Superstore

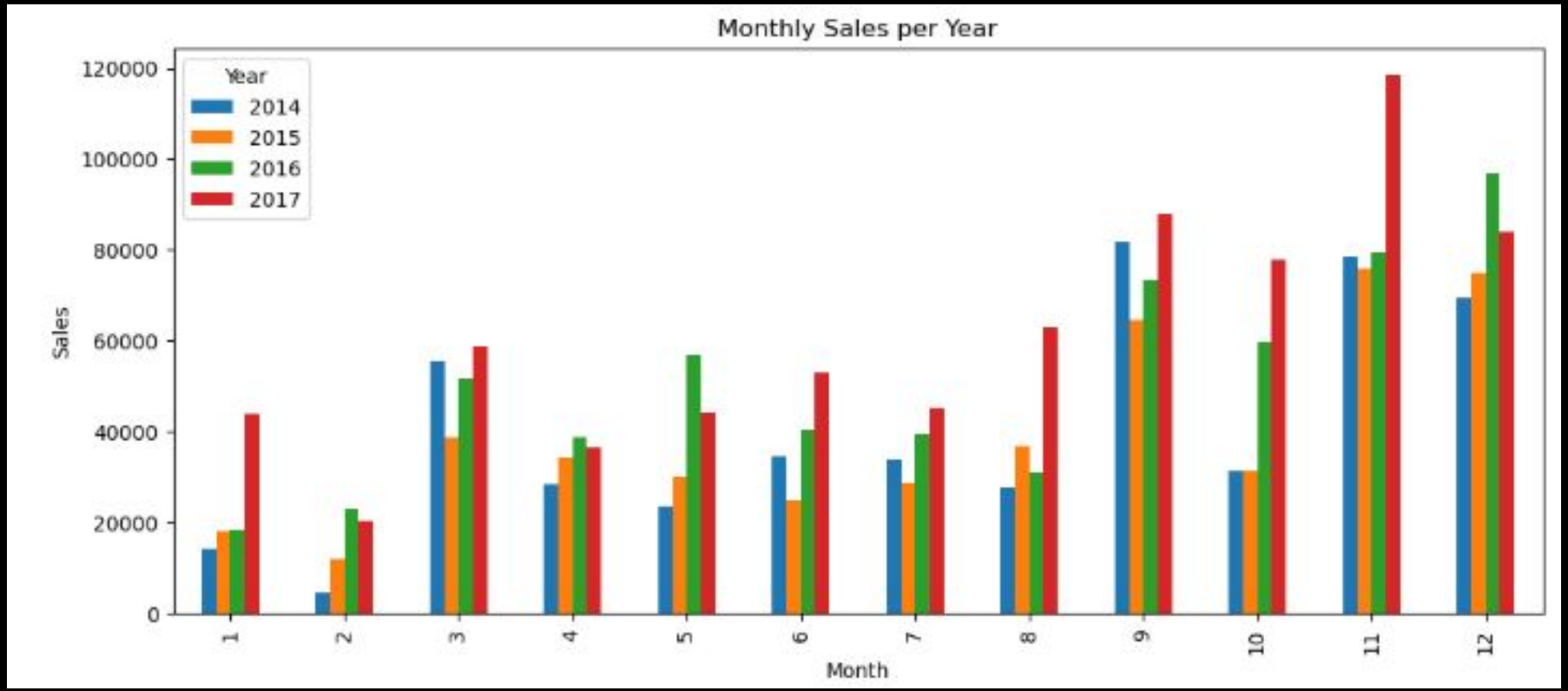
Increasing Profitability Recommendations

Problem Identification: What kind of prediction model with an accuracy of 80% or more, can we make that can help us increase profits in geographical areas and predict future sales trends in the next 30 days?

Criteria For Success:

- Find the patterns: When do people buy the most? Which months or seasons matter?
- Find which products, regions, and customers are truly profitable.
- Which products are dragging us down?
- What type of model can we make to predict future sales trends?
- Can we produce a model that has an accuracy of 80% or more in the next 30 days?

