



Superstore Data Analysis & Objective Setting

Group 2

  Q Team Members 

Group 2



**Kamalika
Das**



**Keerthana
Nemili**



**Manoj
Padmaraju**



**Payal
Pandya**

← → Q Superstore Sales Data

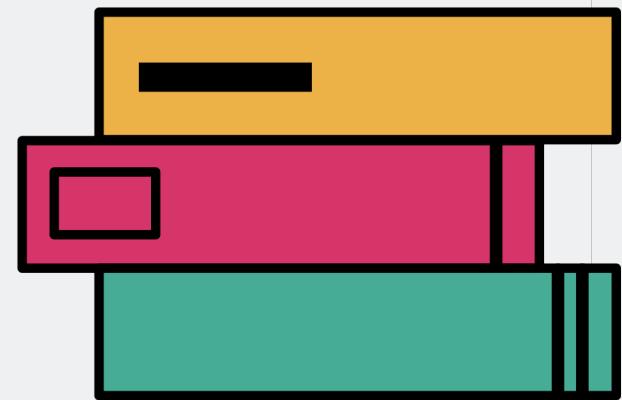
Introduction

Assess the factors driving the profits/loses across various categories and segments pertaining to a retail superstore located across US

First phase will provide an overview of Explanatory analysis and interaction between variables. Phase two involves building a state level goal setting engine for the superstore for each category and segment of its market



Objective



Objective 1

Analyze the Impact of shipping method, segment, categories, discount on sales quantity, and profits in order to determine the best targeting and marketing strategy

Objective 2

Devise a non-predictive methodology to set yearly sales objectives for each segment category for the upcoming calendar year based on historical average growth rate.

Overview of Data Set:

Categorical

- Segment
- State
- Category
- City
- Region
- Subcategory
- Product ID
- Product Name
- Customer ID
- Customer Name
- Ship Mode
- Order Date (MM/DD/YY)

Numerical

- Sales
- Quantity
- Discount
- Profit
- Population

Engineered

- Unit Cost Price
- Gross Sales
- Profit Margin
- Year Order Month

← → Q

Descriptive Statistics

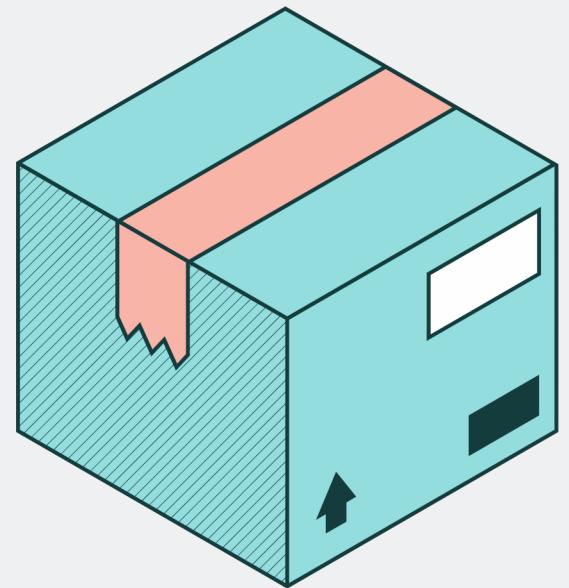
Profit Margin

75% of orders make atleast 36% of the profit margin

Loss incurred

Highest loss on orders goes all the way to -275% that need to be cut off.

| | Postal Code | Sales | Quantity | Discount | Profit | Population | Order year | Profit Margin % | Gross Sales | Unit Cost Price |
|--------------|-------------|----------|----------|----------|----------|-------------|------------|-----------------|-------------|-----------------|
| count | 9993.00 | 9993.00 | 9993.00 | 9993.00 | 9993.00 | 9993.00 | 9993.00 | 9993.00 | 9993.00 | 9993.00 |
| mean | 55191.58 | 229.85 | 3.79 | 0.16 | 28.66 | 14175858.39 | 2015.72 | 12.03 | 286.55 | 73.64 |
| std | 32065.07 | 623.28 | 2.23 | 0.21 | 234.27 | 9880420.11 | 1.12 | 46.68 | 864.14 | 228.35 |
| min | 1040.00 | 0.44 | 1.00 | 0.00 | -6599.98 | 445025.00 | 2014.00 | -275.00 | 0.99 | 0.54 |
| 25% | 23223.00 | 17.28 | 2.00 | 0.00 | 1.73 | 5951832.00 | 2015.00 | 7.50 | 21.36 | 4.37 |
| 50% | 56560.00 | 54.48 | 3.00 | 0.20 | 8.67 | 10167376.00 | 2016.00 | 27.00 | 64.96 | 16.23 |
| 75% | 90008.00 | 209.94 | 5.00 | 0.20 | 29.36 | 21596071.00 | 2017.00 | 36.25 | 251.91 | 69.99 |
| max | 99301.00 | 22638.48 | 14.00 | 0.80 | 8399.98 | 30617582.00 | 2017.00 | 50.00 | 45276.96 | 8149.85 |



Product Demand

The store carries 3 category of items with 17 sub-category. 60% of orders are for Office Supplies and at least 15% purchases are for Binders.

```
sales_data.describe(include='object')
```

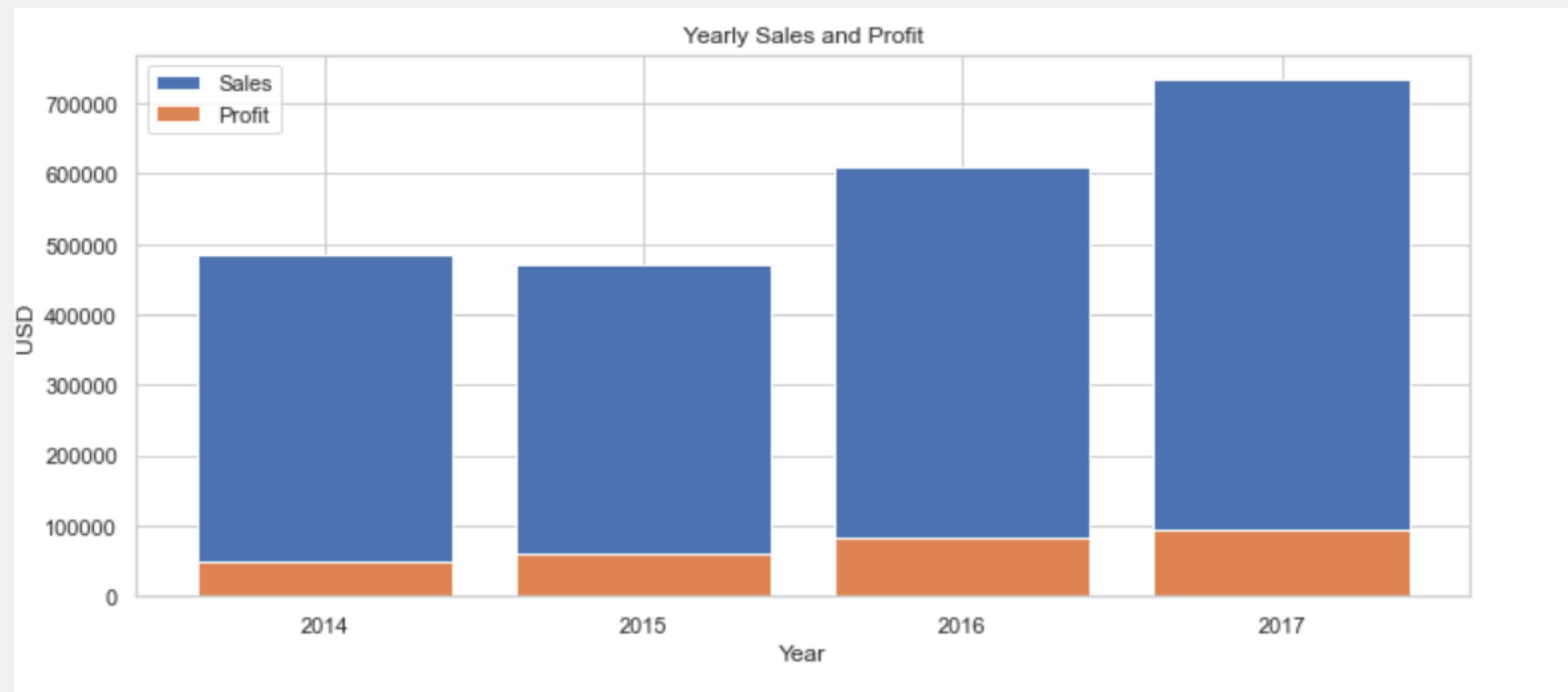
| | Order Date | Ship Date | Ship Mode | Customer ID | Customer Name | Segment | City | State | Region | Product ID | Category | Sub-Category | Product Name | Geographic Area | Order month |
|--------|------------|------------|----------------|-------------|---------------|----------|---------------|------------|--------|-----------------|-----------------|--------------|-----------------|-----------------|-------------|
| count | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 | 9993 |
| unique | 1237 | 1334 | 4 | 793 | 793 | 3 | 531 | 49 | 4 | 1862 | 3 | 17 | 1850 | 49 | 12 |
| top | 9/5/2016 | 12/16/2015 | Standard Class | WB-21850 | William Brown | Consumer | New York City | California | West | OFF-PA-10001970 | Office Supplies | Binders | Staple envelope | California | November |
| freq | 38 | 35 | 5967 | 37 | 37 | 5191 | 915 | 2001 | 3203 | 19 | 6026 | 1523 | 48 | 2001 | 1471 |

Shipping and Customer

There are 4 ship modes and 3 customer segments. Nearly half of the orders are from Consumer segment using Standard Class shipment

[Title Page](#)[Proprietors](#)[Introduction](#)[Review of Literature](#)[Methodology](#)[Results](#)[X](#) [+](#)[← → Q](#)

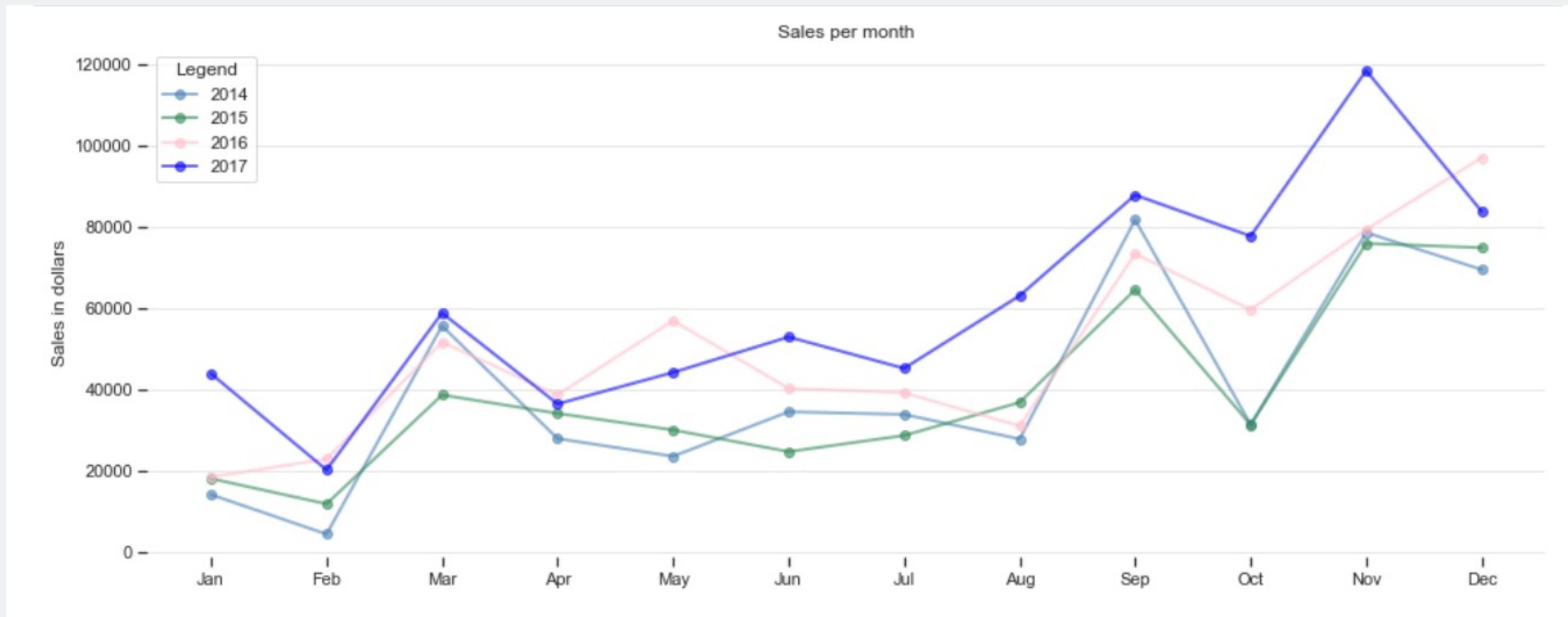
Yearly Sales and Profit (2014 - 2017)





← → Q

Yearly Sales and Profit (2014 - 2017)



