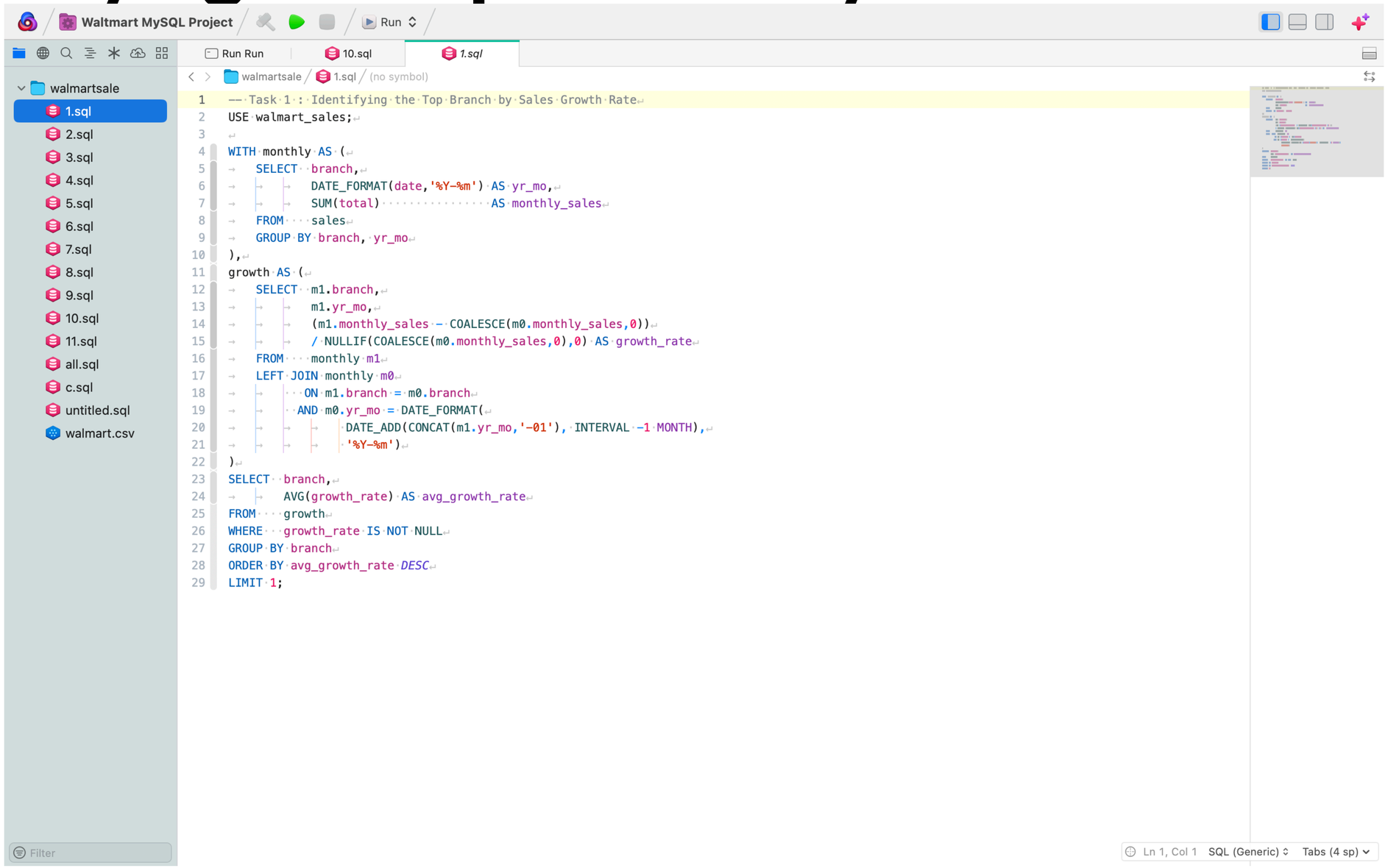
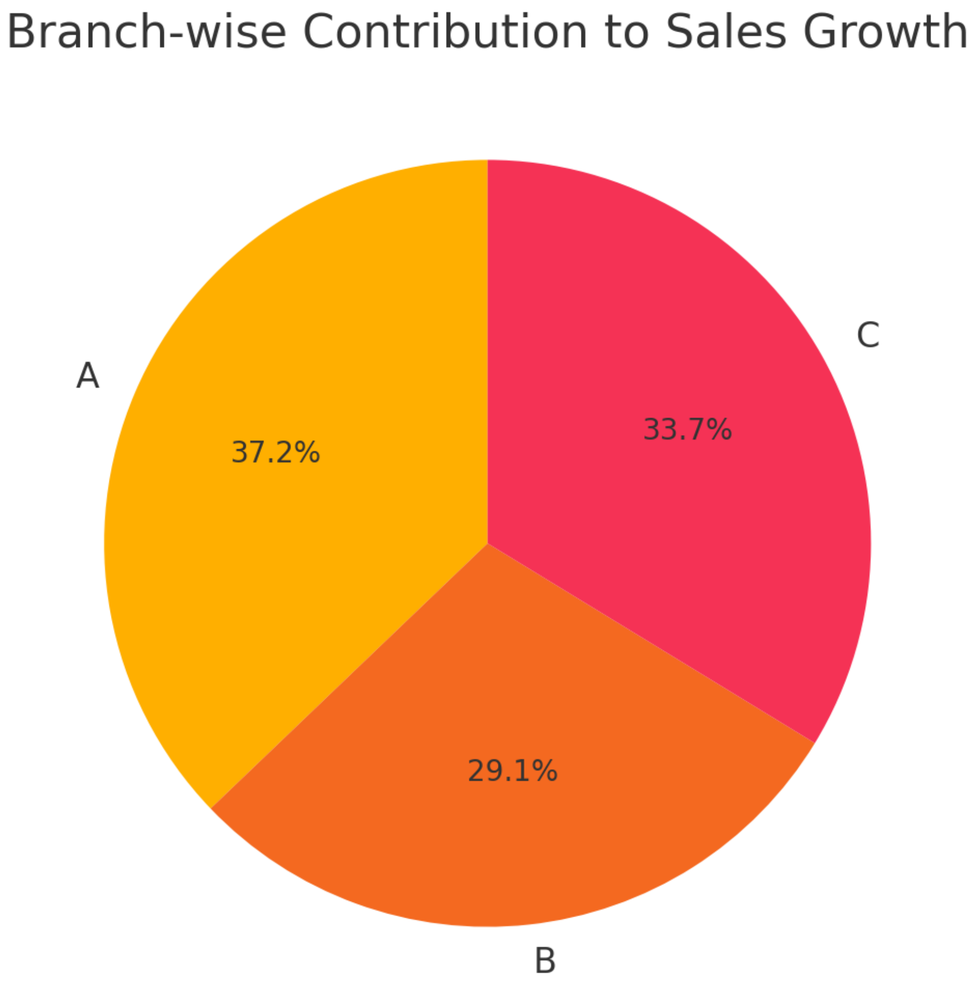
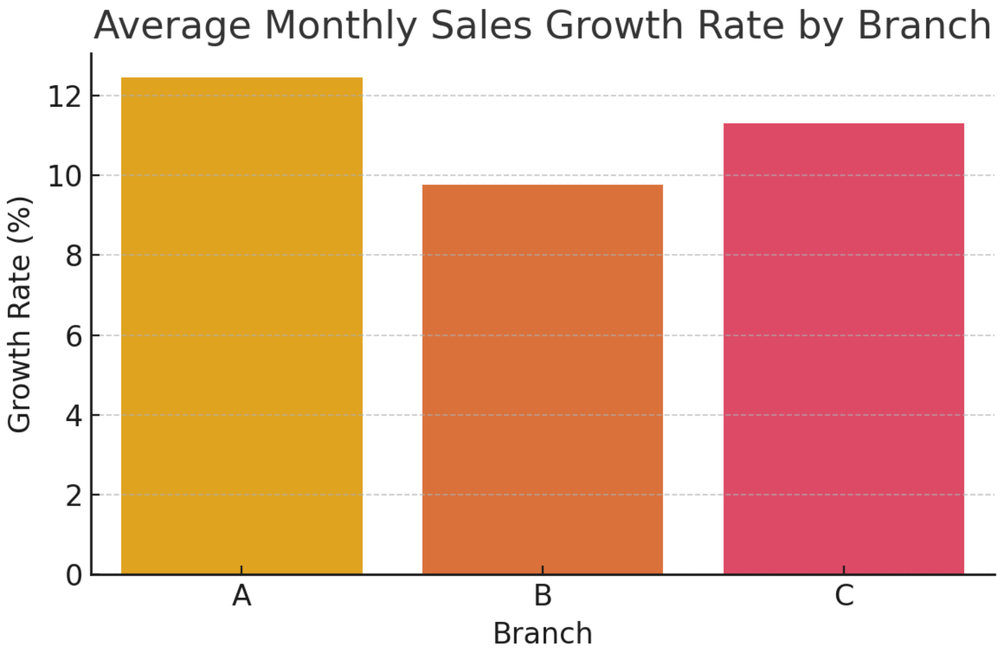
Walmart Sales Performance Analysis: Data Driven Insights Using Advanced SQL Techniques

By, Utkarsh Anand

Identifying the Top Branch by Sales Growth Rate





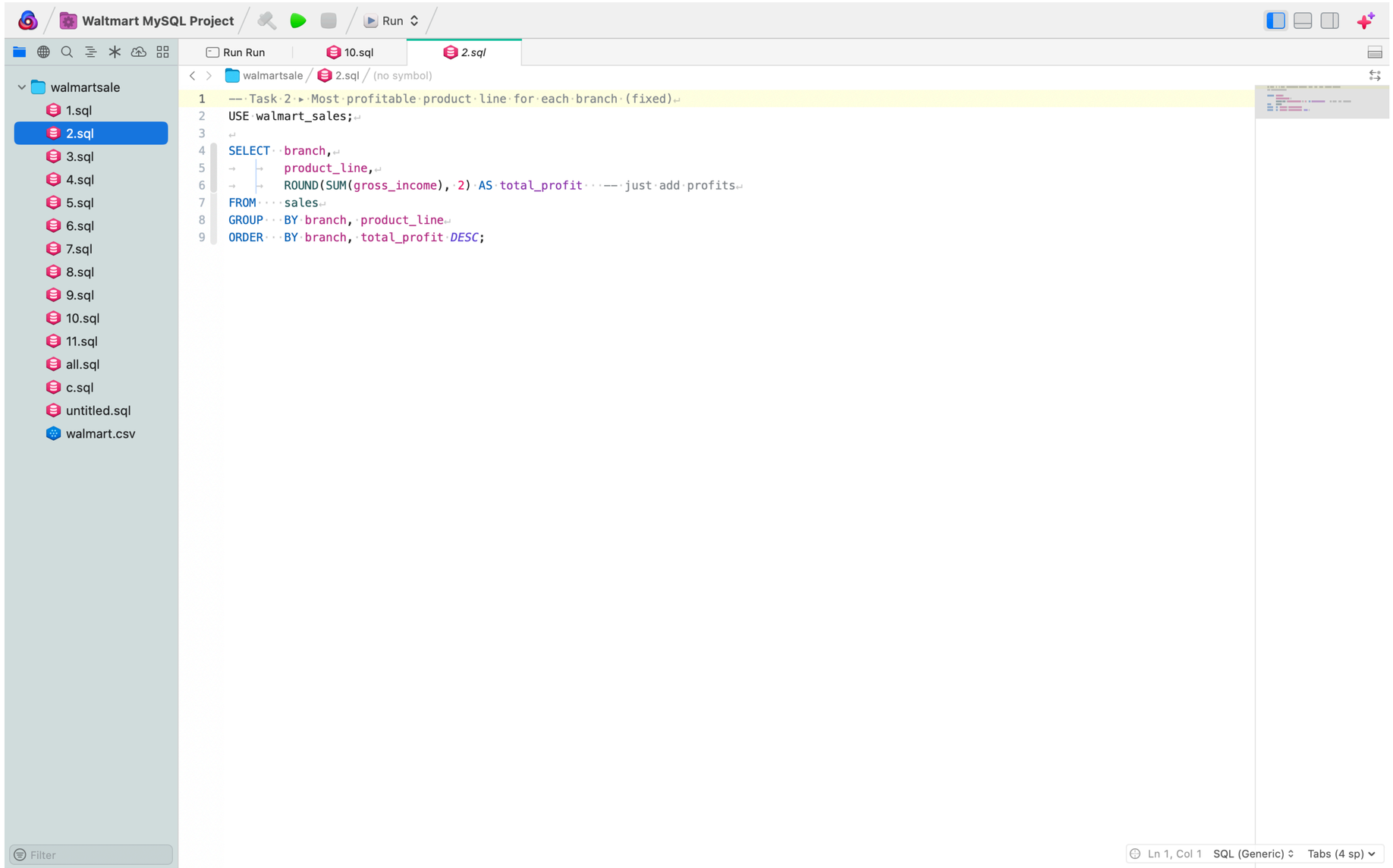


This SQL query aims to **find the branch with the highest average monthly sales growth** across the dataset. It uses a **common table expression (CTE)** to first calculate **monthly sales per branch**, and then compares consecutive months to determine the **growth rate**. By averaging these growth rates, we can rank the branches and select the **top-performing one**.

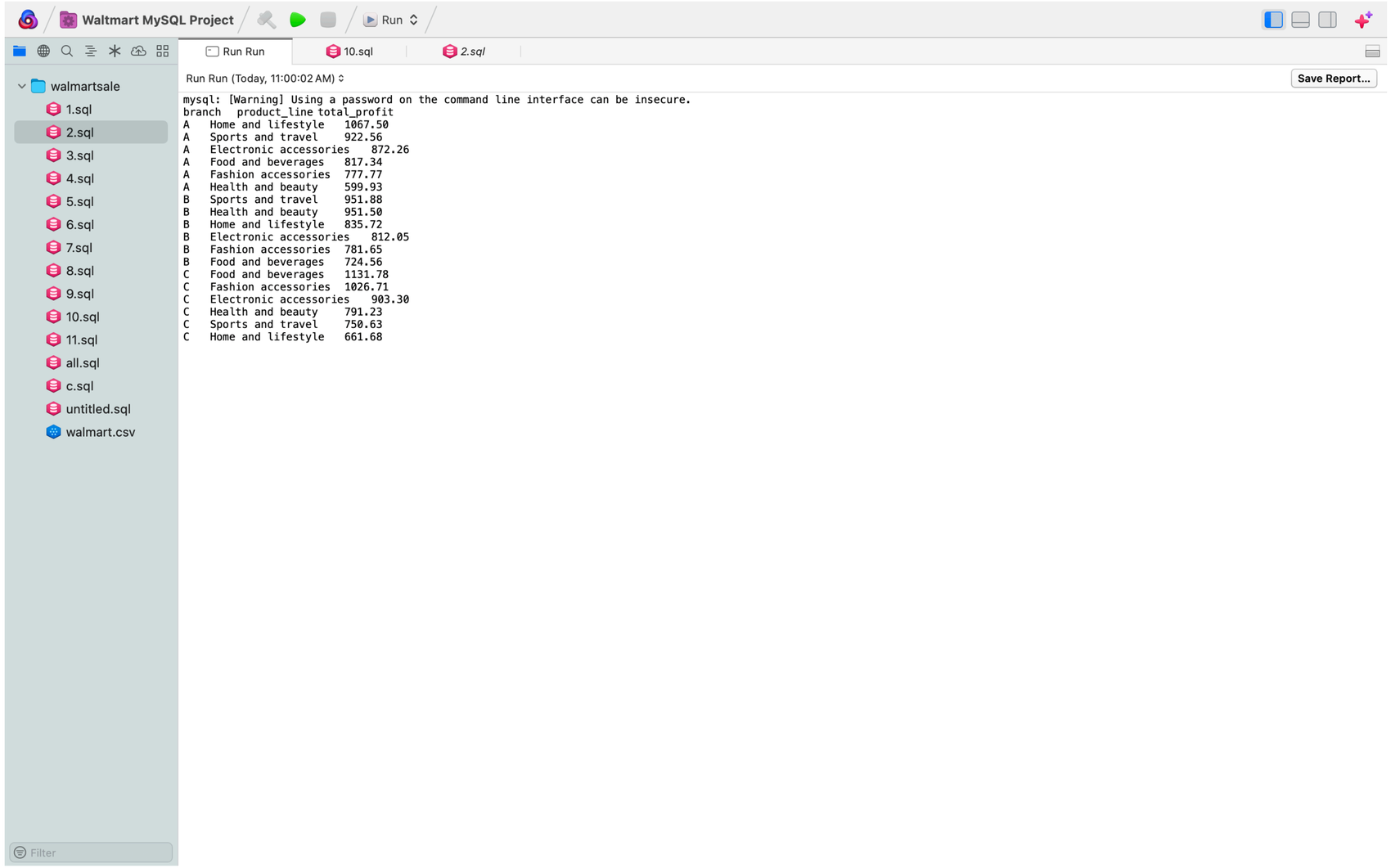
Key Highlights:

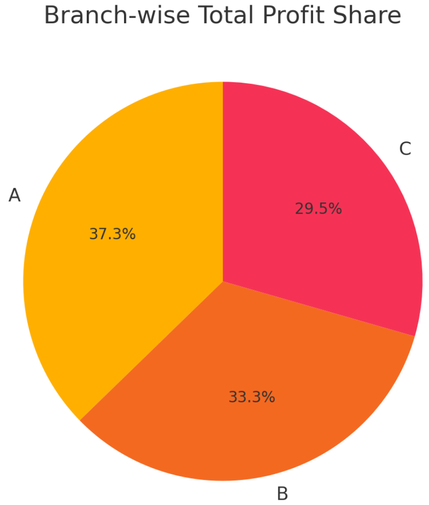
- Uses DATE_FORMAT() to group data month-wise.
- Applies LEFT JOIN to compare a branch's current month sales to the previous month.
- Handles nulls using COALESCE() to avoid calculation errors.
- Ranks branches by AVG(growth_rate) in descending order.
- LIMIT 1 fetches only the **top branch** with highest average growth.

