

## Day 2 -

- **What is data?**

- Data refers to raw, unorganized facts, figures, or observations that lack context but can be processed to derive meaning. It can exist in various formats, such as text, numbers, images, and sounds.

- **Primary types of data**

1. Qualitative (**Descriptive Data**)

- Describes characteristics or attributes (non-numeric).
- Examples: Customer feedback, survey responses, interviews.

2. Quantitative (**Numerical Data**)

- Represents measurable quantities.
- Examples: Temperature readings, sales figures, population counts.

- **Sub types of data :**

- **Transport Data** – Includes vehicle movement, traffic flow, fuel consumption, and logistics data.
- **Geographical Data** – Spatial data, maps, GPS coordinates, elevation levels.
- **Natural Data** – Ecosystem information, biodiversity records, climate impact data.
- **Meteorological Data** – Weather forecasts, temperature patterns, wind speed, precipitation levels.
- **Statistical Data** – Census reports, economic indicators, research statistics.
- **Financial Data** – Stock market trends, company earnings, taxation reports.
- **Scientific Data** – Lab experiment results, DNA sequencing, astronomical observations.
- **Cultural Data** – Artefacts, traditions, linguistic trends, heritage records.

- **What is information?**

Information is processed, structured, and meaningful data that provides insights or knowledge. It is derived from raw data through analysis, filtering, and interpretation.

- **Example:**
- **Raw Data:** 100, 102, 105, 107
- **Processed Data (Information):** The temperature increased by an average of 2°C per hour.

- **Raw data to information**

- **Data Collection** – Gathering unprocessed data.
- **Data Cleaning** – Removing errors, duplicates, inconsistencies.
- **Data Processing** – Organizing and analysing the data.
- **Data Interpretation** – Extracting insights.
- **Data Representation** – Presenting data in graphs, charts, reports.

- **Applications and importance of data :**

- **Business Analytics** – Market trends, customer behaviour.
- **Healthcare** – Patient records, disease prediction.
- **Finance** – Fraud detection, stock analysis.
- **Education** – Student performance tracking, research.
- **Science & Research** – Innovations, experiments.
- **Government & Policy-Making** – Census data, economic planning.

▪ **Data sources :**

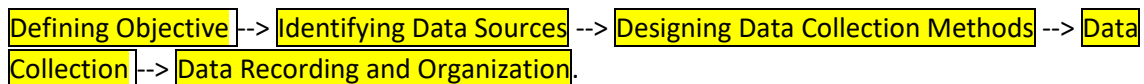
1. **Primary source**

- Door-to-Door Surveys** – Direct feedback collection.
- Student Thesis** – Original academic research.
- Personal Interviews** – One-on-one discussions for insights.

2. **Secondary source** (Pre-collected data from other sources)

- Internet** – Online research papers, reports.
- Books** – Published literature, encyclopaedias.
- Newspapers** – News articles, editorial analysis.

▪ **Data gathering – (data collection\ data acquisition) :**



- **Defining Objective** – Identifying the purpose of data collection.
- **Identifying Data Sources** – Choosing primary or secondary sources.
- **Designing Data Collection Methods** – Surveys, observations, sensors.
- **Data Collection** – Gathering data through selected methods.
- **Data Recording and Organization** – Storing and structuring the data.

▪ **RDBMS – (Relational Database Management System):**

(Tabular format depends on the format of data)

- Stores data in tabular format (rows & columns).
- Ensures data integrity, consistency, and accessibility.
- Examples: MySQL, PostgreSQL, Oracle DB, MS SQL Server.

- **Assignment** – Research and Gather Data of your choice with clear goals and objectives with proper and valid explanation. Decide goal , Find data from secondary source and organize that data on excel

**Goal:** To analyse the impact of climate change over the past decade by studying global temperature anomalies, CO<sub>2</sub> concentration levels, Sea Level Rise, Arctic Ice Extent, and notable climate events.

**Objectives:**

1. **Identify Trends in Climate Change**
  - Analyse how global temperature anomalies have changed over the past decade.
  - Study the increase in atmospheric CO<sub>2</sub> levels and their correlation with rising temperatures.
2. **Understand Climate Patterns and Events**
  - Examine significant climate-related events, such as El Niño, heatwaves, and extreme weather patterns.
  - Assess how these events align with observed temperature and CO<sub>2</sub> trends.
3. **Utilize Reliable Data Sources**
  - Use verified secondary data sources such as NASA, NOAA, and global climate monitoring reports.
  - Ensure accuracy and consistency in data collection.
4. **Organize Data for Better Insights**
  - Structure data in an Excel file with key indicators: Year, Temperature Anomaly, CO<sub>2</sub> Levels, Sea Level Rise, Arctic Ice Extent, and Notable Events.
  - Format the dataset for easy analysis and visualization.

