TARGET DATASET EXPLORATION: AN SQL AND PYTHON APPROACH

OBJECTIVE:

This project analyses an e-commerce dataset, using SQL for data extraction and querying and Python for advanced analysis and visualization. The objective is to generate actionable business insights across sales performance, customer behaviour, and customer retention, supporting data-driven decision-making.

WORKFLOW:

- SQL: Import products.csv, orders.csv, customers.csv, sellers.csv.
- Python: Use pandas to manipulate and visualize data. Query Execution: Run SQL queries on MySQL Workbench. Visualization: Create insightful graphs using Google Colab.

TOOLS USED:

- MySQL Workbench
- Python Libraries: Pandas, Matplotlib
- Google Colab Notebook

FILE NAMES:

- SQL File Name: Python_SQL_Project.sql / SQL_Solutions.pdf
- Python File Name: Python+_SQL_Project.ipynb

KEY ANALYSES:

- **Total Sales by Product Category**: Evaluates revenue contribution by categories to identify primary growth drivers and underperforming segments.
- **Monthly Order Trends**: Analyzes order volume over time to find the seasonality patterns and shifts in customer demand.
- **Average Order Value (AOV) Analysis**: Evaluates changes in order size and customer spending behavior, supporting pricing and cross-sell strategies.
- **Customer Retention and Repeat Purchase Behavior**
  Examines customer loyalty and repeat buying patterns to identify opportunities for improving retention and lifetime value.

BUSINESS INSIGHTS:

- **Revenue Concentration**
  A small number of product categories account for the majority of total revenue,

indicating clear core growth drivers. Health Beauty, Watches present, Bed Table Bath are the categories which generates the highest revenue.

- **Order Volume Growth**
  Over the month Order volumes show consistent growth, reflecting sustained demand and business momentum.
- **Stable Average Order Value (AOV)**
  AOV remains relatively stable, suggesting predictable customer spending behaviour and limited volatility in Order size.
- **Strong Contribution from Repeat Customers**
  Repeat customers contribute a significant share of revenue, highlighting meaningful opportunities to further strengthen customer retention and lifetime value.
- **Correlation Between Product Price and Purchase Count:** The analysis suggests price sensitivity within the customer base, where demand drops sharply as prices increase.

OUTCOME:

This project demonstrates strong proficiency in SQL querying, Python-based data analysis and visualization, and business-focused storytelling—core skills essential for delivering impact as a Data Analyst.