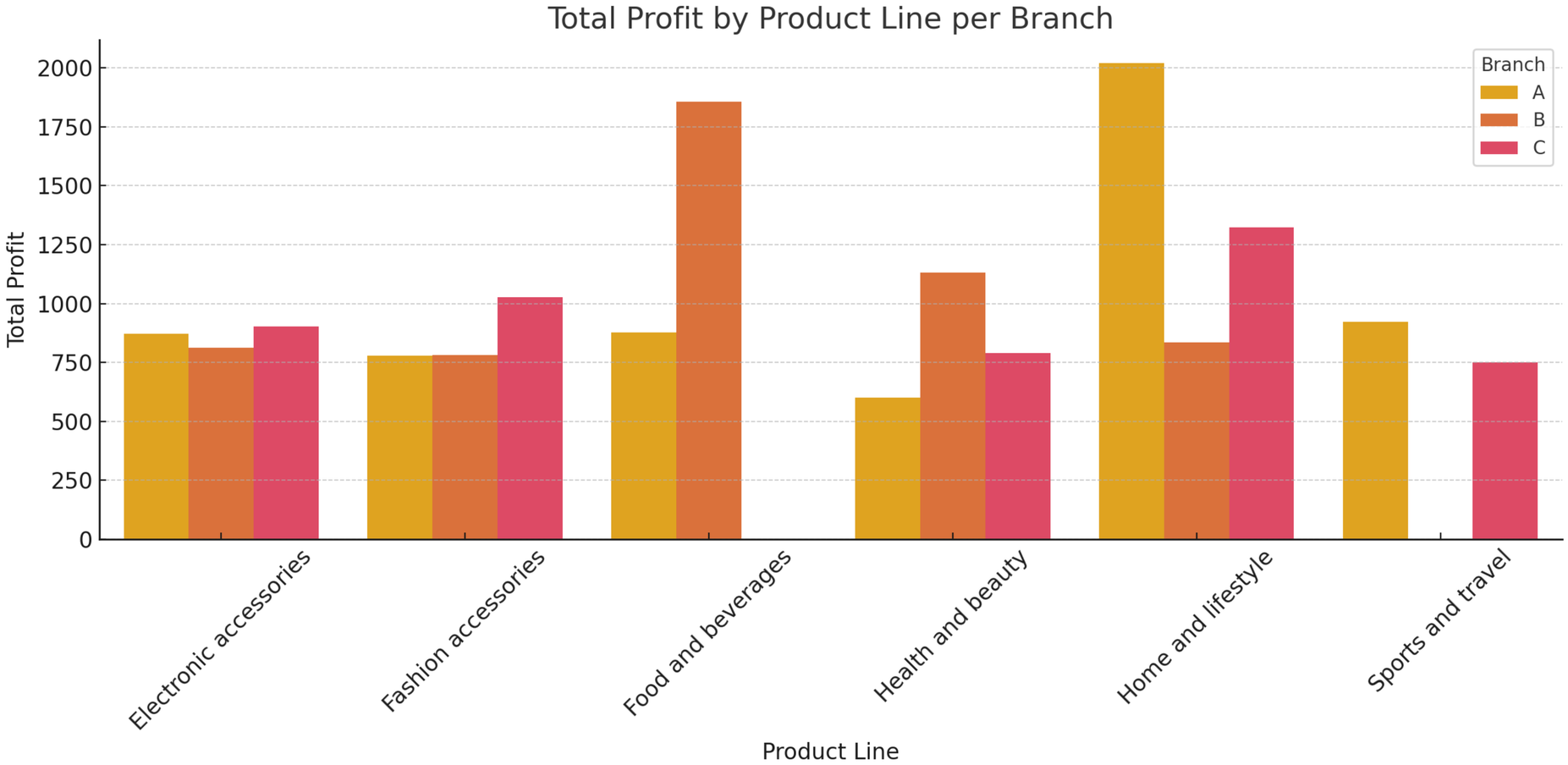
Sales Trend by Gender” or “Gender-wise Purchase Analysis



Output


Waltmart MySQL Project
Run
Run Run
10.sql
2.sql
Run Run (Today, 11:00:02 AM)
Save Report...
walmartsale
1.sql
2.sql
3.sql
4.sql
5.sql
6.sql
7.sql
8.sql
9.sql
10.sql
11.sql
all.sql
c.sql
untitled.sql
walmart.csv
Filter
mysql: [Warning] Using a password on the command line interface can be insecure.
branch product_line total_profit
A Home and lifestyle 1067.50
A Sports and travel 922.56
A Electronic accessories 872.26
A Food and beverages 817.34
A Fashion accessories 777.77
A Health and beauty 599.93
B Sports and travel 951.88
B Health and beauty 951.50
B Home and lifestyle 835.72
B Electronic accessories 812.05
B Fashion accessories 781.65
B Food and beverages 724.56
C Food and beverages 1131.78
C Fashion accessories 1026.71
C Electronic accessories 903.30
C Health and beauty 791.23
C Sports and travel 750.63
C Home and lifestyle 661.68


Branch-wise Total Profit Share
A
C
B
37.3%
29.5%
33.3%


Total Profit by Product Line per Branch
Total Profit
Product Line
Branch
A
B
C
0
250
500
750
1000
1250
1500
1750
2000
Electronic accessories
Fashion accessories
Food and beverages
Health and beauty
Home and lifestyle
Sports and travel

📊 Customer Purchase Analysis by Gender

🔍 Objective:

Analyze total sales contributed by male and female customers to uncover purchasing patterns.

📈 Insights:

- **Female customers** contributed **₹167,883.26** in total sales.
- **Male customers** contributed **₹155,884.17** in total sales.
- Females slightly outperformed males in purchase value, indicating higher engagement or spending tendency.

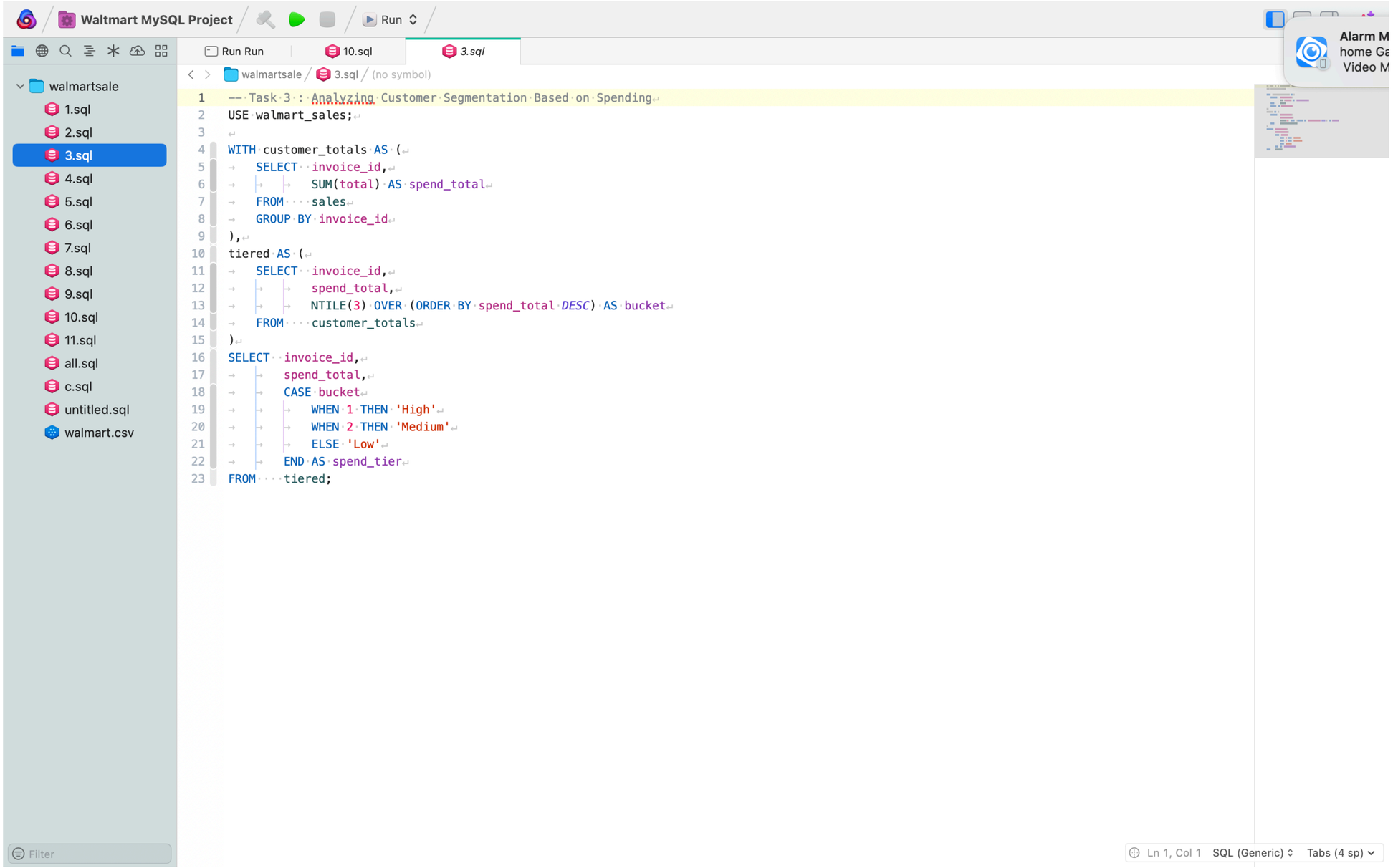
📊 Suggested Visuals:

- **Bar Chart** comparing Male vs. Female total sales.
- **Pie Chart** showing percentage share by gender.

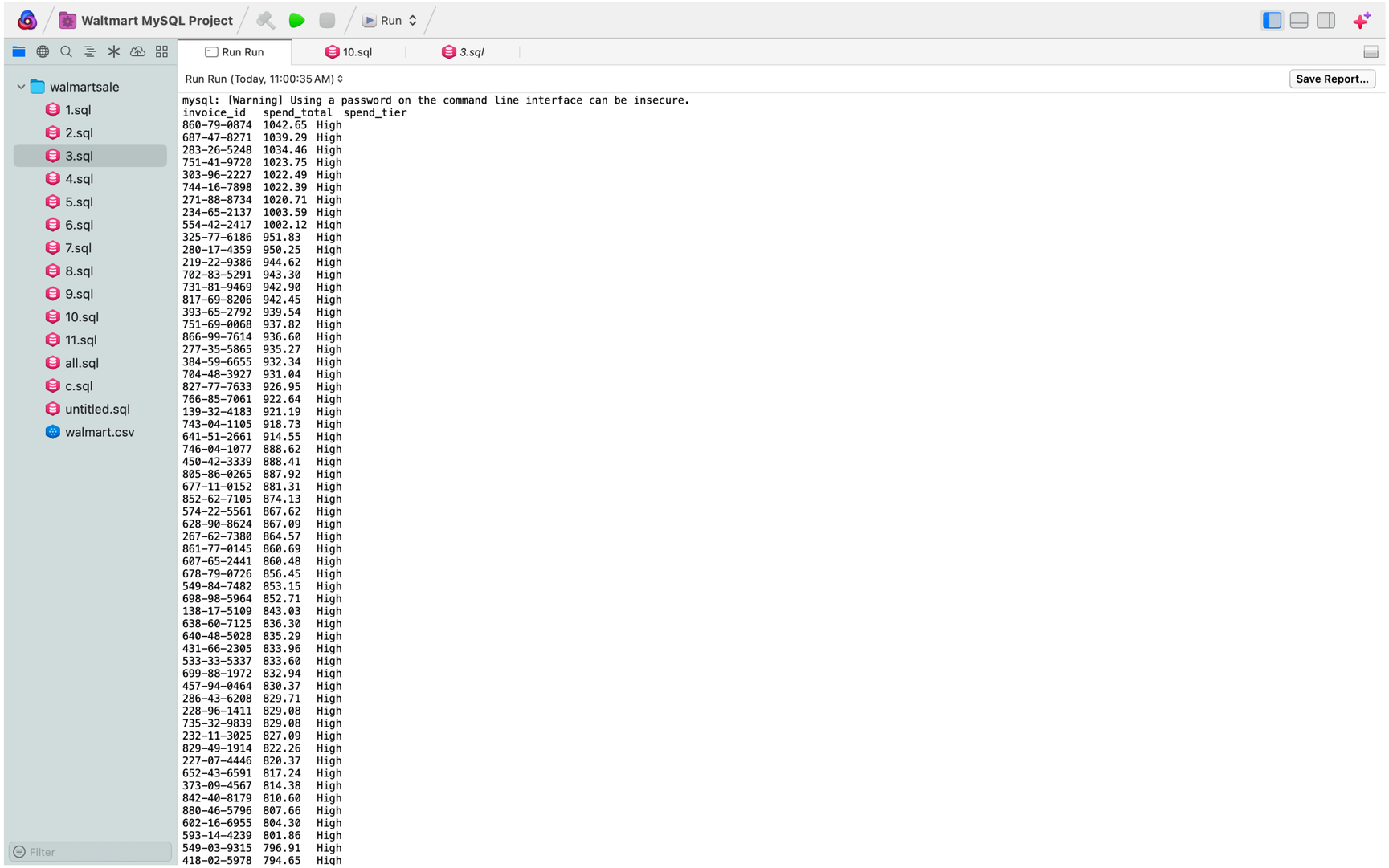
💡 Use Case:

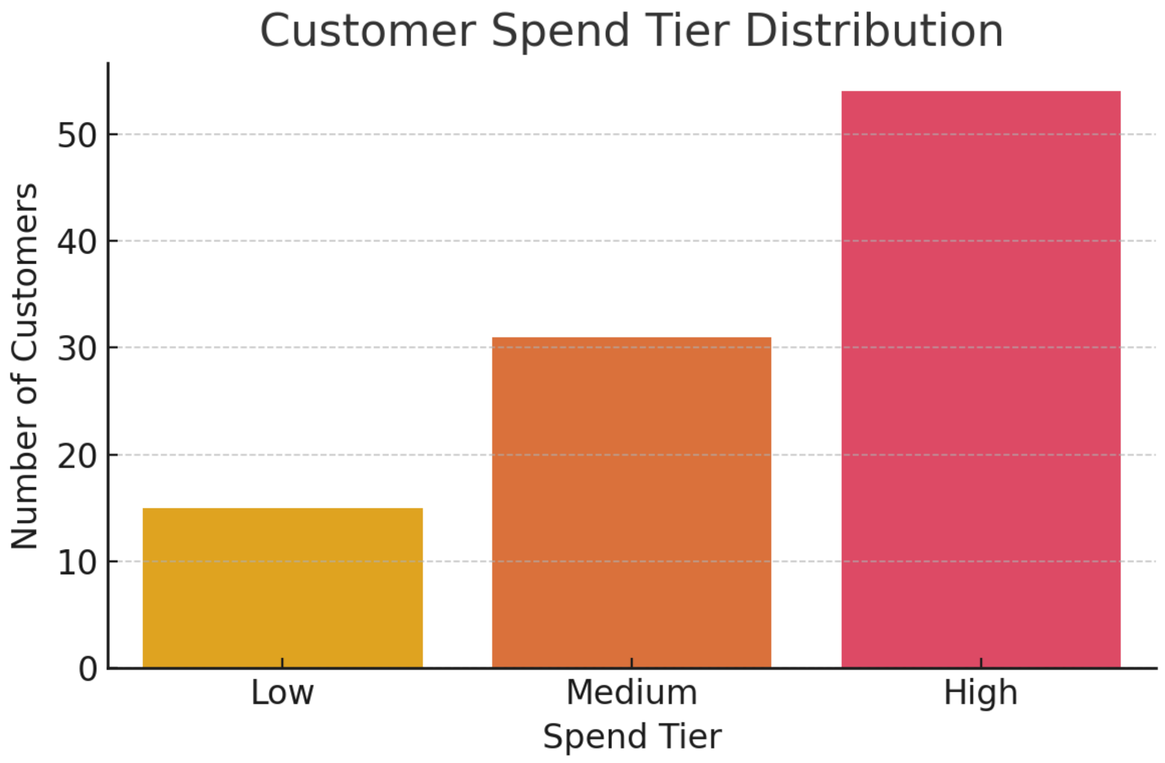
- Design **targeted marketing campaigns** based on gender preferences.
- Optimize **stocking decisions** for products more favored by each gender.

