Advanced Superstore Sales Analysis and Dashboard Development using Tableau

Abstract

This report presents an analysis of Superstore sales data. The superstore data contains information about orders, products, location and customer. It indicates orders made by the customer, their price, location of the store and other valuable information.

This project outlines the specifications for building two dashboards using tableau to help stakeholders, including sales managers and executives to analyze sales performance and customers.

Introduction

Situation: Understanding sales of any company and overall customer demands will be helpful in improving inventory management and overall customer experience. This not only helps the customer but also helps the company to grow and cater the demands of the customers effectively.

Task: This project aims to analyze sales data to identify the key products driving and build interactive dashboards to present these findings to the company.

By understanding these factors, the company can take proactive steps to improve customer experience, improve inventory management and boost overall growth of the company.

Methodology

The analysis was conducted using a different datasets containing various customer information, location, their order information and products delivered at SuperStore.

Dataset Link: GitHub

[1] The steps involved in the analysis are as follows:

1. Analyse Requirements: We have to first understand the problem and understand what kind of requirements we are going to address using this project.

[1] https://github.com/Rajat4445/SuperStore-Analysis/tree/main/datasets

2. Data Cleaning and Merging:Handling missing values, correcting inconsistencies, ensuring data quality and merging data tables using Python or MS Excel (if needed).
3. Exploratory Data Analysis (EDA) : Visualizing data distributions, relationships, and patterns using different charts in Tableau.
4. Dashboarding : Create interactive dashboard and stories to present to the stakeholders and understand sales data better.

Results

The results of the analysis are presented in the form of visualizations and dashboards created in Tableau.

Analyse Requirements

We are required to create two different dashboards serving different purposes namely, **Sales Dashboard** and **Customer Dashboard**, let us understand their purpose and requirements as well.

Sales Dashboard

Dashboard Purpose

The purpose of sales dashboard is to present an overview of the sales metrics and trends in order to analyze year-over-year sales performance and understand sales trends.

Key Requirements

KPI Overview: Display a summary of total sales, profits and quantity for the current year and the previous year.

Sales Trends:

- Present the data for each KPI on a monthly basis for both the current year and the previous year.
- Identify months with highest and lowest sales and make them easy to recognize.

Product Subcategory Comparison:

- Compare sales performance by different product subcategories for the current year and the previous year.
- Include a comparison of sales with profit.

Weekly Trends for Sales & Profit:

- Present weekly sales and profit data for the current year.
- Display the average weekly values.

- Highlight weeks that are above and below the average to draw attention to sales & profit performance.

Customer Dashboard

Dashboard Purpose

The customer dashboard aims to provide an overview of customer data, trends and behaviors. It will help marketing teams and management to understand customer segments and improve customer satisfaction.

Key Requirements

KPI Overview: Display a summary of total number of customers , total sales per customer and total number of orders for the current year and the previous year..

Customer Trends:

- Present the data for each KPI on a monthly basis for both the current year and the previous year.
- Identify months with highest and lowest sales and make them easy to recognize.

Customer Distribution by Number of Orders:

Represent the distribution of customers based on the number of orders they have placed to provide insights into customer behavior, loyalty and engagement.

Top 10 Customers By Profit:

- Present the top 10 customers who have generated the highest profits for the company.
- Show additional information like rank, number of orders, current sales, current profit and the last order date.

Design & Interactivity Requirement

Dashboard Purpose

The customer dashboard aims to provide an overview of customer data, trends and behaviors. It will help marketing teams and management to understand customer segments and improve customer satisfaction.

Dashboard Dynamics:

- The Dashboard should allow users to check historical data by offering them the flexibility to select any desired year.
- Provide users with the ability to navigate between the dashboards easily.
- Make the charts and graphs interactive, enabling users to filter data using the charts.

Data Filters: Allow users to filter data by product information like category and subcategory and by location information like region, state and city

Visualization and Dashboard Creation Process

Fortunately, there were no missing values or inconsistencies in our data, so we do not need to perform any data cleaning. But, since the data is in different tables (orders, customers, products and location), we will have to perform merging, which can be done in Tableau.

Before performing any EDA, we will rename fields (columns), ensure correct datatypes to the columns and overall understand the data. In this section, we will explain all the steps followed to create the visualisations and dashboards. This section can be skipped if someone directly wants to go into **Insights** section, where all the insights have been enlisted. (Page 13)

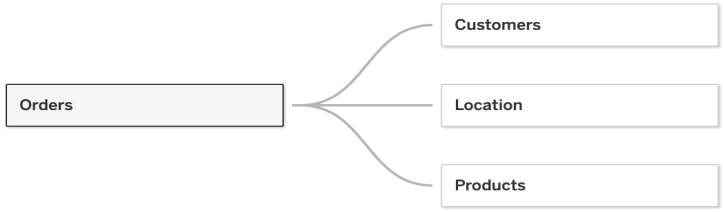

Figure 1: Datasource

- We had four tables namely **Orders** which is related to **Customers** table with common column *Customer ID*. Similarly, other tables such as **Location** and **Products** are also related to **Orders** table with common columns as *Postal Code* and *Product ID* respectively. After joining them we have our main datasource named **Sales DataSource**.

Now, we are first going to create **Sales Dashboard** and for that we will have to create visualisations. Looking at the key requirements from previous sections, we need to display total sales, profits and quantity for current year and previous year. This must be presented on a monthly basis for all with pointers to identify months with highest and lowest values of total sales, profits and quantity.

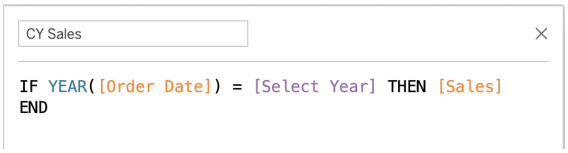

Figure 2: Syntax for Current Year sales

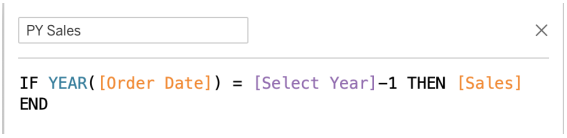

Figure 3: Syntax for Previous Year sales

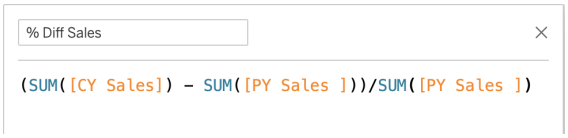

Figure 4: Syntax for % of Sales Difference between Current Year and Previous Year

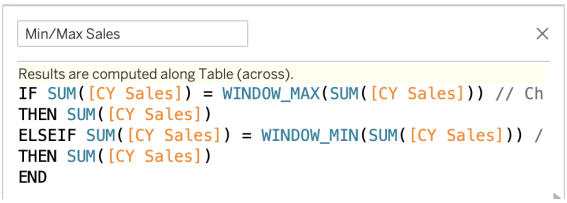

Figure 5: Syntax for months with Min and Max Sales for each year

The above syntax have been used to create calculated fields **CY Sales**, **PY Sales**, **% of Sales Difference** and **Months with Min and Max Sales for each year**. In the above syntax, **Select Year** is a parameter which has been created to ensure dynamic selection of years.

- To create the Sales KPI dashboard, we have created a dual-axis chart with one chart denoting **CY Sales vs PY Sales** and other denoting months with minimum and maximum sales.

