हुए कि वास्तविक डेटा डेटा गवर्नेंस योजना के अनुरूप है, अन्य डेटा संपत्तियों से जुड़ा हुआ है और डेटा गुणवत्ता, अनुपालन या सुरक्षा के संदर्भ में नियंत्रण में है।

डेटा गवर्नेंस के लाभ

एक प्रभावी डेटा प्रशासन रणनीति एक संगठन को कई लाभ प्रदान करती है, जहां इसमें निम्नलिखित शामिल हैं

- डेटा की एक सामान्य समझ है।
 - डेटा की गुणवत्ता में सुधार हुआ है, जैसे डेटा सटीकता, पूर्णता और निरंतरता।
 - डाटा मानचित्र उपलब्ध है।
 - **समग्र दृष्टिकोण:** प्रत्येक ग्राहक और अन्य व्यावसायिक संस्थाओं का 360-डिग्री दृश्य मूल रूप से सत्य का एक संस्करण है।
 - **निरंतर अनुपालन:** डेटा गवर्नेंस सरकारी नियमों की मांगों को पूरा करने के लिए एक मंच प्रदान करता है।
 - **बेहतर डेटा प्रबंधन:** यह मानवीय आयाम को एक अत्यधिक स्वचालित और डेटा-संचालित दुनिया में लाता है। हम इसके लिए प्रौद्योगिकी को समर्थकारी के रूप में उपयोग कर सकते हैं।
- डेटा गवर्नेंस टूल के लिए ओपन सोर्स और क्लाउड बुनियादी रणनीतियाँ हैं। IPaaS भी उनके साथ निकटता से जुड़ा हुआ है। ये उपकरण हमें निम्नलिखित प्राप्त करने में भी मदद करते हैं।
1. **हमारे डेटा को कैप्चर और समझें,**
 2. **अपने डेटा की गुणवत्ता में सुधार करें.**
 3. **डेटा प्रबंधित करना:** मेटाडेटा-संचालित ईटीएल और ईएलटी, और डेटा एकीकरण अनुप्रयोगों के साथ।
 4. **डेटा को नियंत्रित करना।**
 5. **हमारे डेटा का दस्तावेजीकरण करें।**
 6. **उन लोगों को सशक्त बनाएँ जो डेटा को सबसे अच्छी तरह जानते हैं:** डेटा प्रबंधन में योगदान देना।
 7. **संवेदनशील डेटा की सुरक्षा करना**
- हमें यह समझने की आवश्यकता है कि डेटा गवर्नेंस वैकल्पिक नहीं है। 'डेटा लेक' के रूप में जाने जाने वाले कार्यान्वयन के लिए आवश्यक रूप से ऐसी प्रक्रियाओं की आवश्यकता होती है जो आपको आवश्यक डेटा को इस तरह से रखने की अनुमति देती है जो तकनीकी बाधाओं को दूर करती है और उस डेटा को संसाधित करने के लिए नई क्षमताएं प्रदान करती है।

Chapter-5

डेटा इंटरप्रिटेशन

- Q. 1:** निम्न तालिका में एक कंपनी द्वारा लगातार दो वर्षों 2019 और 2020 में निर्मित पाँच विभिन्न प्रकार की कारों (A-E) का प्रतिशत (%) वितरण दर्शाया गया है। वर्ष 2019 में कुल 4,50,000 कारें और वर्ष 2020 में कुल 5,20,000 कारें निर्मित की गई थीं। तालिका में दिए गए आंकड़ों के आधार पर प्रश्न 1-5 के उत्तर दीजिए: (2022)

कारों के उत्पादन का वर्षवार प्रतिशत वितरण।

वर्ष	उत्पादित कार का प्रकार (% में)				
	A	B	C	D	E
2019	15%	25%	30%	10%	20%
2020	10%	30%	25%	25%	10%

