Shopping Trends

Customers Analysis and Segmentation

By Vanshika Rawat

Table of Contents :

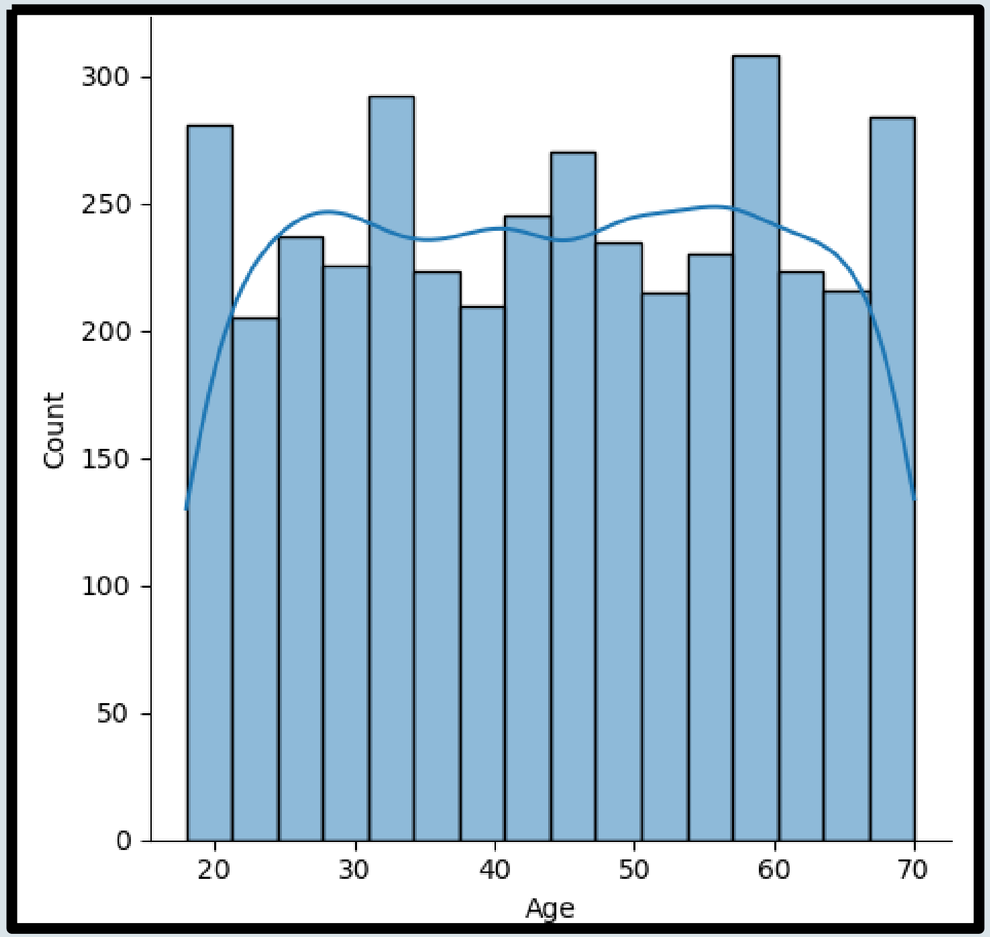
1.Introduction

- The purpose of this project is to learn about customers trends and segmentation and as a result, **improve marketing strategies** and **increase profits** and **customer trust.**
- Dataset contains **3900 rows** of unique purchases described by **18 features** like Customer ID, Age, Gender, Item Purchased etc.
- To analyze the influence of factors such as pricing, discounts, and seasonal changes on shopping behaviors.

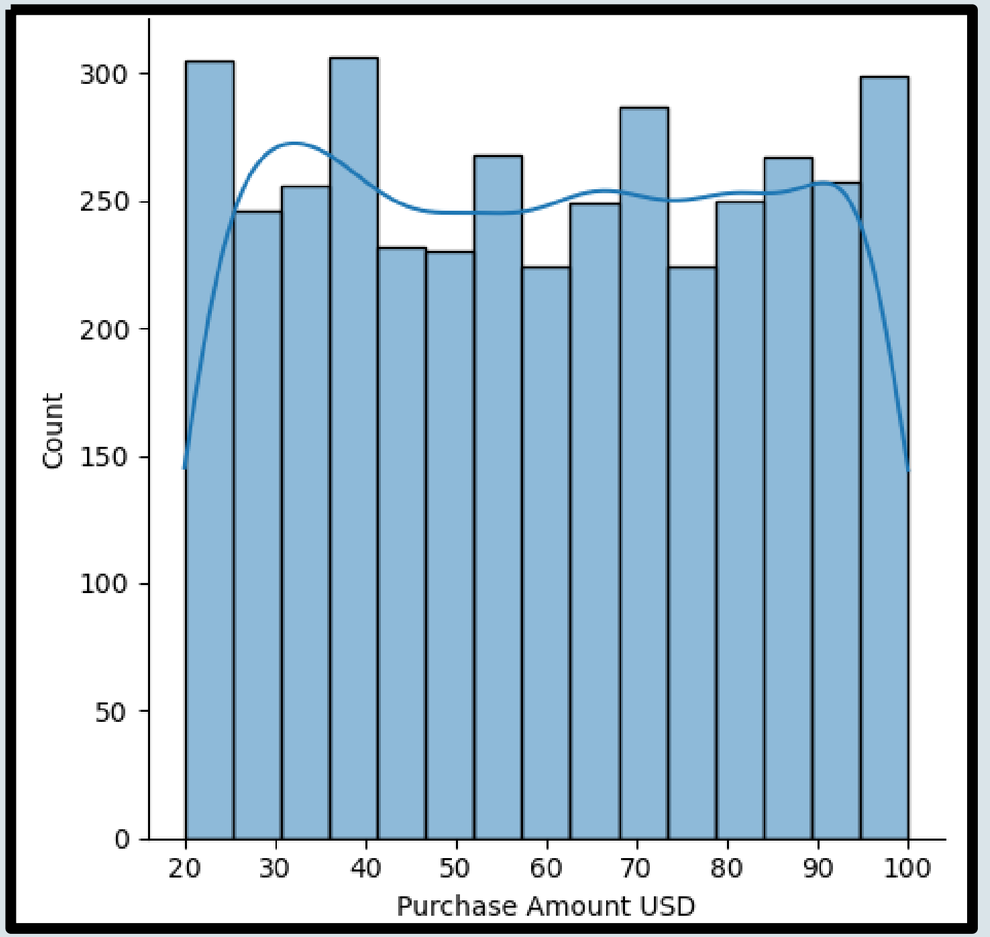
2.Descriptive Analysis

Distributions and basic statistic values of discrete features :

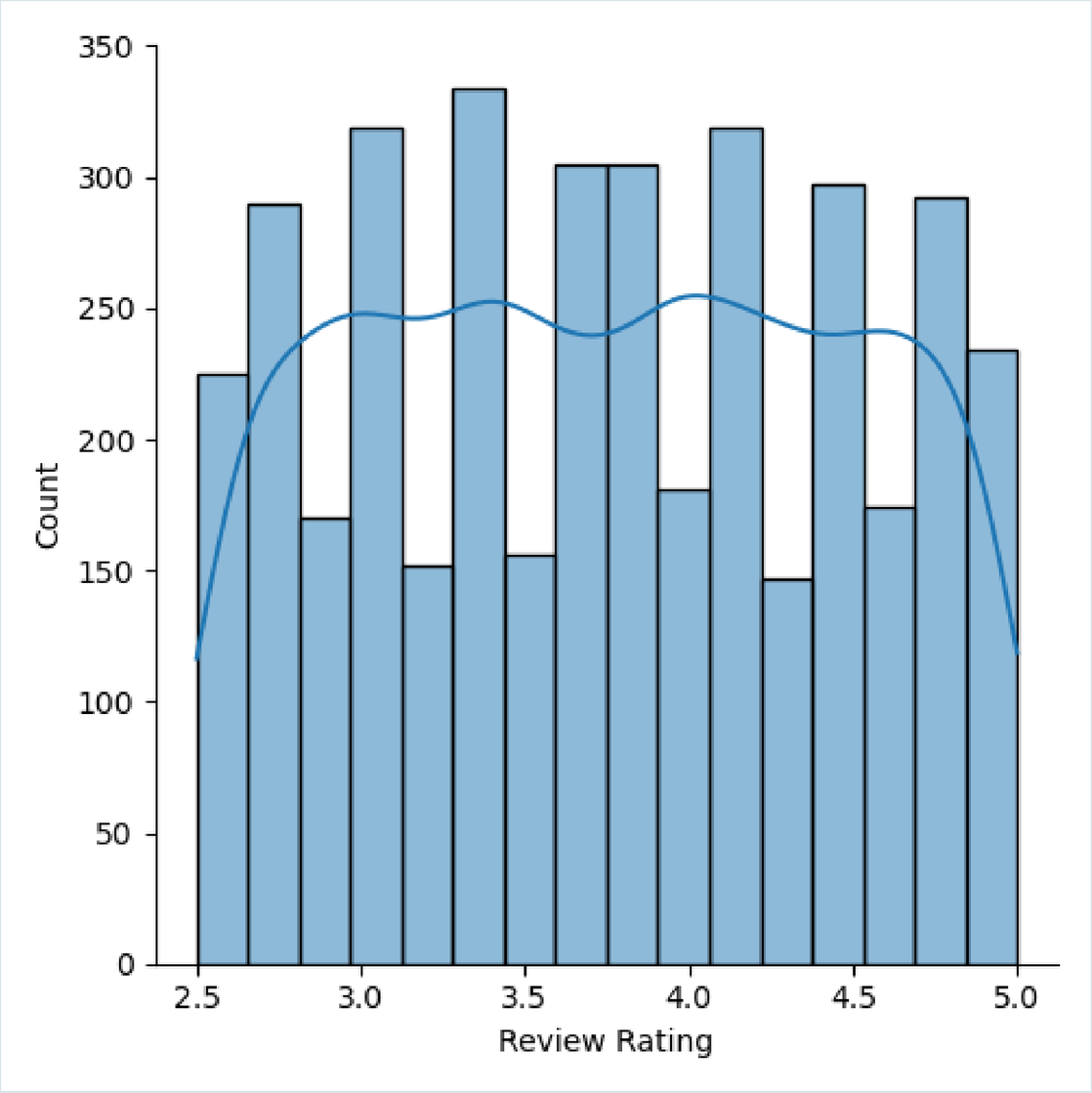
Age: mean:**~44**,
min.**18** max. **70**



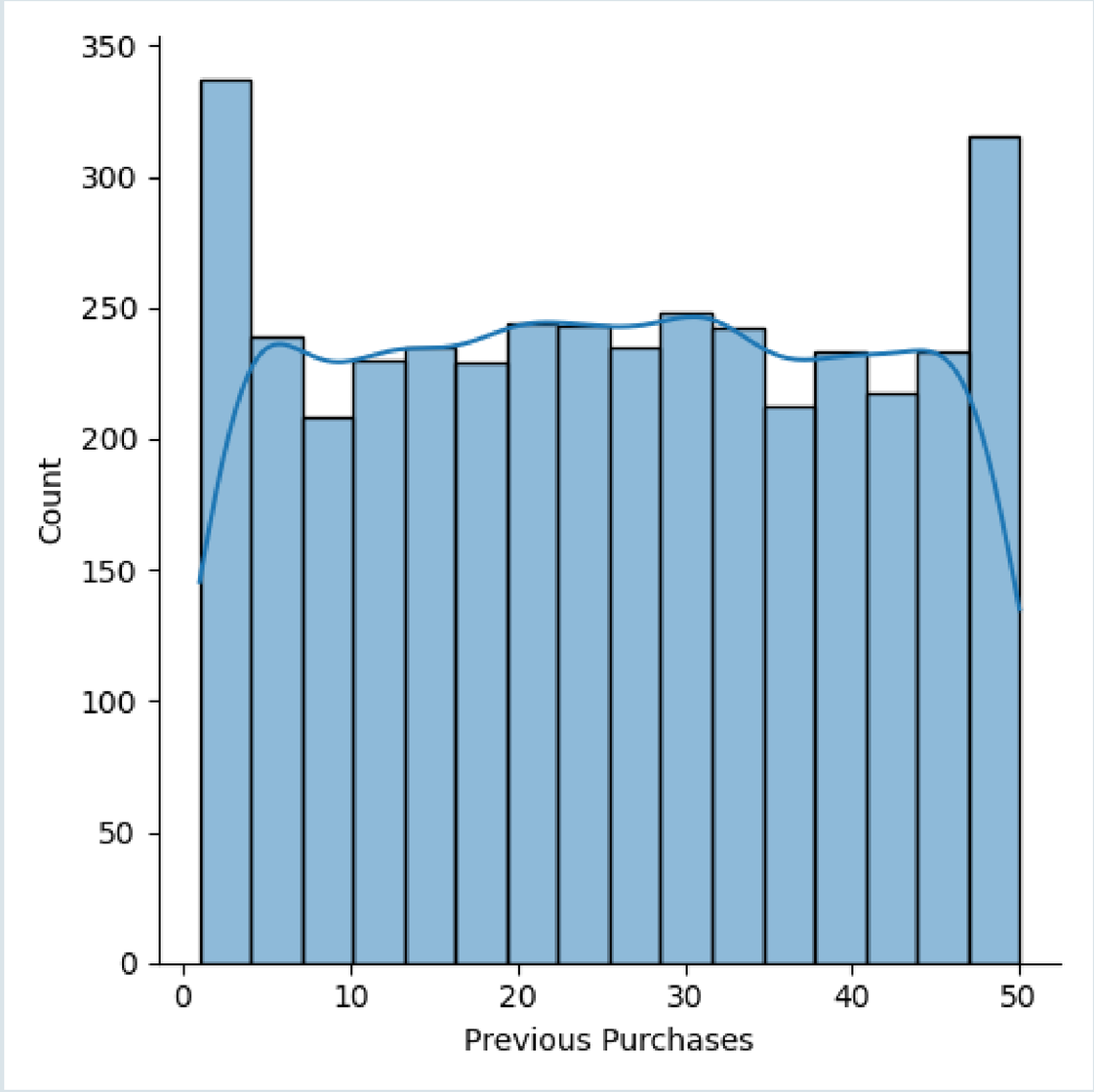
Purchase Amount USD:
mean:**~59.76**, min.**20** max. **100**