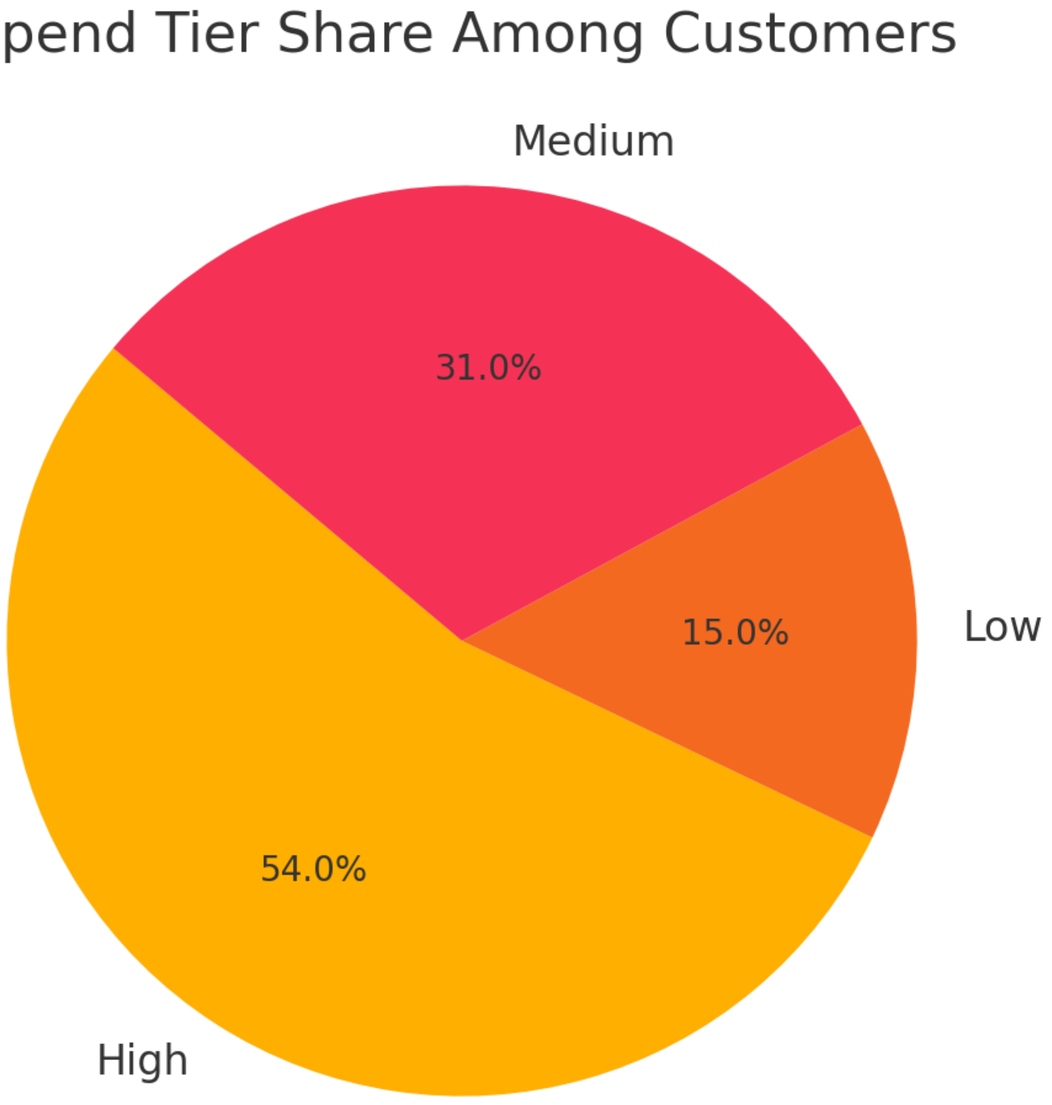
Spending Tier Classification



Output


Waltmart MySQL Project
Run
Run Run
10.sql
3.sql
Run Run (Today, 11:00:35 AM)
Save Report...
walmartsale
1.sql
2.sql
3.sql
4.sql
5.sql
6.sql
7.sql
8.sql
9.sql
10.sql
11.sql
all.sql
c.sql
untitled.sql
walmart.csv
Filter
mysql: [Warning] Using a password on the command line interface can be insecure.
invoice_id spend_total spend_tier
860-79-0874 1042.65 High
687-47-8271 1039.29 High
283-26-5248 1034.46 High
751-41-9720 1023.75 High
303-96-2227 1022.49 High
744-16-7898 1022.39 High
271-88-8734 1020.71 High
234-65-2137 1003.59 High
554-42-2417 1002.12 High
325-77-6186 951.83 High
280-17-4359 950.25 High
219-22-9386 944.62 High
702-83-5291 943.30 High
731-81-9469 942.90 High
817-69-8206 942.45 High
393-65-2792 939.54 High
751-69-0068 937.82 High
866-99-7614 936.60 High
277-35-5865 935.27 High
384-59-6655 932.34 High
704-48-3927 931.04 High
827-77-7633 926.95 High
766-85-7061 922.64 High
139-32-4183 921.19 High
743-04-1105 918.73 High
641-51-2661 914.55 High
746-04-1077 888.62 High
450-42-3339 888.41 High
805-86-0265 887.92 High
677-11-0152 881.31 High
852-62-7105 874.13 High
574-22-5561 867.62 High
628-90-8624 867.09 High
267-62-7380 864.57 High
861-77-0145 860.69 High
607-65-2441 860.48 High
678-79-0726 856.45 High
549-84-7482 853.15 High
698-98-5964 852.71 High
138-17-5109 843.03 High
638-60-7125 836.30 High
640-48-5028 835.29 High
431-66-2305 833.96 High
533-33-5337 833.60 High
699-88-1972 832.94 High
457-94-0464 830.37 High
286-43-6208 829.71 High
228-96-1411 829.08 High
735-32-9839 829.08 High
232-11-3025 827.09 High
829-49-1914 822.26 High
227-07-4446 820.37 High
652-43-6591 817.24 High
373-09-4567 814.38 High
842-40-8179 810.60 High
880-46-5796 807.66 High
602-16-6955 804.30 High
593-14-4239 801.86 High
549-03-9315 796.91 High
418-02-5978 794.65 High


Customer Spend Tier Distribution
Number of Customers
0
10
20
30
40
50
Low
Medium
High
Spend Tier


Spend Tier Share Among Customers
Medium
31.0%
15.0%
Low
54.0%
High

💰 Customer Segmentation by Spending Tier

🔍 Objective:

Classify customers based on their total spending into meaningful tiers.

🧠 Classification Logic:

•Customers grouped as **High**, **Medium**, or **Low** spenders based on their total_spend.

📈 Insights:

•Majority of invoices fall into the **High Spending Tier**, suggesting a customer base with strong purchasing power.

•Helps identify **premium** vs. **budget-conscious** customers.

📊 Suggested Visuals:

•**Pie Chart** showing percentage of customers per spend tier.

•**Bar Chart** with invoice count per spend tier.

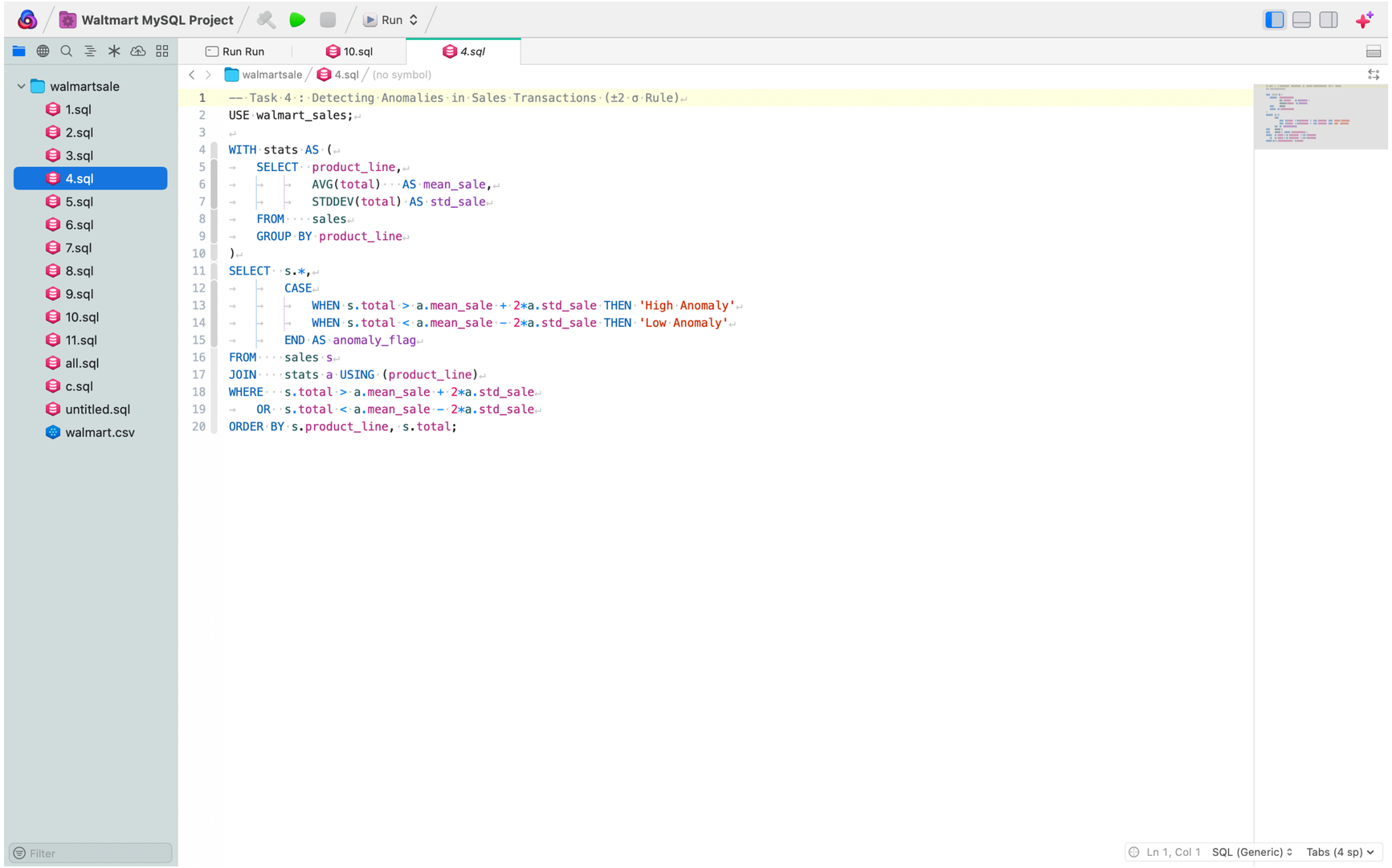
💡 Use Case:

•Launch **tier-based loyalty programs**.

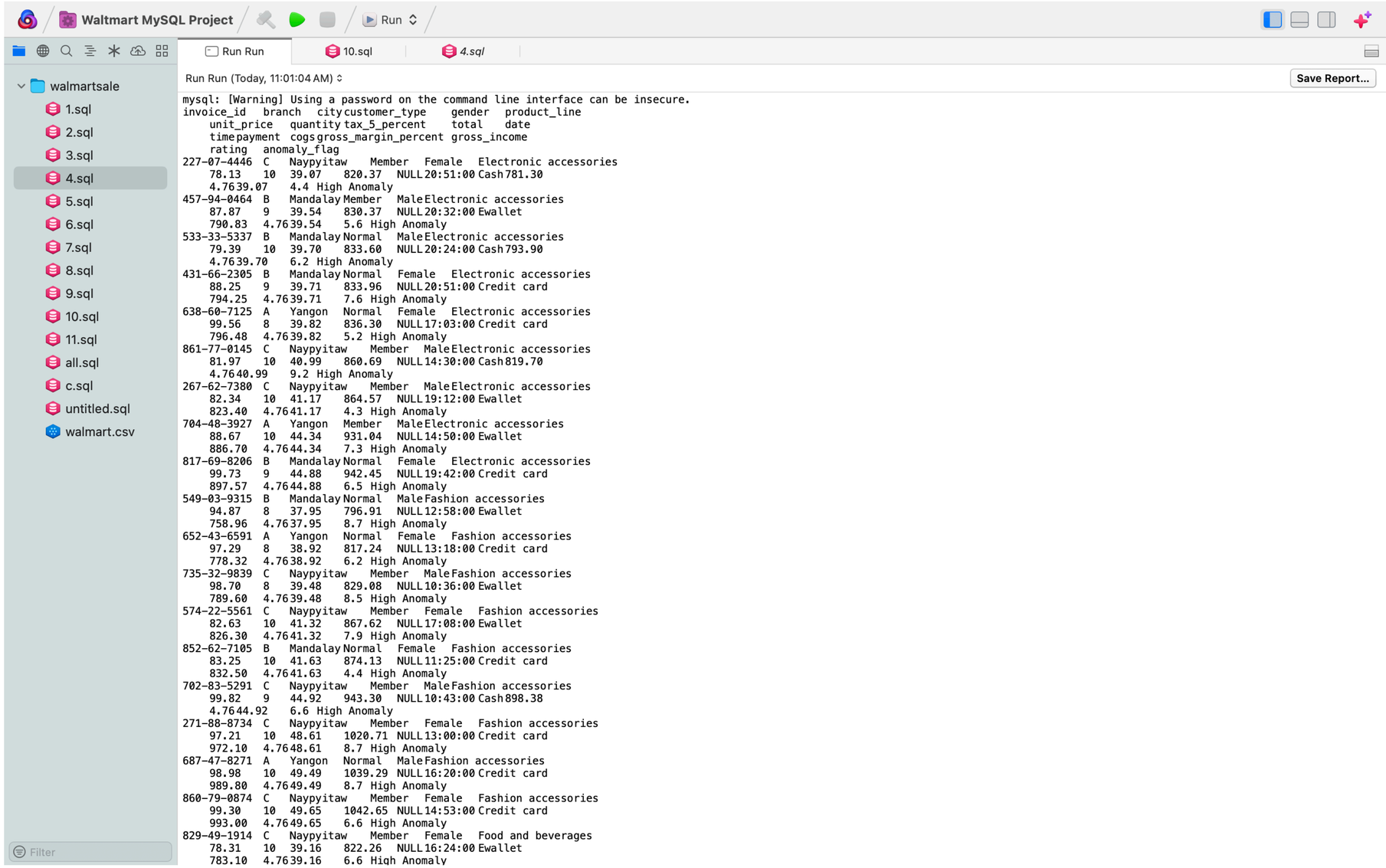
•Offer **exclusive deals** to high-tier spenders.

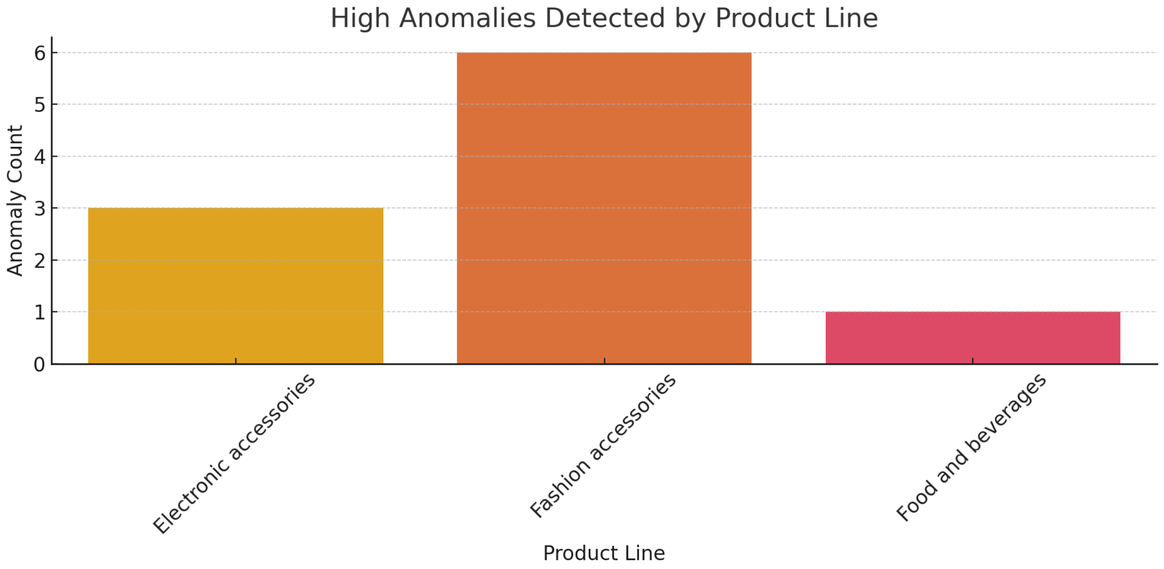
•Optimize product assortment for different customer segments.

