

DATA ANALYST

Intenship Task 8

DESCRIPTION

The task focuses on analyzing sales data using SQL window functions in a database. It includes aggregating total sales per customer, ranking customers by region, calculating running totals, measuring monthly sales growth, and identifying top products per category. The goal is to extract meaningful business insights.

PREPARED BY

Reema Safrin M
(27-01-2026)

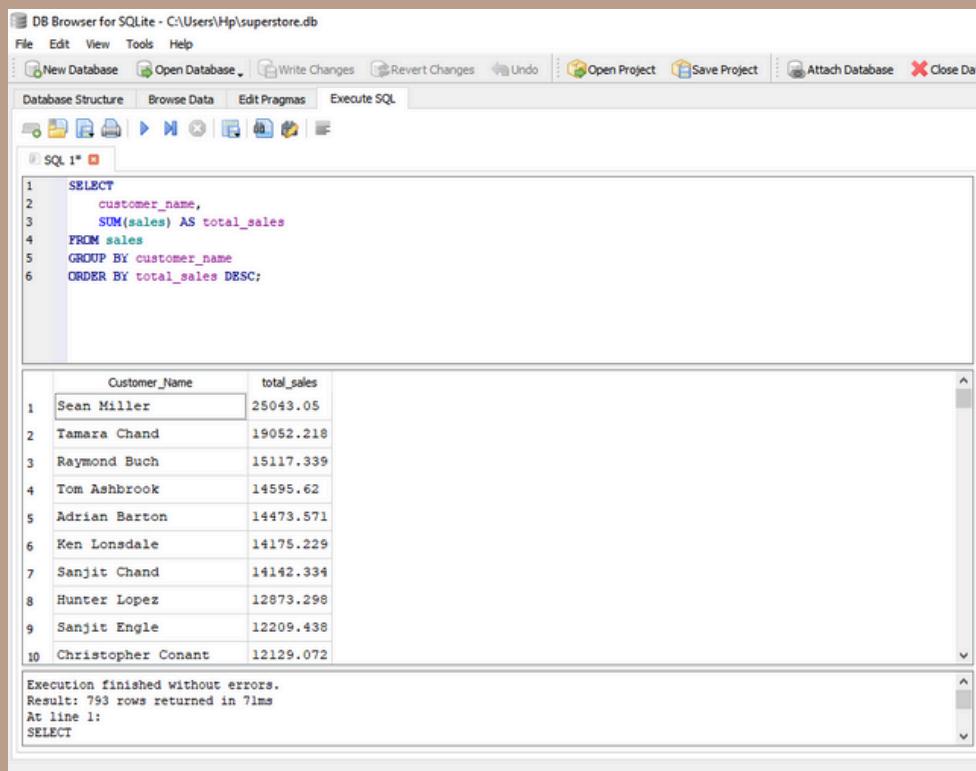
MY WORK

I imported the cleaned Superstore dataset into a database and created the necessary table structure. I executed SQL queries to calculate total sales per customer and ranked customers by region using RANK and DENSE_RANK window functions. I generated running totals to visualize cumulative sales trends, calculated month-over-month sales growth using LAG, and identified the top 3 products in each category. All results were verified, exported as CSV files, and analyzed for patterns. This process allowed me to understand regional performance, customer behavior, product trends, and overall sales dynamics within the dataset, providing actionable insights for business decisions.

DATASET

[superstore_dataset](#)

TOTAL SALES PER CUSTOMER:



The screenshot shows the DB Browser for SQLite interface. The SQL tab contains the following query:

```
1 SELECT
2     customer_name,
3     SUM(sales) AS total_sales
4 FROM sales
5 GROUP BY customer_name
6 ORDER BY total_sales DESC;
```

The results pane displays a table with two columns: Customer_Name and total_sales. The data is as follows:

