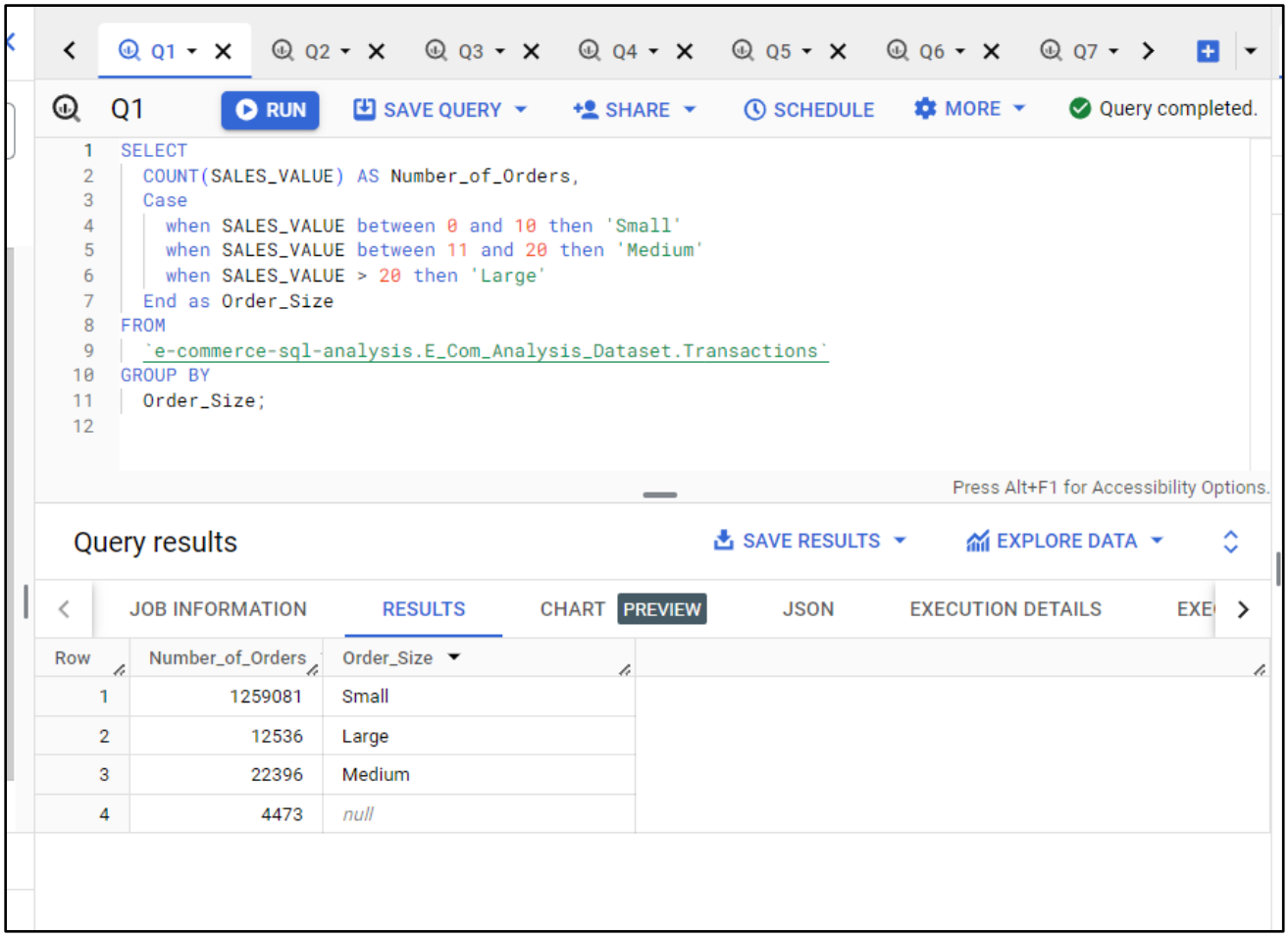
E-commerce SQL Analysis

Problem Statement:

Analysing the sales, product, and customer data for an e-commerce company. Getting various insights and calculating various KPI and data with SQL in Big Query.