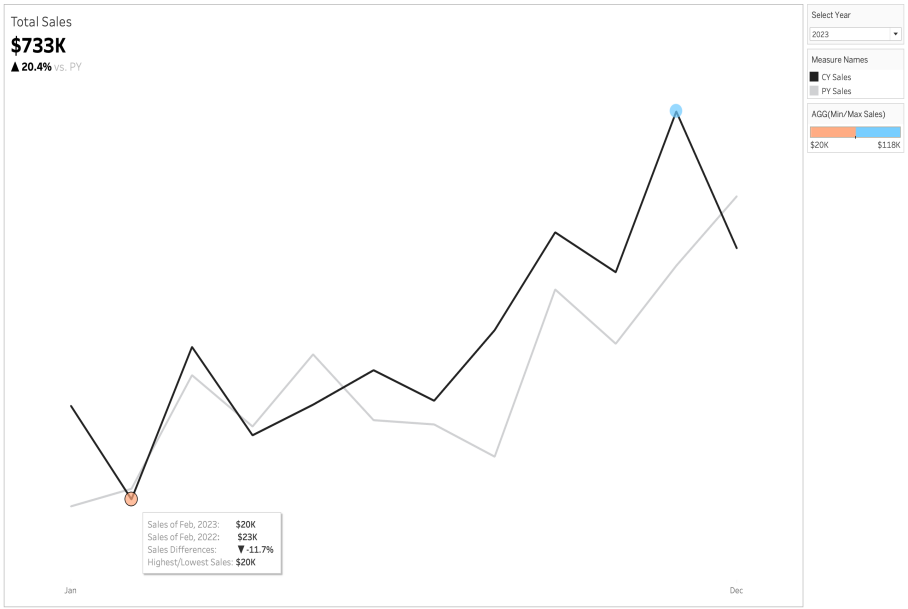

Figure 6: KPI Sales

- In **KPI Sales**, we have used **Orange** color to indicate minimum sales and **blue** color to indicate maximum sales. Similar theme has been used to indicate loss (in **Orange**) and profit (in **blue** in the overall visualisations.

- We also modified our tooltip to indicate sales for current year month, previous year month, sales difference (in percentage)

Similarly, we created **KPI Profits** and **KPI Quantity** following similar color scheme and creation of calculated fields. Below are the snapshots for the same.

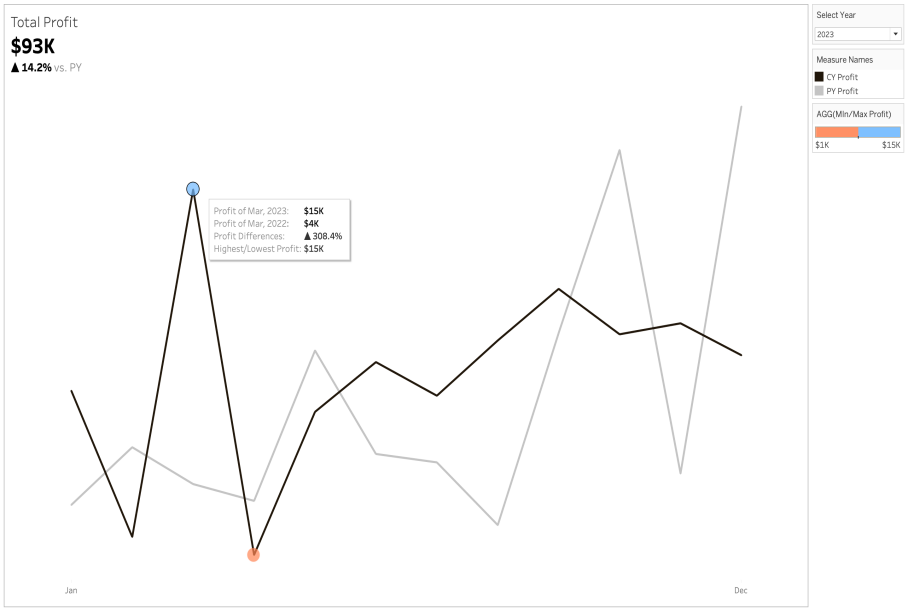

Figure 7: KPI Profits

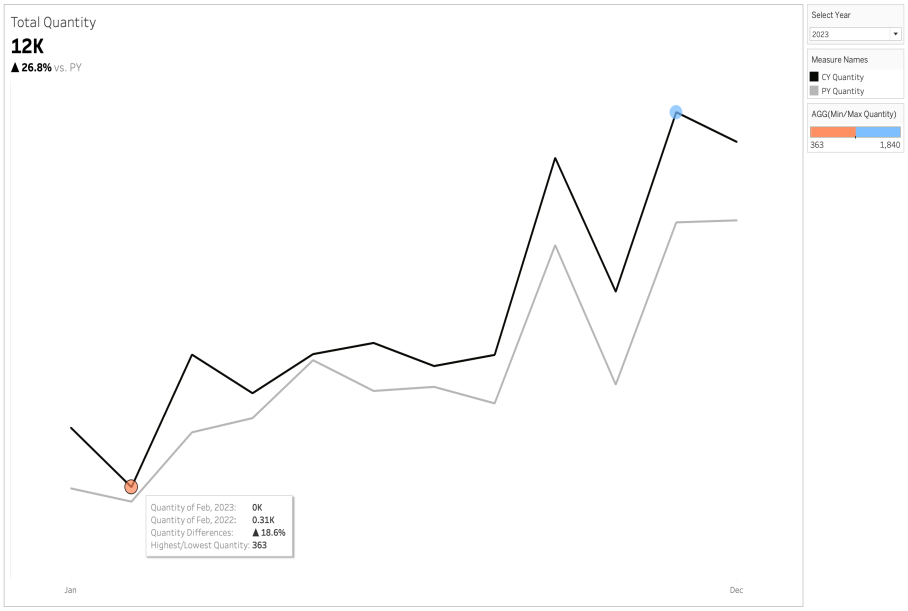

Figure 8: KPI Quantity

We also need to create a visualisation for **Product Subcategory Comparison**, for this we would

be using a dual-axis bar chart (with CY Sales and PY Sales) alongwith a bar chart to indicate profits for those subcategories.

Also, we would be defining another metric which would indicate the subcategories whose CY Sales are lower than PY Sales.

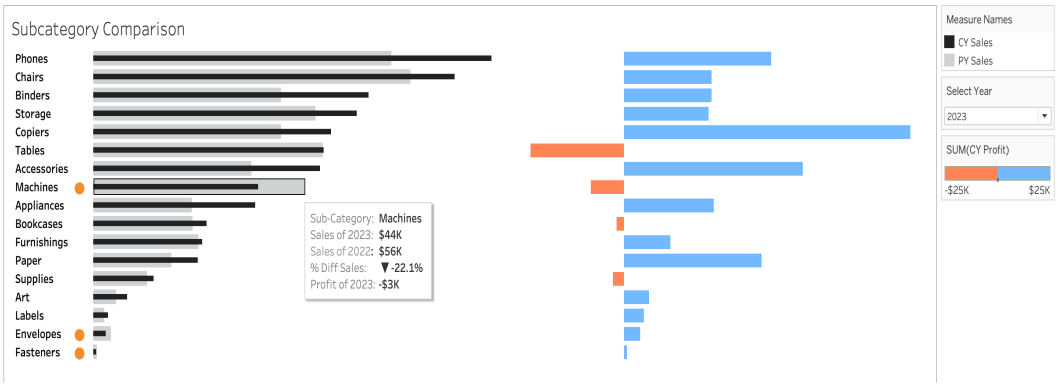

Figure 9: Subcategory Comparison

Note: The **Orange** color dot besides subcategory indicates that CY Sales for that subcategory is lower than PY Sales.

We also need to make a chart to show present weekly sales and profit for current year. Also, we need to the reference line for average weekly value.

