

Business Performance Evaluation (2015 -2018)

Name: Gilbert Ndekwa Simon

Institution: University of Alabama in Huntsville

Course: MSC 622 Analytics for Managers

Professor: Dr. Laird Burns

Executive Summary

The project presents a comprehensive analysis of the U.S. retail sales, a combination of two Kaggle datasets one being national sales dataset that includes the demographics of customers, the type of products, particulars of transactions and financial indicators and a U.S ZIP code dataset that provides the geographic positions to map the area more closely. The main goal was to define the sales distribution of the sales pattern, profits as well as the customer behavior across different regions and states. I managed to clean and combine the datasets in Excel and then import them into Tableau to build an interactive dashboard with bar charts and filled maps that help to visualize the trends in the region and the performance of the states. The analysis approach targeted on three key areas: comparison of total sale and profit within the West, East, Central, and South territories; cross tabulation based on state level purchase and profit to determine geographical groups with a high or poor performance; and the construction of how location affects the pattern of purchase and profitability. The methodology of the project reflects practical knowledge in data preparation, geospatial connection, and storytelling analytically, which is why the project will be a highly valued inclusion in a portfolio as a reflective ability to show technical prowess in business analytics.

The dashboard reveals that there is a high-level regional variation in performance. The West Region has the highest point of about \$17 million of total sales and the profitability is very high and constant, which explains the effective product mix and good operating margins. There is codependency between the East Region and the profit margin as it reveals the area has about sales of about \$15.7 million but with varied profit margin implications, which could indicate the difference in discount structure or cost model. The Central Region has average sales volumes, although the profitability in general is lower, and a number of states have negative profits indicating ineffective operation or non-profitable product classes. The South Region generated the least sales amounting to around \$9 million and had the greatest state-to-state profitability swings. The geographic mapping has revealed some apparent clusters of performance; high achievement states such as California, New York, and Washington record good performances in terms of sales and healthy margins whereas the Midwestern and some Southern states such as Texas, Tennessee, and Oklahoma represent good revenue but significantly tighter margins attributable to either high shipping costs, higher turnover rates, or product concentrations with low margins. States with negative probability are clearly visible on the map, and it is possible to find the areas that need intervention strategies rather quickly. The research successfully incorporates independent datasets to produce a single geographic sales model, which offers actionable insights to safeguard regional strategy and optimize product mix and their net profit. While there were challenges with utilizing ZIP code information to match to state-level data and to balance visual clarity with the quantity of information displayed, the resulting dashboard provides a concise, information-driven view of the sales performance that allows both academic and professional viewers to identify the high-value opportunity and non-performing markets, respectively, in quick time.

1. Introduction

In contemporary organizations, data analytics is a vital core in the context of understanding market dynamics, improving customer experiences and implementing profit-oriented strategies. The project creates a multi modal Tableau analytical system with real world U.S sales data on Kaggle and a national ZIP code coordinated dataset to map the country geographically. The central objective explores how customer demographics, product selection, and geographic differences influence sales and profit in the US.

The Sales Dataset of USA from Kaggle includes a wealth of sales information transacted within the dataset such as customer ID, products, order date, sales, profit, quantity and location down to the city and state of purchase. The ZIP Code List with Coordinates is an excellent supplement that adds wonderful geographic nuance to the ZIP Code List which allows for Tableau's spatial analysis to be spot on. The two combined create a means to assess consumer purchasing behavior in a more comprehensive fashion over time, by product category, customer segment and geographically. The intended audience for these dashboards was professors, classmates and potential employers to show an aesthetically appealing polished piece of work utilizing real world datasets. Combining analytics with visualization, the project provides how organizations can transform raw data into actionable insights to make pricing, marketing and operational decisions.