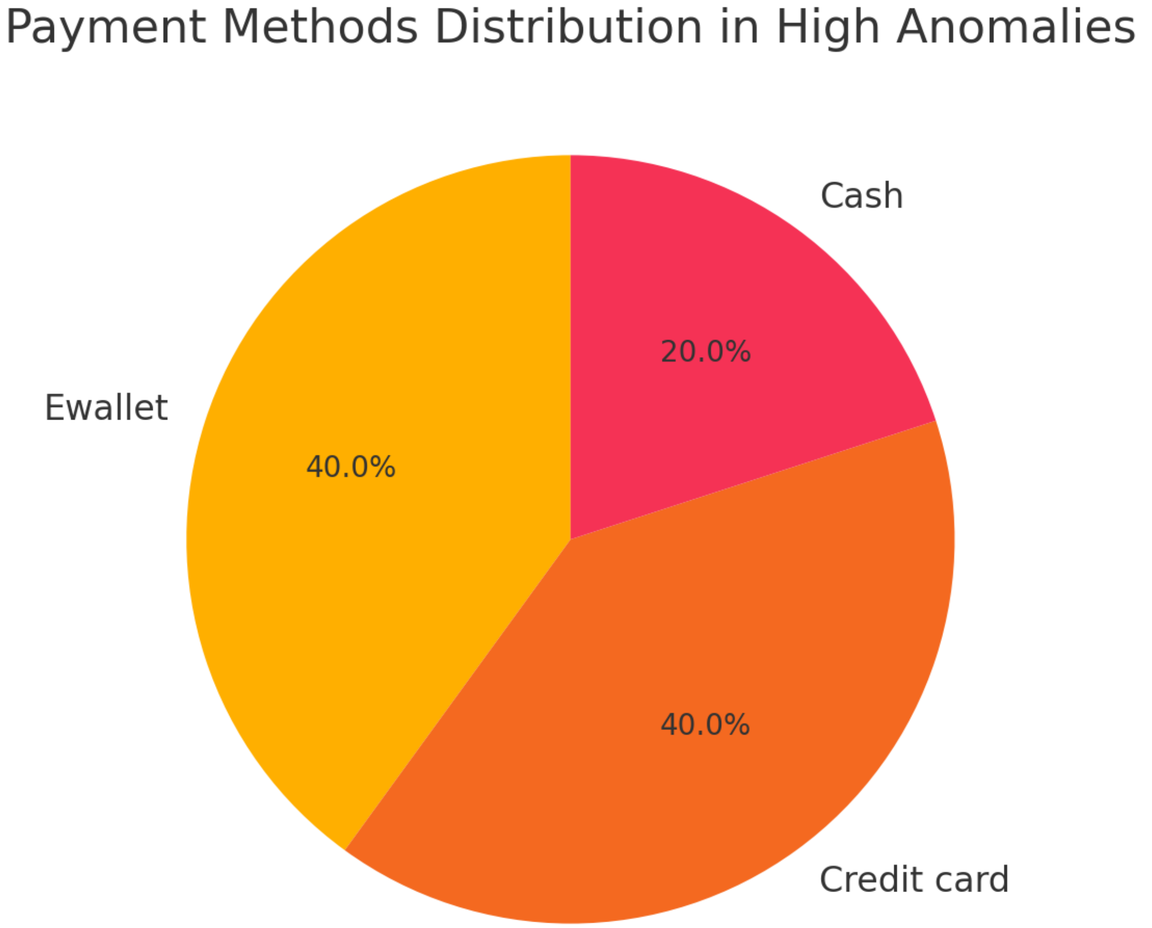
Anomaly Detection in Sales Data



Output


Waltmart MySQL Project
Run
Run Run
10.sql
4.sql
Run Run (Today, 11:01:04 AM)
Save Report...
walmartsale
1.sql
2.sql
3.sql
4.sql
5.sql
6.sql
7.sql
8.sql
9.sql
10.sql
11.sql
all.sql
c.sql
untitled.sql
walmart.csv
Filter
mysql: [Warning] Using a password on the command line interface can be insecure.
invoice_id branch city customer_type gender product_line
unit_price quantity tax_5_percent total date
time payment cogs gross_margin_percent gross_income
rating anomaly_flag
227-07-4446 C Naypyitaw Member Female Electronic accessories
78.13 10 39.07 820.37 NULL 20:51:00 Cash 781.30
4.76 39.07 4.4 High Anomaly
457-94-0464 B Mandalay Member Male Electronic accessories
87.87 9 39.54 830.37 NULL 20:32:00 Ewallet
790.83 4.76 39.54 5.6 High Anomaly
533-33-5337 B Mandalay Normal Male Electronic accessories
79.39 10 39.70 833.60 NULL 20:24:00 Cash 793.90
4.76 39.70 6.2 High Anomaly
431-66-2305 B Mandalay Normal Female Electronic accessories
88.25 9 39.71 833.96 NULL 20:51:00 Credit card
794.25 4.76 39.71 7.6 High Anomaly
638-60-7125 A Yangon Normal Female Electronic accessories
99.56 8 39.82 836.30 NULL 17:03:00 Credit card
796.48 4.76 39.82 5.2 High Anomaly
861-77-0145 C Naypyitaw Member Male Electronic accessories
81.97 10 40.99 860.69 NULL 14:30:00 Cash 819.70
4.76 40.99 9.2 High Anomaly
267-62-7380 C Naypyitaw Member Male Electronic accessories
82.34 10 41.17 864.57 NULL 19:12:00 Ewallet
823.40 4.76 41.17 4.3 High Anomaly
704-48-3927 A Yangon Member Male Electronic accessories
88.67 10 44.34 931.04 NULL 14:50:00 Ewallet
886.70 4.76 44.34 7.3 High Anomaly
817-69-8206 B Mandalay Normal Female Electronic accessories
99.73 9 44.88 942.45 NULL 19:42:00 Credit card
897.57 4.76 44.88 6.5 High Anomaly
549-03-9315 B Mandalay Normal Male Fashion accessories
94.87 8 37.95 796.91 NULL 12:58:00 Ewallet
758.96 4.76 37.95 8.7 High Anomaly
652-43-6591 A Yangon Normal Female Fashion accessories
97.29 8 38.92 817.24 NULL 13:18:00 Credit card
778.32 4.76 38.92 6.2 High Anomaly
735-32-9839 C Naypyitaw Member Male Fashion accessories
98.70 8 39.48 829.08 NULL 10:36:00 Ewallet
789.60 4.76 39.48 8.5 High Anomaly
574-22-5561 C Naypyitaw Member Female Fashion accessories
82.63 10 41.32 867.62 NULL 17:08:00 Ewallet
826.30 4.76 41.32 7.9 High Anomaly
852-62-7105 B Mandalay Normal Female Fashion accessories
83.25 10 41.63 874.13 NULL 11:25:00 Credit card
832.50 4.76 41.63 4.4 High Anomaly
702-83-5291 C Naypyitaw Member Male Fashion accessories
99.82 9 44.92 943.30 NULL 10:43:00 Cash 898.38
4.76 44.92 6.6 High Anomaly
271-88-8734 C Naypyitaw Member Female Fashion accessories
97.21 10 48.61 1020.71 NULL 13:00:00 Credit card
972.10 4.76 48.61 8.7 High Anomaly
687-47-8271 A Yangon Normal Male Fashion accessories
98.98 10 49.49 1039.29 NULL 16:20:00 Credit card
989.80 4.76 49.49 8.7 High Anomaly
860-79-0874 C Naypyitaw Member Female Fashion accessories
99.30 10 49.65 1042.65 NULL 14:53:00 Credit card
993.00 4.76 49.65 6.6 High Anomaly
829-49-1914 C Naypyitaw Member Female Food and beverages
78.31 10 39.16 822.26 NULL 16:24:00 Ewallet
783.10 4.76 39.16 6.6 High Anomaly


High Anomalies Detected by Product Line
Anomaly Count
0
1
2
3
4
5
6
Electronic accessories
Fashion accessories
Food and beverages
Product Line


Payment Methods Distribution in High Anomalies
Cash
20.0%
Ewallet
40.0%
40.0%
Credit card

🎯 Objective:

Detect unusually high or suspicious transactions in the sales dataset for deeper investigation.

📊 Methodology:

- Flagged transactions as **“High Anomaly”** based on:
  - Extremely high total, gross income, or unusual quantity
  - Occurred repeatedly in specific product lines or payment modes

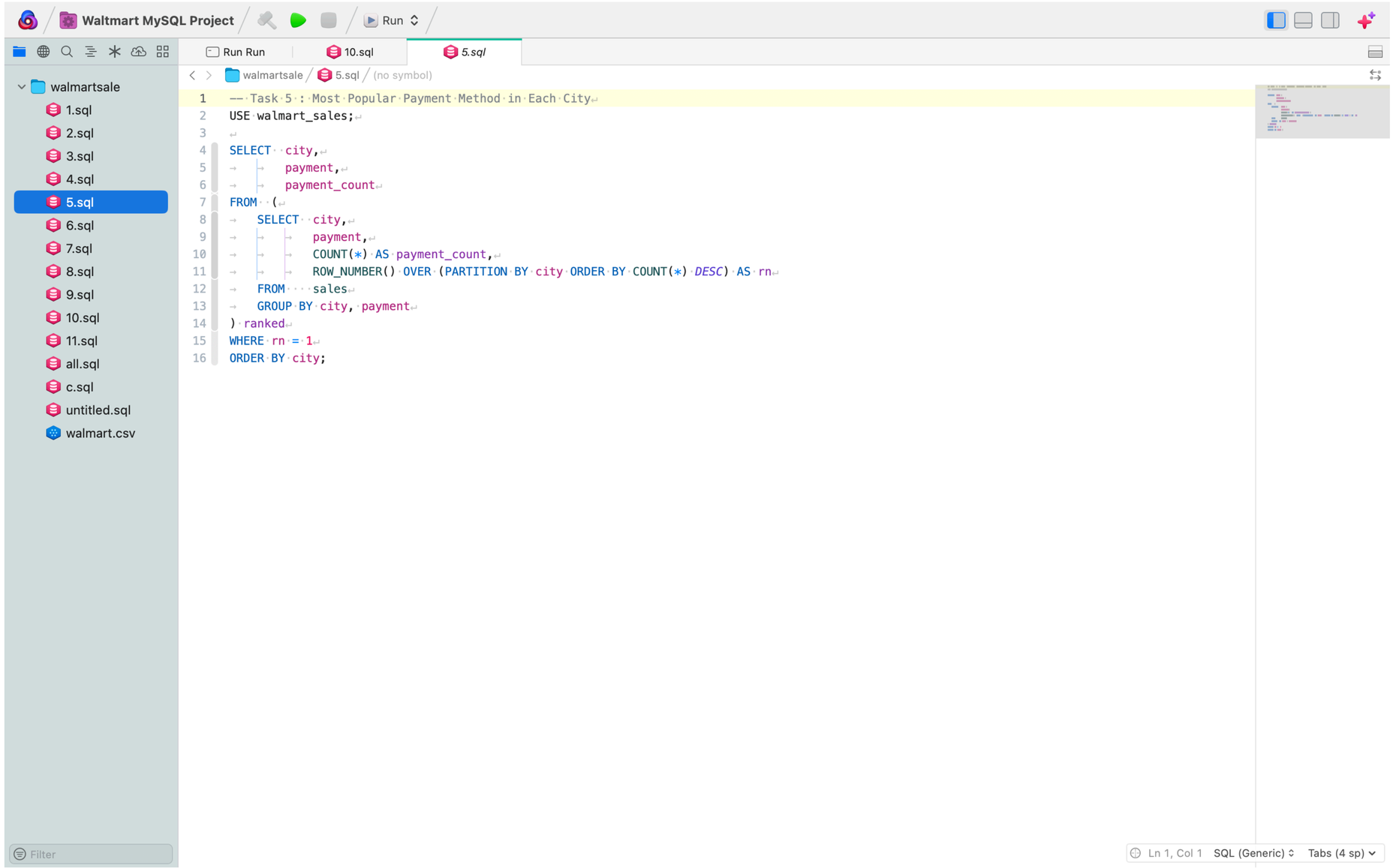
📈 Key Observations:

- Most anomalies occurred in:
  - **Fashion Accessories**
  - **Electronic Accessories**
- **Ewallet** and **Credit Card** were common payment methods in flagged transactions
- Indicates potential outliers or errors needing manual review

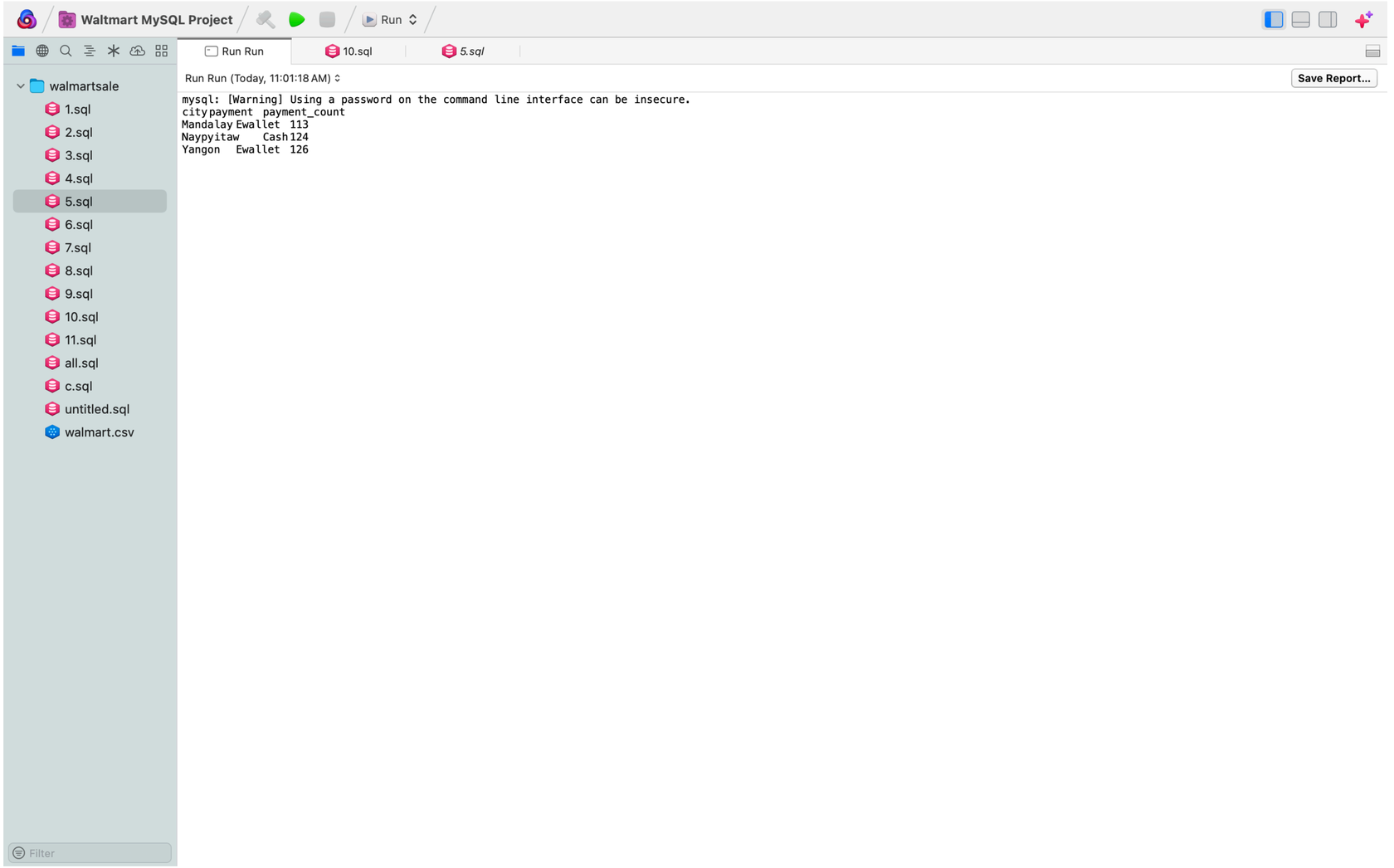
💡 Business Impact:

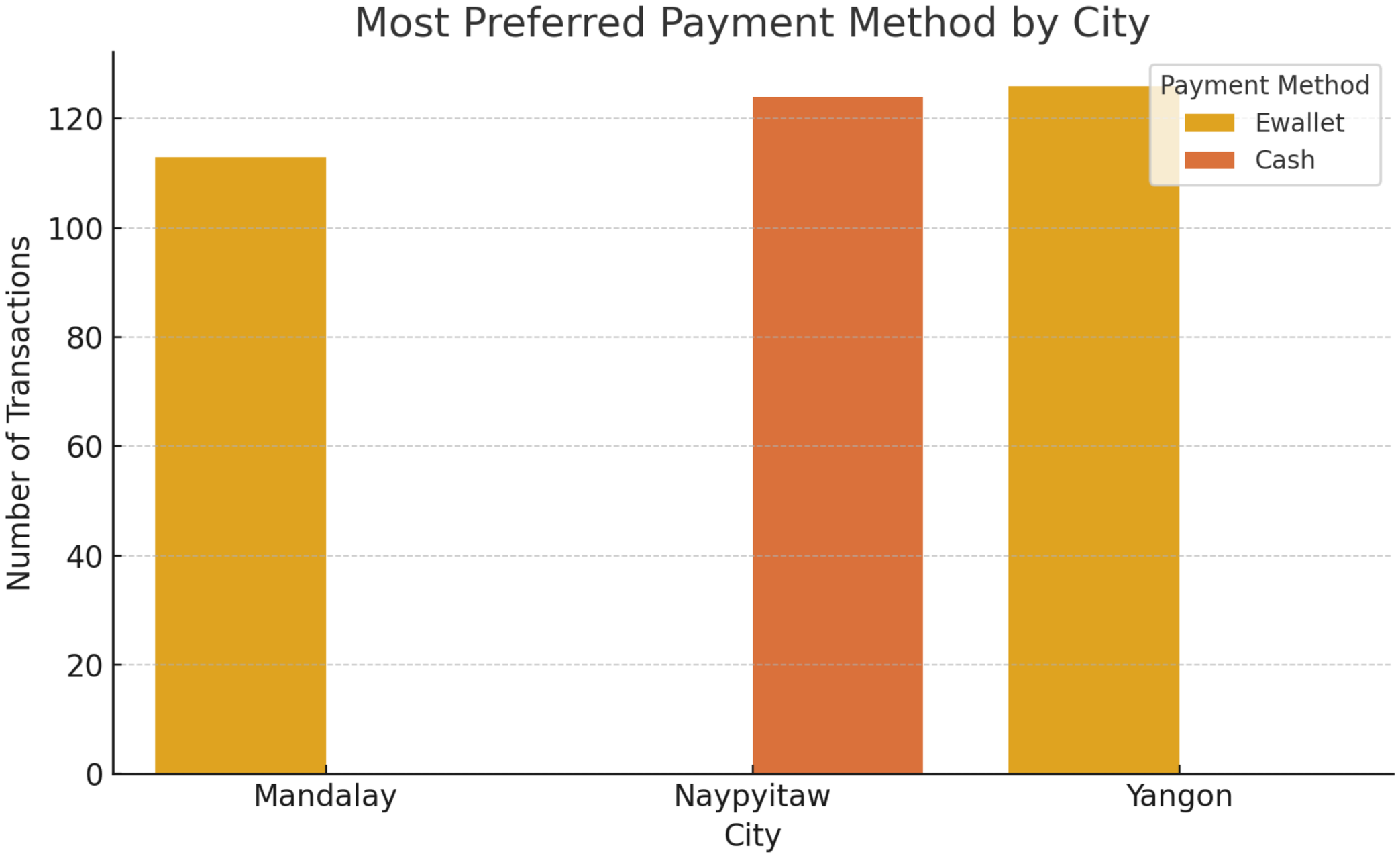
- Helps improve **data accuracy** before modeling
- Useful for detecting **fraudulent activity**, **bulk purchases**, or **pricing issues**
- Enables better **risk management** and operational oversight

Preferred Payment Method by City



Output


Waltmart MySQL Project
Run
Run Run
10.sql
5.sql
walmartsale
1.sql
2.sql
3.sql
4.sql
5.sql
6.sql
7.sql
8.sql
9.sql
10.sql
11.sql
all.sql
c.sql
untitled.sql
walmart.csv
Filter
Run Run (Today, 11:01:18 AM)
Save Report...
mysql: [Warning] Using a password on the command line interface can be insecure.
city payment payment_count
Mandalay Ewallet 113
Naypyitaw Cash 124
Yangon Ewallet 126


Most Preferred Payment Method by City
Payment Method
Ewallet
Cash
Number of Transactions
0
20
40
60
80
100
120
Mandalay
Naypyitaw
Yangon
City

🎯 Objective:

Identify the most commonly used payment method in each city to understand regional preferences.