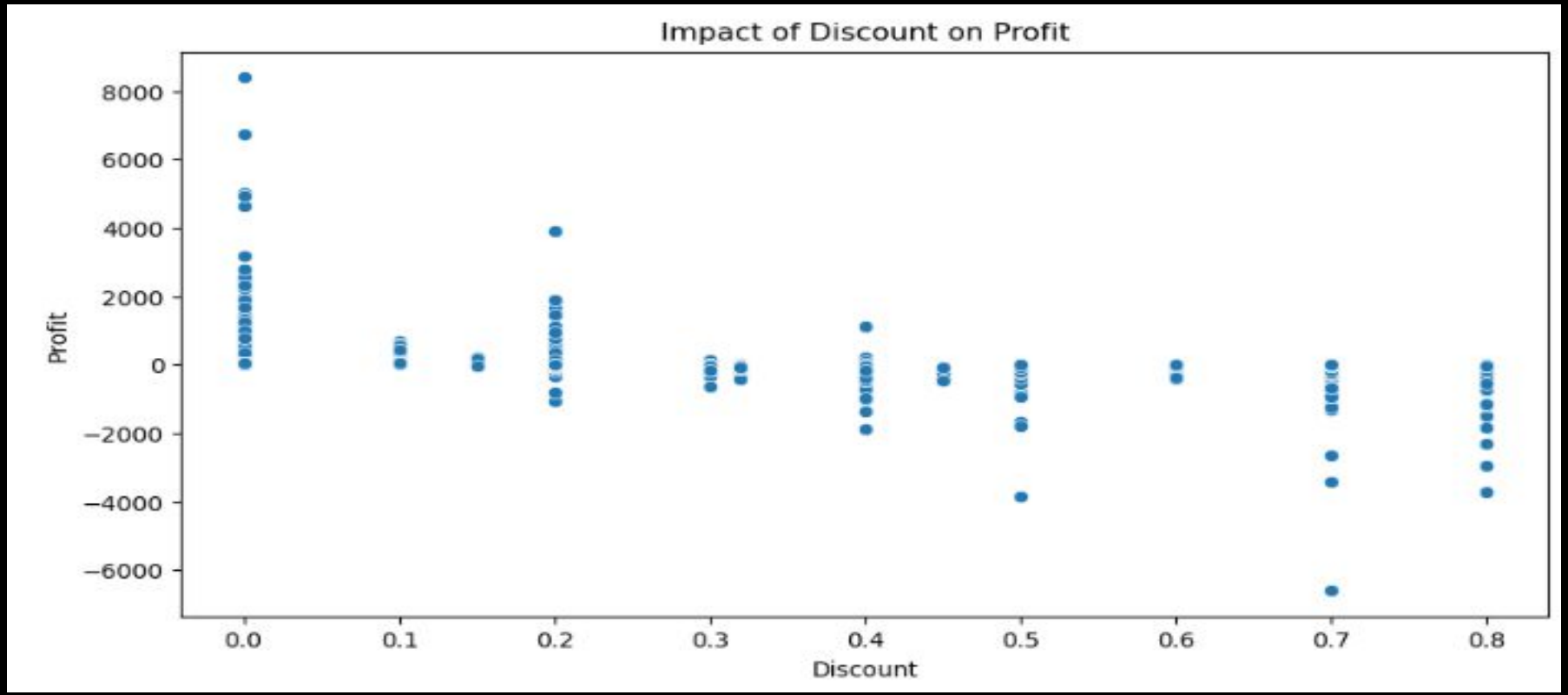
Recommendations



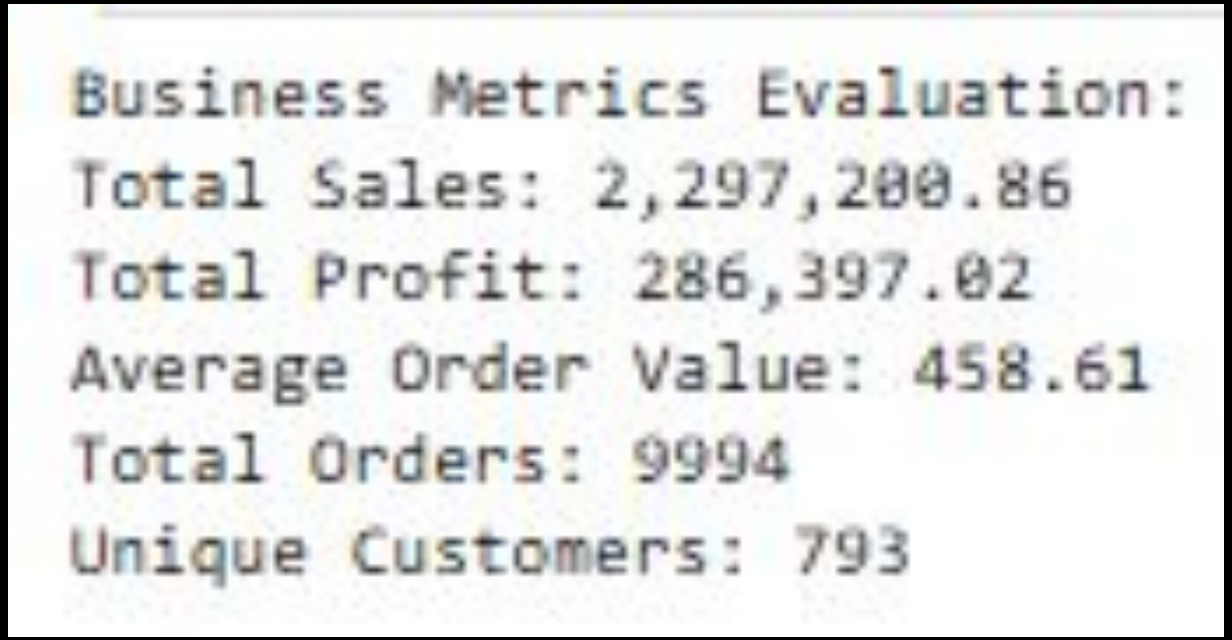
Recommendations

- The monthly sales trend shows sales spikes in December of each year, reaching record highs in December 2017. This pattern strongly suggests that end-of-year holiday promotions and campaigns are the primary drivers of sales growth.
- Additionally, there are noticeable dips in sales during the spring and late summer months, indicating where possible customer demand has decreased. To capitalize on these trends, the company should invest more in marketing of the holiday season to maximize profits.
- For the spring and summer months, launching targeted promotions or new product releases could help smooth out sales fluctuations and maintain steady revenue throughout the year.
- Decreasing heavily discounted items to a 20% or lower discount will potentially increase profitability for the superstore.