Profit/Loss vs Sales

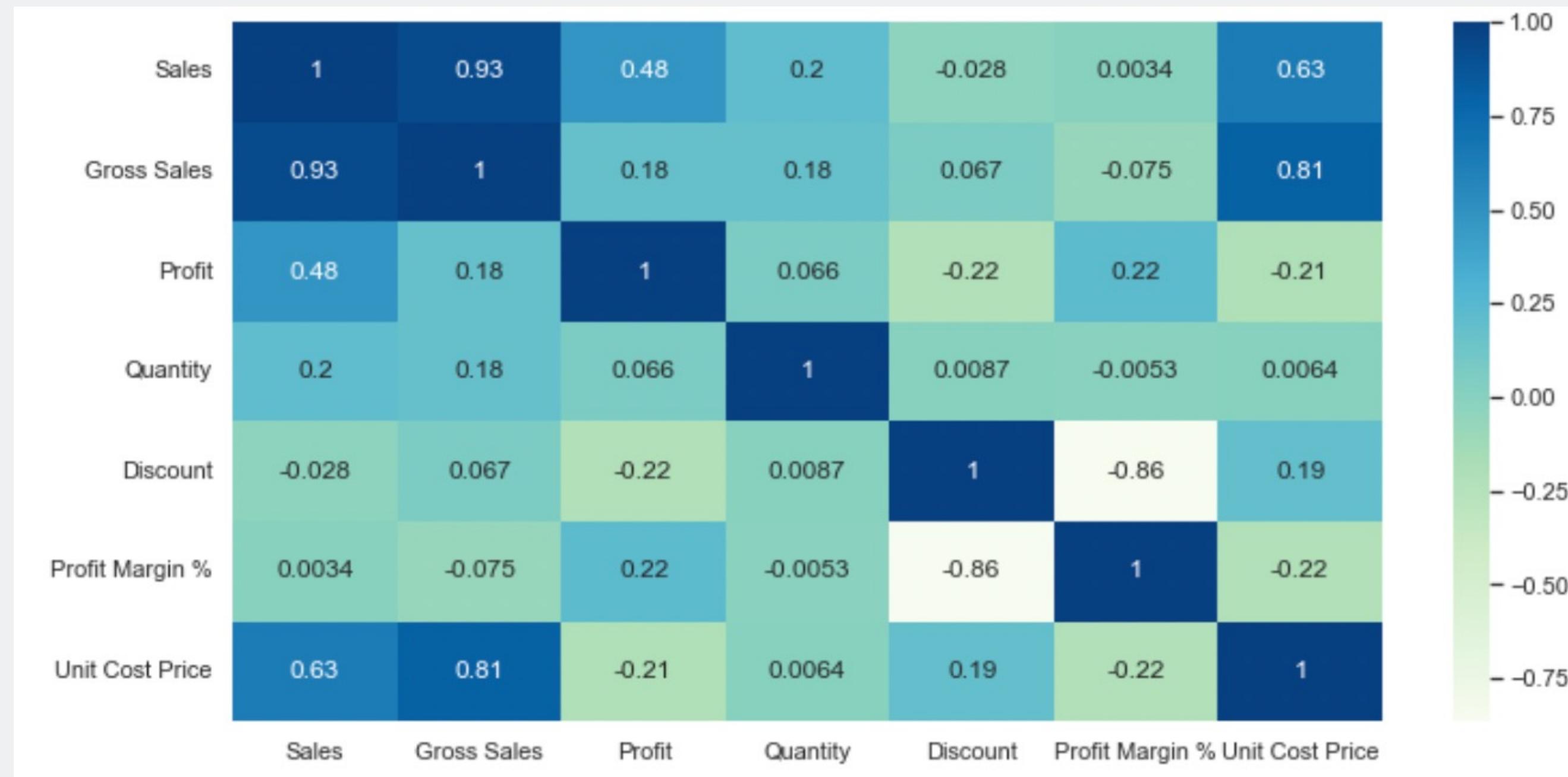


Most sales are less than USD 5k. This is likely because a majority of the Superstore's customers are of the Consumer Segment - individuals who tend to buy and spend less than the other customer Segments comprising of groups and teams from Corporates and Home Offices.

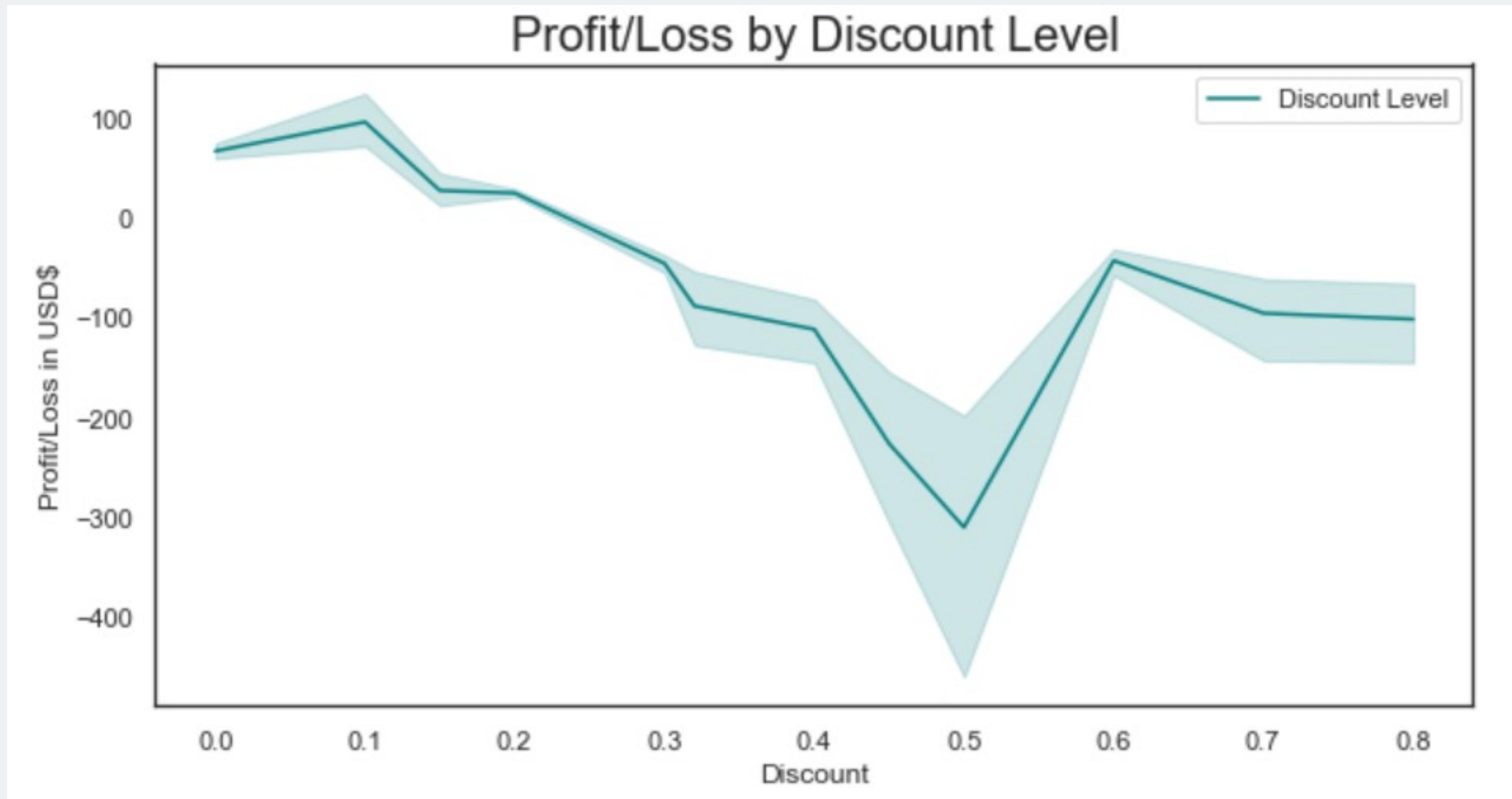
Larger sales above USD 2.5k are very likely to result in a profit. Profit margins may be higher, after economies of scale in cost components such as procurement, packaging, and delivery.

← → Q

CORRELATION MATRIX



Profit vs Discount Level



- ◆ Superstore is more profitable when discount levels are 10% or lower.
- ◆ At 50% discount, profit is the least.

← → G Q

Comparing discounted and non-discounted transactions

Considering both average and total sales, customers tend to spend more when there are discounts. An average sales transaction with discounted products is USD 233 as compared to a lower USD 227 where there are no discounts. In total, the difference amounts to a notable USD 200k increase in sales value when there are discounts

On Average :

Discounted Sales per transaction is USD\$ 232.73 , without is USD\$ 226.74

Discounted Unit Price is USD\$ 103.49 , without is USD\$ 41.33

Discounted Quantity purchased per transaction is 3.77 , without is 3.8

Discounted Profit per transaction is USD\$ -6.66 , without is USD\$ 66.9

Discounted Unit Profit is USD\$ -1.23 , without is USD\$ 17.58

In Total :

Discounted Total Sales is USD\$ 1209011.02 , without is USD\$ 1087908.47

Discounted Total Quantity is 19604 , without is 18267

Discounted Total Profit is USD\$ -34578.52 , without is USD\$ 320987.6

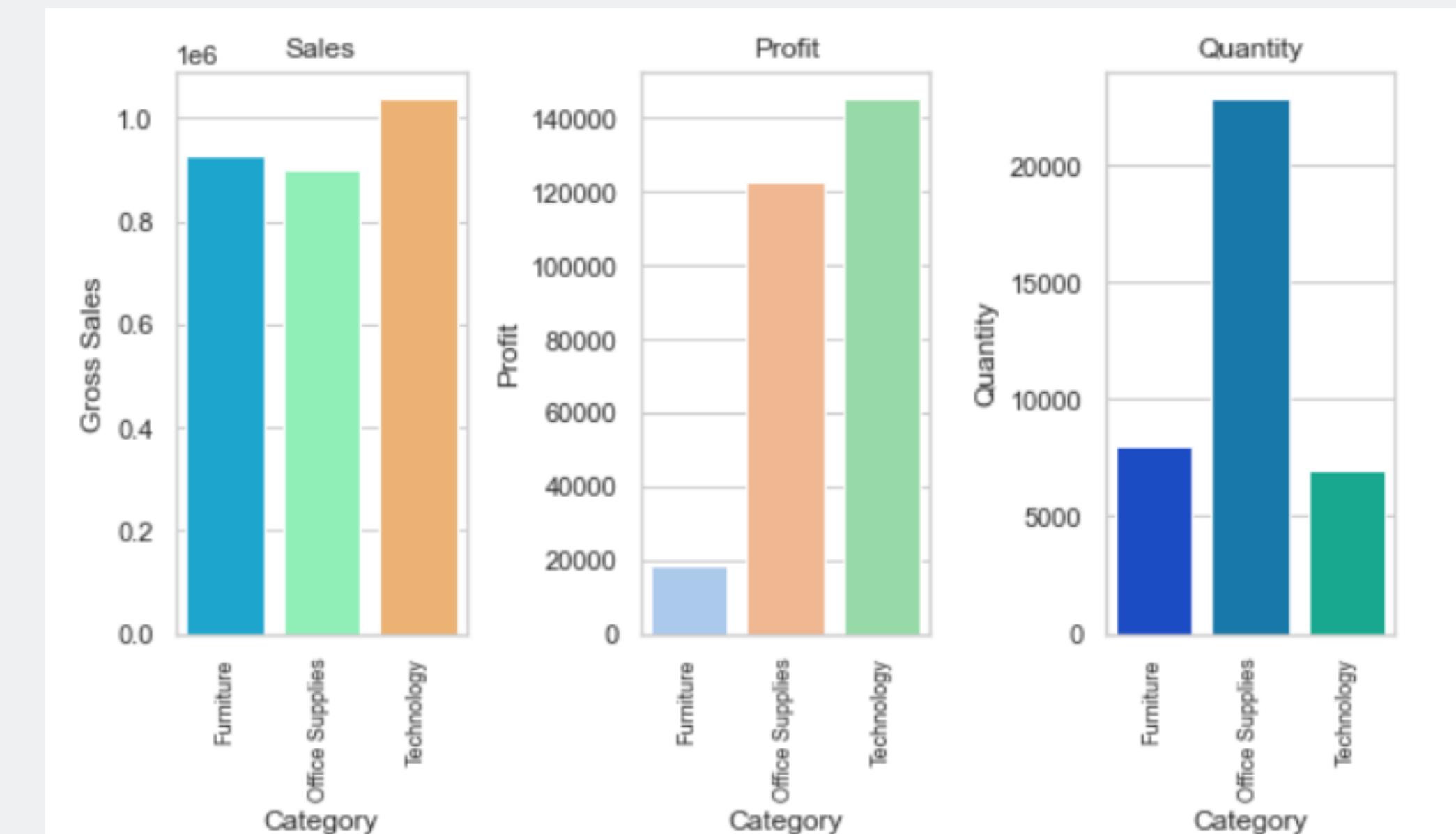
Superstore is advised to reduce overall discount levels to minimise losses, and focus on implementing other promotional strategies to ensure increasing profitable sales and consistent future gains in long term customers

← → Q

Which Category is Best Selling and Most Profitable?

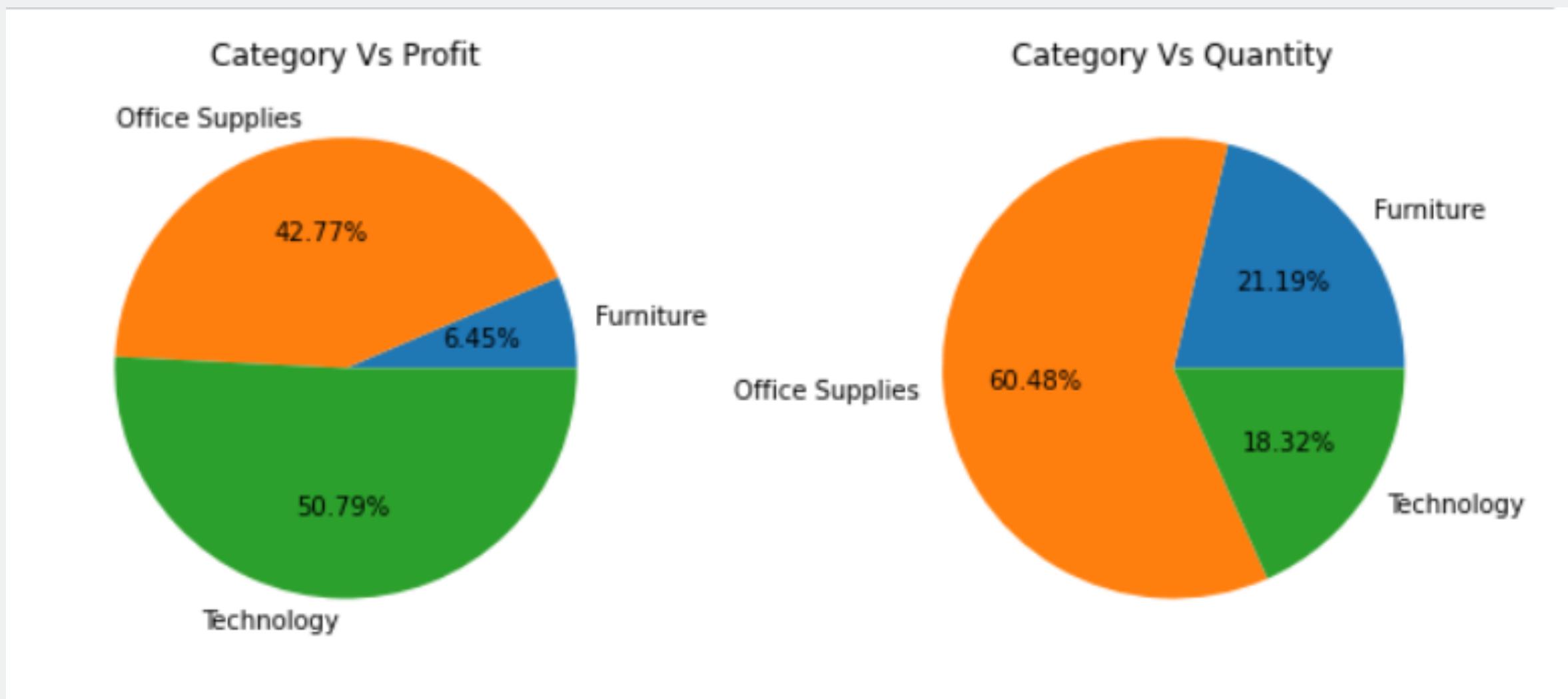
Product category-Technology shows the highest gross sales of \$103K and profit of \$145K.

