**Customer Shopping Trends Analysis**

**Team 14**

**Project Goals**

The project's objective is to utilize Tableau for creating insightful visualizations based on the retail company's sales data from the previous year, intending to present these visual insights to the Marketing & Sales Department. Understanding customer preferences and trends is critical for businesses to tailor their products, marketing strategies, and overall customer experience. The visualizations will provide a comprehensive overview of the purchase of items across different spectral cuts- Demographics, Location, geography, season and frequency of purchase. This will help a business to analyze and decide which market segment to target for maximizing profit.

**Data Description**

●  **Source**:

○  KaggleDataset: <https://www.kaggle.com/datasets/iamsouravbanerjee/customer-shopping-trends-dataset>

●  **Geography**: United States of America

●  **Time Horizon**: Past one-year sales data

●  The sales data includes below information -

* **Customer ID** - Unique identifier for each customer
* **Age** - Age of the customer
* **Gender** - Gender of the customer (Male/Female)
* **Item Purchased** - The item purchased by the customer
* **Category** - Category of the item purchased
* **Purchase Amount (USD)** - The amount of the purchase in USD
* **Location** - Location where the purchase was made
* **Size** - Size of the purchased item
* **Color** - Color of the purchased item
* **Season** - Season during which the purchase was made
* **Review Rating** - Rating given by the customer for the purchased item
* **Subscription Status** - Indicates if the customer has a subscription (Yes/No)
* **Shipping Type** - Type of shipping chosen by the customer
* **Discount Applied** - Indicates if a discount was applied to the purchase (Yes/No)
* **Promo Code Used** - Indicates if a promo code was used for the purchase (Yes/No)
* **Previous Purchases** - The total count of transactions concluded by the customer at the store, excluding the ongoing transaction
* **Payment Method** - Customer's most preferred payment method
* **Frequency of Purchases** - Frequency at which the customer makes purchases (e.g., Weekly, Fortnightly, Monthly)
* Data cleaning:

○  Remove duplicates: Eliminated duplicate records to prevent redundancy and errors in analysis.

○  Handle missing values: Decided to remove missing data points.

○  Standardize data formats: Ensured consistency in formats, prices, and location.

**Team Roles**

***Asha***: Sales data cleaning and dashboard preparation - map and tables

***Trisheeka***: Sales data cleaning and dashboard preparation - bar, line and pie charts

***Vijaya:*** Sales data cleaning and dashboard preparation - map and tables

***Prithvi***: Sales data cleaning and dashboard preparation - bar, line and pie charts

**Dashboard Interactions and Findings:**

**Dashboard Overview**

A screenshot of a computer screen

Description automatically generated

A graph of blue rectangular bars with white text

Description automatically generated

***Average Customer reviews for shipping methods*** displays average customer reviews for various shipping methods. The methods included are Standard, Express, 2-Day Shipping, Next Day Air, Free Shipping, and Store Pickup. The highest-rated method is Standard shipping with a score of 3.82, closely followed by Express at 3.78, and 2-Day Shipping and Next Day Air both at 3.72. Free Shipping is rated slightly lower at 3.71, and Store Pickup is the least favored option, albeit marginally, with a score of 3.71. The range of scores is very narrow, from 3.71 to 3.82, suggesting that customer satisfaction levels are relatively consistent across different shipping methods. This data implies that customers value reliable and predictable shipping options over the speed or cost of delivery.

A graph showing the seasons

Description automatically generated

**Seasonal trend graph** shows a line chart representing a seasonal trend. The trend line indicates a gradual increase in the metric being measured as the seasons progress. There is a slight rise from Summer to Winter, followed by a more pronounced increase from Winter to Fall. The trend continues upward into Spring. The data suggests a seasonal pattern where the value being tracked dips lowest in the summer and peaks in the spring.

