Analyzing Consumer Shopping Trends

Anoushka Amudhan, Sreenidhi Tadepalli, Nihita Kasibhatla

Website Link

https://nidhitadepalli.github.io/ds4200-project

Intro

Our project is focused on analyzing shopping trends and examining consumer behavior. By analyzing these trends, businesses can tweak their marketing strategies as well as manage their inventory levels. Since online shopping is on the rise, analyzing shopping data can provide useful insights into the behavior of different demographics and how they interact with product categories. Analyzing these trends and observing the patterns can also allow the consumer to receive more personalized recommendations.

Introduction to the Data

The dataset that we chose for this project is about shopping trends. It consists of consumer transaction records, including the product category, purchase amount in USD, payment method, and customer demographics like age and gender. The dataset is a CSV file and it was downloaded from Kaggle. The entire dataset consists of 19 different categories and roughly 3000 observations. It has both categorical and continuous variables.

Some of the features specifically used in our analysis are the **category**, showing the type of item purchased (Accessory, Clothing, Outerwear, or Footwear), the **purchase amount**, the amount of USD for the specific purchase, the **location,** the state the purchase was made in, the **frequency of purchases,**

how often the customer makes purchases (Weekly, Fortnightly, Monthly), and the **payment method**, the customer's most preferred payment method.

Link to data: https://www.kaggle.com/datasets/iamsouravbanerjee/customer-shopping-trends-dataset

Design Analysis

Altair Bar Chart → The implemented bar chart visualizes purchase trends across four product categories by gender and season using a small multiple approach. Each season panel contains grouped bars showing categories on the x-axis and purchase counts on the y-axis, with light purple for male and pink for female purchases. The design features consistent y-axis scaling to 350 purchases, light grid lines for readability, clear axis labeling, and alphabetical category sequencing.

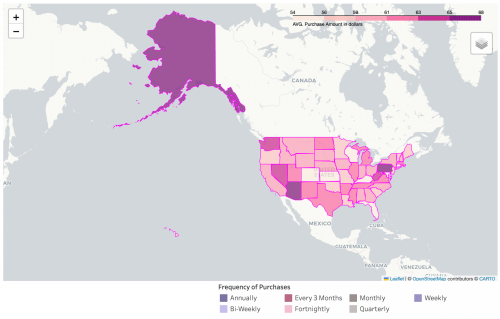
Sankey → The Sankey diagram effectively illustrates the connection between payment methods and customer payment frequency segments using a flow-based design. Positioned on the left, the source nodes represent five payment methods—Credit Card, PayPal, Cash, Venmo, and Bank Transfer—while the target nodes on the right correspond to six payment frequencies: Weekly, Monthly, Annually, Bi-Weekly, Fortnightly, and Every 3 Months. Flowing between these nodes, gray bands of varying widths depict customer volume, with thicker bands indicating more common payment method-frequency combinations. Pink vertical bars frame both the source and target nodes, enhancing visual clarity and separation.

Geospatial → The implemented geospatial visualization presents average purchase amounts across the U.S. states through an interactive choropleth map. The design employs a purple-pink sequential color palette where darker purples represent higher spending states and lighter pinks indicate lower spending regions. State boundaries are clearly delineated with magenta outlines that enhance visibility against the light gray base map. The visualization includes interactive elements, with a comprehensive tooltip

system showing multiple metrics: average purchase amount, average customer rating, top product category, promo code usage, and subscription rate. Navigation controls are positioned in the top-left corner, allowing users to zoom in/out, while a layer control panel appears in the top-right.

Tableau Line Chart → The line chart showcases the distribution of purchase frequencies across the four seasons—Winter, Spring, Summer, and Fall. It features intersecting trend lines representing seven distinct purchase frequency categories, each following its own seasonal pattern. The y-axis ranges from 80 to 160, reflecting customer counts, while the x-axis arranges the seasons in chronological order. An interactive tooltip system, demonstrated for the "Annually" frequency in Spring, highlights four key metrics: the purchase frequency, the season being analyzed, the distinct customer count, and the estimated total annual purchases. Designed keeping readability in mind, the chart incorporates subtle grid lines and well-spaced y-axis values for improved clarity.

