Applied SAP on Profitable Products Creation

Project “Rosemary’s Copier with Advance Wireless Print Technology”

Author: [Teresa Hernandez Ramos](https://calstatela.instructure.com/courses/99816/users/11199), [Rosemary Mosquera](https://calstatela.instructure.com/courses/99816/users/86045), [Long Zhang](https://calstatela.instructure.com/courses/99816/users/72822), [Hoa Pham](https://calstatela.instructure.com/courses/99816/users/13399)

California State University Los Angeles BUS5100-93: Introduction To Business Analytics

**Abstract:** The purpose of this paper is to present a data analysis on an existing data set for the most profitable products from Super Stores all across America. We will be able to extrude visualization data for all the Super Stores’ categories, customer-type purchases, returns and the areas where products are most profitable. We will then transform these datasets into projectable future sales (machine learning) through time series analysis which include linear regression and exponential smooth regression. With the framework provided above we will predict and analyze the frequent buying trends and target additional product creations to benefit the enterprises that are currently running by extending additional products within the Super Stores.

**1. Introduction**  
As a team of business analysts with aspirations to lead the future market in the U.S., we are constantly on the alert for emerging opportunities, particularly those identified through SAP SAC and other forums dedicated to business intelligence. Our enthusiasm for this venture is fueled by the pursuit of exceptional prospects across various channels, determining the optimal launchpad for our startup through strategic locations, and analyzing data trends to forecast future market directions. Our flagship project, the " Rosemary’s Copier with Advance wireless print technology," is a cornerstone of our ambition to develop innovative and impactful products within the market place.

Our approach is methodical and data-driven. We leverage diverse platforms and networks to potentialy breakthrough opportunities. The key to our strategy is the revelation of an outstanding opportunity we've identified. Our analysis extends in determining the ideal base for our operations, focusing on locations that offer not just logistical convenience but also align with our business objectives, facilitated by SAP's geolocation tools.

The "Rosemary’s Copier with Advance wireless print technology" represents more than a product. It enchances the beauty of analysis and copes the direction toward competitive market through the lens of business analysts.

2. Related Work

Generally speaking, data from Super Stores have been widely used in the community for projecting sales trends, analyzing customer segments, and applying strategies based on these insights. Most publications strategize their articles to give readers a comprehensive dashboard of insights about data through visualization, such as Bubble Charts, Bar Charts, Stacked Columns, and Scatter Charts.

One publication on Medium.com, authored by Mreza, conducted a study on the sales and trends of Super Stores from years 2014 to 2017. This publication showed a trend in purchasing based on the category of items at the superstore level by displaying data through pie charts and sales bar charts, highlighting categories and items that are among the top products. Compared to our findings, we present more projectable trends through time series analysis, regression, and exponential smoothing to test for precise risks in creating our products.

Similar to the first publication, the second publication also employed similar techniques for sales trend analysis. However, the main difference was that the second publication, authored by Swasti Khurana, applied projectable trends by analyzing future trends. Compared to our work, we also included which states are most profitable, and we will focus heavily on which states to target.

3. File Specification

The dataset from the distinguished analytics community, details sales data from a hypothetical superstore between years 2014 and 2017. This dataset has served as a fundamental resource for diving into sales trends, understanding customer preferences, and evaluating product performances, offering a solid foundation for data analytics projects.

Despite its richness, it's important to note the dataset's updates concluded in 2017, marking the end of its expansion. This cessation of updates means the analysis has stop for the current Fperiod, yet the dataset remains invaluable for historical trend analysis and educational purposes. Its size, while substantial, ensures a comprehensive exploration of business intelligence concepts, from predictive analytics to market research, within a manageable scope for analysis tools like Tableau.

The data size at 3.20 Megabytes with three different function is as follows: orders, returns, and people.

|  |  |
| --- | --- |
| Data Set | Size: Total 3.2 MGB |
| Orders | 3.19 MGB |
| Returns | 13.7 KB |
| People | 8.63 KB |

Figure 1. Data Sizes

**4. Data Cleaning**

Before uploading to SAP for further analysis, we tried various methods to obtain the location for each of our customer IDs. SAP rejected the geolocation for all of our customers' purchase data. Therefore, we had to organize and use Geocoding as an extension of our data cleaning process to generate the latitude and longitude for each customer base. As we uploaded our CSV file to SAP, we had to create a hierarchy for our customers' locations and the product hierarchy so that, when included in our analysis, it can be further broken down within the process to make it more clear.

