

## **Project Title: Retail Sales Performance & Customer Behavior Analysis**

**Tools Used:** Python (Pandas), PostgreSQL (SQL), Power BI

---

### **1. Objective:**

To analyze a retail dataset using SQL and visualize insights via Power BI to improve sales, customer understanding, and business decision-making

---

### **2. Workflow Overview:**

1. Cleaned raw sales dataset using **Python (Pandas)**
  2. Extracted cleaned data to **PostgreSQL database**
  3. Performed SQL queries for customer, sales, profit, and region-based insights
  4. Connected PostgreSQL to **Power BI** and created visual dashboards
- 

### **3. SQL Analysis Tasks:**

Performed the following analysis using PostgreSQL queries:

- 1 Total number of orders
  - 2 Unique customers count
  - 3 Monthly order distribution
  - 4 Average shipping delay
  - 5 Same-day shipping orders count
  - 6 Sales & profit by region
  - 7 Top-selling products
  - 8 Highest average discount by region
  - 9 Most frequent customers
  - 10 Customers with more than 5 orders
  - 11 Top 3 customers by profit
  - 12 Sales, quantity & profit by category and sub-category
  - 13 Sub-category with highest discount
- 

### **4. SQL Output Snapshots:**

[Click here to view the complete SQL Queries on GitHub](#)

---

### **5. Exploratory Data Analysis (EDA Using Python )**

Before exporting the dataset to SQL, an in-depth exploratory data analysis (EDA) was performed using Pandas, Matplotlib, and Seaborn to understand key patterns in sales, customer behavior, and performance metrics.

#### Key Tasks Performed:

- **Data Cleaning:**
    - Removed duplicates
    - Handled missing/null values
    - Corrected date formats for consistency
  - **Feature Engineering:**
    - *Shipping Delay (Days)* = *Ship Date* – *Order Date*
    - *Profit Margin* = *Profit* ÷ *Sales*
  - **Visual Analysis:**
    - Sales and Profit trends across Category and Region
    - Impact of Discount on Profit
    - Correlation between Shipping Delay and Order Profit
    - Monthly Sales performance tracking
    - Identification of Top Customers by Order Count and Profit Contribution
- 

## 6. Power BI Visualization:

Connected the SQL tables directly into Power BI and developed an interactive dashboard with:

- Region-wise & category-wise sales/profit visualizations
- Customer-wise quantity tracking
- KPI tiles: Total Orders, Average Discount, Top Customers
- Trend analysis (monthly sales, discount vs. profit)

#### . Power BI Dashboard Snapshots:



## 7. Conclusion:

This project demonstrates a complete data analytics pipeline — from data cleaning (Python), database integration (PostgreSQL), SQL-driven analysis, to final dashboarding (Power BI). The insights derived support better sales strategy and customer understanding.