Visualization Analysis



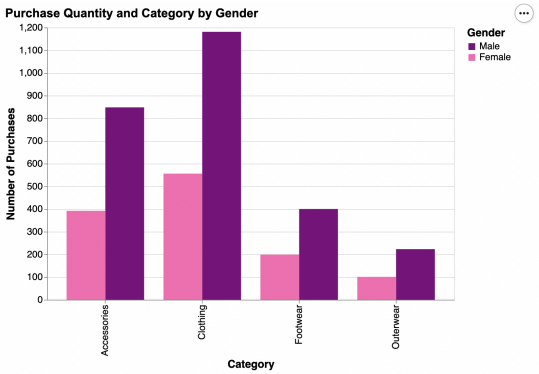
Geospatial Analysis → This geospatial map reveals average purchase values for U.S. states, with each layer differing in that it represents a different product category. Blacker colors indicate greater average spend within the category, allowing users to examine local patterns of Clothing, Accessories, and more. A layer control panel allows direct comparison across categories, raising interactivity and making detailed analysis easier.

Alaska stands alone as the sole state with the highest average purchase value overall, while most states fall in middle to low-spending ranks. The tooltip for a state provides additional information, including

average purchase value, customer ratings, best product category, promo code redemption, and subscription purchase levels. Although "Clothing" is the top category for most states, New Mexico and West Virginia favor "Accessories" individually, suggesting varying regional preferences or market forces.

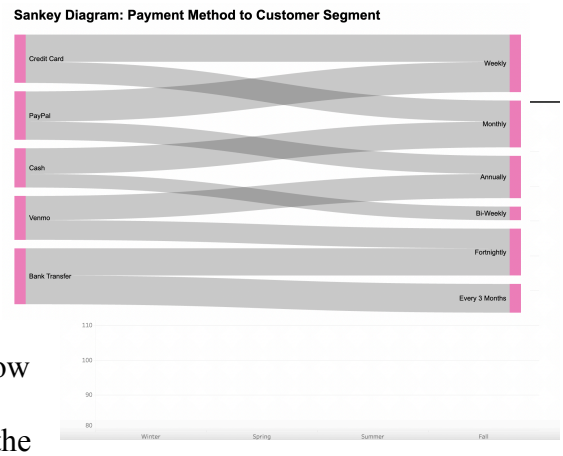
These patterns may suggest broader distinctions in economic conditions, consumer shopping, and success at the point of marketing among states. Stacked style highlights not only where it's high or low expenditure, but also on what consumers are spending—giving better category-level insights into patterns within the U.S. marketplace.

Altair Bar Chart → This bar chart illustrates purchase trends across four product categories—Accessories, Clothing, Footwear, and Outerwear—by gender. Interestingly, males consistently make more purchases than females in all categories. Clothing emerges as the most popular category for both genders, with males making more than twice as many purchases as females. Accessories follow, again with males leading significantly, while Footwear and Outerwear show lower overall volumes but maintain the same gender pattern. The visualization also allows the user to filter by state and see the differences in gender in any state they chose.



These findings challenge common societal stereotypes that suggest females shop more frequently or have a stronger interest in fashion and retail consumption. In contrast, this data shows males not only making more purchases overall but also showing strong category

preferences, particularly for clothing and accessories. Meanwhile, female purchasing behavior appears more balanced across categories.



