

# SQL Business Insights Report

## Superstore Sales Analysis

**Prepared by:** Vikas Navik  
**Course:** Data Analytics Internship  
**Tools Used:** MySQL, Jupyter Notebook



### 1. Introduction

In this project, I performed SQL analysis on the Superstore dataset to answer key business questions. The objective was to extract insights using aggregation, filtering, grouping, and date functions.



### 2. Business Questions & SQL Analysis

- ◆ **Question 1: What is the total revenue generated?**

**SQL Query:**

```
SELECT SUM(Sales) AS Total_Revenue FROM superstore;
```

**Output Screenshot:**

Total_Revenue	
2272449.8562999545	

**Insight:**

The company generated approximately \$2.27M in total revenue during the analyzed period, indicating a strong overall sales performance.

- ◆ **Question 2: What are the top 5 cities by total sales?**

**SQL Query:**

```
SELECT City, SUM(Sales) AS Total_Sales  
FROM superstore  
GROUP BY City  
ORDER BY Total_Sales DESC  
LIMIT 5;
```

**Output Screenshot:**

The screenshot shows a database query results grid. At the top, there are buttons for 'Result Grid' (selected), 'Filter Rows', 'Export' (with icons for CSV, Excel, PDF, and others), 'Wrap Cell Content', and 'Fetch row'. The result grid itself has two columns: 'City' and 'Total\_Sales'. The data rows are: New York City (255248.969), Los Angeles (173168.86500000002), Seattle (117772.58399999999), San Francisco (110917.03599999993), and Philadelphia (107486.47000000006). The 'New York City' row is currently selected.

City	Total_Sales
New York City	255248.969
Los Angeles	173168.86500000002
Seattle	117772.58399999999
San Francisco	110917.03599999993
Philadelphia	107486.47000000006

**Insight:**

The top cities contribute significantly to total revenue, indicating strong regional demand.

- ◆ **Question 3: Which customers generated the highest revenue?**

**Query :**

```
SELECT `Customer Name`, SUM(Sales) AS Total_Spent FROM superstore GROUP BY `Customer Name` ORDER BY Total_Spent DESC LIMIT 5;
```

**Output :**

Result Grid | Filter Rows:  Export:

	Customer Name	Total_Spent
▶	Sean Miller	25043.05
	Tamara Chand	19017.847999999998
	Raymond Buch	15117.339
	Tom Ashbrook	14595.62
	Adrian Barton	14355.610999999997

### Insight:

A small group of customers contributes significantly to total revenue. This indicates customer concentration, and the company should focus on retaining these high-value customers through loyalty programs and personalized offers.

### ◆ Question 4: Which transactions resulted in loss?

#### Query :

```
SELECT * FROM superstore WHERE Profit < 0;
```

#### Output :

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

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State	Postal Code	Region	Product ID
4	US-2015-108966	2015-10-11 00:00:00	2015-10-18 00:00:00	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	FUR-TA-10000577
15	US-2015-118983	2015-11-22 00:00:00	2015-11-26 00:00:00	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-AP-10002311
16	US-2015-118983	2015-11-22 00:00:00	2015-11-26 00:00:00	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-BI-10000756
24	US-2017-156909	2017-07-16 00:00:00	2017-07-18 00:00:00	Second Class	SF-20065	Sandra Flanagan	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-CH-10002774
28	US-2015-150630	2015-09-17 00:00:00	2015-09-21 00:00:00	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-BO-10004834
29	US-2015-150630	2015-09-17 00:00:00	2015-09-21 00:00:00	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10000474
37	CA-2016-117590	2016-12-08 00:00:00	2016-12-10 00:00:00	First Class	GH-14485	Gene Hall	Corporate	United States	Richardson	Texas	75080	Central	FUR-FU-10003664
39	CA-2015-117415	2015-12-27 00:00:00	2015-12-31 00:00:00	Standard Class	SN-20710	Steve Nguyen	Home Office	United States	Houston	Texas	77041	Central	FUR-BO-10002545
40	CA-2015-117415	2015-12-27 00:00:00	2015-12-31 00:00:00	Standard Class	SN-20710	Steve Nguyen	Home Office	United States	Houston	Texas	77041	Central	FUR-CH-10004218
67	US-2015-164172	2015-04-30 00:00:00	2015-05-05 00:00:00	Standard Class	PS-18970	Paul Stevenson	Home Office	United States	Chicago	Illinois	60610	Central	FUR-CH-10001146
73	US-2015-134026	2015-04-26 00:00:00	2015-05-02 00:00:00	Standard Class	JE-15745	Joel Eaton	Consumer	United States	Memphis	Tennessee	38109	South	FUR-CH-10000513
75	US-2015-134026	2015-04-26 00:00:00	2015-05-02 00:00:00	Standard Class	JE-15745	Joel Eaton	Consumer	United States	Memphis	Tennessee	38109	South	OFF-ST-10004123
76	US-2017-118038	2017-12-09 00:00:00	2017-12-11 00:00:00	First Class	KB-16600	Ken Brennan	Corporate	United States	Houston	Texas	77041	Central	OFF-BI-10004182
79	US-2014-147606	2014-11-26 00:00:00	2014-12-01 00:00:00	Second Class	JE-15745	Joel Eaton	Consumer	United States	Houston	Texas	77070	Central	FUR-FU-10003194

