COFFEE SHOP SALES ANALYSIS

A SQL-based Data Analysis Report

You are only one sip away from good mood.

📌 PROJECT OVERVIEW

This project analyzes a coffee shop's sales data using SQL.

The goal is to derive business insights, understand customer behavior, and optimize sales strategies.

The dataset includes transactions, product details, sales amounts, and timestamps.

🎯 Objectives

- Identify key sales trends and revenue drivers
- Analyze customer purchasing patterns
- Determine peak sales hours and seasonal trends
- Evaluate product performance and store efficiency
- Extract insights for data-driven decision-making

❓ Key Business Questions Answered

Basic Questions

1. What is the total sales revenue?
2. How many transactions occurred in total?
3. What are the most frequently purchased products?
4. What are the different product categories available?
5. What is the average unit price of all products?
6. Which store location has the highest number of transactions?

- **Intermediate Questions**

1. What is the total revenue generated per store location?
2. What are the top 5 best-selling products?
3. What is the average number of products purchased per transaction?
4. What is the sales distribution across different product categories?
5. Which time of the day has the highest number of transactions?

- **Advanced Questions**

1. What are the peak sales hours for each store location?
2. How does the sales trend vary across different months?
3. What is the correlation between product price and quantity sold?
4. Are there any seasonal patterns in sales?
5. Which products have the highest revenue contribution across all stores?
6. Can we identify customer purchasing patterns based on time and location?

TOTAL SALES REVENUE

1. What is the total sales revenue?

```
SELECT round( sum(transaction_qty * unit_price),2)
    AS total_sales_revenue
FROM coffee_shop;
```

	total_sales_revenue
▶	573729.79

TOTAL TRANSACTIONS

2. How many transactions occurred in total?

```
select
    count(transaction_id)
    as total_transactions
from coffee_shop;
```

total_transactions
122932

MOST FREQUENTLY PURCHASED PRODUCTS

3. What are the most frequently purchased products?

```
select
    product_detail ,
    count(transaction_id) as frequently_purchased
from coffee_shop
group by product_detail
order by frequently_purchased desc;
```

product_detail	frequently_purchased
Earl Grey Rg	2555
Chocolate Croissant	2537
Morning Sunrise Chai Rg	2532
Dark chocolate Lg	2527
Columbian Medium Roast Rg	2508
Latte	2480

AVAILABLE PRODUCT CATEGORIES

4. What are the different product categories available?

```
select distinct product_category
from coffee_shop;
```

product_category
Coffee
Tea
Drinking Chocolate
Bakery
Flavours
Loose Tea
Coffee beans
Packaged Chocolate
Branded

AVERAGE UNIT PRICE

5. What is the average unit price of all products?

```
select round(avg(unit_price),2)
    as avg_unit_price
from coffee_shop;
```

	avg_unit_price
▶	3.37

STORE LOCATION WITH HIGHEST TRANSACTIONS

6. Which store location has the highest number of transactions?

```
select store_location,
    round(count(transaction_id),2)
    as total_transactions from coffee_shop
group by store_location
order by total_transactions desc;
```

store_location	total_transactions
Hell's Kitchen	41832
Astoria	41823
Lower Manhattan	39277

TOTAL REVENUE PER STORE LOCATION

7.What is the total revenue generated per store location?

```
select store_location,
        round(sum(transaction_id * unit_price),2)
        as total_revenue
from coffee_shop
group by store_location
order by total_revenue desc;
```

store_location	total_revenue
Astoria	8718164204.39
Hell's Kitchen	8646136942.34
Lower Manhattan	8078576855.9

BEST SELLING PRODUCTS

8.What are the top 5 best-selling products?

```
select product_detail,
        count(transaction_id) as sales
from coffee_shop
group by product_detail
order by sales desc
limit 5;
```

	product_detail	sales
▶	Earl Grey Rg	2555
	Chocolate Croissant	2537
	Morning Sunrise Chai Rg	2532
	Dark chocolate Lg	2527
	Columbian Medium Roast Rg	2508

SALES DISTRIBUTION BY PRODUCT CATEGORIES

9.What is the sales distribution across different product categories?

```
select product_category,
        sum(transaction_qty) as total_unit_sold,
        round((sum(transaction_qty)*100 / (select sum(transaction_qty)
from coffee_shop)),2) as percentage_share
from coffee_shop
group by product_category
order by total_unit_sold desc;
```

product_category	total_unit_sold	percentage_share
Coffee	73562	41.64
Tea	57920	32.79
Bakery	19145	10.84
Drinking Chocolate	14461	8.19
Flavours	8201	4.64
Coffee beans	1441	0.82
Loose Tea	942	0.53
Branded	586	0.33
Packaged Chocolate	387	0.22

