Superstore Sales Analysis Using SQL

Project Title: Superstore Sales & Profitability Analysis Using SQL
Done By: Shivanand Metri

---

1. Problem Statement

Retail organizations collect large volumes of transactional data, but raw data alone does not provide business value unless it is properly analyzed. The Superstore dataset contains detailed information about customer orders, products, regions, sales, discounts, and profits.

The main problem is to **identify sales patterns, profitability drivers, customer behavior, and operational inefficiencies** using SQL-based data analysis. Without this analysis, the business risks offering excessive discounts, continuing loss-making products, and misallocating resources across regions and customer segments.

Objective

To use SQL queries on the Superstore dataset to: - Evaluate overall business performance - Identify profitable and loss-making areas - Understand customer and product behavior - Support data-driven decision-making

---

2. Dataset Description

The analysis is performed on a **single Superstore table** with all column names in **lowercase**.

Key Columns

- order_id
- order_date
- ship_date
- ship_mode
- customer_id
- customer_name
- segment
- country
- region
- state
- city
- category
- sub_category

- product_name
- sales
- quantity
- discount
- profit

---

3. Business Questions

Sales & Profit Performance

1. What are the total sales and total profit of the Superstore?
2. How do sales and profit vary by year?
3. Which regions generate the highest sales and profit?
4. Which states contribute the most to overall sales?
5. What are the monthly sales trends?

Product Analysis

6. Which product categories are most profitable?
7. Which sub-categories are consistently loss-making?
8. What are the top 10 products by total sales?
9. Which products have the highest profit margins?

Customer Analysis

10. Which customer segment generates the highest revenue?
11. Who are the top 10 customers by sales value?
12. What is the average order value per customer?

Discount & Operations

13. How does discount level impact profitability?
14. What is the average profit per order?
15. Which shipping mode is used most frequently and is most profitable?

---

4. SQL Methods and Queries

1. Total Sales and Profit

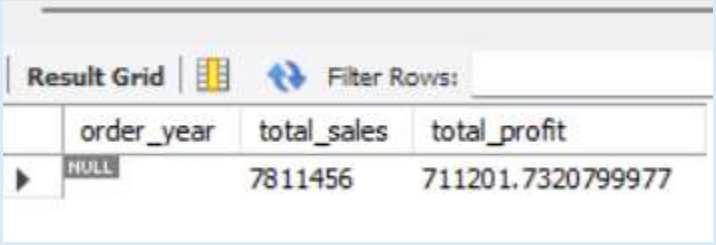
```sql
### 1. Total Sales and Profit
select sum(sales) as total_sales, sum(profit) as total_profit from superstore_sales;
```

Result Grid | Filter Rows:

	total_sales	total_profit
▶	7811456	711201.7320799977

2) Sales and Profit by Year

```
### 2. Sales and Profit by Year
select year(order_date) as order_year, sum(sales) as total_sales, sum(profit) as total_profit from superstore_sales
group by year(order_date) order by order_year;
```

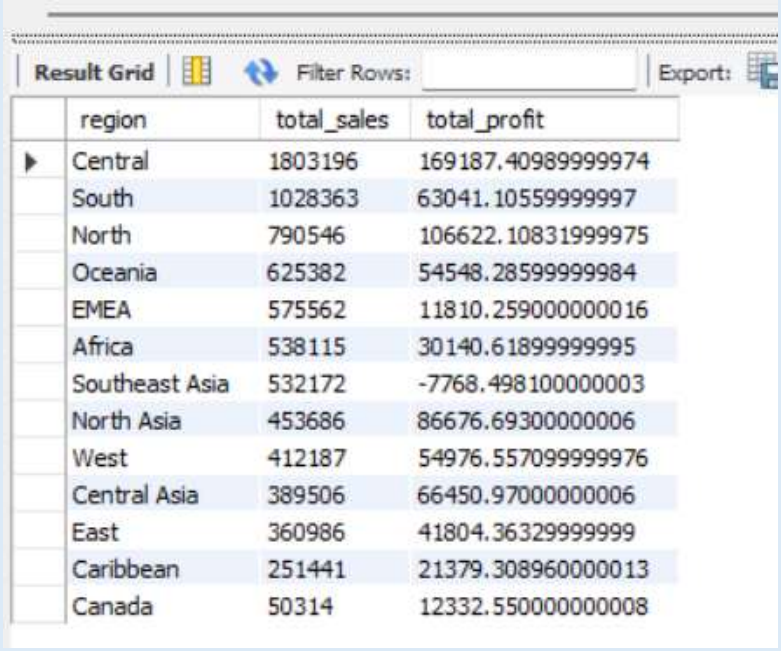


Result Grid | Filter Rows:

order_year	total_sales	total_profit
NULL	7811456	711201.7320799977

3. Sales and Profit by Region

```
### 3. Sales and Profit by Region
select region, sum(sales) as total_sales, sum(profit) as total_profit from superstore_sales group by region
order by total_sales desc;
```