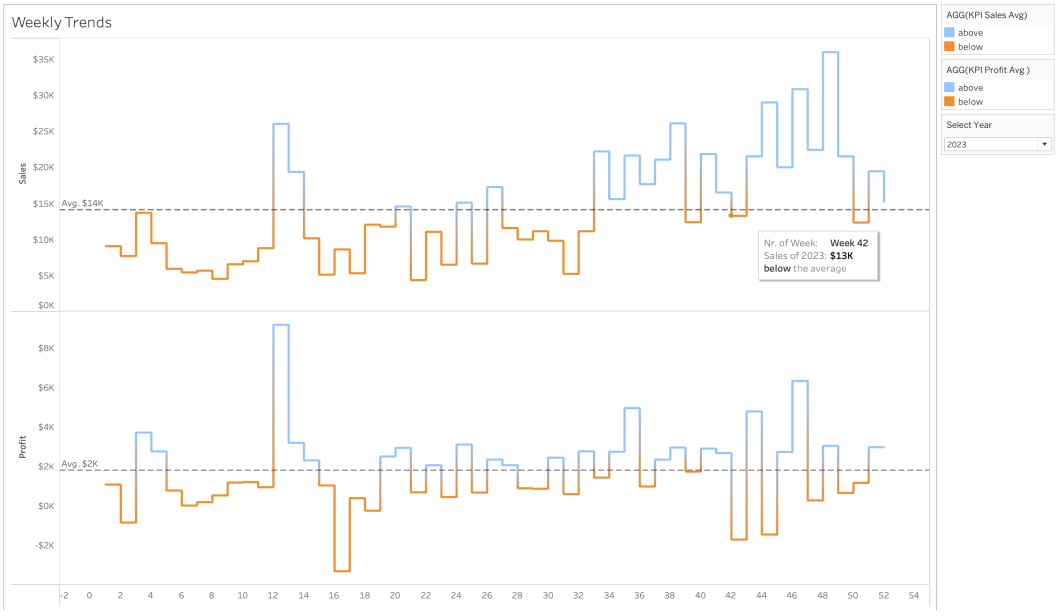

Figure 10: Weekly Trends

Above chart is dynamic where years can be changed using the **Select Year** parameter.

Finally, after creating all the required visualisations and KPIs, we would create our **Sales Dashboard**. Since, the process of creating dashboard requires extensive use of different dashboarding elements, explaining it would be a little long. So, we are refraining for it but similar dashboard (to the one we are attaching) can be created using simpler techniques.

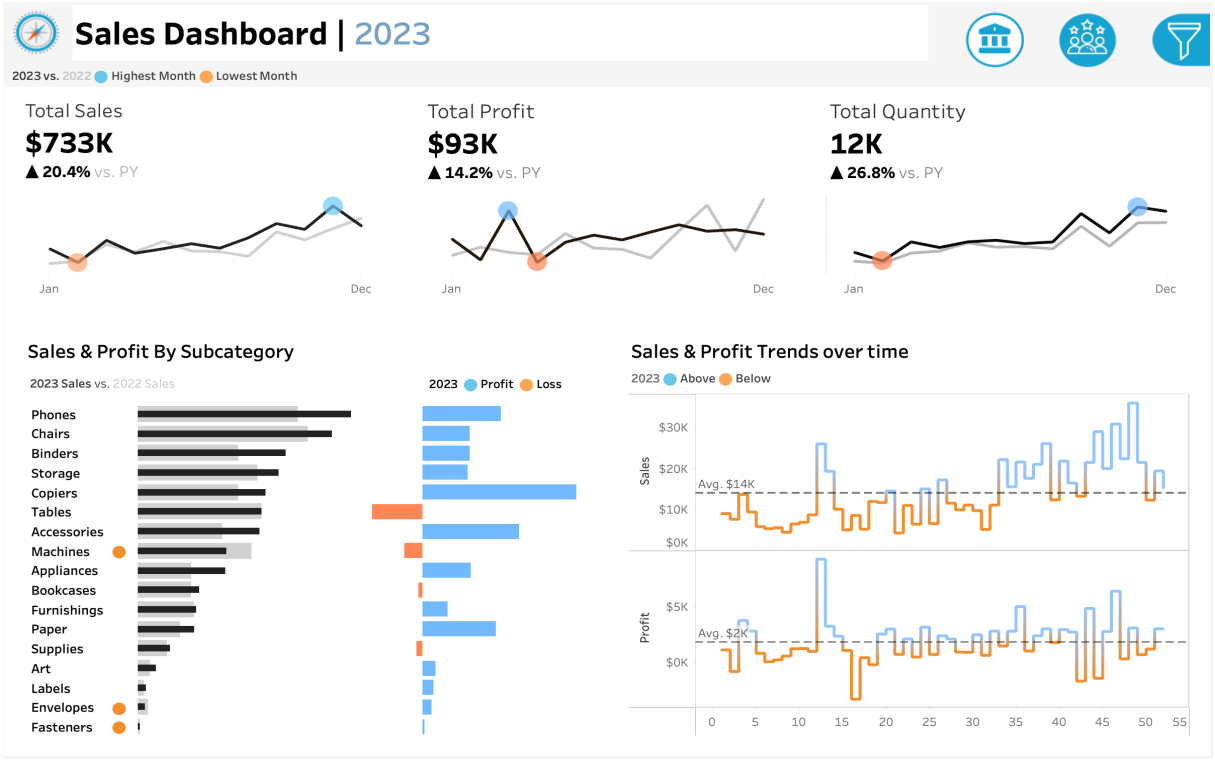

Figure 11: Sales Dashboard

The above dashboard is dynamic with filters to further drill down valuable insights.

Now, we need to create customer dashboard. So, the approach is similar to Sales Dashboard with little changes in the syntax which would now perform calculations for **Total Customers**, **Total Sales per Customer** and **Total Orders**. Also, we have calculated **% of Total Customer difference** and **Min and Max months for each year** (similar to Sales KPIs).

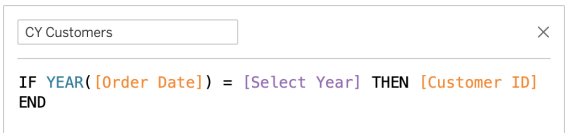

Figure 12: Syntax for total Current Year customers

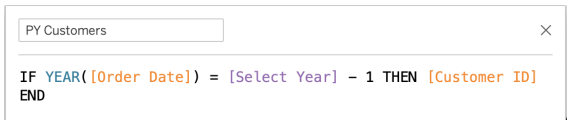

Figure 13: Syntax for total Previous Year customers

Figure 14: Syntax for total Current Year sales per customer

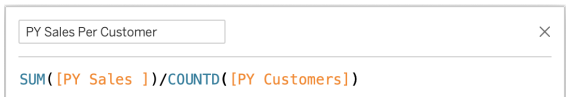

Figure 15: Syntax for total Previous Year sales per customer

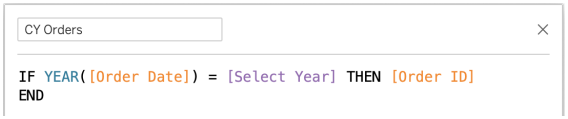

Figure 16: Syntax for total Current Year orders

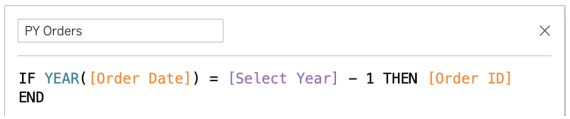

Figure 17: Syntax for total Previous Year Orders

- Similarly, we have created calculated fields for **% of difference** between CY and PY customers, Sales per customer and orders. Also, we have created Minimum and Maximum indicating metrics for each of the previous mentioned KPIs (similar to what we did for Sales Dashboard).
- Next, we created visualisations similar to Sales Dashboard (following similar steps). Below are the snapshots.

