**Optimizing Retail Supply Chains Insights from Sales Data Analytics**

Under the Guidance of Professor : Michail Xintarakis

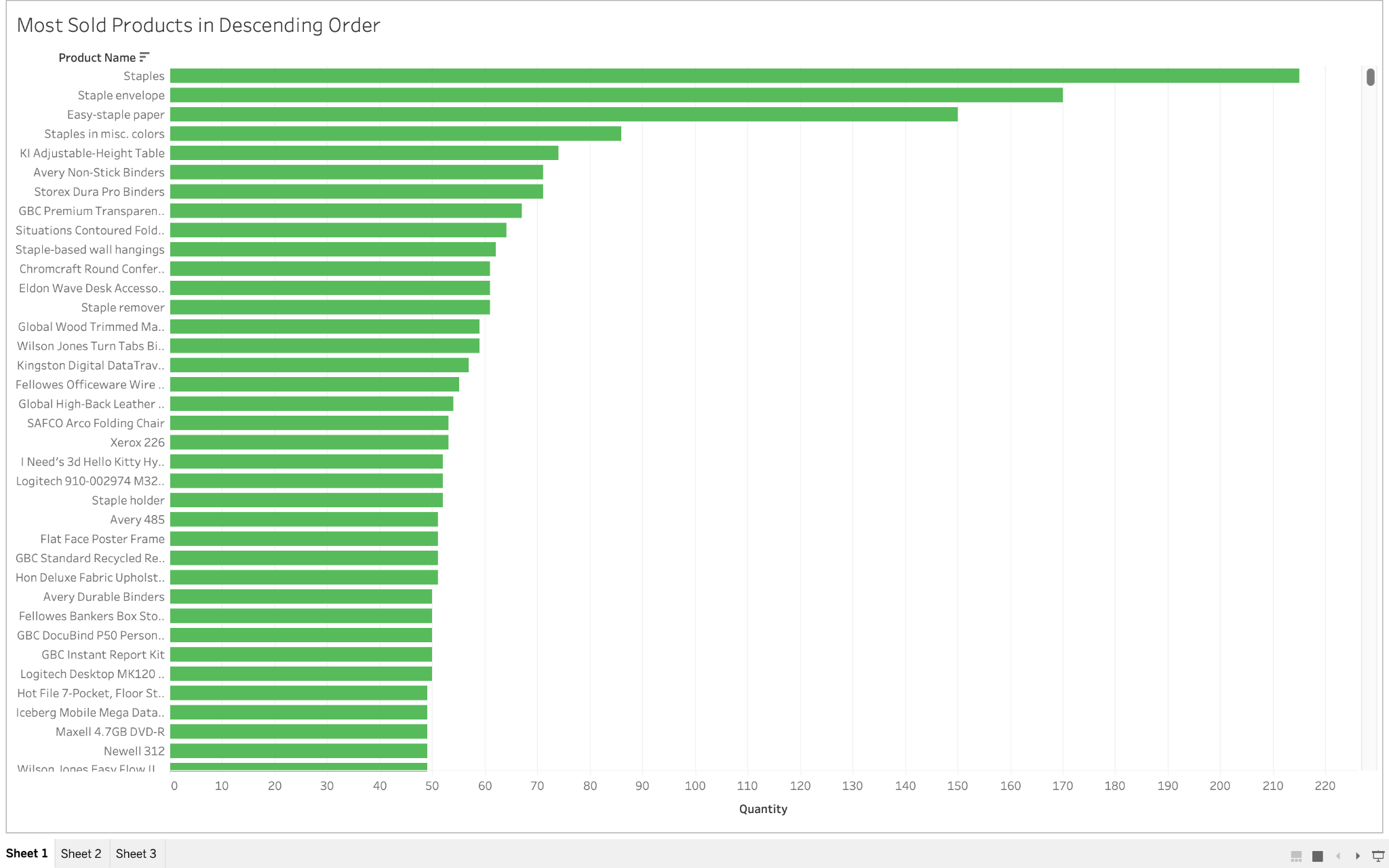
Submitted by: Rusheel Meka, Sid Gowda

Dataset used : <https://www.kaggle.com/datasets/shandeep777/retail-supply-chain-sales-dataset>