Sankey Analysis → The Sankey diagram visualizes the relationship between payment methods and customer payment frequency segments, illustrating how different payment methods correspond to various payment schedules. Credit Card users along with PayPal customers predominantly make "Weekly" payments. Cash payments show a strong connection to "Monthly" payments, as indicated by the label "Cash → Monthly: 85 customers", but also have notable flows toward "Bi-Weekly" payments. Venmo users primarily opt for "Annual" payments, whereas Bank Transfer customers are split between "Fortnightly" and "Every 3 Months" schedules. The width of each flow represents the volume of customers for each payment method and frequency combination. This visualization provides valuable insights into customer payment preferences, helping businesses optimize payment systems and optimize cash flow

Tableau Line Chart Analysis → The Tableau line chart "Distribution of Purchase Frequency Across Seasons" reveals distinct seasonal purchasing patterns. Weekly (purple) purchasing shows steady decline from Winter to Summer, dropping to its lowest in Summer after surging to the highest point in Winter. Bi-weekly (light purple) demonstrates the opposite trend, peaking in Spring at approximately 155 customers before steadily declining to about 120 customers by Fall. Quarterly (light brown) customers increase steadily from Winter through Summer (peaking around 144) before declining ever so slightly in

Fall. Fortnightly (light pink) displays a steady decline from Winter through Summer, its lowest point being Summer at around 127 customers, before spiking back up during Fall. Monthly (dark brown) mainly follows suit with Weekly customers. Every 3 Months (dark pink) and Annual (dark purple) frequencies show more modest fluctuations across seasons. These contrasting patterns suggest customers shift their purchasing frequencies seasonally, with potential opportunities for targeted marketing strategies during transitional periods.

Heatmap Concept → The concept for our fifth visualization will be to produce a heatmap showing the average purchase dollar amount (USD) per category and age range. Using the Age and Category columns in the data set, the visualization will highlight where spending varies between categories and between ages. The age data can be binned (i.e., 15–24, 25–34, etc.), and categories can be derived from the Category column in the data set. The average Purchase Amount USD value per category-age pair will be in each cell in the visualization, and relative spending will be conveyed through color intensity. A visualization such as this would easily and clearly communicate in a readable way where customer segments were spending the most money on what categories and would be a great window to consumer tastes. Features allowing higher levels of interaction with the visualization, such as viewing tooltips with exact values or gender filters, would be useful to allow. As none of the visualizations currently perform spending behavior between categories and ages, a new piece to the overall analysis would be provided by such the heatmap visualization.

Summary & Key Takeaways

Our project uncovers a number of underlying patterns in consumer shopping habits through a series of seasonal, demographic, and geographic visualizations. We found that shopping is high in Spring and low in Summer, especially among frequent shoppers. Gender patterns were exposed, highlighting differences in consumer segments' shopping behaviors. Male consumers had a much higher number of purchases across all categories compared to female consumers. Geographic data patterns revealed that shoppers across the country spent an average of 50-60 dollars and the most popular shopping category is clothing. Consumer payment types also indicate the frequency of purchases. Customers that prefer electronic payment methods tend to purchase more frequently than those who prefer cash.

All the visualizations give insight into where the different factors intersect. The line graph effectively shows trends in buying frequency by seasons, while the grouped bar graph is simple to compare product choices by male and female buyers. The Sankey diagram shows the different payment options together with payment frequencies, and this provides richer insights into the behavior of consumers. These findings not only uncover unique behavioral patterns but also guide key business strategy. With insight into when and how different segments shop, companies can better plan inventory, tailor marketing, and optimize payment options. At the same time, there are opportunities here to drive a better customer experience—personalized offers, location-based promotions, and payment convenience can all be informed by these learnings. Our project can be expanded upon in the future to look at product-level data or to compare several years of trends together. By collecting more data and adding interactive variables such as age or location, the findings can become more personal and actionable to business users.

Reference Papers

Cele O. (2001, April 9). *Perceptions and realities of male shopping behavior*. Journal of Retailing. https://www.sciencedirect.com/science/article/abs/pii/S0022435900000476

Price perceptions and consumer shopping behavior: A field study. (n.d.-b). https://www.researchgate.net/publication/247243361_Price_Perceptions_and_Consumer_Shopping_Behavior_A_Field_Study