**5. Implementation flow chart**

Initially, our plan was to directly upload the CSV file into SAP Analytics Cloud. However, we encountered an unexpected error during this process, necessitating a re-evaluation of our approach. To resolve this, the CSV file required reorganization and the addition of a few critical elements before we could successfully proceed with the upload to SAP. This preparatory step was essential to ensure the data was compatible with SAP's requirements, allowing us to fully leverage the analytical capabilities of the platform thereafter.

These refined models were then transitioned into a PowerPoint format, offering a visual and narrative representation of our analytical journey and findings.

A diagram of a software flowchart

Description automatically generatedFigure 2. Implementation Flow Chart

**6. Analysis and Visualization**

Through the development of our products, the sample data from superstores offers invaluable insights into how our establishment stacks up in a competitive market. This analysis leverages SAP Cloud Analytics for storytelling and utilizes time series predictive modeling to uncover trends. Our findings include identifying profitable states using geolocation data, forecasting future trends with regression analysis in time series data, and conducting a detailed trend analysis by product category within these superstores. This comprehensive approach allows us to strategically position our products for succes.

**6.1 Profitable Target States through Geolocation/Conditional Formatting**

Based on the geolocation data analysis conducted in SAP Analytics, we have located our target market strategy to focus on three highly profitable states: Washington, California, and New York. These states have shown a significant number of points indicating robust sales and potential for market growth, as seen in the visual representation below. Additionally, we have identified a neutral state like Georgia, which presents a steady, albeit less pronounced, opportunity for expansion. This targeted approach allows us to allocate resources more effectively and tailor marketing strategies to regions with the highest returns.