The Unit price of the category-Technology to be highest is conclusive as it has the highest profit despite the least quantity sold.



Insights:

Profit and Quantity grouped by Category

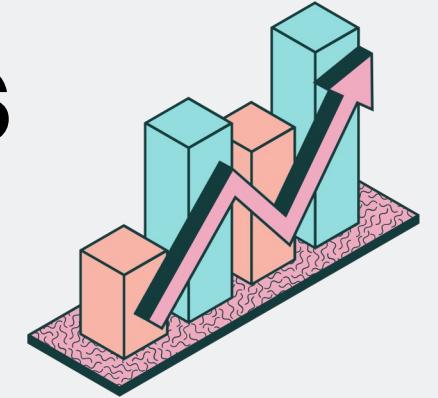


Although Furniture makes similar sales as Technology, it is the least profitable. Technology products are purchased more frequent and at higher price than furnitures.

Office Supplies sells the most in terms of quantity as it is relatively cheap product.



Subcategory Products Performances- Best vs Worst Selling, Most profitable:



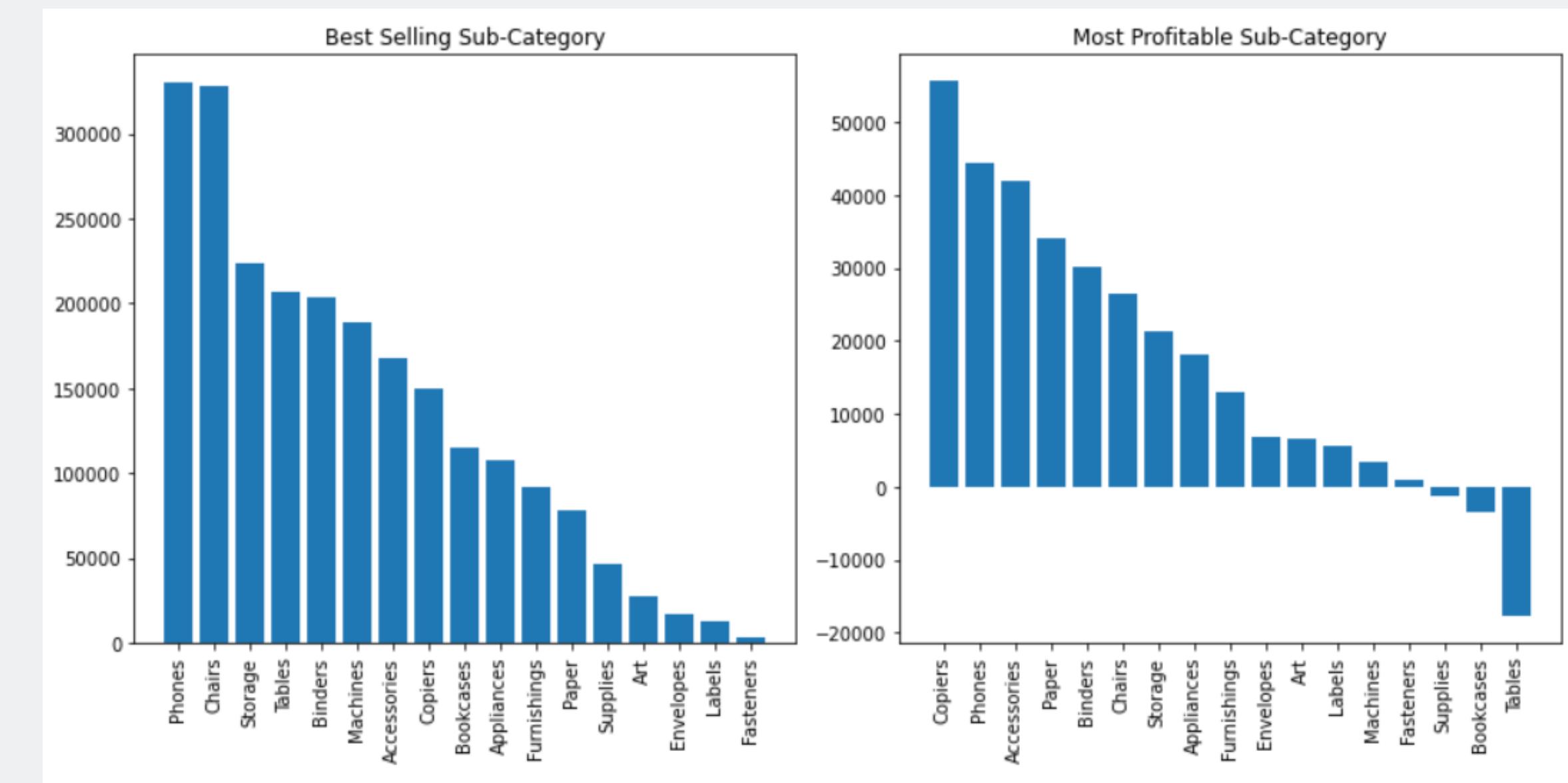
Front-runners

Copiers being a more intricate product sees the least bargain in price, followed by Phones.



Sink-holes

Super Store can consider dropping from the product catalogue of Tables, bookcases etc or bargain for a lower price from the supplier.

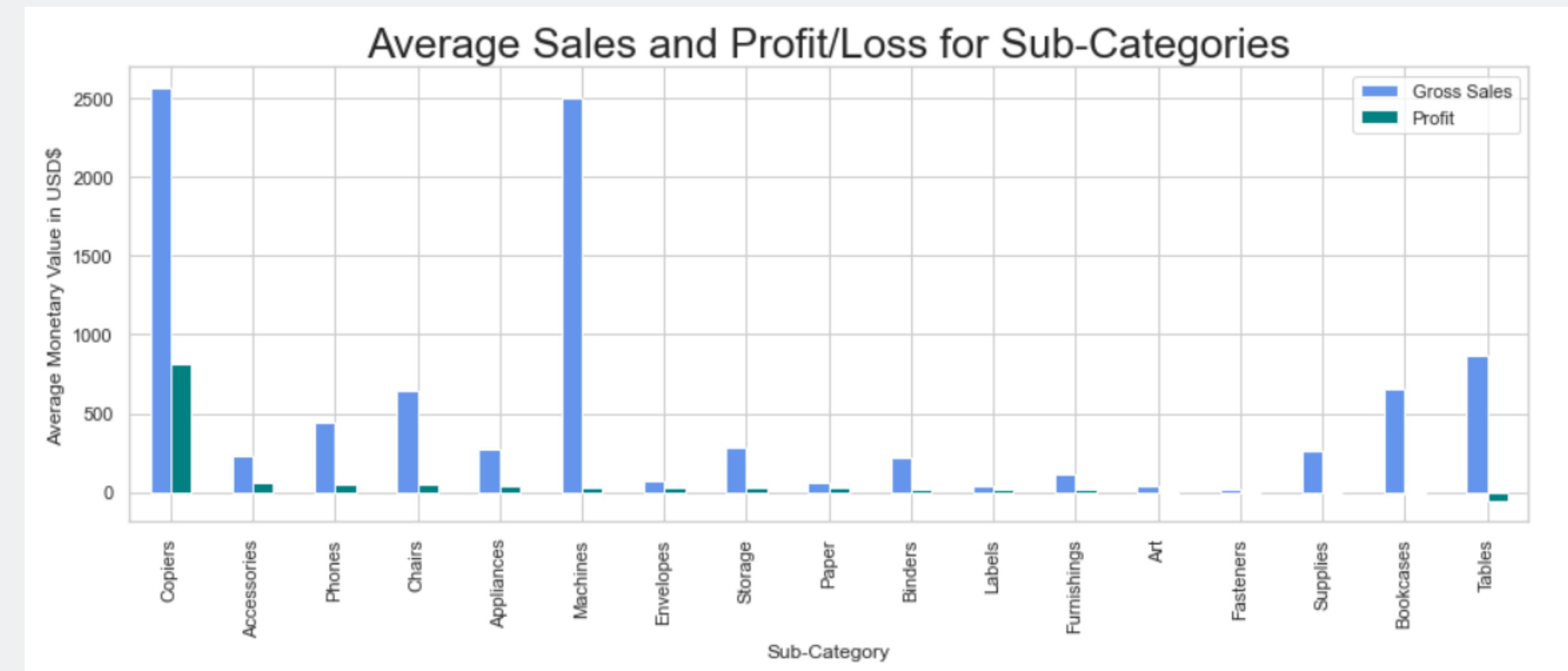




Sales vs Profit/Loss for Subcategories:

Although the sales of Machines is the second highest, the profit incurred is significantly negligible in comparison due to the high average unit price.

Working with suppliers on lowering manufacturing cost or considering vertical integration will help reduce costs across different parts of its production process.

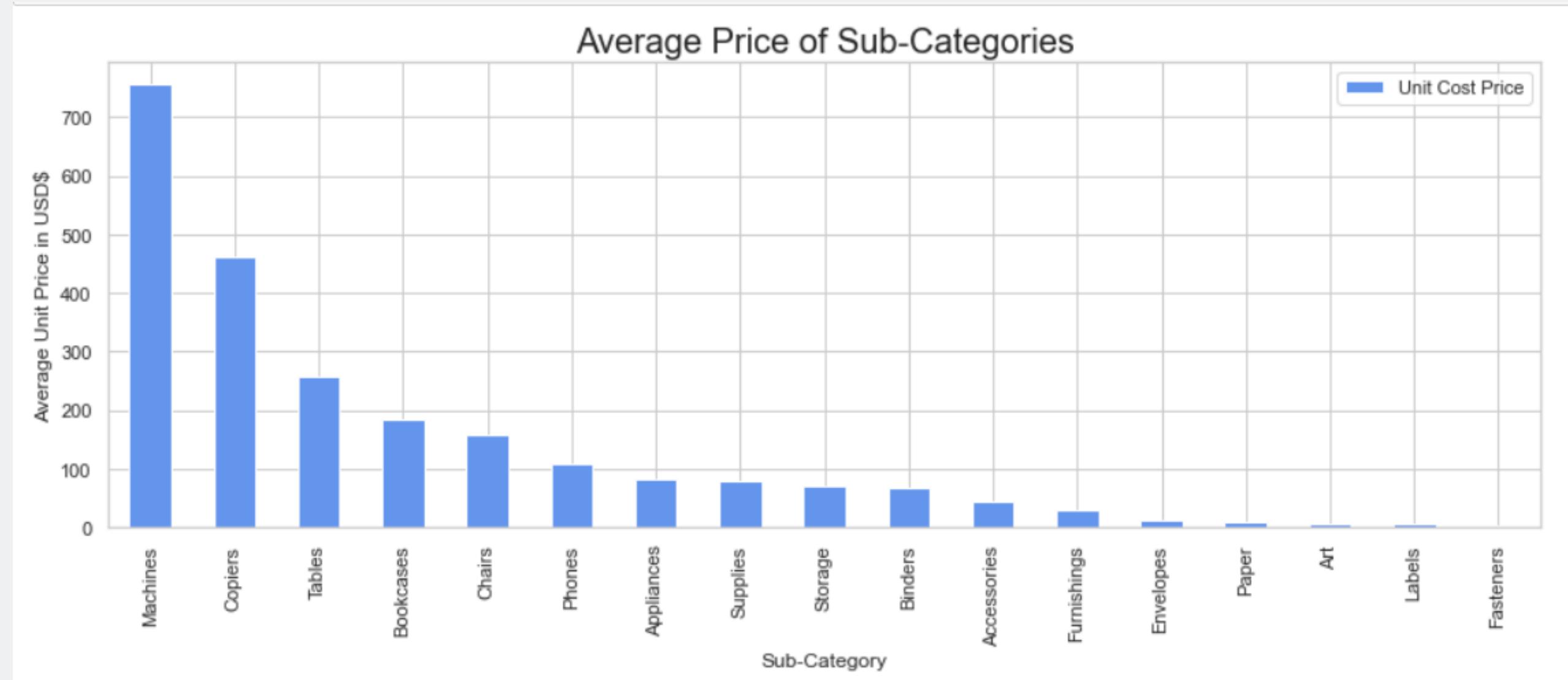


← → Q

Average Unit Price Estimation:



Estimated cost price per unit shows Tables to be the third highest(~\$100) which helps us better understand that the sales of this subcategory is not significant and the superstore could consider discontinuing the product.



Storage and supplies are the most expensive sub-categories of around USD 75. While Storage contributes to the highest profitability for this category, supplies however has a loss incurring. Hence pricing restructure for the same must be considered.



