

---

## ◆ Week 1: Dataset Collection, Installation & Data Cleaning Using Power BI

### ◆ Objective

The main objective of Week 1 was to understand the fundamentals of data visualization tools, collect real-world datasets, and perform initial data cleaning and transformation using **Power BI Power Query Editor**.

---

### ◆ Tools & Technologies Used

- Power BI Desktop
  - Microsoft Excel
  - Open datasets (Call Center & FEMA dataset)
- 

### ◆ Activities Performed

#### 1 Power BI Installation & Environment Setup

- Installed Power BI Desktop and explored its interface

Understood the workflow:

Data Import → Power Query → Data Model → Visualization

- Learned about:
    - Query Editor
    - Relationships
    - Data types
    - Refresh options
- 

#### 2 Dataset Collection

- Collected **Call Center Dataset** for practice
  - Collected **FEMA dataset** from an open government data source
  - Analyzed dataset structure:
    - Rows and columns
    - Categorical and numerical attributes
    - Missing values and inconsistencies
-

### **Data Cleaning Steps Performed in Power Query**

The following data cleaning operations were performed on the dataset:

- Removed **duplicate records**
  - Removed **irrelevant columns** that were not useful for analysis
  - Renamed column headers to meaningful names
  - Changed incorrect **data types** (text → number, date → date format)
  - Handled **missing values**:
    - Replaced null values
    - Removed rows with excessive missing data
  - Trimmed extra spaces and cleaned text columns
  - Split merged columns where necessary
  - Filtered unnecessary rows
- 

### **Data Transformation**

- Created new calculated columns using Power Query
  - Standardized categorical values
  - Sorted and filtered data for consistency
  - Prepared clean dataset ready for analysis and visualization
- 

### **Outcome of Week 1**

By the end of Week 1:

- Successfully installed and explored Power BI
- Understood real-world data issues
- Cleaned and transformed datasets using Power Query
- Prepared structured datasets for analysis