CAPSTONE PROJECT -01

CUSTOMER DATA ANALYSIS

PROJECT ID: PRDA-05 CUSTOMER DATA ANALYSIS

INDEX

|  |  |  |
| --- | --- | --- |
| Sr.No | Content | Pageno |
| 1 | INTRODUCTION   * Project Objective * Dataset Overview * Tools & Technologies Used | 3 |
| 2 | DATA ANALYSIS USING SQL   * Questions / Problem Statements * What to Look For (Business Focus) * SQL Queries * Results & Screenshots of Queries * Business Insights + | 4 |
| 3 | CHALLENGES DURING EXECUTION   * Technical Challenges * Data-Related Challenges * Solutions/Approach to Fix Them * Learnings from Challenges | 30 |
| 4 | BUSINESS INSIGHTS | 34 |
| 5 | POWERBI DASHBOARDS   * Dashboard Objectives * Dashboard Screenshots * Insights from Visuals | 36 |
| 6 | CONCLUSION | 38 |
|  |  |  |

# INTRODUCTION

* The main objective of this project is to understand customer behaviour in a competitive business environment to enhance sales, customer satisfaction, and overall performance. Through customer data analysis, the project aims to uncover patterns and insights that help businesses make data-driven decisions regarding marketing strategies, product offerings, and customer engagement.
* The dataset used in this project consists of customer data collected from multiple shopping malls. It includes information on customer demographics, product categories, purchase quantities, pricing details, payment modes, and store locations. This diverse data enables analysis of customer segments, purchasing behaviour, and mall-level performance.
* SQL was utilized for data extraction and analysis, helping to segment customers based on gender, identify category-wise purchase trends, and examine the influence of age on buying habits. Power BI was employed to design interactive dashboards that visualize insights such as top-selling product categories, preferred payment methods, and mall-wise revenue distribution, supporting better business decisions in marketing, inventory, and pricing strategies.

# DATA ANALYSIS USING SQL

**Q.1. How is the shopping distribution according to gender?**

**What to look for:**

* Identify imbalances between customer groups to support targeted marketing strategies.
* Analyze which gender demonstrates higher purchase volume across different product categories.
* Determine which shopping mall caters to a larger proportion of male or female customers.

Description of query:

This query analyzes the gender distribution within each shopping mall. It groups customers by both shopping\_mall and gender, counts the unique customers in each group, and then calculates what percentage of each mall’s total customer base belongs to each gender.

Query:

SELECT shopping\_mall, gender,

COUNT(DISTINCT customer\_id) AS unique\_customers,

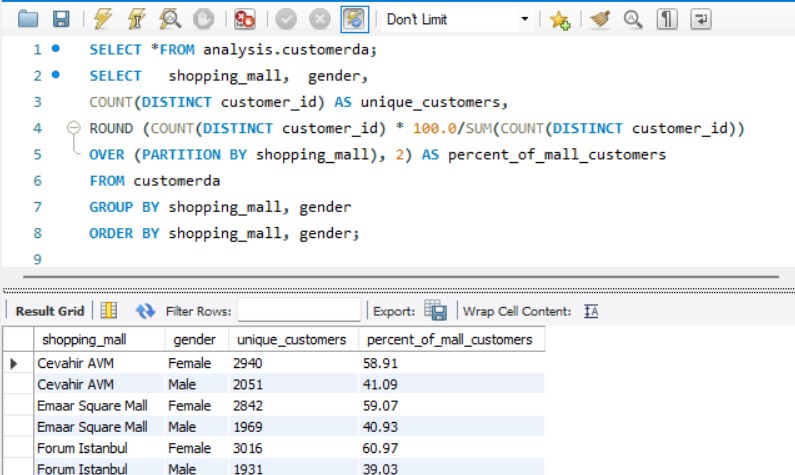
ROUND(COUNT(DISTINCT customer\_id\*100.0/SUM(COUNT(DISTINCT customer\_id))

OVER (PARTITION BY shopping\_mall), 2) AS percent\_of\_mall\_customers

FROM customerda

GROUP BY shopping\_mall, gender

ORDER BY shopping\_mall, gender;



Business insights:

Female customers shop more often and generate higher revenue across all malls and categories. There’s a clear gender imbalance (around 60% female vs. 40% male), with females leading in both purchase count and spending in every category and location

**Q.2. Which gender did we sell more products to?**

What to look for:

* Check which gender has higher total purchases or revenue to determine who buys more.

Description of query:

This query identifies the gender-wise distribution of unique customers in the dataset. It counts how many distinct customers belong to each gender and calculates their percentage share out of the total customer base.

Query:

SELECT \*FROM analysis.customerda;

SELECT gender,

COUNT (DISTINCT customer\_id) AS unique\_customers,

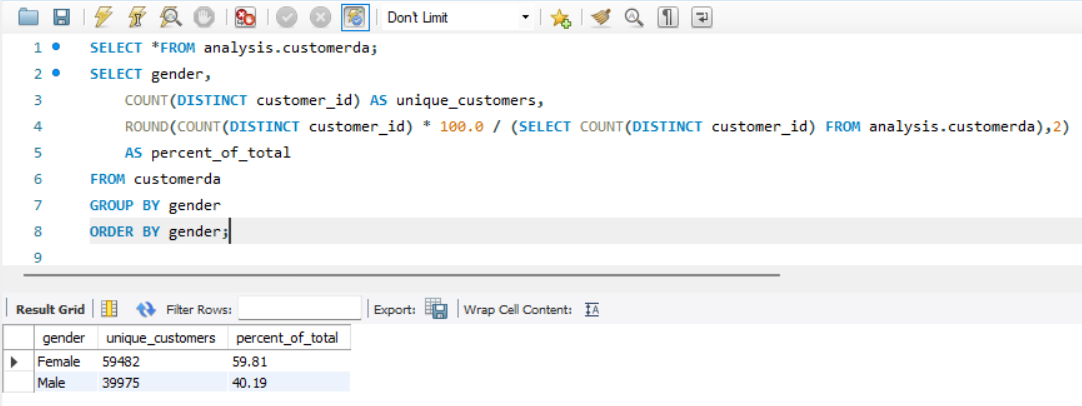
ROUND (COUNT (DISTINCT customer\_id) \* 100.0 / (SELECT COUNT (DISTINCT customer\_id) FROM analysis.customerda),2)

AS percent\_of\_total

FROM customerda

GROUP BY gender

ORDER BY gender;



Business insights:

This helps in understanding the overall customer composition by gender, showing what proportion of total unique customers are male or female.

**Q.3. Which gender generated more revenue?**

What to look for:

* To identify which gender contributes the **highest total revenue** in the dataset.
* To check if there is a **spending imbalance** between genders for targeted marketing strategies.
* To understand **spending behaviour differences** that can guide promotions and product positioning.

Description of query:

This query directly groups the data by gender and sums the revenue, then orders result to show the top contributing gender clearly. It is the most efficient and clear approach for quickly identifying which gender generates more revenue.

Query:

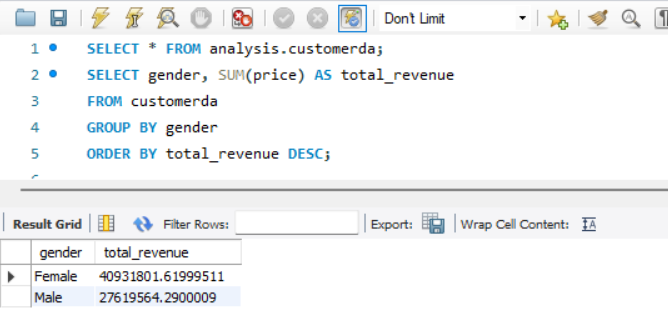
SELECT \* FROM analysis.customerda;

SELECT gender, SUM(price) AS total\_revenue

FROM customerda

GROUP BY gender

ORDER BY total\_revenue DESC;



Business insights:

We want to see which gender spends more overall, so we understand who is contributing more to our sales. Once we know this, we can focus our marketing and offers toward the group that actually buys more. This also helps us plan products and services better, based on the spending habits of each gender.

**Q.4. Distribution of purchase categories relative to other columns?**

What to look for:

We analyze how each product category is purchased across gender, payment methods, malls, and age groups. This helps identify which customer segment prefers which product category and where to focus promotions.

Description of query:

This query groups the purchase data by gender to check how many transactions each gender has made and how much they have spent in total. It also calculates the average purchase value per transaction for each gender. This helps in understanding which gender contributes more to sales and spends more on average.

Query:

In the database there are more than one column on which we can perform purchase distribution on each one of them, following are the queries for the same:

1. For gender column:

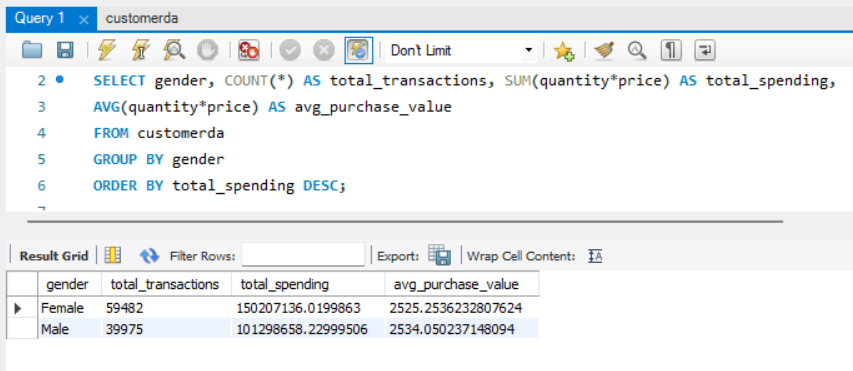
SELECT gender, COUNT (\*) AS total\_transactions, SUM (quantity\*price) AS total\_spending,

AVG (quantity\*price) AS avg\_purchase\_value

FROM customerda

GROUP BY gender

ORDER BY total\_spending DESC;



1. Category:

SELECT \* FROM analysis.customerda;

