Introduction

In today's dynamic marketplace, understanding shopping trends is paramount for businesses aiming to stay competitive and relevant. For this project, I investigated potential shopping trends, which may inspire valuable insights that drive strategic decision-making and enhance customer experiences.

The dataset contains a multitude of information regarding a customer, ranging from demographic details and geographic distributions, payment types, and frequencies. Through these attributes, key goals include identifying emerging trends and consumer preferences, predicting future market dynamics, and aligning strategies.

By deciphering the shopping behaviors, this analysis endeavors to empower businesses across various sectors, enabling adaptation to the evolving consumer demands, optimize marketing strategies, streamline inventory management, and ultimately foster sustainable growth and profitability.

Through this report, we aim to provide stakeholders with invaluable insights into the ever-evolving landscape of consumer behavior, equipping stakeholders with the knowledge necessary to navigate the complexities of modern commerce successfully.

Data Exploration

How was the Data Collected?

The dataset was collected from Kaggle Datasets. Here's the link for the dataset: https://www.kaggle.com/datasets/iamsouravbanerjee/customer-shopping-trends-dataset

Features within the Dataset

- 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)

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Shopping Trend Analysis

```
df2=pd.read_csv('/content/shopping_trends.csv')
df2.head()
```

	Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Season	Review Rating	Subscription Status	Payment Method	Shipping Type	Discount Applied	Promo Code Used	Previous Purchases	Preferred Payment Method	Frequency of Purchases
0	1	55	Male	Blouse	Clothing	53	Kentucky	L	Gray	Winter	3.1	Yes	Credit Card	Express	Yes	Yes	14	Venmo	Fortnightly
1	2	19	Male	Sweater	Clothing	64	Maine	L	Maroon	Winter	3.1	Yes	Bank Transfer	Express	Yes	Yes	2	Cash	Fortnightly
2	3	50	Male	Jeans	Clothing	73	Massachusetts	S	Maroon	Spring	3.1	Yes	Cash	Free Shipping	Yes	Yes	23	Credit Card	Weekly
3	4	21	Male	Sandals	Footwear	90	Rhode Island	M	Maroon	Spring	3.5	Yes	PayPal	Next Day Air	Yes	Yes	49	PayPal	Weekly
4	5	45	Male	Blouse	Clothing	49	Oregon	M	Turquoise	Spring	2.7	Yes	Cash	Free Shipping	Yes	Yes	31	PayPal	Annually

- There's a total of 3900 entries, without null items. Below shows a screenshot of all the column and type:

```
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   Customer ID             3900 non-null   int64
 1   Age                     3900 non-null   int64
 2   Gender                  3900 non-null   object
 3   Item Purchased          3900 non-null   object
 4   Category                3900 non-null   object
 5   Purchase Amount (USD)   3900 non-null   int64
 6   Location                3900 non-null   object
 7   Size                    3900 non-null   object
 8   Color                   3900 non-null   object
 9   Season                  3900 non-null   object
 10  Review Rating           3900 non-null   float64
 11  Subscription Status     3900 non-null   object
 12  Payment Method          3900 non-null   object
 13  Shipping Type           3900 non-null   object
 14  Discount Applied        3900 non-null   object
 15  Promo Code Used         3900 non-null   object
 16  Previous Purchases      3900 non-null   int64
 17  Preferred Payment Method 3900 non-null  object
 18  Frequency of Purchases  3900 non-null   object
```

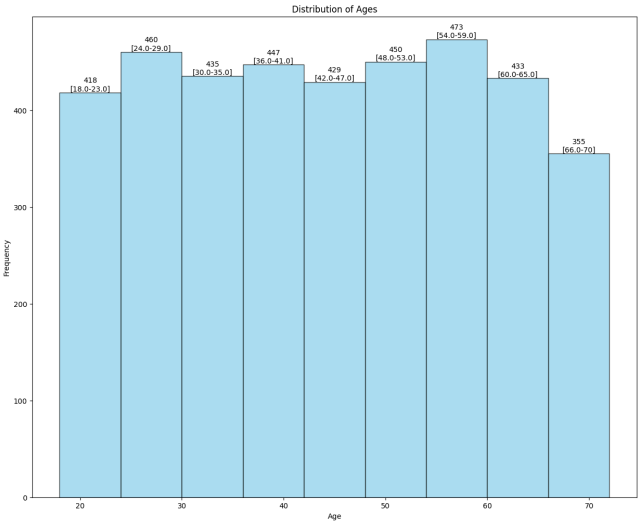
There are: 14 object columns, 4 int64 columns, and 1 float64 column.

