

Introduction

Analyzing e-commerce data helps organizations understand customer behavior, improve sales performance, optimize operations, and reduce return rates. This project focuses on performing a complete end-to-end data analysis using real-world e-commerce datasets to uncover insights related to customer purchasing trends, order patterns, revenue performance, and product returns.

Abstract

This e-commerce analysis project aims to evaluate customer behavior, sales performance, and return patterns using a structured dataset containing Orders, Customers, Products, and Returns information. Through data preprocessing, exploratory analysis, and visualization.

Tools Used

1. SQL (MySQL)

- Used for data extraction, cleaning, and transformation.
- Helps join tables such as Orders, Customers, Returns, and Products.
- Performs aggregations including revenue, order count, AOV, and return rate.

2. Python (Pandas, NumPy, Matplotlib)

- Used for data preprocessing and exploratory data analysis (EDA).
- Generates charts to understand sales trends and customer distribution.

3. Power BI

- Used to build interactive dashboards.
- Visualizes KPIs such as Total Revenue, AOV and Monthly Sales.
- Highlights return patterns, loyalty segments, and category-wise sales.

4. Excel / Google Sheets

- Used for quick validation and data filtering.
- Helps in cleaning raw data before importing into SQL or BI tools.

Steps Involved in Building the Project

Step 1: Data Collection

- Gather datasets for Orders, Customers, Products, and Returns.
- Validate the structure, formats, and consistency of each file.

Step 2: Data Cleaning & Preparation

- Convert dates into proper format (YYYY-MM-DD).
- Remove duplicate records and fix missing values.
- Standardize product categories, return reasons, and customer segments.
- Merge tables using SQL joins for a unified view.

Step 3: Exploratory Data Analysis (EDA)

- Analyze order volume trends, customer demographics, and product performance.

- Compute KPIs like:
- Total Revenue
- Total Orders
- Average Order Value (AOV)
- Return Rate
- Identify top-selling products and high-return categories.

Step 4: Data Modelling

- Create star-schema relationships in the BI tool.
- Build fact tables (Orders, Returns) and dimension tables (Customers, Products).

Step 5: Visualization & Dashboard Creation

- Design dashboards with clear visuals:
- Sales Trend Line Chart
- Revenue by Category Bar Chart
- Customer Loyalty Segmentation
- Return Insights and Reasons
- Add slicers for category, location, and customer segment.

Step 6: Insights & Recommendations

- Summarize key findings such as:
- Peak buying months
- Customer segments driving revenue
- Product performance insights
- Operational gaps due to high returns
- Provide business recommendations for marketing, inventory, and customer engagement.

Conclusion

This e-commerce analysis project demonstrates how raw transactional data can be transformed into impactful business insights using SQL, Python, and BI tools. By exploring customer behavior, sales trends, product performance, and return patterns, the project highlights key areas for growth and optimization.