

Project: Sales store from 2014 : 2017



I performed data transformation using Power Query to organize and analyze the data efficiently. Here are the steps I followed:

- **Converted the "Discount" column into a percentage to**
- **Added a new column named "Discount Value" to calculate the discount amount.**
- **To do this, I navigated to the "Add Column" tab, selected "Custom Column", and entered the formula.**
- **Computed the "Discount Value" using the formula**
Discount Value=Sales*- Discount
- **Repositioned the "Discount Value" column to be placed before the "Profit" column for better readability.**
- **Added a new column named "COGS" to calculate the cost of goods sold**
- **This was done by going to "Add Column", selecting "Custom Column", and entering the formula**
- **Computed the "COGS" value using the formula:**
COGS=-(Sales+Discount Value-Profit)
- **Repositioned the "COGS" column to be placed after the "Discount Value" column to improve data clarity.**

The Data Transformation and Data Cleaning process was successfully completed using Power Query, and now we are moving on to the second part: solving the questions using SQL

1- "Write an SQL query to create a view named **Orders_View** that combines order details from the **Orders** table with region information from the **Regions** table and return status from the **Returns** .

```
CREATE VIEW Orders_View AS
SELECT O.[Row ID], O.[Order ID], O.[Order Date], O.[Ship Date],
O.[Ship Mode], O.[Customer ID], O.[Customer Name], O.[Segment],
O.[Country], O.[City], O.[State], O.[Postal Code], O.[Region],
O.[Product ID], O.[Category], O.[Sub-Category], O.[Product Name],
RE.[Person] AS [Region Manager], R.[Returned] AS [Return Status],
O.[Sales], O.[Quantity], O.[Discount], O.[Discount Value], O.[COGS],
O.[Profit], DATEDIFF(DAY, O.[Order Date], O.[Ship Date])
AS [Delivery Date]
FROM [Orders] O
LEFT JOIN [Returns] R ON O.[Order ID] = R.[Order ID]
LEFT JOIN [Regions] RE ON O.[Region] = RE.[Region];
```

☒ ~~Which categories generate the highest sales? Which are the most profitable?~~

1. Categories Generating the Highest Sales

```
SELECT [Category], SUM([Sales]) AS Total_Sales
FROM Orders_View
GROUP BY [Category]
ORDER BY Total_Sales DESC;
```

| Category | Total_Sales |
|-----------------|-------------|
| Technology | 836154.10 |
| Furniture | 741999.98 |
| Office Supplies | 719046.99 |

2. Most Profitable Categories

```
SELECT [Category], SUM([Profit]) AS Total_Profit
FROM Orders_View
GROUP BY [Category]
ORDER BY Total_Profit DESC;
```

| Categroy | Total_Profit |
|-----------------|--------------|
| Technology | 145455.66 |
| Office Supplies | 122490.88 |
| Furniture | 18451.25 |

Additional Insights: Profit Margin by Category

If you want to analyze the profitability ratio (profit margin) per category

```
SELECT [Category], SUM([Profit]) AS Total_Profit, SUM([Sales]) AS
Total_Sales, (SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS
Profit_Margin
FROM Orders_View
GROUP BY [Category]
ORDER BY Profit_Margin DESC;
```

| Categroy | Total_Profit | Total_Sales | Profit_Margin |
|-----------------|--------------|-------------|---------------|
| Technology | 145455.66 | 836154.10 | 17.395700 |
| Office Supplies | 122490.88 | 719046.99 | 17.035100 |
| Furniture | 18451.25 | 741999.98 | 2.486600 |

- ☒ Which sub-categories have the highest profit margins?

```
SELECT [Sub-Category], SUM([Profit]) AS Total_Profit,
SUM([Sales]) AS Total_Sales, (SUM([Profit]) /
NULLIF(SUM([Sales]), 0)) * 100 AS Profit_Margin
FROM Orders_View
GROUP BY [Sub-Category]
ORDER BY Profit_Margin DESC;
```

- ☒ Which customer segments generate the highest average order value?

```
SELECT [Segment], COUNT( [Order ID]) AS Total_Orders,
SUM([Sales]) AS Total_Sales,
(SUM([Sales]) / NULLIF(COUNT( [Order ID]), 0)) AS
Avg_Order_Value
FROM Orders_View
GROUP BY [Segment]
ORDER BY Avg_Order_Value DESC;
```

| Segment | Total_Orders | Total_Sales | Avg_Order_Value |
|-------------|--------------|-------------|-----------------|
| Home Office | 1783 | 429653.29 | 240.972120 |
| Corporate | 3020 | 706146.44 | 233.823324 |
| Consumer | 5191 | 1161401.34 | 223.733642 |

Financial Performance (Sales, Profit, COGS, Discount)

- ☒ What is the relationship between discounts and profit?
Do higher discounts negatively impact profitability?