A map of the united states

Description automatically generated

Figure 3. Geolocations

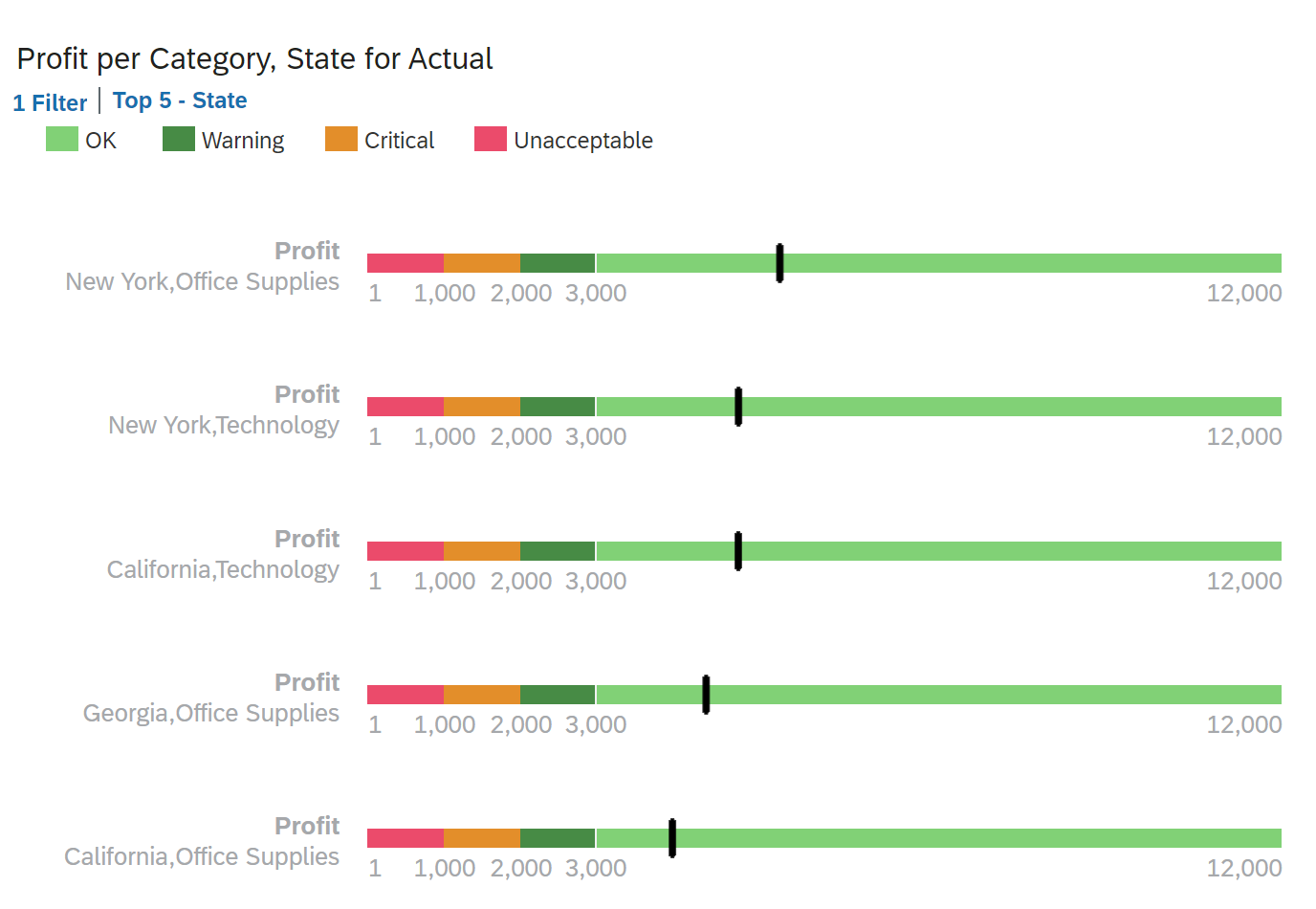


Figure 4. Conditional Formating Filter by States(Category Sales)

**6.2 Linear Regression/Triple Exponential Smoothing Forecast**

In the linear regression forecast, we can observe the trend of profit per order date for the actual recorded data. This form of predictive analysis allows us to establish a relationship between the time of purchase and profit, assuming a linear progression over time. The linear regression highlights key points and predicts future values assuming all of our stores are selling products in all of the categories, suggesting a steady increase in profits. The linear plotted forecast, showed a $52.18. This allow us to not only present our datas to buyers but also hedge us against market risk for super stores if the future prediction is correct, super stores will have a higher chance of profitability assuming the trend is not agaisnt their specific locaiton.

The Triple Exponential Smoothing forecast, is particularly adept at capturing seasonal variations and trends in the data. This method adds layers of smoothing to adjust for fluctuations in the dataset, offering a refined and more nuanced forecast than the linear model. The analysis suggests a more conservative estimate of future profits, as indicated by the forecasted values at $27.52. This model's profit has decrease significantly to the linear regression(-47.25%). However in the perfect market, this forecast would help us in preventing and assesses for risk situations significantly better than linear regression.

**A screenshot of a graph

Description automatically generated**

Figure 5. Linear Regression Forecast

**A graph with green and white lines

Description automatically generated**

Figure 6. Exponential Triple Smoothing

**6.3 Sales by Category**

The analysis by sales by category is an important factor for us in understanding which categories and products are trending the highest in our findings. With these trends being the cursor for our “Rosemary’s Copier with Advance wireless print technology” creation to the competitive market for Super Stores across America.