AVERAGE PRODUCTS PURCHASED PER TRANSACTION

10.What is the average number of products purchased per transaction?

```
select
    avg(transaction_qty) as avg_product_sold
from coffee_shop;
```

	avg_product_sold
▶	1.4369

HIGHEST TRANSACTION BT TIME

11.Which time of the day has the highest number of transactions?

```
SELECT
    CASE
      WHEN EXTRACT(HOUR FROM transaction_time) BETWEEN 6 AND 11 THEN 'Morning'
      WHEN EXTRACT(HOUR FROM transaction_time) BETWEEN 12 AND 17 THEN 'Afternoon'
      WHEN EXTRACT(HOUR FROM transaction_time) BETWEEN 18 AND 21 THEN 'Evening'
    ELSE 'Night'
END AS time_of_day,
COUNT(*) AS total_transactions
FROM coffee_shop
GROUP BY time_of_day
ORDER BY total_transactions DESC;
```

time_of_day	total_transactions
Morning	65798
Afternoon	45125
Evening	12009

PEAK SALES HOURS

12. What are the peak sales hours for each store location?

```
SELECT store_location,
    EXTRACT(HOUR FROM transaction_time) AS Peak_hour,
    COUNT(*) AS Total_Transactions
FROM coffee_shop
GROUP BY store_location, Peak_hour
ORDER BY Total_Transactions DESC ;
```

store_location	Peak_hour	Total_Transactions
Hell's Kitchen	10	5649
Hell's Kitchen	8	5582
Hell's Kitchen	9	5439
Lower Manhattan	10	5088
Lower Manhattan	9	4846
Lower Manhattan	8	4720
Lower Manhattan	7	4719
Astoria	10	4093
Astoria	9	3963

Result 20

SALES TREND ACROSS MONTHS

13.How does the sales trend vary across different months?

```
SELECT
    EXTRACT(MONTH FROM transaction_date) AS MONTH,
    SUM(transaction_qty) AS TOTAL_SALES,
    ROUND(SUM(transaction_qty * unit_price),2)
   AS TOTAL_SALES_REVENUE
FROM coffee_shop
GROUP BY  month
ORDER BY TOTAL_SALES DESC;
```

	MONTH	TOTAL_SALES	TOTAL_SALES_REVENUE
▶	5	48233	156727.76
	4	36469	118941.08
	3	30406	98834.68
	1	24870	81677.74
	2	23550	76145.19
	6	13117	41403.34

SALES TREND BY YEAR AND MONTH

```
SELECT
    EXTRACT(YEAR FROM transaction_date) AS year,
    EXTRACT(MONTH FROM transaction_date) AS month,
    SUM(transaction_qty) AS total_units_sold,
    ROUND(SUM(transaction_qty*unit_price),2) AS total_sales_revenue
FROM coffee_shop
GROUP BY year, month
ORDER BY year, month;
```

year	month	total_units_sold	total_sales_revenue
2023	1	24870	81677.74
2023	2	23550	76145.19
2023	3	30406	98834.68
2023	4	36469	118941.08
2023	5	48233	156727.76
2023	6	13117	41403.34

CORRELATION

14.What is the correlation between product price and quantity sold?

```
SELECT
    ROUND( (SUM(unit_price * transaction_qty) -
            (SUM(unit_price) * SUM(transaction_qty) /
            COUNT(*))) /
            (SQRT(SUM(unit_price * unit_price) -
            (SUM(unit_price) * SUM(unit_price) /COUNT(*))) *
            SQRT(SUM(transaction_qty * transaction_qty) -
            (SUM(transaction_qty) *
            SUM(transaction_qty) / COUNT(*)))),4) AS price_quantity_correlation
FROM coffee_shop;
```

price_quantity_correlation
-0.1257

- **Formula Used:**

  The formula in your query follows the **Pearson correlation coefficient** formula:

$$r = \frac{\sum(XY) - \frac{\sum X \sum Y}{N}}{\sqrt{\sum X^2 - \frac{(\sum X)^2}{N}} \times \sqrt{\sum Y^2 - \frac{(\sum Y)^2}{N}}}$$