Modeling and Results

Here's a summary list of some of the descriptive statistics from the dataset:

- There are 3900 entries, out of all the consumers, the average age is 44 years old, with an average purchase amount $60.

- 2223 purchases didn't apply discounts while 1677 did.
- The correlation coefficient between age and purchasing amount is approximately -0.0104, which is very close to 0: There is almost no linear relationship between age and purchase amount. The negative sign of the correlation coefficient may indicate a slight tendency for purchase amount to decrease slightly with increasing age, but again, this relationship is very weak.
- The range for age goes from 18 to 70.
    - The following showcases the distribution of purchases based on age groups:

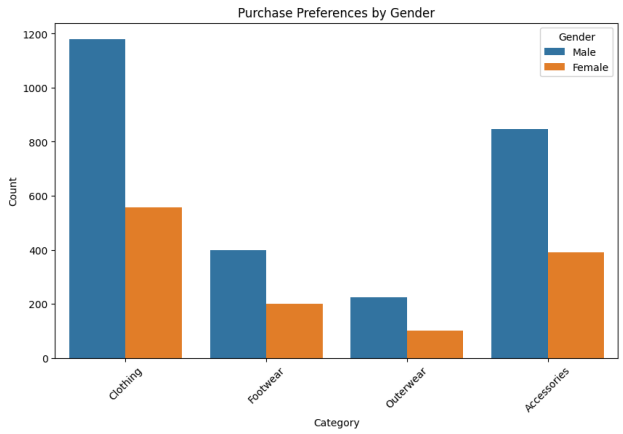


    - Overall, the age groups had relatively similar purchase frequencies, with the age group between 66-70 having only 355 purchases compared to other age groups of 400+ purchases. This may be due to the fact that older age might prefer in-store shops or might not be used to shopping digitally.
- The minimum purchase amount is $20, while the highest purchase amount is $100.
- By organizing the different categories of items and their purchase amount:

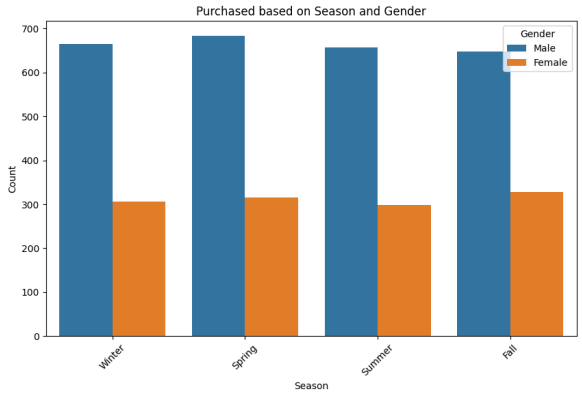
Category	Total Revenue (Purchase Amount in USD)
Clothing	104,264
Accessories	74,200
Footwear	36,093
Outerwear	18,524

Category	Average Purchase Amount (in USD)

Clothing	60.0
Accessories	59.8
Footwear	60.3
Outerwear	57.2



- The different genders seem to have similar purchasing preference, which is that both genders purchased clothing > accessory>footwear>outerwear.



- There was no significant difference with the total amount of revenue between the seasons. Fall has the highest total sales, followed by Spring, Winter, while Summer has the lowest sales.

