# Superstore Analysis using Tableau

**By: Prerana Varsha Ravindra**

## Introduction:

Online shopping has grown in popularity over the years, mainly because people find it convenient and easy to bargain shop from the comfort of their home or office. One of the most enticing factor about online shopping, particularly during a holiday season, is it alleviates the need to wait in long lines or search from store to store for an item. The online sector, referred to as “clicks,” has been slowly eating up market share in the past two decades. Hence, I decided to do analysis on this domain.

**Inspiration:**

We can analyze the data further in 3 different ways

- PRODUCT LEVEL ANALYSIS
- CUSTOMER LEVEL ANALYSIS
- ORDER LEVEL ANALYSIS

So, what questions can be answered looking at this data?

- How much profit is gained for each product over which regions?
- What part of the avg discount% and profit ratio does each product play?
- Which segments contribute the most?

**Dataset:**

The dataset consists of the data from the previous four years (2015-2018) focusing only on the USA. The dataset has 9995 rows and 12 columns. The dataset was considered from Kaggle website. The dataset had no null values and no duplicate values. Hence it was good enough to start off with analysis

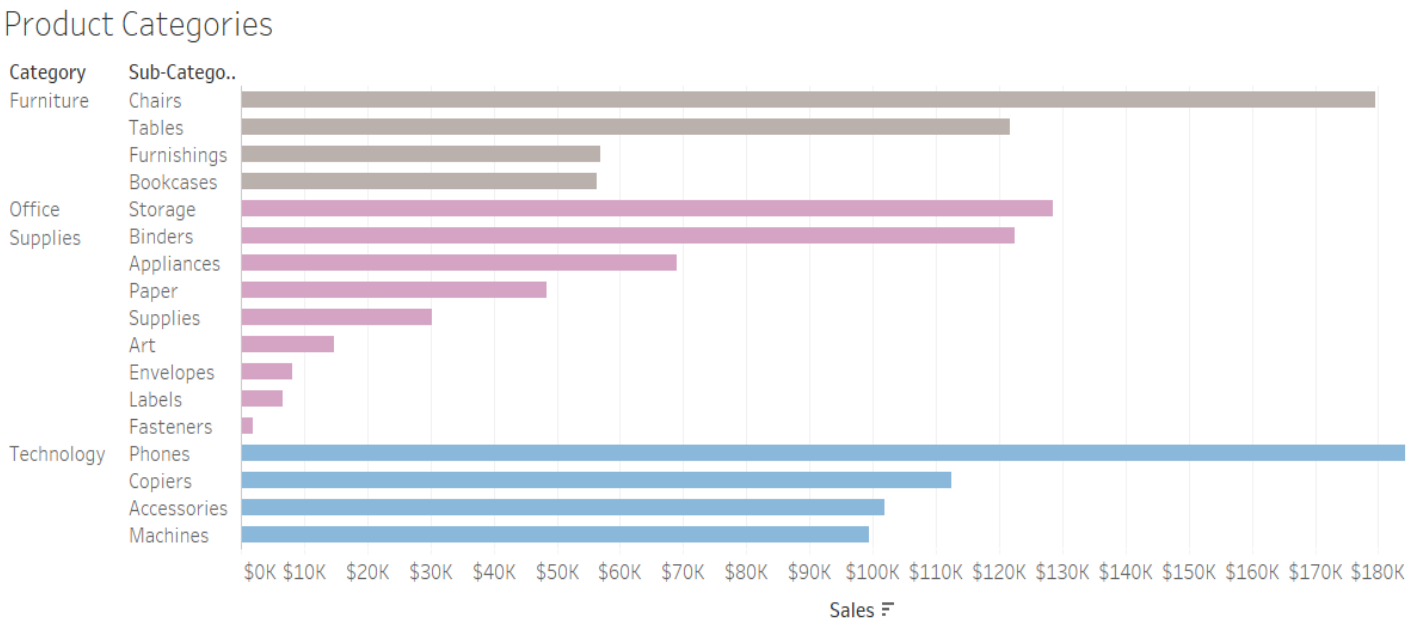
The columns include:

**Category**- Furniture/ Office Supplies/ Technology
**City**-Cities in the united states
**Country**-United states
**Customer Name**
**Discount**
**Number of Records**
**Order Date**
**Order ID**
**Postal Code**
**Manufacturer**
**Product Name**
**Profit**(in $)
**Quantity**
**Region**
**Sales**(in $)
**Segment-**consumer/Home office/corporate
**Ship Date**
**Ship Mode**-Fisrt class/second class/standard class/same day
**State**
**Sub-Category**

