**Project Title: US Superstore – Business Intelligence ETL and Dashboarding Project**

**Objective:**

To build a business intelligence solution using the US Superstore dataset that extracts, transforms, and loads retail sales data into a centralized structure, enabling insight generation on sales performance, profit margins, and customer segmentation through visual analytics.

**Project Phases:**

**1. Data Acquisition**

* Collect raw sales data from the classic **US Superstore dataset** (in Excel).
* Simulate multiple sources if needed (e.g., split by region, year, or department).
* Validate data schema, column consistency, and structure.

**2. ETL Pipeline Development**

**🔹 Extract**

* Load data from multiple files/tables into a Python environment using Pandas.
* Automate reading logic using loops or file path templates.

**🔹 Transform**

* Clean data:
  + Handle null values
  + Standardize categories (e.g., Segment, Region, Product Category)
  + Convert data types (dates, currency fields)
* Feature engineering:
  + Calculate profit margins
  + Derive order week/month/quarter
  + Create flags for high-value customers or products

**🔹 Load**

* Store the cleaned and transformed dataset into a **centralized structure**:
  + Flat file (Excel)
  + Define schema for future scalability

**3. Data Analysis**

* Perform EDA
* Calculate KPIs:
  + Total Sales, Profit, Quantity, and Discounts by dimension (Region, Category, Segment)
  + Profit margin = Profit / Sales

**4. Dashboard Development**

* Use **Tableau** to build interactive dashboards:
  + Sales & Profit Overview
  + Segment-wise Customer Insights
  + Regional Performance Breakdown
  + Time Trend Analysis (Monthly/Quarterly)
* Apply filters and slicers to allow stakeholder interactivity

**Tools & Technologies:**

* **Programming Language**: Python
* **Libraries**: Pandas, Matplotlib
* **Dashboarding**: Tableau
* **IDE**: Jupyter Notebook