

Research Data (शोध आंकड़े)

डाटा का अर्थ -

शोध अध्ययन में क्षेत्रीय या प्रलेखीय आधार पर शोधकर्ता जो डाटा संकलित या एकत्रित करता है वह डाटा (Data) कहलाता है ।

डाटा की परिभाषा -

‘डाटा’ शब्द लैटिन भाषा से उत्पन्न हुआ है जिसका अर्थ है जो कुछ भी दिया जाये (Any thing that is given) इसकी विभिन्न परिभाषा इस प्रकार है :-

- वैबस्टर्स थर्ड न्यू इंटरनेशनल डिक्शनरी (Webster's Third New International Dictionary) ने डाटा इस प्रकार परिभाषित किया है ‘कुछ भी प्रदत्त या स्वीकृत, प्रस्तुत या अनुमत तथ्य अथवा सिद्धान्त, जिसके ऊपर एक अनुमान अथवा तर्क आधारित होता है जिसके आधार पर किसी भी प्रकार की आदर्श प्रणाली का निर्माण किया जाता है ।’

- ऑक्सफोर्ड एनसाइक्लोपीडिया इंग्लिश डिक्शनरी (Oxford Encyclopaedia English Disctionary)के अनुसार - 'निष्कर्षों अथवा अनुमान के लिये आधार के रूप में प्रयुक्त ज्ञात तथ्य एवं वस्तुएँ'
- यूनेस्को (UNESCO) के अनुसार - 'मानव या ऑटोमेटिक माध्यमों के द्वारा सम्प्रेषण, निर्वचन अथवा प्रक्रियाकरण के लिये उपयुक्त औपचारिक रूप में तथ्य, अवधारणा अथवा अनुदेश'
- कोडेटा (Committee on Data for Science and Technology) के अनुसार - अधिकतम परिशुद्ध रूप में वैज्ञानिक ज्ञान के सार का पारदर्शक प्रस्तुतिकरण इस परिभाषा के अनुसार स्पष्टता एवं परिशुद्धता आंकड़े की दो आवश्यक विशेषताएँ हैं ।

डाटा के प्रकार -

सबसे पहले उन्हे विषयों के आधार पर श्रेणीबद्ध कर सकते हैं, जैसे-विज्ञान, सामाजिक विज्ञान, इन विषयों में डाटा को भिन्न-भिन्न प्रकार से विभाजित किया जा सकता है ।

विज्ञान में डाटा के प्रकार

ने 1975 में आंकड़ों को श्रेणीबद्ध करने की आवश्यकता उस समय महसूस की जब वे डाटा उपलब्धता की समस्या के प्रतिवेदन पर कार्य कर रहे थे । इस कार्य समूह ने एक पद्धति विकसित की, जिसके अनुसार वैज्ञानिक आंकड़ों की निम्नलिखित श्रेणियां बनाई जा सकती हैं :-

समय पक्ष के प्रसंग में आंकड़े -

इस समय पक्ष के आधार पर डाटा को दो प्रकारों में वर्गीकृत कर सकते हैं -

समय स्वतंत्र आंकड़े - यह पद उन आंकड़ों से संबंधित है । जिनका मापन पुनः-पुनः किया जा सकता है । उदाहरण के लिये भू-विज्ञान और खगोल विज्ञान (Geo-sciences & Astronomy)से संबंधित आंकड़े ।

समय आश्रित आंकड़े- इसका मापन एक बार ही किया जाता है । जैसे भूभौतिकीय एवं अंतरिक्षीय (Cosmological) घटनाएँ । जैसे ज्वालामुखी का फूटना

आदि । इसी प्रकार दुर्लभ जीवाश्मों से संबंधित आंकड़े भी इसी श्रेणी में आते हैं ।

स्थान पक्ष के प्रसंग में आंकड़े -

स्थान पक्ष के प्रसंग में आंकड़ों को निम्नलिखित श्रेणियों में विभाजित किया जा सकता है ।

स्थान स्वतंत्र आंकड़े - ये वे आंकड़े हैं जो मापन की जा रही वस्तु या पदार्थ की स्थिति से स्वतंत्र होते हैं ।

उदाहरण के लिये शुद्ध भौतिक और रसायन विज्ञान के आंकड़े ।

स्थान आश्रित आंकड़े - ये आंकड़े मापन की जा रही वस्तु या पदार्थ की स्थिति पर आश्रित होते हैं । इस श्रेणी में भू-विज्ञान, खगोल विज्ञान से संबंधित आंकड़े आते हैं इसके साथ-साथ चट्टानों से संबंधित आंकड़े भी स्थान आश्रित आंकड़े हैं ।