- Q. 2** यदि 2020 में निर्मित Type-A कारों की संख्या 2019 के बराबर होती, तो 2020 के कुल उत्पादन में इसका अनुमानित प्रतिशत हिस्सा कितना होता? (2022)
- (A) 11% (B) 13%
(C) 15% (D) 9%
- Q. 3** यदि 2019 और 2020 के दौरान उत्पादित टाइप-ई कारों में से 85% कंपनी द्वारा बेची जाती हैं, तो कंपनी द्वारा कितनी टाइप-ई कारें बिना बिकी बची हैं? (2022)
- (A) 21825 (B) 29100
(C) 25200 (D) 21300
- Q. 4** यदि 2020 में टाइप-बी कारों का प्रतिशत उत्पादन 2019 के समान था, तो 2020 में उत्पादित टाइप-बी कारों की संख्या कितनी रही होगी? (2022)
- (A) 112500 (B) 120000
(C) 130000 (D) 135000
- Q. 5** 2019 में उत्पादित टाइप-सी कारों की संख्या का 2020 में उत्पादित टाइप-डी कारों की संख्या से अनुपात क्या है? (2022)
- (A) 29:27 (B) 23:22
(C) 27:26 (D) 27:23
- Q. 6** 2019 और 2020 के बीच टाइप-सी कारों की संख्या में क्या अंतर आया? (2022)
- (A) 5000 (B) 7500
(C) 8500 (D) 2500
- Q. 7** प्रश्नों के लिए निर्देश (18 से 22): तालिका में डेटा के आधार पर, निम्नलिखित पांच प्रश्नों के उत्तर दें: (2022)
निम्नलिखित तालिका वर्ष 2015 से 2021 के दौरान एक कंपनी द्वारा बेची गई पांच विभिन्न प्रकार की बैटरियों (हजारों में) की बिक्री के आंकड़े देती है।
बैटरियों की वर्षवार बिक्री

वर्ष	बेची गई बैटरियों के प्रकार (हजारों में)				
	4AH	7AH	32AH	35AH	55AH
2015	75	144	114	102	108
2016	90	126	102	84	126
2017	96	114	75	105	135
2018	105	90	150	90	75
2019	90	75	135	75	90
2020	105	60	165	45	120
2021	115	85	160	100	145

- Q.8** 2015 की तुलना में 2021 में 55AH बैटरी की बिक्री में अनुमानित प्रतिशत वृद्धि क्या थी? (2022)
 (A) 28% (B) 33%
 (C) 31% (D) 34%
- Q.9** सभी सात वर्षों की कुल बिक्री किस प्रकार की बैटरी की अधिकतम है? (2022)
 (A) 4AH (B) 32AH
 (C) 7AH (D) 35AH
- Q.10** 2016 और 2020 में बेची गई 35AH बैटरियों की संख्या में क्या अंतर है? (2022)
 (A) 24000 (B) 35000
 (C) 28000 (D) 39000
- Q.11** किसी दिए गए वर्ष में बेची गई 4AH बैटरियों का प्रतिशत उसी वर्ष बेची गई बैटरियों की कुल संख्या से अधिकतम है। (2022)
 (A) 2017 (B) 2019
 (C) 2018 (D) 2020
- Q.12** किस प्रकार की बैटरी की बिक्री में 2015 से 2020 तक लगातार कमी आई है? (2022)
 (A) 4AH (B) 7AH
 (C) 32AH (D) 35AH
- Q.13** निम्नलिखित तालिका छह मित्रों ए-एफ की मासिक आय और विभिन्न व्यय को पूर्ण मूल्य या प्रतिशत (मासिक आय के संदर्भ में) के रूप में दर्शाती है। तालिका में कुछ मान ('-' के रूप में चिह्नित) गायब हैं जिनकी आपसे आवश्यकता पड़ने पर गणना करने की अपेक्षा की जाती है। तालिका में डेटा के आधार पर, प्रश्न का उत्तर दें: (2023)
 आय एवं व्यय विवरण