Scope to Advance:



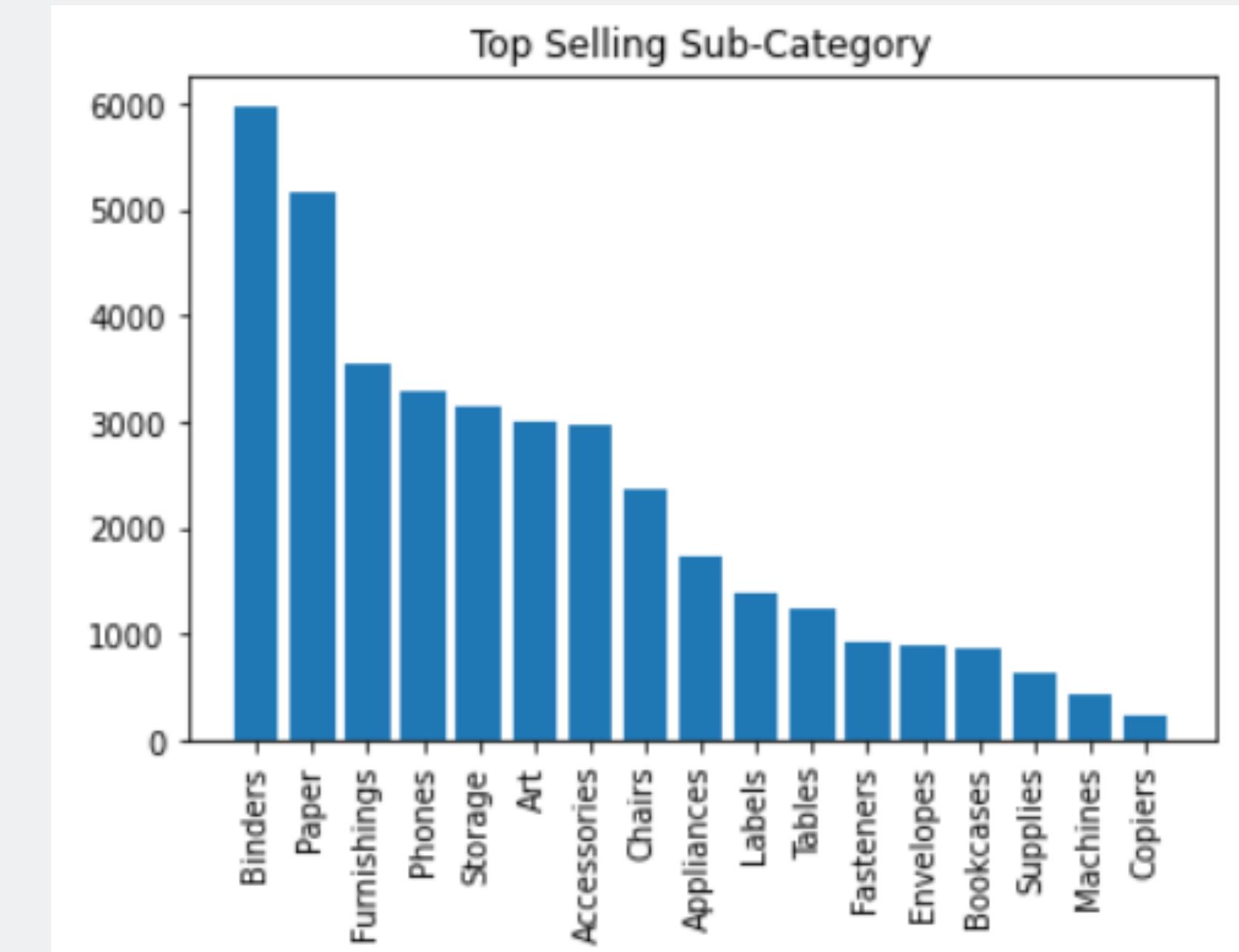
Pull back

Despite being most profitable, Copiers sell the least (only 234), but as it is a relatively expensive office equipment that is usually used for few years, it is understandable that it sells the least among all.



Push back

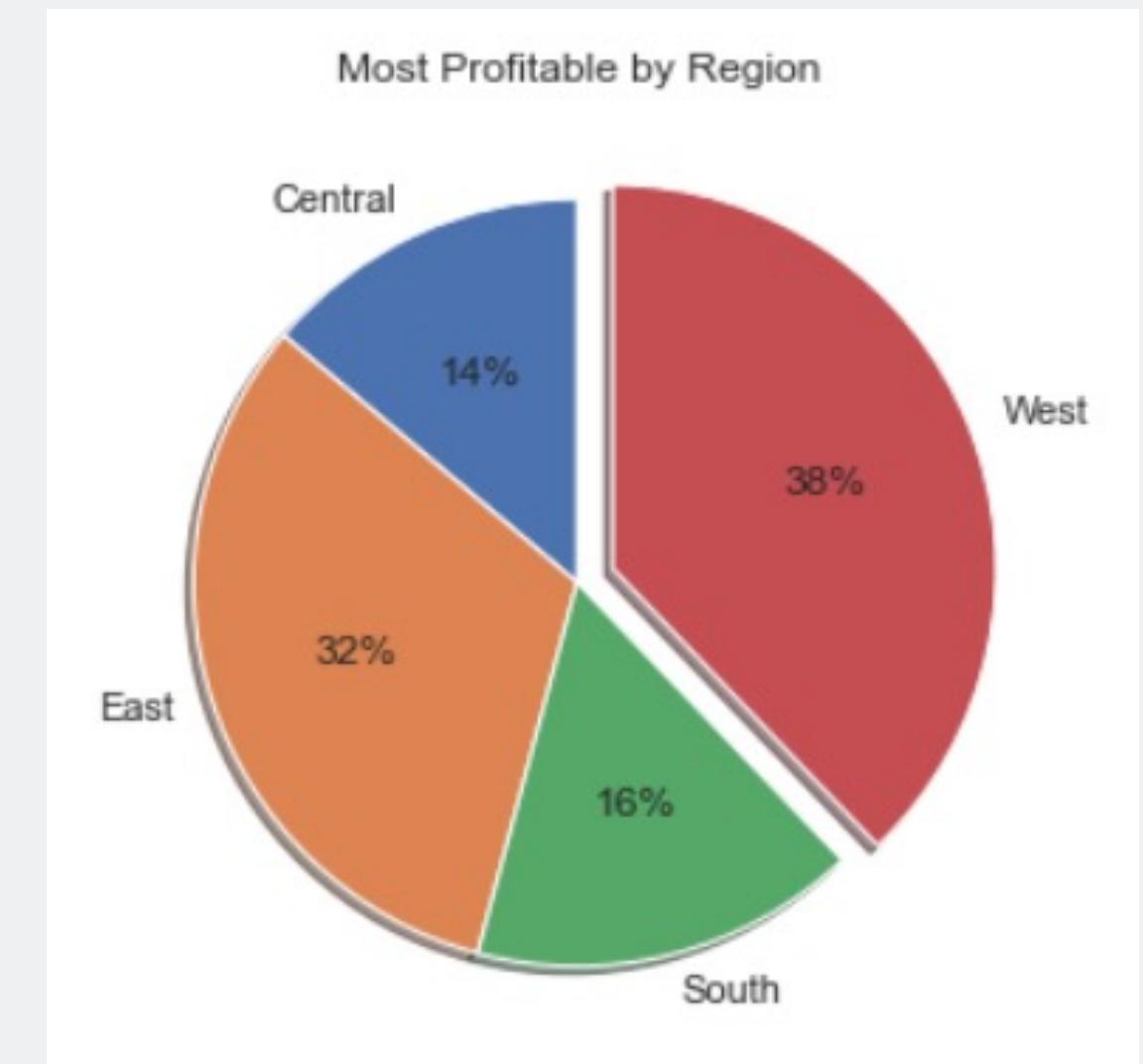
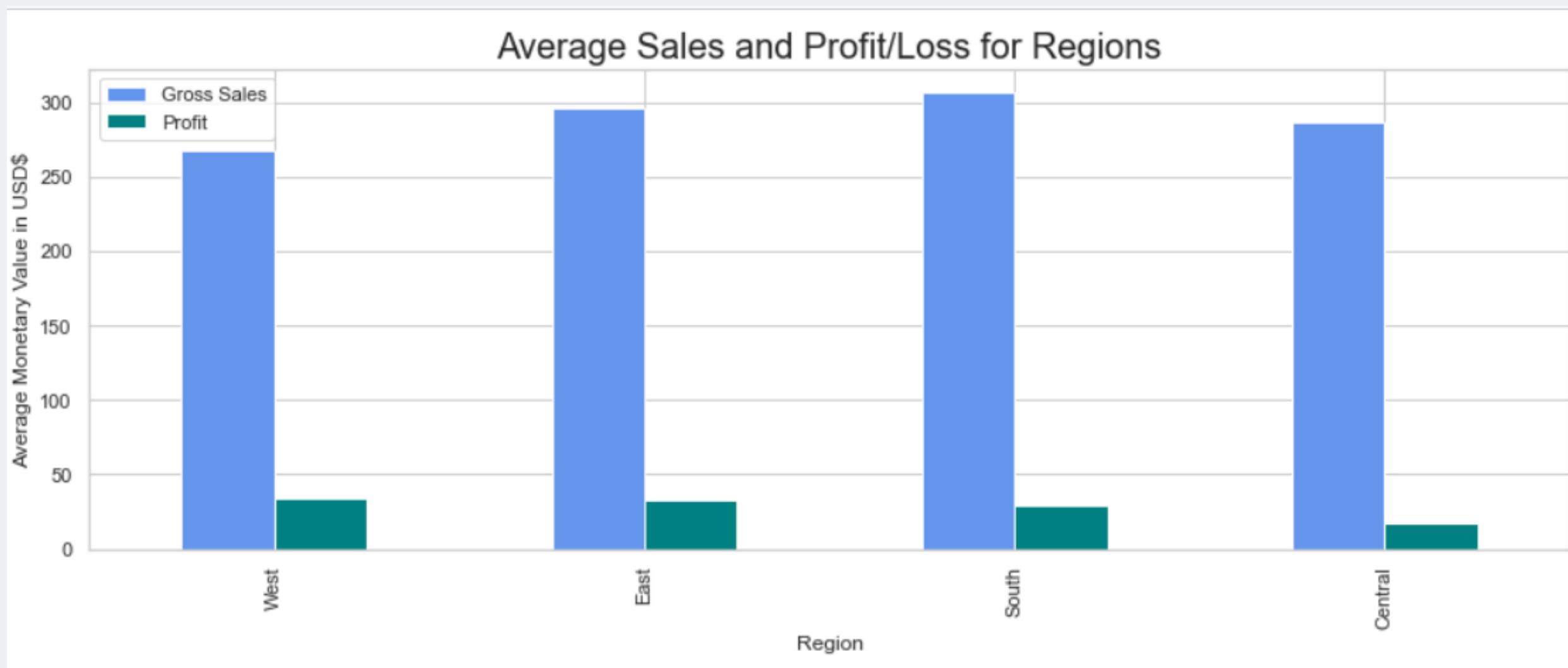
Super Store should ensure inventory are always well-stocked for the top selling sub-category such as Binders, Paper, Furnishings and Phones.





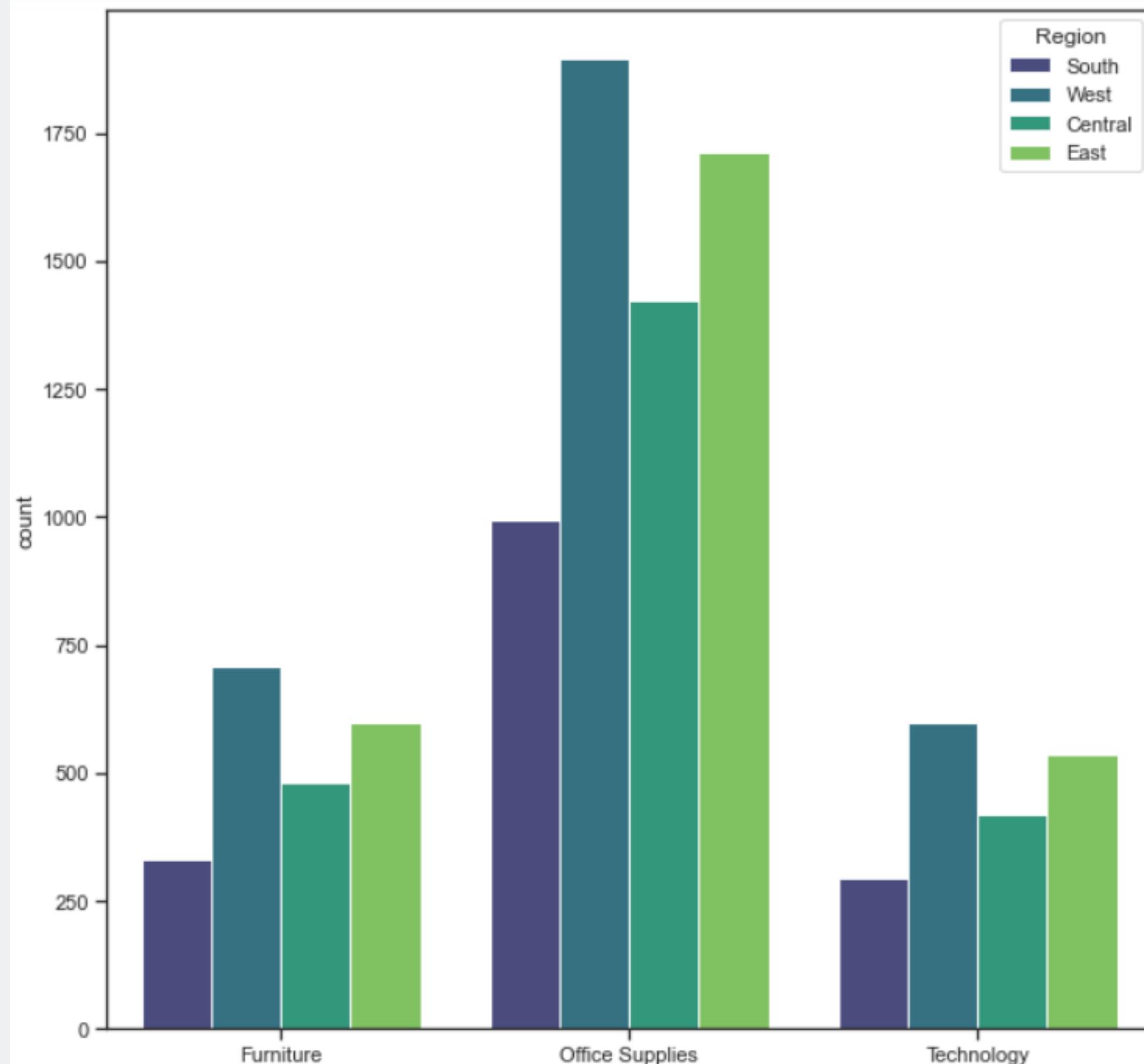
← → Q

Which Region is the Most Profitable?