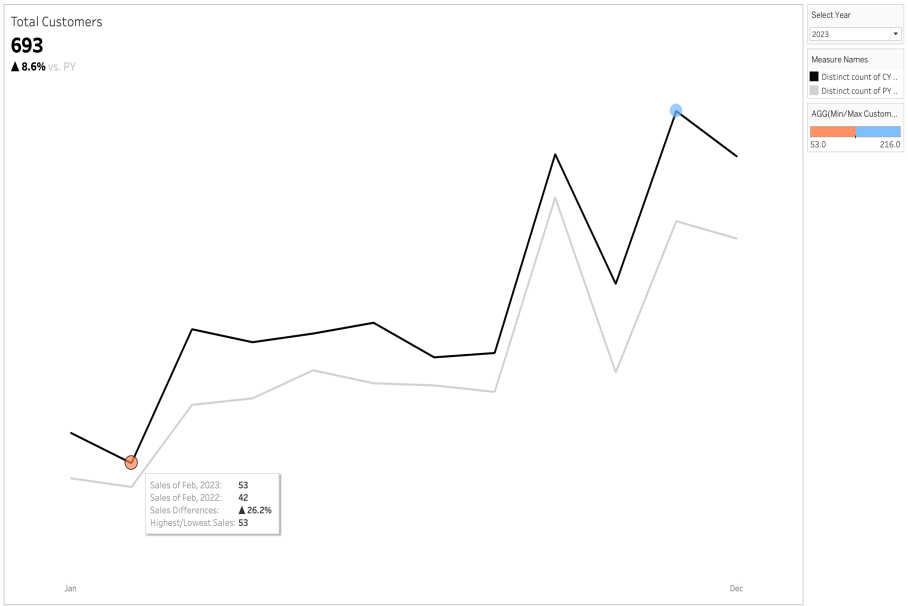

Figure 18: KPI Customers

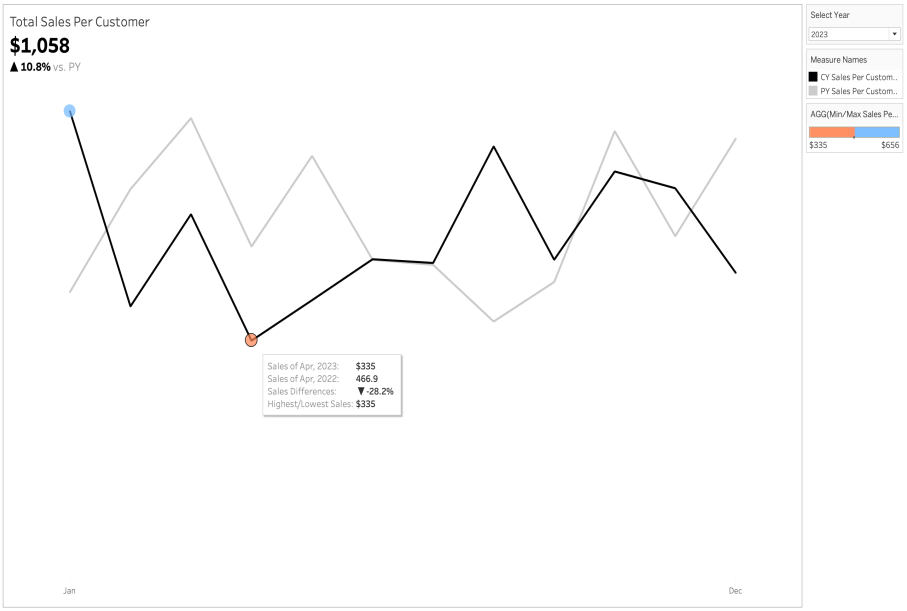

Figure 19: KPI Sales per customer

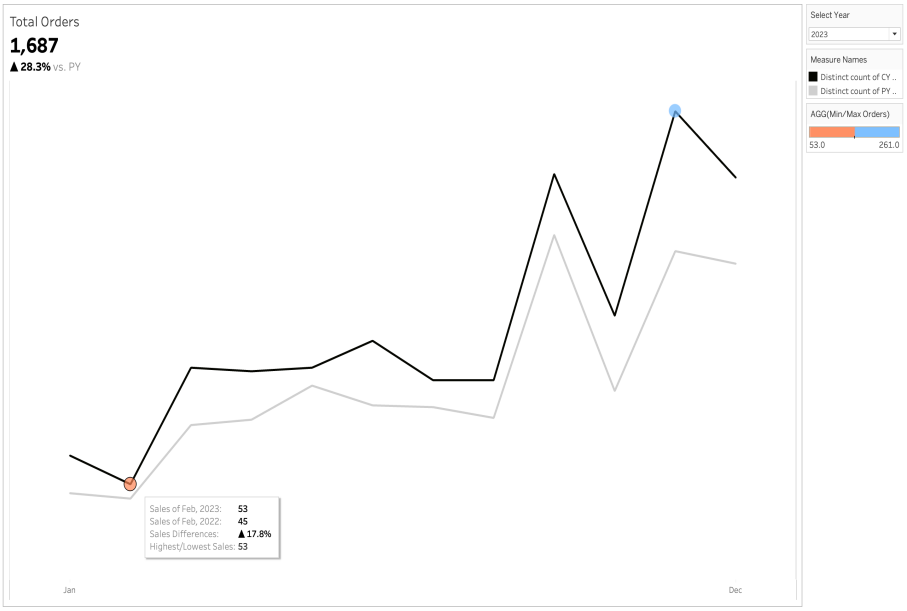

Figure 20: KPI Orders

Also, we have modified our tooltip similar to what we did in KPIs of the Sales Dashboard.

- We have to create a histogram showing the number of orders made by the customer. This would help us knowing customer loyalty and regularity of purchase.
- We would be using LOD Expressions to find this. So, x-axis of the histogram would indicate **Number of orders per customer** and y-axis would indicate **Number of customers** who placed that many orders.

Nr of Orders per Customers

```
{ FIXED [CY Customers]: COUNTD([CY Orders]) }

// Here, for each customer we are finding the number of orders
// they have placed in that (current) year
```

Figure 21: Syntax for finding Number of orders per customer in Current Year

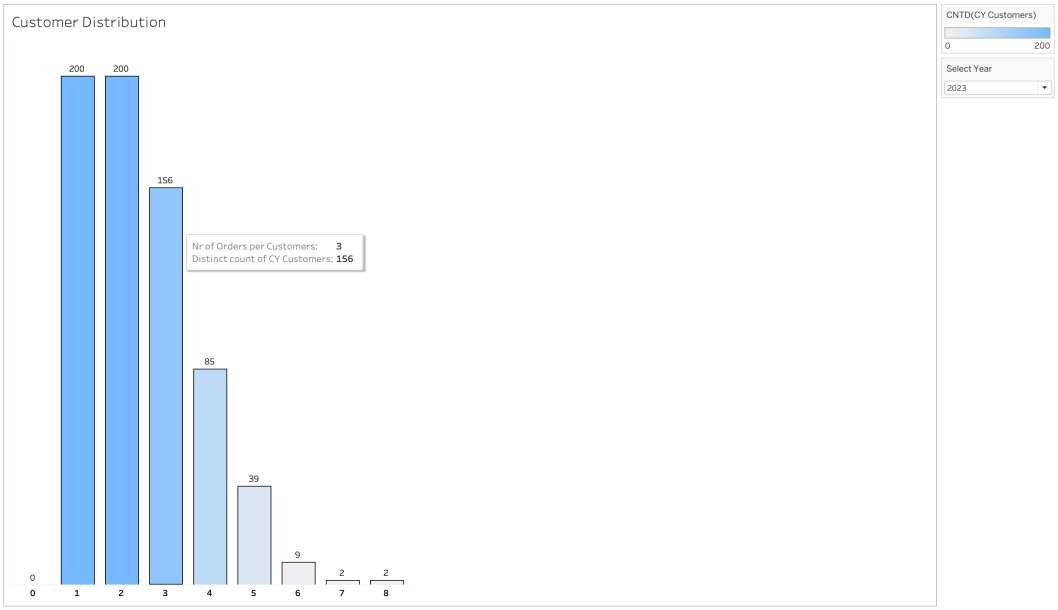

Figure 22: Customer Distribution

- Finally, we have to create a table, which is going to indicate our top 10 Customers by Profit. Additionally, we have to show their ranks, number of orders, current profit and last order date.
- We have used CY Profit, CY Sales and CY Orders, to make this table.