SEASONAL PATTERNS

15.Are there any seasonal patterns in sales?

```
SELECT
    EXTRACT(YEAR FROM transaction_date) AS year,
    EXTRACT(MONTH FROM transaction_date) AS month,
    ROUND(SUM(transaction_qty * unit_price),2) AS total_sales
FROM coffee_shop
GROUP BY year, month
ORDER BY year, month;
```

year	month	total_sales
2023	1	81677.74
2023	2	76145.19
2023	3	98834.68
2023	4	118941.08
2023	5	156727.76
2023	6	41403.34

REVENUE CONTRIBUTION

16.Which products have the highest revenue contribution across all stores?

```
SELECT
    product_detail,
    ROUND(SUM(transaction_qty * unit_price),2)
    AS TOTAL_REVENUE_SALES
FROM coffee_shop
GROUP BY product_detail
ORDER BY TOTAL_REVENUE_SALES DESC
LIMIT 10;
```

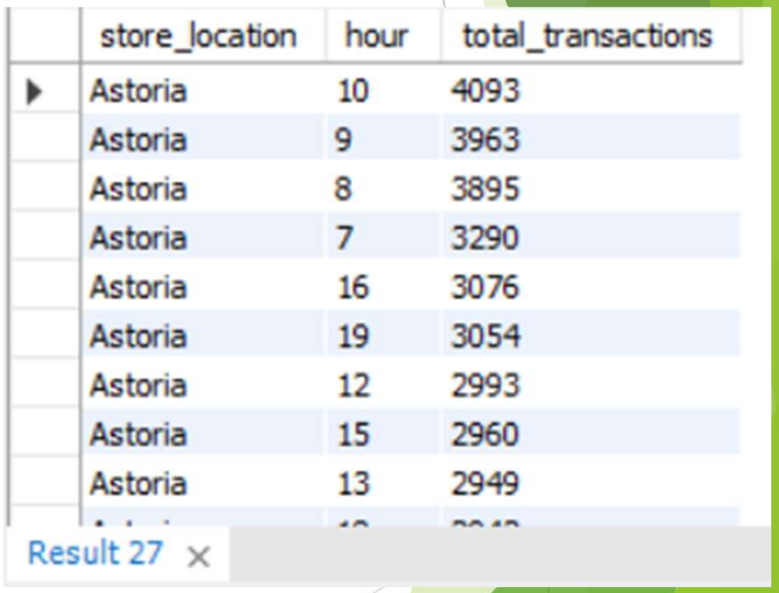
product_detail	TOTAL_REVENUE_SALES
Sustainably Grown Organic Lg	17541.75
Dark chocolate Lg	17446.5
Latte Rg	15759
Cappuccino Lg	14526.5
Morning Sunrise Chai Lg	14364
Latte	14231.25
Jamaican Coffee River Lg	13627.5
Sustainably Grown Organic Rg	13436.25
Cappuccino	13196.25

Result 26

CUSTOMER PURCHASING PATTERNS

17.Can we identify customer purchasing patterns based on time and location?

```
SELECT
    store_location,
    EXTRACT(HOUR FROM transaction_time) AS hour,
    COUNT(*) AS total_transactions
FROM coffee_shop
GROUP BY store_location, hour
ORDER BY store_location, total_transactions DESC;
```



store_location	hour	total_transactions
Astoria	10	4093
Astoria	9	3963
Astoria	8	3895
Astoria	7	3290
Astoria	16	3076
Astoria	19	3054
Astoria	12	2993
Astoria	15	2960
Astoria	13	2949

Result 27

PEAK SHOPPING DAYS

Finding Peak Shopping Days Per Store

```
SELECT
    store_location,
    EXTRACT(DAY FROM transaction_date) AS DAY,
    COUNT(*) AS total_transactions
FROM coffee_shop
GROUP BY store_location, DAY
ORDER BY store_location, total_transactions DESC;
```

store_location	DAY	total_transactions
Astoria	3	1822
Astoria	2	1784
Astoria	5	1752
Astoria	8	1714
Astoria	1	1693
Astoria	4	1653
Astoria	15	1580
Astoria	19	1546
Astoria	7	1522

Result 28

📌 Key Findings & Insights

- Peak Sales Hours: Sales peak in the morning.
- Bestselling Products: Earl Grey Tea and Chocolate Croissant had the highest purchase count.
- Seasonal Trends: Sales increase during April and May.
- Store Performance: The 'Astoria' location generated the highest revenue.
- Revenue Drivers: High-priced items do not always have the highest sales volume.

✅ CONCLUSION

This project provided valuable insights into customer preferences, peak sales times, and product performance. The findings can help coffee shop owners make data-driven business decisions, such as optimizing inventory, pricing strategies, and marketing campaigns.

THANK YOU