2. Methodology

2.1 Data Preparation

Data preparation and cleaning was done in Excel and the datasets were then imported into Tableau. All records were made to be consistent by using standardized formats of especially the date, state, categories and numerical fields. When data completeness was assessed, mismatched pairs of state city or the absence of ZIP codes was observed, and it may lead to poor analysis accuracy. The geographic consistency was to be verified to ensure that Tableau resolved state boundaries on a map successfully, and the dataset on sales was merged with the coordinates of ZIP codes to provide the possibility of visualizing the sales in a geospatial manner.

To identify not only the product revenue results but also the strength of sales and operational cost implications between and among the states, calculated fields such as percent difference measures and contribution margin were essential. These derived measures displayed more profound business patterns as they transformed raw transaction data into performance measures.

2.2 Dashboard Development

All the four analytical dashboards that were developed served a particular strategic purpose. The Sales Dashboard (2015-2018) has provided the seasonal trends, year-on-year, trend analysis, and KPI indicators. The Product Dashboard (2015-2018) emphasizes the relationships between sales and profit with the help of scatterplots to display, evaluates the influence of discounts, and focuses on the profitability of categories and subcategories. The Customer Dashboard (2015-2018) focuses on revenue concentration, profitability distribution, pattern of customer segmentation and repeat purchase of products. The Regional Dashboard (2018) utilizes the profit/loss distribution between all regions of the United States, performance diagnostics of states, and state geographic heatmaps. The visual consistency within the analytics suite was

ensured through the granting of similar color schemes to all the dashboards, where blue replying, and orange-red colors represented offered profitable outcomes and losses respectively. The dashboards had enhanced filters that made them more usable and interactive by allowing the user to zoom on specific parts, toggle years and find segments.

2.3 Analytical Modeling

Most of the modeling techniques incorporated in the dashboards transformed raw data into regional performance stories. Scatterplots revealed the relationship between sales and profit in order to establish structural inefficiencies whereas the line charts revealed seasonal relationship and time trends. Geographic choropleth maps were used to find regional performance clusters and had variations in profitability across states. Although the KPI summary cards displayed conveniently the total sales, profit, quantity sold, and profit ratio, bar charts classified categories, subcategories, and customers to display the best and the worst performers. Trend lines and indicators of comparative change showed abnormalities like a decrease in profit even though there was an increase in sales, which shows that there is need to do more research like finding out cost structures or discounting methodologies.

2.4 Interpretation

The integration of information on all dashboards helped create a cross-functional knowledge about the impact of customer behavior, category-level economics, and regional factors on business performance. The dashboards can be used as a basis of strategic suggestions as they not only show what has happened but also are used to clarify why trends have developed.

3. Dashboard Analyses

3.1 Sales Dashboard Analysis (2015–2018)

The Sales dashboard provided a chronological analysis of the sales, profit and quantity tendencies showing both prospects of growth and profitability. The sales increased by over 65% within the period of four years and order quantities grew by approximately 70%, meaning that the level of customer engagement and demand became bigger. However, Profit was very volatile, and there was a marked rise in profit in 2017 and a fall in profit in 2018 even though sales were at the high point in the entire body of data. Such a difference between the sales and the profit in 2018, shown in the form of line-chart trends, was a strong indicator of the growing profitability as in the year being studied, which could be primarily related to the discounting practices, as well as cost pressures not the sales volume.

The seasonality was identified as a very distinct trend in all the years, where there were a constant low performance of January and February and high sales in November and December. The month-by-month analysis on the dashboard showed that month Q4 generated a disproportionate amount of the annual revenue but the profit curve was not always following these sales peaks. This implies that any promotions made during holidays increased sales volume but compromised margin efficiency so that the growth in revenue and profitability trade-off. The main analytical takeaway is that the sales growth is volume-sensitive, not margin-sensitive, and the periods of heavy discount can be seen graphically as deviation of profit lines despite sales bars not falling.