Top 10 Customers by Profit

Rank	Customers	Last Order	2023 Profit	2023 Sales	# of Orders
#1	Raymond Buch	25/09/2023	$6,781	$14,203	3
#2	Hunter Lopez	17/11/2023	$5,046	$10,523	2
#3	Tom Ashbrook	22/10/2023	$4,599	$13,723	2
#4	Andy Reiter	24/12/2023	$2,608	$5,821	2
#5	Jane Waco	18/11/2023	$1,953	$5,385	4
#6	Helen Wasserman	04/09/2023	$1,947	$8,166	5
#7	Brian Moss	27/11/2023	$1,938	$5,683	5
#8	Alan Dominguez	01/12/2023	$1,867	$5,434	4
#9	Jim Epp	12/11/2023	$1,704	$4,074	4
#10	Steven Roelle	17/06/2023	$1,676	$3,506	1

Figure 23: Top 10 Customers by Profit for Current Year (2023)

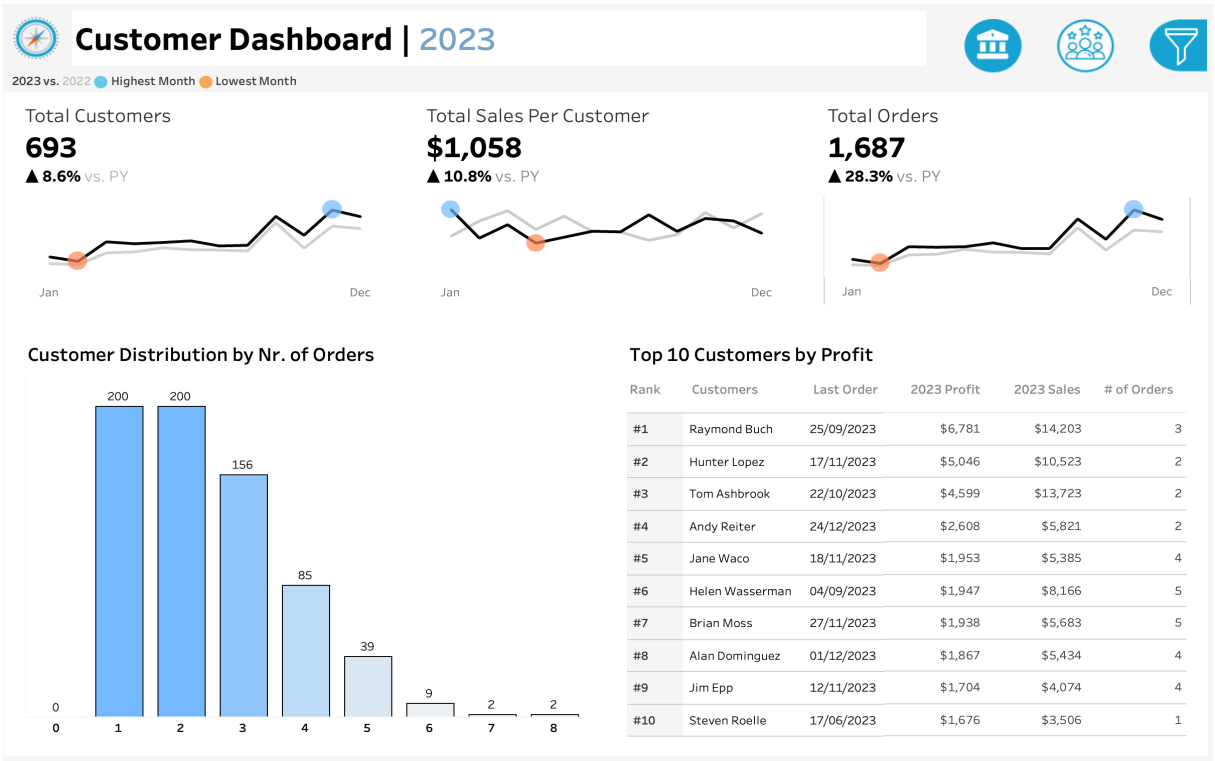

Figure 24: Customer Dashboard

- Now, we would simply make a duplicate of Sales Dashboard and arrange the Customer visualisations, i.e., KPI Customers, KPI Sales per Customer, KPI Orders, Customer Distribution and Top 10 Customers by Profit.
- After rearranging, we would connect this dashboard to the filters made for the Sales Dashboard to enhance consistency across the results.
- Finally, we did some modifications to allow user to toggle between the dashboards using buttons and use filters also.

Insights

Since, our data is for the time range (in years) from 2020 to 2023. Let us look for major insights of each year and also check for the growth (or decline) with the upcoming years.

