

# Sales Performance Analysis Using SQL

(A Data Analysis Project on Customer, Product, and Profit Insights)

## Business problem

SQL is an essential tool for managing and analyzing structured data efficiently, especially when working with large sales datasets containing multiple tables and relationships. It allows users to store data in an organized format, perform complex queries, and extract meaningful insights quickly. In this project, SQL was used to integrate customer, product, and order data to perform detailed business analysis. Using SQL enables accurate calculations, trend identification, and decision-making support for business operations. The primary objectives of this project are:

- To analyze overall sales performance and identify revenue trends across different time periods and regions.
- To evaluate product performance by identifying top-selling products, profitable categories, and loss-making items.
- To understand customer behavior by identifying valuable customers and high-contributing customer segments.
- To assess shipping efficiency and delivery performance to improve operational processes.
- To examine the relationship between discounts and profit to support better pricing and business strategies.

## Data Preparation and Processing

The raw dataset was provided in CSV format and imported into the SQL database by creating structured tables for customers, products, and orders. Appropriate data types were assigned to each column to ensure accurate storage and calculations. Basic data cleaning was performed by checking for missing values, validating date formats, and maintaining consistency in categorical fields. The data was then normalized into separate tables to reduce redundancy and improve data integrity, enabling efficient analysis throughout the project.

```
-- ==STEP 1:CHECKING NULL VALUE== --
```

```
SELECT * FROM superstore_raw WHERE customer_id IS NULL OR product_id IS NULL;
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:														
row_id	order_id	order_date	ship_date	ship_mode	customer_id	customer_name	segment	country	city	state	postal_code	region	product_id	category	sub_category	product_name	sale

## Table Relationships and Data Integration

The tables were connected using primary and foreign key relationships to maintain data consistency and enable meaningful analysis. The orders table contains foreign keys such as customer id and product id, which reference the primary keys in the customers and products tables. These relationships allow data from multiple tables to be combined using SQL joins for comprehensive analysis. This approach ensures accurate linking between customer information, product details, and transaction records.

## Research question and key insight

### 1.analyze sales and revenue for time period and region

```
-- QUE 1: WHAT IS TOTAL SALES AND TOTAL PROFIT FOR THE COMPANY ?
```

```
SELECT  
SUM(sales) AS total_sales,  
SUM(profit) AS total_profit  
FROM orders;
```

	total_sales	total_profit
▶	2297201.07	286397.79

Insight: The total sales and profit values indicate the overall financial performance and profitability of the business.

```
-- QUE 2:WHICH YEAR HAD THE HIGHEST SALES ?  
SELECT YEAR(order_date) AS year, SUM(sales) AS total_sales  
FROM orders GROUP BY YEAR(order_date) ORDER BY total_sales DESC;
```

	year	total_sales
▶	2017	733215.19
	2016	609205.86
	2014	484247.56
	2015	470532.46

Insight: The year with the highest sales represents the period of peak business growth and strong customer demand.

```
-- QUE 3:WHAT IS MONTHLY SALES TREND ?  
SELECT YEAR(order_date) AS year, MONTH(order_date) AS month,  
SUM(sales) AS total_sales FROM orders GROUP BY YEAR(order_date), MONTH(order_date)  
ORDER BY year, month;
```

	year	month	total_sales
▶	2014	1	14236.90
	2014	2	4519.92
	2014	3	55691.04
	2014	4	28295.35
	2014	5	23648.28
	2014	6	34595.14
	2014	7	33946.37
	2014	8	27909.47
	2014	9	81777.34
	2014	10	31453.37
	2014	11	78628.74
	2014	12	69545.64

Insight: The monthly trend reveals seasonal patterns and fluctuations in revenue throughout the year.

```
-- QUE 4: WHICH REGION GENERATES THE HIGHEST REVENUES ?
SELECT c.region, SUM(o.sales) AS total_sales FROM orders o
JOIN customers c ON o.customer_id = c.customer_id
GROUP BY c.region
ORDER BY total_sales DESC;
```

	region	total_sales
▶	West	2104439.11
	South	138571.73
	East	52676.75
	Central	1513.48

Insight: The region generating the highest revenue shows the company's strongest market presence and customer base.