Season	Purchase Amount (in USD)
Fall	60,018
Spring	58,679

Summer	55,777
Winter	58,607

- Here's the top 5 items purchased and the total number of purchases:

Item Purchased	Total Purchases
Jewelry	171
Blouse	171
Pants	171
Shirt	168
Dress	166

    - Out of the items that are most purchased, blouse accounts for $10410 of the $233081, which is 4.4% of the total sales
- The lowest item had a review/rating of 2.5, while the highest had a 5.0 (out of 5).
- The proportion of male customers is 68%, while female customers account for 32%.
- Based on geographic location, here are the top 5 locations with the most number of purchases:

Location	Number of Purchases
Montana	96
California	95
Idaho	93
Illinois	92
Alabama	89

- Meanwhile, these are the locations with the least number of purchases:

Location	Number of Purchases
Rhode Island	63
Kansas	63
Arizona	65

Hawaii	65
New Jersey	67

- Based on the payment methods, here's a breakdown of the various payment options and its popularity:

Payment Method	Count
Credit Card	696
Venmo	653
Cash	648
Paypal	638
Debit Card	633
Bank Transfer	632

- The customers would buy from fortnightly, weekly, bi-weekly, monthly, every 3 months, quarterly, to annually. The most frequent "frequency of purchases" is every 3 months with 584 customers. Meanwhile, the least is weekly with 539. However, the difference between the frequencies is 45, showing there's not a huge gap between the frequencies.

Recommendations

Given the results from conducting exploratory data analysis, here are some of the potential marketing and future growth recommendations:

Target Older Age Groups:

- The dataset indicates that older age groups, particularly those between 66-70, have fewer purchases. Therefore, by implementing strategies to attract these age groups can tap into a potentially lucrative market segment that might be currently underserved. A first step may be conducting market research on the reasons behind the fewer purchases. Possibilities may include providing in-store experiences or personalized assistance for digital shopping can address barriers such as unfamiliarity with technology or preference for traditional shopping methods.

Promote Clothing and Accessories:

- Clothing and accessories are identified as the highest revenue-generating categories in the dataset. Focusing marketing efforts on these categories can capitalize on existing consumer preferences

and drive further sales. By highlighting the latest trends, offering exclusive deals, and leveraging social media influencers, businesses can create buzz and excitement around these products, driving both online and in-store traffic.

Optimize Seasonal Sales:

- The dataset reveals variations in sales across different seasons, with Fall showing the highest total sales. Implementing targeted seasonal promotions and discounts can help optimize revenue generation during off-peak seasons, such as Summer. For example, offering end-of-season clearance sales, bundling related products for special holiday promotions, or introducing limited-time offers aligned with seasonal themes can stimulate demand and drive sales.

Improve Customer Satisfaction:

- Customer satisfaction is crucial for building brand loyalty and fostering long-term relationships with customers. The dataset provides insights into varying review ratings, indicating different levels of satisfaction among customers. Monitoring and addressing customer feedback promptly, ensuring product quality and timely delivery, and providing exceptional customer service are essential for maintaining high review ratings. Additionally, improving customer satisfaction may lead to an increase in purchase frequency. Therefore, implementing a robust feedback management system and actively engaging with customers through surveys, follow-up emails, and social media channels can help identify areas for improvement and demonstrate a commitment to customer satisfaction.

Geographic Targeting:

- The dataset highlights geographic disparities in purchase frequencies, with certain locations showing lower purchase rates compared to others. Tailoring marketing campaigns and promotional offers based on geographic location can help address these disparities and optimize resource allocation. By leveraging location-based data analytics and market segmentation techniques, businesses can identify target markets, customize messaging, and deploy targeted advertising campaigns to increase brand awareness and drive sales in underperforming regions.

Frequency-based Promotions:

- **Importance:** The dataset reveals insights into the frequency of purchases among customers, highlighting varying buying behaviors and preferences. Introducing special promotions or loyalty programs tailored to different purchase frequencies can incentivize repeat purchases and foster customer loyalty. By offering exclusive discounts, rewards, or personalized incentives based on purchase frequency, businesses can encourage customers to engage more frequently and increase their lifetime value. Implementing targeted marketing campaigns and automated email workflows can effectively communicate these promotions and drive customer engagement, ultimately leading to higher retention rates and increased revenue.

Conclusion

In today's fast-paced retail landscape, understanding consumer behavior is crucial for businesses striving to remain competitive and responsive to market demands. Through a comprehensive exploration of shopping trends within our dataset, this report seeks to gain valuable insights that can inform strategic decision-making, drive targeted marketing efforts, and ultimately enhance customer experiences.