Year: 2020

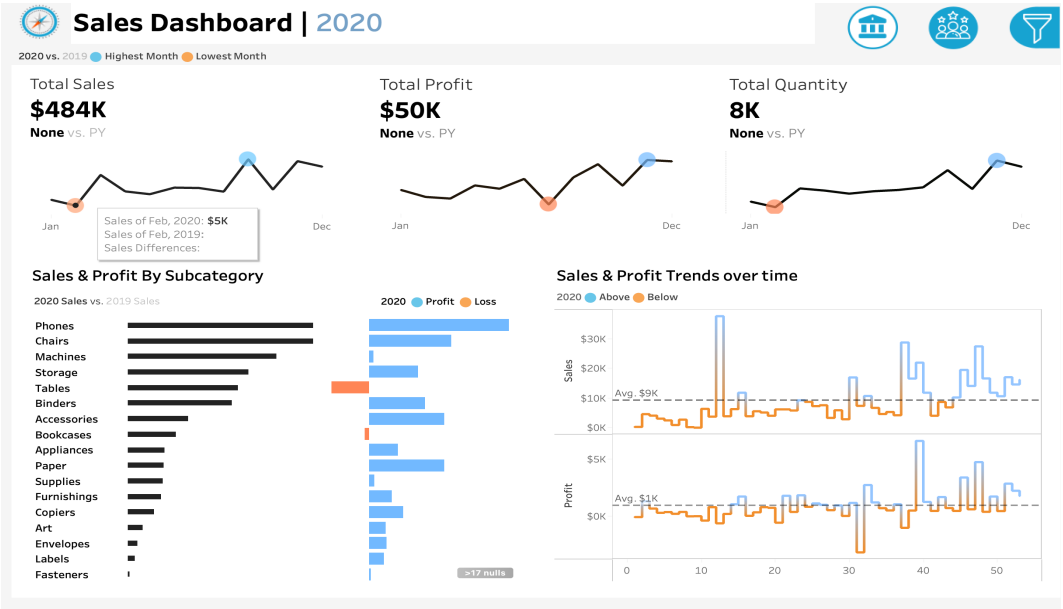

Figure 25: Sales Dashboard 2020

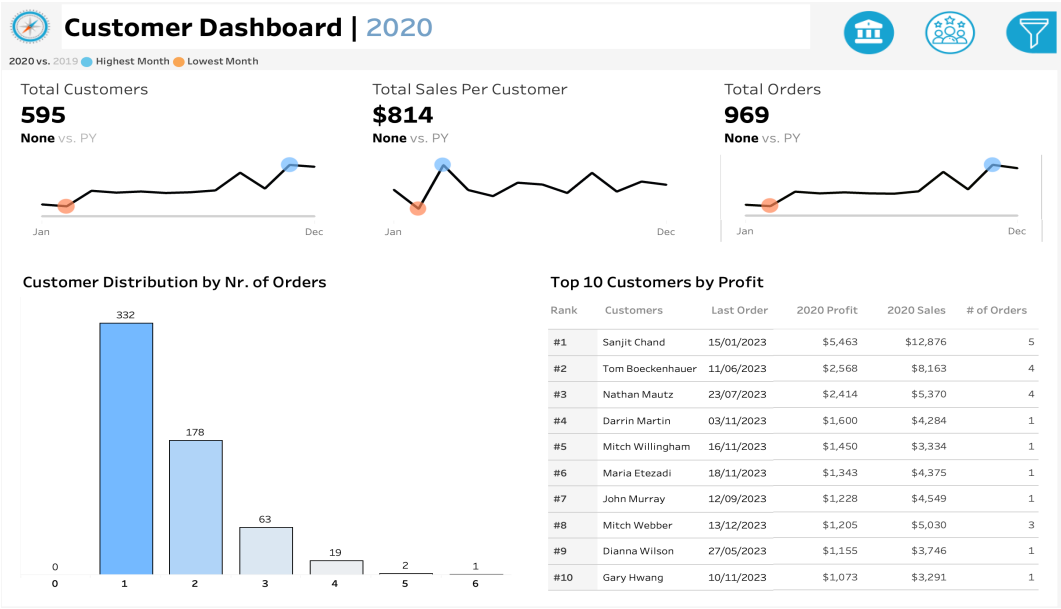

Figure 26: Customer Dashboard 2020

- Company had a total sales of $484K with lowest sales in February and highest in September. Lowest quantity was also sold in February and highest in November. Similarly, highest profits were observed in November.
- Phones were the most popular subcategory in sales while also being the most profitable, followed by Chairs, which sold as much as phones but were a significantly less profitable (around 40% less profitable).
- Looking at sales and profit trends, they are below average for the most part of year, but become significantly higher after week 37 (around September).
- In September (highest sales month), machines make most of sales but create most of loss, while other subcategories remain profitable.
- Machine (subcategory) does not make sales around the year (except March, where it makes loss), but in November and December, where it is significantly profitable.
- Binders in July are not profitable to sell.But drive the highest profits (around $7K) in September, making them more suitable to get sold in September.
- In 2020, Sanjit Chand was the most profitable customer, who only bought binders. Since, his order count was also high (i.e., 5), this might be for his business purpose.
- Other top customers (ranked #2, #3 and #4), bought machines only which led to huge profits for the store.
- Machines, Binders, Tables and Appliances have been a loss making sub categories in Central region over year 2020.

Year: 2021

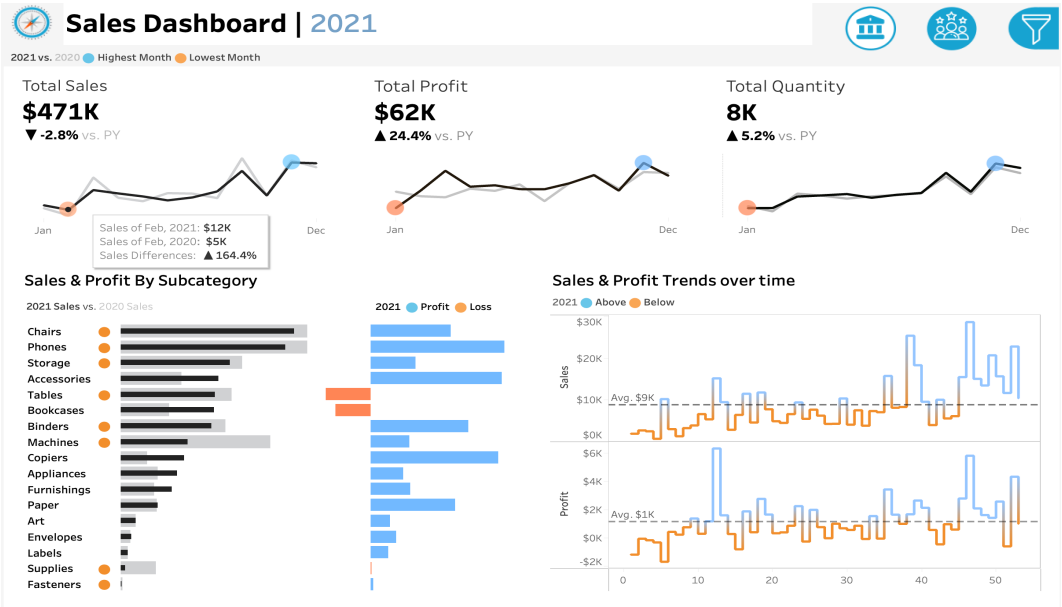

Figure 27: Sales Dashboard 2021

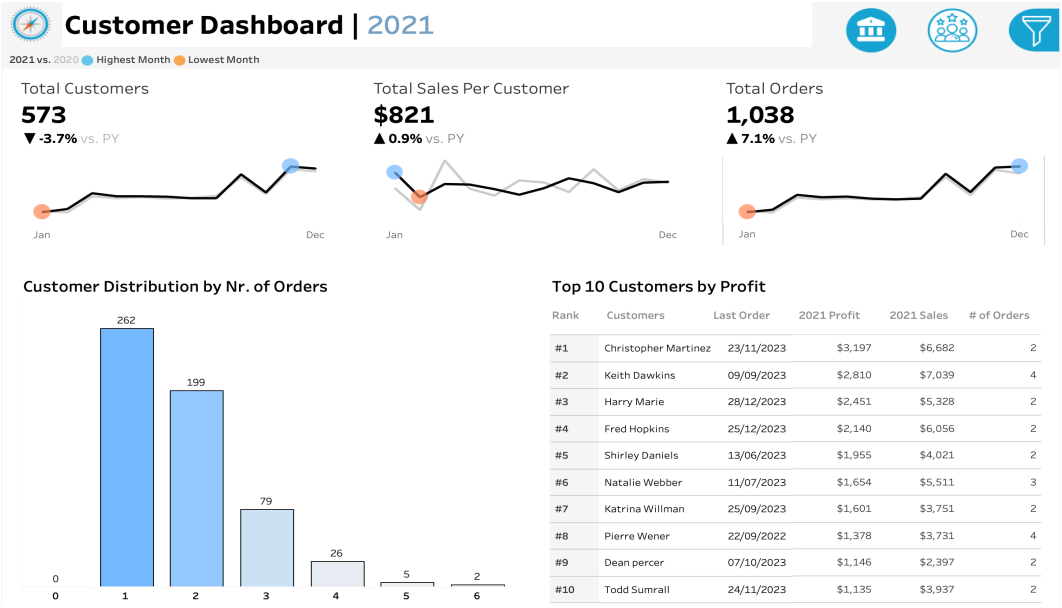

Figure 28: Customer Dashboard 2021

- Although sales decreased by roughly 3%, total profits have increased by 24.4% comparing to previous year. Also, total quantity has increased by 5.2%.
- Many subcategories (with orange dot besides them) have sales lower than previous year, yet they have turned out to be profitable.
- Lowest sales for 2021 were in February (same as 2020) and highest sales were in November (like 2020). But contrary to 2020, sales of nearly all subcategories were higher, especially binders having a sales difference of nearly 700%. This lead to overall profit increase in month by 200% (approx).

  Reason: Binders had more sales in major sales states (New York, California, Michigan and Georgia) comparing to 2020, which turned out to be overall profitable.
- Top 10 Customers in 2021, are mostly buying one of subcategories comprising Binders, copiers and Machines. These customers seem to be buying these for their own businesses as their last order was also in the most recent year, i.e., 2023. Similar is true for top 10 profitable customers of 2020.