## Data Overview

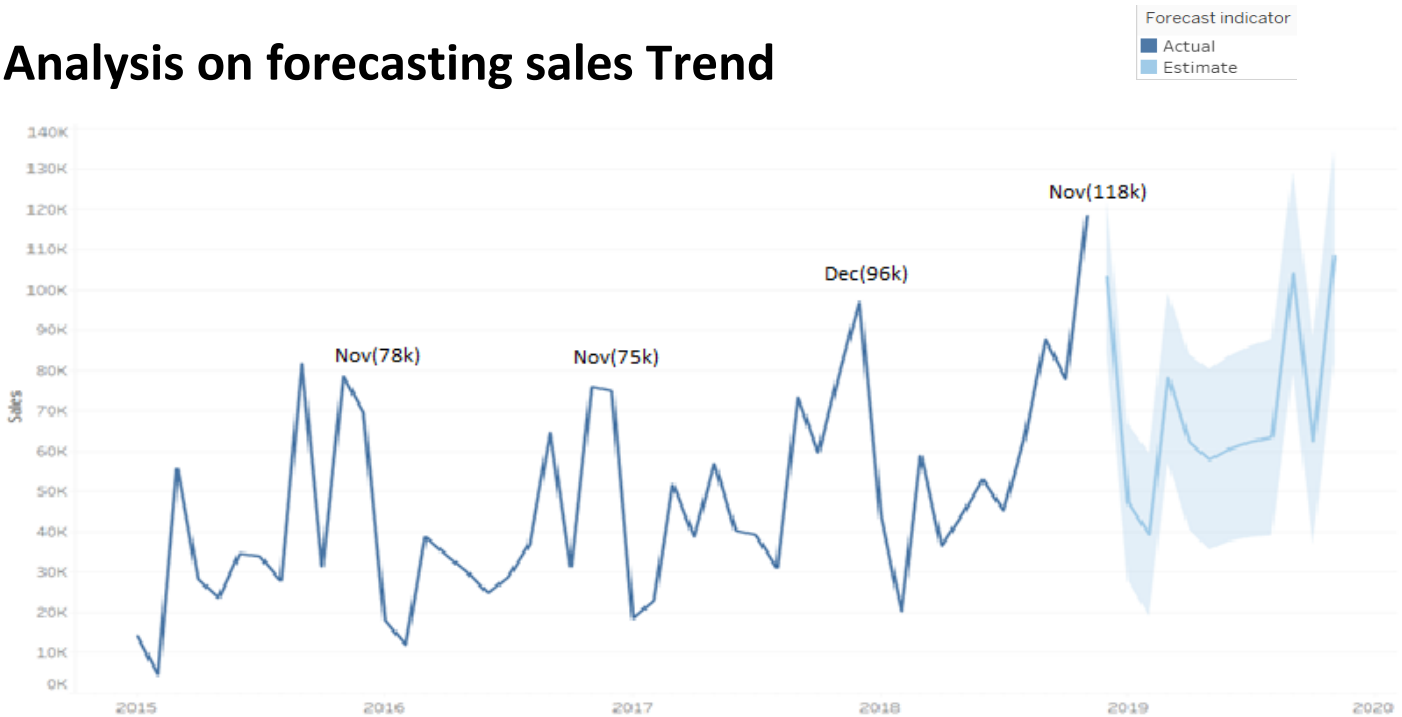
| Category | City | Country | Customer Na... | Discount | Order Date | Order ID | Postal Code | Manufactu... | Product Name | Profit | Qu... | Re... | Sa... | Segment |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Office Sup... | Columbus | United States | Chuck Clark | 0.000000 | 12/30/2018 | CA-2018-126221 | 47201 | Eureka | Eureka The Boss ... | 56.51 | 2 | Cent... | 209.30 | Home Office |
| Office Sup... | New York ... | United States | Patrick O'Donnell | 0.200000 | 12/30/2018 | CA-2018-143259 | 10009 | Wilson Jones | Wilson Jones Leg... | 19.79 | 3 | East | 52.78 | Consumer |
| Furniture | New York ... | United States | Patrick O'Donnell | 0.200000 | 12/30/2018 | CA-2018-143259 | 10009 | Other | Bush Westfield C... | 12.12 | 4 | East | 323.14 | Consumer |
| Technology | New York ... | United States | Patrick O'Donnell | 0.000000 | 12/30/2018 | CA-2018-143259 | 10009 | Other | Gear Head AU37... | 2.73 | 7 | East | 90.93 | Consumer |
| Office Sup... | New York ... | United States | Jennifer Ferguson | 0.000000 | 12/28/2018 | CA-2018-164826 | 10024 | Avery | Avery 473 | 34.78 | 7 | East | 72.45 | Consumer |
| Office Sup... | New York ... | United States | Jennifer Ferguson | 0.000000 | 12/28/2018 | CA-2018-164826 | 10024 | Other | OIC Bulk Pack Me... | 6.42 | 4 | East | 13.96 | Consumer |
| Furniture | Round Rock | United States | Greg Hansen | 0.320000 | 12/28/2018 | CA-2018-136539 | 78664 | Bush | Bush Westfield C... | -11.60 | 2 | Cent... | 78.85 | Consumer |
| Office Sup... | Round Rock | United States | Greg Hansen | 0.200000 | 12/28/2018 | CA-2018-136539 | 78664 | Stanley | Stanley Bostitch ... | 2.72 | 2 | Cent... | 27.17 | Consumer |
| Office Sup... | New York ... | United States | Jennifer Ferguson | 0.200000 | 12/28/2018 | CA-2018-164826 | 10024 | Storex | Storex Dura Pro ... | 11.23 | 7 | East | 33.26 | Consumer |

