

Data Analyst Project

Customer Shopping Analysis (2021-2023)

Team Members : Karunakaran R and Dharun RR

Email id : rkaranakaranraja@gmail.com and
ddharunrio7@gmail.com

Batch : 12th May CDA

Project id : PRDA-05

Project team id : PTID-CDA-SEP-25-722

Time period : 12th sep to 16th sep

Introduction:

This project analyzes customer shopping data from 10 different malls between 2021 to 2023. It understands behavior across gender, age groups, product categories, and payment methods. The objective is to derive insights that help businesses improve sales, optimize inventory, and target customers more effectively.

Dataset details:

The dataset contains 1000 customer data Which contains columns such as invoice_no, customer_id, gender, age, category, quantity, price, payment_method, invoice_date, shopping_mall.

Tools & methodology:

SQL:

Data cleaning and query for 10 tasks

Excel:

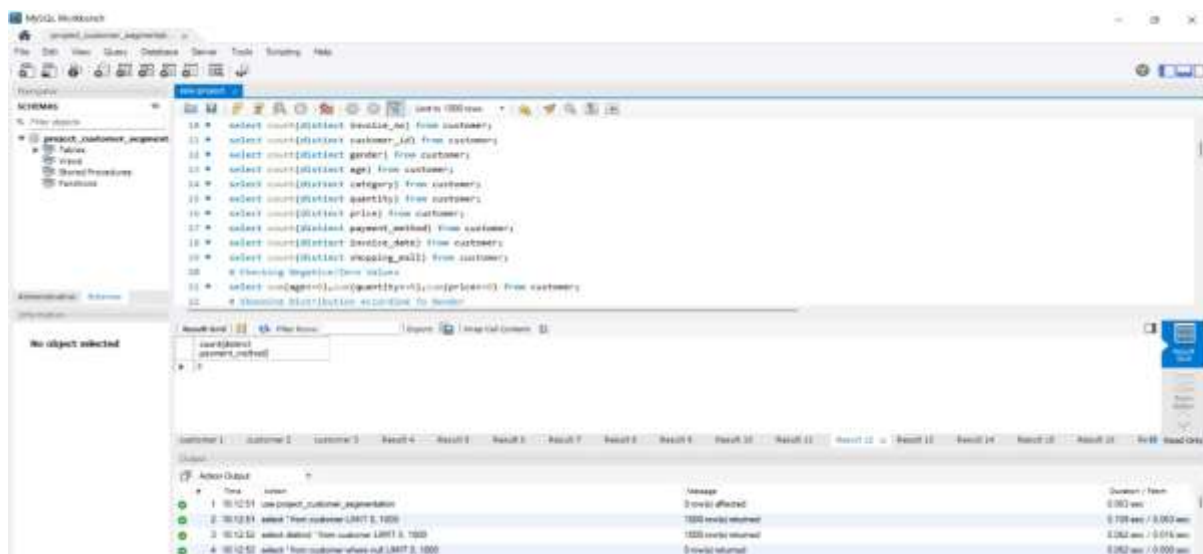
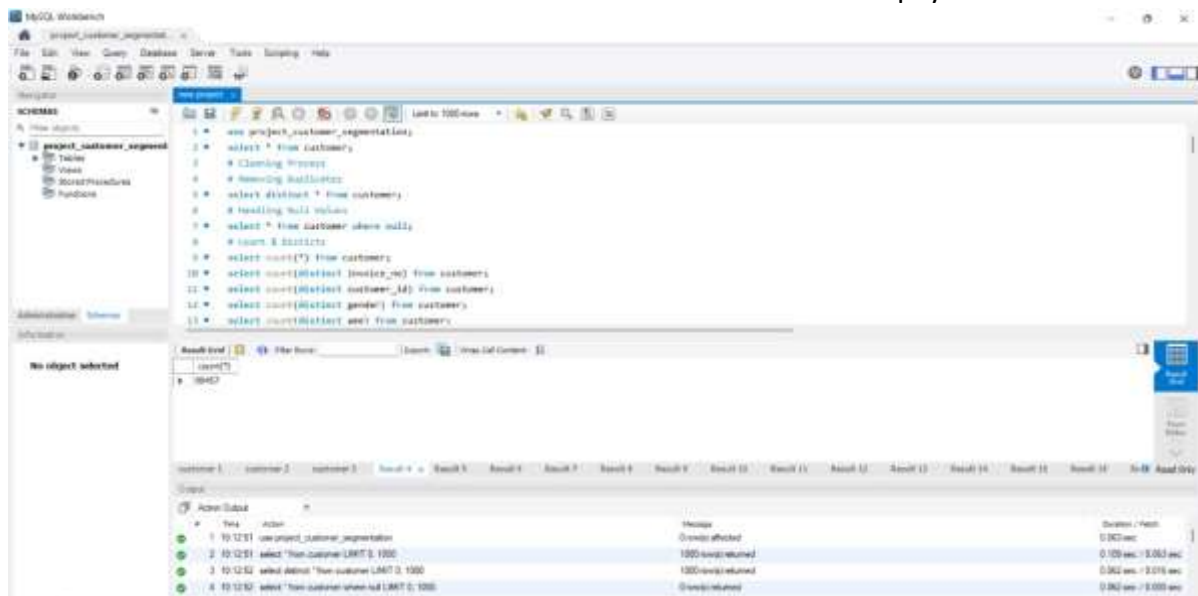
Using pivot tables for quick aggregation and validation