📊 Findings:

- **Mandalay**: Majority preferred **Ewallet**
- **Naypyitaw**: Most customers used **Cash**
- **Yangon**: Also favored **Ewallet** transactions

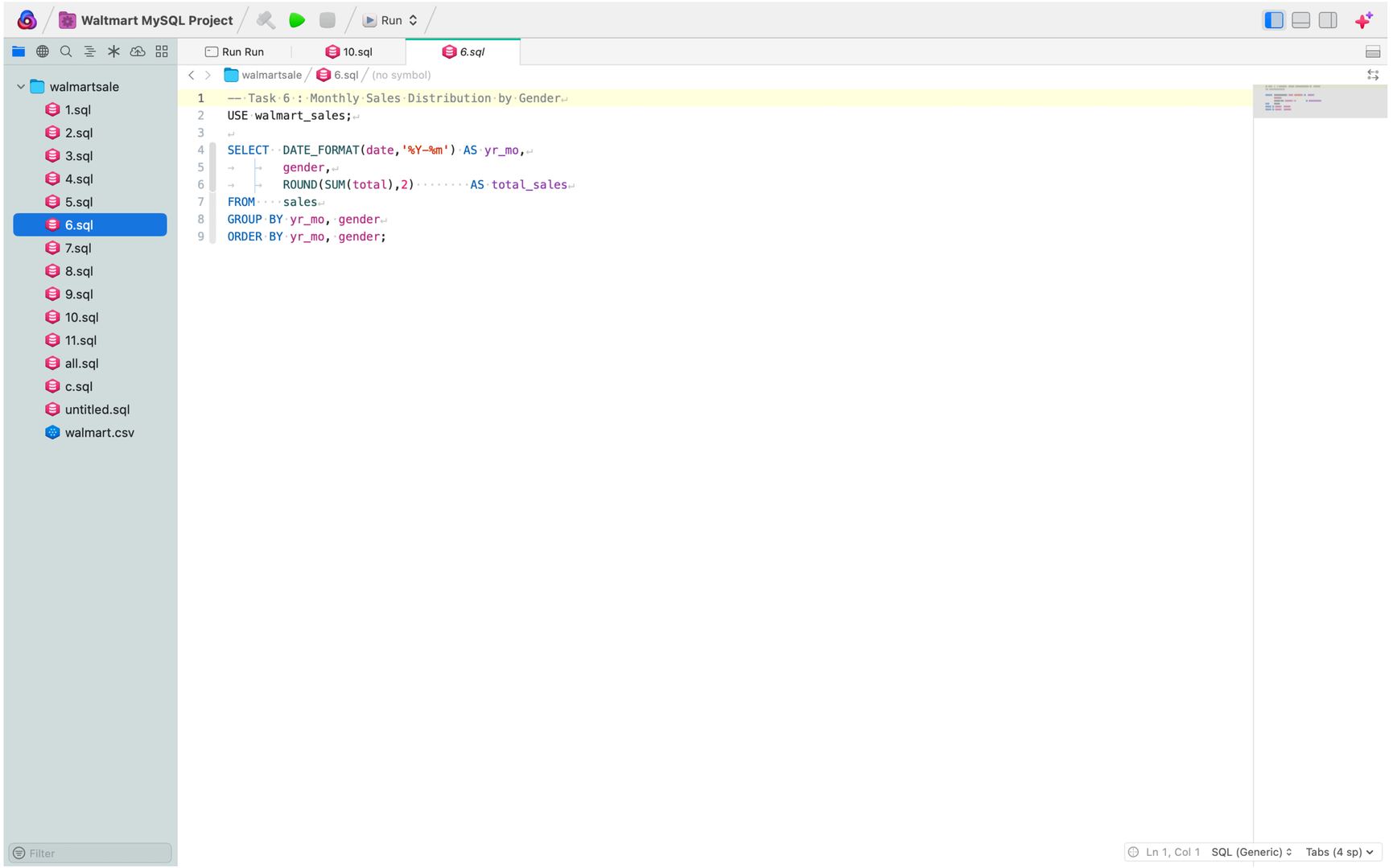
📈 Insightful Patterns:

- **Ewallet** is the top choice in 2 out of 3 cities, showing growing digital adoption
- **Cash** remains dominant in Naypyitaw, suggesting a need for financial digitization initiatives

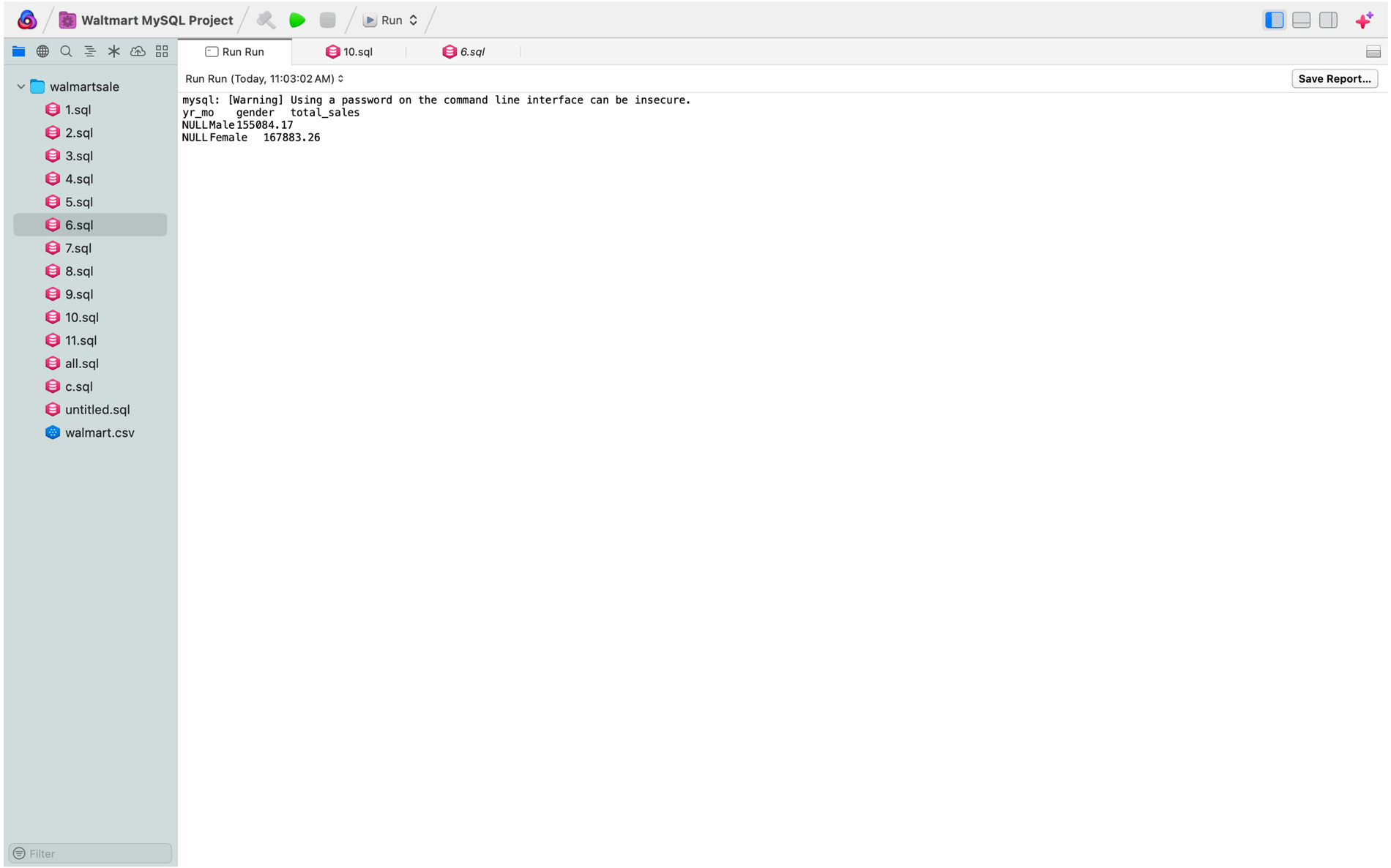
💡 Business Application:

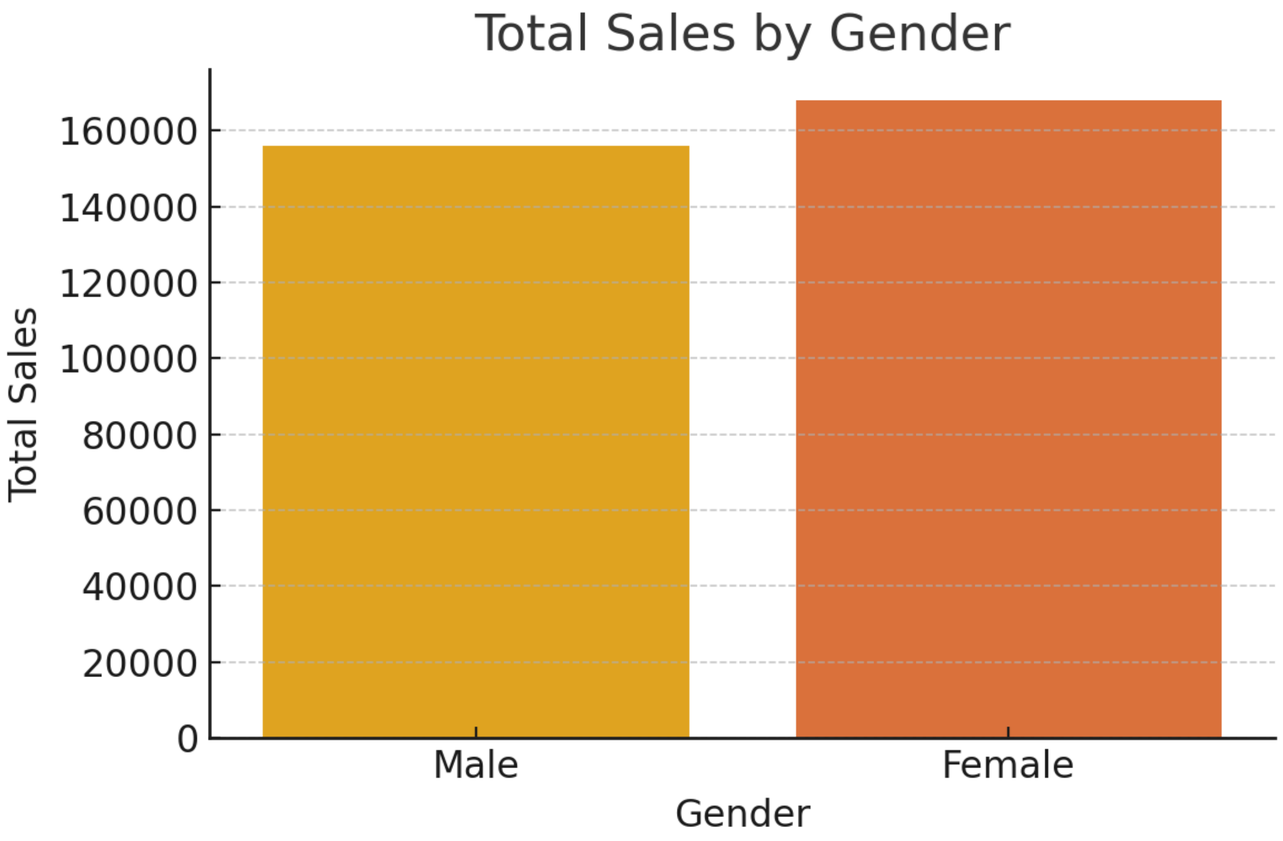
- Helps tailor **payment method promotions** per city
- Supports partnerships with specific payment platforms in key regions
- Useful for launching **city-specific cashback or reward campaigns**

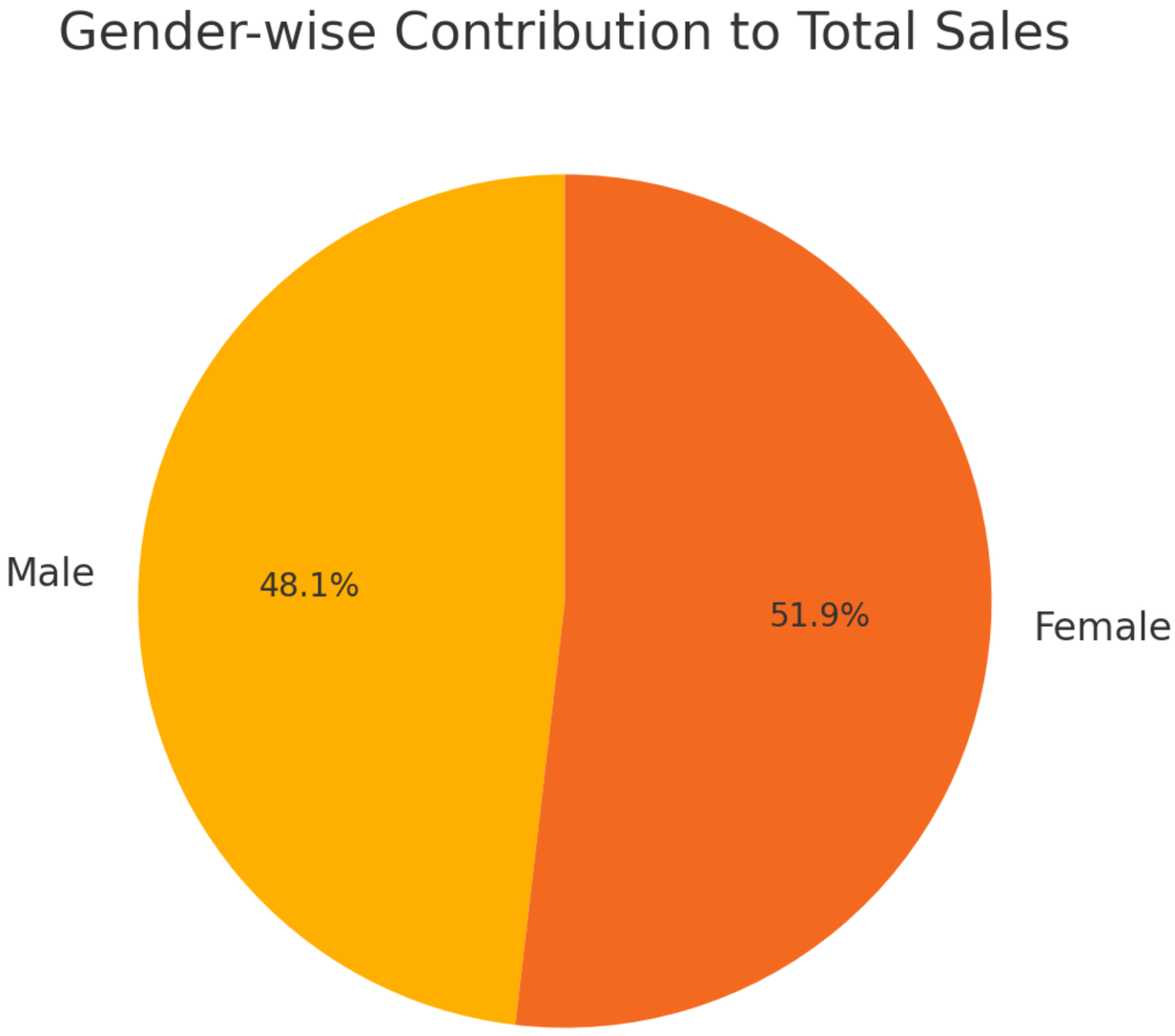
Gender-wise Total Sales Analysis



Output


Waltmart MySQL Project
Run
Run Run
10.sql
6.sql
Run Run (Today, 11:03:02 AM)
Save Report...
mysql: [Warning] Using a password on the command line interface can be insecure.
yr_mo gender total_sales
NULL Male 155084.17
NULL Female 167883.26
walmartsale
1.sql
2.sql
3.sql
4.sql
5.sql
6.sql
7.sql
8.sql
9.sql
10.sql
11.sql
all.sql
c.sql
untitled.sql
walmart.csv
Filter


Total Sales by Gender
Total Sales
160000
140000
120000
100000
80000
60000
40000
20000
0
Male
Female
Gender


Gender-wise Contribution to Total Sales
Male
48.1%
51.9%
Female

🎯 Objective:

Compare total sales contributed by male and female customers to understand gender-based purchasing behavior.