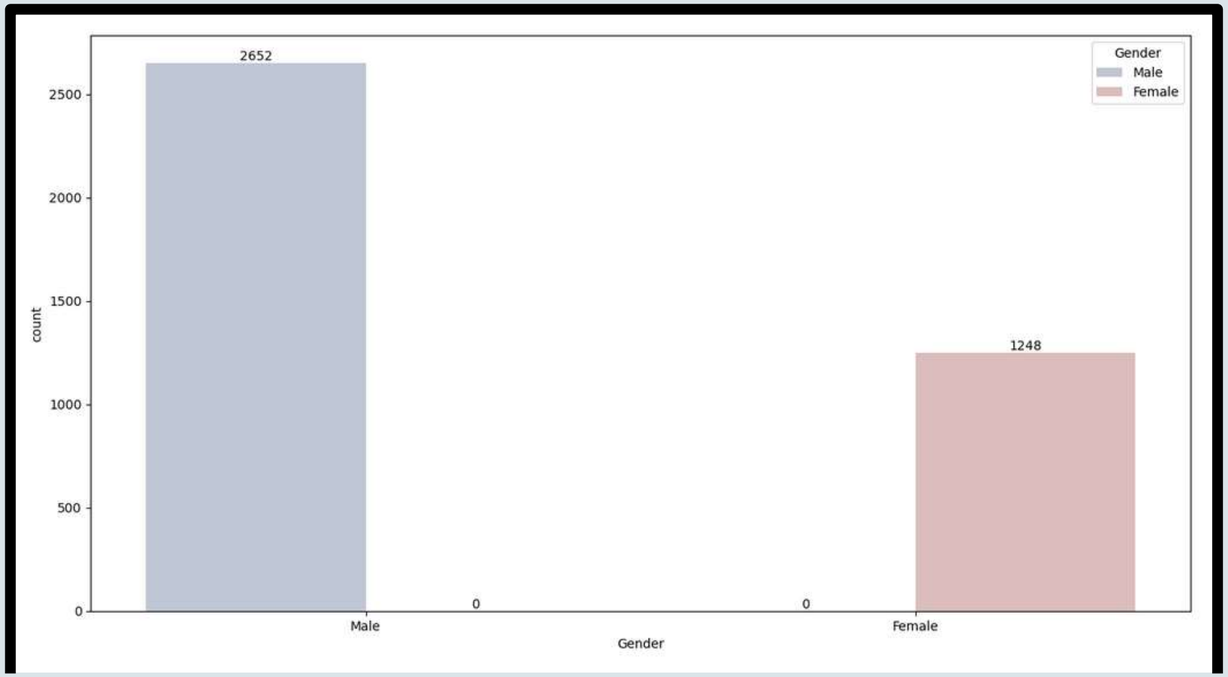
Review Rating: mean:**~3.75**, min.**2.5** max. **5.0**

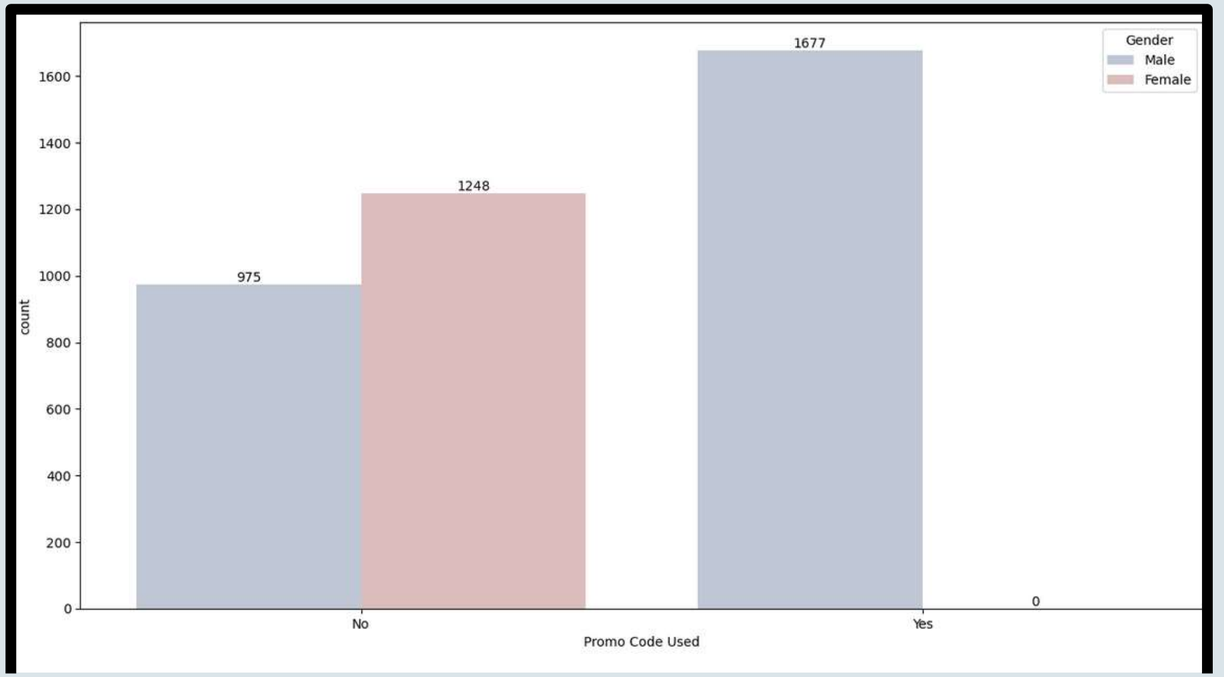


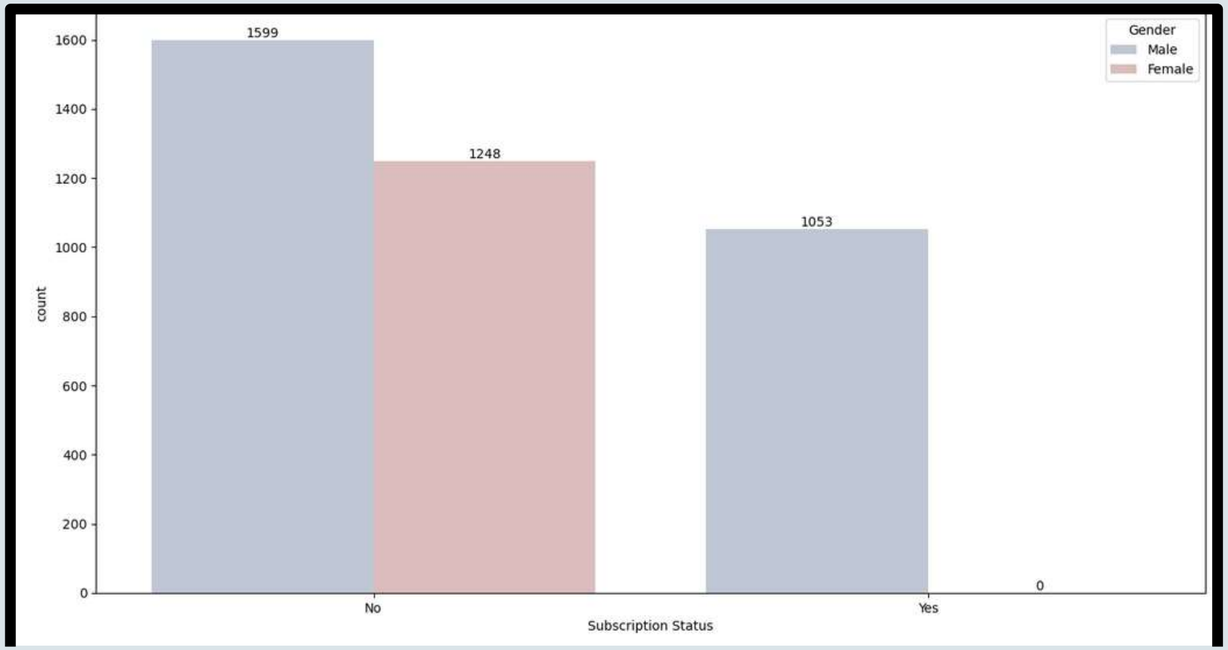
Previous Purchases:
mean:**~25**, min.**1** max. 5**0**

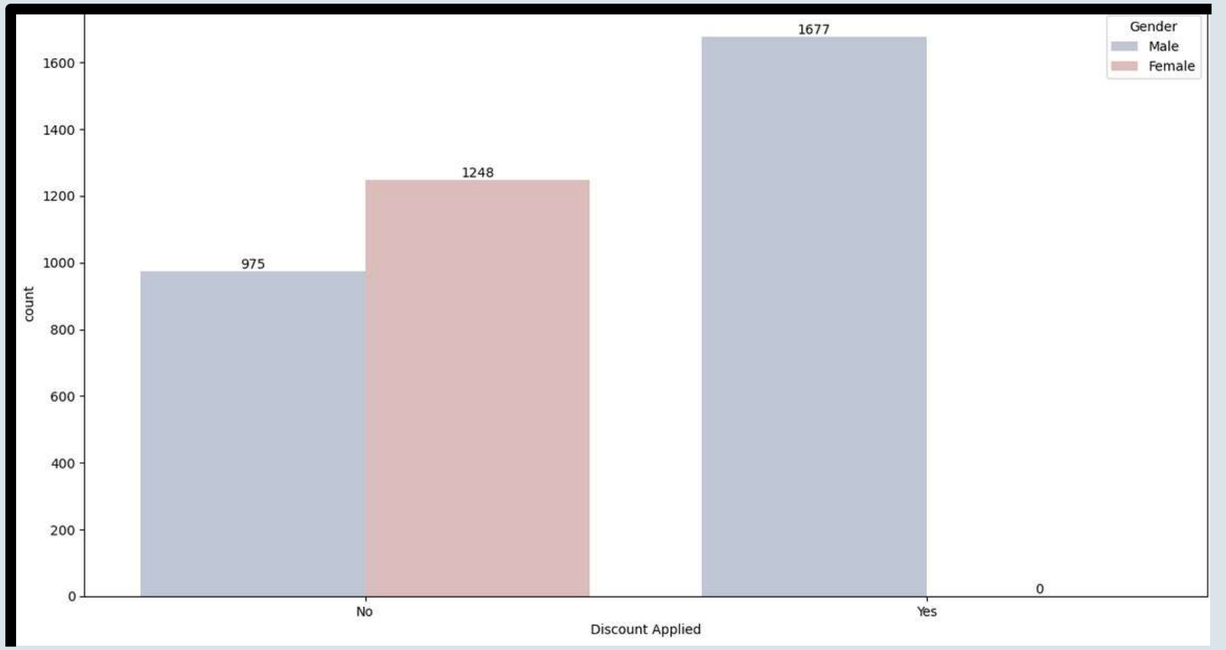


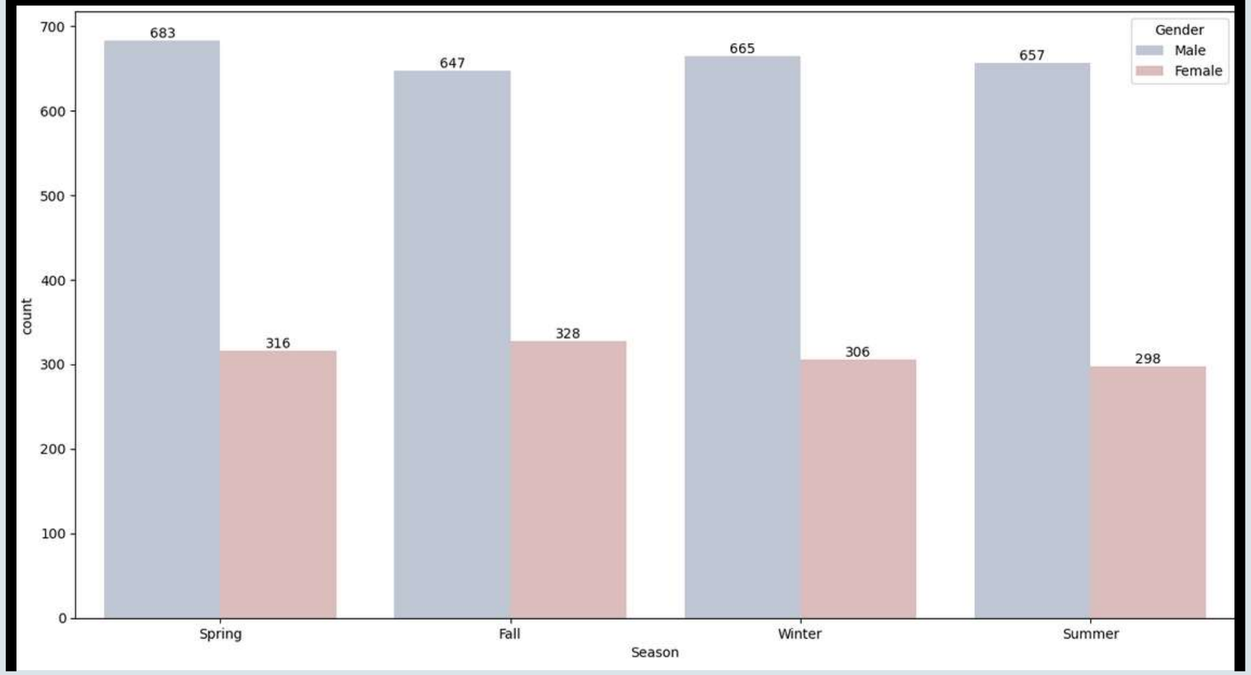
Categorical features and their most common values