दोस्त	मासिक आय		व्यय (₹ में)			
	वेतन (₹ में)	प्रोत्साहन (₹ में)	यात्रा	खाना	आवास	खरीदारी
A	92000	-	-	10960	10%	15%
B	-	14400	15280	17000	12400	-
C	-	12600	12%	8%	-	12%
D	88000	-	-	15120	9%	16800
E	80000	-	5%	-	8400	11240
F	-	11400	8400	8%	-	13720

नोट:

- (A) मासिक आय = वेतन + प्रोत्साहन।
 (B) प्रोत्साहन राशि वेतन का 15% है।
 (C) सभी छह दोस्त अपनी मासिक आय का 40% बचाते हैं।
 (D) तालिका में दिए गए व्यय के अलावा कोई अन्य व्यय नहीं है।

- Q.14** सभी छह दोस्तों द्वारा यात्रा पर एक साथ खर्च की गई कुल राशि (₹ में) क्या है? (2023)
 (A) 85634 (B) 85643
 (C) 84634 (D) 84436
- Q.15** B द्वारा भोजन और खरीदारी पर खर्च की गई राशि और C द्वारा आवास पर खर्च की गई राशि में क्या अंतर है? (2023)
 (A) ₹ 11,412 (B) ₹ 11,532
 (C) ₹ 11,512 (D) ₹ 11,332
- Q.16** सभी छह दोस्तों द्वारा एक साथ बचाई गई राशि (₹ में) क्या है? (2023)
 (A) 223760 (B) 237360
 (C) 237630 (D) 273360

- Q.17** D की वार्षिक आय A की तुलना में ₹ कम है। (2023)
 (A) 4600 (B) 52200
 (C) 55200 (D) 4900

- Q. 18** A द्वारा यात्रा पर किया गया व्यय F के वेतन का लगभग ____% है। (2023)
 (A) 30 (B) 38
 (C) 32 (D) 34

- Q.19** निम्नलिखित तालिका छह शहरों अर्थात् मुंबई, दिल्ली, कोलकाता, चेन्नई, हैदराबाद और चंडीगढ़ के मल्टीप्लेक्स में एक सप्ताह में बेची गई पांच फिल्मों ए-ई के टिकटों की संख्या से संबंधित डेटा प्रदान करती है। तालिका में दिए गए आंकड़ों के आधार पर निम्नलिखित प्रश्नों के उत्तर दीजिए: (2023)

शहर	फिल्म				
	A	B	C	D	E
मुंबई	10000	7500	17500	13000	9000
दिल्ली	8500	9500	10500	12500	14000
कोलकाता	16000	12000	9500	10500	8500
चेन्नई	9000	10500	16000	14000	17000
हैदराबाद	8000	17000	13000	14500	11000
चंडीगढ़	7500	13500	10000	17500	13000

- Q.20** हैदराबाद में बेची गई फिल्म B की टिकटों की संख्या सभी शहरों में बेची गई एक ही फिल्म की कुल टिकटों की संख्या का लगभग ____% थी? (2023)
 (A) 15 (B) 12
 (C) 20 (D) 24

- Q.21** कोलकाता में बेची गई फिल्म D की टिकटों की संख्या चंडीगढ़ में बेची गई फिल्म B की टिकटों की संख्या से लगभग ____% कम है। (2023)
 (A) 22 (B) 21
 (C) 23 (D) 19

- Q.22** सभी छह शहरों में मिलाकर बेची गई फिल्म C की टिकटों की औसत संख्या कितनी है? (2023)
 (A) 12000 (B) 7750
 (C) 12750 (D) 9275

- Q.23** चेन्नई में बेची गई फिल्म E की टिकटों की संख्या मुंबई में बेची गई फिल्म A की टिकटों की संख्या का ____% है। (2023)
 (A) 70 (B) 170
 (C) 130 (D) 30