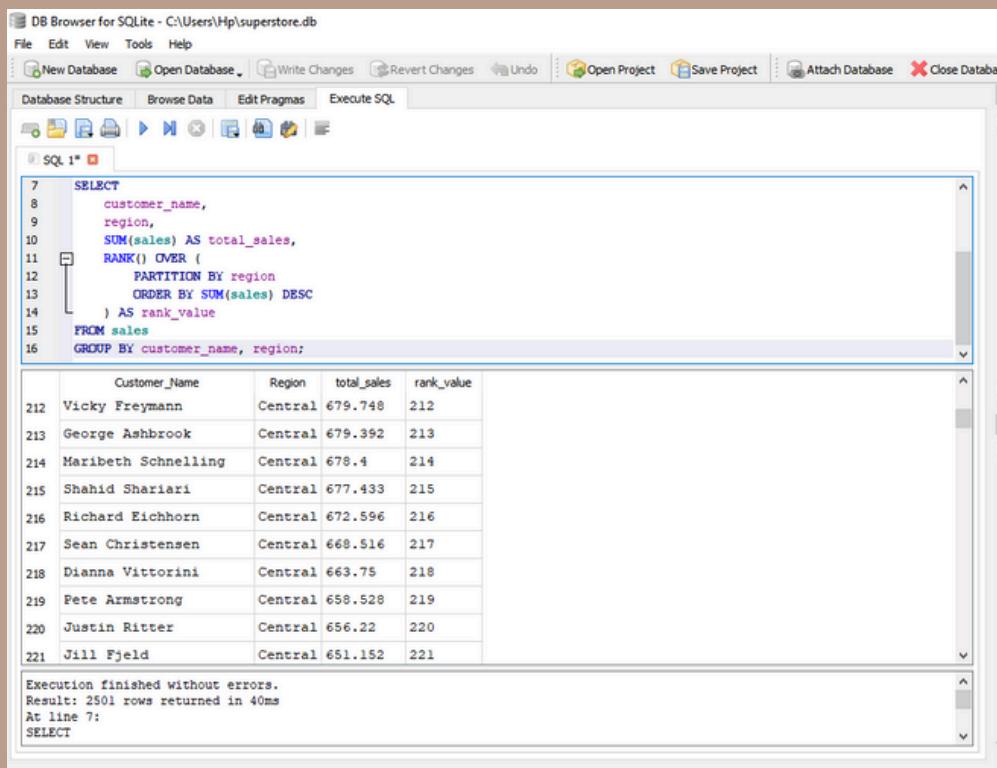
	Customer_Name	total_sales
1	Sean Miller	25043.05
2	Tamara Chand	19052.218
3	Raymond Buch	15117.339
4	Tom Ashbrook	14595.62
5	Adrian Barton	14473.571
6	Ken Lonsdale	14175.229
7	Sanjit Chand	14142.334
8	Hunter Lopez	12873.298
9	Sanjit Engle	12209.438
10	Christopher Conant	12129.072

Execution finished without errors.
Result: 793 rows returned in 71ms
At line 1:
SELECT

output csv

Customers in certain regions contributed significantly higher sales, indicating key high-value clients. Overall, sales were concentrated among top-performing customers, highlighting the importance of focusing on loyal or repeat customers.

RANKING CUSTOMERS BY REGION :



The screenshot shows the DB Browser for SQLite interface. The SQL tab contains the following query:

```
7  SELECT
8      customer_name,
9      region,
10     SUM(sales) AS total_sales,
11     RANK() OVER (
12         PARTITION BY region
13         ORDER BY SUM(sales) DESC
14     ) AS rank_value
15   FROM sales
16  GROUP BY customer_name, region;
```

The results pane displays a table with the following data:

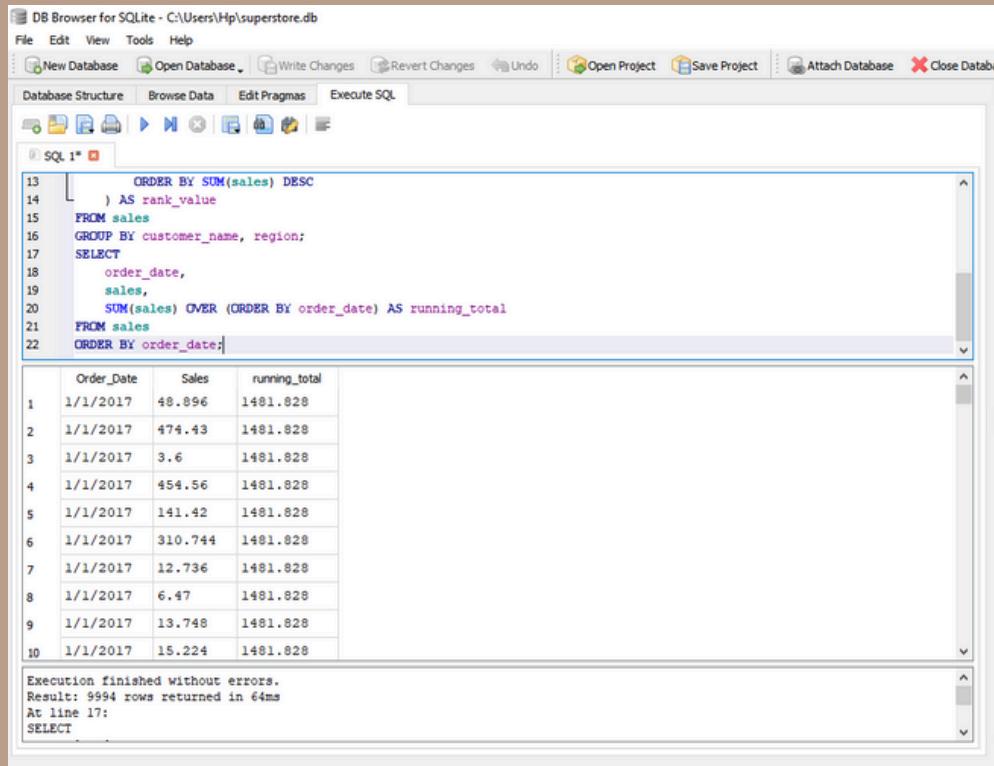
	Customer_Name	Region	total_sales	rank_value
212	Vicky Freymann	Central	679.748	212
213	George Ashbrook	Central	679.392	213
214	Maribeth Schnelling	Central	678.4	214
215	Shahid Shariari	Central	677.433	215
216	Richard Eichhorn	Central	672.596	216
217	Sean Christensen	Central	668.516	217
218	Dianna Vittorini	Central	663.75	218
219	Pete Armstrong	Central	658.528	219
220	Justin Ritter	Central	656.22	220
221	Jill Fjeld	Central	651.152	221

The message pane at the bottom indicates: Execution finished without errors. Result: 2501 rows returned in 40ms At line 7: SELECT

output csv

- Ranking showed that some customers consistently dominated sales in their regions. RANK and DENSE_RANK comparisons revealed ties, emphasizing multiple top contributors in certain areas, which can guide regional marketing and sales strategies.

RUNNING TOTAL :



The screenshot shows the DB Browser for SQLite interface. The SQL tab contains the following code:

```
13 ORDER BY SUM(sales) DESC
14 ) AS rank_value
15 FROM sales
16 GROUP BY customer_name, region;
17 SELECT
18     order_date,
19     sales,
20     SUM(sales) OVER (ORDER BY order_date) AS running_total
21 FROM sales
22 ORDER BY order_date;
```

The results pane displays a table with three columns: Order_Date, Sales, and running_total. The data is as follows:

	Order_Date	Sales	running_total
1	1/1/2017	48.096	1481.828
2	1/1/2017	474.43	1481.828
3	1/1/2017	3.6	1481.828
4	1/1/2017	454.56	1481.828
5	1/1/2017	141.42	1481.828
6	1/1/2017	310.744	1481.828
7	1/1/2017	12.736	1481.828
8	1/1/2017	6.47	1481.828
9	1/1/2017	13.748	1481.828
10	1/1/2017	15.224	1481.828

The message pane at the bottom indicates: Execution finished without errors. Result: 9994 rows returned in 64ms At line 17: SELECT

output csv

- Cumulative sales revealed seasonal peaks and consistent growth trends over time. This helped identify periods of high sales activity, supporting better inventory planning and sales forecasting

MONTHLY GROWTH :

The screenshot shows the DB Browser for SQLite interface. The SQL tab contains the following code:

```
27   FROM sales
28   GROUP BY month
29 )
30
31 SELECT
32   month,
33   total_sales,
34   LAG(total_sales) OVER (ORDER BY month) AS prev_month_sales,
35   total_sales - LAG(total_sales) OVER (ORDER BY month) AS growth
36 FROM monthly_sales;
```