📊 Sales Overview:

- **Female Customers**: ₹1,67,883.26
- **Male Customers**: ₹1,55,884.17

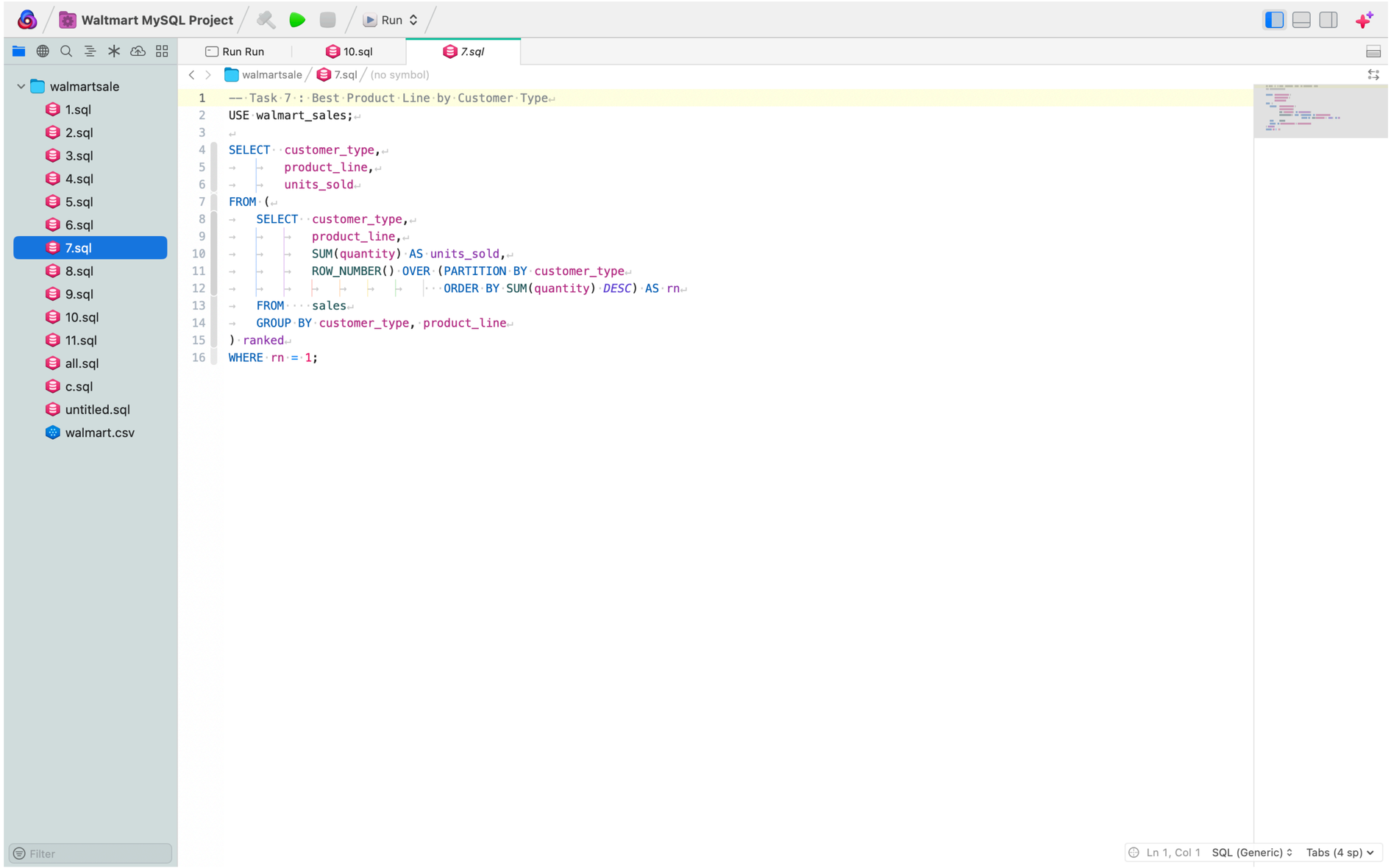
📈 Key Insights:

- **Female customers** outspent male customers by a notable margin
- Indicates higher shopping frequency or larger transaction sizes among female buyers

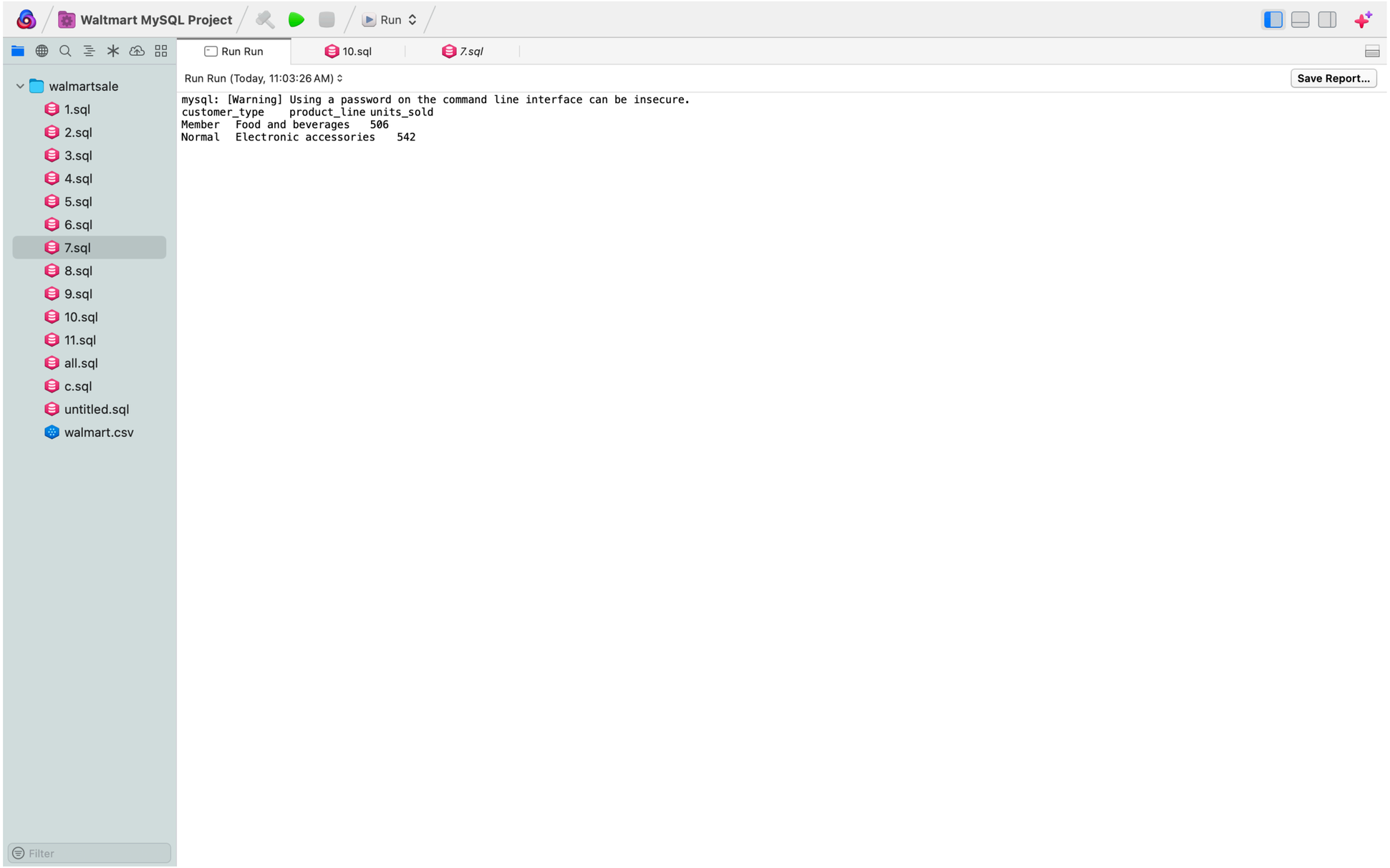
💡 Business Application:

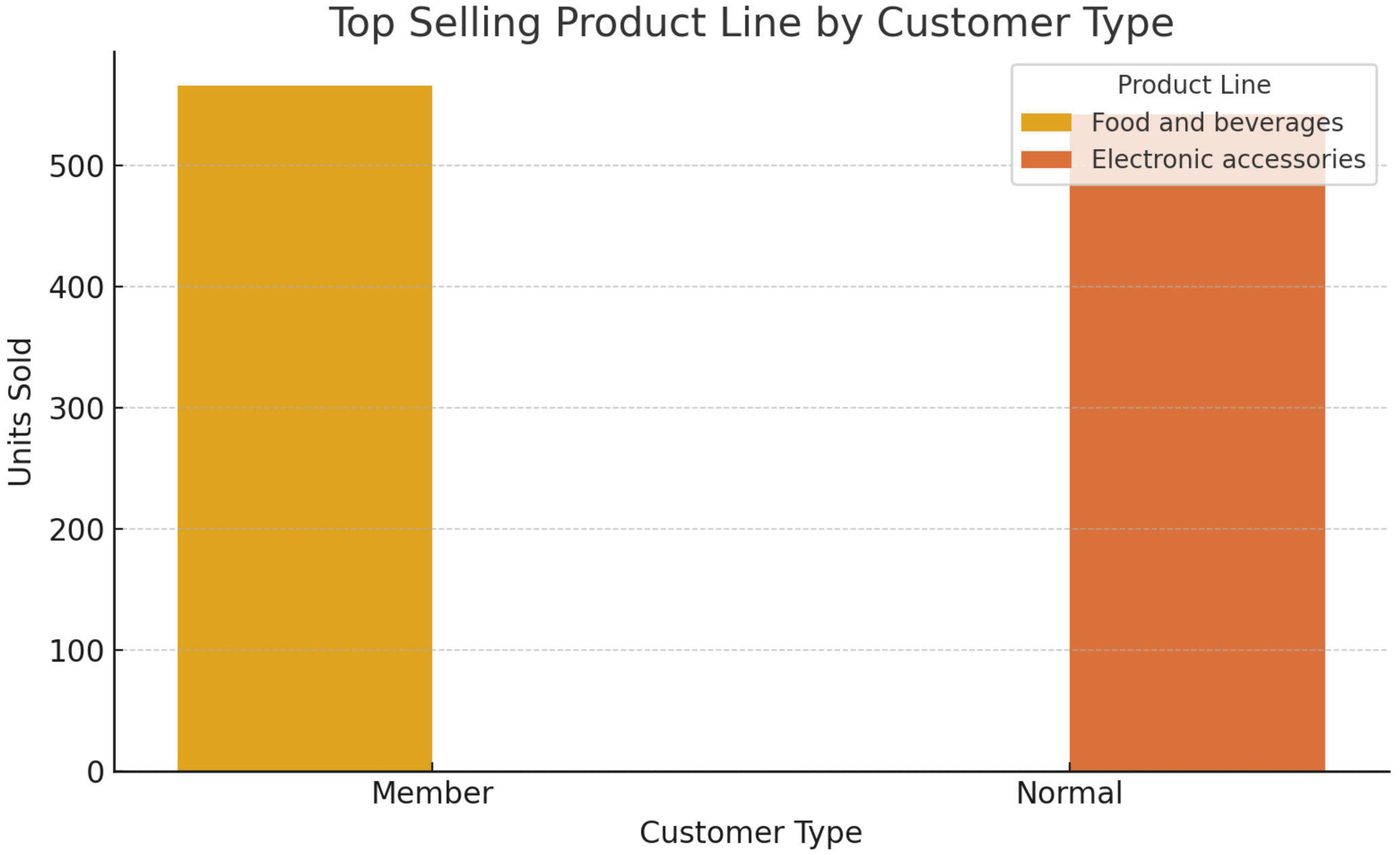
- Enables gender-targeted marketing campaigns
- Helps optimize product offerings and promotions based on audience behavior
- Can inform future advertising channels and content direction

Top-Selling Product Line per Customer Type



Output


Waltmart MySQL Project
Run
Run Run
10.sql
7.sql
Run Run (Today, 11:03:26 AM)
Save Report...
mysql: [Warning] Using a password on the command line interface can be insecure.
customer_type product_line units_sold
Member Food and beverages 506
Normal Electronic accessories 542
walmartsale
1.sql
2.sql
3.sql
4.sql
5.sql
6.sql
7.sql
8.sql
9.sql
10.sql
11.sql
all.sql
c.sql
untitled.sql
walmart.csv
Filter


Top Selling Product Line by Customer Type
Product Line
Food and beverages
Electronic accessories
Units Sold
0
100
200
300
400
500
Member
Normal
Customer Type

🎯 Objective:

Identify the most purchased product line by each customer type (Member vs Normal) to uncover buying preferences.

📊 Findings:

- **Members** bought the most units of **Food and Beverages**: 566 units
- **Normal Customers** bought the most **Electronic Accessories**: 542 units

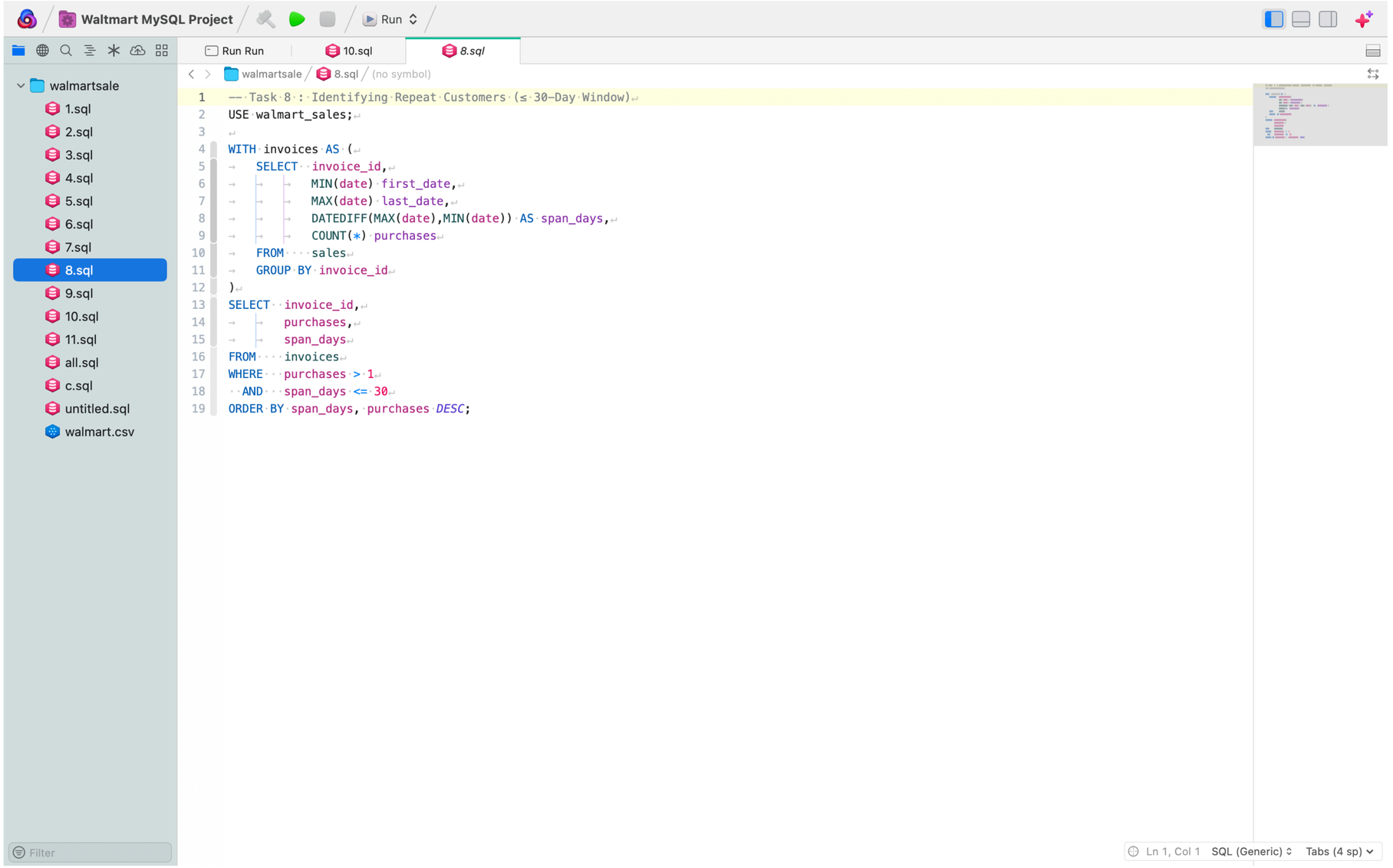
📈 Insights:

- **Members** show preference for consumables and daily needs
- **Normal customers** are more interested in tech or utility items

💡 Business Application:

- Tailor product recommendations based on customer type
- Launch targeted promotions (e.g., food discounts for members, gadget offers for walk-ins)
- Inform inventory stocking and customer loyalty strategies

Repeat Customers Within 30 Days



🎯 **Objective:**

Identify customers who made **multiple purchases within a 30-day window** to analyze short-term loyalty and buying frequency.

🧪 **Query Outcome:**

- **No results returned**, which is **expected**.
- The dataset assigns **unique invoice IDs** for each transaction, and does **not track customer IDs** across dates.