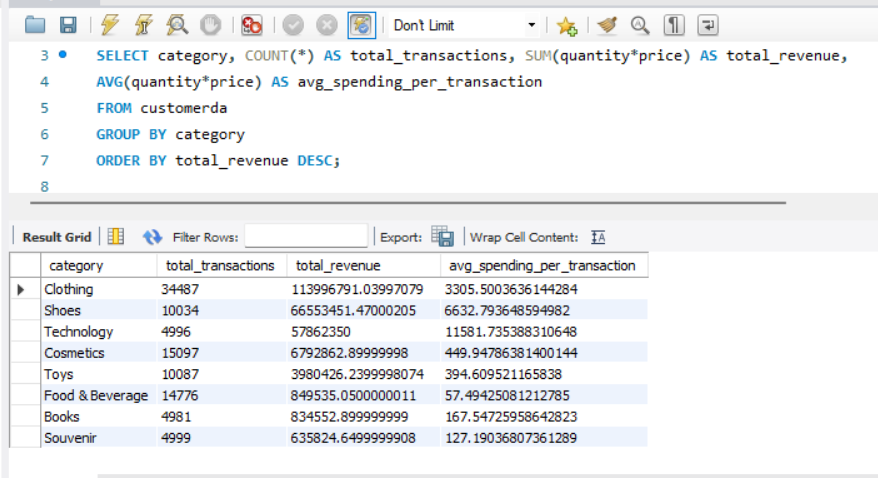
SELECT category, COUNT (\*) AS total\_transactions, SUM (quantity\*price) AS total\_revenue,

AVG (quantity\*price) AS avg\_spending\_per\_transaction

FROM customerda

GROUP BY category

ORDER BY total\_revenue DESC;



1. Age wise group:

SELECT \* FROM analysis.customerda;

SELECT

CASE

WHEN age BETWEEN 18 AND 24 THEN '18-24'

WHEN age BETWEEN 25 AND 34 THEN '25-34'

WHEN age BETWEEN 35 AND 44 THEN '35-44'

WHEN age BETWEEN 45 AND 54 THEN '45-54'

WHEN age > 54 THEN '54+'

END AS age\_group,

COUNT (\*) AS total\_customers,

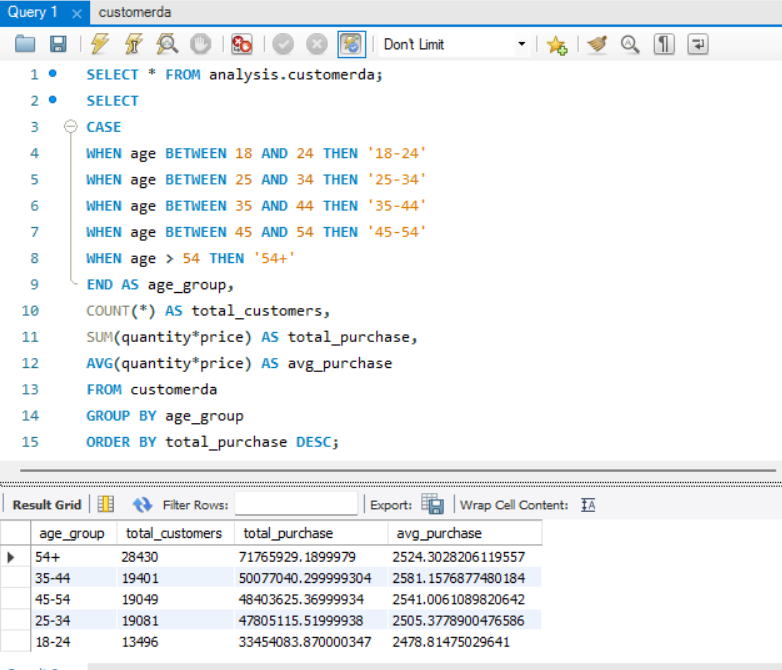
SUM(quantity\*price) AS total\_purchase,

AVG (quantity\*price) AS avg\_purchase

FROM customerda

GROUP BY age\_group

ORDER BY total\_purchase DESC;



1. Payment method:

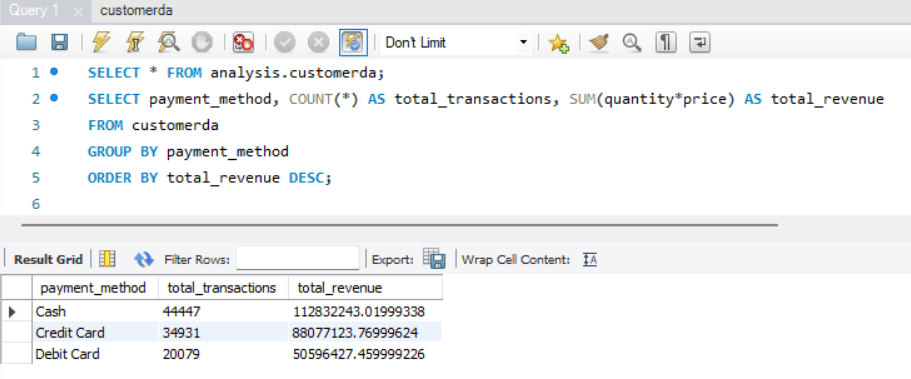
SELECT \* FROM analysis.customerda;

SELECT payment\_method, COUNT (\*) AS total\_transactions, SUM (quantity\*price) AS total\_revenue

FROM customerda

GROUP BY payment\_method

ORDER BY total\_revenue DESC;



1. Shopping mall:

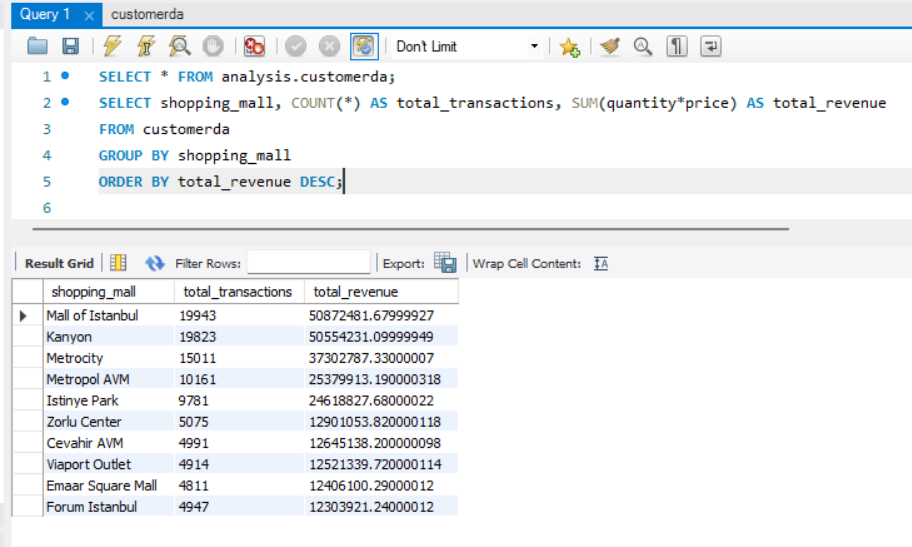
SELECT \* FROM analysis.customerda;

SELECT shopping\_mall, COUNT (\*) AS total\_transactions, SUM (quantity\*price) AS total\_revenue

FROM customerda

GROUP BY shopping\_mall

ORDER BY total\_revenue DESC;



Business insights:

Spending patterns vary across customer segments and product categories, revealing clear opportunities for targeted marketing and smarter inventory planning. These insights help businesses focus on the right customers with the right products to boost sales.

**Q.5. How is the shopping distribution according to age group?**

What to look for:

* Identify which age group is making the highest number of purchases and spending the most.
* Understand which age segment has higher average purchase value, helping in targeted promotions.

Description of query:

This query categorizes customers into age groups, then calculates the total transactions, total spending, and average purchase value for each group to compare their shopping behavior.

Query:

SELECT \* FROM analysis.customerda;

SELECT

CASE

WHEN age BETWEEN 18 AND 24 THEN '18-24'

WHEN age BETWEEN 25 AND 34 THEN '25-34'

WHEN age BETWEEN 35 AND 44 THEN '35-44'

WHEN age BETWEEN 45 AND 54 THEN '45-54'

ELSE '55+'

END AS age\_group,

COUNT (\*) AS total\_transactions,

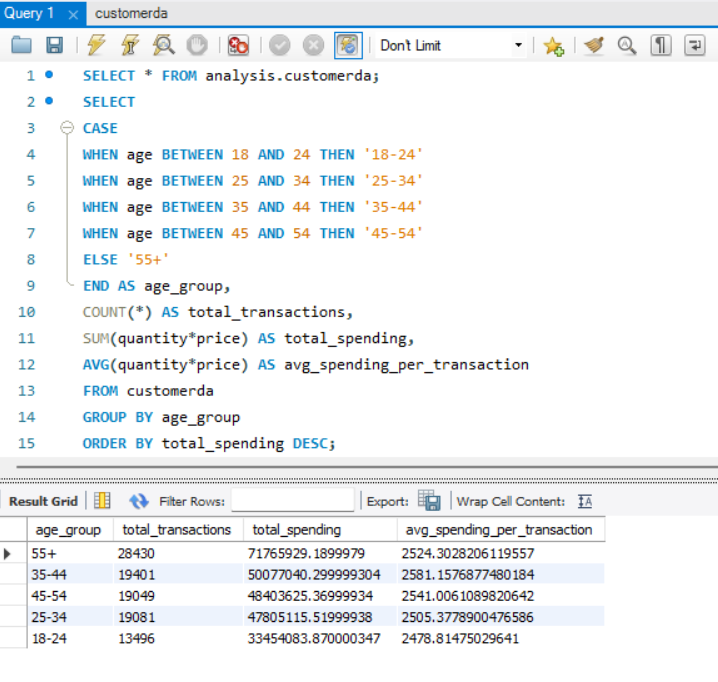
SUM (quantity\*price) AS total\_spending,

AVG (quantity\*price) AS avg\_spending\_per\_transaction

FROM customerda

GROUP BY age\_group

ORDER BY total\_spending DESC;



Business insights:

Older customers, especially those aged 55+, contribute the highest total spending, while the 35–54 age group also shows strong purchasing power. This indicates that premium product offerings, loyalty programs, and personalized promotions for these age groups can significantly boost revenue.