```
-- QUE 5: WHICH STATE HAS LOWEST PROFIT ?
SELECT c.state, SUM(o.profit) AS total_profit FROM orders o
JOIN customers c ON o.customer_id = c.customer_id
GROUP BY c.state
ORDER BY total_profit ASC;
```

	state	total_profit
▶	Ohio	-3776.50
	North Carolina	-2620.30
	West Virginia	-1599.69
	Connecticut	-785.16
	Georgia	58.52
	New Hampshire	116.64
	Missouri	144.87
	Florida	151.32
	Colorado	224.90
	California	258.13
	Wyoming	280.78
	...	---

Insight: The state with the lowest profit highlights potential operational inefficiencies or pricing challenges affecting profitability.

The analysis shows that the company has strong overall sales performance with certain regions contributing significantly to revenue growth. However, variations in monthly trends and low-profit areas indicate opportunities for improvement in pricing strategies and operational efficiency. By focusing on underperforming regions and optimizing resource planning, the company can further enhance profitability and business performance.

## 2. To evaluate product performance

```
SELECT p.product_name, SUM(o.sales) AS total_sales FROM orders o
JOIN products p ON o.product_id = p.product_id
GROUP BY p.product_name
ORDER BY total_sales DESC LIMIT 10;
```

	product_name	total_sales
▶	Canon imageCLASS 2200 Advanced Copier	61599.83
	Fellowes PB500 Electric Punch Plastic Comb Bind...	27453.38
	Cisco TelePresence System EX90 Videoconferen...	22638.48
	HON 5400 Series Task Chairs for Big and Tall	21870.57
	GBC DocuBind TL300 Electric Binding System	19823.48
	GBC Ibimaster 500 Manual ProClick Binding System	19024.50
	Hewlett Packard LaserJet 3310 Copier	18839.68
	HP Designjet T520 Inkjet Large Format Printer -...	18374.90

Insight: The top-selling products contribute significantly to overall revenue and represent strong customer demand.

```
-- QUE 7: WHICH CATEGORY IS MOST PROFITABLE ?
SELECT p.category, SUM(o.profit) AS total_profit FROM orders o
JOIN products p ON o.product_id = p.product_id
GROUP BY p.category
ORDER BY total_profit DESC;
```

	category	total_profit
▶	Technology	145455.66
	Office Supplies	122490.88
	Furniture	18451.25

Insight: The category with the highest profit indicates areas where the company achieves maximum financial returns.

```
-- QUE 8: WHICH SUB-CATEGORY HAS HIGHEST DISCOUNT ?
SELECT p.sub_category, AVG(o.discount) AS avg_discount FROM orders o
JOIN products p ON o.product_id = p.product_id
GROUP BY p.sub_category
ORDER BY avg_discount DESC;
```

	sub_category	avg_discount
▶	Binders	0.372292
	Machines	0.306087
	Tables	0.261285
	Bookcases	0.211140
	Chairs	0.170178
	-----	-----

Insight: Higher discounts in certain sub-categories may impact profitability and require pricing adjustments.

```
-- QUE 9: WHICH PRODUCT ARE SOLD IN HIGHEST QUANTITY ?
SELECT p.product_name, SUM(o.quantity) AS total_qty FROM orders o
JOIN products p ON o.product_id = p.product_id
GROUP BY p.product_name
ORDER BY total_qty DESC
LIMIT 10;
```

	product_name	total_qty
▶	Staples	215
	Staple envelope	170
	Easy-staple paper	150
	Staples in misc. colors	86
	Logitech P710e Mobile Speakerphone	75
	KI Adjustable-Height Table	74
	Storex Dura Pro Binders	71
	Avery Non-Stick Binders	71
	Xerox 1908	70

Insight: Products with high sales volume indicate strong market demand and inventory importance.