उत्पत्ति की विधि के प्रसंग में आंकड़े -

ये तीन प्रकार के होते हैं -

प्राथमिक आंकड़े - ये आंकड़े मापन के लिये बनाये गये प्रयोगों अथवा अवलोकनों के द्वारा प्राप्त किये जाते हैं ।

व्युत्पन्न आंकड़े - ये आंकड़े अनेक प्राथमिक आंकड़ों के सैद्धांतिक प्रतिरूप की सहायता से जोड़े जाते हैं ।

सैद्धांतिक आंकड़े- ये आंकड़े सैद्धांतिक गणनाओं से प्राप्त होते हैं । उदाहरण के लिये सूर्यग्रहण से संबंधित आंकड़े ।

संख्यात्मक मूल्य की प्रकृति के प्रसंग में आंकड़े -

इस श्रेणी के आंकड़ों को दो भागों में विभाजित कर सकते हैं ।

निश्चय आंकड़े - निश्चय आंकड़े वे हैं जो निश्चित परिस्थितियों में सुस्पष्ट मूल्य प्राप्त करने हेतु कल्पना की मात्रा से संबंधित होते हैं ।

स्टॉकैस्टिक आंकड़े - मात्रा से संबंधित आंकड़ा जो कि निश्चित परिस्थिति में एक नमूने से दूसरे नमूने में, एक मापन से दूसरे मापन में स्थिर मूल्य देता है वह स्टॉकैस्टिक आंकड़ा कहलाता है । भूविज्ञान (Geoscience) का अस्टिकांश आंकड़ा इसी श्रेणी में आता है ।

अभिव्यक्ति के संबंध के प्रसंग में आंकड़े -

इस श्रेणी में तीन प्रकार के आंकड़े आते हैं जो इस प्रकार हैं ।

मात्रात्मक आंकड़े - भौतिक विज्ञान का अधिकांश आंकड़ा मात्रात्मक आंकड़ा होता है ।

अर्द्ध-मात्रात्मक आंकड़े - ये आंकड़े नकारात्मक और सकारात्मक उत्तरों से संबंधित होते हैं, जो कि पदों में सम्मिलित विभिन्न विशेषताओं से संबंधित पूछे गये प्रश्नों से सामने आते हैं। उदाहरण के लिये जीव विज्ञान में प्राणियों का वर्गीकरण, जातियों की आकारकीय, जैव रसायन और अन्य विशेषताओं से संबंधित प्रश्नों के उत्तर पर आधारित होता है। 'हां' या 'ना' में होते हैं। इन आंकड़ों को अर्द्ध-मात्रात्मक आंकड़े कहते हैं। 'हां' और 'ना' को 1 और 0 के रूप में कोट किया जाता है, जिससे कि संख्यात्मक आंकड़े प्राप्त होते हैं।

गुणात्मक आंकड़े - जब संबंधित वैज्ञानिक पदार्थ परिभाषात्मक कथनों के रूप में व्यक्त किये जाते हैं तो वे प्रकृति से गुणात्मक होते हैं।

प्रस्तुतिकरण की विधि के प्रसंग में आंकड़े -

ये आंकड़े तीन प्रकार के होते हैं :-

संख्यात्मक आंकड़े - ये आंकड़े संख्यात्मक मूल्य के रूप में व्यक्त किये जाते हैं जैसे अधिकांश मात्रात्मक आंकड़े।

रेखचित्रीय आंकड़े - यहां पर आंकड़े रेखचित्रीय स्वरूप में प्रस्तुत किये जाते हैं। कुछ प्रसंगों में रेखा चित्रों को इसलिए बनाया जाता है ताकि पाठक बड़ी संख्या में

आंकड़ों को दृश्य अनुभूति के द्वारा समझ सके । चार्ट और मानचित्र इसी श्रेणी में आते हैं ।

प्रतीकात्मक आंकड़े- इस तरह के आंकड़ों के प्रतीकात्मक स्वरूपों में प्रस्तुत किया जा सकता है । जिससे दर्शक को भलीभांति समझ सके । उदाहरण के लिये मौसम से संबंधित आंकड़ों का प्रतीकात्मक प्रस्तुतीकरण ।

सामाजिक विज्ञान में आंकड़ों के प्रकार

सामाजिक विज्ञान में भी आंकड़ों को विभिन्न प्रकार से विभाजित किया है जो इस प्रकार है ।

मापन के पैमाने के प्रसंग में आंकड़े

मापन के पैमाने के आंकड़ों को निम्नलिखित श्रेणियों में विभाजित किया जा सकता है

सांकेतिक आंकड़े - वैयक्तिक इकाईयों की पहचान के रूप में संख्याओं को निर्धारित करने के लिये सांकेतिक पैमाने का उपयोग किया जाता है । उदाहरण के लिये