← → Q

Which Category is performing profitably in which Region?

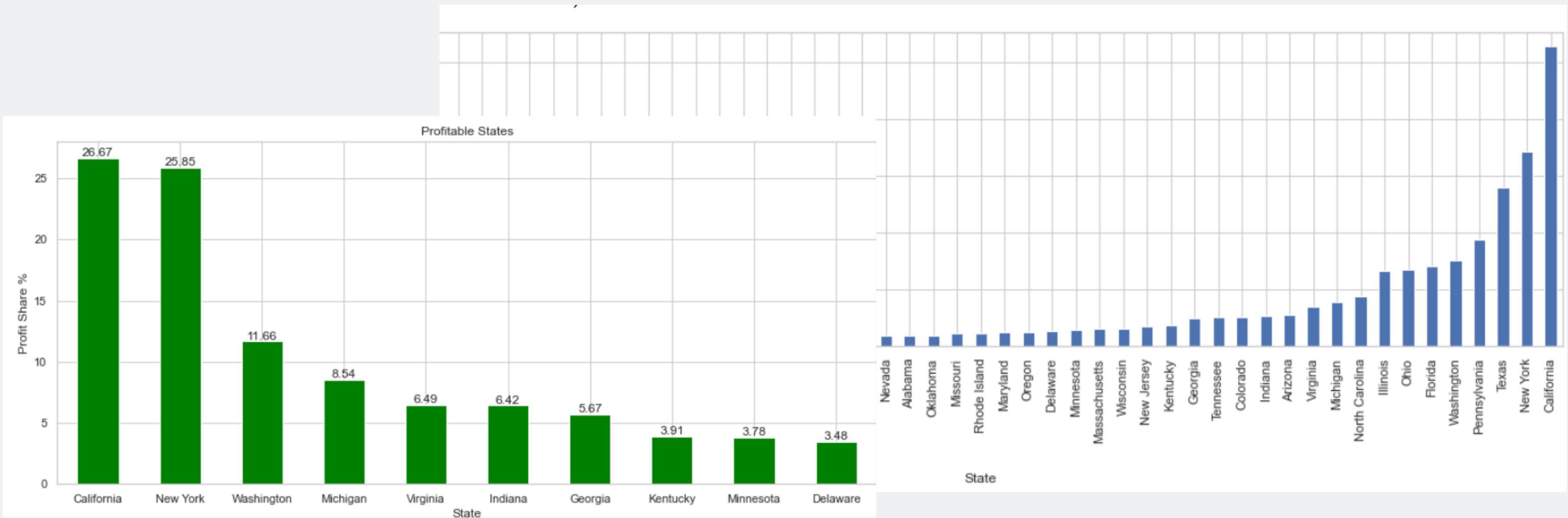
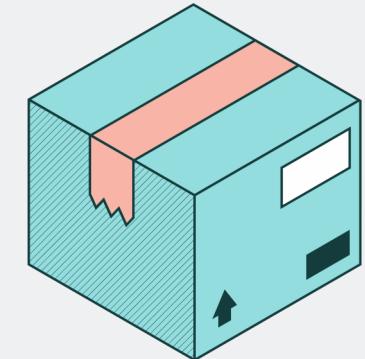


Although West region's average sales are lower than that of east and south regions, profitability accounted with the west region is the maximum as Technology, which is the most profitable category is highest in the west.



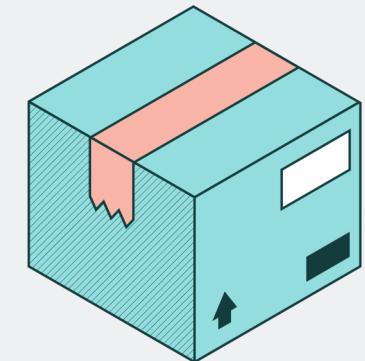
Which state has the Highest Sales?

California, New York and Texas have the highest sales and the first two contributed to more than 50% of the profitability of the store.

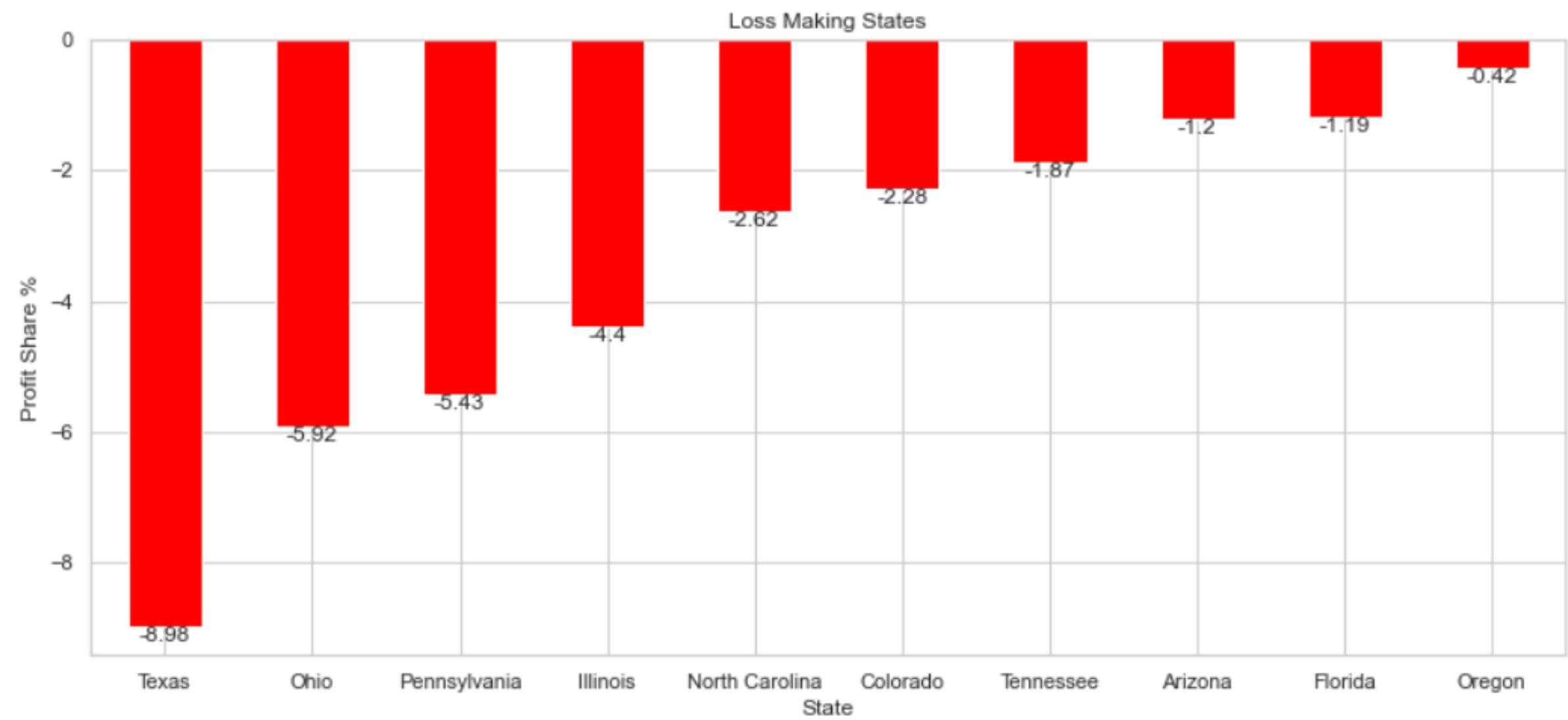
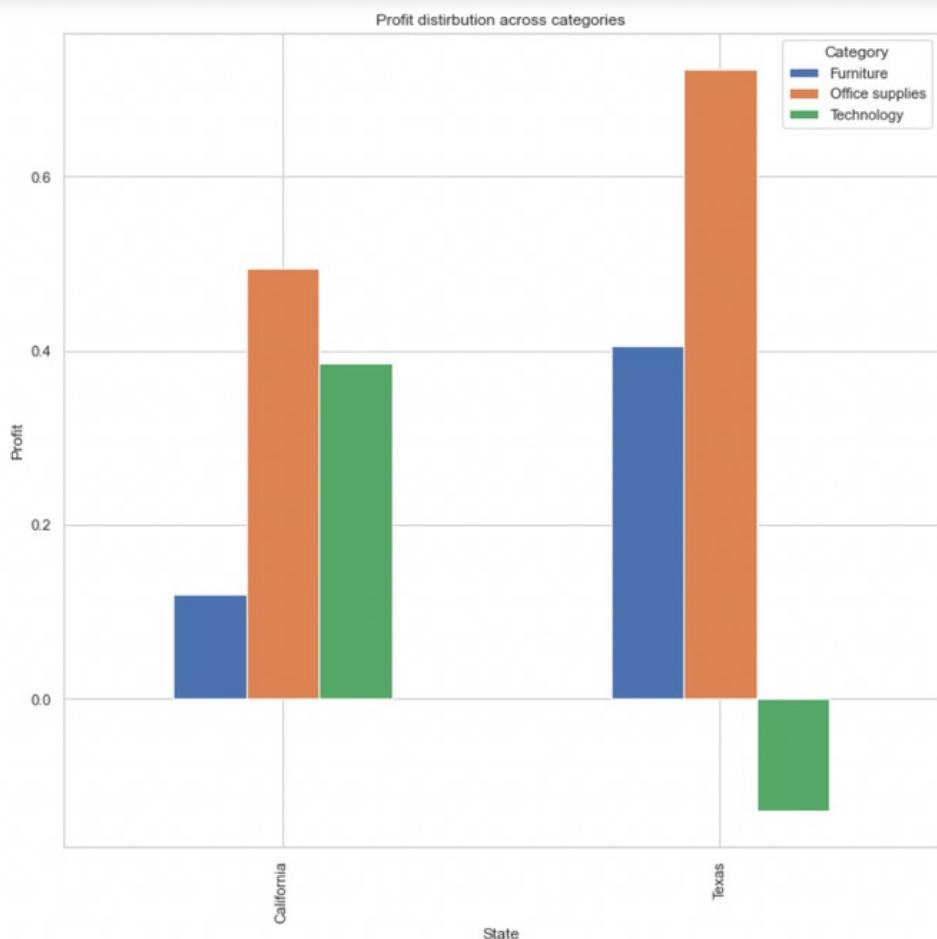


← → Q

Loss Making Trend:



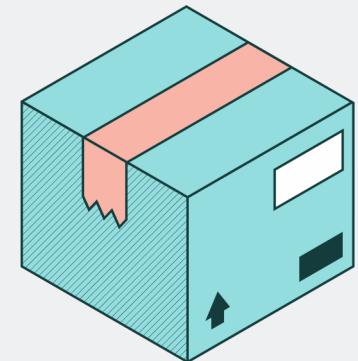
Even marking third highest in the sales, Texas has incurred the highest loss of around USD 25k. When comparing highest profitable with the highest loss incurring states, their sales across all categories are comparable. However, the overall profitability of in Texas is the least due to highest loss in high priced Technology sector



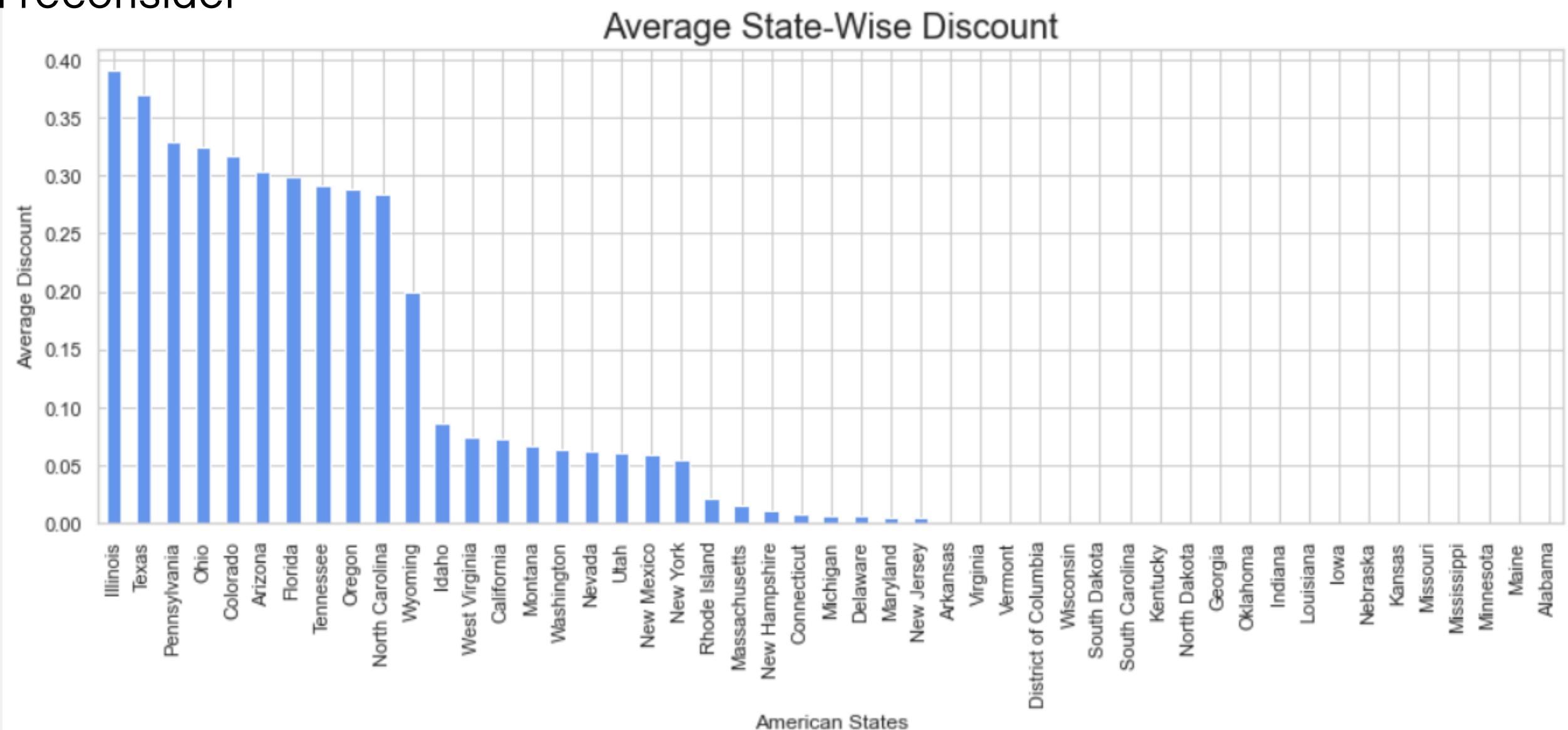


Loss Making Trend:

Pennsylvania, Illinois, and Ohio are the third, first, and fourth state offering larger discounts respectively, and this may be the reason behind them resulting in the third, fourth, and second biggest loss of around USD 15k. As in the previous point, Superstore is advised to switch to a different promotional strategy and reconsider their inventory management.