- Q.24** सभी पांच फिल्मों को मिलाकर बेची गई टिकटों की कुल संख्या शहर में दूसरी न्यूनतम थी (2023)
 (A) दिल्ली (B) मुंबई
 (C) चंडीगढ़ (D) कोलकाता

- Q.25** निम्नलिखित तालिका 2018-2022 तक पांच वर्षों के दौरान एक निश्चित कंपनी की आय और व्यय (लाखों में) को दर्शाती है, साथ ही इन सभी वर्षों में लाभ% और हानि% भी दर्शाती है। तालिका में कुछ मान गायब हैं ('-' द्वारा दर्शाया गया है) जिनकी आवश्यकता पढ़ने पर आपसे गणना करने की अपेक्षा की जाती है। तालिका में दिए गए आंकड़ों के आधार पर निम्नलिखित प्रश्नों के उत्तर दीजिए।
 किसी कंपनी की वर्ष-वार आय, व्यय और लाभ/हानि (2023)

वर्ष	व्यय (मिलियन में)	आय (लाख में)	लाभ %(+)/हानि %(-)
2018	192	-	+15%
2019	248	-	-
2020	210	-	+ 20%
2021	-	359.6	-
2022	-	-	- 12%

Q.26 2020 में लाभ 2018 की तुलना में % अधिक है: (2023)
 (A) 56.76 (B) 44.98
 (C) 45.83 (D) 59.82

प्रश्न 27 2018 और 2020 में कंपनी की आय का औसत (मिलियन में) क्या है? (2023)
 (A) 220.8 (B) 236.4
 (C) 252.2 (D) 246.6

Q.28 2022 में, यदि कंपनी का व्यय 2019 की तुलना में 15% अधिक था, तो 2022 में कंपनी की आय (मिलियन में) थी: (2023)
 (A) 250.976 (B) 248.976
 (C) 250.936 (D) 230.936

Q.29 2021 में, यदि कंपनी का लाभ प्रतिशत (%) 2018 में लाभ प्रतिशत (%) से 60% अधिक था, तो पिछले वर्ष की तुलना में 2021 में कंपनी के व्यय में प्रतिशत वृद्धि (लगभग) होगी: (2023)
 (A) 46% (B) 45%
 (C) 41% (D) 38%

Q.30 यदि 2019 में कंपनी को हुआ घाटा 2018 में कंपनी के लाभ से 33.2 मिलियन अधिक था, तो 2019 में कंपनी को हुआ घाटा प्रतिशत था: (2023)
 (A) 17% (B) 22%
 (C) 25% (D) 23%

Q.31 निम्नलिखित तालिका एक विश्वविद्यालय में एमसीए कार्यक्रम के छह अलग-अलग विषयों ए-एफ में छह अलग-अलग छात्रों द्वारा प्राप्त अंकों का प्रतिशत दर्शाती है, साथ ही प्रत्येक विषय में अधिकतम अंक (कोष्ठक में दिखाया गया है)। तालिका में दिए गए आंकड़ों के आधार पर निम्नलिखित प्रश्नों के उत्तर दीजिए। (2023)
 छह विषयों में छात्र-वार अंकों का प्रतिशत

वर्ष	विषय (अधिकतम अंक)					
	A (150)	B (100)	C (150)	D (125)	E (75)	F (50)
Anil	66%	75%	88%	56%	56%	90%
Arman	82%	76%	84%	96%	92%	88%
Arpit	76%	66%	78%	88%	72%	70%
Gaurav	90%	88%	96%	76%	84%	86%
Gurjit	64%	70%	68%	72%	68%	74%
Priya	48%	56%	50%	64%	64%	58%

प्रश्न 32 अनिल को सभी छह विषयों में मिलाकर कितने अंक मिले? (2023)
 (A) 369 (B) 463
 (C) 558 (D) 496

प्रश्न 33 सभी छह विषयों में प्राप्त कुल अंक _____ छात्रों द्वारा 490 से अधिक हैं। (2023)
 (A) दो (B) तीन
 (C) चार (D) पाँच