Result Grid | Filter Rows: | Export:

region	total_sales	total_profit
Central	1803196	169187.40989999974
South	1028363	63041.10559999997
North	790546	106622.10831999975
Oceania	625382	54548.28599999984
EMEA	575562	11810.259000000016
Africa	538115	30140.61899999995
Southeast Asia	532172	-7768.498100000003
North Asia	453686	86676.69300000006
West	412187	54976.557099999976
Central Asia	389506	66450.97000000006
East	360986	41804.36329999999
Caribbean	251441	21379.308960000013
Canada	50314	12332.550000000008

4. Top States by Sales

```
### 4. Top States by Sales
select state,sum(sales) as total_sales from superstore_sales group by state order by total_sales desc;
```

Result Grid | Filter Rows:

state	total_sales
England	273710
California	263495
Ile-de-France	188108
New York	151943
New South Wales	145380
North Rhine-Westphalia	135097
Queensland	125019
Texas	118199
San Salvador	105245
National Capital	101647
Victoria	95232
Distrito Federal	78419
Provence-Alpes-Côte d...	74183
São Paulo	70388
Guatemala	69372

Result 49

5.Monthly Sales Trend

```
17    ### 5. Monthly Sales Trend
18 •  select year(order_date) as order_year, month(order_date) as order_month, sum(sales) as total_sales
19    from superstore_sales group by year(order_date), month(order_date) order by order_year, order_month;
20
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

order_year	order_month	total_sales
NULL	NULL	7811456

6. Sales and Profit by Category

```
21    ### 6. Sales and Profit by Category
22 •  select category,sum(sales) as total_sales, sum(profit) as total_profit
23    from superstore_sales group by category;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

category	total_sales	total_profit
Office Supplies	2775793	348117.23490000085
Furniture	2397653	95578.87629999963
Technology	2638010	267505.62087999907

7. Loss-Making Sub-Categories

```
28      ### 7. Loss-Making Sub-Categories
29 •    select  sub_category, sum(profit) as total_profit from superstore_sales group by sub_category
30      having sum(profit) < 0 order by total_profit;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

sub_category	total_profit
Tables	-69921.27199999994

8. Top 10 Products by Sales

```
### 8. Top 10 Products by Sales
select product_name, sum(sales) as total_sales from superstore_sales group by product_name order by total_sales desc
limit 10;
```

Result Grid | Filter Rows: | Export: | Wrap C

product_name	total_sales
Eldon File Cart, Single Width	31319
Rogers File Cart, Single Width	22645
Tenex File Cart, Single Width	20778
Smead File Cart, Single Width	20775
Office Star Executive Leather Armchair, Adjust...	19355
Fellowes Lockers, Industrial	19172
Smead Lockers, Industrial	18648
Hewlett Copy Machine, Color	16849
Rogers Lockers, Blue	16494
Fellowes Lockers, Wire Frame	16470

9. Products with Highest Profit Margin

```
### 9. Products with Highest Profit Margin
select product_name,sum(profit) / sum(sales) as profit_margin from superstore_sales group by product_name
order by profit_margin desc;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Conte

product_name	profit_margin
Xerox 20	0.5184
Avery 475	0.5026415094339622
Tops Green Bar Computer Printout Paper	0.4993877551020408
Xerox 1890	0.4993877551020408
Adams Telephone Message Book w/Frequently-...	0.49874999999999997
Southworth Structures Collection	0.4986301369863014
Xerox 223	0.49111578947368423
Strathmore #10 Envelopes, Ultimate White	0.49040316455696203
Xerox 1918	0.49012645161290325
Personal Creations Ink Jet Cards and Labels	0.4891478260869565
Xerox 1984	0.48849230769230767
Xerox 193	0.4883666666666667
Xerox 1983	0.48836666666666667
Rediform S.O.S. Phone Message Books	0.48804
Xerox 1987	0.4872602150537635

Result 54

10. Sales by Customer Segment

```
43    ### 10. Sales by Customer Segment
44    select segment, sum(sales) as total_sales from superstore_sales
45    group by segment;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

segment	total_sales
Consumer	4046879
Home Office	1403776
Corporate	2360801

11. Top 10 Customers by Sales

```
48    ### 11. Top 10 Customers by Sales
49 •  select customer_name,sum(sales) as total_sales from superstore_sales group by customer_name
50    order by total_sales desc
51    limit 10;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

customer_name	total_sales
Eric Murdock	19489
John Grady	19483
Maria Etezadi	18209
Theone Pippenger	17416
Ben Ferrer	16661
Dan Reichenbach	16425
Randy Bradley	16399
Mathew Reese	16125
Muhammed Yedwab	16073
Steven Ward	15984