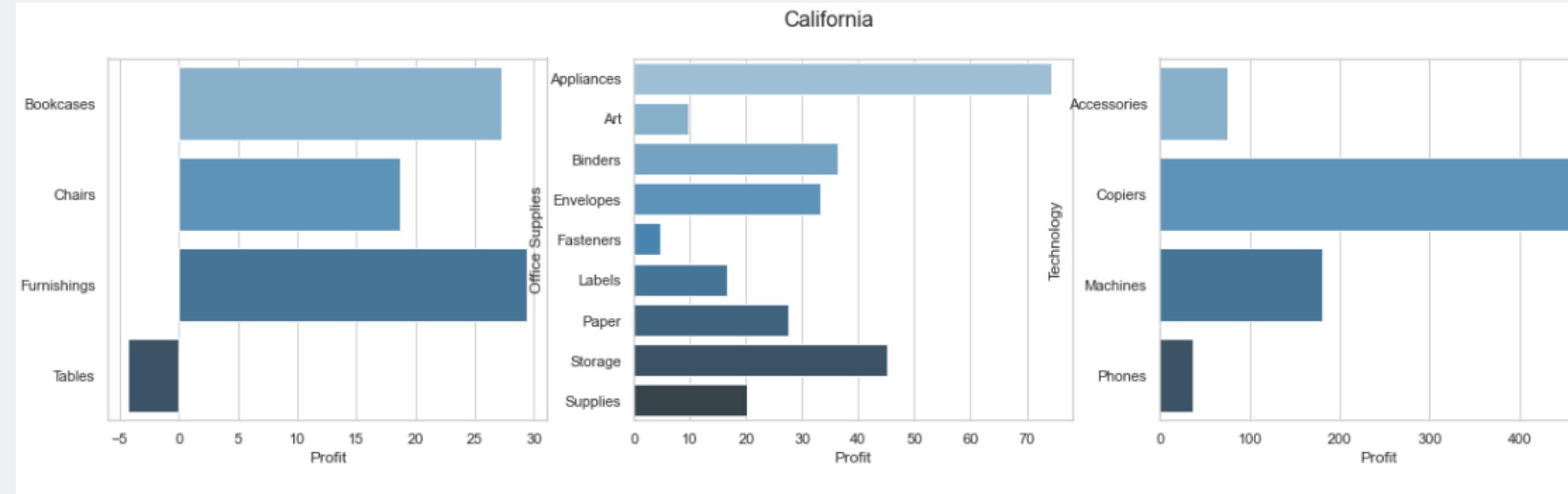
Even marking third highest in the sales, Texas has incurred the highest loss of around USD 25k. This may be due to it having the second highest discount level. Superstore is advised to reduce



← → Q

Assessing profitability across buckets

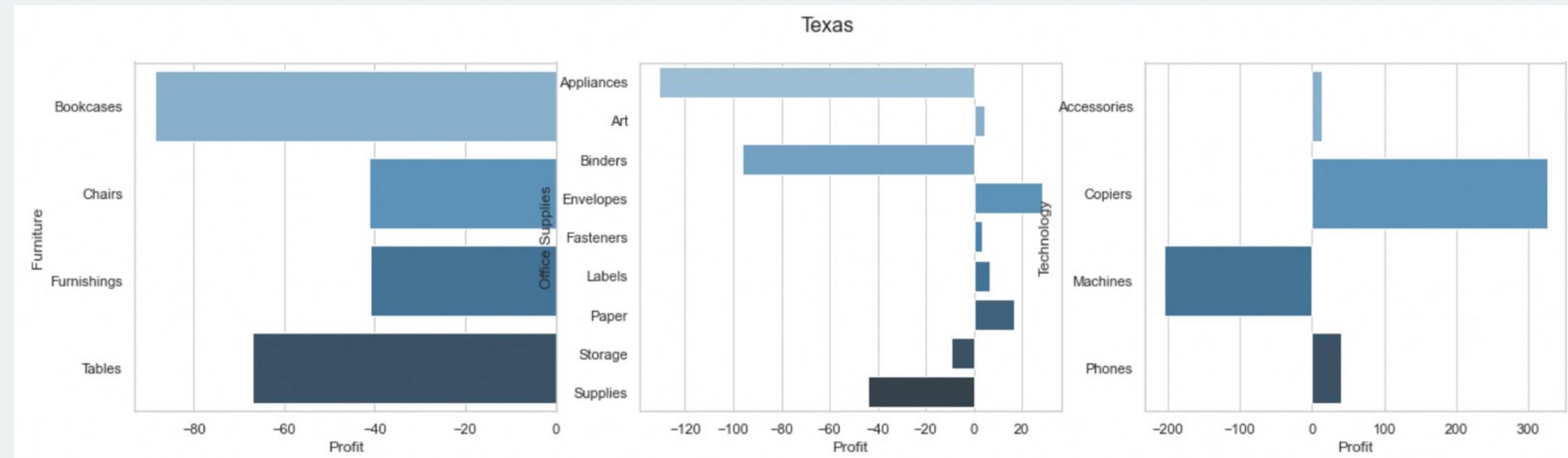
After Bucketing the states into high, medium and low profits, a sample from each of these buckets were chosen and its corresponding profit/loss-making sub categories and categories were studied.



In high profit states such as California, all products sold across all categories and sub-categories are generating comparatively good profits, except for Tables in Furniture category which is suffering from an overall loss of USD 4 per unit sold. Superstore is advised to reduce discounts for Tables. The Technology category has the best performance.

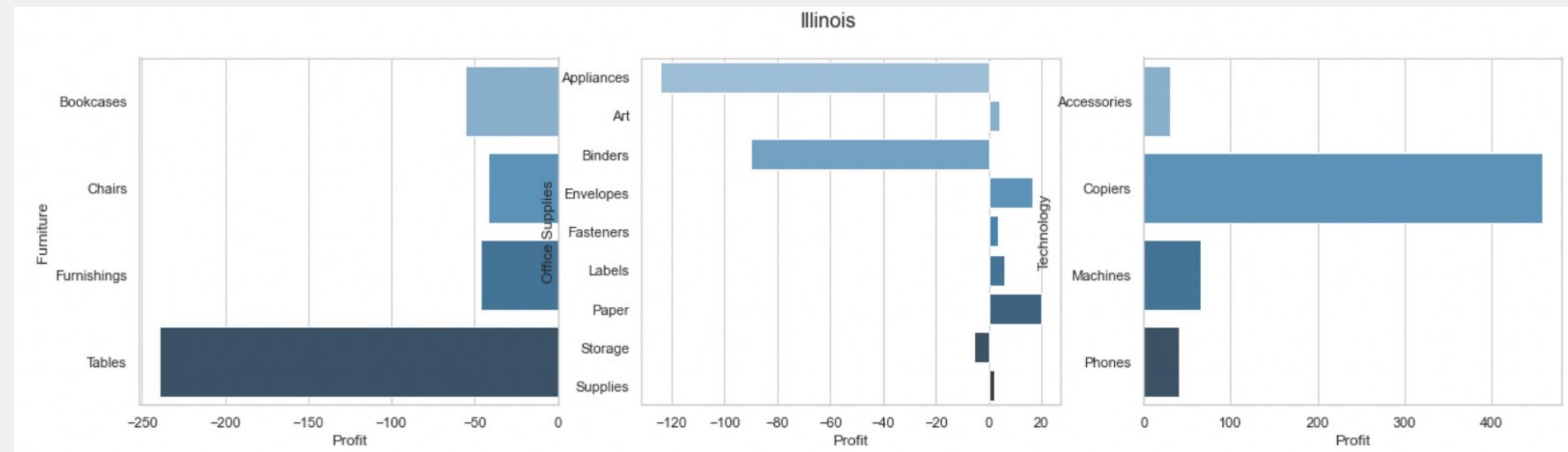
← → Q

Assessing profitability across buckets



In high loss states such as Texas, the Technology category has the best performance with Copiers making the highest profit of around USD 325 profit per unit sold, while the Furniture category's performance is going entirely in loss. Most sub-categories incur significant losses of up to USD 200, including Machines, Appliances, and Bookcases. Superstore is advised to reduce overall discounts, especially for Furniture.

Assessing profitability across buckets



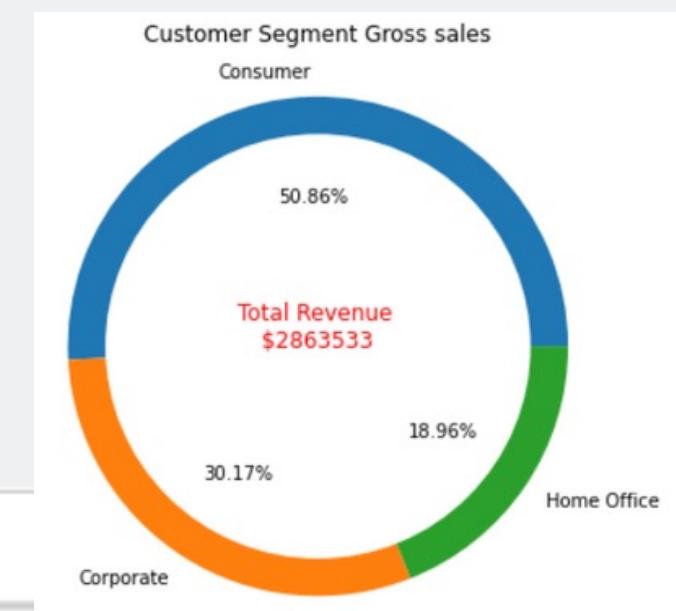
In medium loss states such as Illinois, most sub-categories generate little to no profits, and those that suffer from up to USD 240 of notable losses are Tables, Appliances, and Binders. The Technology category has the best performance with Copiers making the highest profit of around USD 450 profit per unit sold. The Furniture category's performance is going entirely in loss. Superstore is advised to reduce overall discounts, especially for Furniture.

Which Customer Segment is Most Profitable?

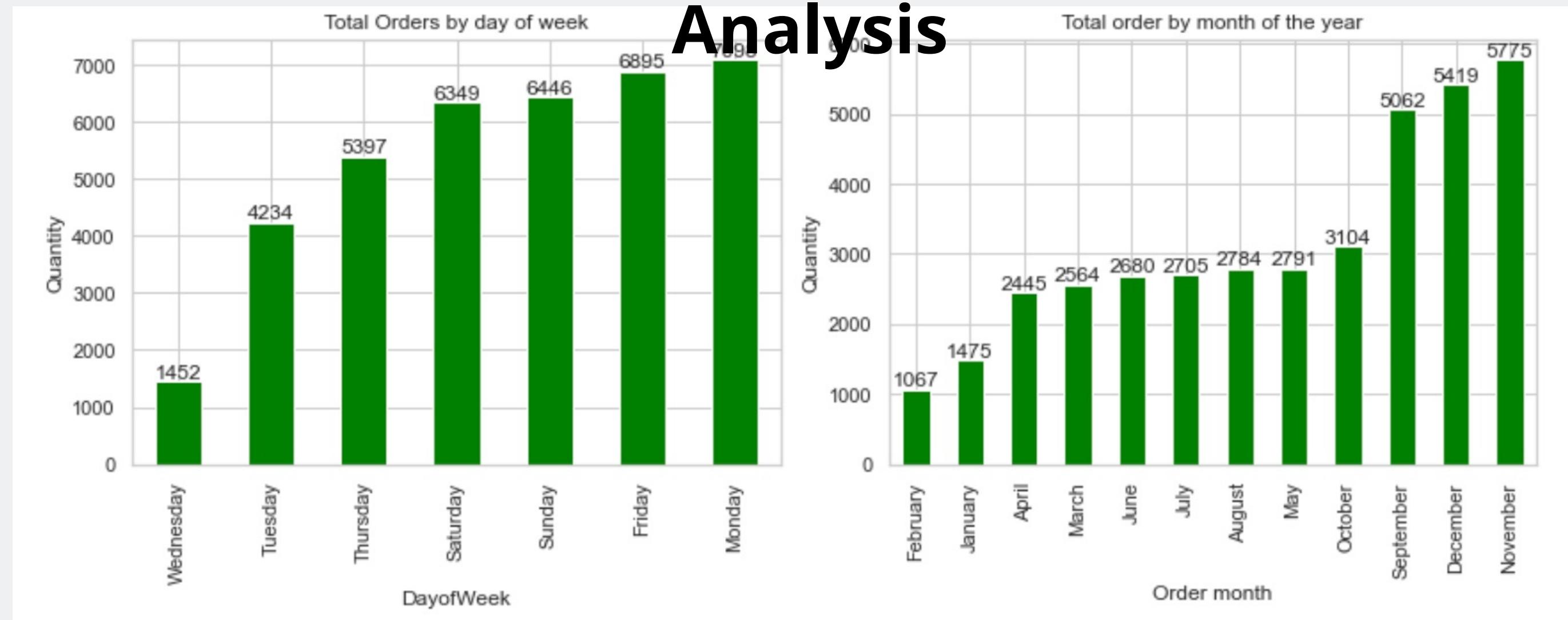


Leading

Consumer and Corporate Segment make up more than 80% of customerbase. Target them, especially customers from the East and West region in the Top 10 cities with Highest Sales by introducing special promotions and bundles for mass Consumer and Home Offices and send promotional emails or flyers.



