Project Phase II: Decision Making

*Submitted by*

[Naveen Kumar Manokaran](https://canvas.asu.edu/groups/607025/users/1076721)

“Ira A. Fulton Schools of Engineering, Arizona State University”

“IFT 533: data Visualization”

(2024 Fall)

*Submitted to:* Prof. Asmaa Elbadrawy

Submitted on: November 24th, 2024.

**Index**

|  |  |  |
| --- | --- | --- |
| **S.No** | **Title** | **Page No** |
| **1** | **Section1: Used Visualization Tools** | **3** |
| **2** | **Section 2: Explanation of Required Data Pre-processing** | **4** |
| **3** | **Section 3: List of Final Sets of Questions** | **9** |
| **4** | **Section 4: Dashboard Plot Drafts** | **10** |
| **5** | **Section 5: Dashboard Interactivity** | **21** |
| **6** | **Section 6: References, Include link to team's Mural board** | **25** |

**SECTION 1:**

We chose Tableau as the primary visualization solution for this project because it allows us to generate dynamic and interactive dashboards. Tableau is an industry-standard platform noted for its ability to handle large information and sophisticated visualization capabilities. It is particularly well-suited for this project due to the dataset's reasonable size (3,900 rows and 18 columns), providing efficient and seamless execution. Tableau Prep Builder was also used to preprocess data before visualizing it.

**Advantages of Tableau:**

Tableau's user-friendly interface allows for the creation of visuals without requiring substantial technical skills.

Tableau allows you to create interactive dashboards with filters, parameters, and drill-down options to enhance the user experience.

Tableau effortlessly interacts with a variety of data sources, including internet servers, cloud platforms, and databases, providing flexibility and accessibility.

Dashboards can be published online or linked to live data sources, allowing users to view and interact with visualizations from anywhere.

**Advantages of Tableau Prep Builder:**

Efficient Data Preparation: Simplifies data pretreatment by providing tools for cleaning, combining, and transforming data.

Allows you to create grouped fields such as Age Group, Purchase Amount Group, and Review Rating Group, ensuring that your data is organized for useful analysis.

Preprocessed data may be sent directly to Tableau for visualization, speeding the entire operation.

These features make Tableau and Tableau Prep Builder excellent options for translating preprocessed data into useful, user-friendly visualizations.

**SECTION 2:**

This Dataset requires only minimal preprocessing since the dataset was already cleaned. The Dataset did not have any null values or missing datapoints. Using Tableau Prep Builder, we created calculated fields to group data for enhanced clarity and usability:

**Age Group:** Categorized into logical segments:

<20, 20s, 30s, 40s, 50s, 60s, 70s.

A screenshot of a computer

Description automatically generated

**Purchase Amount Group:** Grouped by ranges:

20-40, 40-60, 60-80, 80-100, 100+.

A screenshot of a computer

Description automatically generated

**Review Rating Group:** Segmented into intervals:

1-2, 2-3, 3-4, 4-5.

These calculated fields improve analysis by organizing data into meaningful categories, making the dashboard more intuitive and insightful for usersA screenshot of a computer

Description automatically generated.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

**SECTION 3:**

The dashboard will address the following key questions to provide actionable insights:

1. Which product categories generate the highest revenue for the given location?
2. What are the top categories sold for a given season?
3. Which regions contribute most to total sales?
4. What are the top 5 best products as per customer feedback?
5. Which payment method is the most preferred by the customers?
6. What are the customer spending pattern for different age group and Gender?
7. What are the highest-selling products based on gender?
8. Which shipping types are preferred for high-value purchases?
9. Which states have the highest subscription-based customers?
10. How do customer purchase patterns vary based on frequency of purchases?
11. How do review ratings correlate with purchase amounts?
12. What are the spending habits of customers with high purchase counts?
13. Which age group Customer spends the most?
14. Customer in which gender spends the most?
15. Does the review of the products affect the frequency of purchase?