Power Bi:

Dashboard creation and visualization

SQL and Excel results:

We have done cleaning and 10 task in SQL query like shopping distribution according to gender, max of products according to gender, max of revenue according to gender, purchase categories relative to other columns, shopping distribution according to age, max of products according to age, max of revenue according to age, payment methods relative to other columns and distribution of payment methods



In Excel we used pivot table for quick aggregation and for validation. According to gender, age, categories and Payment methods for 10 different malls from 2021 to 2023.

1. Shopping distribution according to gender

Gender	Sum of Quantity	Sum of Price	Sum of Revenue
Male	1128	1000.00	1128.00
Female	1128	1000.00	1128.00
Grand Totals	2256	2000.00	2256.00

2. Which age group generated more revenue

Age Group	Sum of Quantity	Sum of Price	Sum of Revenue
18-24	1128	1000.00	1128.00
25-34	1128	1000.00	1128.00
35-44	1128	1000.00	1128.00
45-54	1128	1000.00	1128.00
55-64	1128	1000.00	1128.00
65+	1128	1000.00	1128.00
Grand Totals	6768	6000.00	6768.00

3. Which gender generated more revenue

Gender	Sum of Quantity	Sum of Price	Sum of Revenue
Male	1128	1000.00	1128.00
Female	1128	1000.00	1128.00
Grand Totals	2256	2000.00	2256.00

4. Distribution of purchase categories relative to other columns

Category	Sum of Quantity	Sum of Price	Sum of Revenue
Electronics	1128	1000.00	1128.00
Clothing	1128	1000.00	1128.00
Home & Garden	1128	1000.00	1128.00
Books	1128	1000.00	1128.00
Food & Beverage	1128	1000.00	1128.00
Health & Beauty	1128	1000.00	1128.00
Travel	1128	1000.00	1128.00
Education	1128	1000.00	1128.00
Other	1128	1000.00	1128.00
Grand Totals	11280	10000.00	11280.00

5. Distribution of payment method

Payment Method	Sum of Quantity	Sum of Price	Sum of Revenue
Cash	1128	1000.00	1128.00
Credit Card	1128	1000.00	1128.00
Debit Card	1128	1000.00	1128.00
Gift Card	1128	1000.00	1128.00
Other	1128	1000.00	1128.00
Grand Totals	5640	5000.00	5640.00

6. Which gender generated more products

Gender	Sum of Quantity	Sum of Price	Sum of Revenue
Male	1128	1000.00	1128.00
Female	1128	1000.00	1128.00
Grand Totals	2256	2000.00	2256.00

1. Shopping distribution according to gender

Gender	Sum of Quantity	Sum of Price	Sum of Revenue
Male	1128	1000.00	1128.00
Female	1128	1000.00	1128.00
Grand Totals	2256	2000.00	2256.00

2. Which age group generated more revenue

Age Group	Sum of Quantity	Sum of Price	Sum of Revenue
18-24	1128	1000.00	1128.00
25-34	1128	1000.00	1128.00
35-44	1128	1000.00	1128.00
45-54	1128	1000.00	1128.00
55-64	1128	1000.00	1128.00
65+	1128	1000.00	1128.00
Grand Totals	6768	6000.00	6768.00

3. Which gender generated more revenue

Gender	Sum of Quantity	Sum of Price	Sum of Revenue
Male	1128	1000.00	1128.00
Female	1128	1000.00	1128.00
Grand Totals	2256	2000.00	2256.00