A map of the united states

Description automatically generated

The **Average Order Value by State** showsa geographic heat map indicating the AOV from sales data across the United States. The intensity of the color correlates with the AOV, where darker shades correspond to higher AOVs. Alaska (AK) is markedly darker than other states, indicating it has a higher AOV. This suggests that on average, customers in Alaska are spending more per order compared to customers in other states. Focusing on market segments in Alaska or similar high AOV states, tailoring marketing strategies regionally based on the AOV data, and investigating underlying factors that might contribute to the higher spending in certain areas to replicate this success in other regions.

A chart of blue bars with white text

Description automatically generated

**AOV based on purchase frequency** suggests a comparison of purchase amounts at different purchase frequencies. The amounts range from $57 to $64, with Annual purchases being the highest at $64 and Fortnightly the lowest at $55. It's notable that the more frequent purchase intervals—Weekly, Fortnightly, and Bi-Weekly—have lower average amounts compared to less frequent intervals like Annually and Every 3 Months, which could indicate that customers spend more per purchase when they buy less frequently. This pattern is useful for the businesses in understanding customer spending habits and tailoring marketing strategies accordingly.

A graph of blue and white colored bars with dollar and x

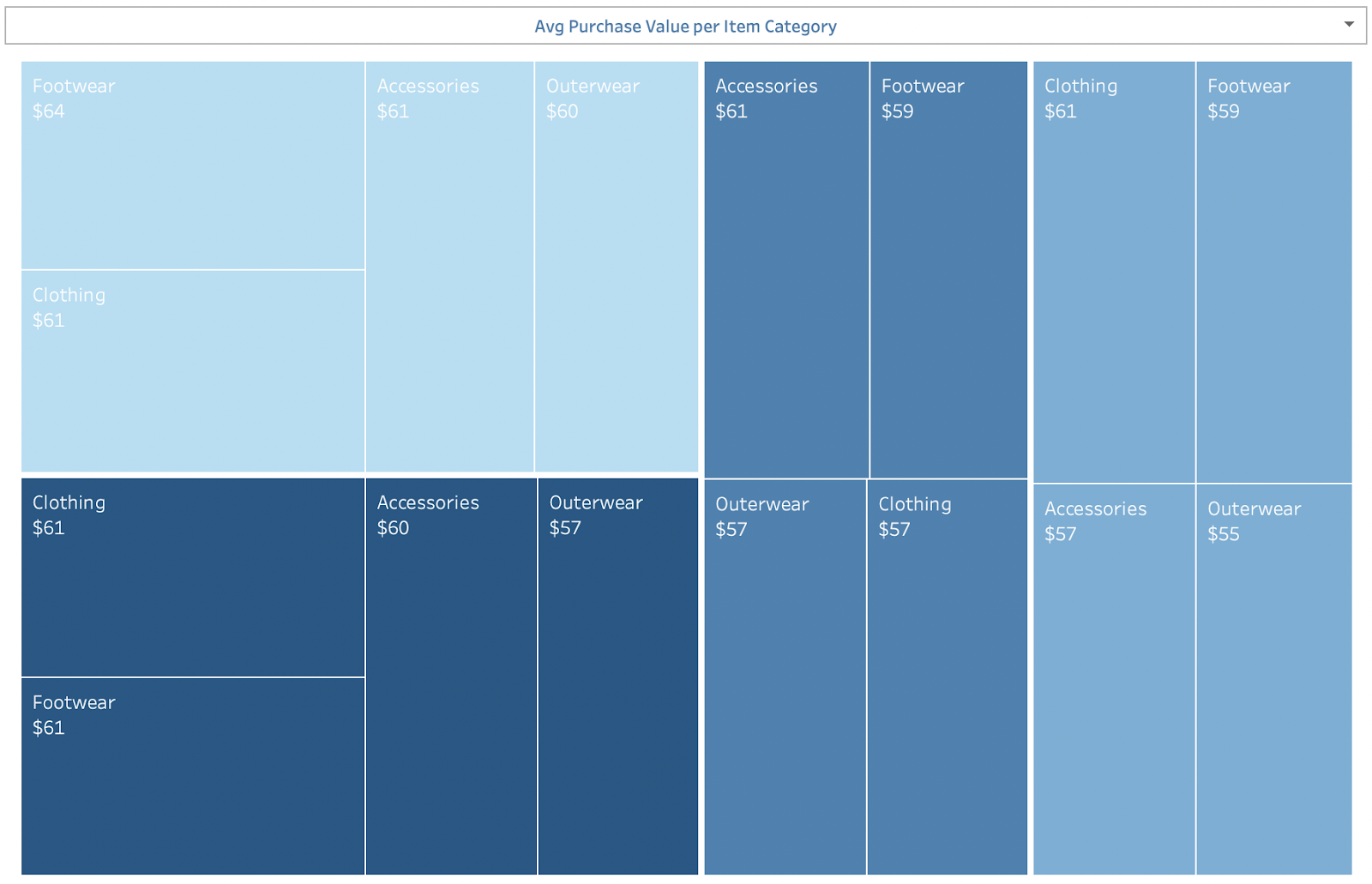
Description automatically generated with medium confidence