# Analysis based on product category



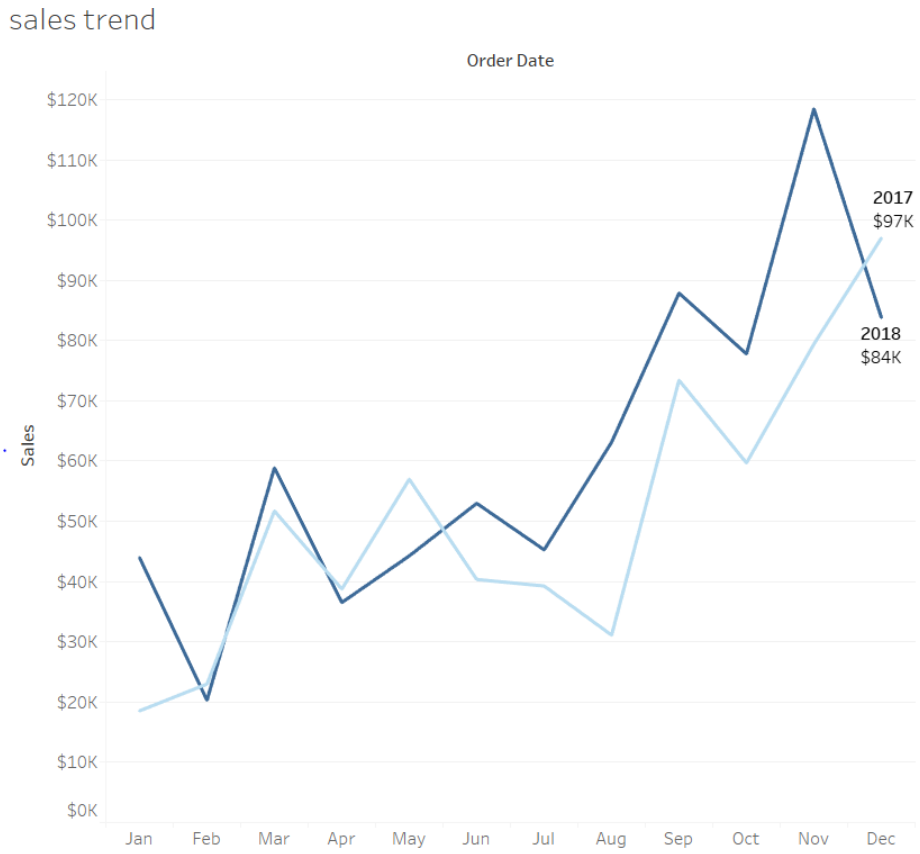

we can observe that the three categories of the products are indicated in different colors and as we see the maximum sales is in the phone under the technology category with more than $184k which is leading among all the products followed by chairs in the furniture category with $179k .Looking at this we can analyses which products should be brought with more demand into the superstore.

# Analysis on forecasting sales Trend




There we can observe that there is a trend pattern followed every year as the sales increases towards the end of each year compared to the initial month of the respective year.We also can observe that the months September, November and December usually have the highest sales. While during the month of october and Febraury we observe decrease in sales.Taking this into consideration we used the forcast indicator to predict the next year's sales and it shows us an estimate and  actual which means that the sales for the year 2019 would remain within the estimated threshold as indicated above.

# Analysis on overall sales Trend

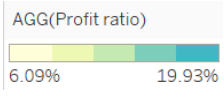



The overall sales compared between the two years 2017 and 2018 has decreased by the by $13k.

We can observe the sales over each month in comparison with the previous year to do an analysis on the trend pattern. The months march, September and November usually have an increase in the sales while the months February and October have decrease in sales.