3.2 Product Dashboard Analysis (2015–2018)

Product dashboards were applied using bar charts, scatter plots and stacked plots to establish the performance at the category level and the effect of discounting on profitability. Technology continued to demonstrate the most profitability and positive sales-to-profit conversion, which is a strategic strength of the organization. On the other hand, Furniture had high turnover of sales but low or negative gross profit especially in the sub categories such as Tables and Bookcases and Office Supplies were comparatively steady with low-margin performance.

Scatterplots were used to depict the structural aspect of the problems of profitability in that they observed how sales volume and the profit margin were connected to each other. Phones, Accessories and Machines were concentrated in the high sales and high profit quadrant, and Tables and Bookcases made it into the negative-profit areas in all years, so the fact that some sub categories are always lowly performing as a result of high discounts or being expensive to serve was observed. Corporate customers gave greater profit per order with less overall sales when the segment filters were used and Consumer customers responded to discounts and gave volume. This effect of the segmentation became visible as dashboard filters were turned on and off, and the distribution of profits among the types of customers was shown to have changed.

3.3 Customer Dashboard Analysis (2015–2018)

The Customer Dashboard gave customer level data in the form of rankings, scatterplots, density charts, and segment filters to analyses revenue distribution and profitability trends. The number of customers had been stable, between the 567 and 635 per year, but increased volumes of orders year by year pointed to a high level of repeat business and customer loyalty. The concentration of revenue was presented dramatically with Pareto-style bar charts and cumulative line visualizations that indicated that around 20% of the customers generated around 60% of the total revenue with the top 10 customers being similar to the findings in several consecutive years.

Nevertheless, scatterplots created a crucial revelation: a great number of high-sales customers were below or along the zero-profit line, meaning that high sales are not always the same as high profit and that transactions high on discounts debase margins. There was a significant variation of profitability at the individual-customer level with scatterplots showing that customers varied extremely in their profitability even when the volume of sales that were made was similar. Segment filters were applied, and the result was that, with the use of segment filters, there was a movement of more points towards positive profit bands with the Consumer customers demonstrating heterogeneous distribution with more points towards negative-profit segments, suggesting a need to implement specific marketing and tailored pricing policies to consider segment specific behavior.

3.4 Regional Dashboard Analysis (2018)

The Regional Dashboard merged sales records and ZIP code location to offer a more detailed look at the revenue distribution on U.S. states using choropleth maps to present geographic performance patterns. Tableau maps showed great contrast during U.S. areas, where West and East states were tinted on dark-blue color that signified high sales and revenue, whereas Central and Southern areas were colored on light or orange color and signified worse performance or losses. State-level map labels showed California, Washington, New York, and Florida to be highest profit center and Texas, Arizona, and Colorado to be good sales but low or negative profit centers evidently as blue sales labels with orange profit labels.

Some of the Southern states recorded consistent losses indicating inefficiencies in logistics, customer discount behavior, or local competition issues. As the map visualization showed, geography can influence the way profitability is developed in significant ways high-revenue states do not always translate into high profitability and the region cost structure, the level of competition, or transportation distance, play significant roles in raising the difference in the margins. This geographic analysis identified regional opportunities in the area of improvement of operations and price.

4. Integrated Key Findings

The integrated dashboard analysis in all the four analytical tools indicated some cross-functional observations that reveal basic business challenges and opportunities. The discounting aspect turns out to be the dominant factor in profit erosion and analysis findings reveal that time with excessive promotion activity always reap in tandem profits even with sales levels going up. The pattern is evident in all customer segments and time periods, which means that there is a systemic overdependence on discounting as a driver of sales on a systematic basis that will hurt the performance of the margins.

Furniture has continued to underperform in all time period and all customer groups, pulling overall profit downward and indicating structural problems in cost-to-serve, pricing strategy or product-market fit. Subcategories such as Tables and Bookcases reoccur in the negative-profit zones which means that these products should be passed to a more crucially important reconsideration than a fringe-adjustment.