### Insight:

Several transactions generated negative profit, which may be due to high discounts or pricing issues. These transactions need review to prevent repeated losses.

### ◆ Question 5: Which products are causing overall negative profit?

#### Query :

```
SELECT `Product Name`, SUM(Profit) AS Total_Profit FROM superstore GROUP BY `Product Name` HAVING Total_Profit < 0 ORDER BY Total_Profit;
```

### Output :

The screenshot shows a database query results grid. The grid has two columns: 'Product Name' and 'Total\_Profit'. The 'Product Name' column lists various products, and the 'Total\_Profit' column shows their respective profit values. The results are sorted by profit in descending order. The grid includes standard SQL navigation buttons like 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Contents'. Below the grid, there's an 'Output' section with a dropdown menu set to 'Action Output'.

Product Name	Total_Profit
Cubify CubeX 3D Printer Double Head Print	-8879.9704
Lexmark MX611dhe Monochrome Laser Printer	-4589.973
Cubify CubeX 3D Printer Triple Head Print	-3839.9904
Chromcraft Bull-Nose Wood Oval Conference T...	-2876.1156
Bush Advantage Collection Racetrack Conferen...	-1934.3975999999998
GBC DocuBind P400 Electric Binding System	-1878.1662000000003
Cisco TelePresence System EX90 Videoconferen...	-1811.0784
Martin Yale Chadless Opener Electric Letter Ope...	-1299.1836
Balt Solid Wood Round Tables	-1201.0581
BoxOffice By Design Rectangular and Half-Moo...	-1148.4375
Riverside Furniture Oval Coffee Table, Oval En...	-1147.4
Epson TM-T88V Direct Thermal Printer - Monoch...	-1057.23
Hon 2090 "Pillow Soft" Series Mid Back Swivel/Til...	-989.0495999999999
O'Sullivan 4-Shelf Bookcase in Odessa Pine	-975.0988
Bretford "Just In Time" Height-Adjustable Multi-...	-964.194
Zebra GK420t Direct Thermal/Thermal Transfer ...	-938.28
3.6 Cubic Foot Counter Height Office Refrigerator	-872.0752
Bevis Oval Conference Table, Walnut	-856.0144
Tennsco Single-Tier Lockers	-825.7479999999999
BPI Conference Tables	-795.9725000000001
StarTech.com 10/100 VDSL2 Ethernet Extender...	-785.7620000000001
O'Sullivan Plantations 2-Door Library in Landver...	-767.7436
O'Sullivan Living Dimensions 5-Shelf Bookcases	-755.7516

### Insight:

Several transactions generated negative profit, which may be due to high discounts or pricing issues. These transactions need review to prevent repeated losses.

### ◆ Question 6: Category-wise Sales and Profit Analysis

#### Query :

```
SELECT Category, SUM(Sales) AS Total_Sales, SUM(Profit) AS Total_Profit FROM superstore GROUP BY Category;
```

#### Output :

The screenshot shows a database query results grid with the following data:

Category	Total_Sales	Total_Profit
Furniture	733046.8612999996	16980.7711999998
Office Supplies	703502.9280000031	120489.8864000001
Technology	835900.0669999964	145387.0965999999

### **Insight:**

While some categories generate high sales, their profit margins may differ. This highlights the importance of analyzing both revenue and profitability together rather than focusing only on sales.