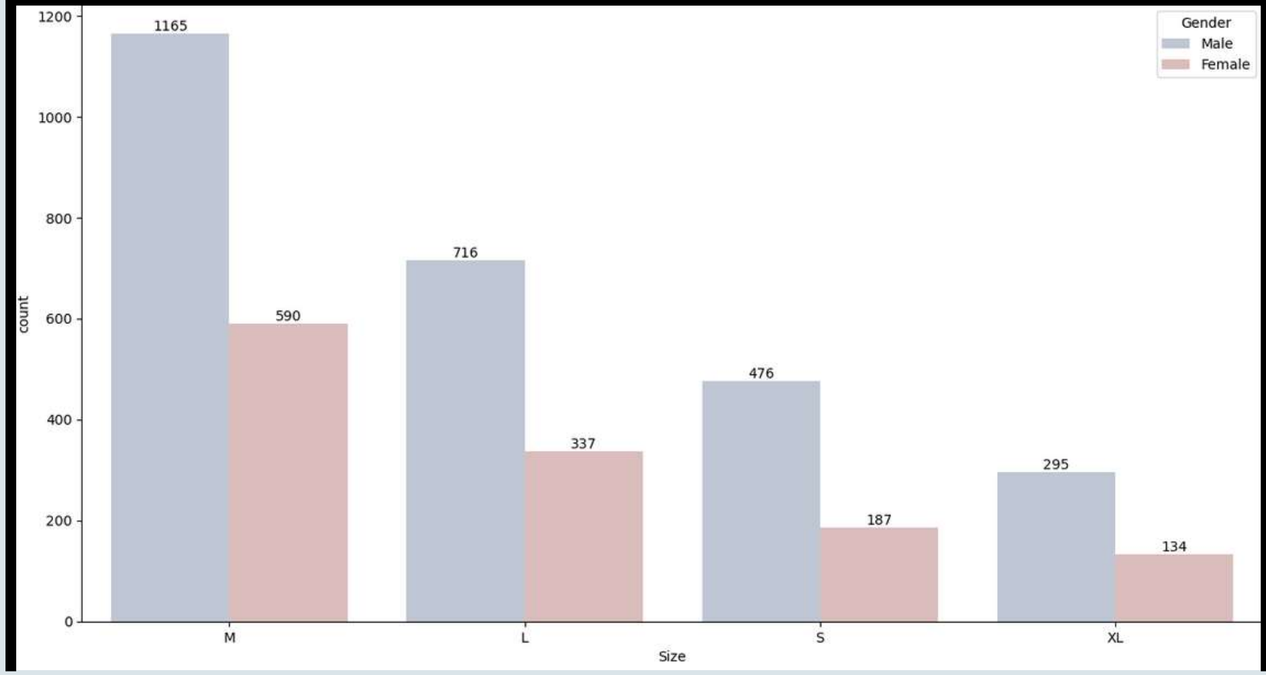


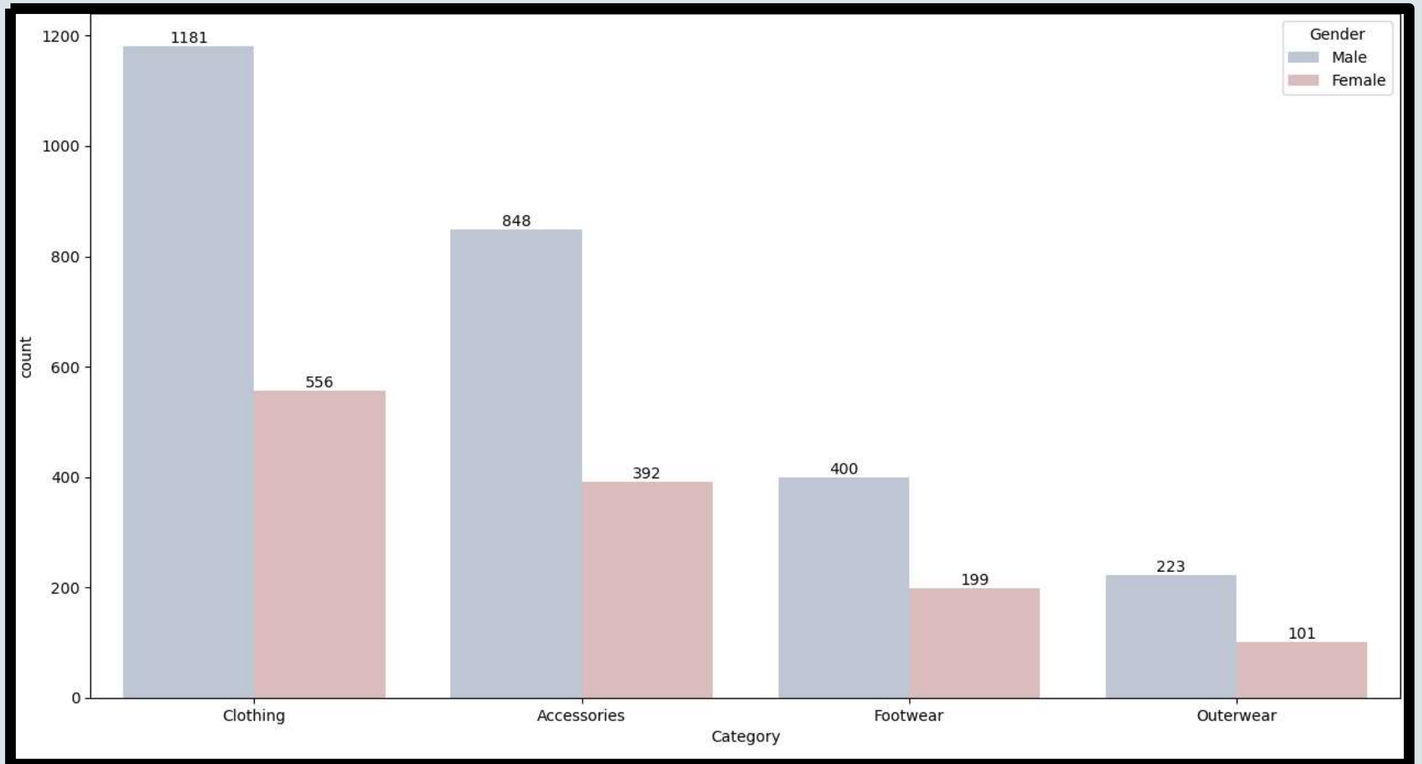


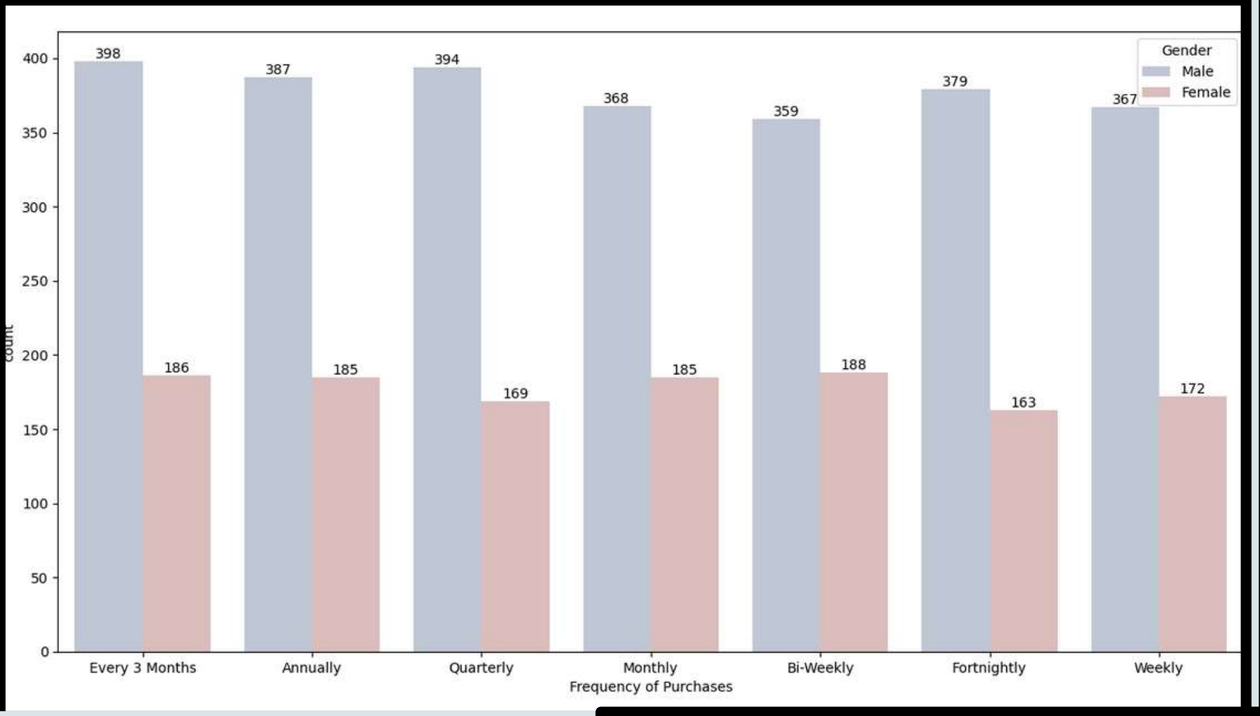


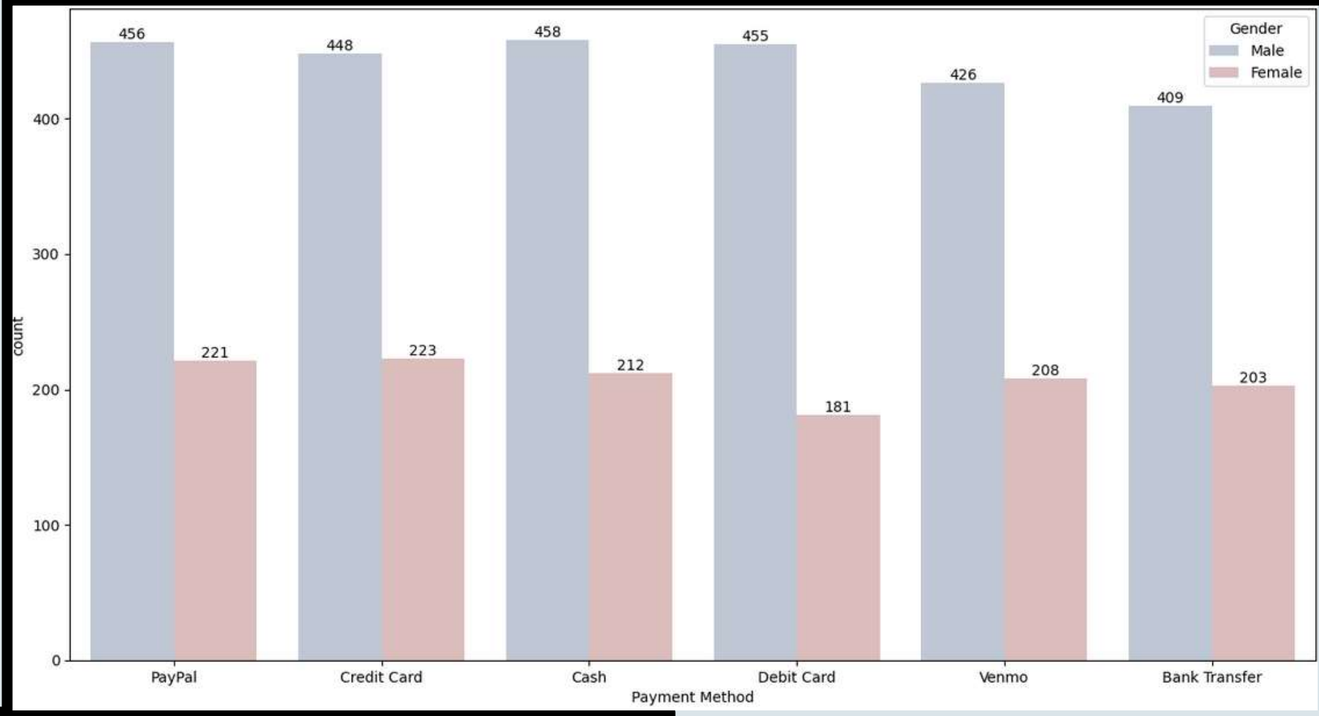


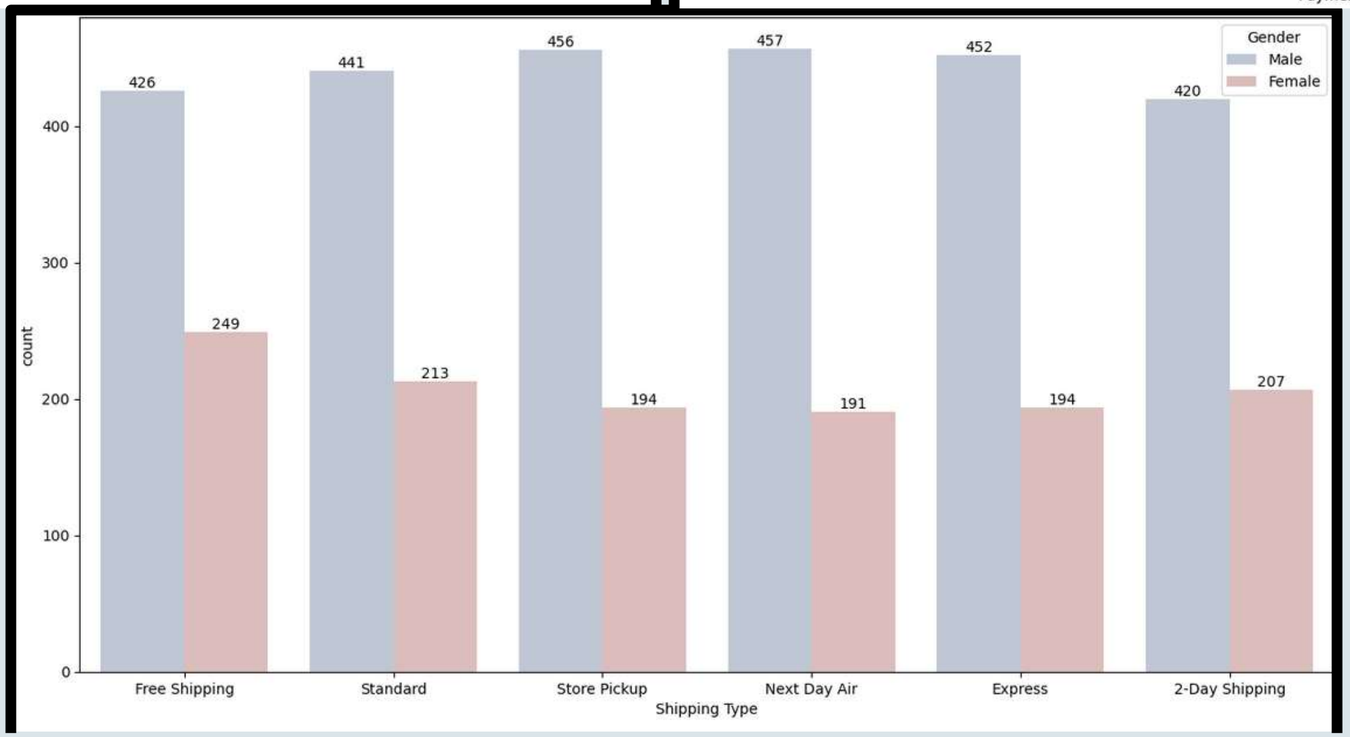

Gender
Male
Female
count
Season
Spring
Fall
Winter
Summer
683
316
647
328
665
306
657
298

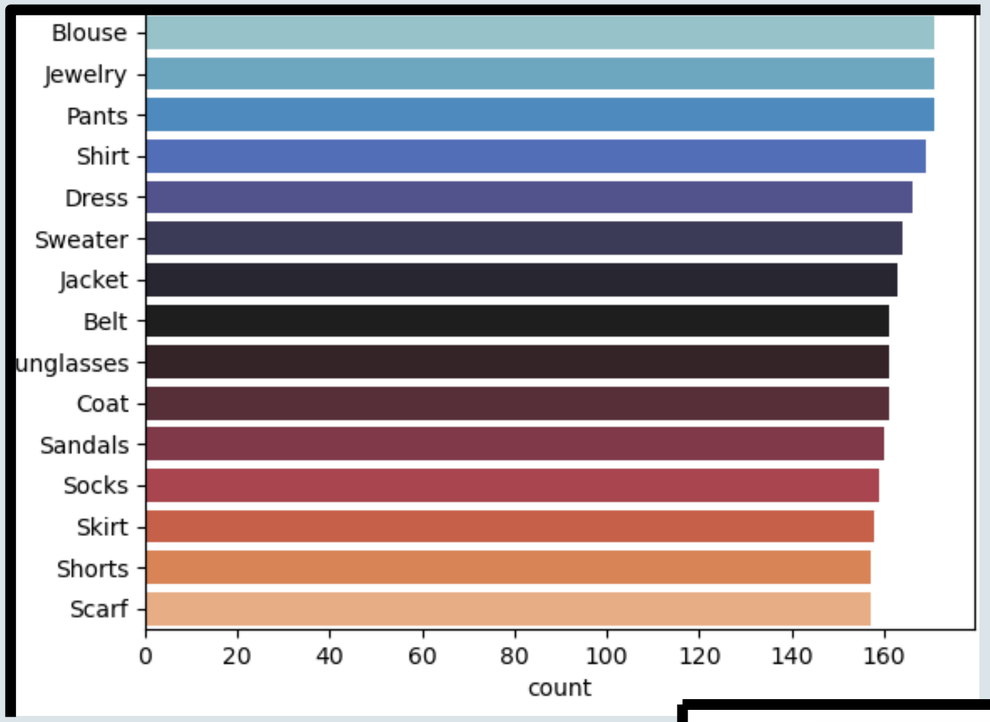

Gender
Male
Female
count
Size
M
L
S
XL
1165
590