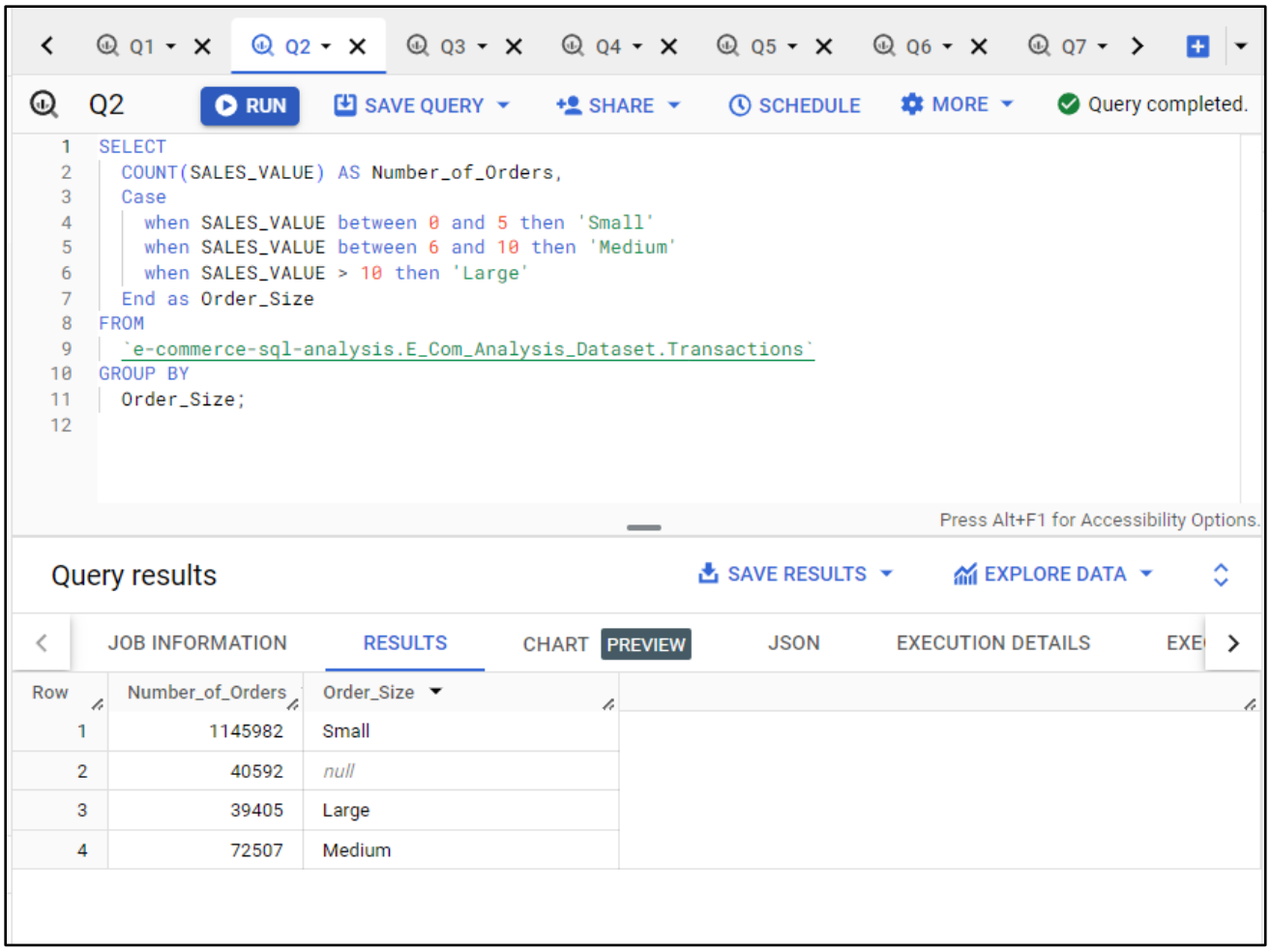
1. Find the number of orders that have small, medium, or large order value (small:0-10 dollars, medium:10-20 dollars, large:20+)

Ans:



