

Retail Sale Analysis Using MySQL On Superstore Dataset



Project By : Pratik Uday Ghag

INTRODUCTION

- Retail sales analysis on Superstore data involves a comprehensive examination of transactional records.
- The primary objective is to uncover patterns and insights that can inform strategic business decisions, optimize operations, and enhance.

PROBLEM STATEMENTS

1. Which categories generate the most sales or profit?
2. Which cities or regions are most profitable?
3. Is discount given on product reducing the profit?
4. Find out top 5 states by sales?
5. Which shipping mode gives the best profit to cost ratio?
6. What is the average delivery time for each shipping mode?
7. How many repeat customers are there?
8. Which customers have returned to order multiple times?
9. Which customer segments are the most valuable?
10. Are there any months with negative profit?

1. Which categories generate the most sales or profit?

```
## Query 1: Which categories generate the most sales or profit?  
select category,round(sum(sales)) as totale_sales, round(sum(profit))  
from `sample - superstore`  
group by category;
```

Result :

	category	totale_sales	round(sum(profit))
▶	Furniture	733047	16981
	Office Supplies	703503	120490
	Technology	835900	145387

Insight :

- This query helps identify which broad product categories (like Furniture, Office Supplies, Technology) bring in the most revenue and profit.

2. Which cities or regions are most profitable?

```
## Query 2: Which cities or regions are most profitables?  
select city ,round(sum(sales)) as total_revenue, round(sum(profit)) as total_profit  
from `sample - superstore`  
group by city  
order by total_revenue desc;
```

Result :

The screenshot shows a database query results grid. The grid has four columns: 'city', 'total_revenue', and 'total_profit'. The 'city' column lists five cities: New York City, Los Angeles, Seattle, San Francisco, and Philadelphia. The 'total_revenue' column shows the total sales for each city, and the 'total_profit' column shows the total profit. The grid includes a header row and five data rows. There are also navigation icons and a 'Filter Rows:' input field at the top.

	city	total_revenue	total_profit
▶	New York City	255249	61624
	Los Angeles	173169	29807
	Seattle	117773	28869
	San Francisco	110917	17177
	Philadelphia	107486	-13732

Insight :

- Reveals top revenue-generating cities, aiding in regional performance analysis.
- This analysis gives us an idea about in which cities people are more interested to buy things so we can add more stores in that areas to increase our business.

3. Is discount given on product reducing the profit?

```
## Query 3: Is discount given on product reducing the profit?  
select discount, avg(profit) as avg_profit, count(*) as order_count  
from `sample - superstore`  
group by discount  
order by discount;
```

Result :

	discount	avg_profit	order_count
▶	0	68.10909261327028	4657
	0.1	96.05507446808508	94
	0.15	27.288298076923077	52
	0.2	25.198563349309303	3547
	0.3	-45.679636123348004	227
	0.32	-88.56065555555558	27
	0.4	-112.51435121951225	205

Insight :

- Shows how different discount levels affect profit.
- If higher discounts correlate with negative profits, you may need to revise the discount strategy.

4. Find out top 5 states by sales?

```
## Query 4: Find out top 5 states by sales?  
select state, round(sum(sales)) as total_sales  
from `sample - superstore`  
group by State  
order by total_sales desc  
limit 5;
```

Result :

	state	total_sales
▶	California	450568
	New York	309454
	Texas	169554
	Washington	136590
	Pennsylvania	114911

Insight :

- Highlights top-performing states by sales.
- These can be targeted for expansion or upselling efforts.