716
337
476
187
295
134

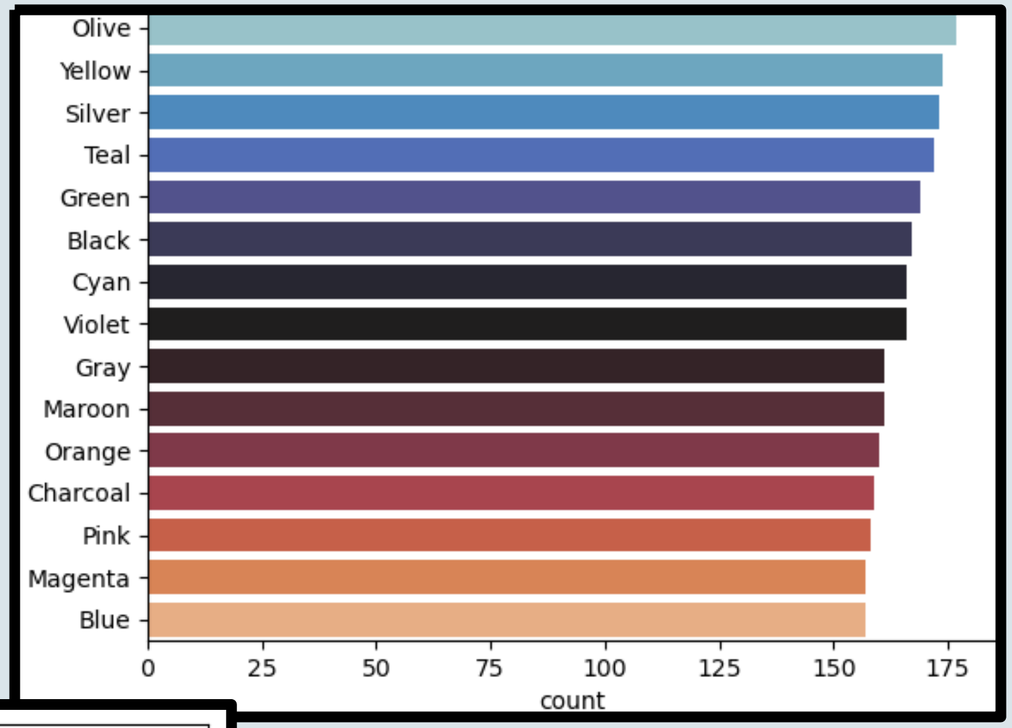

Gender
Male
Female
count
Category
Clothing
Accessories
Footwear
Outerwear
1181
556
848
392
400
199
223
101

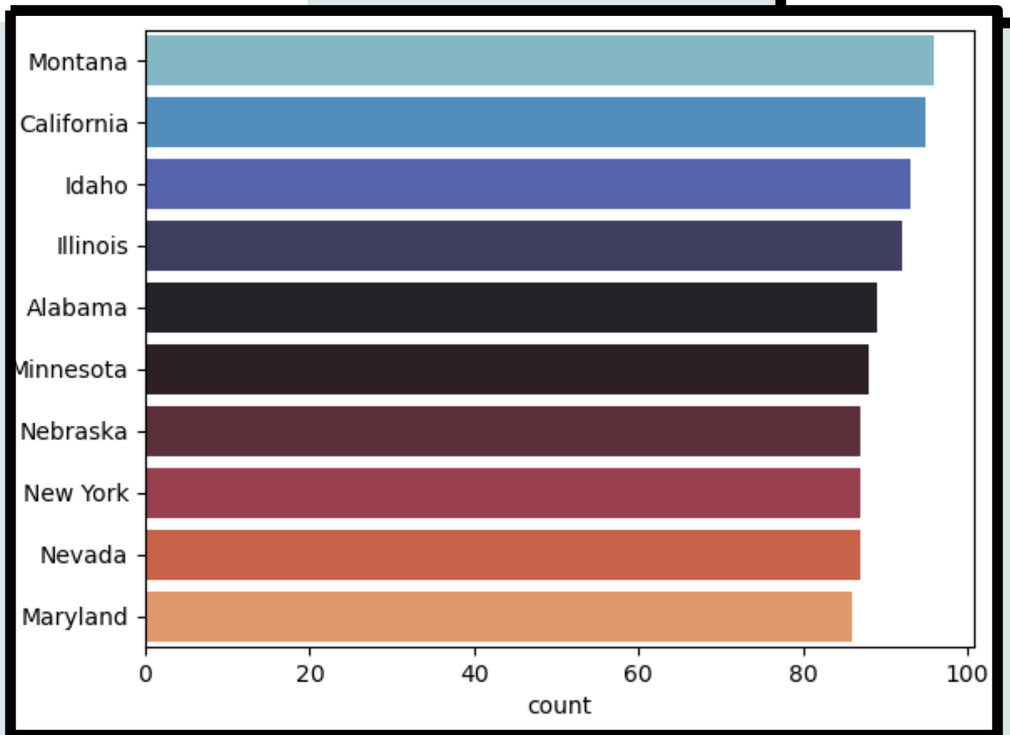

Gender
Male
Female
count
Every 3 Months
Annually
Quarterly
Monthly
Bi-Weekly
Fortnightly
Weekly
Frequency of Purchases
398
186
387
185
394
169
368
185
359
188
379
163
367
172