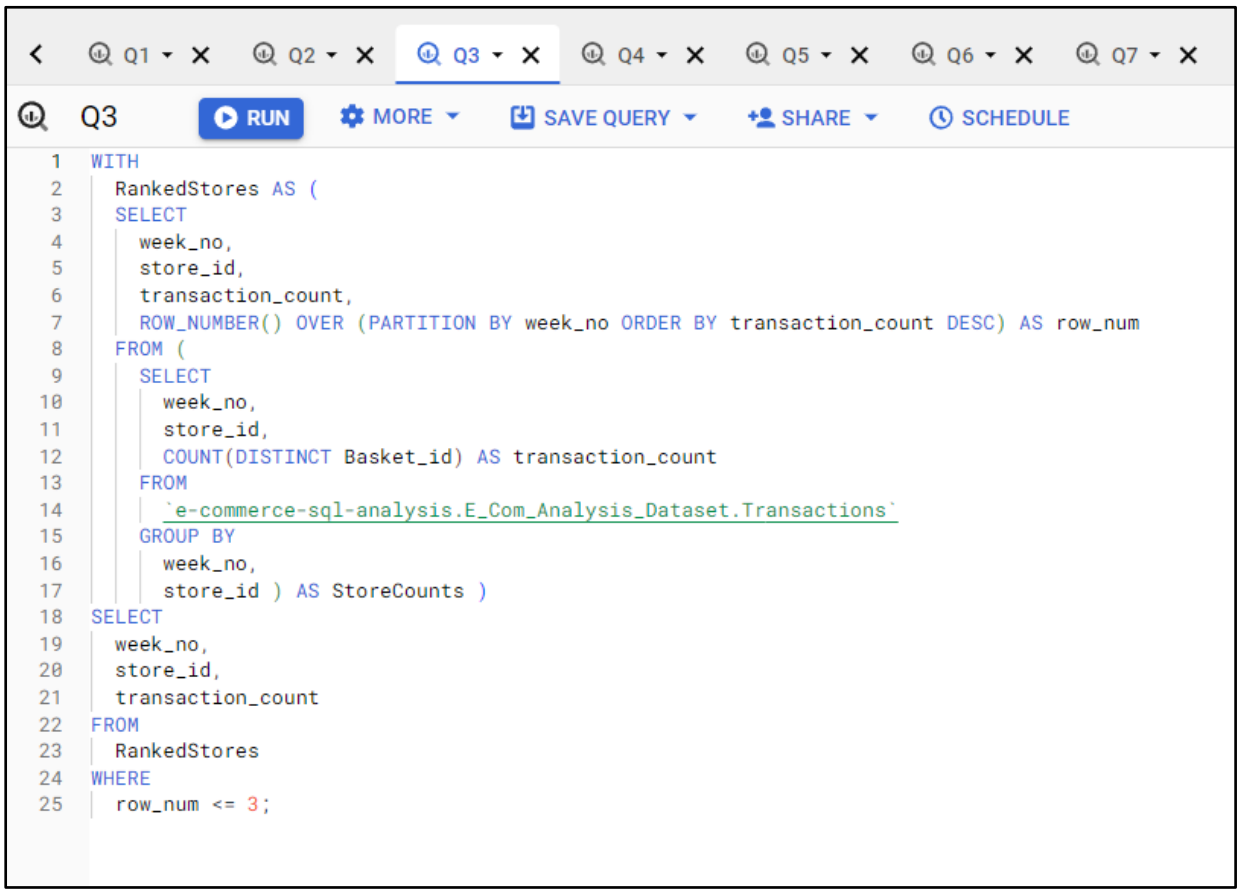
2. Find the number of orders that are small, medium, or large order value (small:0-5 dollars, medium:5-10 dollars, large:10+)

Ans:



3. Find top 3 stores with highest foot traffic for each week (Foot traffic: number of customers transacting)

Ans:



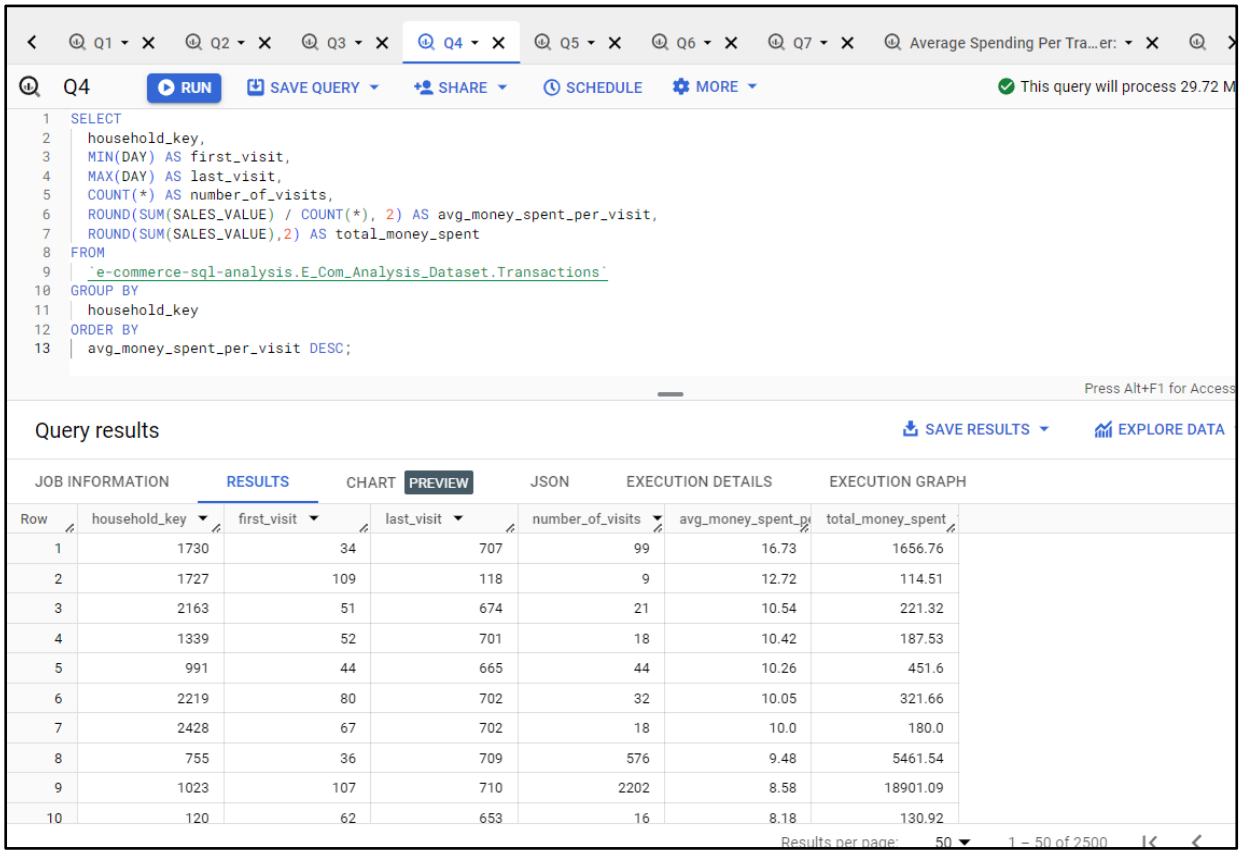
Query results

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON

Row	week_no	store_id	transaction_count
1	28	367	65
2	28	361	60
3	28	406	53
4	62	367	68
5	62	321	52
6	62	406	52
7	15	367	63
8	15	361	59
9	15	381	51
10	42	367	66
11	42	361	58
12	42	343	53
13	11	367	59
14	11	381	41

4. Create a basic customer profiling with first, last visit, number of visits, average money spent per visit and total money spent order by highest average money.

Ans:



5. Do a single customer analysis selecting most spending customer for whom we have demographic information (because not all customers in transaction data are present in demographic table) (show the demographic as well as total spent)

Ans:

Q1 Q2 Q3 Q4 Q5 Q6

Q5 RUN SAVE QUERY SHARE SCHEDULE

```sql
SELECT
  d.Household_key,
  d.Age_desc,
  d.marital_status_code,
  d.income_desc,
  d.homeowner_desc,
  d.HH_comp_desc,
  d.Household_size_desc,
  d.kid_category_desc,
  ROUND(SUM(t.SALES_VALUE),2) AS total_spent
FROM
  `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Demographic` AS d
JOIN
  `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Transactions` AS t
ON
  d.Household_key = t.HOUSEHOLD_KEY
GROUP BY
  d.Household_key,
  d.Age_desc,
  d.marital_status_code,
  d.income_desc,
  d.homeowner_desc,
  d.HH_comp_desc,
  d.Household_size_desc,
  d.kid_category_desc
ORDER BY
  total_spent DESC
LIMIT
  1;
```

Query results SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	Household_key	Age_desc	marital_status_code	income_desc	homeowner_desc	HH_comp_desc	Household_size_desc	kid_category_desc	total_spent
1	1609	45-54	A	125-149K	Homeowner	2 Adults Kids	5+	3+	13804.38

6. Find products (product table: SUB_COMMODITY_DESC) which are most frequently bought together and the count of each combination bought together. do not print a combination twice (A-B / B-A)

Ans:

Q1 Q2 Q3 Q4 Q5 Q6

Q6 RUN SAVE QUERY SHARE SCHEDULE MORE

```sql
SELECT
  p1.SUB_COMMODITY_DESC AS product_1,
  p2.SUB_COMMODITY_DESC AS product_2,
  COUNT(*) AS combination_count
FROM
  `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Transactions` AS t1
JOIN
  `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Transactions` AS t2
ON
  t1.BASKET_ID = t2.BASKET_ID
  AND t1.PRODUCT_ID < t2.PRODUCT_ID
JOIN
  `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Product` AS p1
ON
  t1.PRODUCT_ID = p1.PRODUCT_ID
JOIN
  `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Product` AS p2
ON
  t2.PRODUCT_ID = p2.PRODUCT_ID
GROUP BY
  p1.SUB_COMMODITY_DESC,
  p2.SUB_COMMODITY_DESC
ORDER BY
  combination_count DESC;
```

Query results

SAVE RESULTS

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	product_1	product_2	combination_count
1	YOGURT NOT MULTI-PACKS	YOGURT NOT MULTI-PACKS	15947
2	BABY FOOD - BEGINNER	BABY FOOD - BEGINNER	10080
3	SS ECONOMY ENTREES/DINNERS ALL	SS ECONOMY ENTREES/DINNERS ALL	6633
4	SOFT DRINK POWDER POUCHES	SOFT DRINK POWDER POUCHES	6375
5	FRZN SS PREMIUM ENTREES/DNRS/N	FRZN SS PREMIUM ENTREES/DNRS/N	6340
6	SFT DRNK 2 LITER BTL CARB INCL	SFT DRNK 2 LITER BTL CARB INCL	5459
7	SOFT DRINKS 12/18&15PK CAN CAR	SOFT DRINKS 12/18&15PK CAN CAR	5173
8	CANDY BARS (SINGLES)(INCLUDING	CANDY BARS (SINGLES)(INCLUDING	4194
9	BABY FOOD JUNIOR ALL BRANDS	BABY FOOD JUNIOR ALL BRANDS	3751
10	FLUID MILK WHITE ONLY	SOFT DRINKS 12/18&15PK CAN CAR	3580
11	FLUID MILK WHITE ONLY	YOGURT NOT MULTI-PACKS	3520
12	CANNED CAT FOOD (9 LIVES/FRISK	CANNED CAT FOOD (9 LIVES/FRISK	3477
13	MAINSTREAM WHITE BREAD	FLUID MILK WHITE ONLY	3158
14	FLUID MILK WHITE ONLY	BANANAS	2971

Results per page: 50 1 – 5

7. Find the weekly change in Revenue Per Account (RPA) (difference in spending by each customer compared to last week) (use lag function).

Ans:

Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Average Spending Per Tra...er:

Q7 RUN SAVE QUERY SHARE SCHEDULE MORE

```sql
WITH
  WeeklySpending AS (
  SELECT
    Household_key,
    WEEK_NO,
    ROUND(SUM(SALES_VALUE),2) AS weekly_spending
  FROM
    `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Transactions`
  GROUP BY
    Household_key,
    WEEK_NO
  ORDER BY
    Household_key,
    WEEK_NO )
SELECT
  Household_key,
  WEEK_NO,
  weekly_spending,
  LAG(weekly_spending) OVER (PARTITION BY Household_key ORDER BY WEEK_NO) AS previous_week_spending,
  ROUND(weekly_spending - LAG(weekly_spending) OVER (PARTITION BY Household_key ORDER BY WEEK_NO),2) AS weekly_change_in_RPA
FROM
  WeeklySpending
ORDER BY
  Household_key,
  WEEK_NO;
```