**Q.6. Which age category did we sell more products to?**

What we look for:

* . Which age group had highest numbers of products sales
* Which age group prefers to purchase which category?

First approach:

Description of query:

We first divide the customer’s age into groups (age categories) and then, we calculate the total number of products sold within each age groups.

Query:

SELECT \* FROM analysis.customerda;

SELECT

CASE

WHEN age BETWEEN 18 AND 24 THEN '18-24'

WHEN age BETWEEN 25 AND 34 THEN '25-34'

WHEN age BETWEEN 35 AND 44 THEN '35-44'

WHEN age BETWEEN 45 AND 54 THEN '45-54'

WHEN age >= 55 THEN '55+'

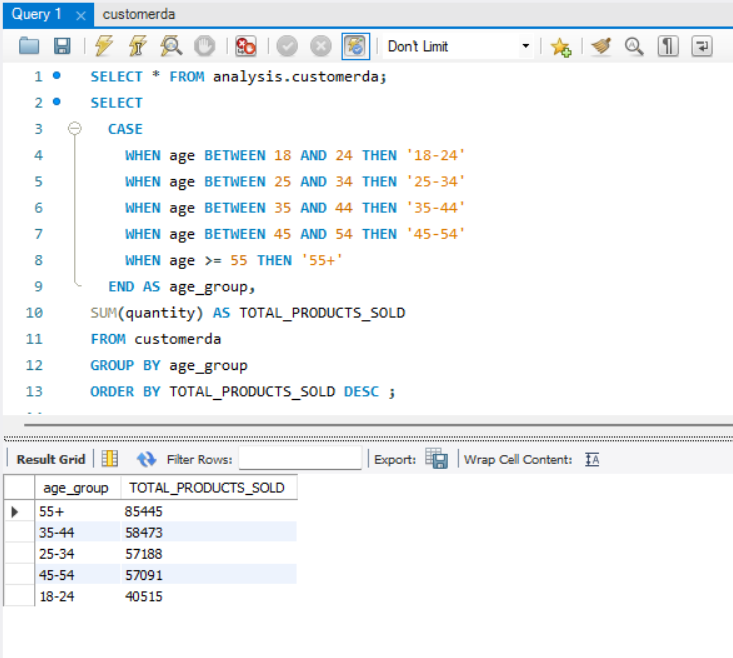
END AS age\_group,

SUM (quantity) AS TOTAL\_PRODUCTS\_SOLD

FROM customerda

GROUP BY age\_group

ORDER BY TOTAL\_PRODUCTS\_SOLD DESC;



Second approach:

Description of the query:

We divide the customers in age groups and then find the total quantity of products sold per category for each age group. This helps identify which product category is most preferred by each age segment.

Query:

SELECT \* FROM analysis.customerda;

SELECT

CASE

WHEN age BETWEEN 18 AND 24 THEN '18-24'

WHEN age BETWEEN 25 AND 34 THEN '25-34'

WHEN age BETWEEN 35 AND 44 THEN '35-44'

WHEN age BETWEEN 45 AND 54 THEN '45-54'

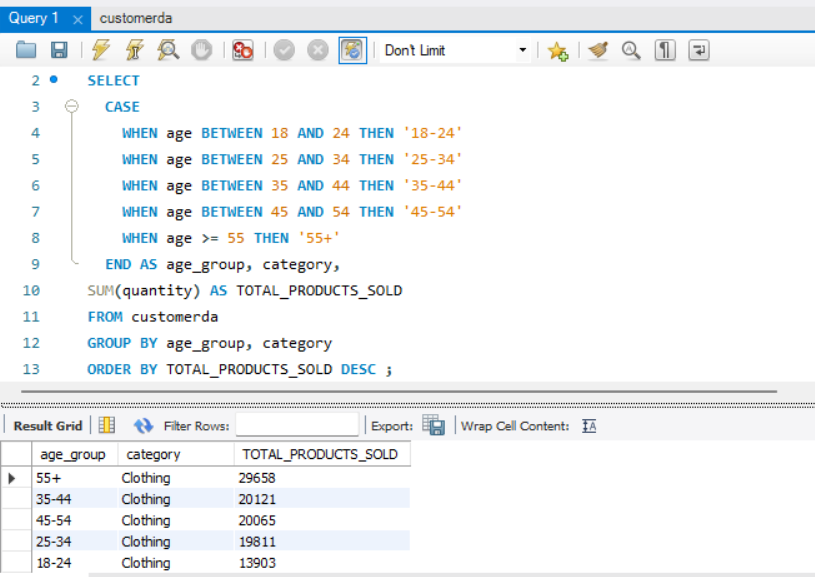
WHEN age >= 55 THEN '55+'

END AS age\_group, category,

SUM (quantity) AS TOTAL\_PRODUCTS\_SOLD

FROM customerda

GROUP BY age\_group, category

ORDER BY TOTAL\_PRODUCTS\_SOLD DESC;  
Business insights:  
•In first approach, the highest number of products were sold in the 55+ age group.

•In second approach, we analyzed clothing category was most purchased by each age group.

**Q.7. Which age cat generated more revenue?**

What we look for:

* Which age group spent the most money on purchases and generated more revenue.

Description of the query:

This query groups customers into age ranges and calculates the total revenue contributed by each age group to identify which segment spends the most.

Query:

SELECT \* FROM analysis.customerda;

SELECT

CASE

WHEN age BETWEEN 18 AND 24 THEN '18-24'

WHEN age BETWEEN 25 AND 34 THEN '25-34'

WHEN age BETWEEN 35 AND 44 THEN '35-44'

WHEN age BETWEEN 45 AND 54 THEN '45-54'

WHEN age >= 55 THEN '55+'

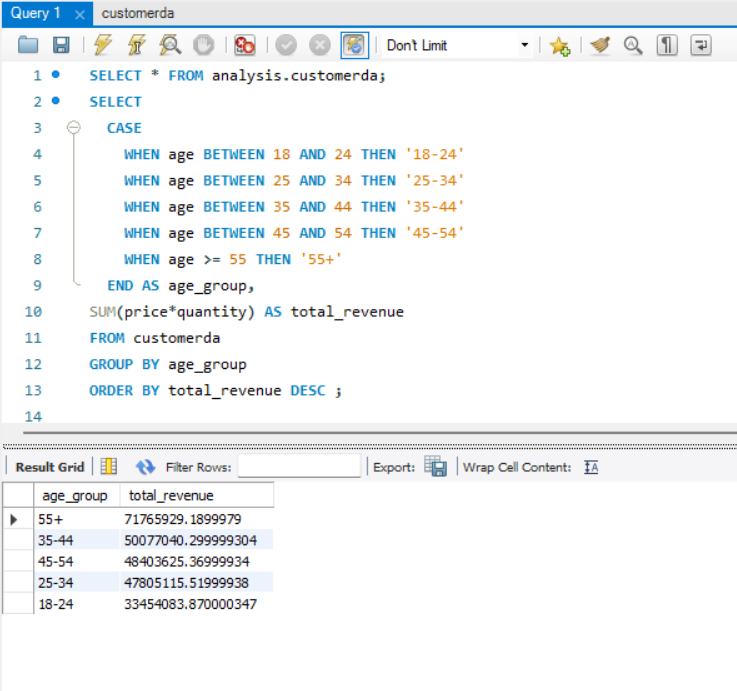
END AS age\_group,

SUM (price\*quantity) AS total\_revenue

FROM customerda

GROUP BY age\_group

ORDER BY total\_revenue DESC;



Business insights:

• The 55+ age group helps generate the highest revenue among all age categories.

**Q.8. Distribution of purchase categories relative to other columns?**

What we look for:

Category popularity shifts across malls by revenue.

Description of query:

This query compares the revenue generated from different product categories across each shopping mall. It helps identify which mall performs best in specific categories and highlights category-wise strengths or gaps. These insights can guide decisions on product placement, store promotions, and mall-specific sales strategies.

Query:

SELECT shopping\_mall,

SUM (CASE WHEN category = 'Clothing' THEN revenue ELSE 0 END) AS Clothing,

SUM (CASE WHEN category = 'Shoes' THEN revenue ELSE 0 END) AS Shoes,

SUM (CASE WHEN category = 'Technology' THEN revenue ELSE 0 END) AS Technology,

SUM (CASE WHEN category = 'Cosmetics' THEN revenue ELSE 0 END) AS Cosmetics,

SUM (CASE WHEN category = 'Toys' THEN revenue ELSE 0 END) AS Toys,

SUM (CASE WHEN category = 'Books' THEN revenue ELSE 0 END) AS Books,

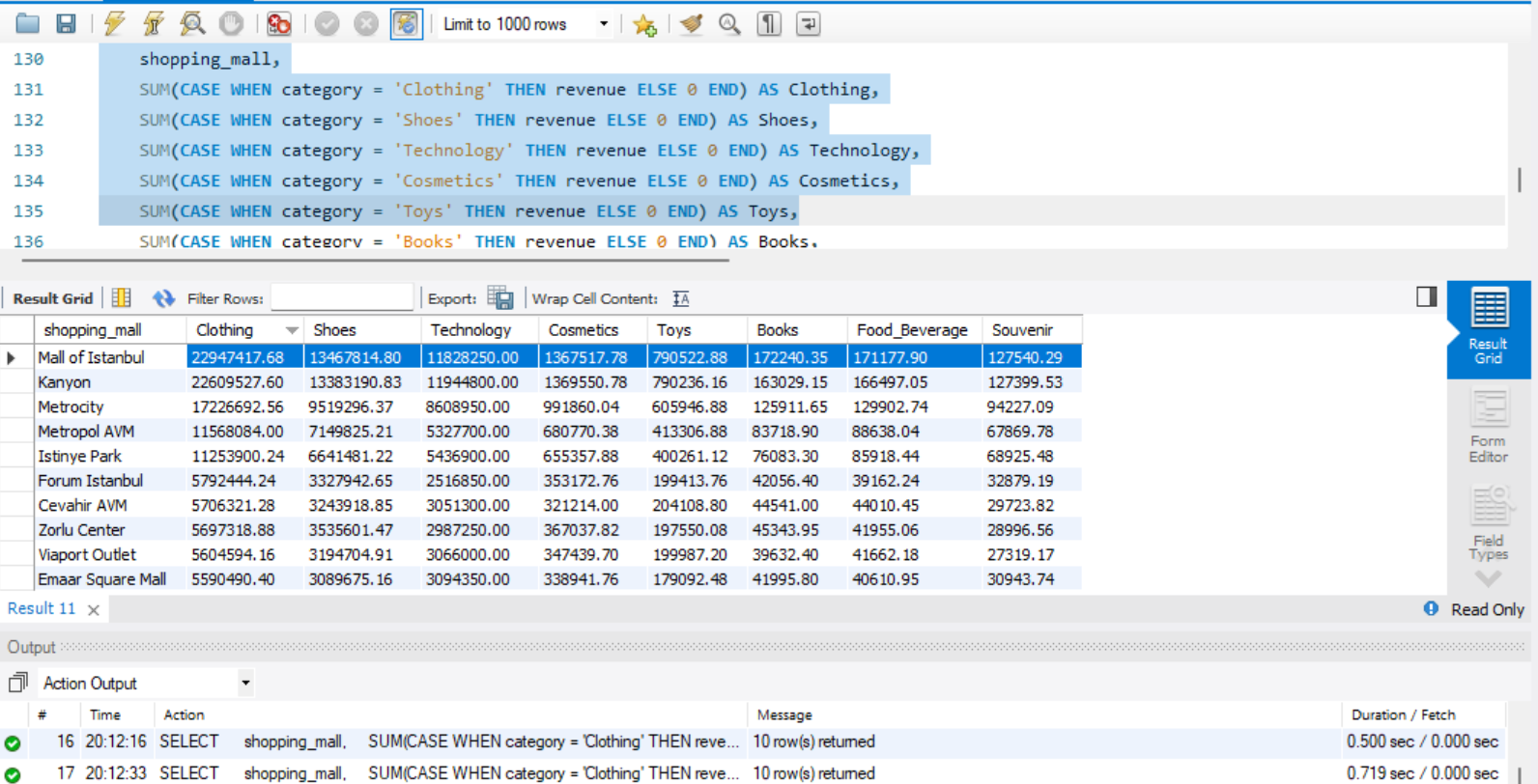
SUM (CASE WHEN category = 'Food & Beverage' THEN revenue ELSE 0 END) AS Food\_Beverage,

SUM (CASE WHEN category = 'Souvenir' THEN revenue ELSE 0 END) AS Souvenir

FROM project.cda\_copy2

GROUP BY shopping\_mall

ORDER BY shopping\_mall;



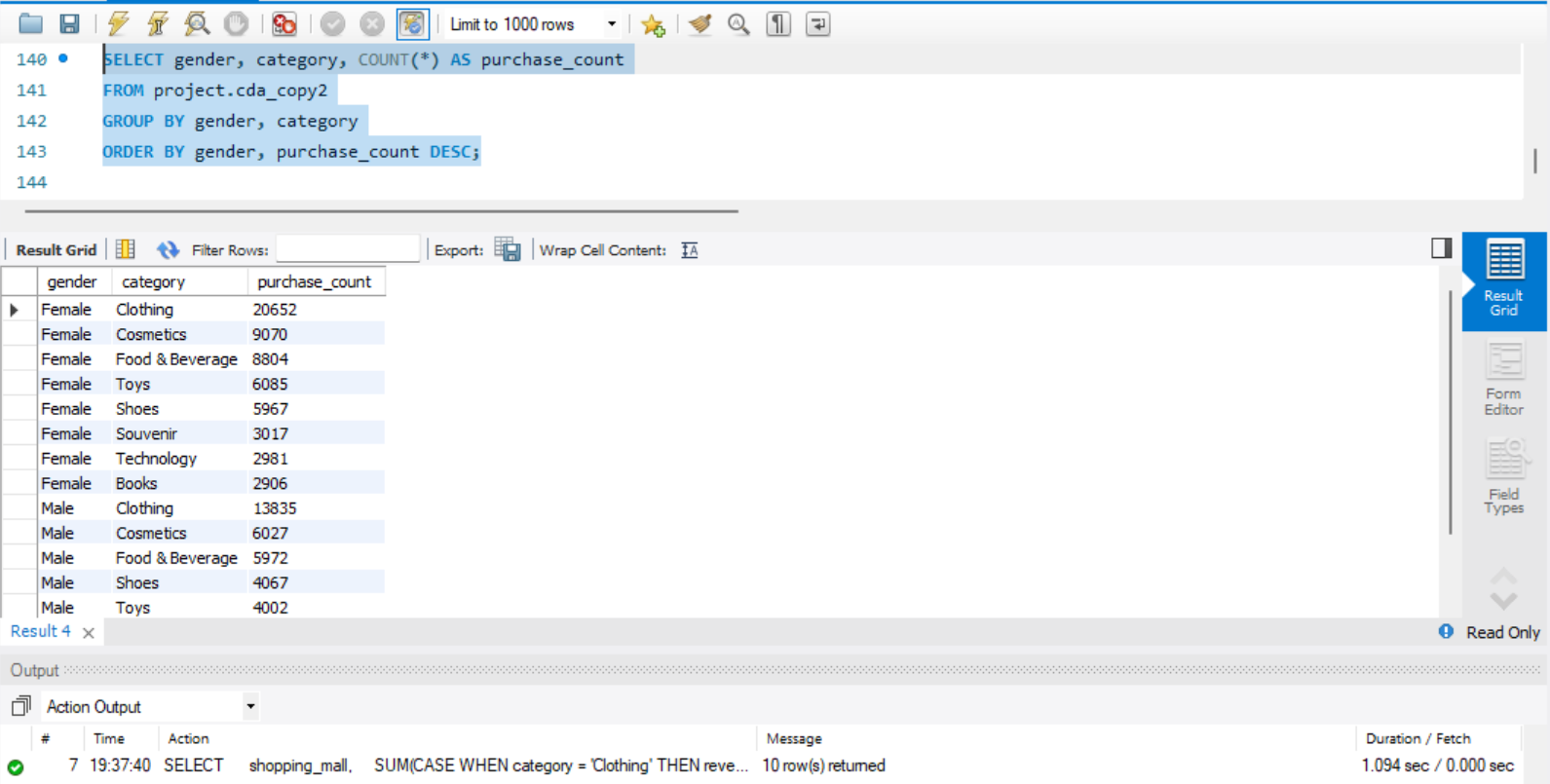
Most Popular Category by Gender:

SELECT gender, category, COUNT (\*) AS purchase\_count

FROM project.cda\_copy2

GROUP BY gender, category

ORDER BY gender, purchase\_count DESC;

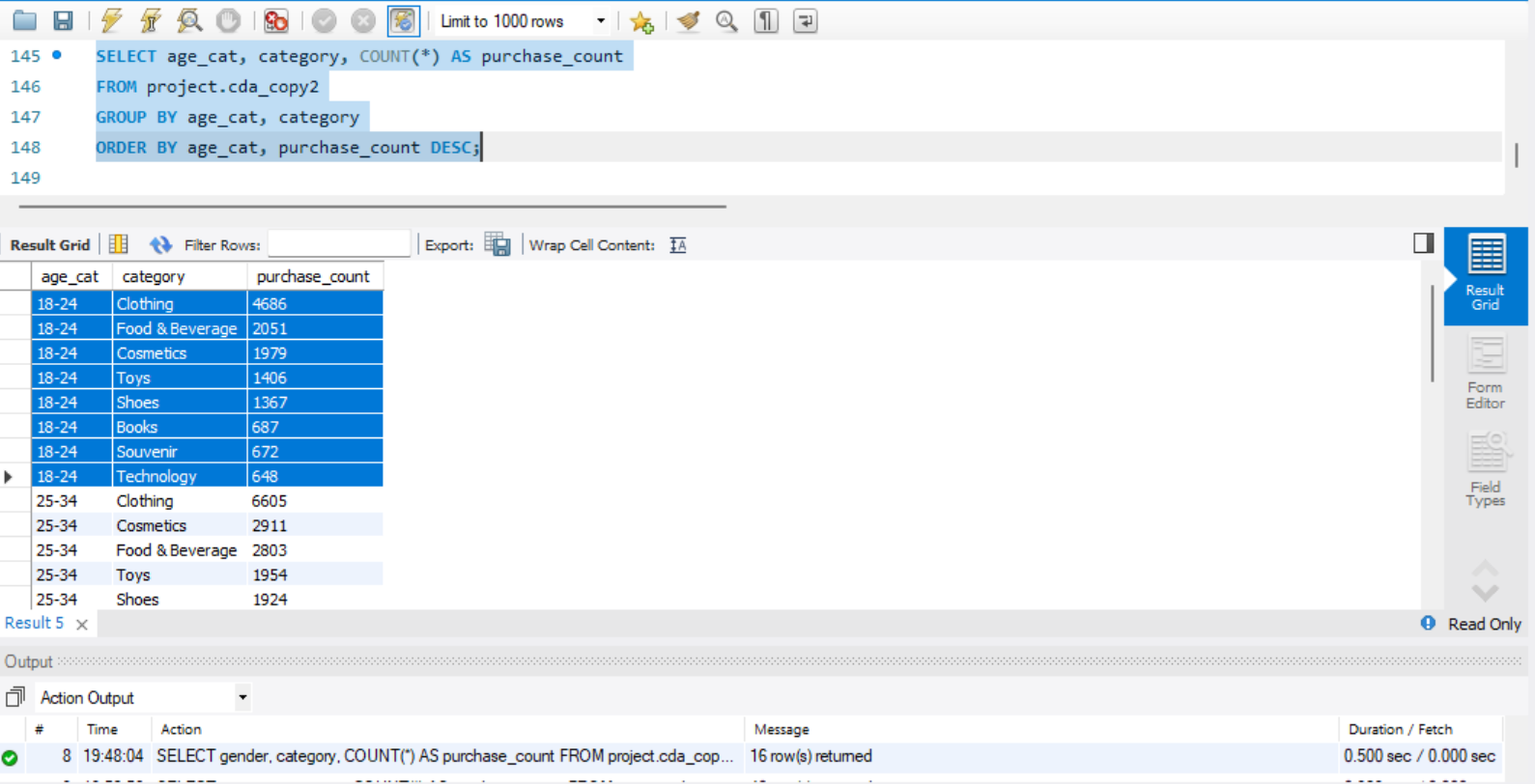


Most Purchased Category by Specific Age Group  
  
SELECT age\_cat, category, COUNT (\*) AS purchase\_count

FROM project.cda\_copy2

GROUP BY age\_cat, category

ORDER BY age\_cat, purchase\_count DESC;