4. Distribution of purchase categories relative to other columns

Category	Sum of Quantity	Sum of Price	Sum of Revenue
Electronics	1128	1000.00	1128.00
Clothing	1128	1000.00	1128.00
Home & Garden	1128	1000.00	1128.00
Books	1128	1000.00	1128.00
Food & Beverage	1128	1000.00	1128.00
Health & Beauty	1128	1000.00	1128.00
Travel	1128	1000.00	1128.00
Education	1128	1000.00	1128.00
Other	1128	1000.00	1128.00
Grand Totals	11280	10000.00	11280.00

5. Distribution of payment method

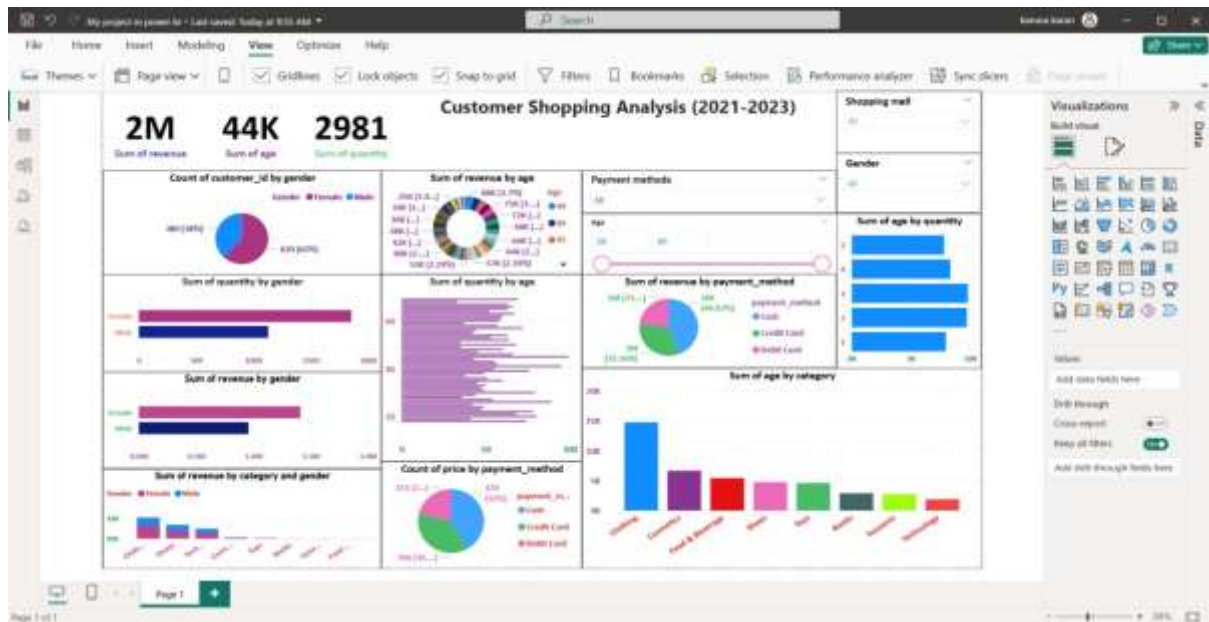
Payment Method	Sum of Quantity	Sum of Price	Sum of Revenue
Cash	1128	1000.00	1128.00
Credit Card	1128	1000.00	1128.00
Debit Card	1128	1000.00	1128.00
Gift Card	1128	1000.00	1128.00
Other	1128	1000.00	1128.00
Grand Totals	5640	5000.00	5640.00

6. Which gender generated more products

Gender	Sum of Quantity	Sum of Price	Sum of Revenue
Male	1128	1000.00	1128.00
Female	1128	1000.00	1128.00
Grand Totals	2256	2000.00	2256.00

Power Bi dashboard:

Finally we used power Bi for visualization. We create dashboard in power BI using which I have find queries in SQL and this dashboard explain insights, trends are according gender, age, payment methods and quantity.



Insights & recommendations:

- Females spent 20% more than males → marketing campaigns can focus on them
- Age group 21–30 is most profitable → introduce student offers
- Credit Card is the most popular payment → give cashback partnerships
- Product Category Clothing had highest sales → stock more variety

Conclusion:

This project demonstrates the complete data analysis lifecycle using SQL, Excel, and Power BI. The findings can help businesses make data-driven decisions in marketing, inventory, and customer engagement.