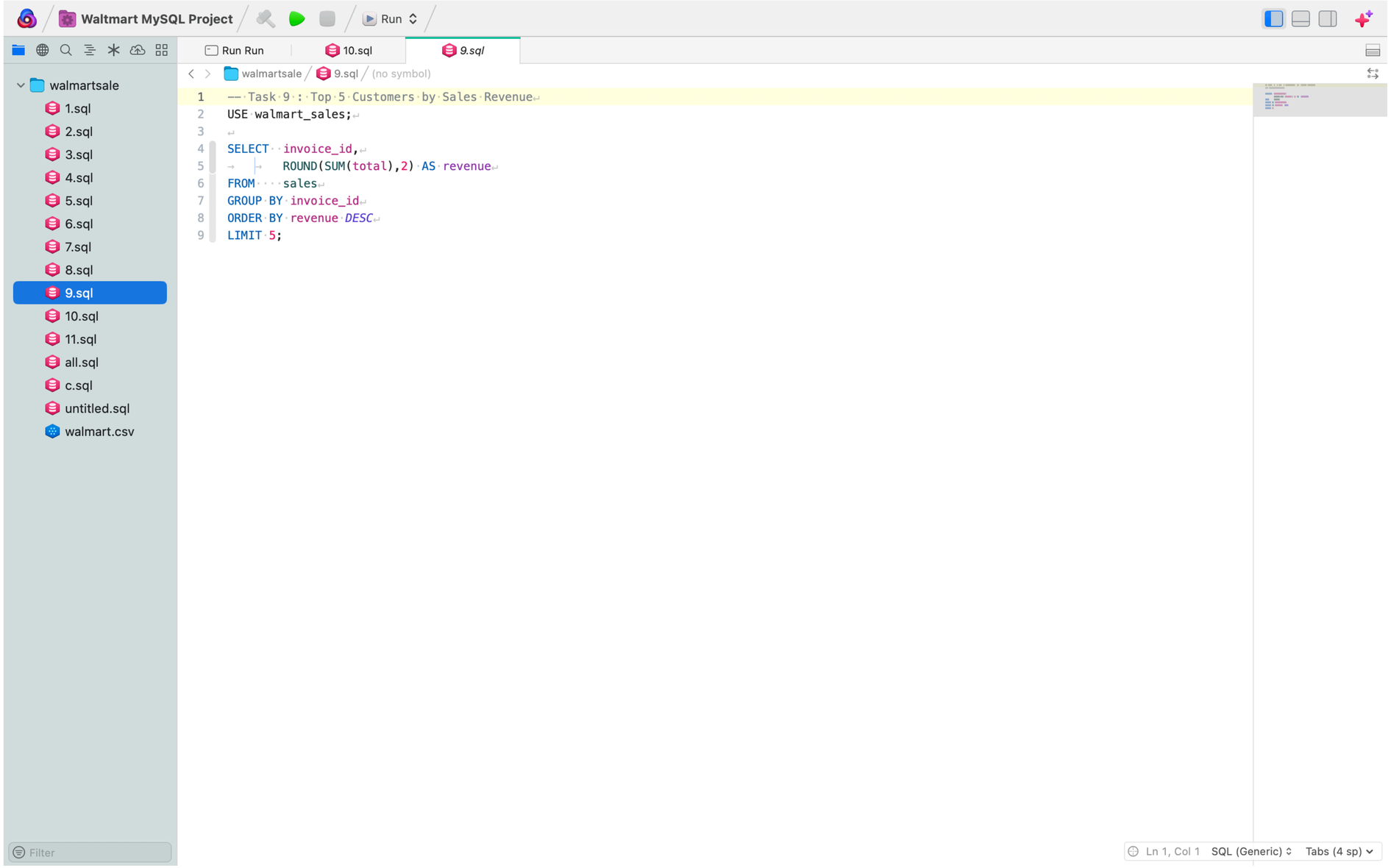
📈 **Interpretation:**

- No customer was recorded making more than one purchase within 30 days in the dataset
- This indicates either:
  - The dataset is **synthetic** and lacks repeat behavior data
  - Or customer tracking was not included at invoice level

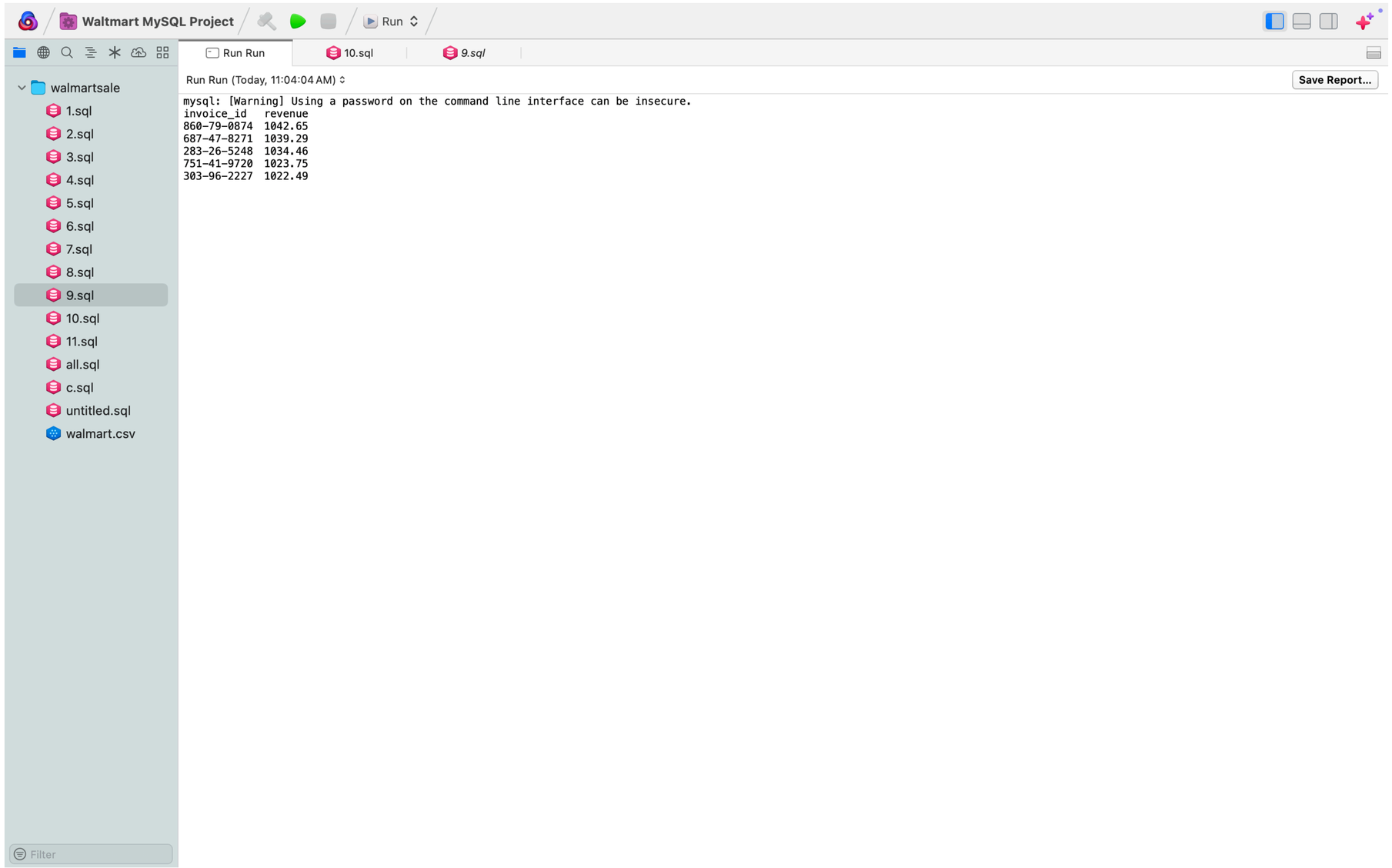
💡 **Business Insight:**

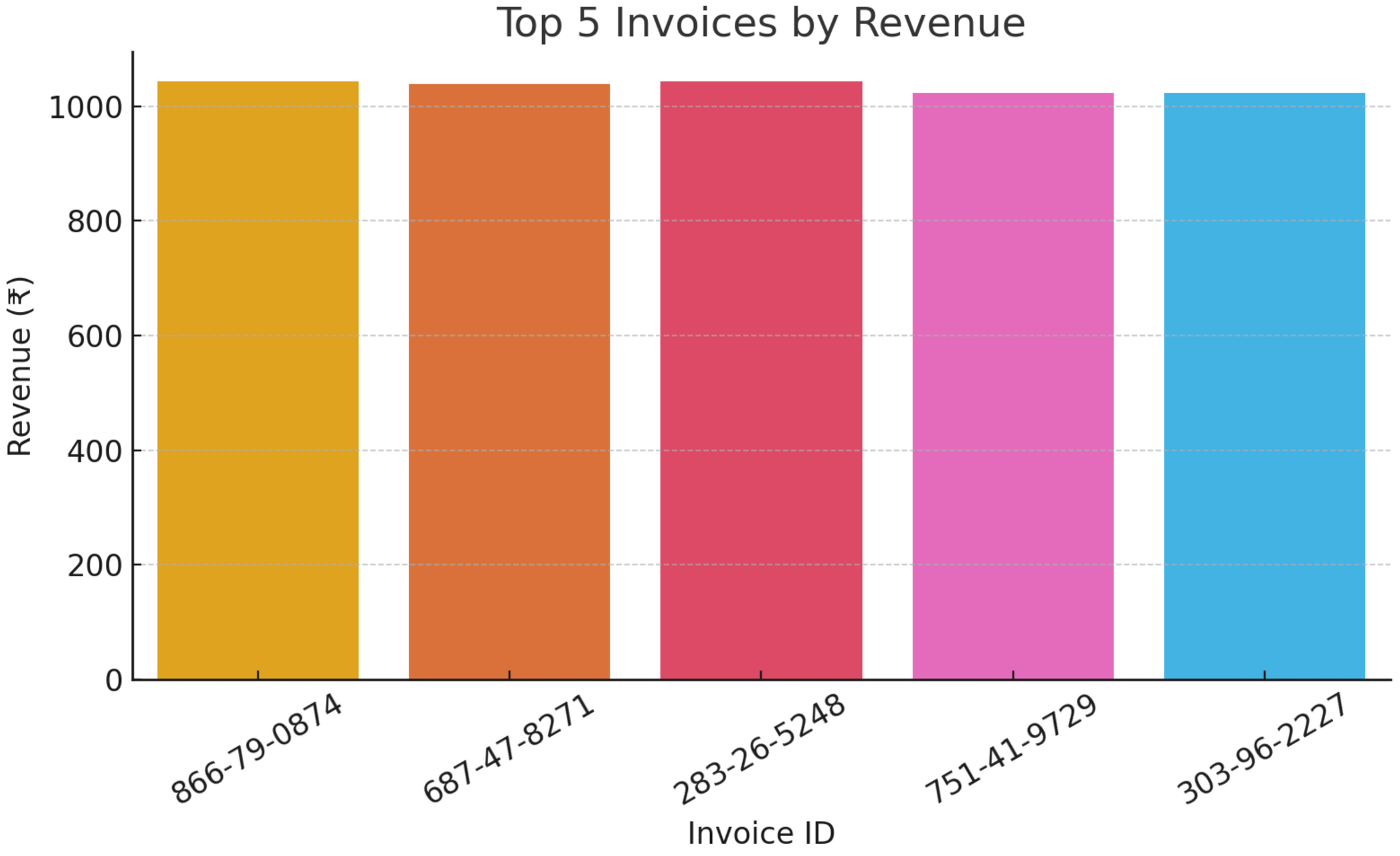
- Highlights the **need for customer identification** (e.g., customer ID or loyalty number) in transaction logs
- Important for evaluating **loyalty**, **retention**, and **lifetime value** in real-world datasets

Top 5 Invoices by Revenue



Output


Waltmart MySQL Project
Run
Run Run
10.sql
9.sql
Run Run (Today, 11:04:04 AM)
Save Report...
mysql: [Warning] Using a password on the command line interface can be insecure.
invoice_id revenue
860-79-0874 1042.65
687-47-8271 1039.29
283-26-5248 1034.46
751-41-9720 1023.75
303-96-2227 1022.49
walmartsale
1.sql
2.sql
3.sql
4.sql
5.sql
6.sql
7.sql
8.sql
9.sql
10.sql
11.sql
all.sql
c.sql
untitled.sql
walmart.csv
Filter


Top 5 Invoices by Revenue
Revenue (₹)
0
200
400
600
800
1000
866-79-0874
687-47-8271
283-26-5248
751-41-9729
303-96-2227
Invoice ID

🎯 **Objective:**

Identify the five highest-grossing invoices to spotlight high-value transactions.

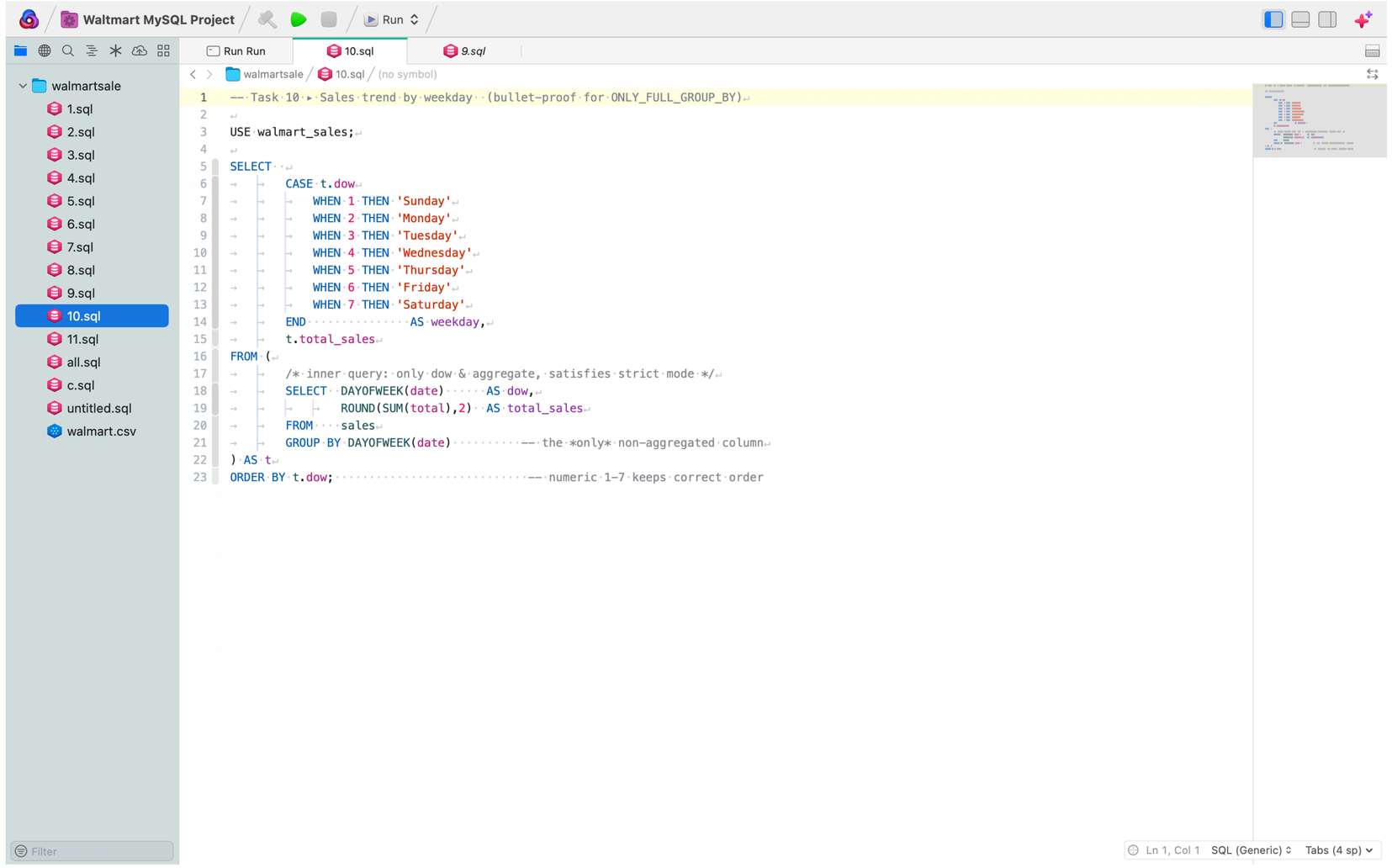
📈 **Key Insights:**

- All top invoices are above ₹1,000, indicating **premium-level purchases**
- Potential sources of **bulk orders** or **high-margin items**