Press Alt+F

Query results

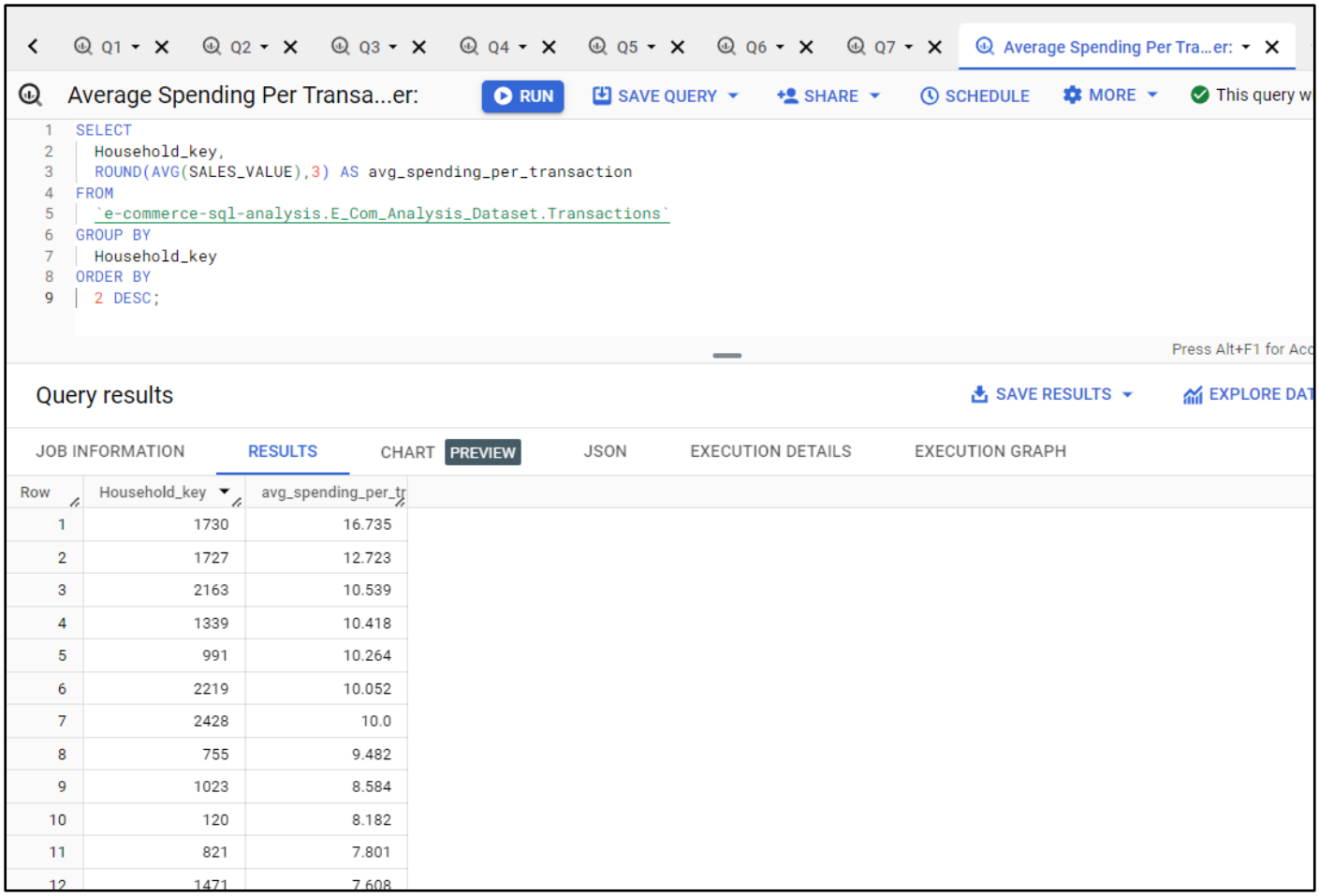
JOB INFORMATION | RESULTS | CHART PREVIEW | JSON | EXECUTION DETAILS | EXECU

Row	Household_key	WEEK_NO	weekly_spending	previous_week_spen	weekly_change_in_R
1	1	8	42.58	null	null
2	1	10	14.01	42.58	-28.57
3	1	13	14.03	14.01	0.02
4	1	14	25.71	14.03	11.68
5	1	15	10.98	25.71	-14.73
6	1	16	9.09	10.98	-1.89
7	1	17	13.98	9.09	4.89
8	1	19	47.35	13.98	33.37
9	1	20	31.77	47.35	-15.58
10	1	22	38.98	31.77	7.21
11	1	23	26.36	38.98	-12.62

