CUSTOMER SHOPPING BEHAVIOUR ANALYSIS

1.Project Overview

This project analyzes customer shopping behavior using transactional data from 3,900 purchases across various product categories. The goal is to uncover insights into spending patterns, customer segments, product preferences, and subscription behavior to guide strategic business decisions.

2.Dataset Summary

- Rows: 3,900

- Columns: 18

- Key Features:

  - Customer demographics (Age, Gender, Location, Subscription Status)

  - Purchase details (Item Purchased, Category, Purchase Amount, Season, Size, Color)

  - Shopping behavior (Discount Applied, Promo Code Used, Previous Purchases, Frequency of Purchases, Review Rating, Shipping Type)

- Missing Data: 37 values in Review Rating column

3.Exploratory Data Analysis using Python

Data processing starts with Python.

- **Data Import-** Importing the data with the help of Panda.
- **Dataset Check-** Checked the datatype & statistics of dataset by df.info() and df.describe() respectively.

```
Data columns (total 19 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   Customer ID             3900 non-null   int64
 1   Age                     3900 non-null   int64
 2   Age Group               3900 non-null   object
 3   Gender                  3900 non-null   object
 4   Item Purchased          3900 non-null   object
 5   Category                3900 non-null   object
 6   Purchase Amount (USD)   3900 non-null   int64
 7   Location                3900 non-null   object
 8   Size                    3900 non-null   object
 9   Colour                  3900 non-null   object
 10  Season                  3900 non-null   object
 11  Review Rating           3863 non-null   float64
 12  Subscription Status     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  Payment Method          3900 non-null   object
 18  Frequency of Purchases  3900 non-null   object
dtypes: float64(1), int64(4), object(14)
memory usage: 579.0+ KB
```

	Customer ID	Age	Purchase Amount (USD)	Review Rating	Previous Purchases
count	3900.000000	3900.000000	3900.000000	3863.000000	3900.000000
mean	1950.500000	44.068462	59.764359	3.750065	25.351538
std	1125.977353	15.207589	23.685392	0.716983	14.447125
min	1.000000	18.000000	20.000000	2.500000	1.000000
25%	975.750000	31.000000	39.000000	3.100000	13.000000
50%	1950.500000	44.000000	60.000000	3.800000	25.000000
75%	2925.250000	57.000000	81.000000	4.400000	38.000000
max	3900.000000	70.000000	100.000000	5.000000	50.000000

- **Handling Missing Values-** Found 37 missing value from review rating. Replaced them with median value of their respective category for better efficiency.
- **Snake Casing-** Used lower case for column name and replace “space” between two words with “_” for better analysis.
- **Create new column-** Created new column purchase_frequency_days.
- **Drop column-** Dropped unnecessary column. Ex-promo_code_used.
- **Data download-** Downloaded the updated data from python & uploaded it into MySQL.

4.Data Analysis using MySQL.

Performed analysis in MySQL for answering business questions.

- **Revenue by Gender-** Revenue generated by Male vs Female customers.

gender	revenue
Male	157890
Female	75191

- **Spent high amount with discounts-** Found those customers who spent more than average amount despite applying discount coupon.

customer_id	purchase_amount
2	64
3	73
4	90
7	85
9	97
12	68
13	72
16	81
20	90
22	62
24	88
29	94
32	79
33	67
35	91
37	69
40	60
41	76
43	100

- **Highest average reviews-** Customers love these products as those products got highest average reviews.

item_purchased	average_review_ratings
Gloves	3.86
Sandals	3.84
Boots	3.82
Hat	3.8
Handbag	3.78

- **Revenue by shipping type-** This analysis shows average purchase amount of different types of shipping mode.

shipping_type	average_purchase_amount
2-Day Shipping	60.73
Express	60.48
Free Shipping	60.41
Store Pickup	59.89
Next Day Air	58.63
Standard	58.46

- **Subscription wise comparison-** This shows comparison of average purchase amount and total spending of Subscribers vs Non-Subscribers.

subscription_status	average_purchase_amount	total_revenue
No	59.87	170436
Yes	59.49	62645

- **Top items of each category-** Shows top 3 purchased items of each category.

category	item_purchased	rnk
Accessories	Jewelry	1
Accessories	Sunglasses	2
Accessories	Belt	3
Clothing	Blouse	1
Clothing	Pants	2
Clothing	Shirt	3
Footwear	Sandals	1
Footwear	Shoes	2
Footwear	Sneakers	3
Outerwear	Jacket	1
Outerwear	Coat	2

- **Customer Classification-** Classified customers as returning(1 purchase), loyal(2-15 purchase) and new(More than 15 purchase) on the basis of number of times purchased.

customer_category	count(*)
Returning	1096
Loyal	2721
New	83

- **No of Colored products-** This shows most popular colored products which are mostly liked by customers.

colour	count(*)
Olive	177
Yellow	174
Silver	173
Teal	172
Green	169

- **Seasonal products-** This shows season wise top 3 selling products along with the revenue generated through them.

