Exploratory Data Analysis of Multi-Mall Shopping Data Using SQL

Author: Emmanuel Nwanguma

Role: Data Analyst

Date: 20th May 2025

GitHub: https://github.com/Emart29/EDA-Multimall-shopping-sql

LinkedIn: https://linkedin.com/in/nwangumaemmanuel

Email: nwangumaemmanuel29@gmail.com

Table of Contents

1. Project Overview
2. Dataset Description
3. Objectives
4. Tools & Technologies
5. SQL Analyses Performed
6. Key Insights
7. Conclusion
8. Appendix

1. Project Overview

This project presents an exploratory data analysis (EDA) of a multi-mall shopping dataset using SQL. The goal is to extract meaningful insights about customer demographics, purchasing behaviors, sales trends, and mall performance to support data-driven decision-making.

2. Dataset Description

The dataset **customer_shopping_data** contains shopping records across multiple malls. Below is a description of key fields:

Column Name	Description
invoice_no	Unique ID per transaction
customer_id	Unique customer identifier
gender	Gender of the customer
age	Age of the customer
category	Product category
quantity	Quantity purchased
price	Price per unit
total_amount	Total price paid (price × quantity)
payment_method	Mode of payment used
invoice_date	Date of purchase
shopping_mall	Name of the mall where purchase was made

3. Objectives

- ☐ Determine overall and segment-specific revenue
- ☐ Analyze customer demographics and behavior

- ☐ Identify top-performing malls and product categories
- ☐ Understand trends in payment method usage
- ☐ Detect peak sales periods and shopping patterns

4. Tools & Technologies

- ☐ **SQL**-for querying and data analysis
- ☐ **Microsoft SQL Server**-database engine
- ☐ **GitHub**-version control and project hosting

5. SQL Analyses Performed

- ☐ **Customer Analysis**
  - ☐ Count of unique customers
  - ☐ Gender-wise customer distribution
  - ☐ Age group segmentation
- ☐ **Revenue Insights**
  - ☐ Total revenue
  - ☐ Revenue by gender, age group, and shopping mall
  - ☐ Monthly and weekday revenue trends
- ☐ **Product and Category Insights**
  - ☐ Top categories by quantity and revenue
  - ☐ Gender-wise product preferences
- ☐ **Payment Method Trends**

- ☐ Usage by age group
- ☐ Most common payment types

- ☐ **Shopping Behavior**
  - ☐ Average basket size
  - ☐ Peak shopping days by mall
  - ☐ Spending trends by age

6. Key Insights

- ☐ Customers aged **40–49** are the highest spenders overall.
- ☐ **Clothing** and **Cosmetics** dominate category sales.
- ☐ **Cash** is the most used payment method among all customers.

7. Conclusion

This SQL-based EDA uncovered several actionable insights:

- ☐ Age and gender significantly influence purchasing behavior.
- ☐ Malls have varying peak performance days, suggesting marketing opportunities.

These insights can help mall managers improve inventory planning, targeted promotions, and customer experience.

8. Appendix

- ☐ **Full SQL Script:** See **exploratory_analysis.sql** in the GitHub repo
- ☐ **Data Source:** Synthetic dataset for educational purposes