### **◆ Question 7: Region-wise Profit Contribution**

#### **Query :**

```
SELECT Region, SUM(Profit) AS Total_Profit FROM superstore GROUP BY Region ORDER BY Total_Profit DESC;
```

#### **Output :**

The screenshot shows a database query results grid with the following data:

Region	Total_Profit
West	106021.1495000007
East	90672.0127000022
South	46035.6890000004
Central	40128.903

### **Insight:**

While some categories generate high sales, their profit margins may differ. This highlights the importance of analyzing both revenue and profitability together rather than focusing only on sales.

### **◆ Question 8: Monthly Sales Trend**

#### **Query :**

```
SELECT DATE_FORMAT(`Order Date`, '%Y-%m') AS Month, SUM(Sales) AS Monthly_Sales FROM superstore GROUP BY Month ORDER BY Month;
```

**Output :**

The screenshot shows a MySQL Workbench interface with a results grid. The title bar says "SELECT `Sub-Category` , SUM(Profit) AS Total\_Profit". The results grid has two columns: "Month" and "Monthly\_Sales". The data starts from 2016-01 and ends at 2017-05. The "Monthly\_Sales" column contains values such as 18432.590999999993, 22706.415, etc. The bottom of the window shows "Result 11" and an "Output" tab.

Month	Monthly_Sales
2016-01	18432.590999999993
2016-02	22706.415
2016-03	50832.54900000002
2016-04	38587.140999999974
2016-05	56457.319999999985
2016-06	39628.19199999998
2016-07	39108.00799999997
2016-08	31014.4023
2016-09	71848.25089999994
2016-10	58120.42300000001
2016-11	78454.90179999999
2016-12	96075.06700000002
2017-01	43860.28600000001
2017-02	20262.322399999997
2017-03	58739.223799999985
2017-04	36020.17310000001
2017-05	44095.3522

**Insight:**

Profit varies significantly across regions. High-performing regions can be used as benchmarks, while low-performing regions may require targeted marketing or pricing adjustments.

◆ **Question 9: How does discount affect average profit?**

**Query :**

```
SELECT Discount, AVG(Profit) AS Avg_Profit FROM superstore GROUP BY Discount ORDER BY Discount;
```

**Output :**

The screenshot shows a data grid titled "Result Grid" with two columns: "Discount" and "Avg\_Profit". The "Discount" column lists values from 0 to 0.8 in increments of 0.1. The "Avg\_Profit" column lists corresponding values starting at 68.10909261327028 for a 0% discount and decreasing to -105.74120982456134 for an 80% discount. There are also some intermediate values like 0.15, 0.2, 0.3, etc.

	Discount	Avg_Profit
▶	0	68.10909261327028
▶	0.1	96.05507446808508
▶	0.15	27.288298076923077
▶	0.2	25.198563349309303
▶	0.3	-45.679636123348004
▶	0.32	-88.56065555555558
▶	0.4	-112.51435121951225
▶	0.45	-226.64646363636365
▶	0.5	-310.703456060606
▶	0.6	-42.03333030303031
▶	0.7	-101.390593606138
▶	0.8	-105.74120982456134

### Insight:

As discount increases, average profit tends to decrease. This confirms that aggressive discount strategies negatively impact overall profitability.

### ◆ Question 10: Sub-Category Profitability Analysis

#### Query :

```
SELECT `Sub-Category`, SUM(Profit) AS Total_Profit FROM superstore GROUP BY `Sub-Category`  
ORDER BY Total_Profit DESC;
```

#### Output :

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

Sub-Category	Total_Profit
Copiers	55617.82490000001
Phones	44447.87910000006
Accessories	41936.63569999993
Paper	32712.16899999995
Binders	29983.021299999986
Chairs	26590.166300000026
Storage	21527.90829999996
Appliances	18138.00539999995
Furnishings	11588.64199999987
Art	6527.78699999998
Envelopes	6460.869100000035
Labels	5546.25399999998
Machines	3384.7569
Fasteners	942.437799999997
Supplies	-1348.565499999986
Bookcases	-3472.555999999978
Tables	-17725.481100000008

Result 13 ×

### Insight:

As discount increases, average profit tends to decrease. This confirms that aggressive discount strategies negatively impact overall profitability.

### ✓ 4. Conclusion

SQL analysis helped identify performance trends, profit drivers, and loss-making areas. These insights can assist in improving pricing strategy, discount planning, and regional targeting.