# US Superstore Sales Report (Project)

**Submitted BY:**

**Muzafar Ali**

**Submitted To:**

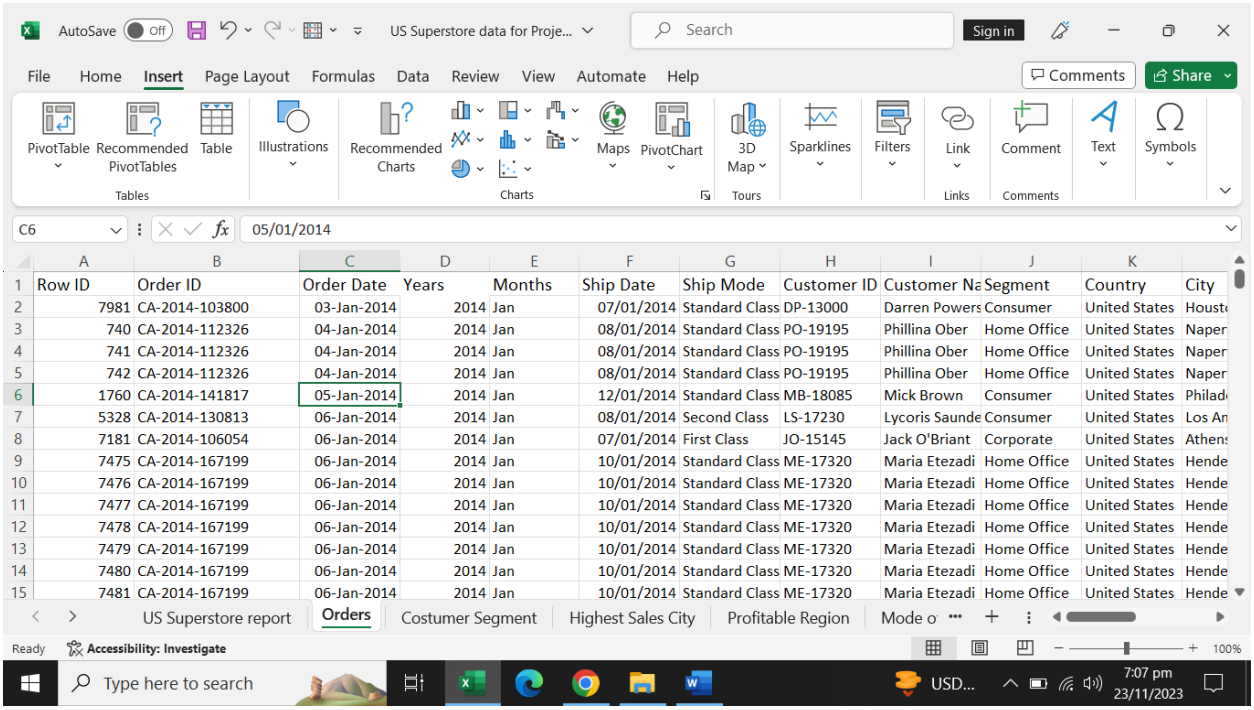
**Sir Awais Ather**

# TABLE OF CONTENTS

## 1) Dataset

- US Superstore data

I used the popular superstore data which is the sales record for a retail business located in the United States. I search and select an appropriate dataset on Kaggle which will match a standard superstore requirement. It shows the records of sold furniture, office supplies and technological products to a customer set of consumers, home offices and corporate entities.