```
-- QUE 10: WHICH CATEGORY HAS THE LOWEST PROFIT MARGIN ?
SELECT p.category, SUM(o.profit)/SUM(o.sales) AS profit_margin FROM orders o
JOIN products p ON o.product_id = p.product_id
GROUP BY p.category
ORDER BY profit_margin ASC;
```

	category	profit_margin
▶	Furniture	0.024867
	Office Supplies	0.170352
	Technology	0.173958

Insight: Products generating losses highlight the need for cost control or pricing strategy improvements.

The product performance analysis shows that certain products and categories contribute significantly to revenue and profitability, indicating strong market demand. However, the presence of high-discount and loss-making products suggests opportunities for pricing optimization and cost management. By focusing on profitable products and improving underperforming ones, the company can enhance overall financial performance.

### 3.Understanding customer behaviour

```
-- QUE 11: TOP 10 CUSTOMERS BY TOTAL PURCHASE VALUE
SELECT c.customer_name, SUM(o.sales) AS total_spent FROM orders o
JOIN customers c ON o.customer_id = c.customer_id
GROUP BY c.customer_name
ORDER BY total_spent DESC LIMIT 10;
```

	customer_name	total_spent
▶	Sean Miller	25043.07
	Tamara Chand	19052.22
	Raymond Buch	15117.35
	Tom Ashbrook	14595.62
	Adrian Barton	14473.57
	Ken Lonsdale	14175.23
	Sanjit Chand	14142.34
	Hunter Lopez	12873.30
	Sanjit Engle	12209.44

Insight: High-spending customers contribute significantly to total revenue and represent valuable business relationships.

```
-- QUE 12: WHICH CUSTOMERS SEGMENT GENERATES THE MOST SALES?  
SELECT c.segment, SUM(o.sales) AS total_sales FROM orders o  
JOIN customers c ON o.customer_id = c.customer_id  
GROUP BY c.segment  
ORDER BY total_sales DESC;
```

	segment	total_sales
▶	Consumer	1161401.34
	Corporate	706146.44
	Home Office	429653.29

Insight: The segment generating the highest sales indicates the primary target market for the company.

```
-- QUE 13: HOW MANY UNIQUE CUSTOMERS ARE THERE IN EACH REGION ?  
SELECT region, COUNT(DISTINCT customer_id) AS customer_count FROM customers GROUP BY region;
```

	region	customer_count
▶	Central	3
	East	36
	South	68
	West	686

Insight: Regions with more customers show stronger market penetration and growth opportunities.

```
-- QUE 14:WHICH CUSTOMERS ARE MAKING LOSSES FOR THE COMPANY ?
SELECT c.customer_name, SUM(o.profit) AS total_profit FROM orders o
JOIN customers c ON o.customer_id = c.customer_id
GROUP BY c.customer_name
HAVING total_profit < 0
ORDER BY total_profit;
```

	customer_name	total_profit
▶	Cindy Stewart	-6626.37
	Grant Thornton	-4108.66
	Luke Foster	-3583.97
	Sharelle Roach	-3333.91
	Henry Goldwyn	-2797.95
	Nathan Cano	-2204.81
	Sean Braxton	-2082.75
	Sean Miller	-1980.75
	Christine Phan	-1850.31

Insight: Customers generating losses may require pricing adjustments or cost control measures.

The customer behavior analysis reveals that a small group of customers and segments contribute significantly to overall revenue, highlighting key target markets. Regional distribution patterns show opportunities for expansion in underperforming areas. By focusing on customer retention and personalized strategies, the company can strengthen relationships and increase long-term profitability.

#### 4. To assess shipping efficiency and delivery performance

```
-- QUE 15:WHAT IS THE AVERAGE DELIVERY TIME?
SELECT AVG(DATEDIFF(ship_date, order_date)) AS avg_delivery_days FROM orders;
```

	avg_delivery_days
▶	3.9582

Insight: The company delivers orders in an average of about 4 days, indicating generally efficient shipping performance.

