**Assignment-01  
  
Topic- DATA:**

**According to the Oxford *“Data is distinct pieces of information, usually formatted in a special way”*.**

* Data is a collection of facts, numbers, or observations that can be used to learn about something
* e.g. Data can be used as the basis for calculation, or discussion in any organization or the individual.
* It can be of two types either Primary or Secondary.

**DATA SOURCES:**

**Primary Data:** Collected first-hand from sources.

* E.g. Surveys (door-to-door interviews, online forms)
* Student theses and research projects
* Personal interviews (PI)

**Secondary Data**: Gathered from existing sources.

* Internet databases, reports, research papers
* Government statistics and industry publications
* Books and journals

**Qualitative Data (Categorical Data/ Non- Numeric):**

* Describes characteristics, qualities, or attributes. It focuses on meanings, descriptions, and experiences rather than numbers.
* e.g., customer feedback, colour preferences, behavioural traits**, etc*.***

**Quantitative Data (Numerical Data)**

* This is **numerical data** that can be measured, counted, and analysed mathematically.
* It is used to represent quantities, trends, and patterns in research, statistics, and business analysis
* e.g., height, length, temperatures, test scores, number of users, sales in a month, etc.

**Why Data is important?**

* Data helps in making better decisions.
* Data helps solve problems by finding the reason for underperformance.
* Data helps one to evaluate the performance.
* Data helps one improve processes.
* Data helps one understand consumers and the market**.**

## **WHAT IS INFORMATION?**

* Information is data that has been processed, organized, or structured in a way that makes it meaningful, valuable and useful.
* It gives knowledge, understanding and insights that can be used for decision-making, problem-solving, communication and various other purposes.

**Data vs. Information:**

* Raw Data: Unprocessed numbers, symbols, or text (e.g., 34, 56, "rainy").
* Information: Processed data that conveys meaning (e.g., "The average temperature for the month is 28°C")

**Importance and Applications of Data:**

* **Insurance:** Risk assessment and fraud detection.
* **Tax Authorities:** Identifying tax payers, and monitoring financial transactions.
* **Business & Marketing:** Customer behaviour analysis, trend forecasting.
* **Healthcare:** Disease tracking, patient records management.
* **Government:** Policy-making, Plans the better cities, roads, and policies, etc.

**Data Collection Methods:**

**Steps for Data Collection:**

1. **Define Objective:** Decide what you want to find out and why?
2. **Identify Data Source:** Choose from where to collect the data
3. **Design Data Collection Method:** Pick how you'll collect data (surveys, experiments, or observations)
4. **Collect Data:** Gather the information in a structured way
5. **Record & Organize Data:** Store data in a proper format (Excel, database) for easy analysis

1. **Geographical Data**:

This data is information about locations, places, and features on Earth. It helps us understand where things are and how they relate to each other.

* Examples:
* GPS coordinates – Helps in navigation (like Google Maps).
* Maps & satellite images – Shows cities, forests, and landscapes.
* Land use patterns – Tracks how land is used (farms, cities, forests)
* Applications:
* **City Planning** – Helps design roads, buildings, and public spaces
* **Disaster Management** – Predicts floods, earthquakes, and helps in rescue operations.
* **Climate & Environment** – Tracks temperature, pollution, and deforestation.  
    
  *Geographical data makes life easier by helping us plan, protect, and understand the world around us*

1. **Cultural Data**: This data captures information related to traditions, languages, social behaviours, and human interactions.

* Examples:
* Linguistic patterns – What languages people speak and how they change over time.
* Festival and tradition records – Information about cultural celebrations and rituals.
* Population migration trends – How and why people move from one place to another.
* Applications:
* Cultural Research – Helps preserve traditions and study social changes
* Global Business – Companies use it to understand different markets and customer preferences.
* Policy Making – Governments use it to create policies for diverse populations.

*Cultural data helps connect people, preserve heritage, and guide decisions in a globalized world!*

**3. Scientific Data**:This Data is collected from experiments, observations, and research findings in various scientific fields.

* Examples:
* Laboratory experiment results– Data from testing chemicals, reactions, or physics experiments
* Chemical compositions– Information about elements and compounds in substances.
* Medical trial data– Research on new medicines and treatments.
* Applications:
* Healthcare & Medicine – Helps develop new drugs and treatments.
* Leads to innovations like AI, robotics, and materials science.
* Environmental studies– Tracks pollution, climate change, and ecosystem health.

**4. Financial Data**: Financial data consists of monetary transactions, investments, economic indicators, and business financial records.

* Examples:
* Stock market prices
* Company revenue and expenditures
* Inflation and GDP statistics
* Applications:
* Investment analysis and risk management
* Financial fraud detection
* Economic policy development

**5. Statistical**: Statistical data involves numerical information derived from surveys and research studies.

* Examples:
* Population census reports
* Consumer survey results
* Industry growth rates
* Applications:
* Data-driven policy-making
* Business and market research
* Predictive analytics in various fields

**6. Meteorological Data:** Meteorological data pertains to atmospheric conditions, climate patterns, and weather forecasting.