| Row ID | Order ID | Order Date | Years | Months | Ship Date | Ship Mode | Customer ID | Customer Na | Segment | Country | City |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 7981 | CA-2014-103800 | 03-Jan-2014 | 2014 | Jan | 07/01/2014 | Standard Class | DP-13000 | Darren Powers | Consumer | United States | Houst |
| 740 | CA-2014-112326 | 04-Jan-2014 | 2014 | Jan | 08/01/2014 | Standard Class | PO-19195 | Phillina Ober | Home Office | United States | Naper |
| 741 | CA-2014-112326 | 04-Jan-2014 | 2014 | Jan | 08/01/2014 | Standard Class | PO-19195 | Phillina Ober | Home Office | United States | Naper |
| 742 | CA-2014-112326 | 04-Jan-2014 | 2014 | Jan | 08/01/2014 | Standard Class | PO-19195 | Phillina Ober | Home Office | United States | Naper |
| 1760 | CA-2014-141817 | 05-Jan-2014 | 2014 | Jan | 12/01/2014 | Standard Class | MB-18085 | Mick Brown | Consumer | United States | Philad |
| 5328 | CA-2014-130813 | 06-Jan-2014 | 2014 | Jan | 08/01/2014 | Second Class | LS-17230 | Lycoris Saunde | Consumer | United States | Los An |
| 7181 | CA-2014-106054 | 06-Jan-2014 | 2014 | Jan | 07/01/2014 | First Class | JO-15145 | Jack O'Briant | Corporate | United States | Athens |
| 7475 | CA-2014-167199 | 06-Jan-2014 | 2014 | Jan | 10/01/2014 | Standard Class | ME-17320 | Maria Etezadi | Home Office | United States | Hende |
| 7476 | CA-2014-167199 | 06-Jan-2014 | 2014 | Jan | 10/01/2014 | Standard Class | ME-17320 | Maria Etezadi | Home Office | United States | Hende |
| 7477 | CA-2014-167199 | 06-Jan-2014 | 2014 | Jan | 10/01/2014 | Standard Class | ME-17320 | Maria Etezadi | Home Office | United States | Hende |
| 7478 | CA-2014-167199 | 06-Jan-2014 | 2014 | Jan | 10/01/2014 | Standard Class | ME-17320 | Maria Etezadi | Home Office | United States | Hende |
| 7479 | CA-2014-167199 | 06-Jan-2014 | 2014 | Jan | 10/01/2014 | Standard Class | ME-17320 | Maria Etezadi | Home Office | United States | Hende |
| 7480 | CA-2014-167199 | 06-Jan-2014 | 2014 | Jan | 10/01/2014 | Standard Class | ME-17320 | Maria Etezadi | Home Office | United States | Hende |
| 7481 | CA-2014-167199 | 06-Jan-2014 | 2014 | Jan | 10/01/2014 | Standard Class | ME-17320 | Maria Etezadi | Home Office | United States | Hende |

## 2) Problem Statement

The store named “The US Superstore” generated a huge amount of data from sales over the past 4 years which has lied dormant and without use. The owners hope to leverage on it to help drive the growth of the store over the coming years.

## 2.1) Questions:

- **What is the total revenue generated by the store?**

### Total Revenue

| 3 | Sum of Sales |
|---|---|
| 4 | $2,297,201 |

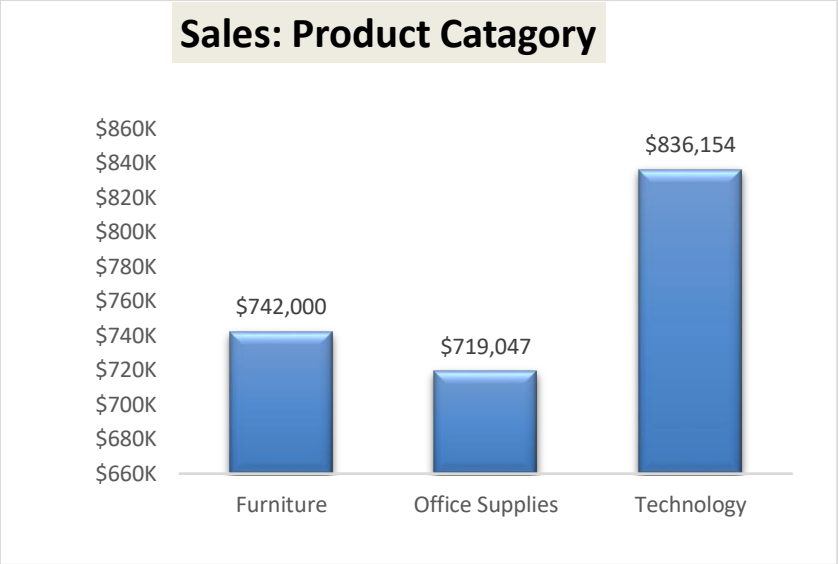
The total revenue generated by the store is $2,297,201.