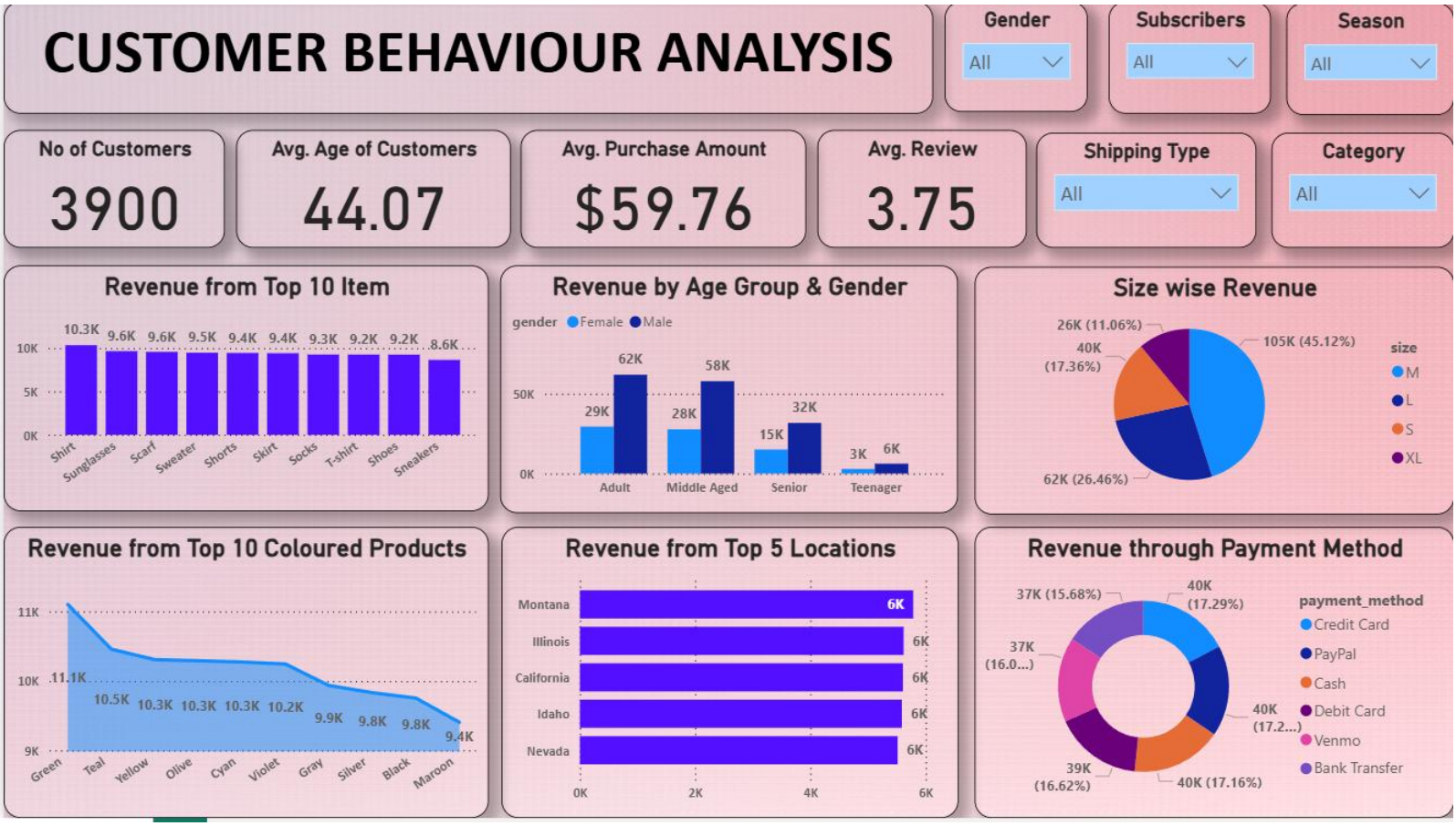
season	item_purchased	sum(purchase_amount)	rnk
Fall	Hat	3224	1
Fall	Jacket	3106	2
Fall	Handbag	2782	3
Fall	Sandals	2741	4
Fall	Blouse	2690	5
Spring	Sweater	3145	1
Spring	Skirt	2794	2
Spring	Blouse	2771	3
Spring	Shorts	2704	4
Spring	Shirt	2669	5
Summer	Jewelry	3006	1
Summer	Pants	2886	2
Summer	Shoes	2781	3
Summer	Backpack	2780	4
Summer	Scarf	2759	5
Winter	Shirt	3102	1
Winter	Sunglasses	3085	2
Winter	Pants	2999	3
Winter	Hoodie	2850	4
Winter	Jewelry	2708	5

- **Sell of products by review-** From this we can compare the selling of products by their rating status.

review_status	no_of_customers
Very Good	1634
Good	1586
Average	680

5.Dashboard creation using Power Bi.

Created an interactive dashboard for highlighting visual representation of analysis.



6.Business Recommendation

Boost Subscription Sign-Ups — Provide exclusive perks such as free shipping, early access, and reward points to encourage more customers to subscribe.

Enhance Loyalty Programs — Offer tier-based rewards to motivate returning buyers to increase their purchase frequency.

Optimize Discount Strategy — Reduce discount dependency on specific products by improving product value perception and running limited-time offers.

Promote Top-Rated Products — Highlight high-rating items in marketing campaigns to leverage customer trust and increase conversion.

Personalized Marketing by Age Group — Target high-spending age segments with customized product suggestions and relevant offers.

Express Shipping Upsell — Encourage customers to choose express delivery by bundling faster shipping with small incentives.

Bundle Complementary Products — Increase average order value by offering product bundles based on category purchase patterns.

Improve Low-Rated Items — Analyze customer feedback and enhance product quality or features for items with lower ratings.

Season-Based Campaign Planning — Increase inventory and run category-focused promotions during peak seasonal demand.

Focus on Re-Engagement — Target new or dormant customers with personalized emails, first-purchase discounts, or limited offers.

Expand Popular Color Variants — Increase supply and variety of top-selling colors to capitalize on customer preference trends.

Cross-Sell to High-Frequency Buyers — Suggest related premium products to loyal and frequent customers to maximize lifetime value.

Review Shipping Experience — Monitor feedback on shipping types and improve delivery satisfaction to boost repeat orders.

Customer Feedback Collection — Encourage more review submissions to improve product rating accuracy and guide product development.