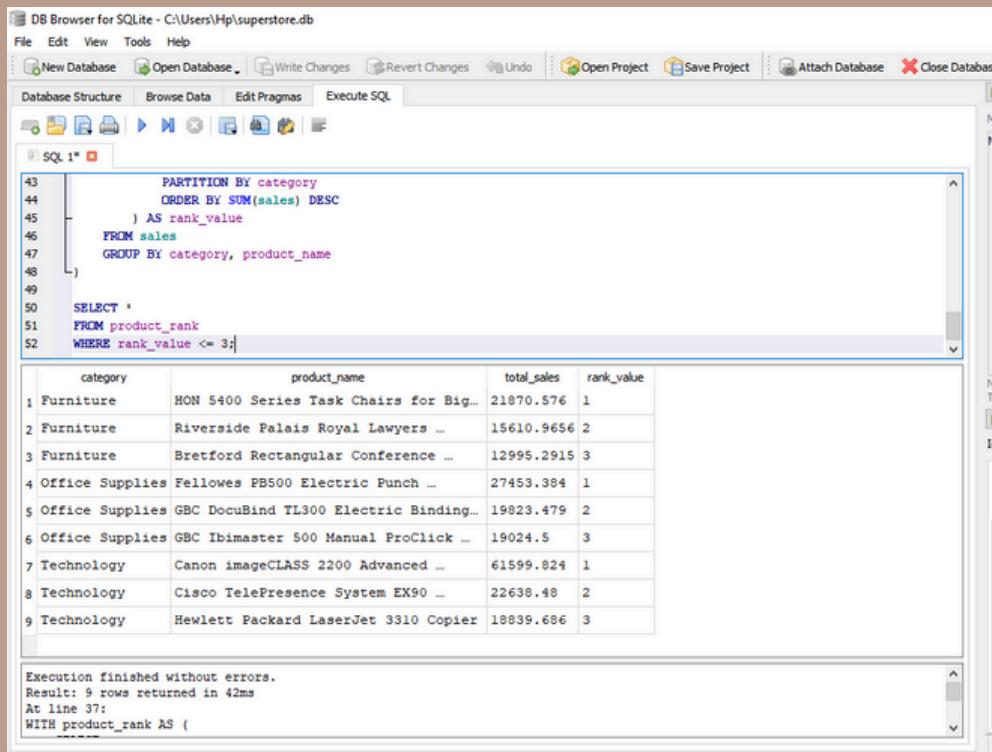
The results pane displays a single row of data:

month	total_sales	prev_month_sales	growth
1	2297200.8603	NULL	NULL

Below the results, the message "Execution finished without errors." is displayed.

- Month-over-month growth calculation showed fluctuations in sales, identifying months with significant increases or decreases. This insight helps pinpoint trends, seasonal effects, and areas needing promotional focus.

TOP 3 PRODUCTS PER CATEGORY :



The screenshot shows the DB Browser for SQLite interface. The SQL tab contains the following code:

```
43   PARTITION BY category
44     ORDER BY SUM(sales) DESC
45   ) AS rank_value
46   FROM sales
47   GROUP BY category, product_name
48
49
50   SELECT *
51   FROM product_rank
52   WHERE rank_value <= 3;
```

The results pane displays a table with the following data:

category	product_name	total_sales	rank_value
1 Furniture	HON 5400 Series Task Chairs for Big...	21870.576	1
2 Furniture	Riverside Palais Royal Lawyers ...	15610.5656	2
3 Furniture	Bretford Rectangular Conference ...	12995.2915	3
4 Office Supplies	Fellowes PB500 Electric Punch ...	27453.384	1
5 Office Supplies	GBC DocuBind TL300 Electric Binding...	19823.479	2
6 Office Supplies	GBC Ibimaster 500 Manual ProClick ...	19024.5	3
7 Technology	Canon imageCLASS 2200 Advanced ...	61599.824	1
8 Technology	Cisco TelePresence System EX90 ...	22638.48	2
9 Technology	Hewlett Packard LaserJet 3310 Copier	18839.686	3

The message pane at the bottom indicates:

```
Execution finished without errors.
Result: 9 rows returned in 42ms
At line 37:
WITH product_rank AS (
```

output csv

- Certain products consistently drove the most revenue in their categories. Focusing on these high-performing products can optimize marketing, inventory, and cross-selling opportunities for maximum profitability.