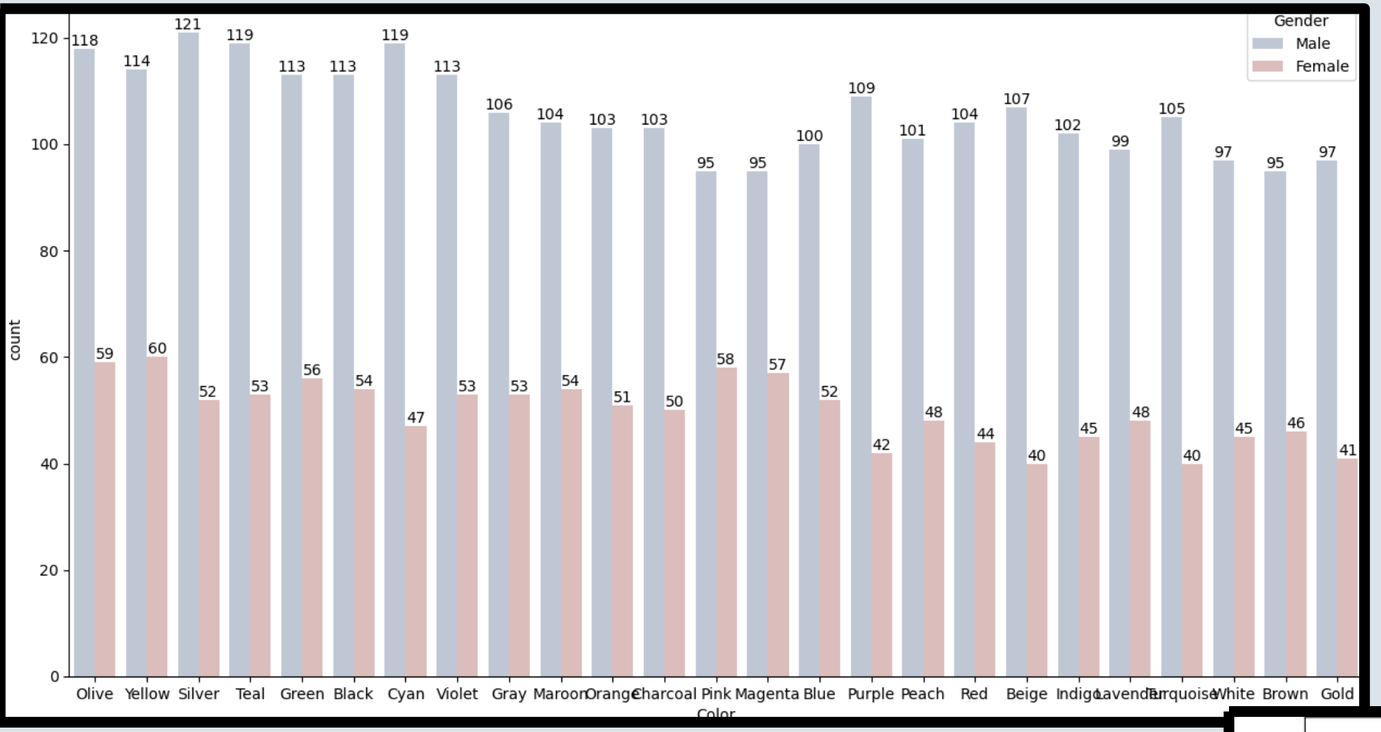

Gender
Male
Female
count
PayPal
Credit Card
Cash
Debit Card
Venmo
Bank Transfer
Payment Method
456
221
448
223
458
212
455
181
426
208
409
203

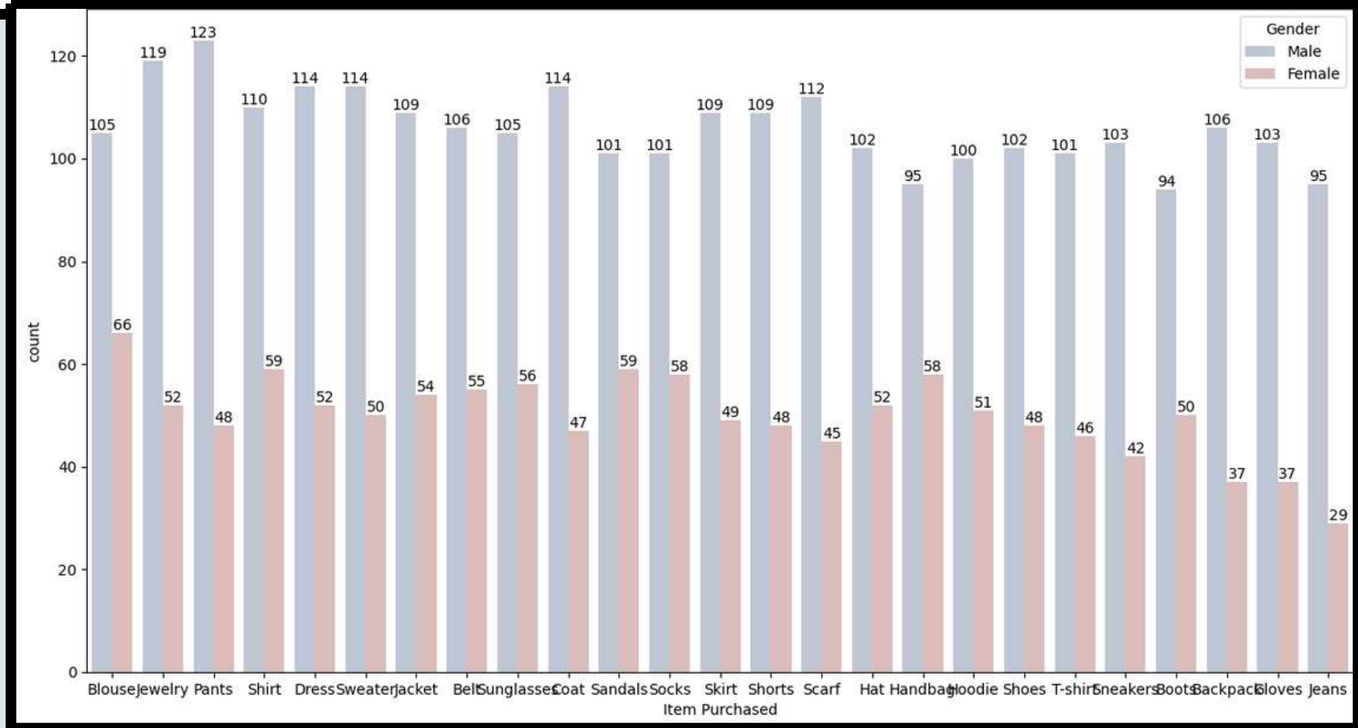

Gender
Male
Female
count
Free Shipping
Standard
Store Pickup
Next Day Air
Express
2-Day Shipping
Shipping Type
426
249
441
213
456
194
457
191
452
194
420
207


Blouse
Jewelry
Pants
Shirt
Dress
Sweater
Jacket
Belt
unglasses
Coat
Sandals
Socks
Skirt
Shorts
Scarf
0
20
40
60
80
100
120
140
160
count

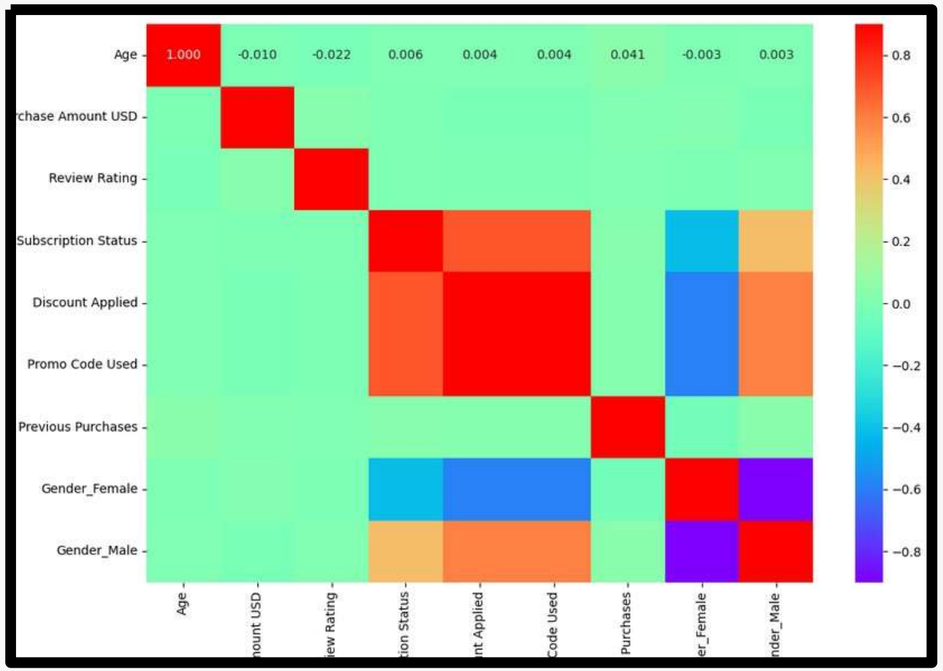

Olive
Yellow
Silver
Teal
Green
Black
Cyan
Violet
Gray
Maroon
Orange
Charcoal
Pink
Magenta
Blue
0
25
50
75
100
125
150
175
count

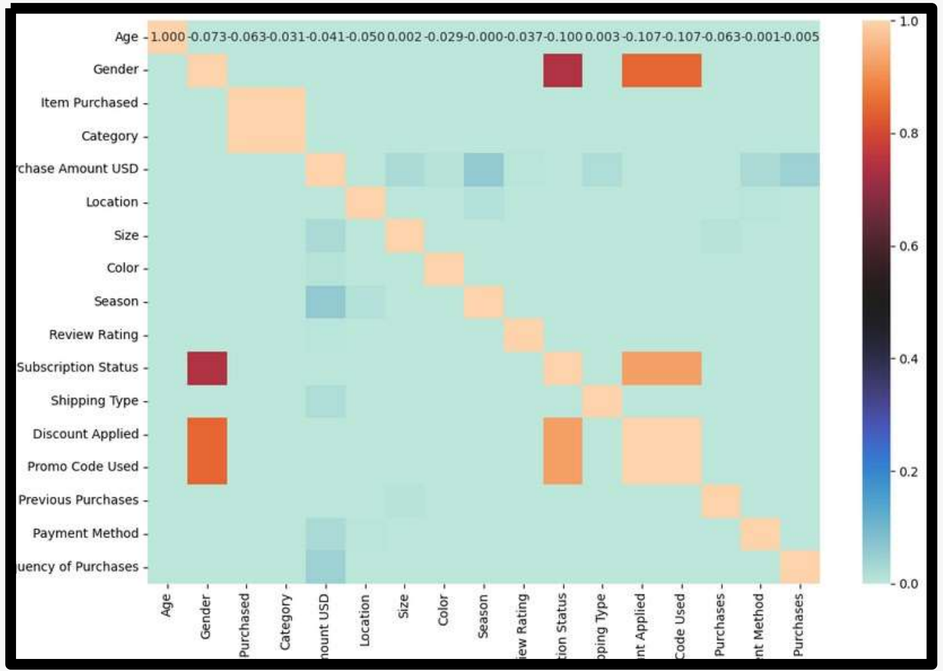

Montana
California
Idaho
Illinois
Alabama
Minnesota
Nebraska
New York
Nevada
Maryland
0
20
40
60
80
100
count


Gender
Male
Female
count
Color
Olive Yellow Silver Teal Green Black Cyan Violet Gray Maroon Orange Charcoal Pink Magenta Blue Purple Peach Red Beige Indigo Lavender Turquoise White Brown Gold
118 59
114 60
121 52
119 53
113 56
113 54
119 47
113 53
106 53
104 54
103 51
103 50
95 58
95 57
100 52
109 42
101 48
104 44
107 40
102 45
99 48
105 40
97 45
95 46
97 41


Gender
Male
Female
count
Item Purchased
Blouse Jewelry Pants Shirt Dress Sweater Jacket Belt Sunglasses Coat Sandals Socks Skirt Shorts Scarf Hat Handbag Hoodie Shoes T-shirt Sneakers Boots Backpack Gloves Jeans
105 66
119 52
123 48
110 59
114 52
114 50
109 54
106 55
105 56
114 47
101 59
101 58
109 49
109 48
112 45
102 52
95 58
100 51
102 48
101 46
103 42
94 50
106 37
103 37
95 29

Possible Correlations





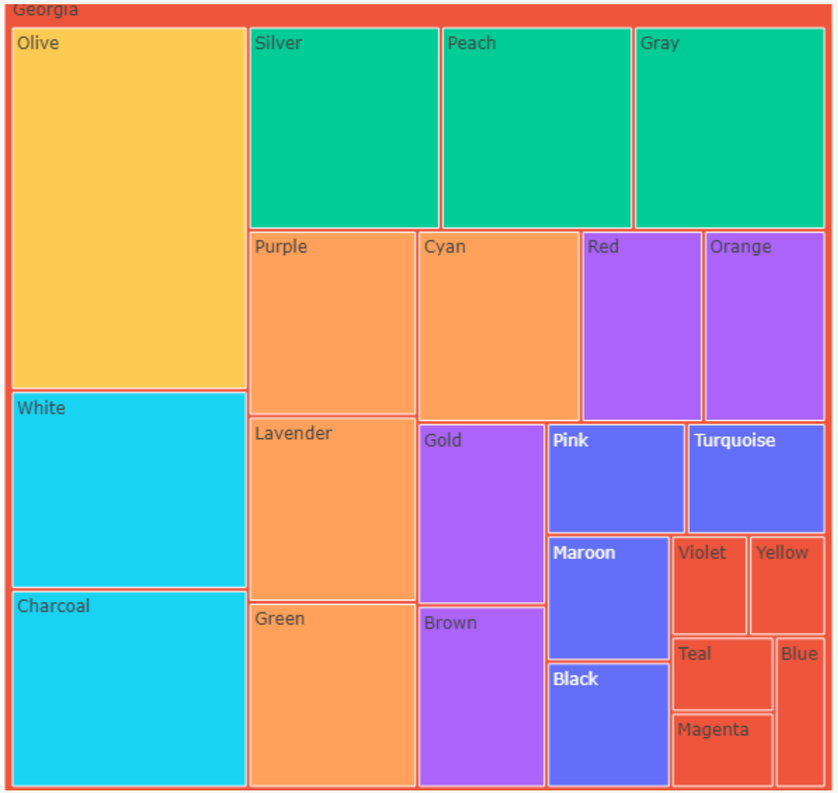
--> For discrete data Spearman's correlation :