Recommendations



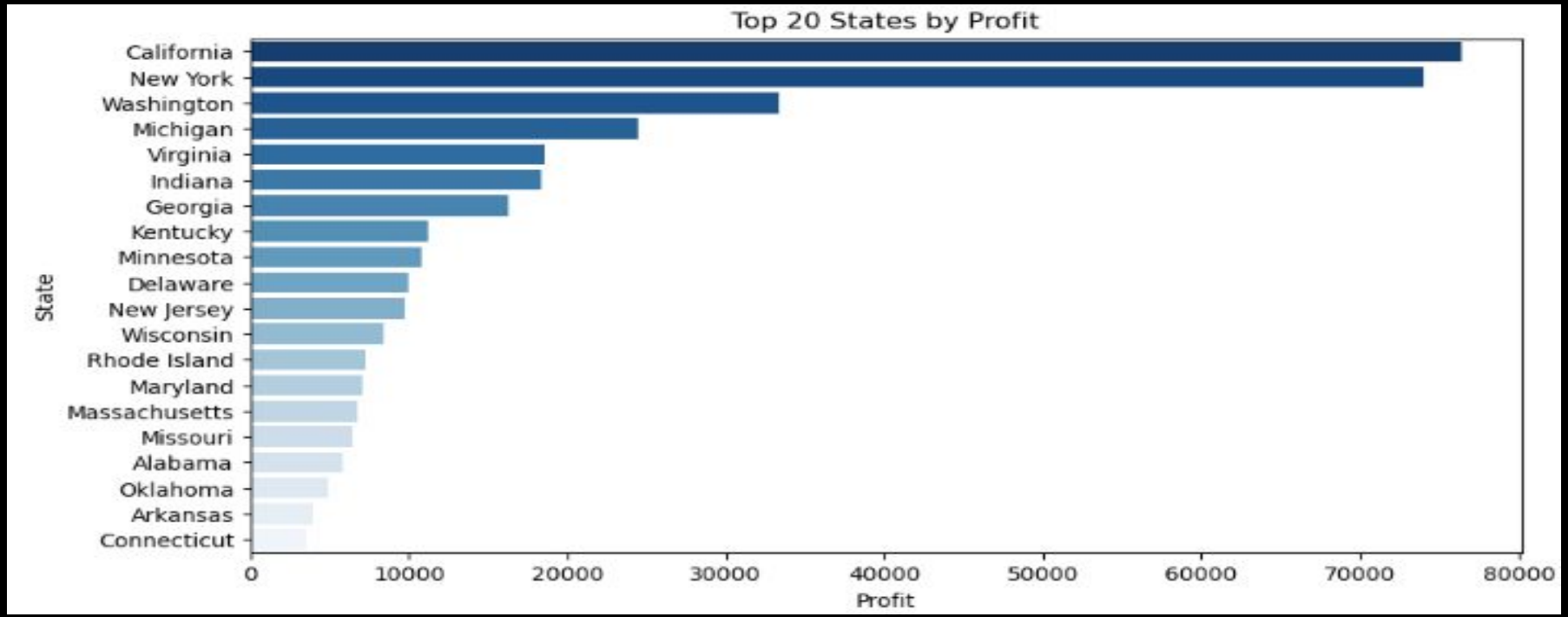
Business Metrics Evaluation



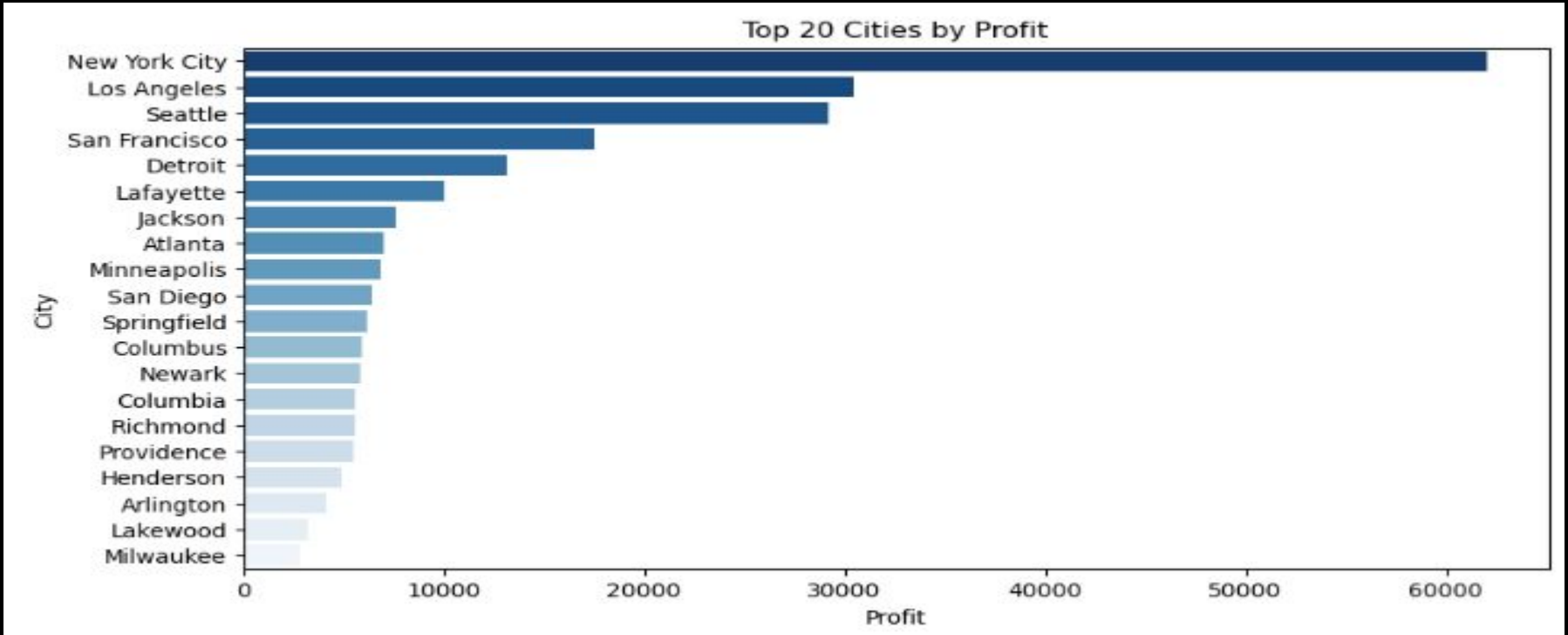
Key Findings

- The business is already operating profitably at about a 10% profit.
- The most sales for the superstore are produced seasonally. The last 4 months of the year produce the most sales.
- New York City and Los Angeles had the most product sales for the company compared to other cities.
- West region has the highest volume of sales and is the most profitable region
- We found that the superstore made most of its profits selling copiers and the “Technology” category was the most profitable out of every other category.
- Our most profitable customers were “Home Office” and “Corporate” customers.
- Heavily discounted items were negatively impacting our profits. Any discounted item with a 30% or more discount is likely to produce unprofitable results.
- “Tables” are the items that are hurting our profitability the most.