1. Correlation Between Discount and Profit

```
SELECT ROUND([Discount], 2) AS Discount_Level,  
COUNT(*) AS Order_Count,  
SUM([Sales]) AS Total_Sales,  
SUM([Profit]) AS Total_Profit,  
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS  
Profit_Margin  
FROM Orders_View  
GROUP BY ROUND([Discount], 2)  
ORDER BY Discount_Level;
```

| Discount_Level | Order_Count | Total_Sales | Total_Profit | Profit_Margin |
|----------------|-------------|-------------|--------------|---------------|
| 0.00 | 4798 | 1087908.47 | 320987.88 | 29.505000 |
| 0.10 | 94 | 54369.30 | 9029.21 | 16.607100 |
| 0.15 | 52 | 27558.59 | 1418.98 | 5.148900 |
| 0.20 | 3657 | 764594.28 | 90338.16 | 11.815100 |
| 0.30 | 227 | 103226.76 | -10369.34 | -10369.34 |
| 0.32 | 27 | 14493.45 | -2391.16 | -16.498200 |
| 0.40 | 206 | 116417.83 | -23057.08 | -19.805400 |
| 0.45 | 11 | 5484.98 | -2493.12 | -45.453500 |
| 0.50 | 66 | 58918.65 | -20506.51 | -34.804700 |
| 0.60 | 138 | 6644.68 | -5944.64 | -89.464600 |
| 0.70 | 418 | 40620.40 | -40075.46 | -98.658400 |
| 0.80 | 300 | 16963.68 | -30539.13 | -180.026500 |

Insights from the Discount vs. Profitability Analysis

1. No Discount (0.00) Has the Highest Profit Margin (29.51%)

a. 4798 orders generated \$1,087,908.47 in sales and \$320,987.88 in profit.

b. This confirms that non-discounted sales are the most profitable.

2. Moderate Discounts (10-20%) Still Generate Profits, But with Lower Margins

a. 10% discount: 16.61% profit margin (lower but still positive).

b. 15% discount: Profit margin drops to 5.15%, indicating declining profitability.

c. 20% discount: 11.82% profit margin, still profitable but significantly lower than no discount.

3. Higher Discounts (30% and Above) Lead to Losses

30% discount: -10.05% profit margin (turns negative).

40% discount: -19.81% profit margin, indicating significant losses.

50% discount: -34.80% profit margin, showing unsustainable discounting.

70-80% discounts: Massive losses (-98.66% to -180.03%), meaning these discounts are completely unprofitable and should be avoided.

Key Takeaways & Recommendations

✓ **Keep discounts at or below 10% to**

✓ ****Avoid high discounts (30% and above), as they lead to significant losses.**

✓ **Re-evaluate pricing strategy**

✓ ****Test small targeted discounts (5-10%)**

2. Profitability Comparison: Discounted vs. Non-Discounted Orders?

```
SELECT
CASE WHEN [Discount] > 0 THEN 'Discounted Orders' ELSE
'Non-Discounted Orders' END AS Discount_Type,
COUNT(*) AS Order_Count,
SUM([Sales]) AS Total_Sales,
SUM([Profit]) AS Total_Profit,
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS
Profit_Margin
FROM Orders_View
GROUP BY CASE WHEN [Discount] > 0 THEN 'Discounted Orders'
ELSE 'Non-Discounted Orders' END;
```

| Discount_Type | Order_Count | Total_Sales | Total_Profit | Profit_Margin |
|-----------------------|-------------|-------------|--------------|---------------|
| Non-Discounted Orders | 4798 | 1087908.47 | 320987.88 | 29.505000 |
| Discounted Orders | 5196 | 1209292.60 | -34590.09 | -2.860300 |

Key Insights: Impact of Discounts on Profitability

1. Non-Discounted Orders Are Highly Profitable (29.51% Profit Margin)

- 4,798 orders w/\$1,087,908.47 in \$320,987.88 in
- The profit margin is 29.51%, non-discounted sales contribute significantly to profitability.