**Section 4:**

**PLOT 1: "Revenue by Product Category for a Given Location"**

**Chart Type**: Bar Chart  
**Rows**: SUM(Purchase Amount(USD))  
**Column**: Category  
**Filters**: State Select (Calculated Filter):True, Filter by State (Parameter)   
**Pre-attentive Attributes**: Length  
**Description**: Addresses Question 1 by showing total revenue generated by product categories for the selected location. The location can also be all giving the total revenue generated for each category.  
**Connected Questions**: 1

A screenshot of a computer

Description automatically generated

**PLOT 2: "Top Categories by Season"**

**Chart Type**: Stacked Bar Chart  
**Rows**: COUNT(Shopping\_Trends\_Final.csv)  
**Column**: Season  
**Marks Color Card**: Category  
**Pre-attentive Attributes**: Length, Position, Color  
**Color Distribution**: Distinct colors for each category  
**Description**: Addresses Question 2 by highlighting seasonal trends for top-selling categories.  
**Connected Questions**: 2

A screenshot of a computer

Description automatically generated

**PLOT 3: "Regional Sales Contributions"**

**Chart Type**: Geographic Map  
**Rows**: Latitude  
**Column**: Longitude  
**Marks Color Card**: SUM (Purchase Amount(USD))  
**Pre-attentive Attributes**: Geographic Position, Color Intensity  
**Color Distribution**: Gradient (Light Blue to Dark Blue)  
**Description**: Addresses Question 3 by visualizing regional contributions to total sales.  
**Connected Questions**: 3

A map of the united states

Description automatically generated

**PLOT 4: "Best Products by Feedback"**

**Chart Type**: Stacked Bar Chart  
**Rows**: AVG(Review Rating)  
**Column**: Item Purchased  
**Filters**: Top Rated Products (Parameter), Filter Top Ratings (Calculated Field)  
**Pre-attentive Attributes**: Length  
**Description**: Addresses Question 4 by ranking products based on customer feedback, segmented into best categories.  
**Connected Questions**: 4

A screenshot of a computer

Description automatically generated

**PLOT 5: "Payment Method Preferences"**

**Chart Type**: Bar Chart  
**Rows**: COUNT(Transactions)  
**Column**: Payment Method  
**Pre-attentive Attributes**: Length, Stack Position, Color  
**Description**: Addresses Question 5 by visualizing the popularity of payment methods across all regions.  
**Connected Questions**: 5

A screenshot of a computer

Description automatically generated

**PLOT 6: "Spending patterns by Age Group and Gender"**

**Chart Type**: Bubble Chart  
**Size**: SUM (Purchase Amount (USD))  
**Color**: Gender  
**Label:** Age Group, Gender and SUM (Purchase Amount (USD))  
**Pre-attentive Attributes**: Size, Color  
**Color Distribution**: Distinct colors for each Gender  
**Description**: Addresses Questions 6, 13 and 14 by showing spending pattern across age groups and Gender  
**Connected Questions**: 6, 13 and 14

A screenshot of a computer

Description automatically generated

**PLOT 7: "Gender-Based Product Preferences"**

**Chart Type**: Stacked Bar Chart  
**Rows**: COUNT (Item Purchased)  
**Column**: Item Purchased  
**Marks Color Card**: Gender  
**Pre-attentive Attributes**: Length, Position, Color  
**Color Distribution**: Male (Orange), Female (Blue)  
**Description**: Addresses Question 7 by highlighting the highest-selling products for each gender.  
**Connected Questions**: 7

A screenshot of a computer

Description automatically generated

**PLOT 8: "Shipping Preferences for High-Value Purchases"**

**Chart Type**: Grouped Bar Chart  
**Rows**: CNT (Shopping\_Trends\_Final.csv)  
**Column**: Shipping Type, Purchase Amount Group

**Marks Color Card**: Purchase Amount Group  
**Pre-attentive Attributes**: Length

**Color Distribution**: Categorical colors for different Purchase Amount Group  
**Description**: Addresses Question 8 by showing preferred shipping types for high-value purchases.  
**Connected Questions**: 8

A screenshot of a computer screen

Description automatically generated

**PLOT 9: "Subscription-Based Customers by State"**

**Chart Type**: Stacked Bar Chart  
**Rows**: CNT(Shopping\_Trends\_Final.csv)  
**Column**: Location  
**Marks Color Card**: Subscription Status  
**Pre-attentive Attributes**: Length, Color  
**Color Distribution**: Subscription Status - Yes (Orange), No (Blue)  
**Description**: Addresses Question 9 by visualizing states with the highest subscription-based customers.  
**Connected Questions**: 9

A screenshot of a computer screen

Description automatically generated

**PLOT 10: "Purchase Frequency Analysis"**

**Chart Type**: Line Chart  
**Rows**: SUM (Purchase Amount (USD))  
**Column**: Frequency of Purchases  
**Marks Color Card**: Review Rating Group  
**Pre-attentive Attributes**: Line Length, Position, Color  
**Color Distribution**: categorical colors for review ratings.  
**Description**: Addresses Questions 10, 11 and 15 by analyzing the relationship between purchase frequency, amounts and review ratings.  
**Connected Questions**: 10, 11, 15