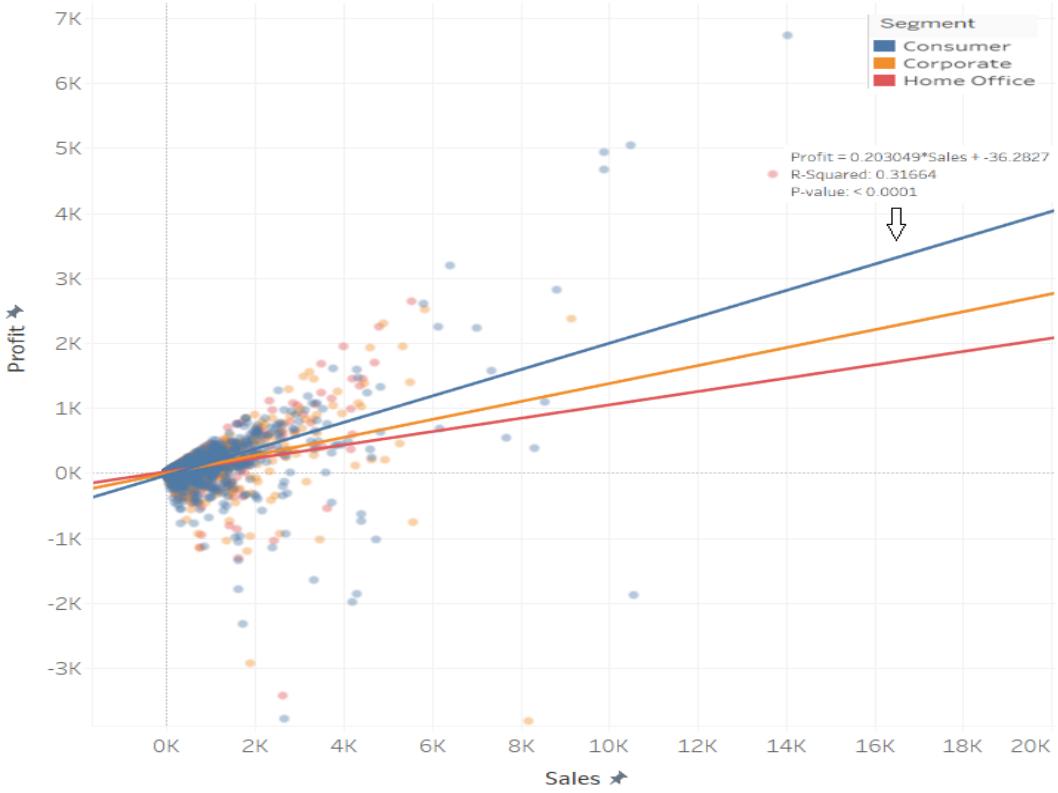
# Highlight Table

| Segment | Central | East | South | West |
|---|---|---|---|---|
| Consumer | 6.09% | 10.42% | 15.77% | 18.60% |
| Corporate | 12.23% | 13.92% | 8.13% | 14.66% |
| Home Office | 11.68% | 19.93% | 8.83% | 10.65% |



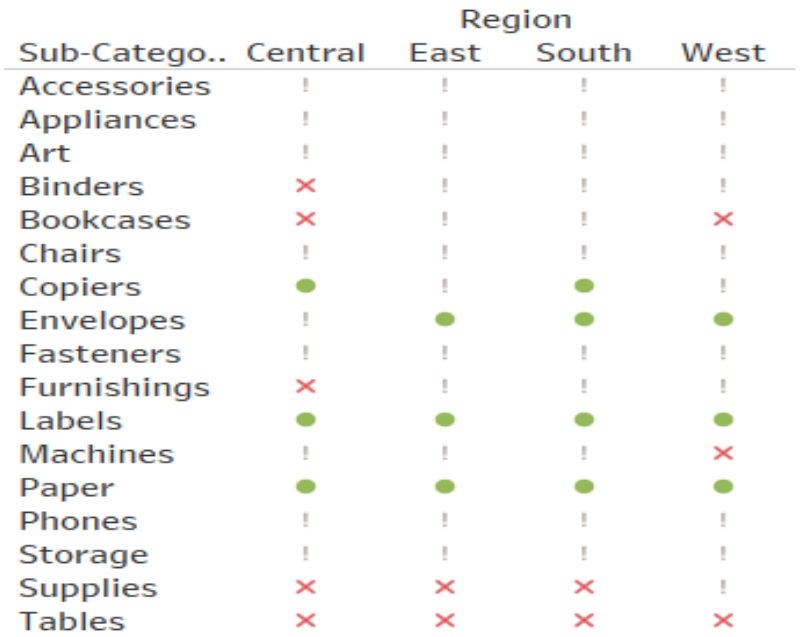

Considering only the positive Profit ratio of the supermarket against the segments.Profit ratio ranges from 6% to 19.93%. The highest profit ratio has been from the East Region contributed by the Home Office followed by the consumer in the west region.While the least profits are from the consumer in the central region.

# Customer Analysis




In this basic scatter plot, we analyze the correlation between sales and Profit in consideration with the segments. We simulate the relationship by the linear model. From the statistical variables provided by Tableau, we can see that P-value is less than 0.001 and R-Squared is 0.316. This indicates their linear correlation is good enough. We can also clearly distinguish the outliers and further analyze the detail information of them. Blue the Consumer segment has the strongest relationship with the sales and profit which means that more sales from this segment brings out more profit. Steeper the line better the relationship between the sales and profit entities. The Home Office has a weaker relationship compared to the other segments.