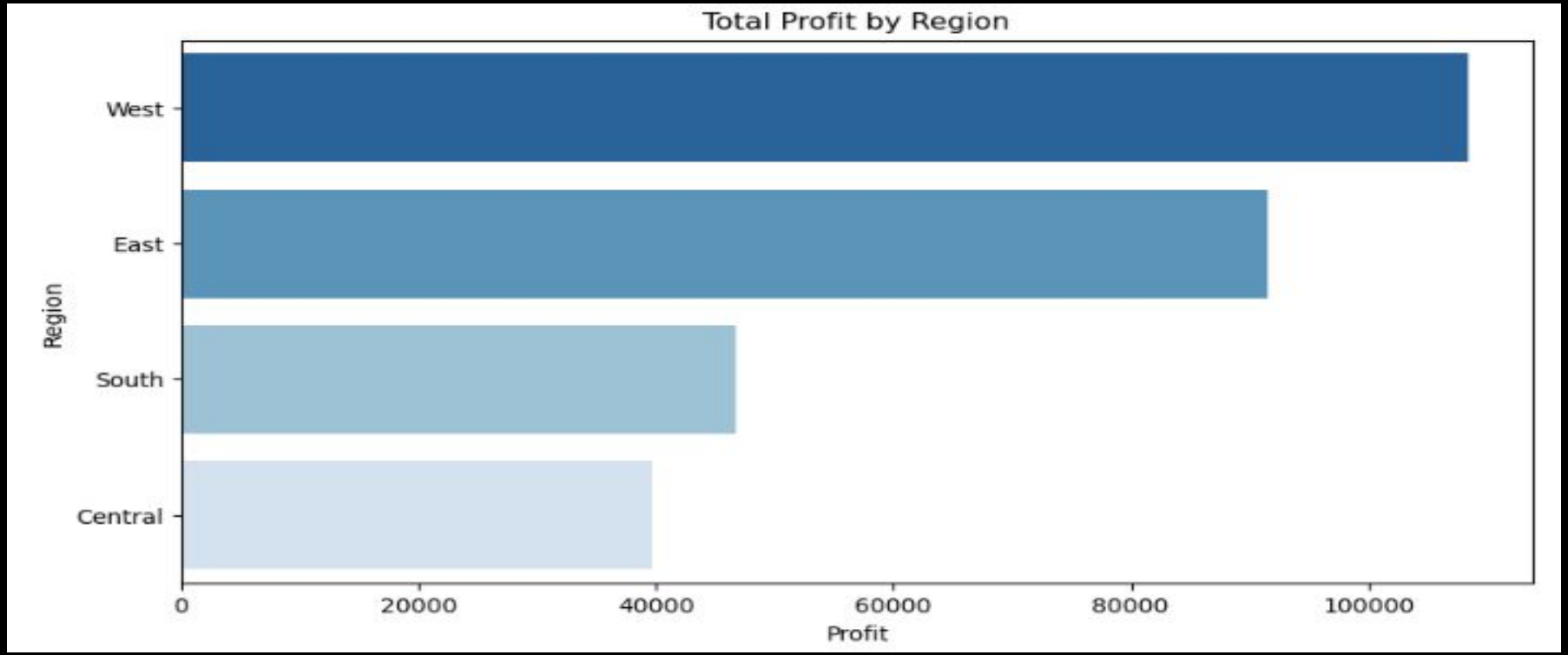
Key Findings



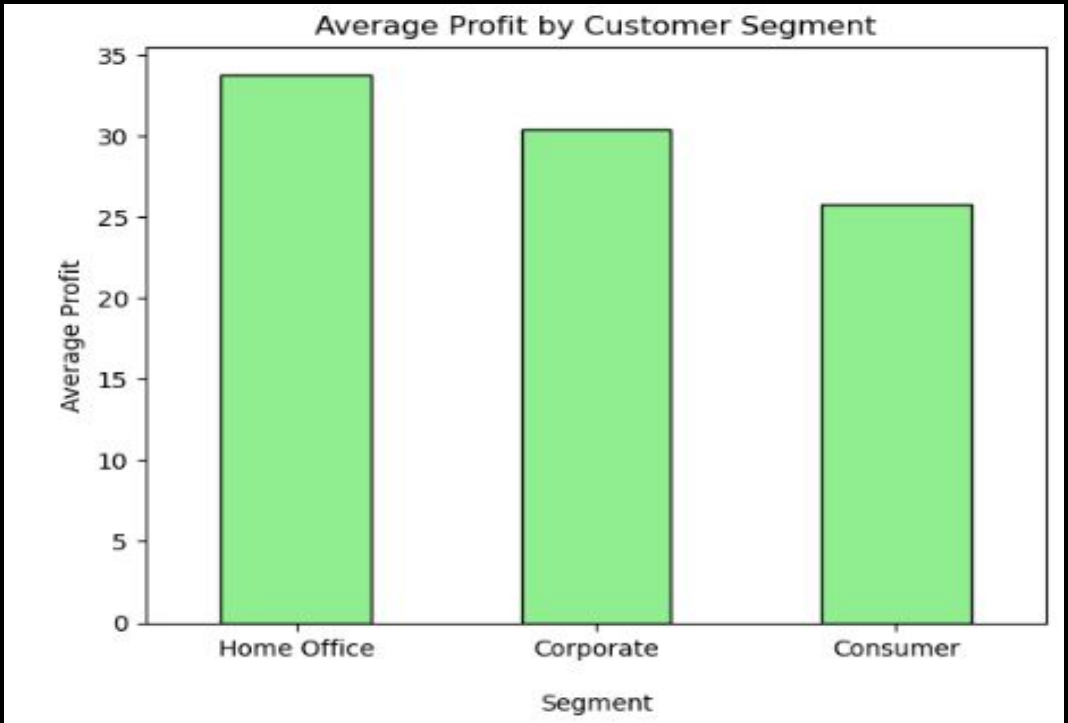
Key Findings



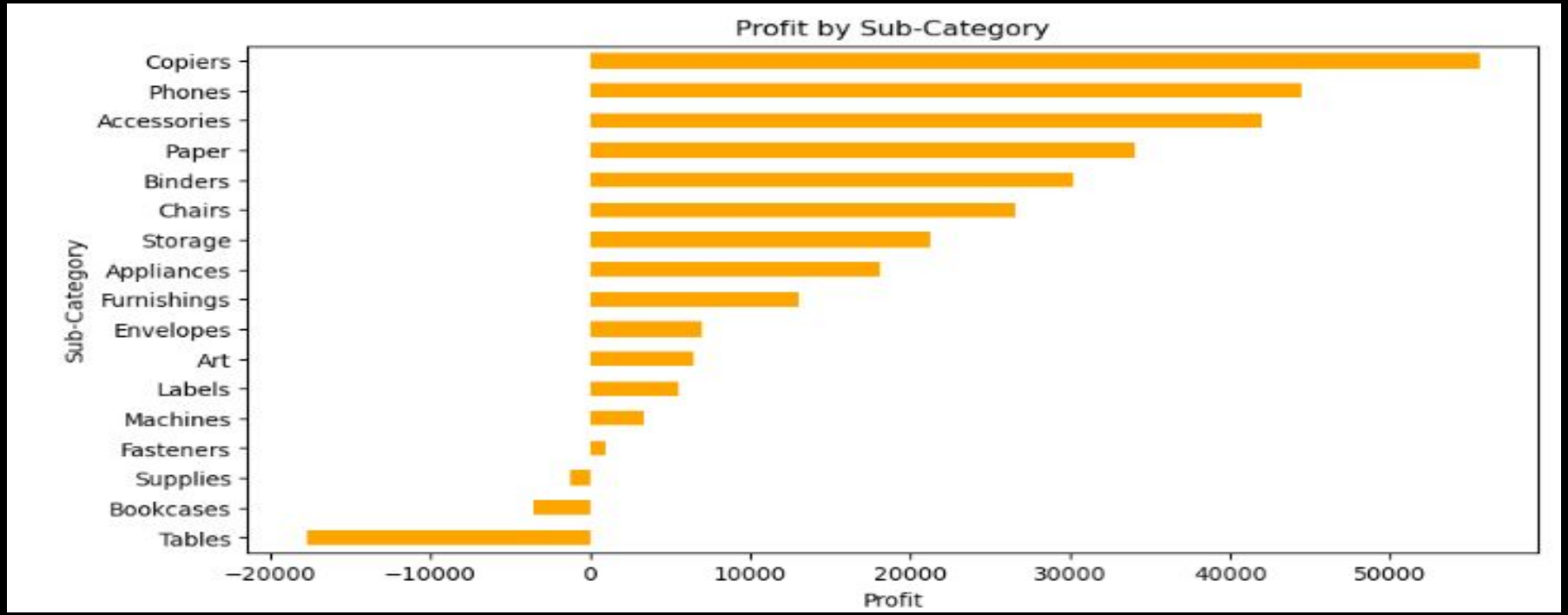
Key Findings



Key Findings



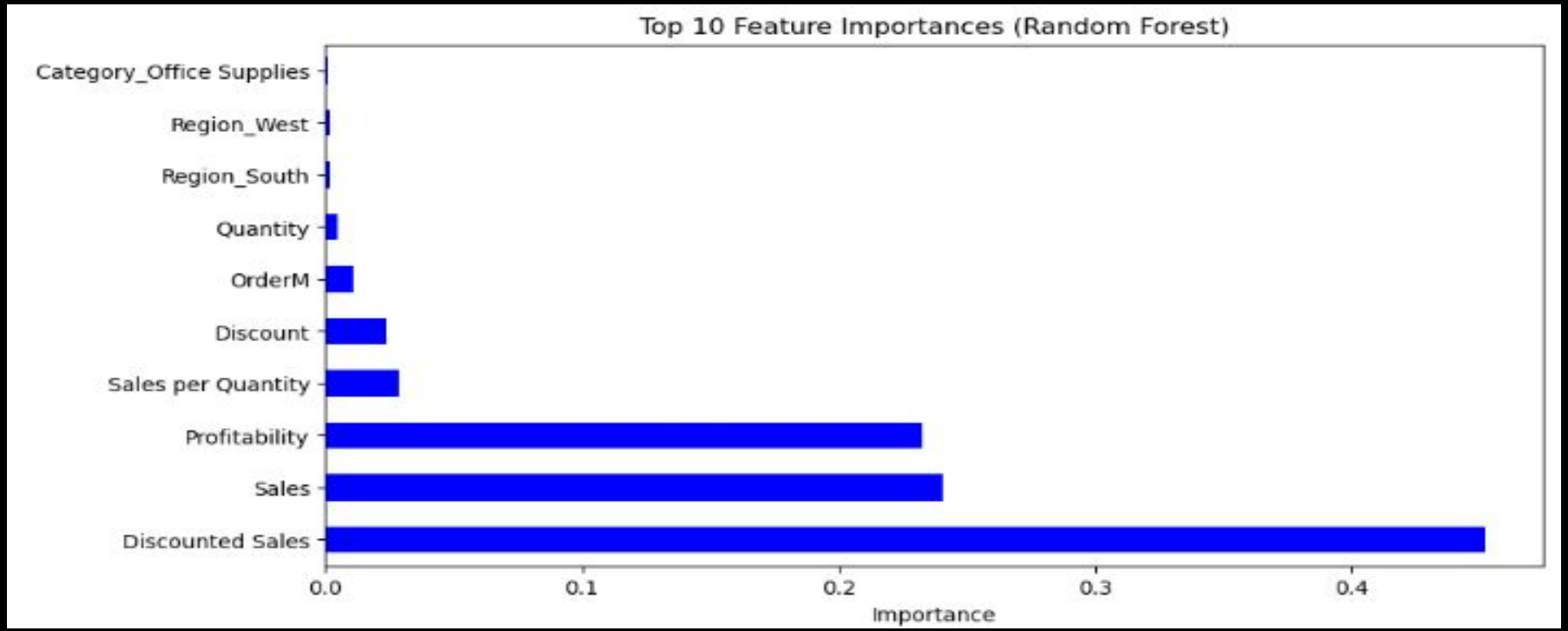
Key Findings



Modeling Analysis

- We wanted to find a prediction model that will help us increase profits in geographical areas and predict future sales trends in the next 30 days so we needed to find the important features in the dataset that are impacting our profitability.
- For our prediction model testing we decided to test Random Forest Regression and XG Boost since the sales data is large and has a lot of features. In the Random Forest model we used a Grid Search to perform hyperparameter tuning and find the best parameters for our model. We found that the best tree depth for our model was a max depth of 10.
- After training the data and fitting it to the model, the test data performance accuracy came out to be 94%. The Random Forest model helped us conclude that “Discounted Sales” was the most important feature impacting our profitability. The XG Boost model produced a result of 75% accuracy performance.