- significant correlation **between 'Subscription Status' and 'Gender'** probably because of zero 'female' with 'Subscription Status'=Yes
- significant correlation between 'Subscription Status' and 'Promo Code Used','Discount Applied'
- significant correlation **(pvalue<.05) between "Age" and "Previous Purchases"**

--> For categorical data Cramer's V correlation :

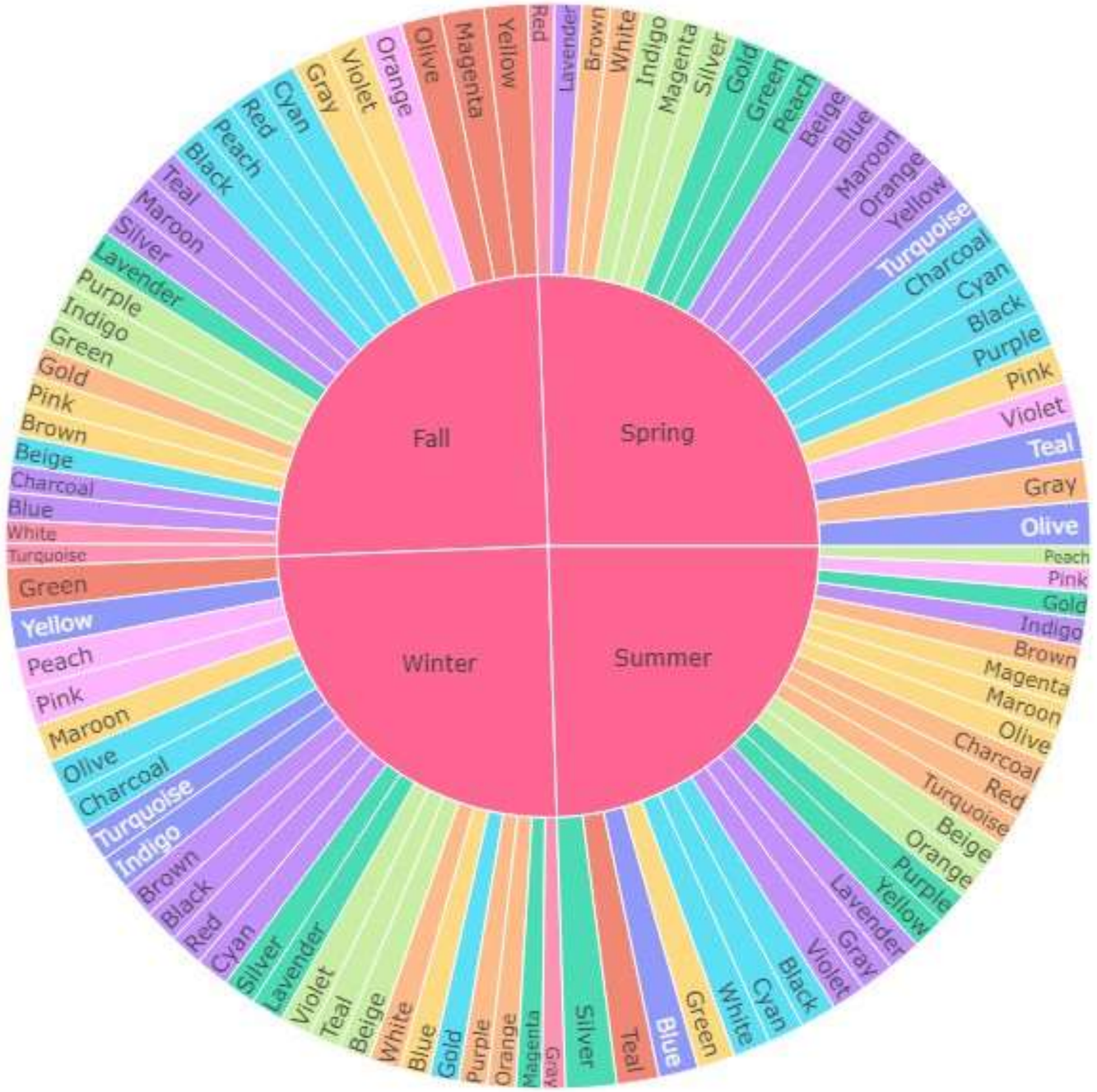
- Some changes **across column "Purchase amount USD"**

3.Shopping Trends



Most common colors grouped by locations :

- 0 Georgia Olive 11
- 1 California Turquoise 11
- 2 Tennessee Cyan 10
- 3 Idaho Black 10
- 4 North Carolina Magenta 9
- 5 Nevada Orange 9
- 6 Minnesota Purple 9
- 7 Maine Yellow 9
- 8 New Mexico Olive 8
- 9 Illinois Gray 8
- 10 Kansas White 8

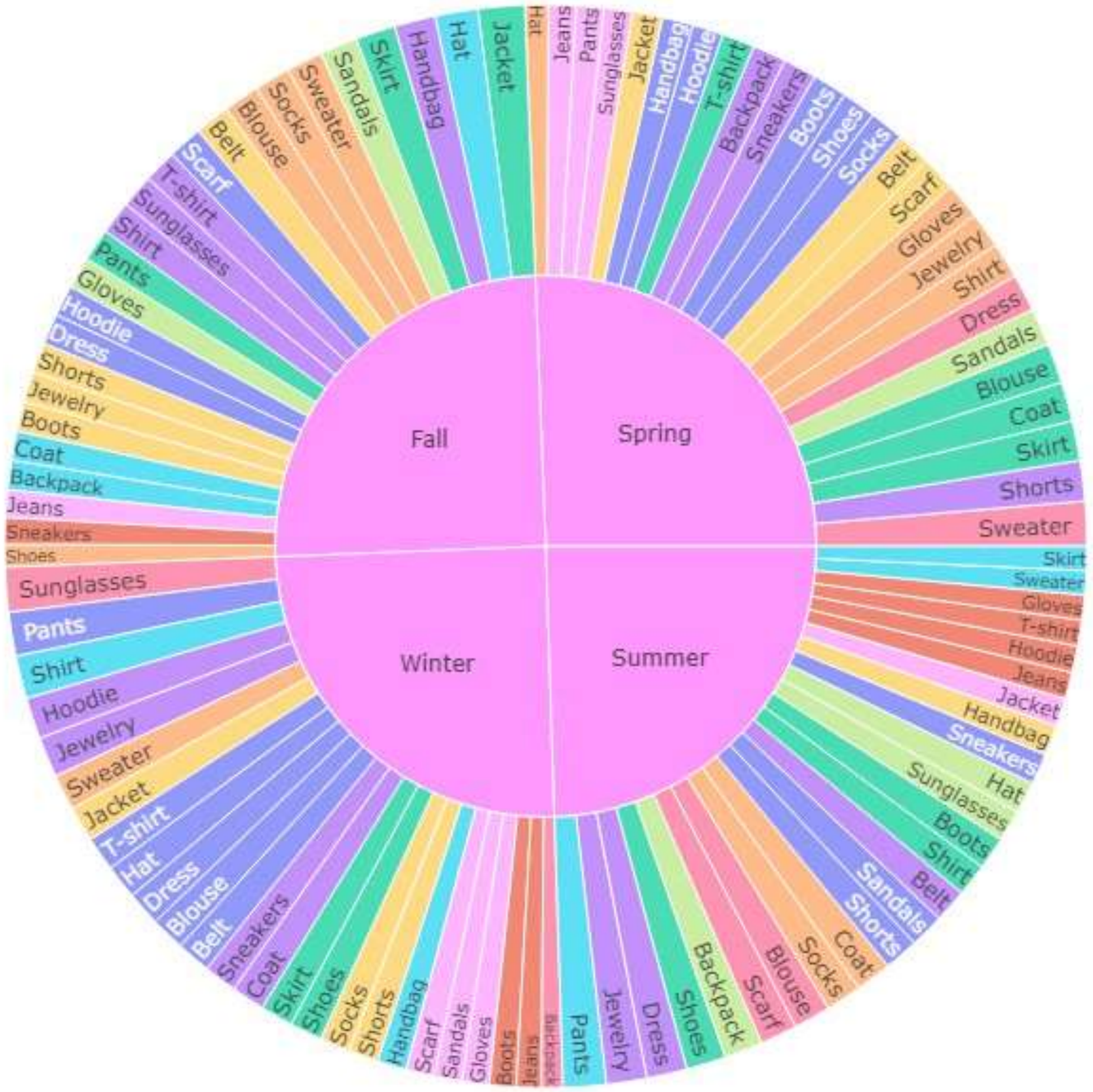


Most common colors grouped by seasons :

- 0 Summer Silver 59
- 1 Spring Olive 52
- 2 Winter Green 50
- 3 Fall Magenta 50
- 4 Fall Yellow 50
- 5 Summer Teal 49
- 6 Spring Gray 48
- 7 Fall Olive 47
- 8 Winter Yellow 46
- 9 Summer Blue 46
- 10 Spring Teal 46
- 11 Winter Peach 45
- 12 Spring Violet 45
- 13 Winter Pink 45
- 14 Fall Orange 45

Most common colors grouped by genders :

- 0 Silver Male 121
- 1 Cyan Male 119
- 2 Teal Male 119
- 3 Olive Male 118
- 4 Yellow Male 114
- 5 Green Male 113
- 6 Violet Male 113
- 7 Black Male 113
- 8 Purple Male 109
- 9 Beige Male 107
- Yellow Female 60
- Olive Female 59
- Pink Female 58
- Magenta Female 57
- Green Female 56
- Maroon Female 54
- Black Female 54
- Teal Female 53
- Violet Female 53
- Gray Female 53



Most common items grouped by seasons :

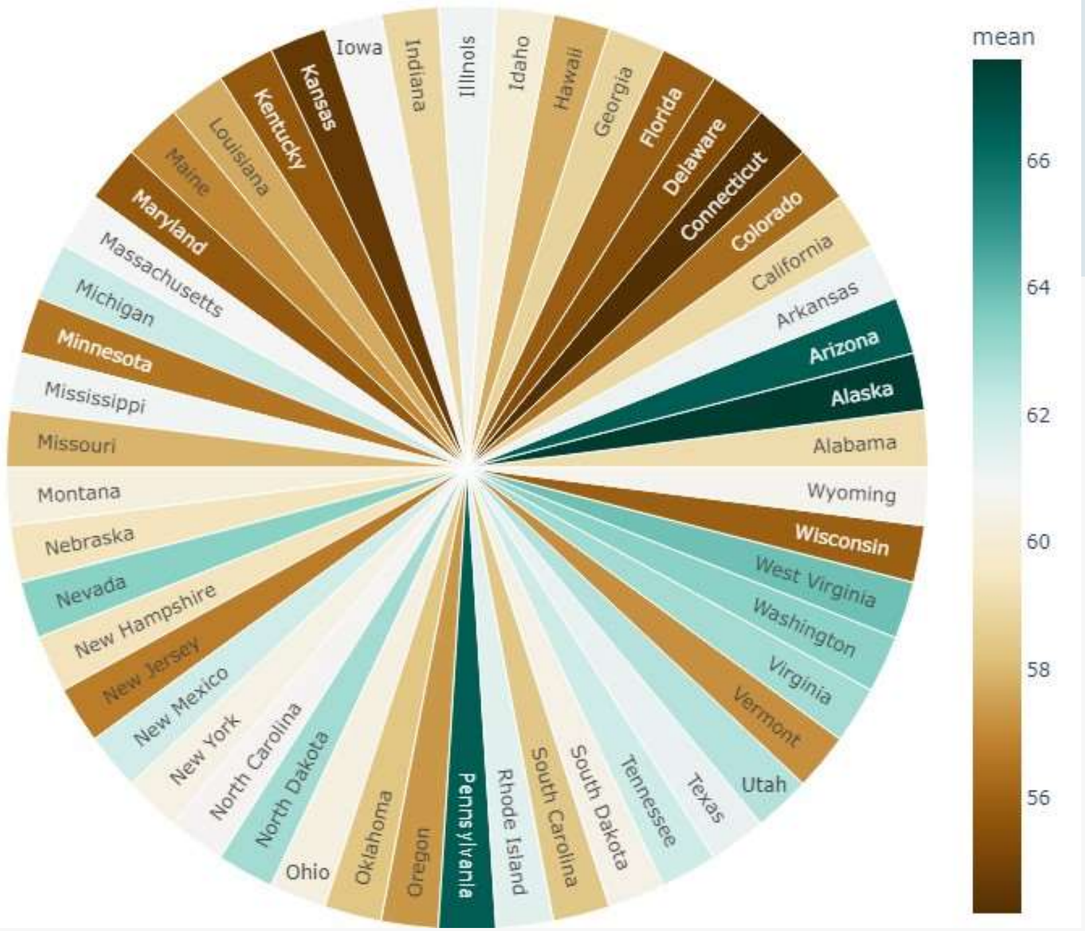
- 0 Fall Jacket 54
- 1 Spring Sweater 52
- 2 Winter Sunglasses 52
- 3 Winter Pants 51
- 4 Fall Hat 50
- 5 Winter Shirt 50
- 6 Summer Pants 50
- 7 Fall Handbag 48
- 8 Winter Hoodie 48
- 9 Summer Jewelry 47
- 10 Spring Shorts 47
- 11 Summer Dress 47
- 12 Winter Jewelry 47
- 13 Spring Skirt 46
- 14 Spring Coat 46