2. Discounted Orders Are Operating at a Loss (-2.86% Profit Margin)

- 5,196 orders with higher total sales (\$1,209,292.60) but overall loss of -\$34,590.09.
- The negative profit margin (-2.86%) suggests

Discounts Do Not Necessarily Drive Profitable Growth

- Even though discounted orders contributed higher total sales,
- This suggests that the business may be over-discounting, lead

Conclusion & Recommendations:

- ✓ Minimize or eliminate discounts where possible to ma
 - ✓ *Reevaluate discount strategies, esp
 - ✓ **CConsider targeted, low-percentage discounts (1-10%)

☒ ~~Which products or categories have the highest profit margin after deducting COGS?~~

1. Profit Margin by Category

```
SELECT [Category],SUM([Sales]) AS Total_Sales,  
SUM([COGS]) AS Total_COGS,SUM([Profit]) AS Total_Profit,  
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS  
Profit_Margin  
FROM Orders_View  
GROUP BY [Category]  
ORDER BY Profit_Margin DESC
```

2. Profit Margin by Sub-Category

```
SELECT [Sub-Category],SUM([Sales]) AS Total_Sales,  
SUM([COGS]) AS Total_COGS,SUM([Profit]) AS Total_Profit,  
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS  
Profit_Margin  
FROM Orders_View  
GROUP BY [Sub-Category]  
ORDER BY Profit_Margin DESC;
```

3. Profit Margin by Product

```
SELECT TOP 10 [Product Name], [Category],  
SUM([Sales]) AS Total_Sales, SUM([COGS]) AS Total_COGS,  
SUM([Profit]) AS Total_Profit,  
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS  
Profit_Margin  
FROM Orders_View  
GROUP BY [Product Name], [Category]  
ORDER BY Profit_Margin DESC;
```

- ☒ ~~Is there a correlation between high discounts and lower profit margins?~~

1: Calculate Correlation Between Discount & Profit Margin

```
WITH DiscountProfit AS (  
    SELECT Discount,   
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS Profit_Margin  
    FROM Orders_View  
    GROUP BY [Discount]),  
Averages AS (  
    SELECT  
    AVG(Discount) AS Avg_Discount,  
    AVG(Profit_Margin) AS Avg_Profit_Margin  
    FROM DiscountProfit)  
SELECT  
    SUM((dp.Discount - a.Avg_Discount) * (dp.Profit_Margin -  
    a.Avg_Profit_Margin)) /  
    (SQRT(SUM(POWER(dp.Discount - a.Avg_Discount, 2))) *  
    SQRT(SUM(POWER(dp.Profit_Margin - a.Avg_Profit_Margin, 2))))  
    AS Correlation_Coefficient  
    FROM DiscountProfit dp  
    CROSS JOIN Averages a;
```

Correlation_Coefficient = -0.941191971270226

- ☒ ~~What is the average Cost of Goods Sold (COGS) per category or sub-category?~~

To calculate the average Cost of Goods Sold (COGS) per category or sub-category, you can use

```
SELECT [Category],  
AVG([COGS]) AS Avg_COGS  
FROM Orders_View  
GROUP BY [Category]  
ORDER BY Avg_COGS DESC;
```

For sub-categories:

```
SELECT [Sub-Category],  
AVG([COGS]) AS Avg_COGS  
FROM Orders_View  
GROUP BY [Sub-Category]  
ORDER BY Avg_COGS DESC;
```

Quantity & Demand Trends

- ☒ ~~Which products or categories have the highest sales volume? Do they align with profitability?~~

1: Identify the Top-Selling Products and Categories by Sales Volume

```
SELECT [Category], SUM([Quantity]) AS Total_Quantity_Sold,  
SUM([Sales]) AS Total_Sales, SUM([Profit]) AS Total_Profit,  
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS Profit_Margin  
FROM Orders_View  
GROUP BY [Category]  
ORDER BY Total_Quantity_Sold DESC;
```

- **✓ Office Supplies** has the **highest sales volume** and a **good profit margin (17.04%)**, making it a strong category.
- **✓ Technology** has the **highest total sales and profit**, with a strong **profit margin (17.40%)**—suggesting it's the most profitable category.
- **✓ Furniture** has **high sales but the lowest profit margin (2.49%)**, indicating **high costs or heavy discounts**.

Query for Top-Selling Products by Volume:

```
SELECT TOP 10 [Product Name],  
SUM([Quantity]) AS Total_Quantity_Sold,  
SUM([Sales]) AS Total_Sales,  
SUM([Profit]) AS Total_Profit,  
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS Profit_Margin  
FROM Orders_View  
GROUP BY [Product Name]  
ORDER BY Total_Quantity_Sold DESC;
```

- ☒ ~~What is the average number of days between the order date and the ship date?~~

```
SELECT AVG(DATEDIFF(DAY, [Order Date], [Ship Date]))  
AS Avg_Processing_Time  
FROM Orders_View;
```

Avg_Processing_Time = 3

- ☒ ~~1 Which shipping methods (Ship Mode) are the most commonly used, and which provide the fastest delivery?~~

```
SELECT  
[Ship Mode],  
COUNT(*) AS Total_Orders,  
(COUNT(*) * 100.0) / SUM(COUNT(*)) OVER () AS  
Percentage_Of_Total  
FROM Orders_View  
GROUP BY [Ship Mode]  
ORDER BY Total_Orders DESC;
```

| Ship Mode | Total_Orders | Percentage_Of_Total |
|----------------|--------------|---------------------|
| Standard Class | 5968 | 59.715829497698 |
| Second Class | 1945 | 19.461677006203 |
| First Class | 1538 | 15.389233540124 |
| Same Day | 543 | 5.433259955973 |

