**DATA VISUALIZATION: LAB 3**

**Name: Anushareddy Ramachandra Reddy**

**Module7:**

**Question:** Are there any noticeable differences in the Shipping Cost-Profit relationship across the customer segments (colored clusters)?

A screenshot of a computer

Description automatically generated

**Explanation about the visualization:**

The graphic displays a scatter plot with colored clusters that illustrate the relationship between profit and shipping costs for various client categories. With the x-axis representing the shipping cost and the y-axis representing the profit, each data point on the scatter plot represents a distinct transaction. To facilitate the identification and comparison of trends among different client categories, the data points have been colored according to their respective segments. The color coding for the various clusters or customer segments is shown in the legend in the upper right corner.

The downward-sloping distribution of data points initially indicates a general pattern wherein increased shipping costs are linked to decreased profitability. This negative link makes sense because increased shipping costs generally reduce total business margins.  
  
On closer examination, though, it seems that the distribution and dispersion of data points among the various consumer segments varied noticeably. Cluster 1 (shown in blue) appears to have a more focused distribution, for example, with data points grouped around a certain range of shipping costs and profits. This may indicate that the delivery cost-profit relationship for this specific consumer category is more stable or predictable.

Contrarily, Cluster 4 (shown in orange) shows a more evenly distributed distribution of data points, suggesting a higher degree of fluctuation in the shipping cost-profit dynamic for this particular customer base. While some data points in Cluster 4 have lower shipping costs but also lower earnings, others have quite high shipping costs but still manage to maintain respectable profit levels.  
  
The distributions of Clusters 2 and 3 (shown in green and red, respectively) seem to lie halfway between the extremes of Clusters 1 and 4, showing a considerable degree of variability in the shipping cost-profit relationship.

The visualization also makes it possible to spot any anomalies or possible outliers within each client group. For instance, a few data points in Cluster 3 indicate both relatively high shipping costs and large profits; these observations may call for additional research to determine the underlying causes of this peculiar combination.  
  
All things considered, this graphic successfully draws attention to the variations in the shipping cost-profit relationship between client categories, offering insightful information for future optimization or customized tactics. Through the examination of distribution patterns and variability within certain segments, enterprises can enhance their comprehension of the influence of shipping expenses on profitability and devise focused strategies to efficiently handle these dynamics.

**Module 8: Dashboard**

**Question:** In what ways does the geographic distribution of sales and profit alter when a different time period is chosen for different category?

A screenshot of a computer

Description automatically generated

**Explanation about the visualization:**

The dashboard representation that is supplied sheds light on how, for different product categories, the geographic distribution of sales and profit varies over different time periods. It offers several interactive elements that make it possible to investigate these processes.   
  
The distribution of profits by state is shown on the map to the left. The profit level of each state is represented by the color intensity, with darker hues denoting higher earnings. A summary of the states that contributed the most to total profits over the chosen time period is given by this geographic depiction.

The sales numbers for each state are displayed in a bar chart labeled "Sales by States," which is located just beneath the map. One can determine any possible disparities or relationships between sales volumes and profitability across various regions by contrasting the profit map with the state-level sales data.  
  
Based on the x-axis labels, the "Profit Over Time" line chart in the bottom right corner shows the variations in profit over a given time period, which looks to be from 2013 to 2017. This graphic lets viewers see how profit levels have changed over time. It may show seasonal changes or other temporal patterns that have an impact on corporate strategy.

The "Category (set actions)" function is one of this dashboard's most intriguing features. With the help of this filter, users may choose particular product categories and see how the visualizations change to show the data for those categories. Users can spot possible variations in the geographic distribution of sales and profits across different product lines by switching between different categories.

When the "Office Machines" category is chosen, for instance, data unique to that category will be displayed on the map, sales bar chart, and profit line chart. This could provide important information, such as the states or areas that do exceptionally well or poorly in terms of office machine sales and earnings. By comparing these trends across categories, tailored marketing campaigns, inventory control, and price decisions that take into account local demand and preferences can all be improved.

This strategy may reveal insightful information like which regions show seasonal variations in product category profitability or demand. Additionally, it might draw attention to prospective chances for focused advertising campaigns or stock modifications depending on the temporal trends seen in various markets.