- **Which category of products contributes the most to sales?**

### Sales By Product Category

| 3 | Row Labels | Sum of Sales |
|---|---|---|
| 4 | Furniture | $742,000 |
| 5 | Office Supplies | $719,047 |
| 6 | Technology | $836,154 |

Pivot table was used to show different Product Category and the total sales made by each category. It was found that Technology generated the highest revenue.

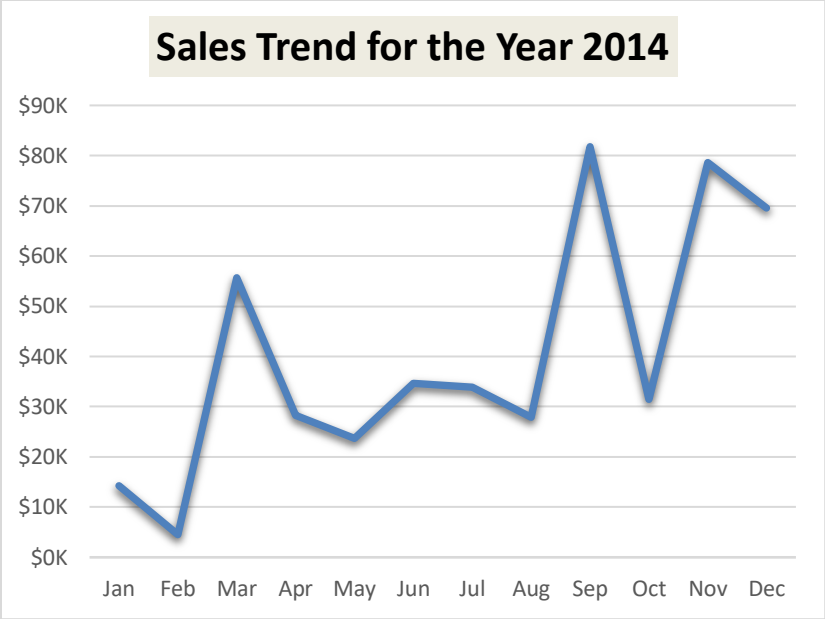



- **How has the sales trend been for the year 2014?**

## Sales trend for the year 2014

| | | |
|---|---|---|
| 3 | Years | 2014 |
| 4 | | |
| 5 | **Row Labels** | **Sum of Sales** |
| 6 | Jan | $14,237 |
| 7 | Feb | $4,520 |
| 8 | Mar | $55,691 |
| 9 | Apr | $28,295 |
| 10 | May | $23,648 |
| 11 | Jun | $34,595 |
| 12 | Jul | $33,946 |
| 13 | Aug | $27,909 |
| 14 | Sep | $81,777 |
| 15 | Oct | $31,453 |
| 16 | Nov | $78,629 |
| 17 | Dec | $69,546 |

This pivot table shows the total sales made in each month of the year 2014. Line chart is used to show the trend of sales in the year 2014.