☒ Does shipping mode influence sales or profitability?

```
SELECT [Ship Mode],COUNT(*) AS Total_Orders,  
SUM([Sales]) AS Total_Sales,SUM([Profit]) AS Total_Profit,  
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS  
Profit_Margin  
FROM Orders_View  
GROUP BY [Ship Mode]  
ORDER BY Total_Sales DESC;
```

| Ship Mode | Total_Orders | Total_Sales | Total_Profit | Profit_Margin |
|----------------|--------------|-------------|--------------|---------------|
| Standard Class | 5968 | 1358216.08 | 164089.45 | 12.081200 |
| Second Class | 1945 | 459193.44 | 57446.49 | 12.510300 |
| First Class | 1538 | 351428.43 | 48969.95 | 13.934500 |
| Same Day | 543 | 128363.12 | 15891.90 | 12.380400 |

Analysis of Shipping Mode Influence on Sales & Profitability

Key Insights:

- Standard Class dominates in total orders and sales, likely because it is the most cost-effective and widely used. However, its profit margin is the lowest.
- First Class has the highest profit margin, suggesting that premium shipping services generate better profitability despite fewer customers choosing them.
- Same Day shipping has the least impact on sales but maintains a reasonable profit margin, indicating that it might be a niche yet profitable segment.

Final Conclusion:

Shipping mode does influence both sales and profitability:

- More affordable shipping (Standard Class) drives higher sales volume but at lower profitability.
- Premium shipping options (First Class) generate better profit margins, but fewer customers choose them.
- Same Day shipping has limited usage but offers better profitability than Standard Class.

Recommendation:

- If the business aims for higher profit margins, promoting First Class shipping may be beneficial.
- If the goal is higher sales volume, keeping Standard Class competitive is crucial.
- Consider targeted promotions for Same Day shipping to increase adoption while maintaining profitability.

Geographical Insights (City & State)

- ☒ ~~Which cities or states generate the highest sales?~~
~~Which are the most profitable?~~

```
SELECT [State],[City],SUM([Sales]) AS Total_Sales,
SUM([Profit]) AS Total_Profit,
(SUM([Profit]) / NULLIF(SUM([Sales]), 0)) * 100 AS
Profit_Margin
FROM Orders_View
GROUP BY [State], [City]
ORDER BY Total_Sales DESC, Total_Profit DESC;
```


☑ Which year achieved the highest total sales and profit, and which month achieved the highest total sales and profit?

-- Year with the highest total sales and profit

```
SELECT YEAR([Order Date]) AS Sales_Year,
WITH YearlySales AS (
SELECT
YEAR([Order Date]) AS Sales_Year,
SUM(Sales) AS Total_Sales,
SUM(Profit) AS Total_Profit
FROM Orders_View
GROUP BY YEAR([Order Date]))
SELECT Sales_Year, Total_Sales, Total_Profit,
LAG(Total_Sales) OVER (ORDER BY Sales_Year) AS
Previous_Year_Sales,
CASE
WHEN LAG(Total_Sales) OVER (ORDER BY Sales_Year) IS
NULL THEN NULL
ELSE ROUND(((Total_Sales - LAG(Total_Sales) OVER
(ORDER BY Sales_Year)) / LAG(Total_Sales) OVER (ORDER
BY Sales_Year)) * 100, 2)
END AS Sales_Growth_Percentage
FROM YearlySales
ORDER BY Total_Sales DESC;
```

| Sales_Year | Total_Sales | Total_Profit | Previous_Year_Sales | Sales_Growth_Percentage |
|------------|-------------|--------------|---------------------|-------------------------|
| 2017 | 733215.19 | 93439.77 | 609205.86 | 20.360000 |
| 2016 | 609205.86 | 81795.27 | 470532.46 | 29.470000 |
| 2014 | 484247.56 | 49544.06 | NULL | NULL |
| 2015 | 470532.46 | 61618.69 | 484247.56 | -2.830000 |

-- Month with the highest total sales and profit

```
SELECT FORMAT([Order Date], 'yyyy-MM') AS Sales_Month,  
SUM(Sales) AS Total_Sales,  
SUM(Profit) AS Total_Profit  
FROM Orders_View  
GROUP BY FORMAT([Order Date], 'yyyy-MM')  
ORDER BY Total_Sales DESC, Total_Profit DESC;
```