Results per page

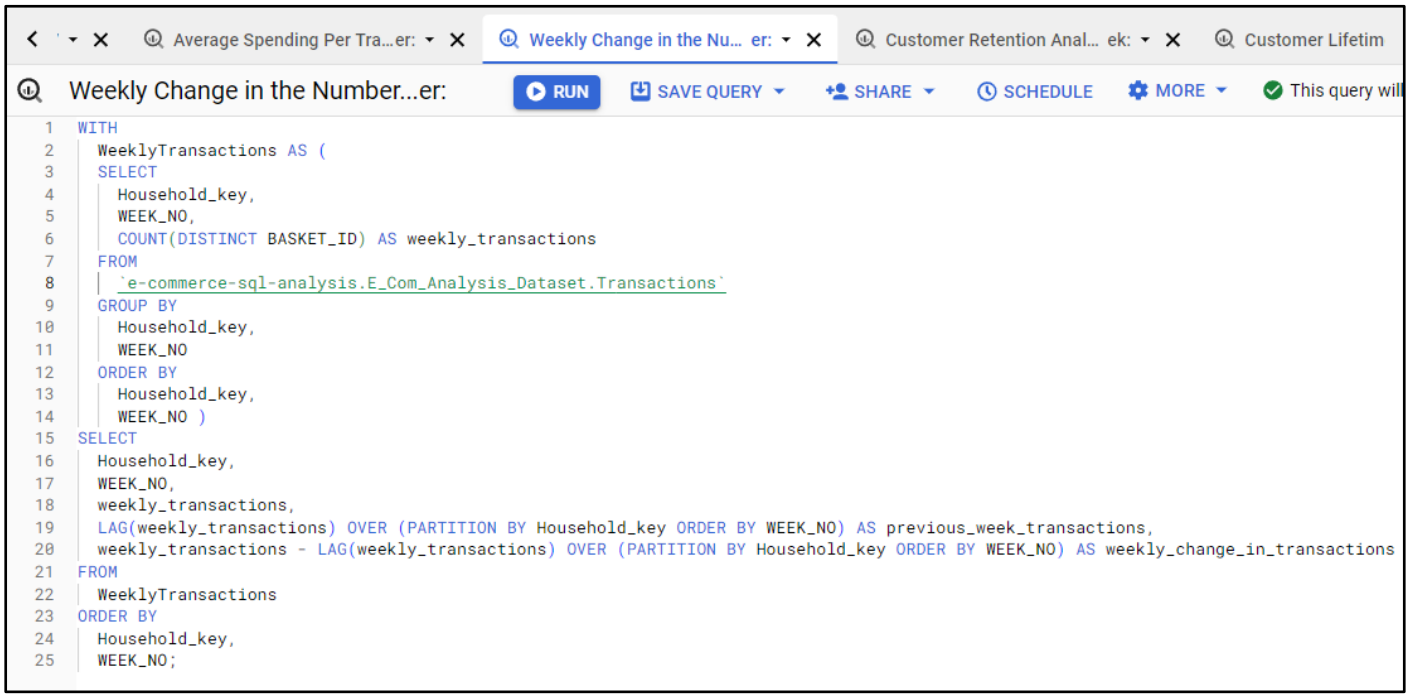
8. Average Spending Per Transaction for Each Customer.

Ans:



9. Weekly Change in the Number of Transactions for Each Customer.

Ans:



JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	Household_key	WEEK_NO	weekly_transactions	previous_week_transactions	weekly_change_in_transactions
1	1	8	1	*null*	*null*
2	1	10	1	1	0
3	1	13	1	1	0
4	1	14	1	1	0
5	1	15	1	1	0
6	1	16	1	1	0
7	1	17	1	1	0
8	1	19	1	1	0
9	1	20	1	1	0
10	1	22	1	1	0
11	1	23	1	1	0
12	1	24	1	1	0
13	1	25	1	1	0
14	1	26	1	1	0
15	1	28	1	1	0
16	1	30	1	1	0
17	1	31	1	1	0

10. Customer Retention Analysis Over Week on Week:

Ans:

Average Spending Per Tra...er: | Weekly Change in the Nu... er: | Customer Retention Anal... ek:

Customer Retention Analysis...ek: RUN | SAVE QUERY | SHARE | SCHEDULE | MORE

```sql
WITH
  Retention AS (
  SELECT
    Household_key,
    WEEK_NO,
    COUNT(DISTINCT BASKET_ID) AS weekly_transactions
  FROM
    `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Transactions`
  GROUP BY
    Household_key,
    WEEK_NO
  ORDER BY
    Household_key,
    WEEK_NO )
SELECT
  Household_key,
  WEEK_NO,
  weekly_transactions,
  LAG(weekly_transactions, 1, 0) OVER (PARTITION BY Household_key ORDER BY WEEK_NO) AS previous_week_transactions,
  CASE
    WHEN LAG(weekly_transactions, 1, 0) OVER (PARTITION BY Household_key ORDER BY WEEK_NO) = 0 THEN 0
  ELSE
  (weekly_transactions / LAG(weekly_transactions, 1, 0) OVER (PARTITION BY Household_key ORDER BY WEEK_NO)) * 100
END
  AS retention_percentage
FROM
  Retention
ORDER BY
  Household_key,
  WEEK_NO;
```

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON | EXECUTION DETAILS | EXE