Most common items grouped by genders :

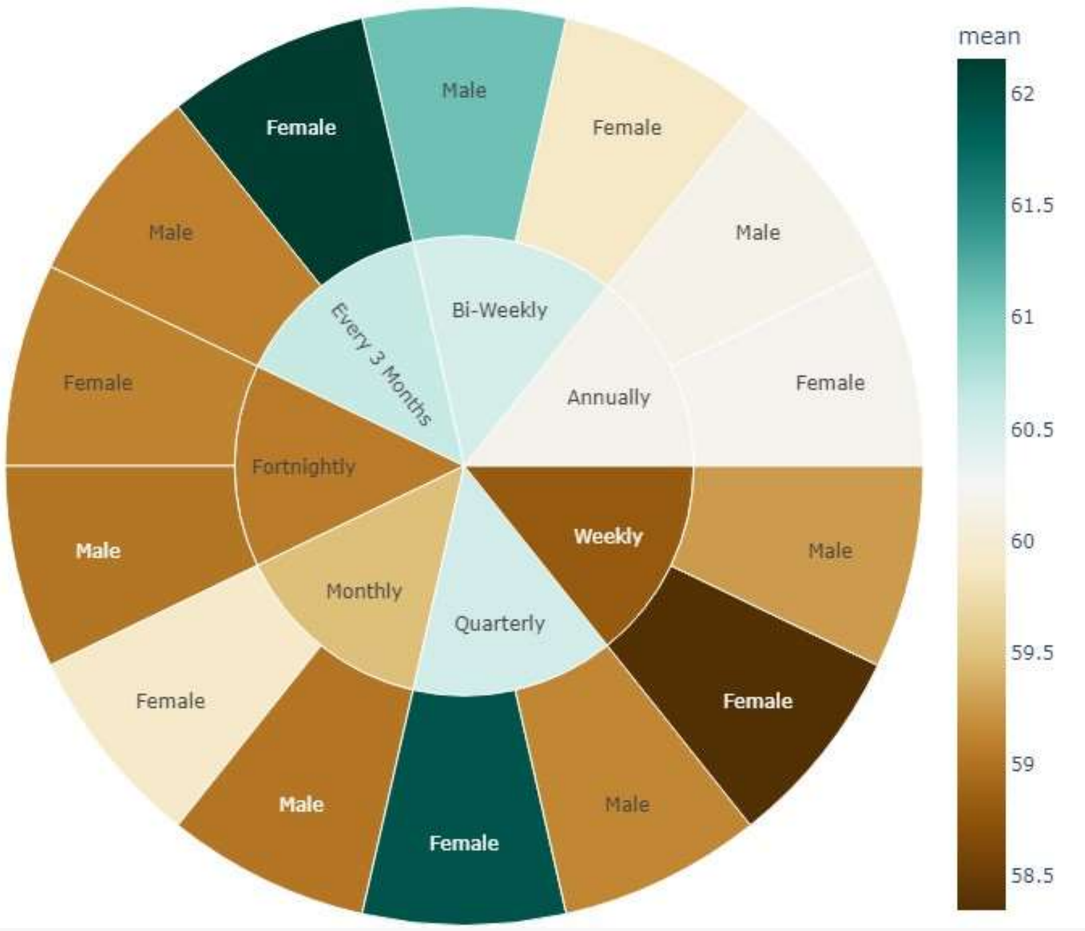
• 1	Pants	Male	123	• Blouse Female 66
• 2	Jewelry	Male	119	• Sandals Female 59
• 3	Sweater	Male	114	• Shirt Female 59
• 3	Coat	Male	114	• Handbag Female 58
• 4	Dress	Male	114	• Socks Female 58
• 5	Scarf	Male	112	• Sunglasses Female 56
• 6	Shirt	Male	110	• Belt Female 55
• 7	Shorts	Male	109	• Jacket Female 54
• 8	Skirt	Male	109	• Dress Female 52
• 9	Jacket	Male	109	• Hat Female 52

Purchase Amount(USD) across different groups

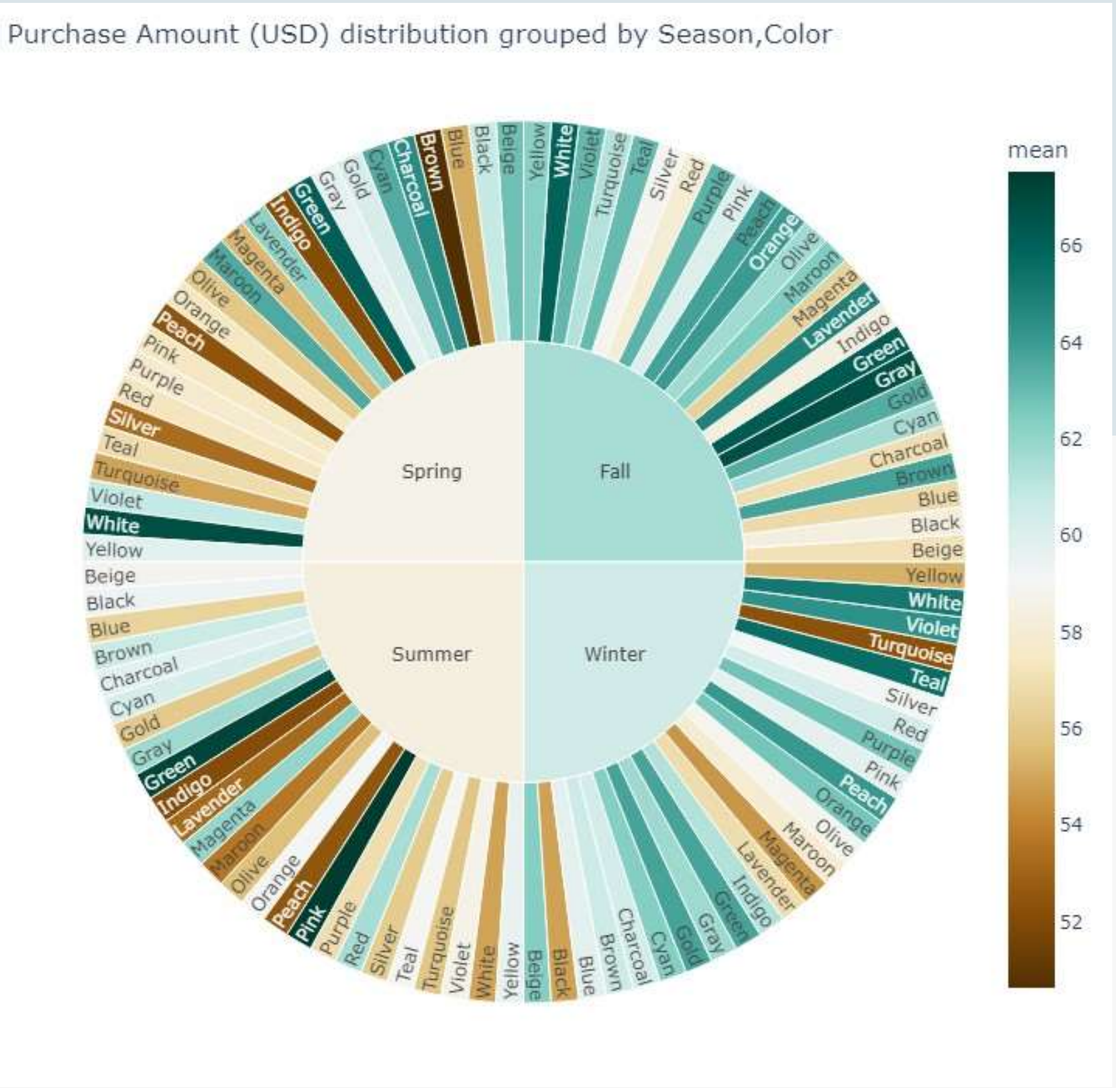
Purchase Amount (USD) distribution grouped by Location