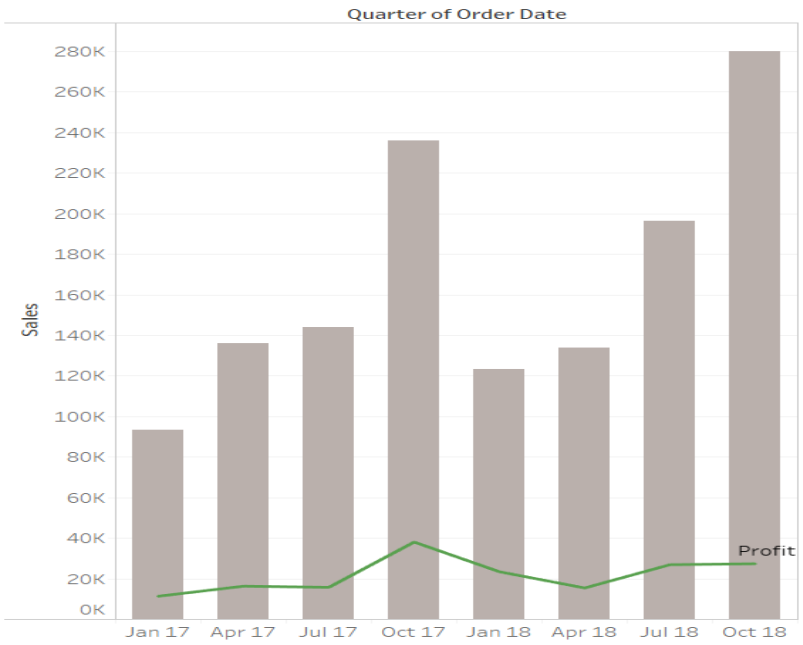
# KPI's of Profit ratio




```
IF[Profit ratio]>= 0.4 THEN "GOOD"
ELSEIF [Profit ratio]>= 0 THEN "OKAY"
ELSE "BAD"
END
```

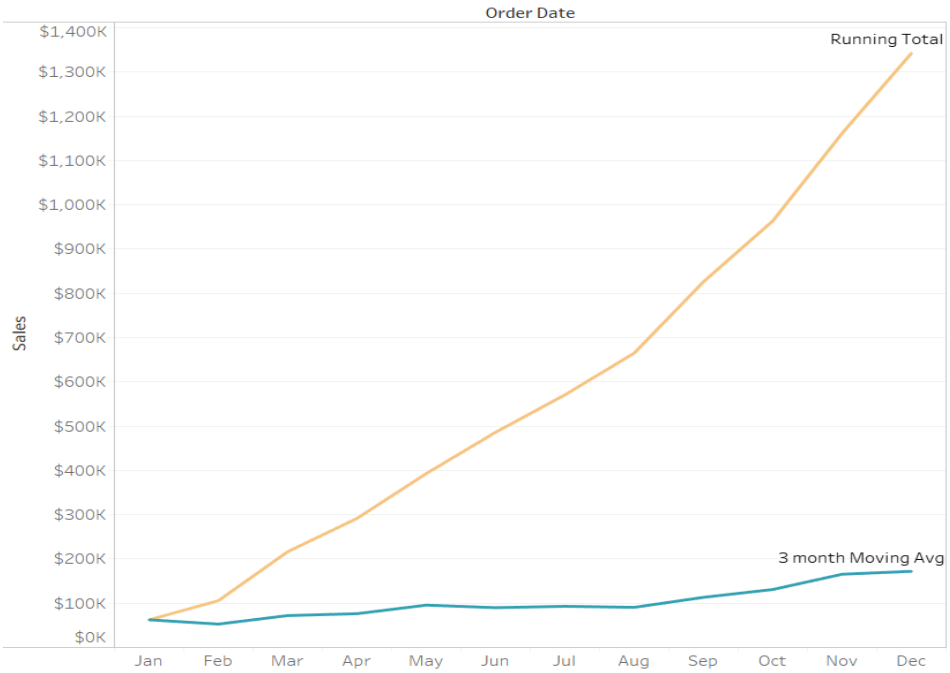

I decided to analyze each subcategory with the region by creating a KPI. The Good, okay and bad are represented with their respective symbols. The condition I used to represent them is as shown above. By this chart we can know what categories must be under more consideration to improve the profit margins of the superstore