**ABSTRACT**

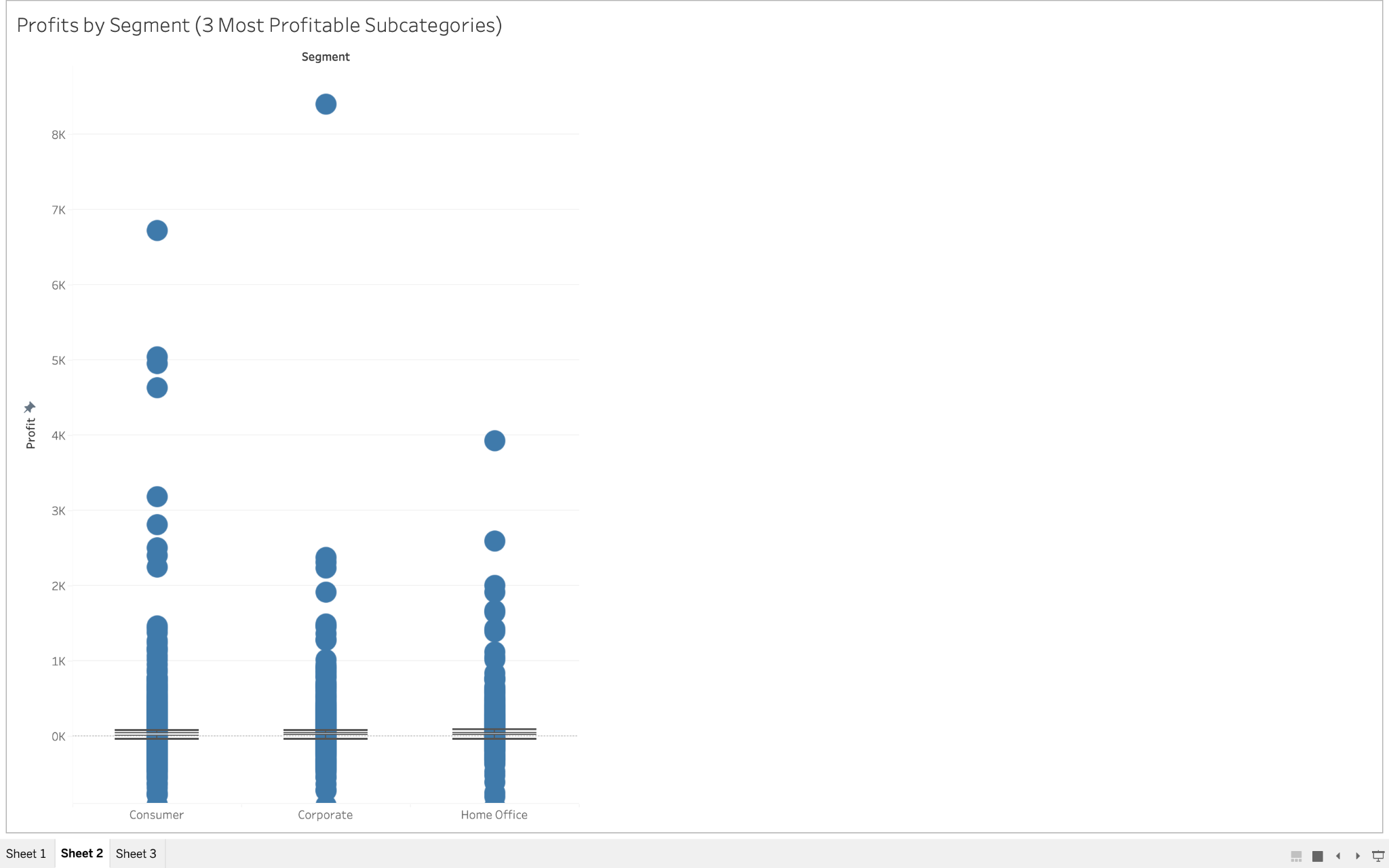
The Retail Supply Chain Sales Dataset offers comprehensive insights into retail sales performance, customer demographics, and shipment logistics. Consisting of 9,994 rows and 23 columns, this dataset provides granular details about individual orders, including order and shipment dates, sales metrics, and product specifics. Key attributes include Order ID, Sales, Profit, and Discount, alongside geographic information such as Region, State, and City. Customer data, including Customer ID, Segment, and assigned Retail Sales People, allow for targeted analysis of purchasing behavior. The dataset also incorporates product information, categorizing items by Category and Sub-Category, while the Returned column highlights product returns. With its rich time-series potential and diverse variables, this dataset is ideal for visualizing trends, analyzing sales performance, and identifying factors driving profitability. The project leverages data visualization techniques, including boxplots, bar charts, and interactive dashboards, to uncover insights into sales trends, product performance, and regional differences.