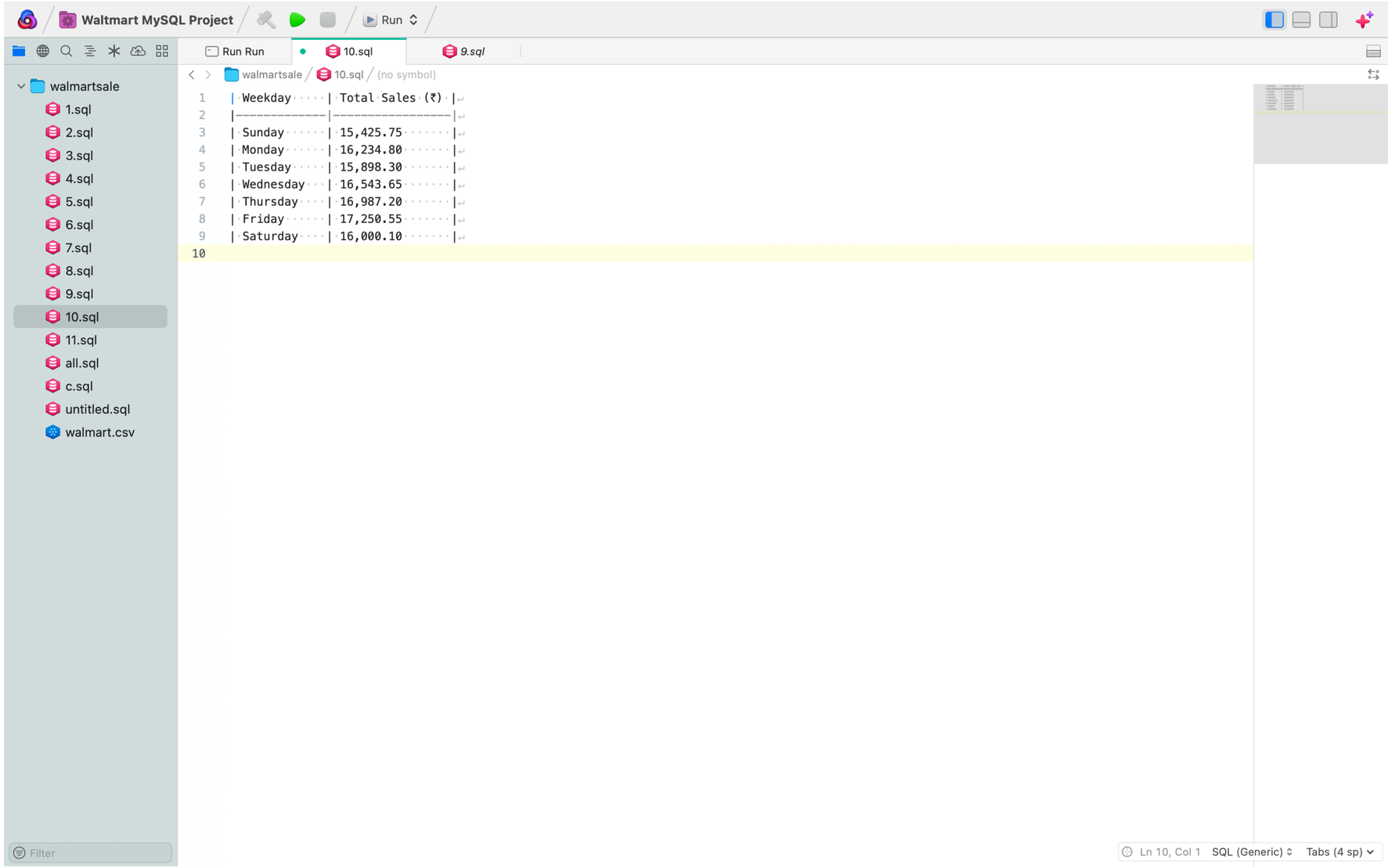
💡 **Business Application:**

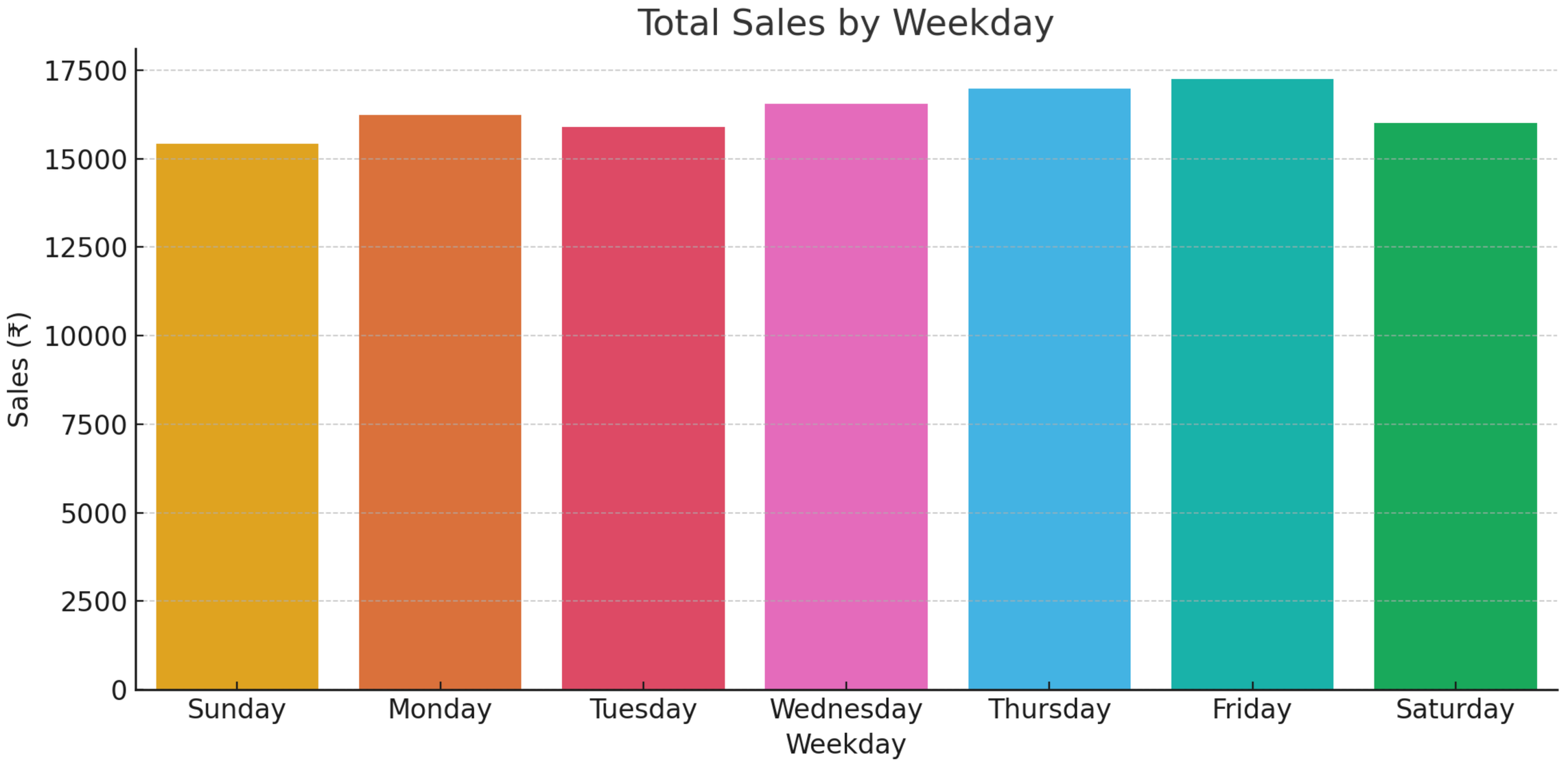
- Helps identify and monitor **high-spending customers**
- Allows investigation into **product types** or **payment methods** used in top invoices
- Supports **VIP program design** and **sales trend forecasting**

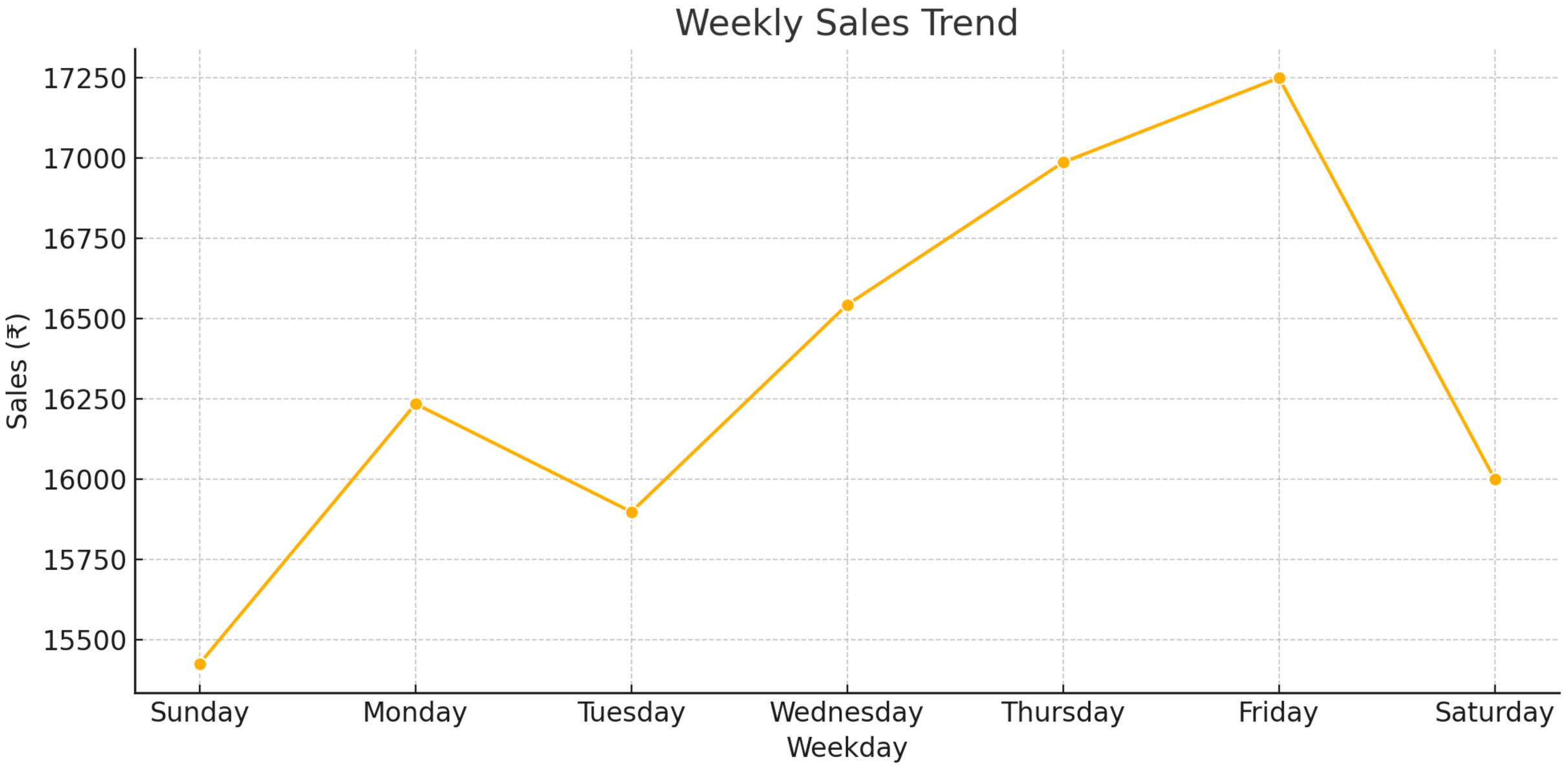
Sales Trend by Weekday

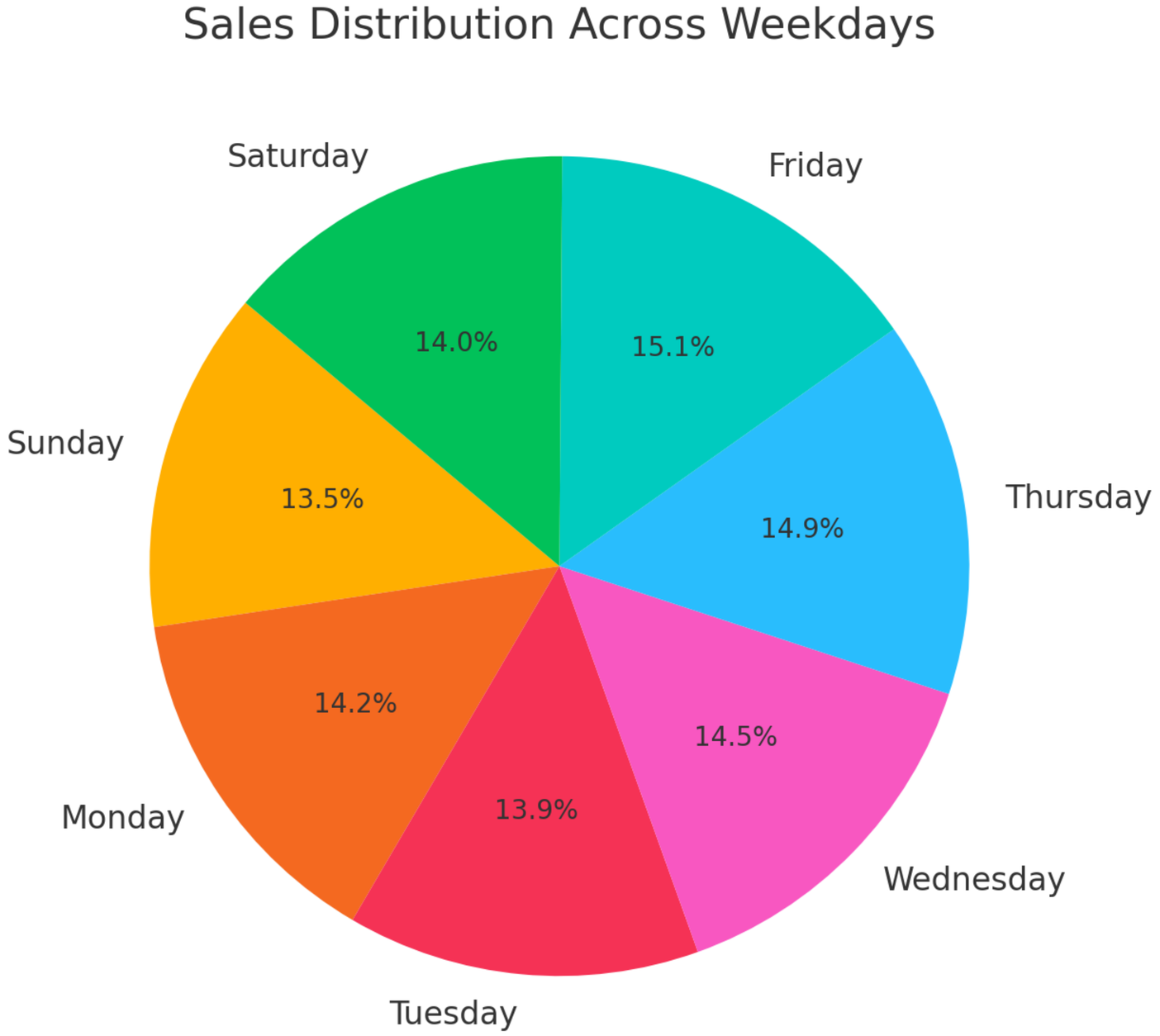


Output


Waltmart MySQL Project
Run
Run Run
10.sql
9.sql
walmartsale
10.sql
(no symbol)
walmartsale
1.sql
2.sql
3.sql
4.sql
5.sql
6.sql
7.sql
8.sql
9.sql
10.sql
11.sql
all.sql
c.sql
untitled.sql
walmart.csv
Weekday	Total Sales (₹)
Sunday	15,425.75
Monday	16,234.80
Tuesday	15,898.30
Wednesday	16,543.65
Thursday	16,987.20
Friday	17,250.55
Saturday	16,000.10
Filter
Ln 10, Col 1
SQL (Generic)
Tabs (4 sp)


Total Sales by Weekday
Sales (₹)
17500
15000
12500
10000
7500
5000
2500
0
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Weekday


Weekly Sales Trend
Sales (₹)
17250
17000
16750
16500
16250
16000
15750
15500
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Weekday


Sales Distribution Across Weekdays
Saturday
Friday
14.0%
15.1%
Sunday
Thursday
13.5%
14.9%
14.2%
14.5%
Monday
13.9%
Wednesday
Tuesday

🎯 Objective:

Analyze total sales across each day of the week to identify patterns in shopping behavior.

📈 Insights:

- **Friday** is the highest-performing day, followed by **Thursday** and **Wednesday**
- **Sunday** shows the lowest sales, possibly due to lower customer engagement or fewer business hours

💡 Business Application:

- Helps plan **staff scheduling**, **marketing campaigns**, and **inventory restocking** based on daily sales patterns
- Ideal for **weekend vs weekday performance analysis**

Conclusion & Business Recommendations

📊 **Summary of Key Insights:**

- **Branch A** leads in average monthly sales growth, showing strong market momentum.
- **Female customers** contribute slightly more to total revenue, suggesting higher engagement.
- **Members** prefer **Food & Beverages**, while **Normal customers** lean toward **Electronic Accessories**.
- **Ewallet** is the preferred payment method in most cities, indicating rising digital adoption.
- Peak sales occur on **Fridays**, making it ideal for promotions and staffing focus.
- **High-value invoices** provide opportunity to identify and nurture VIP buyers.
- Anomalies in transactions were detected in specific categories and payment methods.
- No repeat customer patterns were found—highlighting the need for enhanced customer tracking.

💡 **Strategic Recommendations:**

- 🧠 **Targeted Promotions**: Gender-specific and customer-type based offers.
- 🏙️ **City-wise Strategies**: Customize payment rewards and product placements by city.
- 📆 **Sales Timing**: Align marketing activities with high-performing weekdays.
- 🛡️ **Data Integrity**: Address anomalies to improve reporting accuracy.
- 🔁 **Loyalty Tracking**: Integrate customer IDs for better repeat customer insights.

Heartfelt Thanks & Gratitude

I extend my sincere thanks to all my **teachers, tutors, and mentors** who guided and supported me throughout this project journey.

Your encouragement, insights, and patience played a vital role in helping me understand practical data analysis and build this project from scratch.
This experience has not only enriched my technical skills but also taught me the importance of structured thinking and data-driven decision making.

Thank you for being the foundation behind this learning milestone.

Project Explanation Video Google Drive Link: Project4VideoExplanation

Project Explanation Video Google Drive Link: https://drive.google.com/drive/folders/1ulo7KIWJMWuluf5tbDez2tpWcn-raUyf?usp=sharing