# Combo axis and mark:

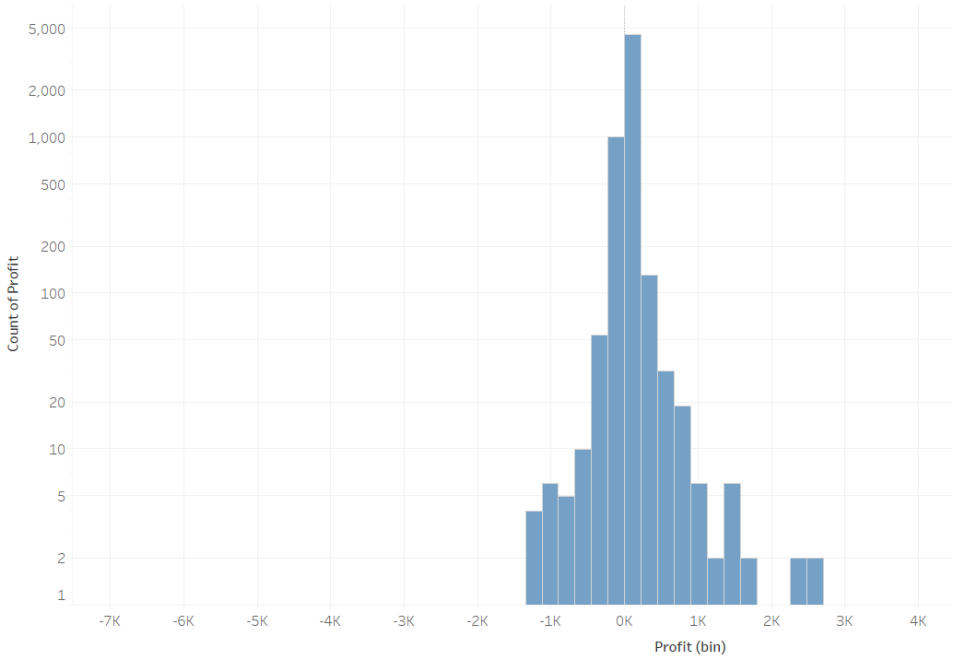



This is a dual axis of sales and profit and it is a visualization based on the quarterly period over the two years .From this analysis we can conclude that the last quarter of the years usually have increase in both sales and profit. There is also a pattern observed in the quarters of each year, every quarter the sales seem to be increasing while the profit varies.