**BAR CHART**



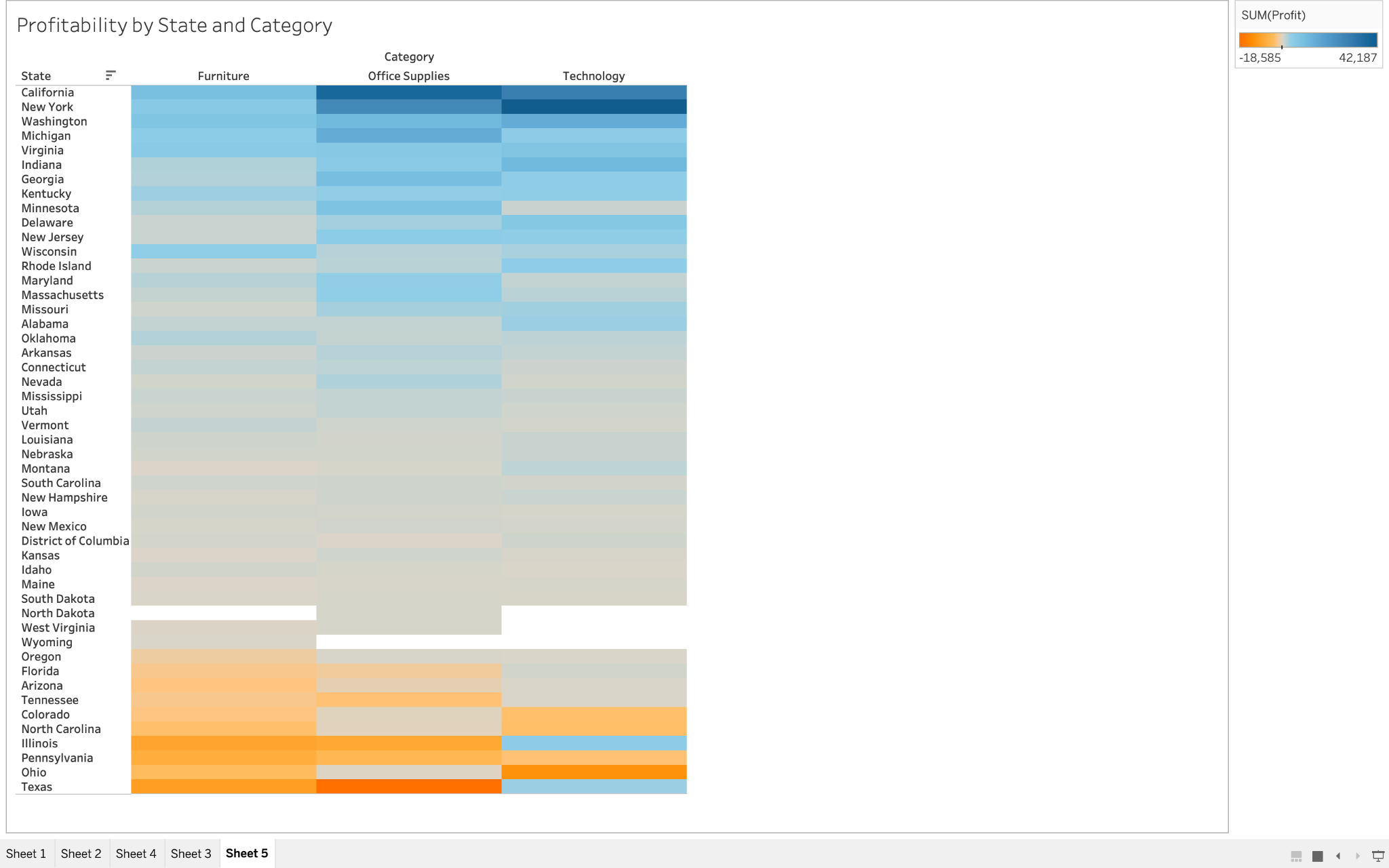
This bar chart highlights the top-selling products, with staples leading by a significant margin, followed by staple envelopes and easy staple paper. The high demand for these items can be attributed to their essential nature in office and home environments, their low cost, and frequent need for replenishment. Staples and related products are consumable, inexpensive, and often purchased in bulk, making them popular among businesses, schools, and households. Additionally, their ease of storage and distribution ensures accessibility for customers. This trend reflects a strong demand for everyday office supplies, indicating opportunities to optimize inventory and bundle related products to drive further sales.