Year: 2022

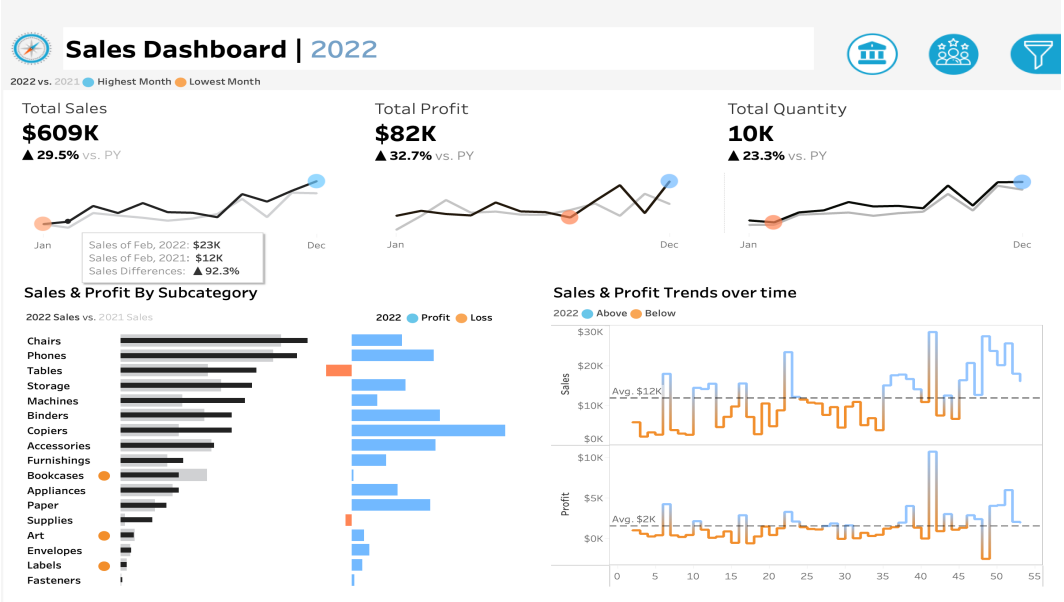

Figure 29: Sales Dashboard 2022

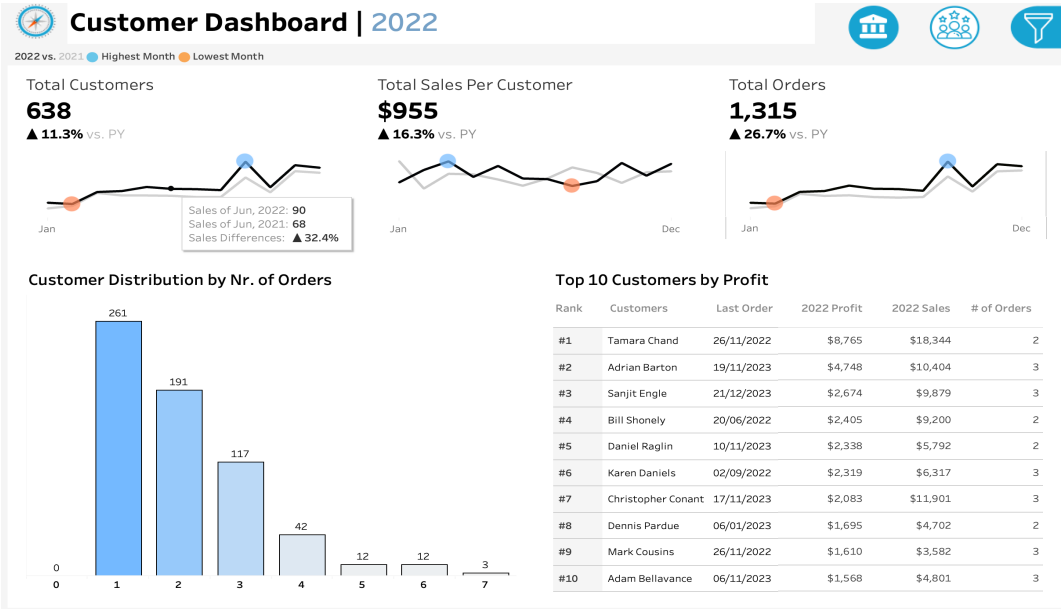

Figure 30: Customer Dashboard 2022

- 2022, has been overall the most successful year as it has seen significantly higher growth in sales (30%), profits (33%) and total quantity (23%) in comparison to 2021.
- Number of customers, total sales per customer and total orders have also increased significantly by 11%, 16% and 27% respectively comparing to 2021.

- January and December have been months for lowest and highest sales respectively, but did not report any loss.
- During the month of lowest profit, i.e., August, there have been less sales in many subcategories comparing to previous year which had lead to overall less profit.

  Reason: In all regions (except South), there has been a decrease in sales of different subcategories (except some states like New York, California etc) which lead to overall decrease in profits in August.
- Top 10 Customers in 2022, are mostly buying one of subcategories comprising Binders, copiers and Machines. These customers seem to be buying these for their own businesses as their last order was also in the most recent year, i.e., 2023. Similar is true for top 10 profitable customers of 2021.
- Customer distribution for number of orders follows a right skewed distribution with majority of customers having only 1 order while very few having more than 4 orders.
- California and New York are top states making most overall sales in United States.

Year: 2023

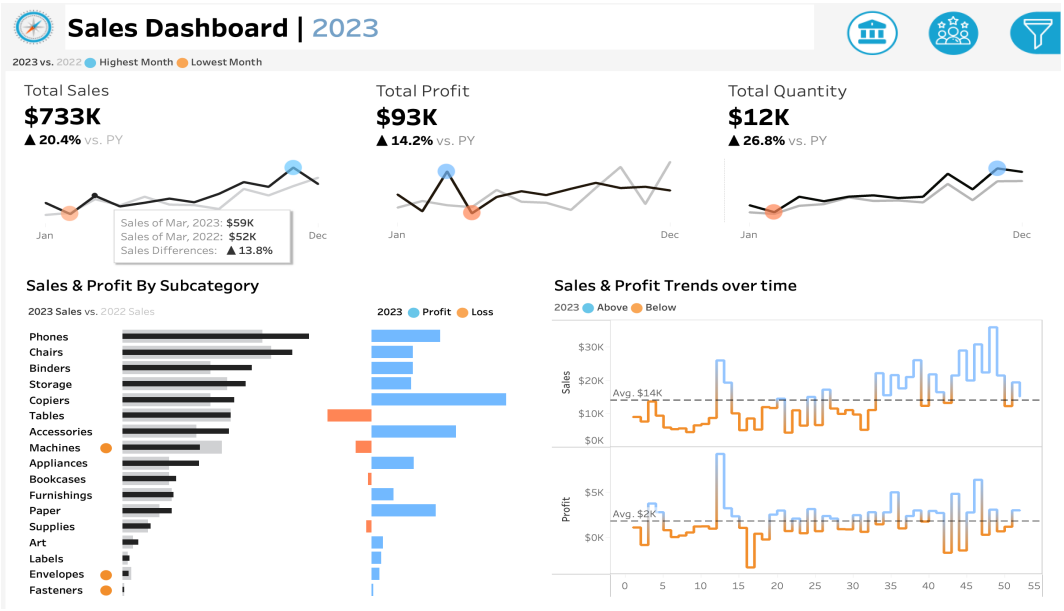

Figure 31: Sales Dashboard 2023

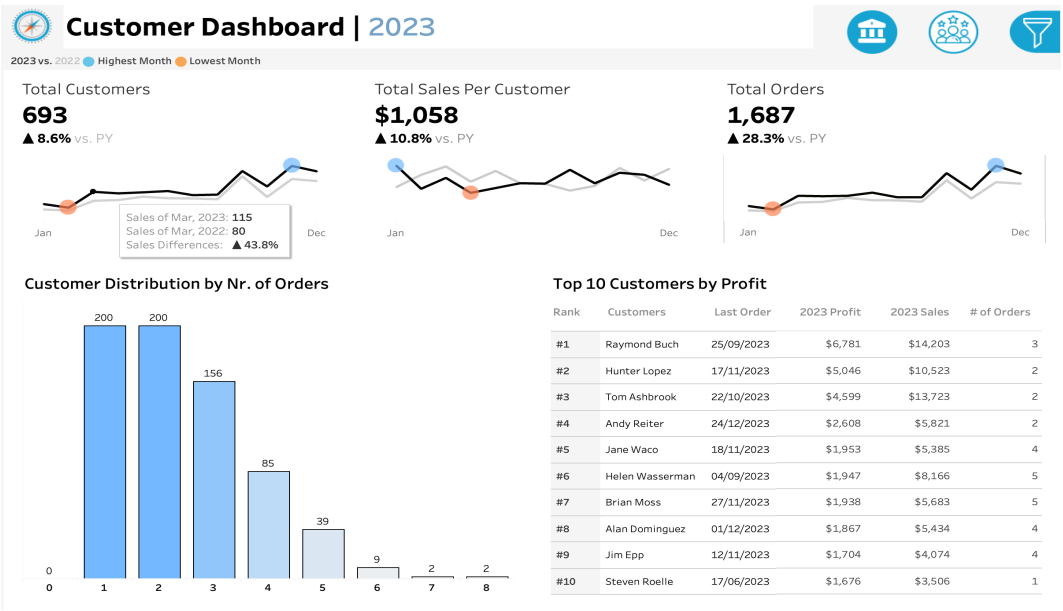

Figure 32: Customer Dashboard 2023

- 2023, has increased growth in sales (20%), profits (14%) and total quantity (27%) in comparison to 2021.
- Number of customers, total sales per customer and total orders have also increased significantly by 9%, 11% and 28% respectively comparing to 2022.
- February and November have been months for lowest and highest sales respectively, but February has been month with a loss of 68% in profits.