# Running Total

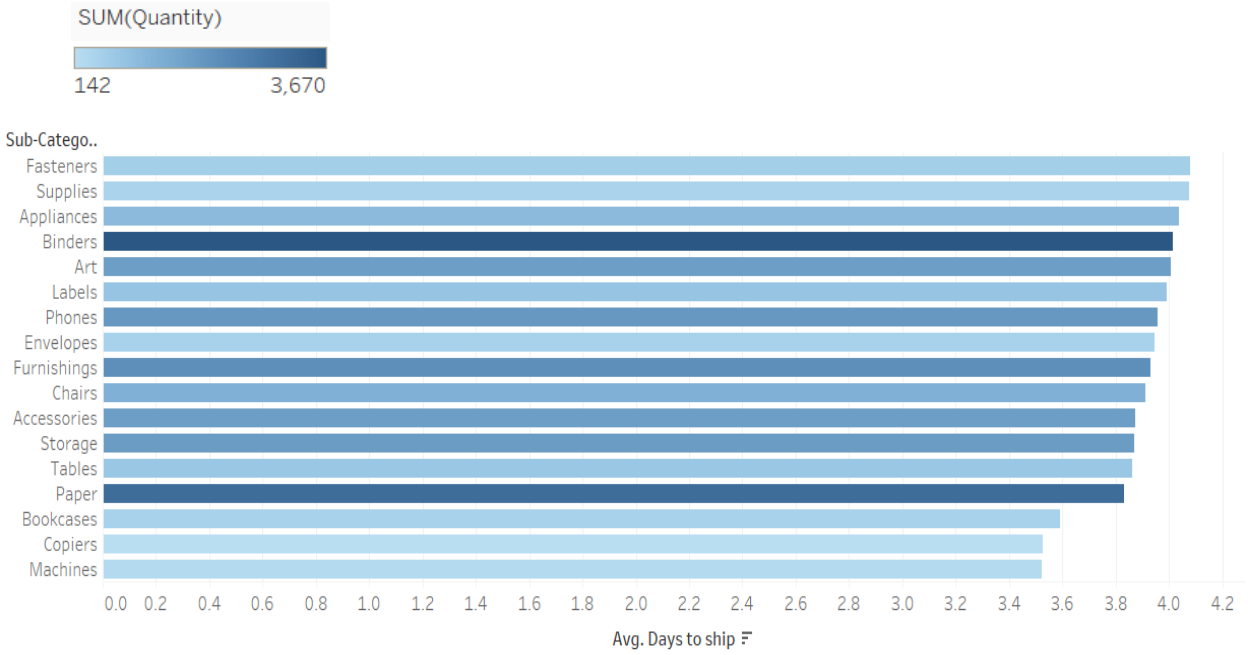


# Distribution of Profit:



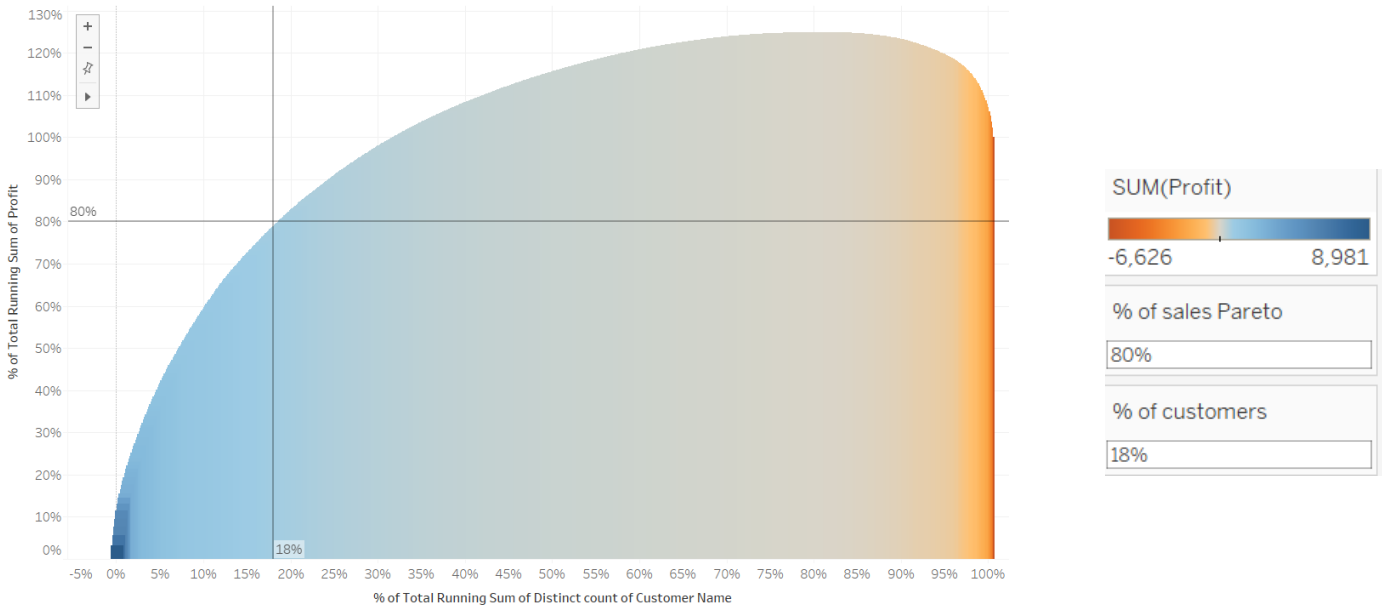

The Distribution of the profit can be analyzed by using bins of prof it with their count. Here I have used a logarithmic function the count of profit to show their distribution

# Average shipping days



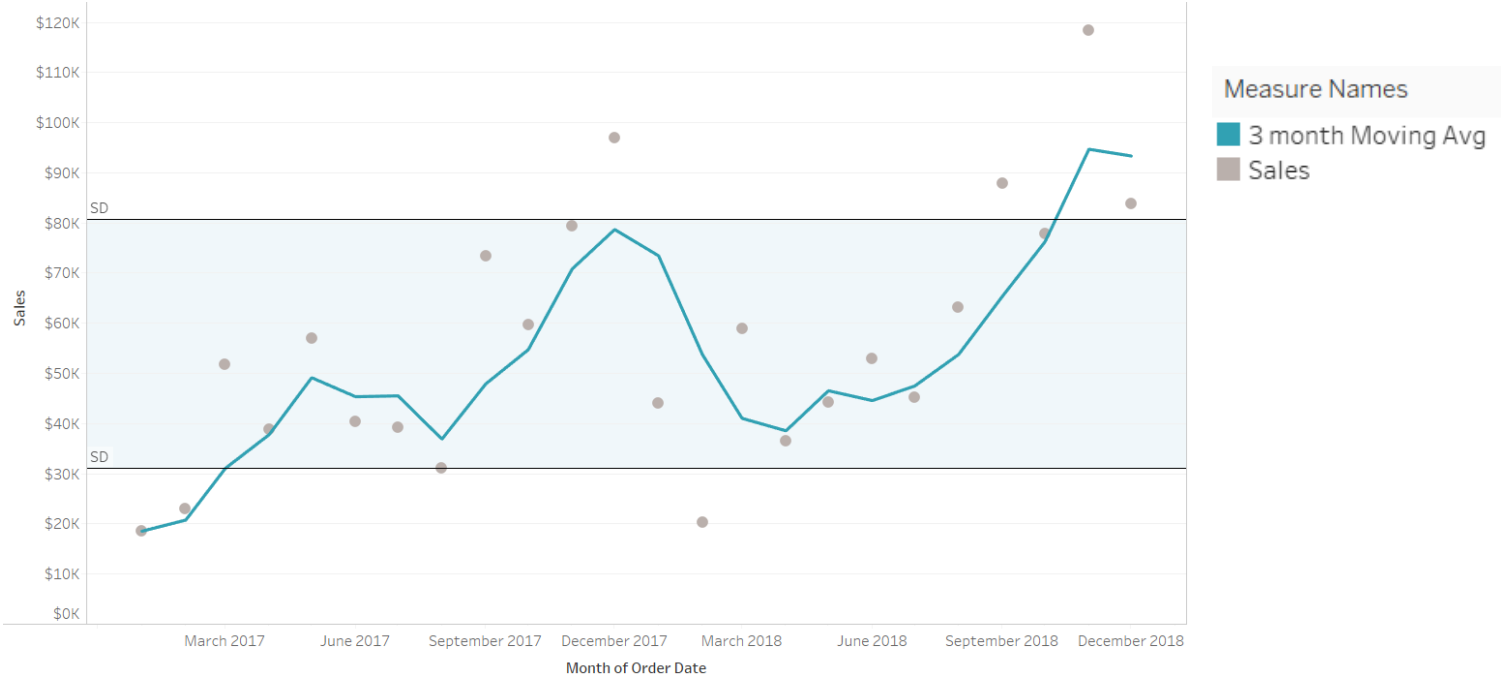

we can see that the Binders have the maximum order quantity with an average shipping day of 4.0153 days. While copiers and Machines have least quantity order with less average shipping days. The calculation for Avg shipping days was done as :DATEDIFF('day',[Order Date],[Ship Date])

# Pareto




According to Pareto's rule 20% of the customers bring in 80% of profit but according to our data 18% of customers bring in 80% of the profit. So, if we want to do some promotions we have to be focusing on those customers.