**BOX AND WHISKER PLOT**



The box plot analyzing Profit across the Consumer, Corporate, and Home Office segments reveals key insights. Interestingly, all three segments share a median profit of 9, which is highly unlikely and suggests potential rounding or data distribution quirks. While the distributions are fairly similar across segments, Consumer exhibits numerous data points in the high thousands, indicating a higher mean profit despite the median. This suggests that while most profits are modest, significant high-profit orders pull the mean upward. In contrast, the Corporate segment shows the largest single outlier, suggesting occasional extremely high-profit transactions. Overall, while all segments display similar profit spreads, outliers and high-value transactions notably influence the Consumer and Corporate segments.

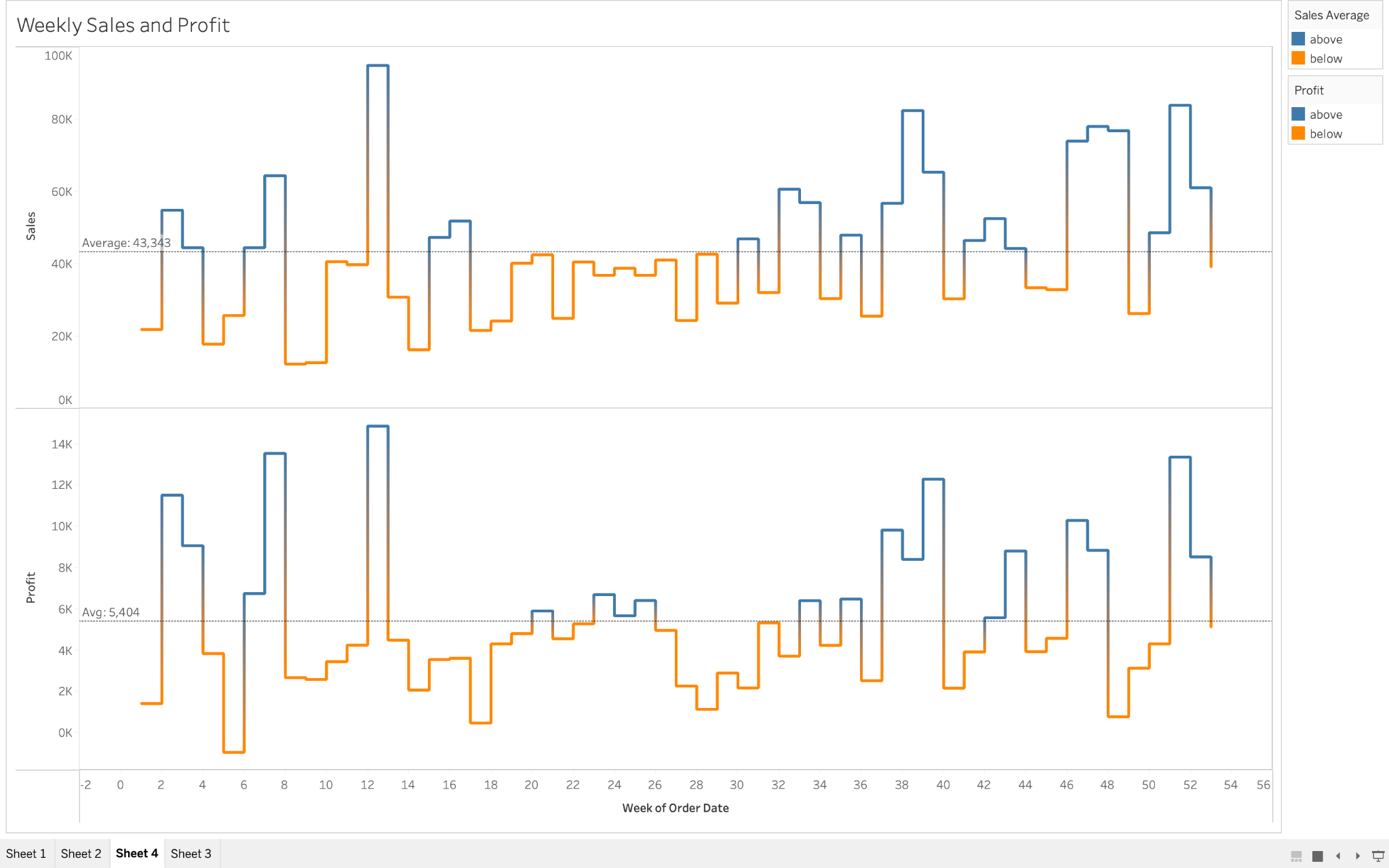
**HEAT MAP**

This heat map provides a detailed breakdown of Profit by State and Category, ranked in order of total profitability per state. The color scale ranges from dark orange (least profitable) to dark blue (most profitable), allowing for quick identification of trends. California stands out as the top-performing state, particularly excelling in Office Supplies and Technology, indicating strong demand and efficient operations in these categories. Similarly, New York achieves the highest profit in the Technology category, likely driven by high-value sales in this segment.

Moving down the heat map, we observe a growing presence of gray areas, indicating states with neutral or mixed performance across categories. For example, Illinois struggled significantly in Furniture and Office Supplies but managed to perform fairly well in Technology, suggesting that certain product categories may be more resilient in specific markets. In contrast, Pennsylvania consistently underperformed across all three categories, reflecting systemic challenges in profitability within the state.

Texas emerges as the least profitable state overall, driven by massive losses in Office Supplies, which overshadow its moderate success in Technology. This indicates that while Technology sales offer some cushion, substantial losses in other categories can drastically impact a state’s total profitability.

Overall, the heat map reveals clear geographic disparities in profitability, highlighting opportunities for targeted strategies, such as improving category performance in underperforming states or capitalizing further on high-performing categories in states like California and New York.