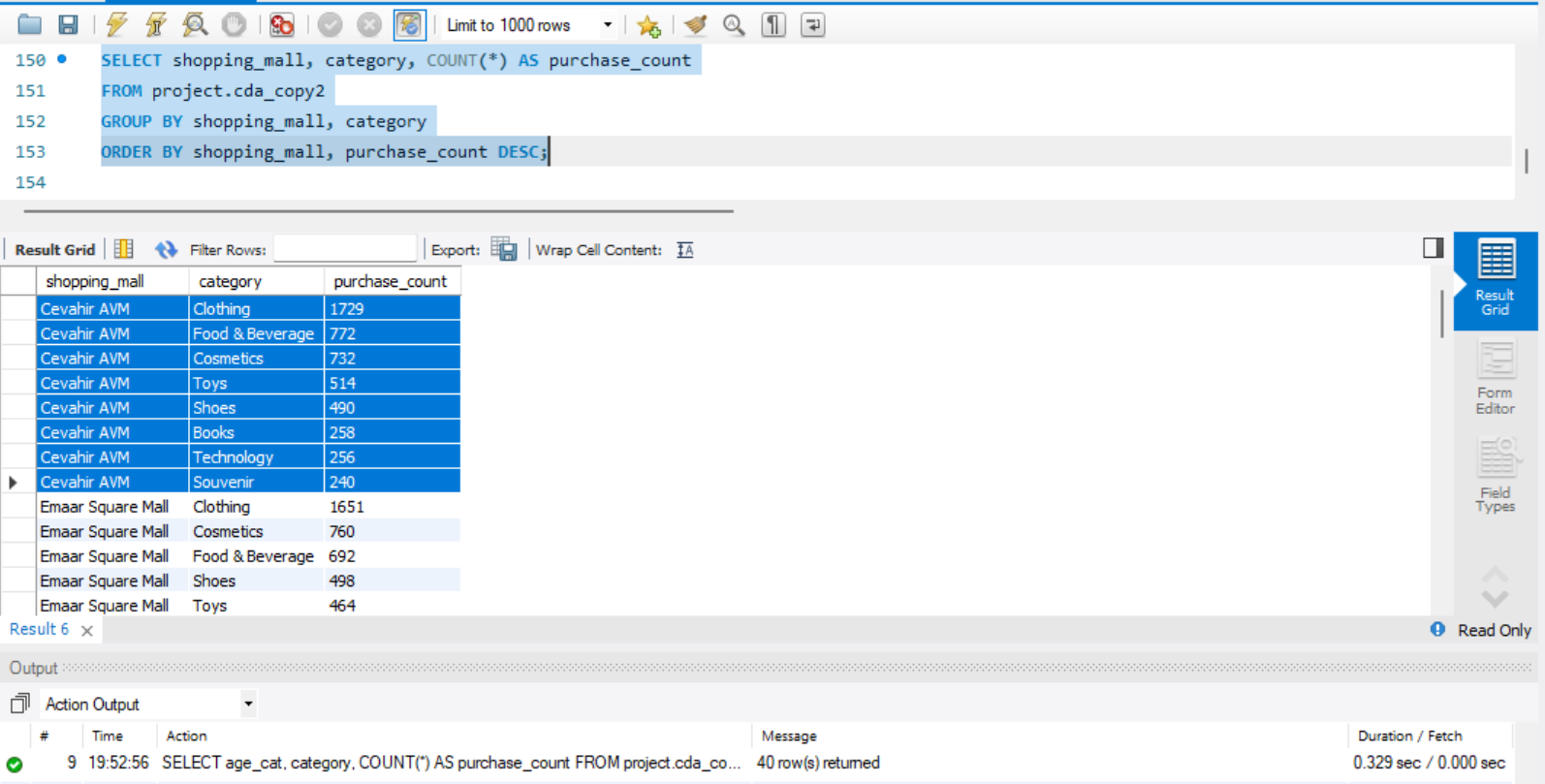
Which Mall Sells More of Each Category

SELECT shopping\_mall, category, COUNT (\*) AS purchase\_count

FROM project.cda\_copy2

GROUP BY shopping\_mall, category

ORDER BY shopping\_mall, purchase\_count DESC;



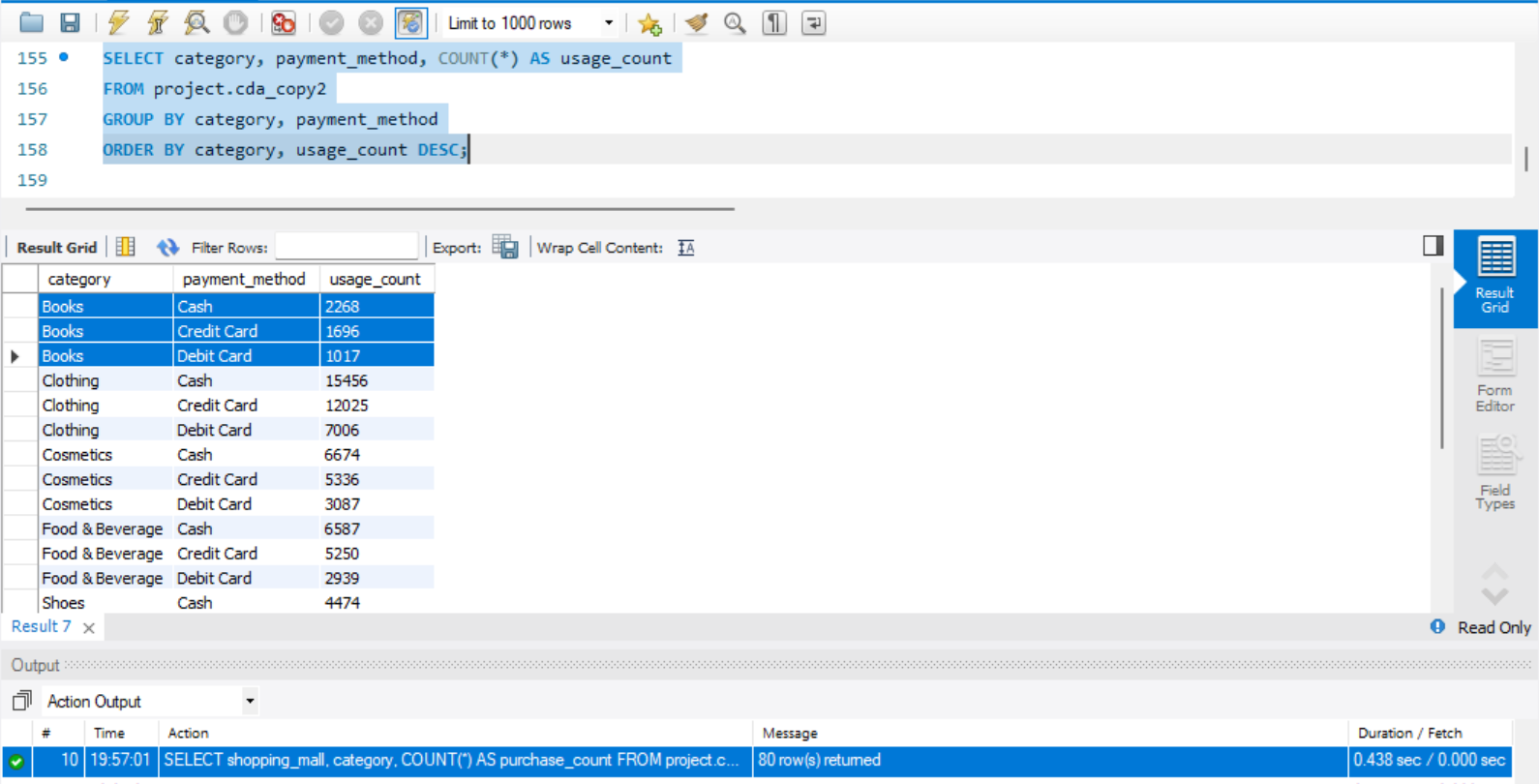
Most Used Payment Method for Each Category

SELECT category, payment\_method, COUNT (\*) AS usage\_count

FROM project.cda\_copy2

GROUP BY category, payment\_method

ORDER BY category, usage\_count DESC;



Business insights

Clothing is the top revenue and purchase category across all malls, age groups, and genders, with cash being the most used payment in every category

**Q.9. Does the payment method have a relation with other columns?**

What we look for:

1. Payment Method Contingency Table vs Gender

2. Payment Method Contingency Table vs Age Category

3. Payment Method Contingency Table vs Category

4. Payment Method Contingency Table vs Shopping Mall

5. Compare Average Revenue by Payment Method.

Query:

SELECT gender,

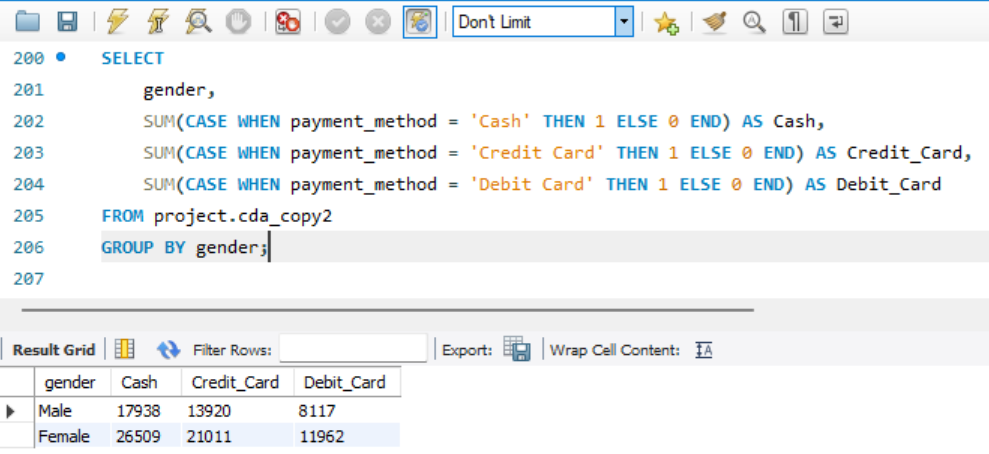
SUM (CASE WHEN payment\_method = 'Cash' THEN 1 ELSE 0 END) AS Cash,

SUM (CASE WHEN payment\_method = 'Credit Card' THEN 1 ELSE 0 END) AS Credit\_Card,

SUM (CASE WHEN payment\_method = 'Debit Card' THEN 1 ELSE 0 END) AS Debit\_Card

FROM project.cda\_copy2

GROUP BY gender;



Payment Method Contingency Table vs Age Category

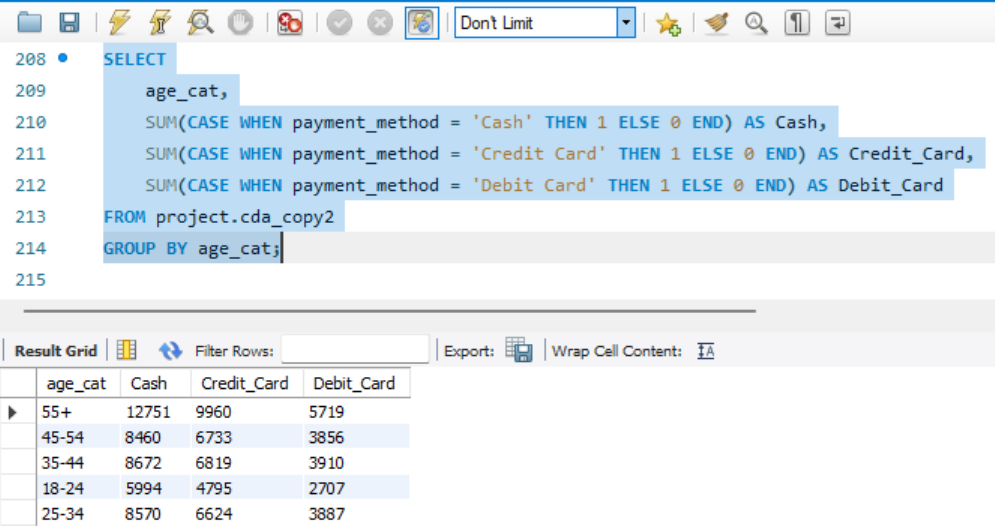
SELECT age\_cat,

SUM (CASE WHEN payment\_method = 'Cash' THEN 1 ELSE 0 END) AS Cash, SUM (CASE WHEN payment\_method = 'Credit Card' THEN 1 ELSE 0 END) AS Credit\_Card

SUM (CASE WHEN payment\_method = 'Debit Card' THEN 1 ELSE 0 END) AS Debit\_Card

FROM project.cda\_copy2

GROUP BY age\_cat;



Payment Method Contingency Table vs Category

SELECT

category,

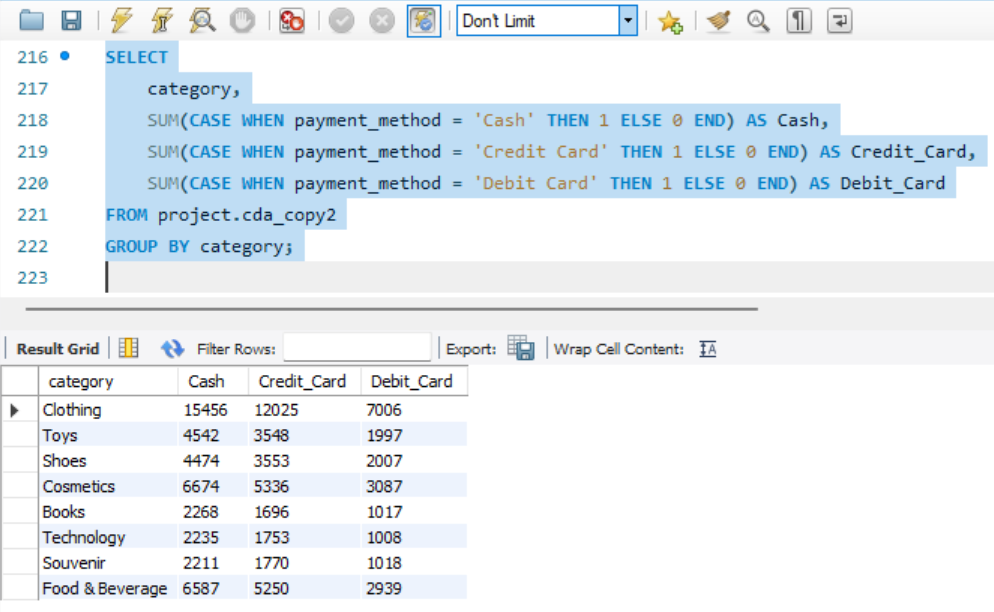
SUM (CASE WHEN payment\_method = 'Cash' THEN 1 ELSE 0 END) AS Cash,

SUM (CASE WHEN payment\_method = 'Credit Card' THEN 1 ELSE 0 END) AS Credit\_Card,

SUM (CASE WHEN payment\_method = 'Debit Card' THEN 1 ELSE 0 END) AS Debit\_Card

FROM project.cda\_copy2

GROUP BY category;



Payment Method Contingency Table vs Shopping Mall

SELECT

shopping\_mall,

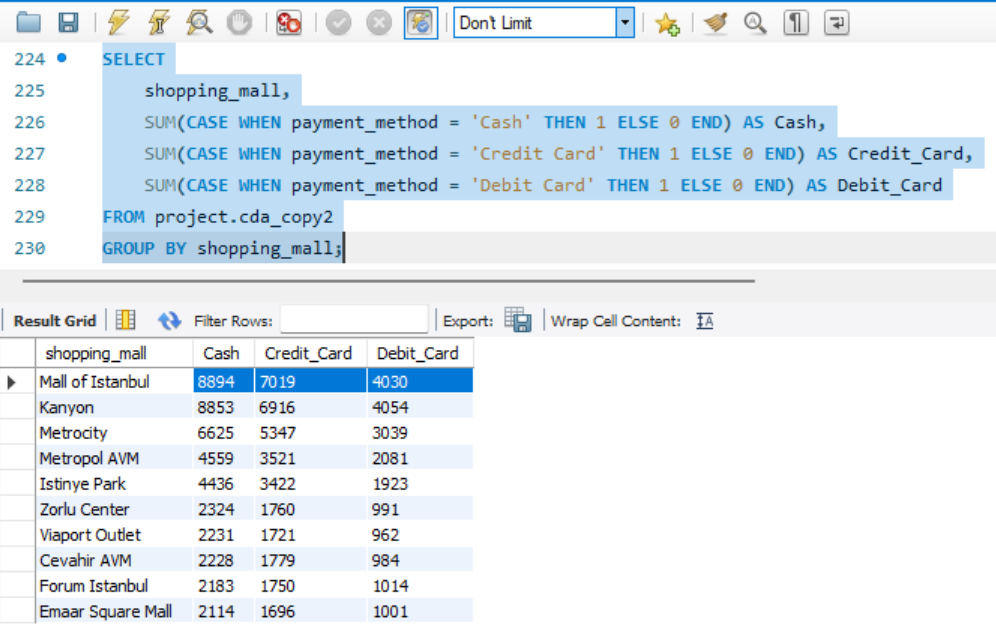
SUM (CASE WHEN payment\_method = 'Cash' THEN 1 ELSE 0 END) AS Cash,

SUM (CASE WHEN payment\_method = 'Credit Card' THEN 1 ELSE 0 END) AS Credit\_Card,

SUM (CASE WHEN payment\_method = 'Debit Card' THEN 1 ELSE 0 END) AS Debit\_Card

FROM project.cda\_copy2

GROUP BY shopping\_mall;



Compare Average Revenue by Payment Method

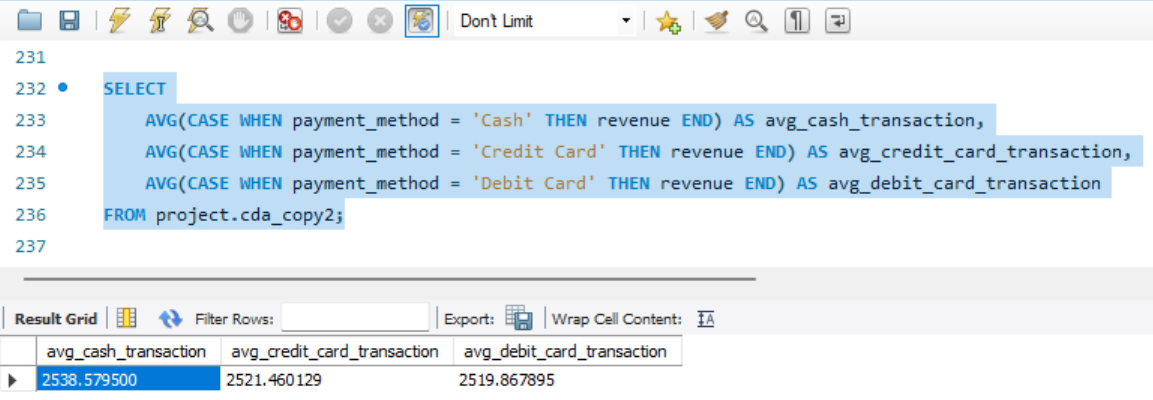
SELECT

AVG (CASE WHEN payment\_method = 'Cash' THEN revenue END) AS avg\_cash\_transaction,

AVG (CASE WHEN payment\_method = 'Credit Card' THEN revenue END) AS avg\_credit\_card\_transaction,

AVG (CASE WHEN payment\_method = 'Debit Card' THEN revenue END) AS avg\_debit\_card\_transaction

FROM project.cda\_copy2;



Business insights:  
Cash’s dominance suggests customer preference for immediate, fee-free, and universally accepted transactions, but limits digital record-keeping and remote sales options. If certain payment methods associate with higher spend, consider promoting those payment channels (e.g., credit card offers).