# Statistics in visualization



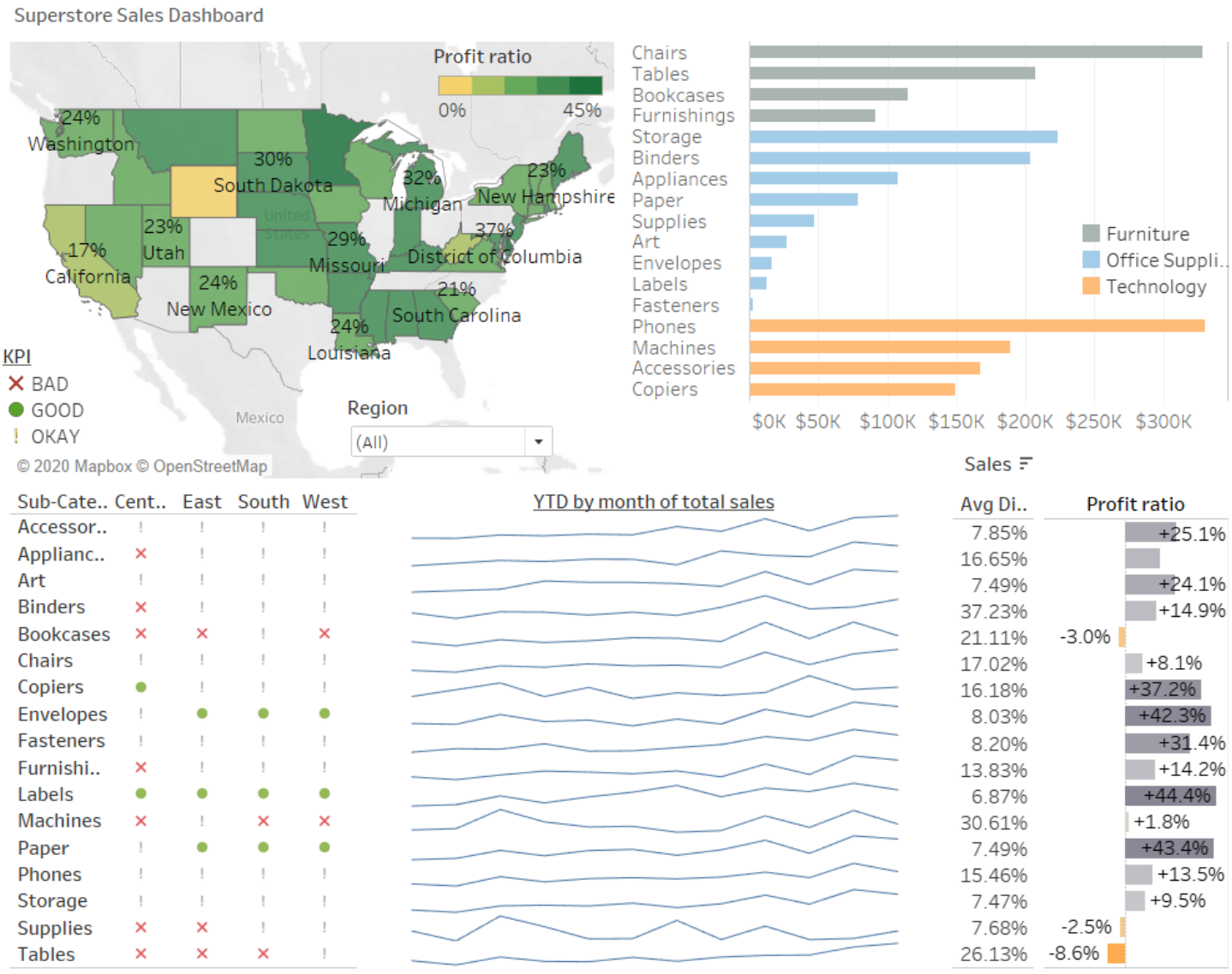

This shows the 3month moving avg of sales with actual sales, 1 standard deviation of the population drawn anything outside these are considered as outliers and we can observe how they pull the trend above as well. This plot uses stats to give the plot more context usually used in answering business related questions.

***Now let us look at the Dashboard for further better understanding.***

***Overview of Dashboard***

***Link to view the Dashboard:*** https://public.tableau.com/profile/prerana.varsha.ravindra#!/

This is just the overview of how my dashboard looks, it is an interactive one. We shall see the interactive part using the Tableau.




----------Thank you 😊 ----------