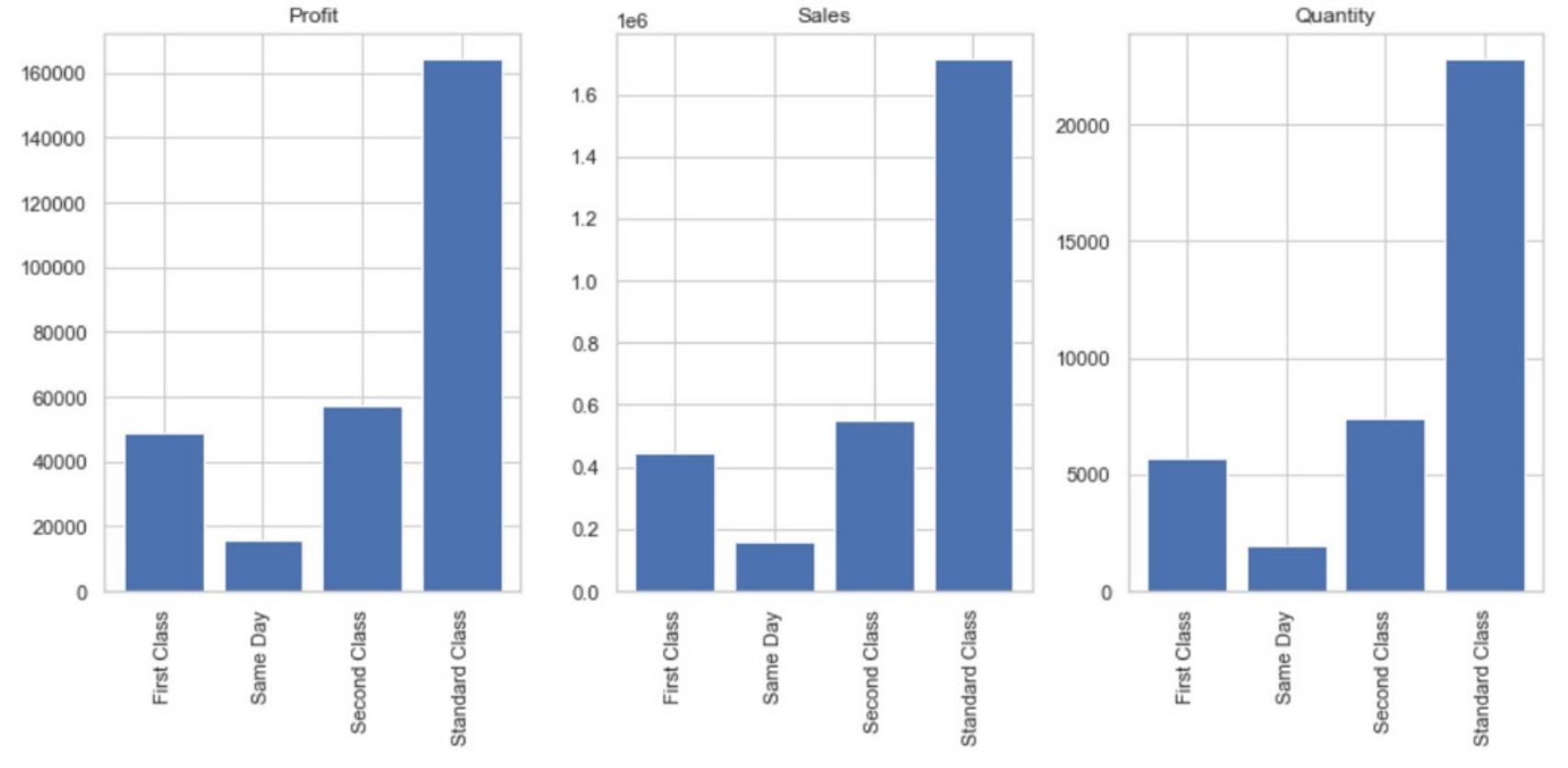
Time Stamped Analysis



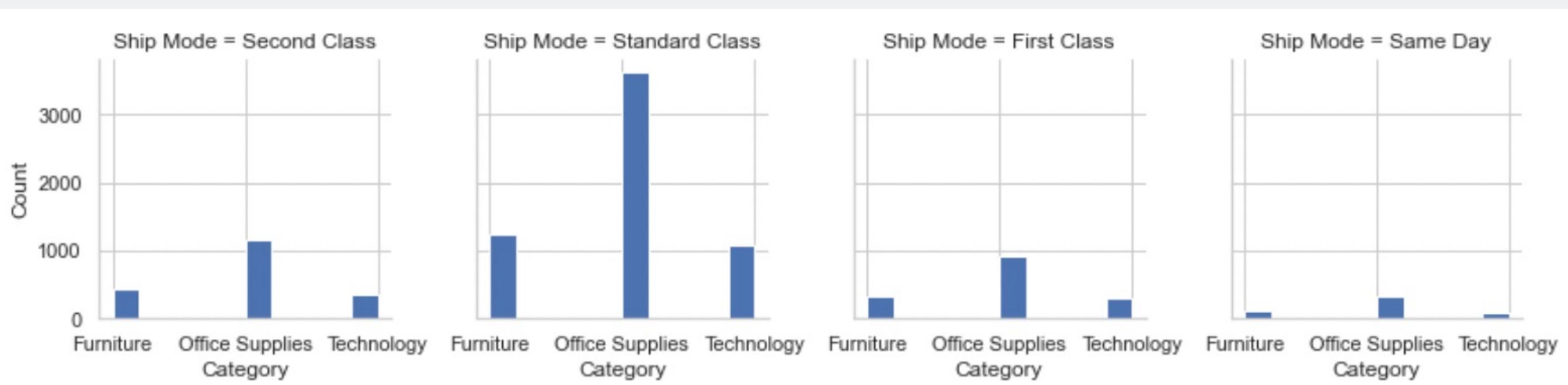
In each year, there were two month spikes whether customer ordered the most November and December. Monday was the day that customers mainly ordered from the Superstore throughout the week.

← → Q

Shipping Modes



Standard Class is the preferred method of shipment and perhaps the cheapest one too. The other modes are not popular among the customers and may be too costly.



Conclusion - Data Analysis

Targeting Strategy:

- The superstore should target Consumer segment customers as they represented the most sales & quantity of items ordered across all segments.
- They should focus on selling Staples-related products as most customers ordered these from Superstores.
- From a geographical point of view, they should target on US central and south Region, and run fewer discounts in high volume states like Texas and Illinois to minimize losses

Marketing and Promotion:

- The superstore should spend additional resources on ads in Q1 as we observed these months were the Months with the least Orders. They should put ads and special promotions on Wednesdays and Tuesdays as they recorded fewer sales

Conclusion - Data Analysis

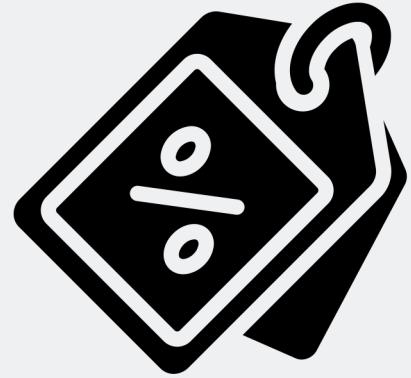
- Superstore is most profitable when discount levels lower than 10% are offered. This allows fewer occurrences of lowered prices after discount deductions, and thus larger profit margins and profits
- Run Promotional Sales in October owing to its declining seasonality
- The overall observed negative trend is that all Furniture sub-categories, as well as Appliances and Binders under the Office Supplies category, contribute to a majority of losses. Superstore is advised to improve on the promotional strategies for these, as well as reduce overall discounts

[Title Page](#)[Proprietors](#)[Introduction](#)[Review of Literature](#)[Methodology](#)[Results](#)[Conclusion](#)[References](#)[X](#)[+](#)[←](#) [→](#) [Q](#) 

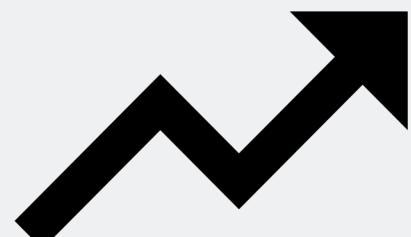
OBJECTIVE SETTING

Compute sales objectives of each state for the subsequent calendar year (2018) based on the sales data for each Category and Segment combination.

Factors used to calculate Objectives



Historical Sales

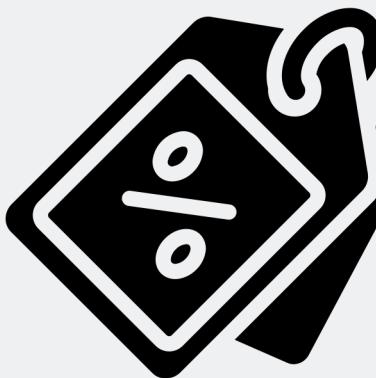


Trends



Untapped Potential

← → G Q



Historical sales

- The dataset comprises daily sales data starting from Jan 2014 to Dec 2017 . Based on the input baseline year, the historical component is calculated for each Segment-Category-State combination
- The historical contribution is calculated for each of these states for a given Segment and Category

| Segment | State | Category | Quantity | Gross Sales | Total_Quantity | Total_GrossSales | Hist_contri_GrossSales | Hist_contri_Quantity |
|----------|---------|-----------------|----------|-------------|----------------|------------------|------------------------|----------------------|
| Consumer | Alabama | Office Supplies | 5 | 27.42 | 3922 | 147292.17 | 0.000186 | 0.001275 |
| Consumer | Alabama | Technology | 8 | 239.92 | 1141 | 123984.92 | 0.001935 | 0.007011 |
| Consumer | Arizona | Furniture | 20 | 2710.95 | 1219 | 138034.00 | 0.019640 | 0.016407 |
| Consumer | Arizona | Office Supplies | 74 | 1846.58 | 3922 | 147292.17 | 0.012537 | 0.018868 |

← → G 🔎



Trended sales

- Year on Year Growth rate analysis was performed for each of the States for a given Segment and Category**
- The average growth rate across years (2015-2017) was computed and applied to historical sales to determine the trend component for each state**
- The trend contribution is calculated for each of these states for a given Segment and Category**

| Order year | Segment | State | Category | Gross Sales | | | | Quantity | | | | GrossSalesGR (2016Vs2015) | GrossSalesGR (2017Vs2016) | QuantityGR (2016Vs2015) | QuantityGR (2017Vs2016) |
|---------------|----------|---------|-----------------|-------------|---------|---------|--------|----------|------|------|------|------------------------------|------------------------------|----------------------------|----------------------------|
| | | | | 2014 | 2015 | 2016 | 2017 | 2014 | 2015 | 2016 | 2017 | | | | |
| 0 | Consumer | Alabama | Furniture | 1828.82 | 1858.82 | 350.98 | 0.00 | 16.0 | 9.0 | 1.0 | 0.0 | -0.811181 | -1.000000 | -0.888889 | -1.000000 |
| 1 | Consumer | Alabama | Office Supplies | 174.70 | 498.93 | 1098.13 | 27.42 | 10.0 | 7.0 | 30.0 | 5.0 | 1.200970 | -0.975030 | 3.285714 | -0.833333 |
| 2 | Consumer | Alabama | Technology | 180.93 | 59.97 | 1218.92 | 239.92 | 3.0 | 3.0 | 10.0 | 8.0 | 19.325496 | -0.803170 | 2.333333 | -0.200000 |

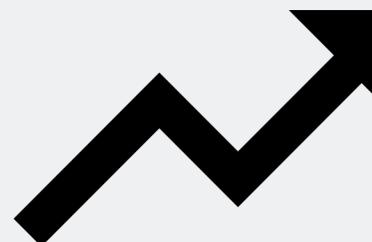
← → G Q



Untapped Potential

- Untapped Potential is calculated by taking the population belonging to the age group of 20-60 years from each state and its contribution is calculated against the total US population belonging to the same age group

| Segment | State | Category | Population | Total_population | UP_contri |
|-------------|-----------|-----------------|------------|------------------|-----------|
| Consumer | Alabama | Office Supplies | 3814879 | 141635147042 | 0.000027 |
| Consumer | Alabama | Office Supplies | 3814879 | 141635147042 | 0.000027 |
| Consumer | Alabama | Office Supplies | 3814879 | 141635147042 | 0.000027 |
| Consumer | Alabama | Office Supplies | 3814879 | 141635147042 | 0.000027 |
| Consumer | Alabama | Office Supplies | 3814879 | 141635147042 | 0.000027 |
| ... | ... | ... | ... | ... | ... |
| Home Office | Wisconsin | Technology | 4555837 | 139477314755 | 0.000033 |
| Home Office | Wisconsin | Technology | 4555837 | 139477314755 | 0.000033 |
| Home Office | Wisconsin | Technology | 4555837 | 139477314755 | 0.000033 |
| Home Office | Wisconsin | Technology | 4555837 | 139477314755 | 0.000033 |
| Home Office | Wisconsin | Technology | 4555837 | 139477314755 | 0.000033 |

← → G
Q
🔍


Forecast Calculation

- At the Nation level (US) , the average year-on-year growth rate is computed for each Segment-Category combination**
- The computed growth rate is applied on historical sales to obtain the forecast for each Segment-Category set**