**10. How is the distribution of the payment method?**

What to compute:

counts and share of transactions by payment method also revenue share by payment method avg revenue by payment method.

1. Count and % Share of Transactions by Payment Method

SELECT

payment\_method,

COUNT (\*) AS transaction\_count,

CONCAT (ROUND (COUNT (\*) \* 100.0 / (SELECT COUNT (\*) FROM project.cda\_copy2), 2), '%') AS percent\_share

FROM project.cda\_copy2

GROUP BY payment\_method

ORDER BY transaction\_count DESC;



2. Revenue Share by Payment Method

SELECT

payment\_method,

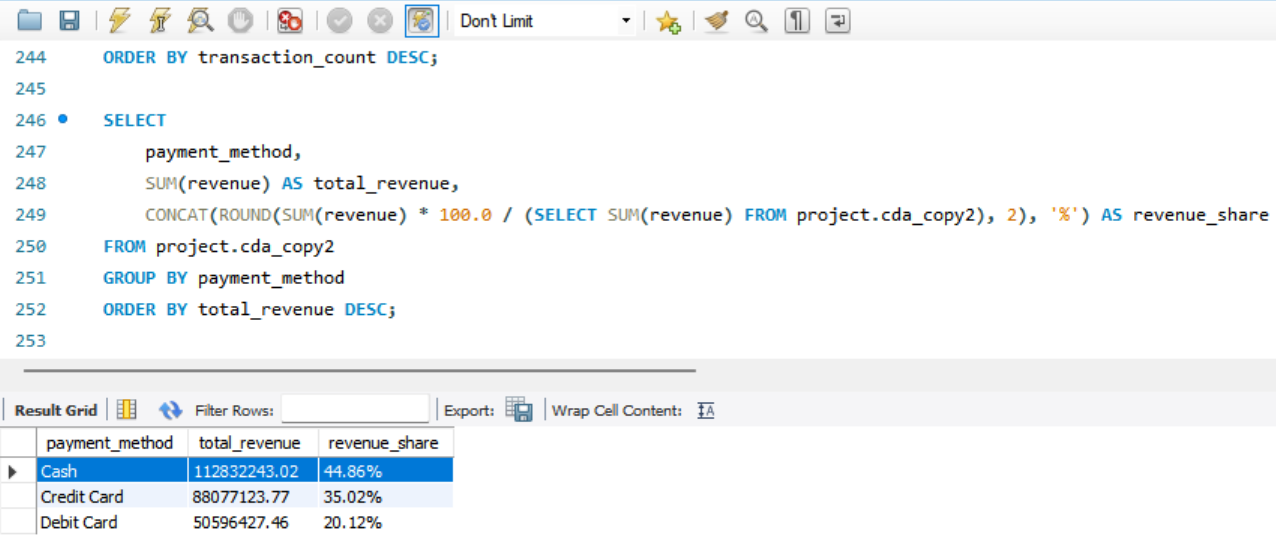
SUM (revenue) AS total\_revenue,

CONCAT (ROUND (SUM (revenue) \* 100.0 / (SELECT SUM (revenue) FROM project.cda\_copy2), 2), '%') AS revenue\_share

FROM project.cda\_copy2

GROUP BY payment\_method

ORDER BY total\_revenue DESC;



3. Average Ticket (Revenue) by Payment Method

SELECT

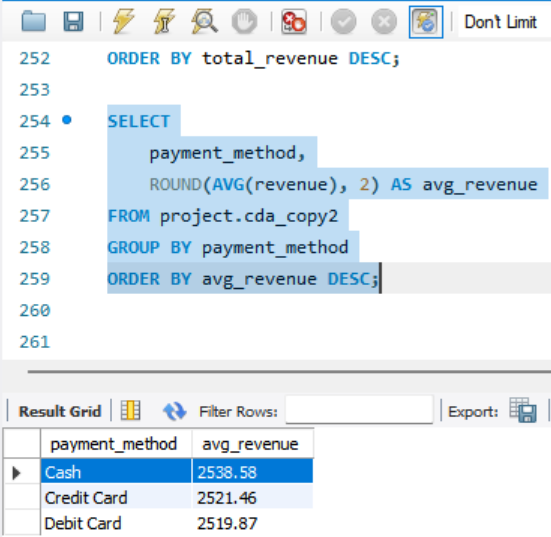
payment\_method,

ROUND (AVG (revenue), 2) AS avg\_revenue

FROM project.cda\_copy2

GROUP BY payment\_method

ORDER BY avg\_revenue DESC;



Business insights:

Cash is the leading payment method, with 45% share in both transaction count and revenue, and also delivers the highest average sale value.

# CHALLENGES DURING EXECUTION

**Challenge-1:**

* **Technical Challenge**  
  Making direct changes in the main database was risky, as multiple operations were linked to the same dataset. Any accidental modification, missing record, or misplaced entry could have led to incorrect analysis and inaccurate results.
* **Data-Related Challenge**  
  The main dataset was large and interconnected, containing 99,457 records across various tables. Performing operations directly on it posed a high risk of data inconsistency or loss, which could impact the validity of the entire analysis.
* **Solution/Approach to Fix It**  
  Initially, a plan was made to create a clone of the dataset within the same database server to ensure safe experimentation. However, due to restricted student authentication access, database creation or cloning was not permitted. To overcome this, the entire dataset was exported and imported into a separate “Project-Operation-Test” server, allowing all required operations to be performed safely without affecting the main database.
* **Learnings from the Challenge**  
  This challenge emphasized the importance of data safety, access control, and version management in analytical projects. It provided practical experience in securely handling databases and taught the value of maintaining a separate test environment to prevent data corruption and ensure accurate analysis.

**Challenge-2:**

* **Technical Challenge:**  
  Frequent MySQL errors occurred during large data operations — notably Error Code 2013: Lost connection to MySQL server during query (30.015 sec), caused by complex queries exceeding the server timeout limit.
* **Solution:**  
  Queries were executed in smaller chunks, repeated until completion, and verified for accuracy, ensuring stable performance without connection loss.
* **Learnings from the Challenge:**  
  Learned the importance of query optimization and efficient resource handling to maintain server stability and ensure smooth execution on large datasets.

**Challenge-3:**

* **Technical Challenge**  
  Large data modifications made it difficult to manage and represent multiple analyses within a single table. Performing additional operations or filters on already filtered tables often led to confusion and data overlap.
* **Solution:**  
  Created multiple sub-tables for different analysis scenarios, each representing specific filtered datasets based on project questions and requirements. This ensured clarity, flexibility, and easier data handling.
* **Learnings from the Challenge**  
  Learned the value of data segregation and modular analysis to maintain accuracy, simplify operations, and improve analytical efficiency.

**Challenge-4:**

* **Technical Challenge:**  
  Finding an optimized solution that could execute queries efficiently and deliver results in minimum time was a major challenge during analysis.
* **Solution:**Applied different SQL approaches for each task, testing and comparing query performance to identify the fastest and most efficient method.
* **Learnings from the Challenge:**Learned the importance of query optimization, logical structuring, and performance testing to enhance execution speed and overall efficiency in SQL-based analysis.

**Challenge-5:**

* **Technical Challenge:**  
  It was challenging to design and implement multiple visuals for various analytical questions. Selecting the right chart type, determining how many visuals to use, and presenting multiple insights clearly on one dashboard required careful planning and creativity.
* **Solution:**  
  To address this, common analytical questions were grouped and divided into three visualization patterns. Different Power BI visuals were tested to identify the most suitable ones, ensuring each dashboard communicated insights effectively and efficiently.
* **Learnings from the Challenge:**  
  These processes improved understanding of data visualization strategy, chart selection, and information balance, emphasizing the importance of clarity, simplicity, and alignment with business objectives.

**Challenge-6:**

**Technical Challenge:**  
Ensuring visually appealing dashboards and a cohesive presentation layout was difficult. Choosing colour combinations, making minor visual adjustments, and organizing PPT segments to include technical details with business context posed significant design challenges.

**Solution:**  
Different colour themes, layouts, and presentation structures were experimented with to enhance the visual appeal and professionalism of dashboards and slides. Each presentation segment was refined to combine visuals with concise technical explanations.

**Learnings from the Challenge:**  
Learned the importance of visual consistency, colour harmony, and storytelling in business reporting. The experience reinforced how thoughtful design and presentation can significantly improve the clarity and impact of analytical insights.

# BUSINESS INSIGHTS

Customer Data Analysis – Business Insights Summary

The analysis of shopping mall customer data revealed key behavioral and sales patterns across demographics, categories, and payment methods. Most revenue is concentrated in a few top-performing malls and product segments, while certain customer groups and locations show limited engagement. Older customers display strong spending power, but younger segments remain underrepresented. Product demand also varies widely by age and gender, leading to uneven stock movement and missed cross-selling opportunities.

**Business Challenges:**

* Sales are highly dependent on limited customer groups and a few malls, increasing business risk.
* Younger customers (18–34) show low purchase frequency, affecting long-term brand loyalty.
* Overreliance on specific product categories reduces revenue diversity.
* Limited digital payment adoption and customer incentives restrict modern buying convenience.
* Seasonal and uneven demand patterns make inventory and pricing decisions complex.

**Future Improvement Strategies:**

* Launch targeted promotions and loyalty programs for younger customers to build lasting engagement.
* Diversify product offerings to balance category-wise revenue and reduce dependency on a few items.
* Encourage digital and contactless payments with cashback or reward schemes to modernize the shopping experience.
* Use personalized marketing based on customer purchase history and demographics to boost conversion.
* Strengthen mall-wise performance strategies, focusing on underperforming malls with localized campaigns and better category mix.