```
-- QUE 16:WHICH SHIPPING MODE IS MOST USED?
```

```
SELECT ship_mode, COUNT(*) AS total_orders FROM orders  
GROUP BY ship_mode  
ORDER BY total_orders DESC;
```

	ship_mode	total_orders
▶	Standard Class	5968
	Second Class	1945
	First Class	1538
	Same Day	543

Insight: Standard Class is the most preferred shipping method by customers, showing it is the company's primary delivery channel.

```
-- QUE 17:WHICH ORDERS TOOK THE LONGEST TIME TO DELIVER ?
```

```
SELECT order_id,DATEDIFF(ship_date, order_date) AS delivery_days FROM orders  
ORDER BY delivery_days DESC LIMIT 10;
```

	order_id	delivery_days
▶	CA-2016-128867	7
	CA-2014-105893	7
	US-2015-108966	7
	US-2015-108966	7
	US-2017-107272	7
	CA-2017-155558	7
	CA-2016-128867	7
	US-2017-107272	7
	CA-2017-107720	7

Insight: Some orders take up to 7 days to deliver, showing occasional delays compared to the average delivery time.

## 5. examine the relationship between discounts and profit

```
SELECT discount, AVG(profit) AS avg_profit FROM orders  
GROUP BY discount  
ORDER BY discount;
```

	discount	avg_profit
▶	0.00	66.900350
	0.10	96.055426
	0.15	27.288077
	0.20	24.702806
	0.30	-45.679912
	0.32	-88.561481
	0.40	-111.927573
	0.45	-226.647273
	0.50	-310.704697

Insight: Profit decreases as discount increases, and high discounts (30% and above) lead to significant losses.

```
-- QUE 19: WHICH REGION GIVES THE HIGHEST AVERAGE PROFIT PR ORDER?  
SELECT c.region, AVG(o.profit) AS avg_profit FROM orders o  
JOIN customers c ON o.customer_id = c.customer_id  
GROUP BY c.region  
ORDER BY avg_profit DESC;
```

	region	avg_profit
▶	West	29.510292
	East	26.414848
	South	16.961983
	Central	-14.400000

Insight: The West region generates the highest average profit per order, while the Central region is running at a loss.

```
-- QUE 20: WHICH PRODUCT ARE CAUSING OVERALL LOSS?
```

```
SELECT p.product_name, SUM(o.profit) AS total_profit FROM orders o
JOIN products p ON o.product_id = p.product_id
GROUP BY p.product_name
HAVING total_profit < 0
ORDER BY total_profit;
```

	product_name	total_profit
▶	3.6 Cubic Foot Counter Height Office Refrigerator	-872.07
	36X48 HARDFLOOR CHAIMAT	-40.49
	3D Systems Cube Printer, 2nd Generation, White	-468.00
	3M Replacement Filter for Office Air Cleaner for...	-20.47
	Acco 6 Outlet Guardian Basic Surge Suppressor	-3.66
	Acco 6 Outlet Guardian Premium Plus Surge Sup...	-113.78
	Acco D-Ring Binder w/DublLock	-102.84
	Acco Economy Flexible Poly Round Ring Binder	-5.74
	Acco Pressboard Covers with Storage Hooks, 1...	-13.90

Insight: Some specific products consistently generate negative total profit, meaning they are being sold at an overall loss.

The company performs well in delivery and some regions, but profitability is heavily impacted by high discounting, loss-making products, and poor performance in the Central region, so reducing excessive discounts and reviewing unprofitable products and regions is necessary to improve overall profitability.

## ★ Final Business Recommendations

### 1. Control Discount Strategy

Since higher discounts are causing profit loss, the company should limit excessive discounts and apply them only to slow-moving or seasonal products.

### 2. Improve Loss-Making Products

Products generating negative profit should be reviewed for pricing, supplier cost reduction, or possible discontinuation.

### 3. Focus on Low-Performing Region (Central)

The Central region needs strategic improvement through better pricing, targeted marketing, and cost optimization to turn losses into profit.

### 4. Optimize Shipping Performance

Although delivery time is generally good, reducing delays in Standard Class shipping can improve customer satisfaction and retention.

### 5. Strengthen High-Performing Regions

Regions like the West that already generate strong profit should receive more investment and promotional focus to maximize revenue growth.