  Reason: Machines sales decreased by 98% comparing to 2022, as Machines were a major contributor to overall profits.
- During the month of lowest profit, i.e., August, there have been less sales in many subcategories comparing to previous year which had lead to overall less profit.

  Reason: In all regions (except South), there has been a decrease in sales of different subcategories (except some states like New York, California etc) which lead to overall decrease in profits in August.
- Top 10 Customers in 2023, are mostly buying one of subcategories comprising Copiers, Storage and Binders. Surprisingly, only 1 of top 10 customers bought machines.
- California, New York and Washington are top 3 states in terms of sales and profits.

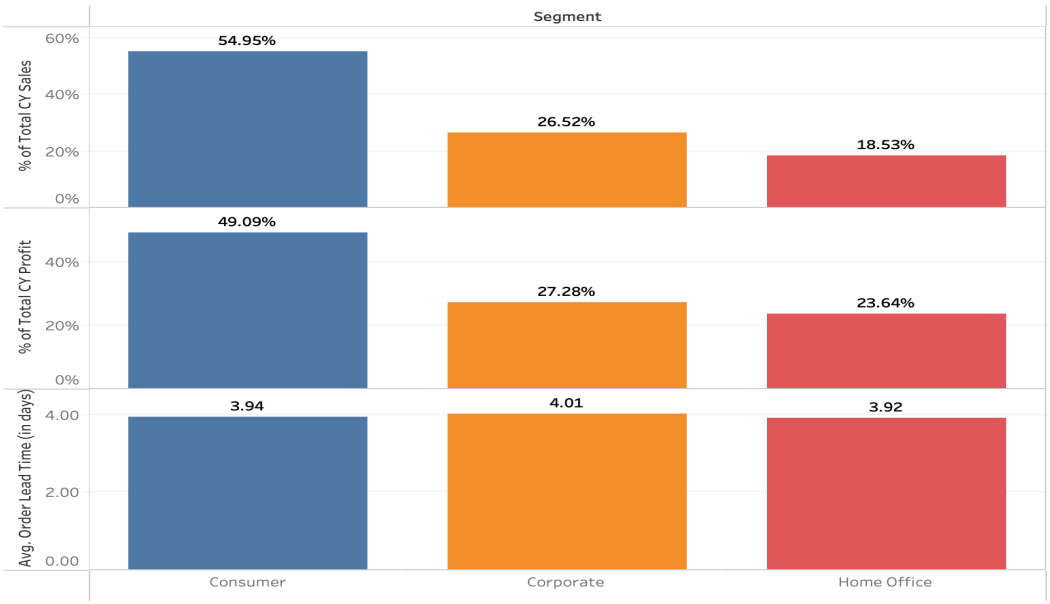

Figure 33: Segment vs CY Sales, CY Profits, Avg Order Lead Time

- Over years, Consumer Segment (55% sales, 49% profits) has been the biggest in terms of sales and profits followed by Corporate (27% sales, 27% profits) and Home Office (19% sales, 24% profits) (for 2020).
- Average Order Lead Time (difference between order date and ship day) is approximately same across all segments, i.e., 4 days.

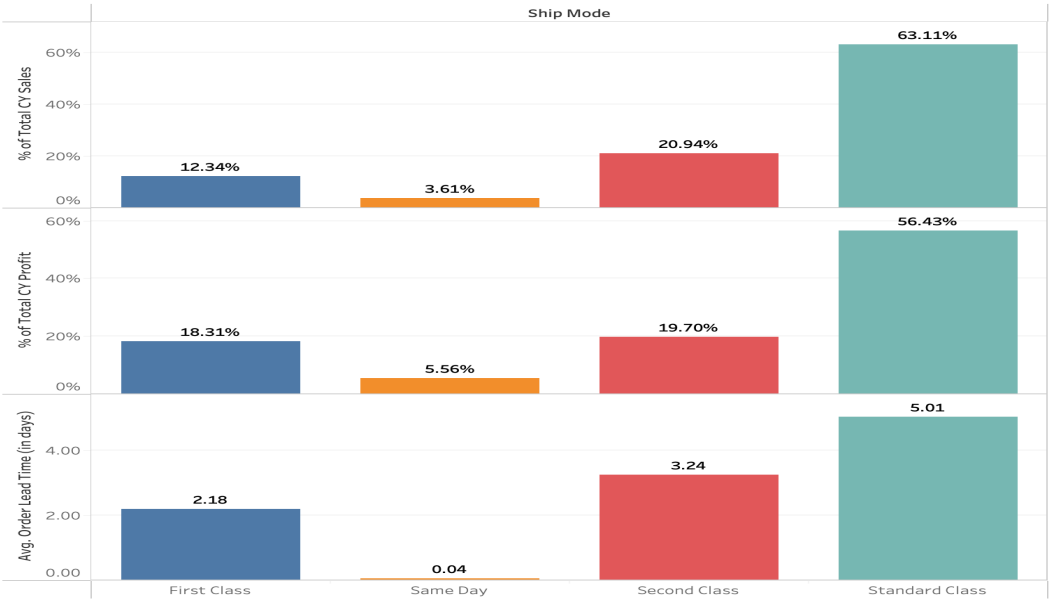

Figure 34: Ship Mode vs CY Sales, CY Profits, Avg Order Lead Time

- Over years, Standard Class (63% sales, 56% profits) ship mode followed by Second Class

(21% sales, 20% profits), First Class (12% sales, 18% profits) and Same day (4% sales, 6% profits), have been ranked 1, 2, 3 and 4 in terms Sales and profits (for 2020).

- Average Order Lead Time (difference between order date and ship day) is different across all segments, i.e., Standard (5 days), Second (3 days), First (2 days) and Same (0 days), has been same over years.

REGIONAL ANALYSIS

Central Region

- Central Region, Texas state has highest sales but still is not profitable in many categories for all years.
- Central Region, Michigan has increased its profits over years, but its majority profit comes from Binders sale.
- Central Region, Indiana has increased its sales as well as profits over year, but major reason for not enough growth seems to be price. Superstore has been selling items at a lower price which is eventually leading to losses. Price modulations with careful discounting can be promoted to ensure more sales as well as profits.

East Region

- New York state has been highest sales as well as profitable state over the years, not once showing any overall losses.
- All other states in this region are consistently performing same over the years not showing good growths.

South Region

- Southern states have been showing great sales over years, but not so good with profit. Further inspections can be performed by Superstore's team to understand deeper reasons for higher sales but no profit.

West Region

- West region has highest number of loss making subcategories such as Chairs, Tables, Furnishing, Appliances, Art, Machines etc.
- California and Washington have been making highest sales and overall most profitable for this region.
- West region has overall highest sales and profits with minimal losses over years.
- Superstore can look deeper into understanding why sales and profits have been on a positive side for this region.