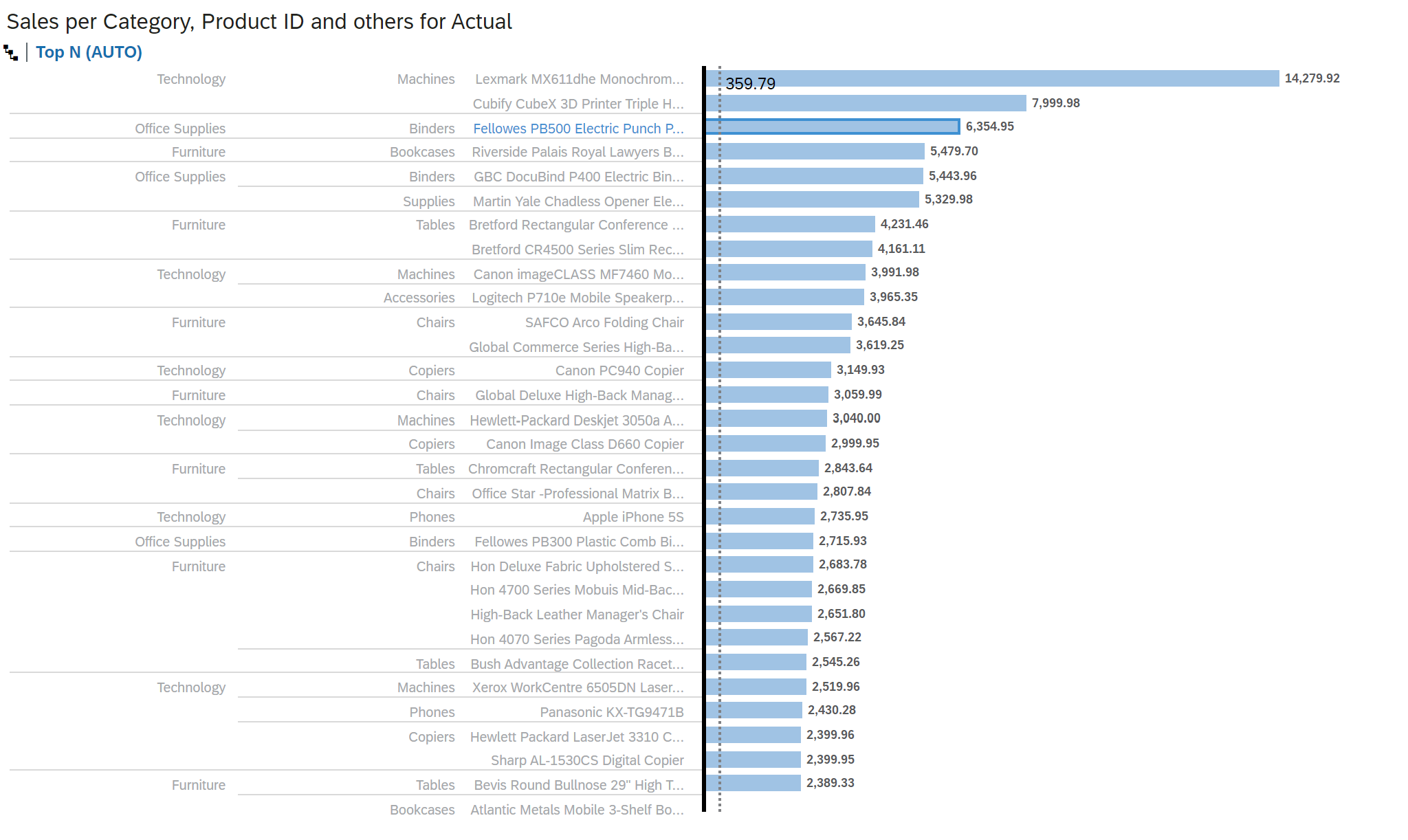
****

Figure 7. Sales By Categories

**6.4 Predictive Analysis- Time Series Trained Model**

Based on the predictive analysis we've conducted, the Mean Absolute Percentage Error (MAPE) has emerged as one of the key indicators throughout the analysis. Copiers are the only category with a MAPE of less than 50%. This perspective allows us to filter out categories with a MAPE higher than 100% for further analysis. Furthermore, a higher MAPE indicates lower accuracy in our predictive analysis.

In the datasets, most of the categories have a MAPE higher than 100%, indicating that focusing on a different product category might have led to failure, as MAPE tends to be biased towards negative integers in its analysis. At higher percentages, this suggests that these categories are most likely not profitable.

**A graph with lines and numbers

Description automatically generated with medium confidence**

Figure 8. Time Series Predictive Analysis Actual vs Forecast vs Trend

**A screenshot of a computer

Description automatically generated**

Figure 9. Mean Absolute Percentage Error.   
The data comprises the highest accuracy for Copiers (49.65%), Fasteners (107.33%), and Bookcases (116.07%). To gauge its accuracy, a MAPE higher than 20% is considered invalid. However, Copiers have the most accurate percentage worth reckoning with.

**7. Conclusion**

Throughout the analysis for the creation of “Rosemary’s Copier with Advance wireless print technology” we can conclude that:

1. The product should only be sold in Washington State, New York, California and expand to Georgia after a successful campaign.
2. Even though stores have massive inventory on products, choosing a copier machine with printing functionality might be the best bet as datas is leaning towards these two categories.
3. Launch date would have to be around October, as sales are starting to accelerate from October until March.

References

MReza. (2023, November 13). *Analyzing sales and profits trends at Superstore (2014–2017) using SQL and Looker Studio*. Medium. https://medium.com/@rezafahlopysd/analyzing-sales-and-profits-trends-at-superstore-2014-2017-using-sql-and-looker-studio-ebf045907b50

Khurana, S. (2021, January 18). *Superstore sales use case data analytics and visualization*. Medium. <https://medium.com/clique-org/superstore-sales-use-case-data-analytics-and-visualization-62afacd0777>

Tableau Community Forums. (n.d.). https://community.tableau.com/s/question/0D54T00000CWeX8SAL/sample-superstore-sales-excelxls