Shopping _Trends Data Analysis

Summary and recomendations

Detailed Dataset Summary

Overview

The dataset serves as the foundation for analyzing shopping trends, with features focusing on consumer demographics, purchasing behavior, and associated shopping patterns. It supports exploratory data analysis and visualizations aimed at identifying patterns and trends in retail or e-commerce.

Dataset Structure

1. **Columns**: Based on the visualizations prepared in the notebook, the dataset contains at least the following columns:
   - **Gender**: Indicates the gender of the customer (e.g., Male, Female, Others).
   - **Subscription Status**: Tracks whether the customer is subscribed to a service (e.g., Premium/Basic).
   - **Payment Method**: Captures the mode of payment (e.g., Credit Card, PayPal).
   - **Shipping Type**: Specifies the type of shipping chosen (e.g., Standard, Express).
   - **Discount Applied**: Boolean or categorical value indicating whether a discount was used during the purchase.
   - **Promo Code Used**: Details whether a promotional code was applied.
   - **Item Purchased**: The specific product or product category purchased.
   - **Size**: May refer to the size of the item (e.g., Small, Medium, Large).
   - **Season**: Represents the season when the purchase occurred (e.g., Winter, Summer).
   - **Category**: Likely a grouping feature used for analyzing subsets of data (e.g., Electronics, Clothing).
2. **Potential Missing Columns**: There may be additional columns not explicitly referenced in the notebook outputs, such as timestamps or geographic information, which are typical in shopping datasets.

3. **Data Size**: While the exact number of rows (observations) or columns is not directly mentioned, the dataset is sufficient for generating count plots and other aggregate visualizations, implying it covers a substantial number of entries.

Data Content and Examples

Although the first few rows (`head()`) are not explicitly shown, the notebook suggests:

- The dataset contains both categorical and numerical features.
- Most columns likely represent qualitative attributes (e.g., `Gender`, `Shipping Type`).
- Numerical attributes may include quantities such as item counts, discount percentages, or prices.

Key Insights for Analysis

- **Consumer Segmentation**: Attributes such as `Gender`, `Subscription Status`, and `Season` are leveraged to segment consumer groups.
- **Promotional Impact**: Fields like `Discount Applied` and `Promo Code Used` can reveal the influence of discounts and promotions on purchasing behavior.
- **Behavioral Patterns**: Columns such as `Payment Method`, `Shipping Type`, and `Size` help analyze customer preferences and habits.
- **Seasonality**: The `Season` column provides scope for identifying time-based trends in shopping activity.

Utilization

The dataset is used for:

1. **Count Plots**: Visualizing the frequency of categorical attributes (e.g., `Gender`, `Season`) across different categories (e.g., `Category`).
2. **Trends Analysis**: Exploring temporal and demographic factors influencing shopping.
3. **Behavioral Analysis**: Understanding customer interactions with promotions and discounts.

To visualize all the columns mentioned in the notebook :

1. Generate **count plots** for categorical columns like `Gender`, `Subscription Status`, etc.

2. Include potential **hue-based visualizations** by splitting data using the `Category` column.
3. Display a grid of plots to make all visualizations accessible.

＊ ＊ ＊