**Data Source:**

- **NASA Global Temperature Data:** [NASA SVS](#)
- **NOAA CO<sub>2</sub> Concentration Data:** [NOAA Global Monitoring Laboratory](#)
- **IPCC Climate Reports:** [IPCC](#)

Notable climate events							
Sr. No.	Year	Global Temperature Anomaly (°C)	CO <sub>2</sub> Concentration (ppm)	Notable Climate Events	Sea Level Rise (mm)	Arctic Ice Extent (Million Sq Km)	Number of Extreme Weather Events
1)	2013	0.62	395.4	Heat waves	3.1	5.35	78
2)	2014	0.74	397.2	Polar ice melt	3.2	5.29	82
3)	2015	0.93	399.4	Extreme drought	3.4	5.15	90
4)	2016	1	404.2	Strong El Niño	3.5	4.98	97
5)	2017	0.91	406.5	Hurricane activity	3.7	4.82	105
6)	2018	0.82	408.5	Wildfires	3.9	4.67	115
7)	2019	0.95	411	Ocean acidification	4.1	4.5	120
8)	2020	0.98	414.2	Sea level rise	4.3	4.35	130
9)	2021	0.84	416.5	Arctic ice loss	4.5	4.2	140
10)	2022	0.86	419	Flooding	4.7	4.05	150
11)	2023	1.17	420	Record global heat	5	3.95	160
12)	2024	1.29	422.2	Extreme storms	5.2	3.8	170

## ➤ Assignment -Sub types of data and their examples

### 1. Transport Data -

**Definition:** Data related to the movement of vehicles, traffic flow, fuel consumption, and logistics.

**Example:**

- **Vehicle Movement:** GPS tracking data of Uber cars to optimize routes.
- **Traffic Flow:** Google Maps analysing real-time traffic congestion.
- **Fuel Consumption:** Airlines monitoring fuel usage for cost efficiency.
- **Logistics Data:** Amazon tracking delivery trucks for on-time shipments.

### 2. Geographical Data -

**Definition:** Data that represents spatial locations, terrain, and geographic features.

**Example:**

- **Spatial Data:** Google Earth mapping urban and rural areas.
- **Maps:** Open Street Map showing roads and landmarks.
- **GPS Coordinates:** A hiking app providing real-time location tracking.
- **Elevation Levels:** NASA satellite data measuring mountain heights.

### 3. Natural Data -

**Definition:** Data related to ecosystems, biodiversity, and environmental conditions.

**Example:**

- **Ecosystem Information:** Wildlife conservation tracking tiger populations.
- **Biodiversity Records:** Marine biologists studying coral reef species.
- **Climate Impact Data:** Deforestation analysis using satellite imagery.

### 4. Meteorological Data -

**Definition:** Data that helps monitor and predict weather patterns and atmospheric conditions.

**Example:**

- **Weather Forecasts:** The Indian Meteorological Department (IMD) predicting monsoons.
- **Temperature Patterns:** NASA tracking global warming trends.
- **Wind Speed:** Wind energy farms using sensors for turbine efficiency.
- **Precipitation Levels:** Farmers using rainfall data for crop planning.

### 5. Statistical Data -

**Definition:** Numerical data used for analysis, decision-making, and research.

**Example:**

- **Census Reports:** Government collecting population data for policy-making.
- **Economic Indicators:** GDP growth rate published by the World Bank.
- **Research Statistics:** WHO tracking COVID-19 case numbers globally.

## 6. Financial Data -

**Definition:** Data related to economic transactions, market trends, and financial performance.

**Example:**

- **Stock Market Trends:** NSE/BSE providing real-time stock prices.
- **Company Earnings:** Tesla publishing quarterly revenue reports.
- **Taxation Reports:** Government tracking GST collection statistics.

## 7. Scientific Data -

**Definition:** Data generated from experiments, observations, and research.

**Example:**

- **Lab Experiment Results:** COVID-19 vaccine trials by Pfizer.
- **DNA Sequencing:** Human Genome Project mapping genetic structures.
- **Astronomical Observations:** NASA recording black hole images.

## 8. Cultural Data -

**Definition:** Data related to traditions, heritage, languages, and historical artefacts.

**Example:**

- **Artefacts:** Museum databases cataloguing ancient Egyptian sculptures.
- **Traditions:** UNESCO preserving intangible cultural heritage like Indian classical dance.
- **Linguistic Trends:** Google Translate analysing global language patterns.
- **Heritage Records:** Historical manuscripts digitized by the British Library.