* Examples:
* Temperature and humidity levels
* Rainfall and snowfall records
* Wind speed and direction
* Applications:
* Weather prediction and forecasting
* Climate change analysis
* Disaster preparedness and mitigation

**7. Natural Data**: Includes environmental and ecological observations gathered from nature.

* Examples:
* Wildlife population records
* Forest and vegetation mapping
* Water quality reports
* Applications:
* Biodiversity conservation efforts
* Sustainable resource management
* Environmental impact assessments

**8. Transport Data**: This data provides information about movement patterns, logistics, and infrastructure usage.

* Examples:
* Traffic flow and congestion reports
* Public transport schedules
* Shipping and logistics data
* Applications:
* Urban mobility and transportation planning
* Supply chain optimization
* Traffic control and safety measures

**9. Demographic Data**: This data refers to information about population characteristics and social statistics.

* Examples:
* Age, gender, and ethnicity distribution
* Employment and income levels
* Educational background
* Applications:
* Targeted marketing and consumer insights
* Public health and welfare policies
* Workforce planning and development

**10. Behavioral Data**: This captures human actions, interactions, and decision-making patterns.

* Examples:
* Online browsing history
* Purchase behaviors
* Social media engagement
* Applications:
* Personalized marketing strategies
* AI and machine learning model training
* Cybersecurity threat detection

**11. Sensor Data**: This data is collected from IoT devices and smart systems for real-time monitoring and automation.

* Examples:
* Temperature and pressure readings from industrial sensors
* Heart rate and step count from smartwatches
* Traffic sensor data from smart cities
* Applications:
* Predictive maintenance in industries
* Smart home automation
* Real-time health monitoring systems

**Conclusion**

Data is a powerful tool across multiple industries, providing insights that drive informed decision-making. Effective data collection methods ensure reliability and accuracy, making it essential to define clear research objectives and use appropriate source

**Assignment-02**

* ***Collect real data***

1. ***Define Objective:***

* *In this Assignment I have collected* ***“Cyber Security Incident Data.”***
* *The objective of this assignment is focus on identifying, analysing and mitigating cyber threats that also include following objectives:*

1. *Identify Cyber Threats*
2. *Assess the impact of these threats*
3. *Improve response strategies*
4. *Enhance threat detection and prevention*
5. *Regulatory compliance and reporting*
6. **Identifying data sources**: Here I have Used secondary data sources.  
   * + The data is source out data from various news articles, blogs and research reports that are available on internet.
7. **Designing Data collection methods**: Here I have collected data manually from various sources, which includes:

* Incident reports
* Case studies
* Searching relevant news articles

1. **Data collection**: Whatever data we have collected is given below*:*

# **Cyber Incidents Report:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Incident Name** | **Date** | **Affected Users/Data** | **Impact** |
| 1 | BharatPay Hacked | 2022 | 37,000 users | Personal data exposure |
| 2 | Swachhta Platform Hacked | 2022 | 16 million users | Data sold on the Dark Web |
| 3 | Cyberattack on AIIMS | 2022 | 1.3 TB data | Healthcare infrastructure attack |
| 4 | RailYatri Data Breach | 2022 | 30 million users | User records leaked online |
| 5 | CloudSEK Data Breach | 2022 | Confidential company data | Cybersecurity firm targeted |
| 6 | Zivame Data Breach | 2022 | 1.5 million customers | Personal data sold online |
| 7 | Motilal Oswal Cyber Incident | 2023 | Employee systems affected | Extortion tactics used |
| 8 | Polycab Ransomware Attack | 2023 | IT infrastructure targeted | Operations unaffected |
| 9 | Sun Pharma Cyber Attack | 2023 | Operational disruptions | Potential healthcare data risk |
| 10 | MoChhatua Data Breach | 2023 | Ration distribution data | User credentials leaked |
| 11 | Cyberabad Police Data Leak | 2023 | 66.9 crore records | Financial and personal data exposed |
| 12 | Rentomojo Cyber Attack | 2023 | User personal data | Cloud misconfiguration issue |
| 13 | SPARSH Data Breach | 2023 | Defence pension details | Data leaked on dark web |
| 14 | Hathway ISP Data Breach | 2023 | 41.5 million customers | ISP customer data exposed |
| 15 | Telangana Police’s Hawk Eye App Breach | 2023 | 200,000 citizens | Hacker arrested |
| 16 | Tamil Nadu’s Facial Recognition Portal Breach | 2023 | 6 million records | Security lapse in government portal |
| 17 | NDMA Data Breach | 2023 | 93,000 volunteers | Data sold on the dark web |
| 18 | boAt India Data Breach | 2023 | 7.5 million users | Personal information leaked online |
| 19 | Hyundai Motor India Data Leak | 2023 | Customer & vehicle details | Fixed vulnerability |
| 20 | Burger Singh Website Hack | 2023 | Website defaced | Playful hacker response |
| 21 | WazirX Crypto Exchange Breach | 2024 | $230 million stolen | Multisig wallet compromised |
| 22 | Indian Government Cyberattacks | 2024 | Defence & critical sectors | Phishing attacks detected |
| 23 | BSNL Data Breach | 2024 | Millions of users | 278GB of sensitive data leaked |
| 24 | Indian Energy Sector Cyberattack | 2024 | 8.81GB of data stolen | Espionage campaign |
| 25 | UP Marriage Assistance Scheme Fraud | 2024 | Rs. 1 crore misused | Unauthorized payments made |