# POWERBI DASHBOARDS

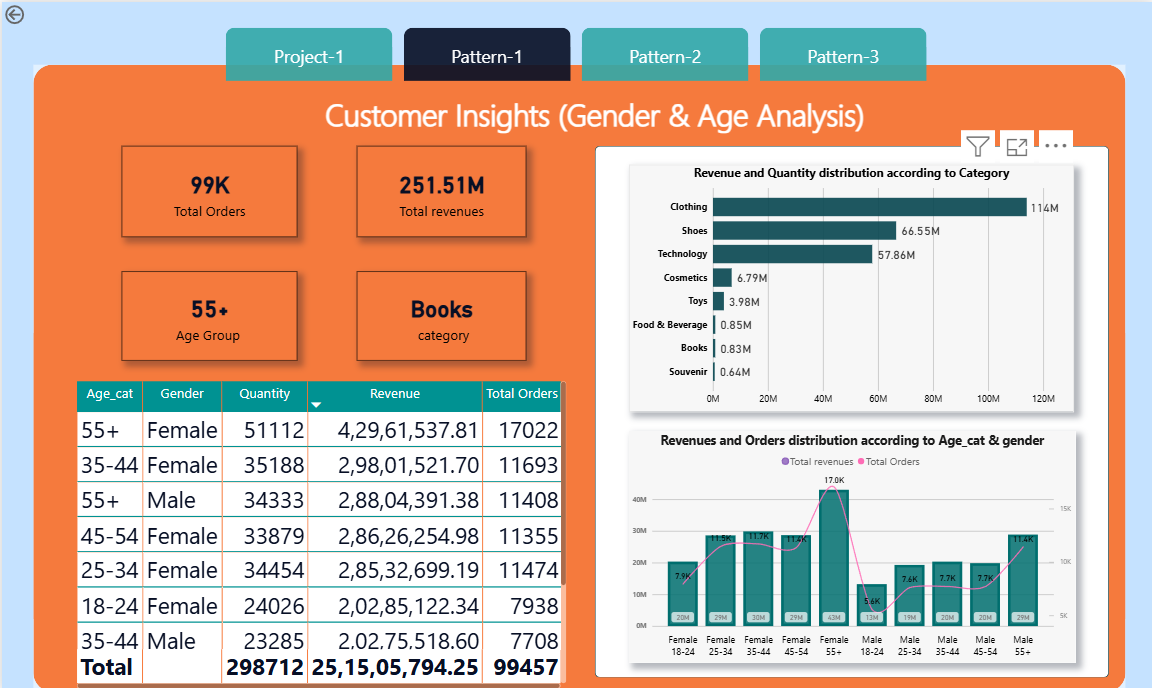
Following are the three dashboards created in Power BI for visualization purpose.

1. **Dashboard-1:**

* **Objective:**

This dashboard aims to analyse customer purchase behaviour based on gender and age groups. It provides insights into total orders, revenue contribution, and category-wise performance.  
The goal is to identify the most profitable customer segments for better targeting and marketing strategies.

* **Screenshot:**

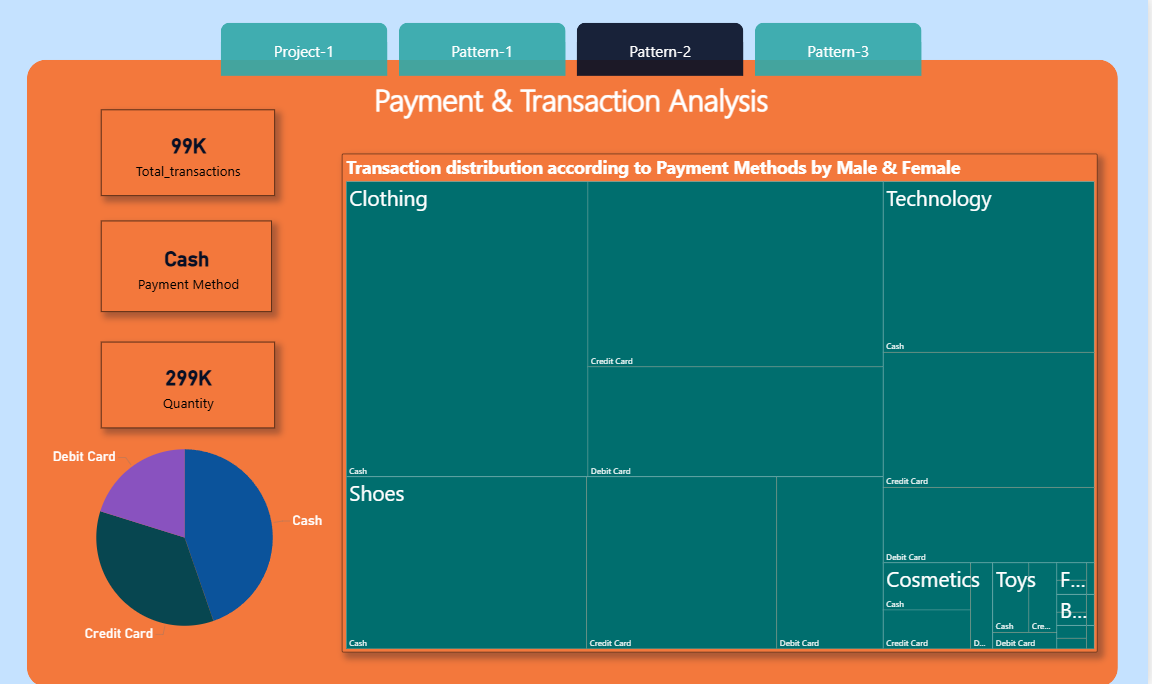


* **Insights from visuals:**

55+ Female customers generate the highest revenue and order volume, making them the most valuable segment. Clothing, Shoes, and Technology are the top revenue-driving categories across all groups. Females consistently outperform males in purchases across most age groups, indicating a stronger buying influence.

1. **Dashboard-2:**

* **Objective:**This dashboard analyses how customers prefer to pay across different product categories. It highlights the share of transactions made via Cash, Credit Card, and Debit Card by both genders. The goal is to understand payment behaviour to plan suitable offers and checkout strategies.
* **Screenshot:**

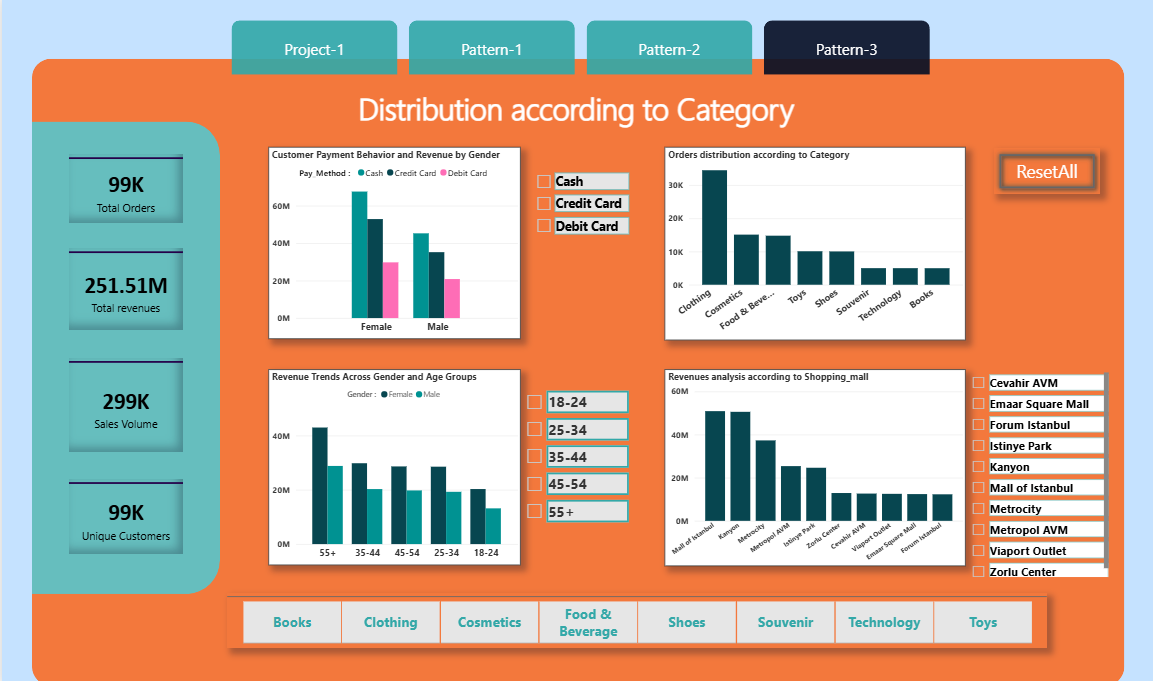


* **Insights from visuals:**

Cash and Credit Card are the most preferred payment modes, especially in high-purchase categories like Clothing and Technology. Debit Card usage remains significantly lower, indicating limited dependency on bank-linked payments.

1. **Dashboard-3:**

* **Dashboard objective:**This dashboard examines how orders and revenues are distributed across product categories, shopping malls, genders, and age groups. It tracks spending trends and payment behaviour across different customer segments.  
  The aim is to identify top-performing categories and high-revenue customer demographics.
* **Screenshot:**



* **Insights from visuals:**Clothing dominates both orders and revenue, with the 55+ and 35-44 age groups being the highest contributors. Mall of Istanbul and Cevahir AVM lead in revenue generation, especially from female customers.

# CONCLUSION

1. Compared to female customers, male customers generate more overall revenue, indicating stronger purchasing activity.

2. The most lucrative market segment is senior customers (55+), who exhibit greater brand loyalty and spending power.

3. Although the use of digital payments is steadily growing, cash payments are still the most popular option.

4. The Pareto principle (80/20 rule) the majority of profits come from a small number of key products, is confirmed by the fact that a few top categories dominate total revenue.

5. Due to their high customer traffic and purchase volume, the Mall of Istanbul and Kanyon stand out as the top revenue generators.

6. Overall, the business shows concentrated revenue sources, by targeting underperforming malls, younger age groups, and female customers, total revenue can grow significantly.