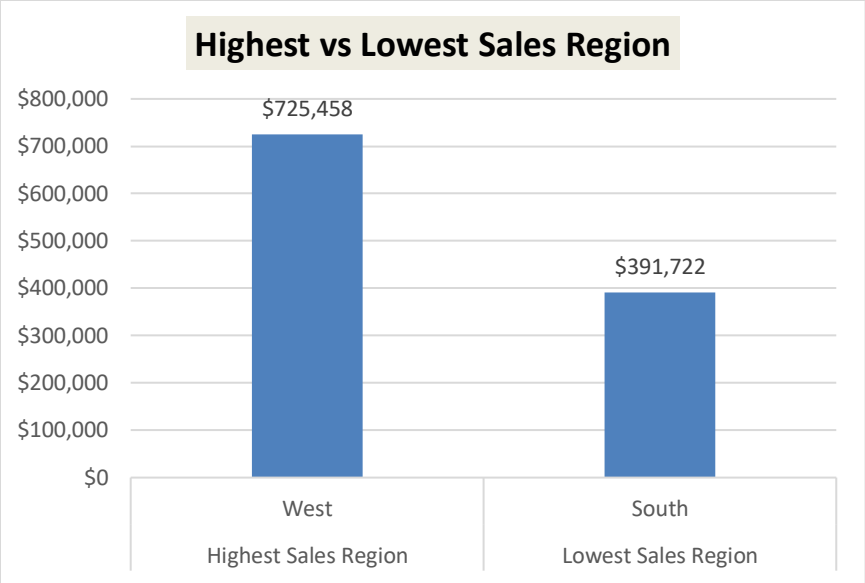



- **Which region has the highest sales, and which one has the lowest?**

## Highest vs Lowest sale region

| 3 | Row Labels | Sum of Sales |
|---|---|---|
| 4 | West | $725,458 |
| 5 | East | $678,781 |
| 6 | Central | $501,240 |
| 7 | South | $391,722 |

This pivot table shows the sales made by different regions. By sorting from highest to lowest, it was found that West region produced the highest sales while South region made the lowest sales.




- **What is the average profit margin of the store?**

Average Profit Margin

| | Average of Profit Margin |
|---|---|
| 3 | |
| 4 | $12.03 |

The average profit margin of the store is $12.03.

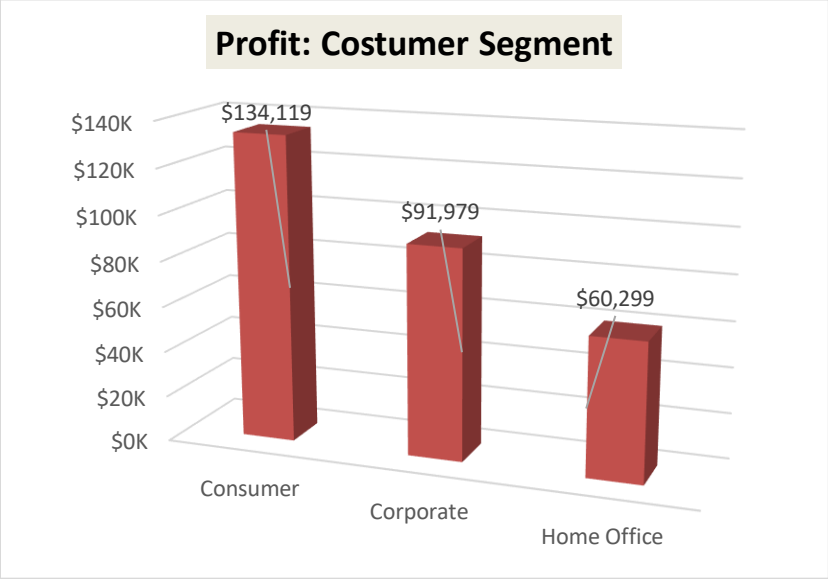
- **Which costumer segment is the most profitable?**

## Profit: By Costumer Segment

Customer segment was classified into consumers, home offices and corporate for the average person with needs, people with workspaces in their homes and business entities.

| 3 | Row Labels | Sum of Profit |
|---|---|---|
| 4 | Consumer | $134,119 |
| 5 | Corporate | $91,979 |
| 6 | Home Office | $60,299 |
| 7 | | |

The consumer segment was the most profitable. It is followed by corporate segment and home office.