Row	Household_key	WEEK_NO	weekly_transactions	previous_week_transactions	retention_percentage
1	1	8	1	0	0.0
2	1	10	1	1	100.0
3	1	13	1	1	100.0
4	1	14	1	1	100.0
5	1	15	1	1	100.0
6	1	16	1	1	100.0
7	1	17	1	1	100.0
8	1	19	1	1	100.0
9	1	20	1	1	100.0
10	1	22	1	1	100.0
11	1	23	1	1	100.0
12	1	24	1	1	100.0
13	1	25	1	1	100.0

11. Customer Lifetime Value (CLV) Calculation.

Ans:

e Spending Per Tra...er: | Weekly Change in the Nu... er: | Customer Retention Anal... ek:

Customer Lifetime Value (CL...on: RUN | SAVE QUERY | SHARE | SCHEDU

```sql
SELECT
  Household_key,
  COUNT(DISTINCT BASKET_ID) AS total_transactions,
  ROUND(SUM(SALES_VALUE),2) AS total_spending,
  ROUND(SUM(SALES_VALUE) / COUNT(DISTINCT BASKET_ID),2) AS avg_transaction_value,
  ROUND(COUNT(DISTINCT BASKET_ID) / COUNT(DISTINCT WEEK_NO),2) AS frequency,
  ROUND((SUM(SALES_VALUE) / COUNT(DISTINCT BASKET_ID)) * (COUNT(DISTINCT BASKET_ID) / COUNT(DISTINC
FROM
  `e-commerce-sql-analysis.E_Com_Analysis_Dataset.Transactions`
GROUP BY
  Household_key
ORDER BY
  CLV DESC
LIMIT
  5;
```

Query results

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON | EXECUTION DETAILS | EXECUTION

Row	Household_key	total_transactions	total_spending	avg_transaction_value	frequency	CLV
1	1023	422	18901.09	44.79	5.02	225.01
2	1609	324	13804.38	42.61	3.41	145.31
3	2322	290	11934.66	41.15	3.19	131.15
4	1453	591	10720.72	18.14	6.79	123.23
5	973	80	6875.89	85.95	1.36	116.54

Insights and Recommendations:

1. *Leverage Complementary Product Insights:*
   Identify frequently paired products like yogurt variants or baby foods. Use this knowledge to craft bundled promotions, strategically place related items together, and optimize marketing strategies to encourage cross-selling and enhance customer shopping experiences.
2. *Optimize Inventory and Marketing Strategies:*
   Utilize observed frequent product combinations to refine inventory management, ensuring adequate stock levels for frequently paired items. Tailor marketing campaigns to promote cross-category sales, exploring packaging options and pricing strategies to capitalize on these commonly paired products.
3. *Monitor Weekly Spending Trends:*
   Track household spending week-over-week to identify fluctuations in spending patterns. Analyze changes, especially drastic shifts, to understand potential factors driving increased or decreased spending per household.
4. *Customer Spending Stability Assessment:*
   Assess household spending stability using the calculated weekly change in Revenue Per Account (RPA). Identify customers with consistent spending versus those displaying significant spending variations, aiding in targeted retention or engagement strategies.
5. *Identify High-Value Customers:*
   Highlight customers with significantly higher average spending per transaction. Target these high-value customers for personalized offers, loyalty programs, or exclusive perks to foster continued engagement and increase their lifetime value.
6. *Segment Customers by Spending Behaviour:*
   Segment customers based on their average spending per transaction to tailor marketing strategies. For instance, focus on different messaging or incentives for segments with varying spending levels to optimize conversions and enhance customer satisfaction.
7. *Retaining Consistent Transaction Frequency:*
   Monitor households displaying consistent or increasing transaction frequency week-over-week. Focus on retaining this trend to ensure sustained customer engagement and potentially higher retention rates.
8. *Addressing Fluctuating Transaction Patterns:*
   Identify households with fluctuating transaction frequencies. Investigate and implement targeted strategies to stabilize or improve their engagement, aiming for consistent transaction behaviour to enhance customer loyalty.
9. *Leveraging High CLV Customers:*
   Identify and focus on customers with higher Customer Lifetime Value (CLV). Allocate resources for personalized experiences, loyalty programs, or targeted campaigns to nurture and retain these valuable customers.
10. *Improving Engagement for Low-Frequency Customers:*
    Engage with customers having lower transaction frequency but potential for increased CLV. Implement strategies to increase their transaction frequency, such as personalized offers or improved customer experiences, aiming to enhance their long-term value to the business.