

# E-Commerce Data Analysis Project

## Project Overview

This project focuses on analyzing an e-commerce business dataset to extract meaningful insights related to **sales performance, customer behavior, product trends, and city-level revenue distribution**. The analysis was conducted using Python for exploratory data analysis and SQL for answering real-world business questions. The project demonstrates an end-to-end data analytics workflow on transactional e-commerce data.

## Tools & Technologies Used

- Python (Pandas, Matplotlib, Seaborn) for data cleaning, analysis, and visualization
- SQL (PostgreSQL / MySQL compatible) for business-driven analytical queries
- Jupyter Notebook for Python-based analysis
- CSV file containing transactional e-commerce data

## Dataset Description

The dataset contains transaction-level information from an e-commerce platform, including order details, customer location, product information, pricing, quantity sold, customer ratings, and platform data. This dataset supports both operational analysis and strategic decision-making.

## Python Analysis (Exploratory Data Analysis)

- Cleaned and preprocessed the dataset by handling missing values and formatting columns
- Calculated revenue using quantity × price
- Performed exploratory analysis on sales trends, city-wise performance, and product categories
- Created visualizations to represent revenue distribution, order trends, and customer behavior

## SQL Analysis (Business Queries)

- Calculated total orders and total revenue generated by the platform
- Identified top cities by customer ratings and by total revenue
- Computed average order value (AOV) at the city level
- Determined top revenue-generating product categories
- Identified the most popular product category in each city using window functions
- Analyzed recent orders, high-value orders, and top-rated products

## Key Insights Generated

- High-revenue cities contributing significantly to overall sales were identified
- Top-performing product categories were determined based on revenue contribution
- Customer rating patterns were analyzed to understand product and platform performance
- Average order value analysis provided insights into customer purchasing behavior
- City-wise category preferences were identified to support targeted marketing strategies

## Outcome

This project demonstrates the ability to transform raw e-commerce data into actionable business insights. It highlights practical skills in Python-based data analysis, SQL querying, and analytical thinking, making it suitable for showcasing on GitHub and professional portfolios.