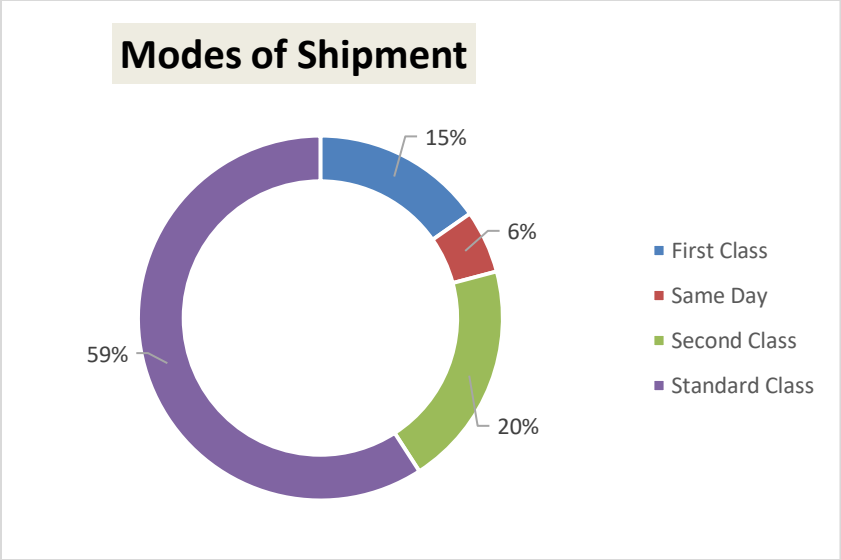

- **Determine the portion of sales generated through different modes of shipment?**

## Mode Of Shipment

The store had 4 systems of shipping ordered products namely first class, same day, second class and standard class.

| 3 | Row Labels | Sum of Sales |
|---|---|---|
| 4 | First Class | 15% |
| 5 | Same Day | 6% |
| 6 | Second Class | 20% |
| 7 | Standard Class | 59% |
| 8 | | |

This pivot table used to find portion of sales generated through different mode of shipment. It was found that the 59% of the sales generated thorough standard class.



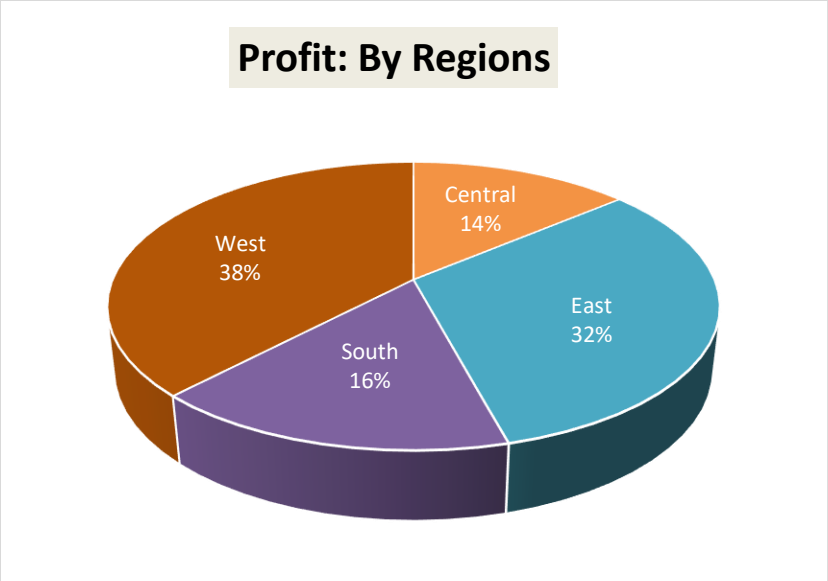

- **Determine the most profitable region?**

## Profit: By Regions

The store sold items to all regions of the United States, Central, East, South and West.

| 3 | Row Labels | Sum of Profit |
|---|---|---|
| 4 | Central | $39,706 |
| 5 | East | $91,523 |
| 6 | South | $46,749 |
| 7 | West | $108,418 |

The West and East regions came up as the most profitable regions while the Central made the least profit.