विषयानुसार पुस्तकों का वर्गीकरण । इसमें श्रेणी को बताने के लिये संख्या निर्धारित कर दी जाती है । क्योंकि संख्या श्रेणी के मात्र नाम का प्रतिनिधित्व करती है ।

क्रमसूचक आंकड़े - यह किये गये अवलोकनों के लिये नियत की गई संख्याओं के मध्य आपसी संबंध के क्रम का संकेत करता है । उदाहरण के लिये ग्रंथालय

कर्मचारियों के व्यवहार के अध्ययन के लिये शोधार्थी निम्न अनुसार क्रम निर्धारित कर सकता है । 1. खराब को संकेत करने हेतु 2. साधारण को संकेत करने हेतु 3. अच्छे को संकेत करने हेतु 4. बहुत अच्छे को संकेत करने हेतु । यहां 1, 2, 3, 4, संख्याएँ क्रमसूचक आंकड़े हैं ये इस बात का संकेत करते हैं कि संख्या 4, 3 से बेहतर है और 3, 2, से बेहतर है आदि यहां यह ध्यान रखना चाहिये कि क्रमसूचक आंकड़े भिन्नता की दिशा बताते हैं । ये भिन्नता की सही मात्रा का वर्णन नहीं करते ।

अंतराल आंकड़े - अंतराल आंकड़े, आंकड़ों एवं समान मापन की विभिन्न श्रेणियों के मध्य अंतर की श्रेणियां हैं । उदाहरण के लिये बच्चों के समूह प्राप्तांकों का मापन कर सकते हैं । प्रत्येक बच्चे के प्राप्तांक का संख्यात्मक मान निर्धारित करने के बाद इन आंकड़ों को 10 के अंतराल के समूह में रखा जाता है । जैसे - 0-10, 10-20, 20-30 और 30-40 आदि आदि ।

अनुपात आंकड़े - ये आंकड़े चर का परिमाण/मात्रात्मक के अर्थ में मात्रात्मक मापन है उदाहरण के लिये वजन, ऊंचाई, दूरी, मूल्य आदि ।

निरन्तरता के प्रसंग में आंकड़े :-

इस श्रेणी के आंकड़े निम्न प्रकार के हैं ।

सतत् आंकड़े - सतत् आंकड़े संभावित मूल्यों का अपरिमित सेट हैं। व्यक्तियों के मध्य अपरिमित संभावित मूल्य होते हैं। उदाहरण के लिये किसी व्यक्ति का ऊंचाई का कथन एक निश्चित मूल्य जैसे 160 से.मी. और उसके बाद 161 से.मी. में नहीं किया जा सकता। यह 160.59 अथवा 160.89 से.मी. या 161.59 से.मी. में एक सतत क्रम में हो सकता है।

विवेचित आंकड़े - विवेचित आंकड़े परिमित अथवा गणना योग्य मूल्य/मात्रा में होते हैं। उदाहरण के लिये गं्र्थालय की सदस्य संख्या। यह 2525 अथवा 2561 हो सकती है लेकिन 2561.8 नहीं हो सकती।

विशेषताओं की संख्या के प्रसंग में आंकड़े - इसे निम्नलिखित श्रेणियों में विभाजित किया जा सकता है

एक चर आंकड़े - एक चर आंकड़े उस समय प्राप्त होते हैं जब अवलोकन के लिये एक विशेषता का उपयोग किया जाता है।

द्विचर आंकड़े - द्विचर आंकड़े उस समय प्राप्त होते हैं जब एक विशेषता के स्थान पर, एक ही समय पर दो विशेषताओं का मापन किया जाता है। उदाहरण के लिये कक्षा 12 वीं के छात्रों की ऊंचाई और वजन।

बहुचर आंकड़े - जब अवलोकन तीन या उससे अधिक विशेषताओं को ध्यान में रखकर किया जाता है तो उससे प्राप्त आंकड़ों को बहुचर आंकड़े कहते हैं जैसे छात्रों की योग्यता, वजन, ऊँचाई ।

समय के प्रसंग में आंकड़े -

समय के प्रसंग में आंकड़ों को दो भागों में बांटा गया है-

समय-श्रृंखला आंकड़े - जो आंकड़े समय के आस पास कालक्रमानुसार अभिलिखित किये जाते हैं उन्हे

समय-श्रृंखला आंकड़ा कहते हैं । उदाहरण के लिये विभिन्न वर्षों में ग्रंथालय में अर्जित की गई