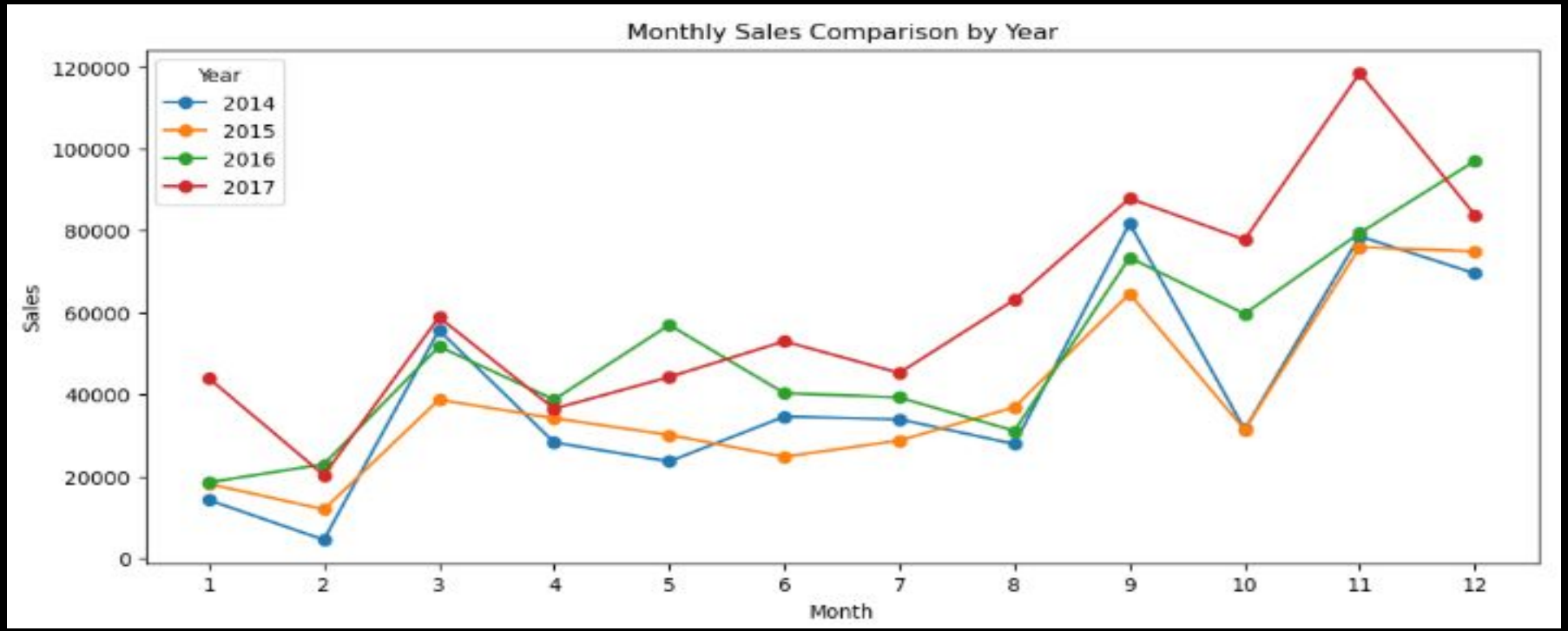
Modeling Results: Random Forest Model



Modeling Results

- Our goal was to produce a predictive model of 80% accuracy so our winning model is the Random Forest Model with a prediction accuracy of 94%.
- The XG Boost Model did not produce the results we were hoping for with a 75% accuracy.
- We ran a Cross Validation score and our Random Forest Model explains 70% of the variance in our target “Profit” on the test data which means our Random Forest Model is an efficient predictive model.

Superstore Monthly Sales Year by Year



Conclusion

- It is important to note that the Superstore is already operating profitably with sales increasing every year.
- The Random Forest Model was our winning predictive model scoring a 94% performance accuracy.
- After exploring the data the trend seems to be that more sales equals more profits and decreasing our heavily discounted items can boost profitability.
- Our most profitable customers are in the “Home Office” and "Corporate" segments so offering those customers a discount of 20% or lower can potentially produce more sales.
- If the business executives find the Random Forest Model to be useful in their decision making to increase profitability, we could automate the model’s use and create an excel plugin for efficient usability.