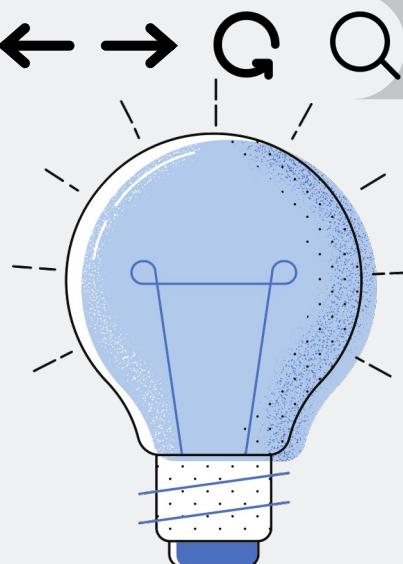
| Segment | Category | Gross Sales | Quantity | GrossSalesAVG | QuantityAVG | GrossSales_Forecast | Quantity_Forecast |
|---------|-----------------------------|--------------|----------|---------------|-------------|---------------------|-------------------|
| 0 | Consumer Furniture | 105195085.91 | 1021192 | 0.07 | 0.14 | 112205635.22 | 1163369.60 |
| 1 | Consumer Office Supplies | 115287210.56 | 3246278 | 0.22 | 0.26 | 140423863.24 | 4091663.68 |
| 2 | Consumer Technology | 82583456.31 | 941843 | 0.08 | 0.17 | 88843464.22 | 1106136.77 |
| 3 | Corporate Furniture | 54483850.48 | 565328 | 0.26 | 0.19 | 68379663.87 | 675110.65 |
| 4 | Corporate Office Supplies | 97023211.53 | 1853099 | 0.64 | 0.25 | 158675402.90 | 2321241.05 |
| 5 | Corporate Technology | 96196128.08 | 606990 | 0.40 | 0.31 | 134415636.87 | 796633.54 |
| 6 | Home Office Furniture | 47057243.74 | 410979 | 0.38 | 0.28 | 64824333.49 | 525214.28 |
| 7 | Home Office Office Supplies | 29501474.45 | 1080333 | 0.52 | 0.38 | 44902123.65 | 1487783.04 |
| 8 | Home Office Technology | 76082584.34 | 348935 | 0.93 | 0.52 | 146555315.10 | 530626.26 |

← → G Q



Objectives Calculation

- Input weights for each of the three factors i.e., History, Trends, and Untapped potential.
- Based on input weights, weighted contributions are calculated and accordingly applied to the final forecast for each Segment and Category
- The final yearly objectives are obtained for each Segment, Category, and State



Objectives Calculation - Example

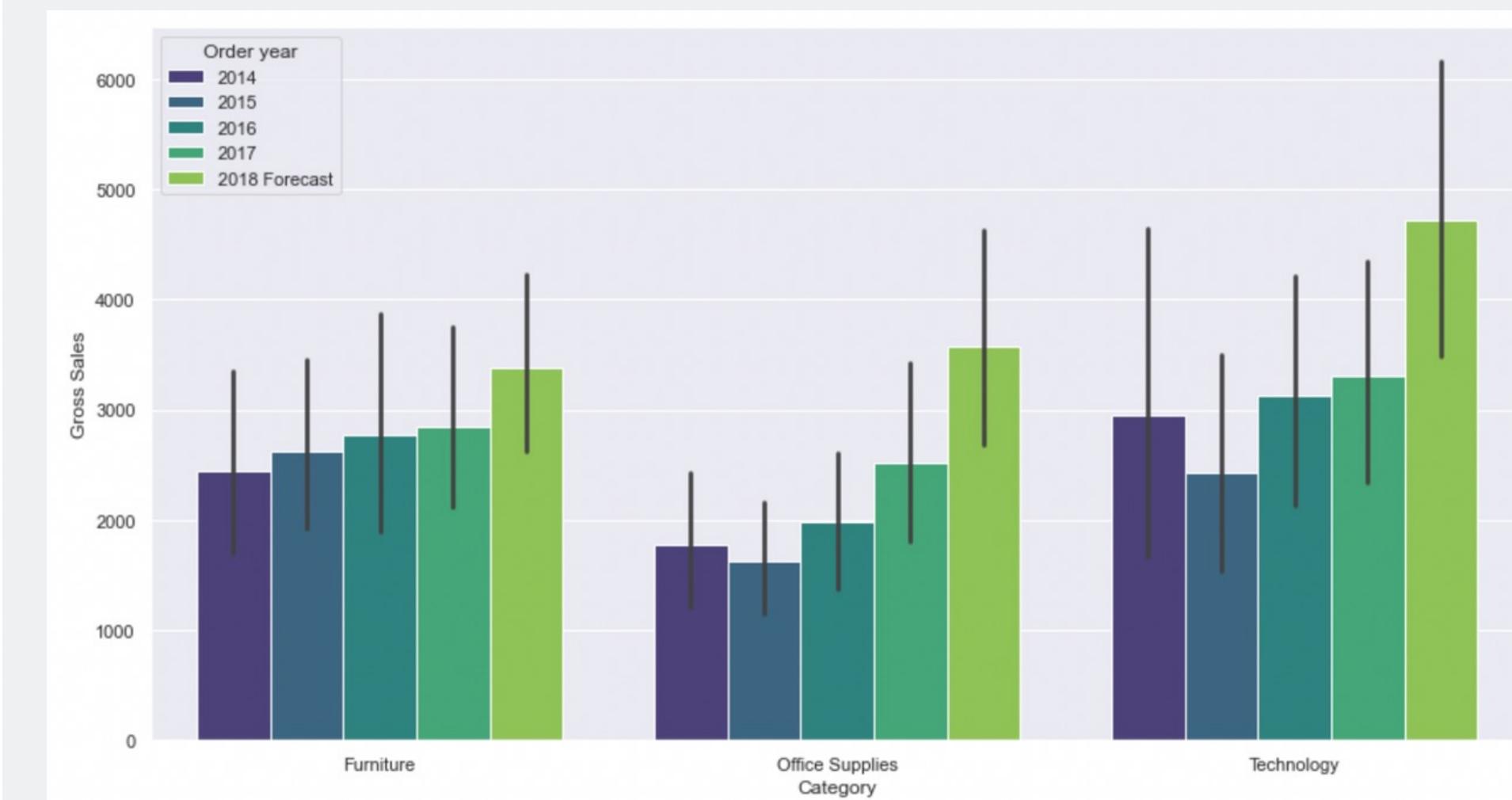
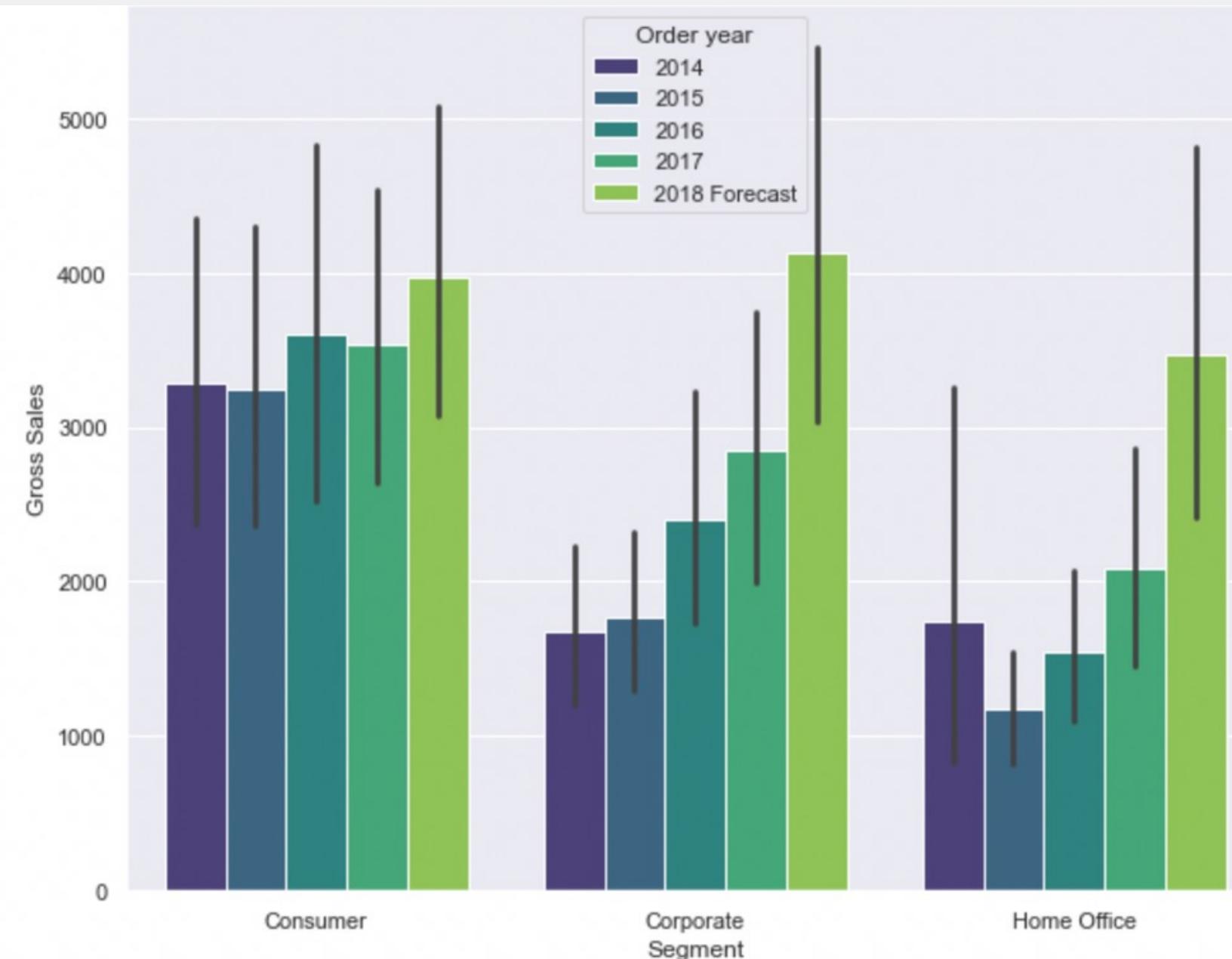
Objectives (A) = (History(A)*0.5 + Trend(A)*0.3 + UP(A)*0.2)*Forecast

| Parameters | Weightage |
|--------------------|-----------|
| History | 50% |
| Trend | 30% |
| Untapped Potential | 20% |
| Forecast | 10,000 |

| State | History Contribution | Trend Contribution | UP Contribution | Objectives |
|-------|----------------------|--------------------|-----------------|------------|
| A | 6.67% | 10.00% | 14.29% | 919.05 |
| B | 13.33% | 15.00% | 17.14% | 1,459.52 |
| C | 20.00% | 20.00% | 20.00% | 2,000.00 |
| D | 26.67% | 25.00% | 22.86% | 2,540.48 |
| E | 33.33% | 30.00% | 25.71% | 3,080.95 |
| Total | 1 | 1 | 1 | 10,000.00 |

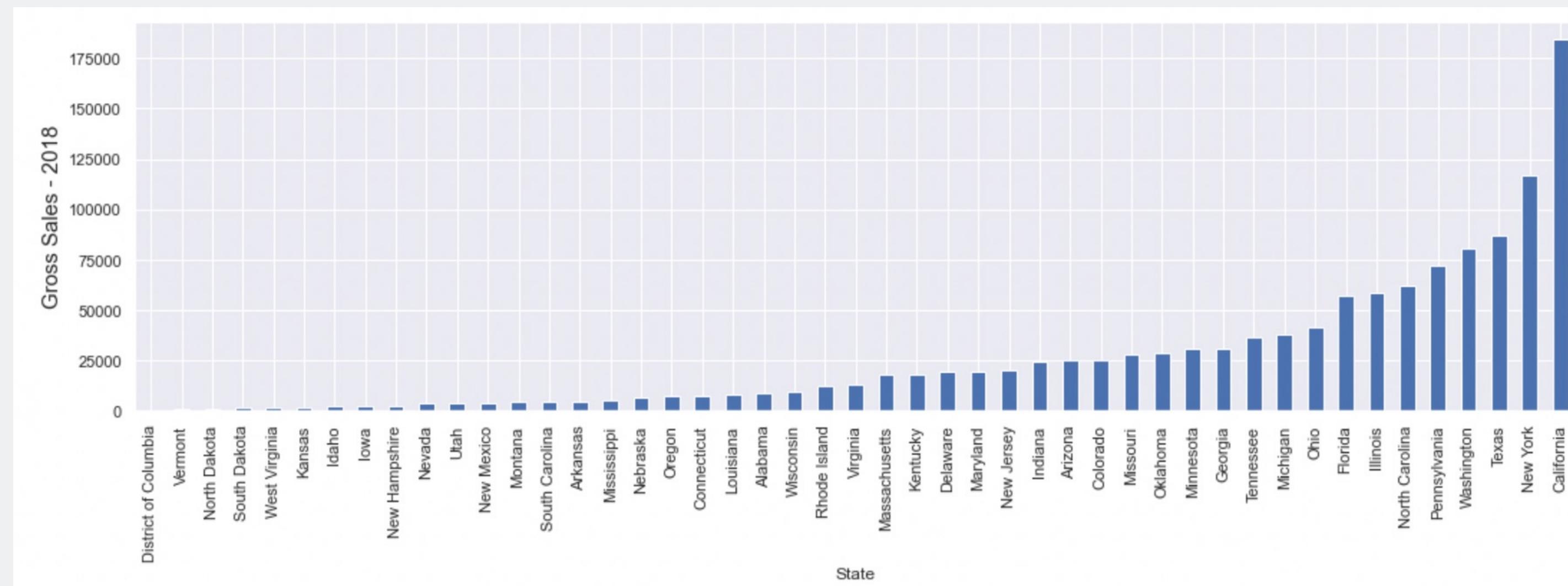
Forecast Analysis - 2018

On comparing the sales year-wise across Segments and categories, we can observe that 2018 has assumed a nominal growth rate considering 2017 as the baseline



Forecast Analysis - 2018

The 2018 sales forecast follow the same trend across states as that of the previous three years on average assuming no major market fluctuations





← → Q

Forecast Analysis -2018

For the top 3 cities, we can observe the 2018 numbers are on an increasing trend across various categories except the furniture segment in new york which saw a steep decline from 2015 to 2016