5. Which shipping mode gives the best profit to cost ratio?

```
## Query 5: Which shipping mode gives the best profit to cost ratio?  
select `Ship Mode`,  
round(sum(profit)) as total_profit,  
round(sum(sales)) as total_sales,  
round((sum(profit) / sum(sales)) * 100 , 2) as profit_margin_percentage  
from `sample - superstore`  
group by `Ship Mode`  
order by profit_margin_percentage desc;
```

Result :

	Ship Mode	total_profit	total_sales	profit_margin_percentage
▶	First Class	48779	349495	13.96
	Same Day	16025	127353	12.58
	Second Class	56506	453342	12.46
	Standard Class	161548	1342260	12.04

Insight :

- The higher the profit-to-cost ratio, the more efficient that shipping mode is in terms of generating profit from incurred cost.

6. What is the average delivery time for each shipping mode?

```
## Query 6: What is the average delivery time for each shipping mode?  
select `Ship Mode`,  
round(avg(datediff(`Ship Date`, `Order Date`)),2) as Avg_Delivery_Days  
from `sample - superstore`  
where `Ship Date` is not null and `Order Date` is not null  
group by `Ship Mode`  
order by Avg_Delivery_Days desc;
```

Result :

	Ship Mode	Avg_Delivery_Days
▶	Second Class	HULL
	Standard Class	HULL
	First Class	HULL
	Same Day	HULL

Insight :

- This helps you understand shipping efficiency.
- Faster delivery modes may cost more but improve customer satisfaction.

7. How many repeat customers are there?

```
## Query 7: Hoe many repeat customers are there?  
select count(*) as RepeatCustomersCount  
from(  
    select `Customer ID`  
    from `sample - superstore`  
    group by `Customer ID`  
    having count(distinct `Order ID`) > 1  
) as RepeatCustomers;
```

Result :

Result Grid	
	RepeatCustomersCount
▶	781

Insight :

- Repeated customer count shows us interest of our customers buying that product repeatedly so we can increase product availability.

8. Which customers have returned to order multiple times?

```
## Query 8: Which customers have returned to order multiple times?  
select `Customer ID`,  
`Customer Name`,  
count(distinct `Order ID`) as TotalOrders  
from `sample - superstore`  
group by `Customer ID`, `Customer Name`  
having count(distinct `Order ID`) > 1  
order by TotalOrders desc;
```

Result :

	Customer ID	Customer Name	TotalOrders
▶	EP-13915	Emily Phan	17
	CK-12205	Chloris Kastensmidt	13
	EA-14035	Erin Ashbrook	13
	NS-18640	Noel Staavos	13
	PG-18820	Patrick Gardner	13

Insight :

- This can give us an idea about interested customers to buy products repeatedly.
- We can give some offers to that customers so they can order more and continue to shopping with us.

9. Which customer segments are the most valuable?

```
## Query 9: Which customer segments are the most valuable?  
select Segment,  
round(sum(sales),2) as total_sales,  
round(sum(profit),2) as total_profit,  
round((sum(profit) / sum(sales)) * 100 , 2) as Profit_Margin_Percentage  
from `sample - superstore`  
group by Segment  
order by total_profit desc;
```

Result :

	Segment	total_sales	total_profit	Profit_Margin_Percentage
▶	Consumer	1150166.18	132669.78	11.53
	Corporate	696604.51	90366.3	12.97
	Home Office	425679.16	59821.68	14.05

Insight :

- Total Sales shows revenue contribution.
- Total Profit reveals actual value gained.
- Customer Count shows how many unique customers are in each segment.

10. Are there any months with negative profit?

```
## Query 10: Are there any months with negative profit?  
select  
date_format(`Order Date`, '%y-%m') as month,  
sum(profit) as Total_Profit  
from `sample - superstore`  
group by month  
having Total_Profit < 0;
```

Result :

	month	Total_Profit

Insight :

- There are no months with negative profit. It is good for business growth.

CONCLUSION

- This project effectively transforms raw retail data from the Superstore dataset into actionable business insights using MySQL.
- By analyzing product performance, customer behavior, regional trends, discount impact, and shipping efficiency, it helps identify key areas for growth and optimization.
- The insights gained support data-driven decisions that enhance profitability, customer targeting, and operational strategy in a retail environment.