A screenshot of a computer

Description automatically generated

**PLOT 11: "Customer Spending Patterns by Purchase Count"**

**Chart Type**: Bubble Chart  
**Size**: CNT (Shopping\_Trends\_Final.csv)  
**Color**: Gender  
**Label:** Purchase Amount Group, Gender and CNT (Shopping\_Trends\_Final.csv)  
**Pre-attentive Attributes**: Size, Color  
**Color Distribution**: Distinct colors for each Gender  
**Description**: Addresses Questions 12 by visualizing spending patterns for customers with varying purchase counts and gender segmentation.  
**Connected Questions**: 12

A screenshot of a computer

Description automatically generated

**SECTION 5:**

**List of Interactive Controls**

The following interactive controls will allow users to customize the visualizations and extract meaningful insights from the dashboard:

**State Select Dropdown**

* **Purpose**: Allows users to filter data by specific states or view data for all states.
* **Connected Plots**:
  + Plot 1: "Revenue by Product Category for a Given Location"
  + Plot 3: "Regional Sales Contributions"
  + Plot 9: "Subscription-Based Customers by State"
* **Value Range**: List of unique states from the dataset. Includes "All States" as an option.
* **Loaded From**: Location attribute in the dataset.

Created a parameter “Filter by State” to filter the revenue based on the state selected.

A screenshot of a computer

Description automatically generated

For this we created a calculated field State Select to filter and display only the states selected by the user in “Filter by State” drop down.

A screenshot of a computer

Description automatically generated

**Top Rated Products Parameter**

* **Purpose**: Filters the top x products based on review ratings
* **Connected Plots**:
  + Plot 4: "Best Products by Feedback"
* **Value Range**: Numeric range from 1 to 100 (but actually corresponds to number of products available).

Below is the screenshot for the “Top Rated Products” parameter created.

A screenshot of a computer

Description automatically generated

Below is the screenshot for the “Filter Top Ratings” calculated field created to filter the top x ratings selected from the “Top Rated Products” parameters.

A screenshot of a computer

Description automatically generated

**Gender Filter**

* **Purpose**: Filters data by gender to analyze product preferences or spending patterns.
* **Connected Plots**:
  + Plot 6: "Spending Patterns by Age Group and Gender"
  + Plot 7: "Gender-Based Product Preferences"
  + Plot 11: "Customer Spending Patterns by Purchase Count"
* **Value Range**: Male, Female.
* **Loaded From**: Gender attribute in the dataset.

**Purchase Amount Group Selector**

* **Purpose**: Filters data by purchase value ranges for detailed analysis.
* **Connected Plots**:
  + Plot 8: "Shipping Preferences for High-Value Purchases"
  + Plot 11: "Customer Spending Patterns by Purchase Count"
* **Value Range**: Defined purchase amount groups (e.g., 20-40, 40-60, 60-80, 80-100, 100+).
* **Loaded From**: Purchase Amount Group Column calculated field created using Purchase Amount.

**Category Selector Dropdown**

* **Purpose**: Allows users to filter and analyze data for specific product categories.
* **Connected Plots**:
  + Plot 1: "Revenue by Product Category for a Given Location"
  + Plot 2: "Top Categories by Season"
  + Plot 7: "Gender-Based Product Preferences"
* **Value Range**: List of unique categories present in the dataset.
* **Loaded From**: Category attribute in the dataset.

**Item Purchased Selector Dropdown**

* **Purpose**: Enables users to focus on specific items purchased for detailed feedback and analysis.
* **Connected Plots**:
  + Plot 4: "Best Products by Feedback"
  + Plot 7: "Gender-Based Product Preferences"
* **Value Range**: List of unique items purchased in the dataset.
* **Loaded From**: Item Purchased attribute in the dataset.

These controls enhance interactivity, enabling users to explore the data across multiple dimensions and extract insights tailored to their interests. Let me know if further details are needed!

**Section 6: References**

1. Sourav Banerjee. *Customer Shopping Trends Dataset*. Kaggle, 2024, <https://www.kaggle.com/datasets/iamsouravbanerjee/customer-shopping-trends-dataset>
2. Mural Board Link: <https://app.mural.co/t/project373258/m/project373258/1731619110561/91ce729475358157ad3ba50e93d507db527e9d4c>