

## **Milestone 1: Data Collection, Exploration, and Preprocessing**

### **Objectives**

- Collect and explore historical sales data.
  - Preprocess the data for analysis and model building.
- 

### **Tasks**

#### **1. Data Collection**

- Acquire a dataset containing historical sales data (e.g., daily or weekly sales from retail or e-commerce platforms).
- Ensure dataset includes relevant features such as:
  - Sales amount
  - Date
  - Promotions
  - Holidays
  - Weather, etc.

#### **2. Data Exploration (EDA)**

- Perform exploratory data analysis (EDA) to understand:
  - Trends
  - Seasonality
  - Missing values in the dataset
- Generate:
  - Summary statistics
  - Outlier checks
  - Correlation analysis (e.g., sales vs. promotions/holidays).

### 3. Data Preprocessing

- Handle missing values, remove duplicates, and resolve inconsistencies.
  - Engineer time-based features (e.g., day of the week, month, seasonality, promotional periods).
  - Apply data scaling and transformations (e.g., normalization) for model readiness.
- 

#### Deliverables

- **Data Exploration Report** → Summary of findings from data exploration, highlighting trends, seasonality, and data quality issues.
- **EDA Notebook (Jupyter)** → Visualizations such as line plots, histograms, and correlation heatmaps.
- **Cleaned Dataset** → A processed dataset ready for analysis and model training.