पाठ्य-सामग्री की संख्या, महाविद्यालय में प्रवेश किये गये छात्रों की वार्षिक आधार पर संख्या ।

प्रतिनिध्यात्मक आंकड़े - ये आंकड़े एक ही इकाई या विभिन्न इकाइयों से एक ही समय बिन्दु से संबंधित होते हैं ।

उत्पत्ति के प्रसंग में आंकड़े -

इस श्रेणी में आंकड़े निम्न प्रकार के होते हैं :-

प्राथमिक आंकड़े - प्राथमिक आंकड़े वे आंकड़े हैं, जिन्हे शोधार्थी के द्वारा प्रत्यक्ष अवलोकन, गणना मापन अथवा साक्षात्मकार, प्रश्नावली के द्वारा एकत्रित किया जाता है । उदाहरण के लिये पाठक सर्वेक्षण के द्वारा प्राप्त आंकड़े ।

द्वितीयक आंकड़े - ये आंकड़े प्राथमिक उद्देश्य से एकत्रित किये गये थे और किसी प्रतिवेदन में प्रकाशित हो चुके हैं। बाद में इन आंकड़ों का उपयोग किसी अन्य उद्देश्य के लिये किया जाता है। तो इसे द्वितीयक आंकड़े कहते हैं उदाहरण के लिये जनसंख्या प्रतिवेदन, ग्रंथ आदि के प्राप्त आंकड़े या संकलित आंकड़े।

विशेषता के प्रसंग में आंकड़े -

इसी स्वयं की विशेषता के आधार पर निम्नलिखित भागों में विभाजित किया गया है।

मात्रात्मक आंकड़े - जब अवलोकनों की विशेषताओं की मात्राओं का निर्धारण किया जाता है तो हमें मात्रात्मक आंकड़े मिलते हैं। मात्रात्मक आंकड़े प्रयुक्त विशेषताओं के मात्रा के मापन के द्वारा प्राप्त होते हैं। उदाहरण के लिये वस्तु का मूल्य एवं वजन आदि।

गुणात्मक आंकड़े - जब अवलोकन की विशेषता गुण होते हैं तो हमें गुणात्मक आंकड़े मिलते हैं। उदाहरण के लिये व्यक्ति का रंग अथवा बुद्धिमत्ता। आंकड़ों की प्रकृति को, जिस वर्ग से वह संबंध रखता है, उसके आधार पर समझा जा सकता है। विज्ञान में 6 मूलभूत प्रकार के आंकड़े मिलते हैं और जिन्हे 15 विभिन्न वर्गों में विभाजित किया गया है। मात्रात्मक, संख्यात्मक आंकड़े होते हैं और अधिकांश आंकड़े मात्रात्मक होते हैं।

इसी तरह सामाजिक विज्ञानों में भी आंकड़ों की प्रकृति का निर्धारण संभव है सामाजिक विज्ञानों में आंकड़ों की प्रकृति या तो वर्णात्मक होती है या गणनात्मक । जब भी गणनात्मक रूप से आंकड़े प्राप्त होते हैं तो उनकी प्रकृति गणनात्मक होती है और जो आंकड़े घटना या वस्तु का वर्णन करते हैं उनकी प्रकृति वर्णात्मक होती है ।

डाटा का क्षेत्र -

डाटा का क्षेत्र काफी विस्तृत होता है और विविध विषयों के अंतर्गत शोध अध्ययनों में इसका विशेष महत्व है ।

डाटा के अभाव में सूचना एवं ज्ञान की उत्पत्ति संभव नहीं है । डाटा के क्षेत्र का अध्ययन निम्नलिखित दृष्टिकोणों के आधार पर किया जा सकता है :-

डाटा की उपयोगिता - ज्ञान के विकास में डाटा की विशेष उपयोगिता होती है । डाटा के अभाव में कोई भी शोध, अन्वेषण, निरीक्षण, परीक्षण आदि संभव नहीं हो सकता । उपयुक्त डाटा के उपयोग के अभाव में निर्णयन प्रणाली कार्य नहीं कर सकती । किसी समस्या का समाधान नहीं हो सकता और किसी भी नियोजन को कार्य रूप नहीं दिया जा सकता ।

डाटा का विस्तार - अध्ययन विषय क्षेत्र से संबंधित डाटा का विस्तार डाटा के तत्व, डाटा बैंक, सर्वेक्षण विधि, साक्षात्कार, अवलोकन विधि आदि सम्मिलित होते हैं ।

डाटा का काल- किसी भी शोध समस्या के लिये डाटा का संग्रह करते समय उसके समय को अवश्य इंगित करना चाहिये । यह स्पष्ट रूप से बताना चाहिये कि डाटा सामयिक या संचयी ।