All things considered, this dashboard visualization provides a thorough and interactive means of examining the interactions amongst product categories, geographical areas, sales volumes, earnings, and time intervals. Users can gain important insights that could guide resource allocation, strategic decision-making, and focused marketing campaigns catered to particular geographic areas and product lines by utilizing the many visualizations and filters available.

**Module 9:**

**Question:** What is the relationship between the number of bathrooms and various sales kinds in relation to the distribution of property sizes (in square feet) among various address clusters?

A screen shot of a graph

Description automatically generated

**Explanation about the visualization:**

The graphic shows the correlation between the quantity of bathrooms and the two types of sales (real estate) for a range of square footage property sizes, categorized by address clusters.   
Plotted on the same axis, the chart consists of three line graphs combined. The y-axis shows the property's square footage, and the x-axis shows the number of bathrooms. A particular kind of sale is represented by each line graph

The general trend of property size growing as the number of bathrooms increases is the first noteworthy pattern. The foreclosure and realtor sales lines, which exhibit a consistent upward tendency, both demonstrate this. The assumption that larger properties typically have more bathrooms is supported by this pattern.  
  
Nonetheless, there are differences in the distribution of property sizes between foreclosure and realtor sales for a given number of bathrooms. The foreclosure sales typically have a greater range of property sizes for properties with fewer bathrooms (around 1-2 bathrooms), with some properties being noticeably larger than their counterparts in realtor sales. This implies that foreclosure residences with fewer bathrooms could come in a variety of sizes, possibly indicating different situations that resulted in foreclosure.

The property size distributions for foreclosure and realtor sales converge and both lines follow a similar rising trajectory as the number of bathrooms increases (about 3–4). This convergence may suggest that the size differences between foreclosure and realtor sales become less noticeable for pricier residences with additional bathrooms.this graphic successfully depicts the connection between the size of the property, the number of bathrooms, and the manner of sale, offering insights into the distribution patterns and possible distinctions between sales through foreclosure and those through realtors in different address clusters. It makes it possible to spot patterns, anomalies, and places that can benefit from more research or study.

There where 3 Trend Lines in the above visualization.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Module 10: Sales by Category (Lollipop Chart)**

**Question:** Using a lollipop chart, do categories stand out as exceptionally strong or bad performers relative to the rest?

A screenshot of a computer

Description automatically generated

**Explanation about the visualization:**

A lollipop chart is used to show the sales performance of several categories. A horizontally oriented circle marker, the size of which is determined by the total sales for each category, is used to symbolize each one. With the biggest circle marker, which indicates noticeably higher sales than other categories, the "Office Supplies" category clearly stands out as the best performer from the visualization.   
  
However, the comparatively smaller circle markers in categories such as "Tables," "Phones and Tablets," and "Telephones and Communication" indicate worse sales performance in these areas.

The lollipop chart makes it simple to compare the marker sizes (sales numbers), which effectively illustrates the relative performance of each category. In the context of sales, categories with larger marks outperformed those with smaller markers in terms of sales numbers.  
  
Overall, this representation makes it simple and easy to determine which categories, based on sales data, are performing well and poorly. This makes it possible to quickly discover possible areas that could benefit from additional research or strategic focus.

By using visualization's text and color features to provide further details or context for each group. Color is used to visually distinguish poor performers from good performers by highlighting the former in one color and the latter in another. The actual sales numbers for any category can be shown on text labels, facilitating accurate comparisons and analysis.

**References**

1. Tableau Software. (2022). Dashboard Design Best Practices. <https://help.tableau.com/current/pro/desktop/en-us/dashboards_best_practices.htm>
2. Knaflic, C. N. (2015). Storytelling with Data: A Data Visualization Guide for Business Professionals. Wiley.
3. Visual Best practices, <https://help.tableau.com/current/pro/desktop/en-us/visual_best_practices.htm>
4. Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for discovering Clusters in Large Spatial Databases with Noise.
5. Edward Tufte's "The Visual Display of Quantitative Information"
6. Tableau Reference material, <https://www.tableau.com/resources/reference-materials>