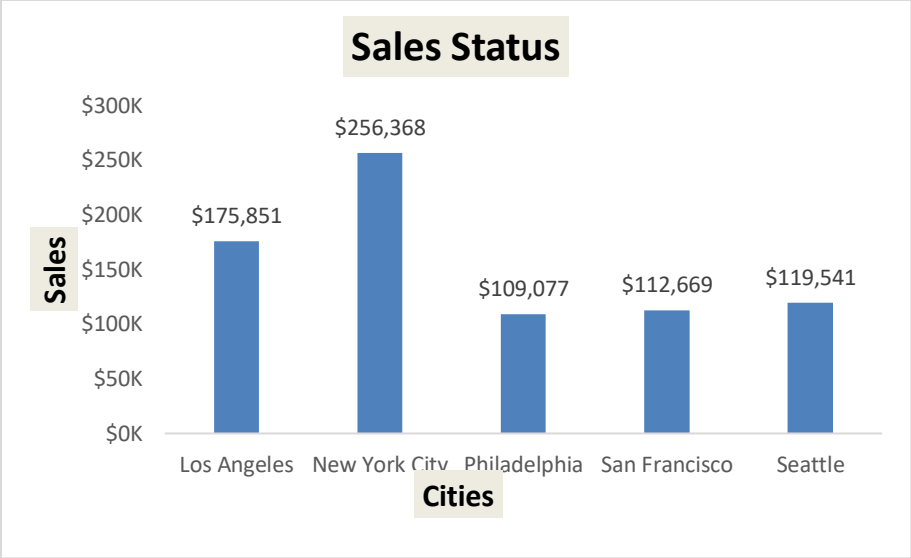



- **Which city generated the highest revenue?**

## Revenue: By City

| 3 | Row Labels | Sum of Sales |
|---|---|---|
| 4 | Los Angeles | $175,851 |
| 5 | New York City | $256,368 |
| 6 | Philadelphia | $109,077 |
| 7 | San Francisco | $112,669 |
| 8 | Seattle | $119,541 |

New York city generated the highest revenue, and this is backed by the fact that it is the most populated city in the United States, and a city with a lot of organizations and corporate entities.




## 2.2) Analysis:

The project is to provide in-depth analysis to help the store owners understand their data trends and patterns and provide valuable insights to drive better business decisions. The following criteria been used for analysis:

- performance of product categories in terms of revenue and profit.
- performance of sub-categories of products.
- the preferred ship mode among customers
- the biggest revenue generator among all customer segments
- the regions and cities with revenue generated and profit made.

## 2.3) Issues:

Common issues such as identifying missing data (empty cells), misspellings, duplicate data were sorted to ensure the integrity of the data.

Columns were formatted with prices set to the currency, order and ship dates columns set to the short date format and others converted to correlate with their corresponding data types.

New fields were created to get the important numbers as the numbers given could not generate the insight needed alone.

# 3) Methodology

For this project, I used Microsoft Excel for cleaning, analysis/data exploration, and Visualization.

## 3.1) Data Cleaning

Before embarking on any analysis, it is essential to ensure the dataset is clean and reliable. The data cleaning process involves handling missing values, correcting inconsistencies, and formatting the data for analysis. In this project, I carefully cleaned the dataset to ensure the accuracy and integrity of findings.

## 3.2) Excel Table

Excel tables provide an array of features to effectively analyze and manage data such as calculated columns, total row, auto-filter and sort options, automatic expansion of a table, and more.

## 3.3) Checking for duplicates

I highlighted the entire worksheet to locate duplicates. I then looked at the Data ribbon to see if there were any duplicate values, but none were found.

## 3.4) Data Exploration

This phase entails providing answers to business inquiries to help Client (Superstore) improve their understanding their data.

## 3.5) Pivot Table

A pivot table is a sophisticated data summarizing tool that sorts, counts, and sums up data contained in tables automatically. In this section, I developed some pivot tables to compare and identify trends and relationships between some indicators in our dataset.

# 4) Results

1) Technology sub-category contributes the most to sales.

2) West is the highest sales region and South is the Lowest sales region.

3) $12.03 is the average profit margin of the store.

4) Consumer costumer segment is the most profitable segment.

5) 59% sale generated thorough standard class mode of shipment.

6) West is the most profitable region.

7) New York City generated the highest sales.

## References


- Kaggle.com
- Open.ai