**Generational Buying Power** compares the average amount of money spent by different generations. Generation Z and Generation X both show the highest average spend at $63, followed by Baby Boomers at $61 and Millennials at $60. The data indicates a relatively uniform spending power across the generations, with a slight variation of $3 between the highest and lowest averages. This suggests that each generation has significant purchasing power, and there is no stark difference between the youngest (Generation Z) and the oldest (Baby Boomers) in terms of spending. Our Business might use this information to tailor marketing campaigns that are inclusive of all age groups, given the similar spending levels.

A graph of blue dots

Description automatically generated with medium confidence

**Number of items previously purchased** is a scatter plot with various types of apparel and accessories. Visual analysis suggests items like jackets, shoes, and T-shirts have been purchased more frequently than items like gloves and sandals. The plot shows a varied distribution, indicating that certain items are more popular or necessary purchases among consumers. For instance, essentials or fashion staples may account for higher purchase frequencies. Our Business could use this data to adjust inventory levels, plan marketing strategies, or offer promotions on popular items to boost sales.



**Average Purchase Value per Item Category** tells us that footwear has the highest value at $64, followed by Clothing and Accessories, both at $61. Outerwear shows values ranging from $55 to $60. This suggests that customers have spent the most on Footwear and the least on Outerwear. The consistent $61 value for Clothing and Accessories might indicate a balanced spending behavior between these categories. Businesses could infer that while customers are willing to spend more on shoes, they may be more conservative with their outerwear purchases based on the data. This data helps our business in stocking decisions and promotional activities, focusing on the higher spending categories to drive sales.

A blue circle with a number of percentages

Description automatically generated

**Gender distribution** shows the proportion of males to females in the dataset. Males make up a larger portion of the chart at 68%, while females account for 32%. This substantial difference indicates a significant gender disparity which is a critical insight for targeted marketing strategies.

We have enriched the analysis by incorporating additional datasets to enhance the depth and relevance of our insights. Recognizing the necessity for a comprehensive view of our consumer base, we have integrated census demographic data by geography. This allows us to calculate sales penetration rates within the general population and, more specifically, within specific age and gender brackets across different regions. This added layer of demographic context enables us to paint a more accurate picture of our market and to identify untapped segments that could be targeted through tailored marketing strategies.

Furthermore, we have refined our dashboard design to better serve the decision-making process for our business. By doing so, we have developed metrics and charts that directly inform potential changes in marketing approaches. Each component of the dashboard has been carefully crafted to reflect the strategic objectives, with added interactivity to allow users to drill down into different demographic layers. This interactive capability ensures our business can explore various scenarios and extract precise data points that are most pertinent to their marketing decisions. The purpose behind each chart is now clearly articulated, ensuring that the data visualization aligns with the overarching goal of optimizing marketing efforts to boost sales.