Customer profitability is highly uneven and concentrated where 20% of the customers provide 60% of the revenue and most of the high-volume customers work just below or across the zero-profit mark. This is a risk and an opportunity in the sense that, on the one hand, it is important to retain top customers, and on the other, the interaction methods with lower-ranking customers are different. The customer level profit scoring and segmentation according to the analysis, however, allows predicting profitability based on sales volume alone is not possible.

Geographic differences demonstrate the high level of inefficiencies in working of the companies, and certain states are always unprofitable although their sales volumes are reasonable. Even the high-revenue states such as Texas and Arizona do not always deliver the proportionate profits implying that the regional cost structures, competition pressure, and logistics inefficiencies need to be investigated and handled regionally.

Growth in sales is high at 65% over the four-year period but the profit margin does not keep pace with that growth which means that there is a basic misalignment between volume growth and profit management. This deviation became more pronounced in 2018, as sales were at the highest level but profits fell, showing that current growth strategies are more about volume than profitability.

The seasonality is strong and annual performance is mostly dominated by Q4, however profits are not always following the spikes of sales. Promotion intensive during peak season causes the loss of profitability in favor of volume, which causes quarterly volatility that makes it difficult to forecast and plan operations. The stability of this trend in terms of years indicates the use of counterproductive promotional efforts

5. Recommendations

5.1 Optimize Pricing and Discounting

In order to narrow the pricing and discounting tools, the introduction of discount ceilings within certain product lines will be useful in avoiding a decline in the margin. It is recommended to minimize the application of a broad-based discounting in the months of holiday seasons and integrate customer-specific discounting policies that process score according to profitability. This dual approach will assist in sustaining sales and save margins.

5.2 Improve Product Strategy

When making product strategy improvements, it is important to reconsider unprofitable Furniture subcategories. It is recommended that more emphasis should be laid on the investments in Technology categories which are more profitable. Moreover, the research on the supply chain expenditure on supply items that make losses can show areas where expenses can be reduced.

5.3 Enhance Customer Segmentation

In customer segmentation, it is advisable to introduce loyalty programs to the top customers and target the customers with potential growth who are the mid-tier ones. Segmenting the marketing plans based on the segments of the consumer, corporate and home office will make sure that they are matched to the different profitability rates of these segments.

5.4 Develop Region-Specific Approaches

The creation of region-specific strategies would entail exploring the problem of logistics and the cost-to-serve in low profit state which would be used to formulate specific regional pricing outcome. It should also invest in the underpenetrated regions with high potential so as to grow.

5.5 Boost Operational Efficiency

Lastly, increasing operational performance needs to be achieved through alignment of inventory and staffing according to the seasonal performance, optimization of service routes, and quality of service. These measures will lift customers to a higher level of satisfaction and greatly cut down costs that will facilitate the profitability.

6. Conclusion

This project demonstrates that integrated data analytics, enabled by Tableau visualization, can be effectively used to extract meaningful business information out of real-world sales data. The dashboards offer a multi-layered view of customer behavior, performance of different products, and regional inequalities nationwide because of trying to merge sales reports and ZIP-code geography. The results reveal the following critical structural issues: the loss of margins through the process of discounting, disproportionate profitability of various states and product segments, dependence on a limited number of key customers. Meanwhile, the analysis has shown that there is a robust increase in sales and steady consumer demand, which provides substantial strategic opportunities. Ultimately, the project also demonstrates graduate critical thinking, where one can clean and combine data, visualize it into interactive dashboards, debug business issues with visual analytics, and make actionable data-driven recommendations.

7. Bibliography

- Ahmed, S. (2024). *Sales Dataset of USA [Updated]*. Kaggle.
<https://www.kaggle.com/datasets/sulaimanahmed/sales-dataset-of-usa-updated>
- Tableau Software, LLC. (2024). *Tableau Desktop*. Version 2024.
- Wible, J. (2023). *US Zip Code List with Coordinates*. Kaggle.
<https://www.kaggle.com/datasets/jedwible/uszipcodes-20231227>
- Youtube, <https://youtu.be/dahrmqT5GD4>