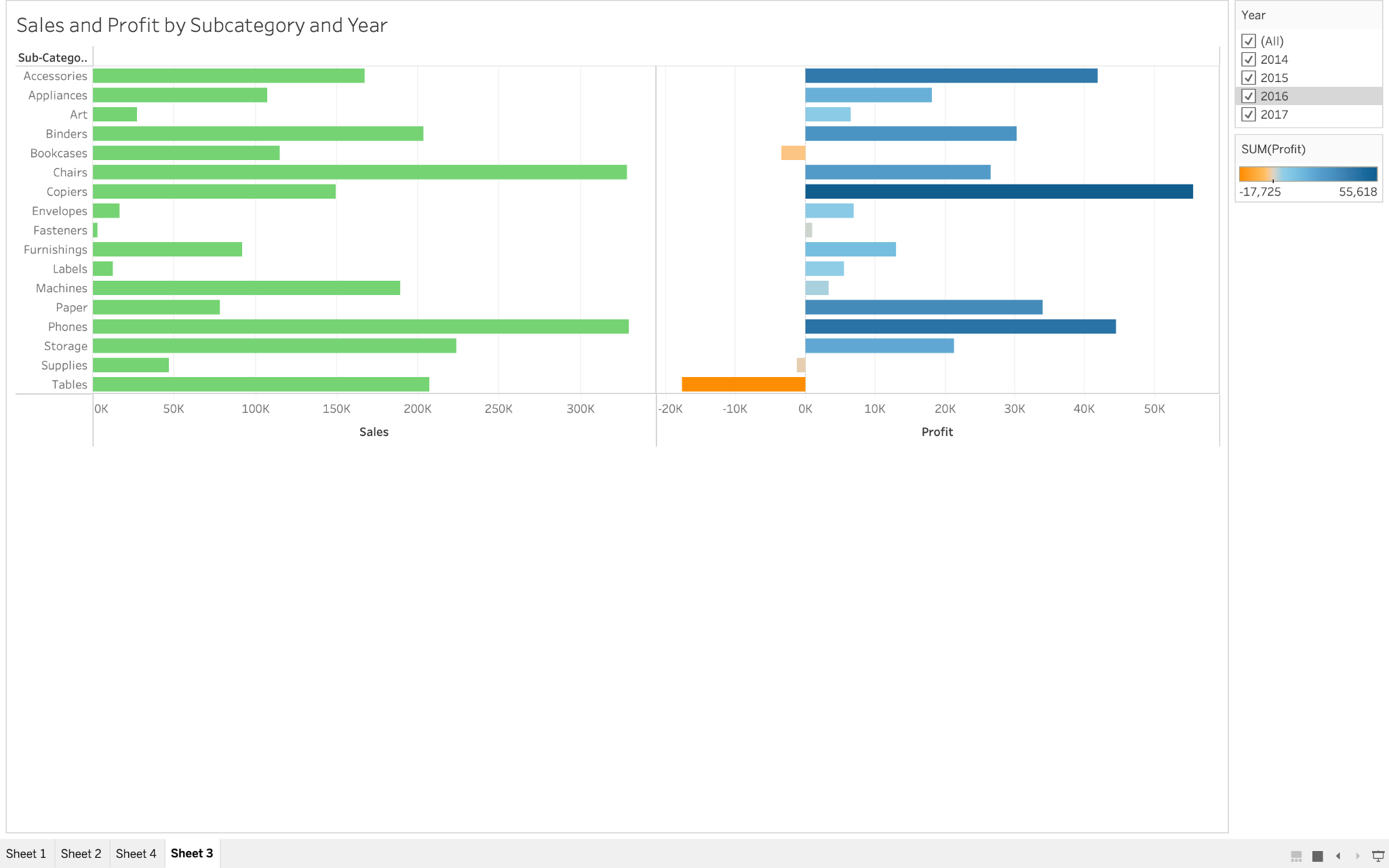
**STEP CHART**

This step chart, which tracks Profit and Sales over the weeks (2014–2018), provides valuable insights into seasonal performance trends. The reference lines for averages help highlight deviations across the weeks. A key observation is the divergence between Profit and Sales in specific periods. For example, during weeks 15–17, while sales remained positive, profits dipped below average, suggesting high sales volume but thin margins, possibly due to discounts, promotions, or low-profit product sales. Conversely, in weeks 22–26, sales were below average, yet profits climbed above average, indicating that fewer but more profitable transactions occurred, likely driven by higher-margin products or cost optimizations.

The store’s most profitable period appears to be week 12, where both profit and sales exceeded their averages, hinting at strong demand and operational efficiency during this time. However, week 5 stands out as the period with the lowest profit and the largest losses, which could be attributed to factors like poor demand, excessive costs, or inventory inefficiencies early in the year.

Overall, this step chart reveals important patterns, such as periods of disconnect between sales and profit, highlighting opportunities to analyze pricing strategies, product mix, and cost control measures during critical weeks to maximize profitability year-round.

**SIDE BY SIDE CHART (WITH INTERACTIVITY)**

This side-by-side bar chart with interactivity provides a dynamic view of Sales and Profit across different Subcategories, with the ability to filter by year or compare multiple years. The color-coded legend (dark blue for high profit and dark orange for heavy losses) highlights performance trends that are both consistent and concerning across subcategories.

A notable trend is the performance of Tables, which consistently generate high sales year over year but result in significant negative profit. This discrepancy suggests issues like high production or shipping costs, low margins, or frequent discounts that erode profitability despite strong demand. Addressing these inefficiencies could transform Tables into a more profitable subcategory.

In contrast, Copiers stand out as a high-performing subcategory despite relatively low to modest sales. The consistently high profits suggest that copiers are sold at strong margins, likely due to their premium pricing or lower associated costs. This subcategory represents a clear opportunity to further capitalize on a reliable source of profit by increasing sales volume through targeted promotions or bundling strategies.

Bookcases, on the other hand, show consistently high sales but tend to result in a slight loss. This performance indicates that while demand for bookcases is strong, low margins or operational costs are limiting profitability. A strategic price adjustment or focus on cost optimization could help make this subcategory more profitable without sacrificing demand.

Overall, the chart emphasizes a clear need to analyze the profit-to-sales relationship across subcategories. Products like Tables and Bookcases reveal inefficiencies that should be addressed, while Copiers highlight opportunities for leveraging strong margins. By analyzing trends over time with the interactivity provided, decision-makers can identify underperforming areas and optimize strategies to maximize overall profitability.

**DASHBOARD**