12. Average Order Value per Customer

```
4     ### 12. Average Order Value per Customer
5 •   select customer_id, avg(sales) as avg_order_value from superstore_sales
6     group by customer_id;
```

Result Grid | Filter Rows: | Export:

customer_name	total_sales
Eric Murdock	19489
John Grady	19483
Maria Etezadi	18209
Theone Pippenger	17416
Ben Ferrer	16661
Dan Reichenbach	16425
Randy Bradley	16399
Mathew Reese	16125
Muhammed Yedwab	16073
Steven Ward	15984

13. Discount Impact on Profit

```
### 13. Discount Impact on Profit
select discount,sum(profit) as total_profit from superstore_sales group by discount
order by discount;
```

Result Grid | Filter Rows: | Export:

discount	total_profit
0	1068788.1821999983
0.002	39232.35124000004
0.07	9792.8781
0.1	140779.30150000038
0.15	12700.86819999999
0.17	13847.55030000001
0.2	67267.76159999998
0.202	-379.23072000000013
0.25	-137.97000000000037
0.27	-983.0975999999999
0.3	-12514.021799999997
0.32	-1570.2289000000003
0.35	-6078.121500000003
0.37	-4911.983699999999
0.4	-114339.05809999994

Result 58

14. Average Profit per Order

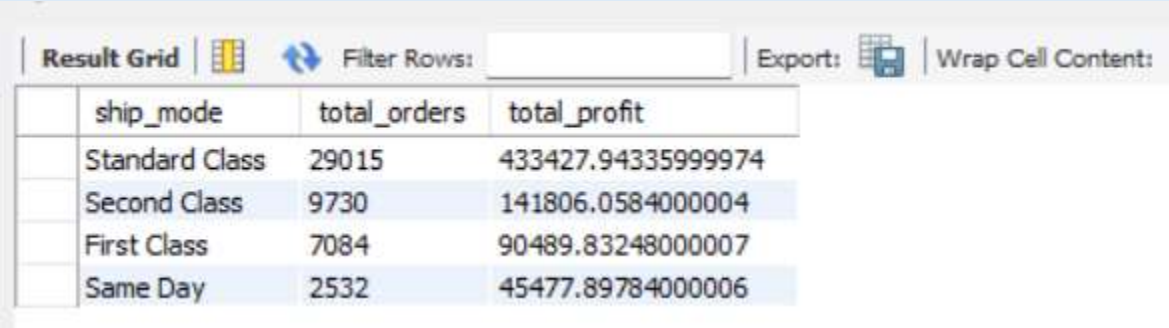
```
64
65      ### 14. Average Profit per Order
66 •    select avg(profit) as avg_profit_per_order from superstore_sales;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

avg_profit_per_order
14.706100619920962

15. Shipping Mode Analysis

```
### 15. Shipping Mode Analysis
select ship_mode, count(order_id) as total_orders, sum(profit) as total_profit from superstore_sales group by ship_mode
order by total_orders desc;
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

ship_mode	total_orders	total_profit
Standard Class	29015	433427.94335999974
Second Class	9730	141806.0584000004
First Class	7084	90489.83248000007
Same Day	2532	45477.89784000006

5. Analysis Summary

- **Overall Performance:** The Superstore achieves strong revenue, but profit margins vary significantly across products and regions.
- **Regional Performance:** A few regions and states dominate sales contribution, while others underperform or generate losses.
- **Product Insights:** Technology and Office Supplies categories tend to be more profitable, whereas certain Furniture sub-categories consistently incur losses.
- **Customer Insights:** Consumer and Corporate segments contribute the majority of revenue. A small group of customers accounts for a large portion of total sales.
- **Discount Strategy:** Higher discounts are strongly associated with reduced or negative profits, highlighting the need for controlled discounting.
- **Operational Efficiency:** Standard shipping is the most frequently used shipping mode, while faster shipping methods do not always result in higher profit.

6. Business Recommendations

1. Reduce or restructure discounts on low-margin products.
2. Discontinue or reprice consistently loss-making sub-categories.
3. Focus sales and marketing efforts on high-profit regions and customer segments.
4. Optimize shipping strategies to balance cost and customer satisfaction.
5. Mon!itor top customers and introduce loyalty programs to retain high-value clients.

7. Conclusion

This SQL-based Superstore analysis demonstrates how structured querying can convert raw transactional data into actionable business insights. The findings support strategic decision-making in pricing, product management, customer targeting, and operational optimization.