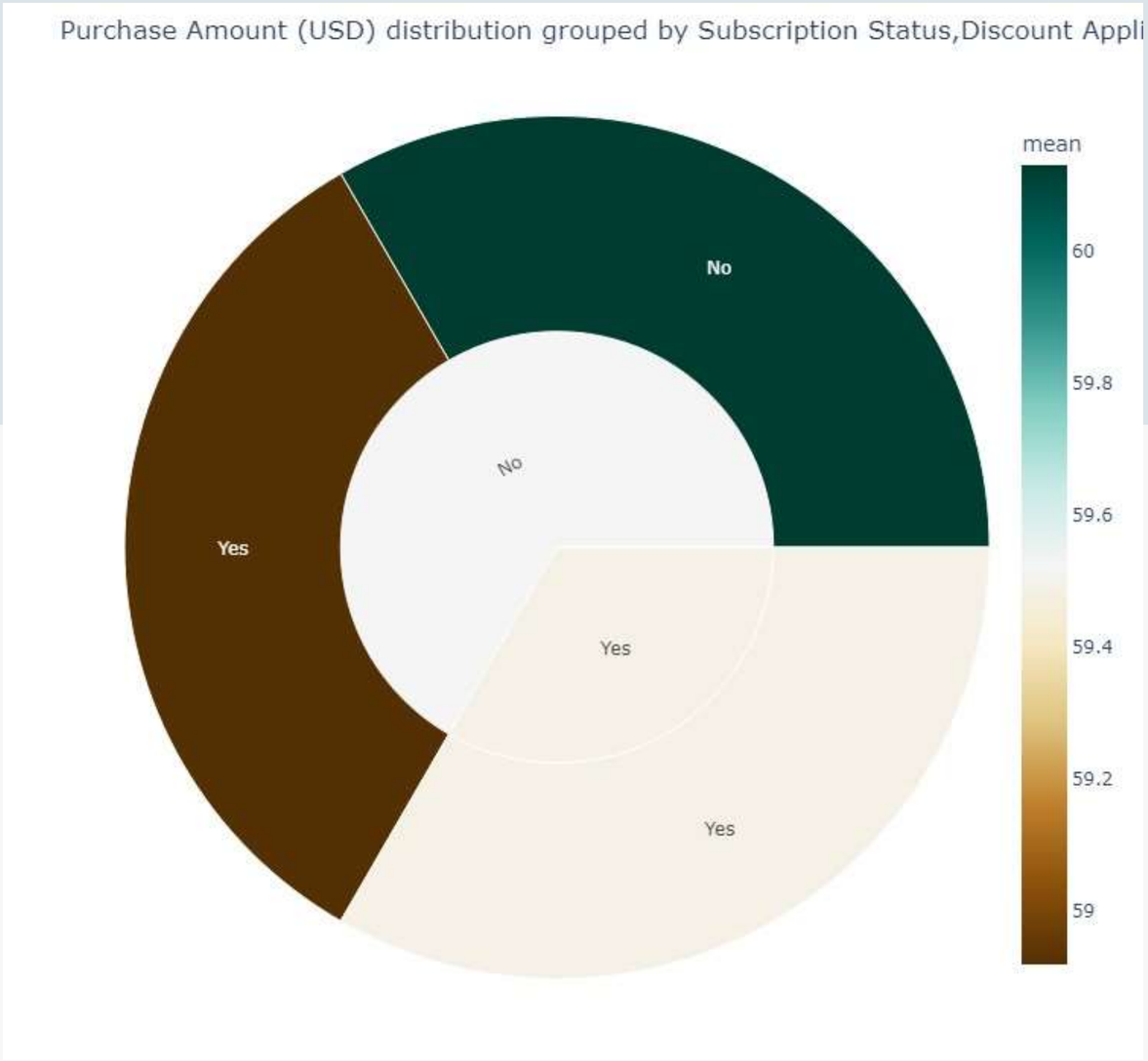


Purchase Amount (USD) distribution grouped by Frequency of Purchases,Gender

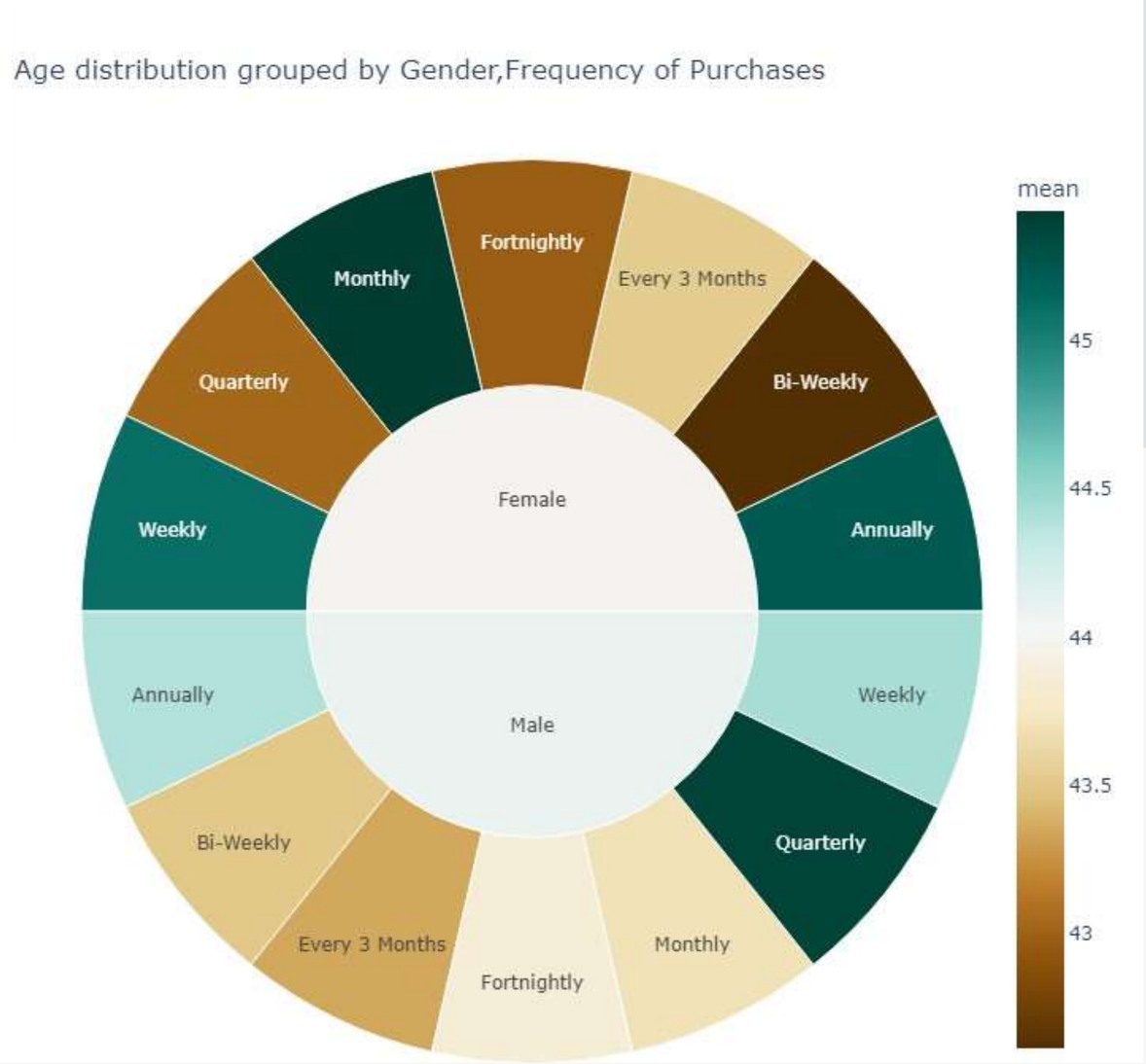


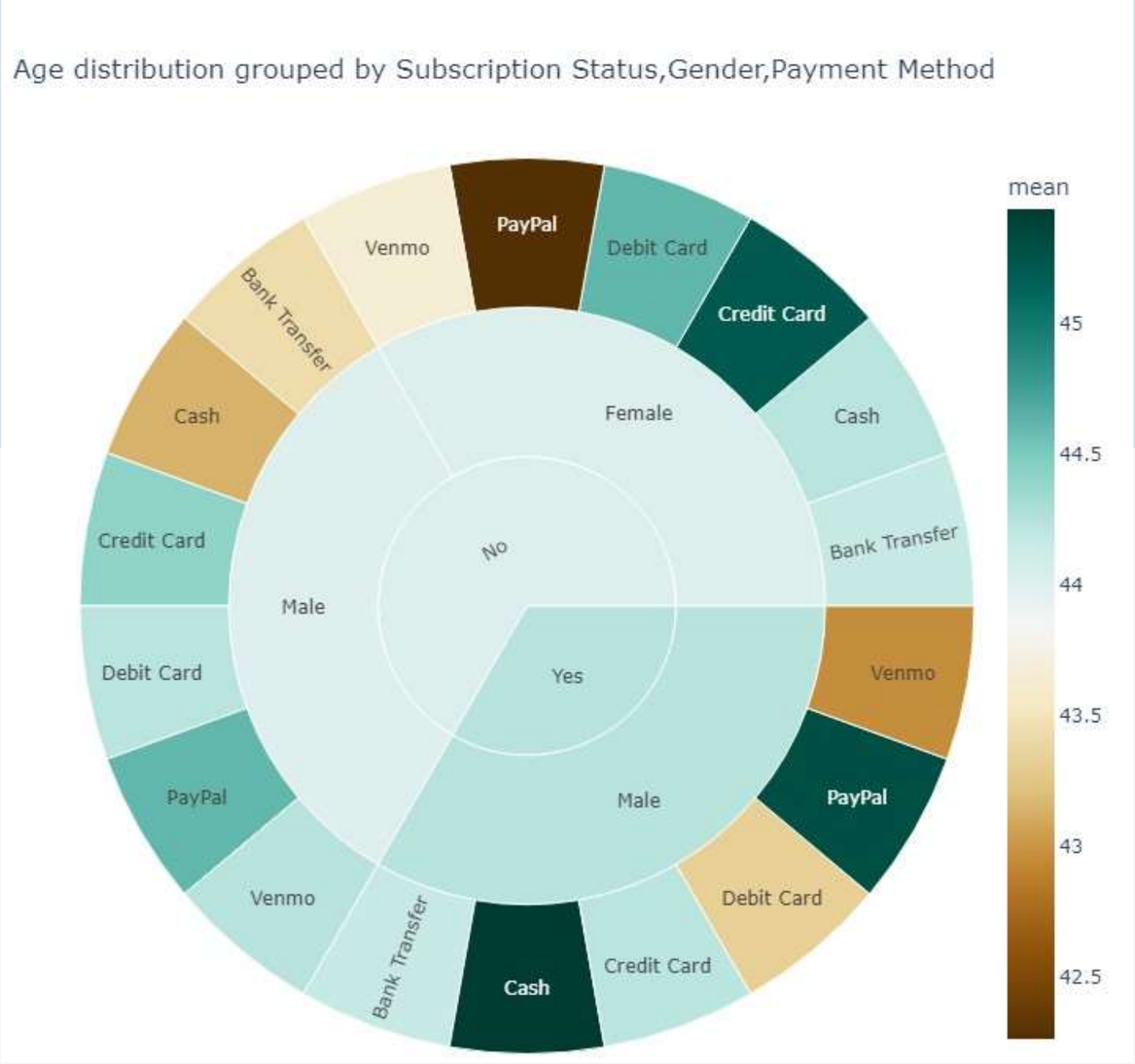
Purchase Amount(USD) across different groups





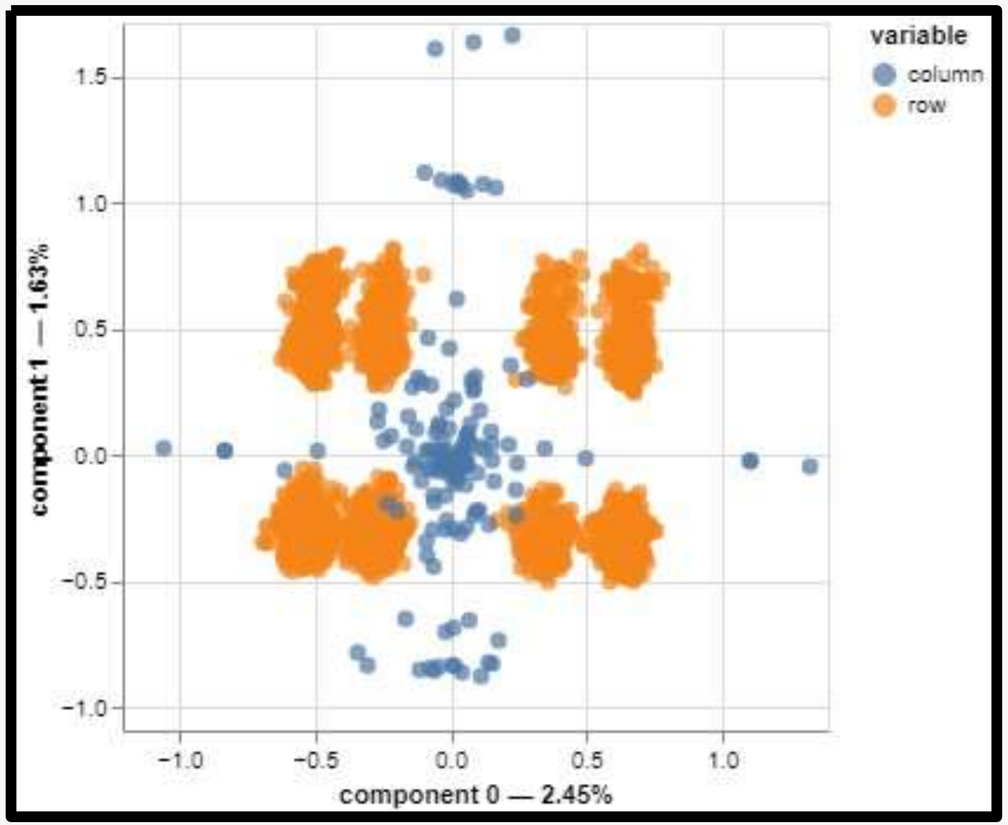
Age across different groups



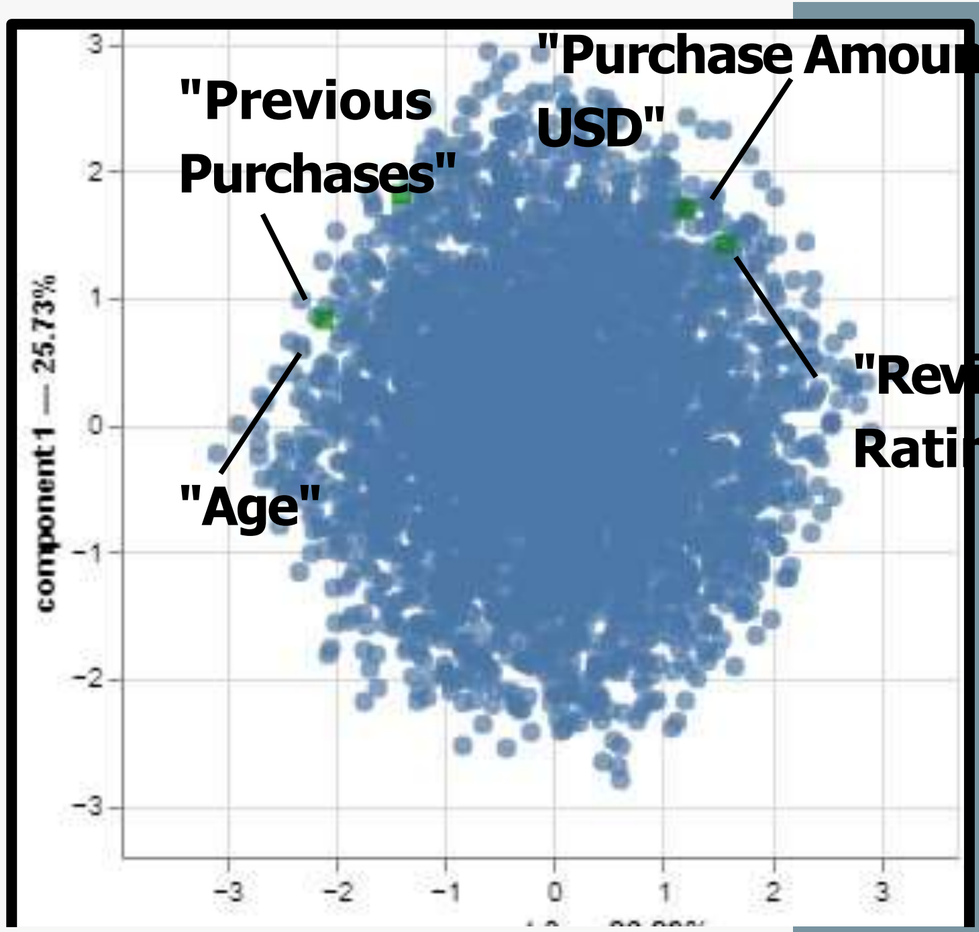


4. Additional Relations And Segmentations

PCA and MCA for patterns and clusters in data :

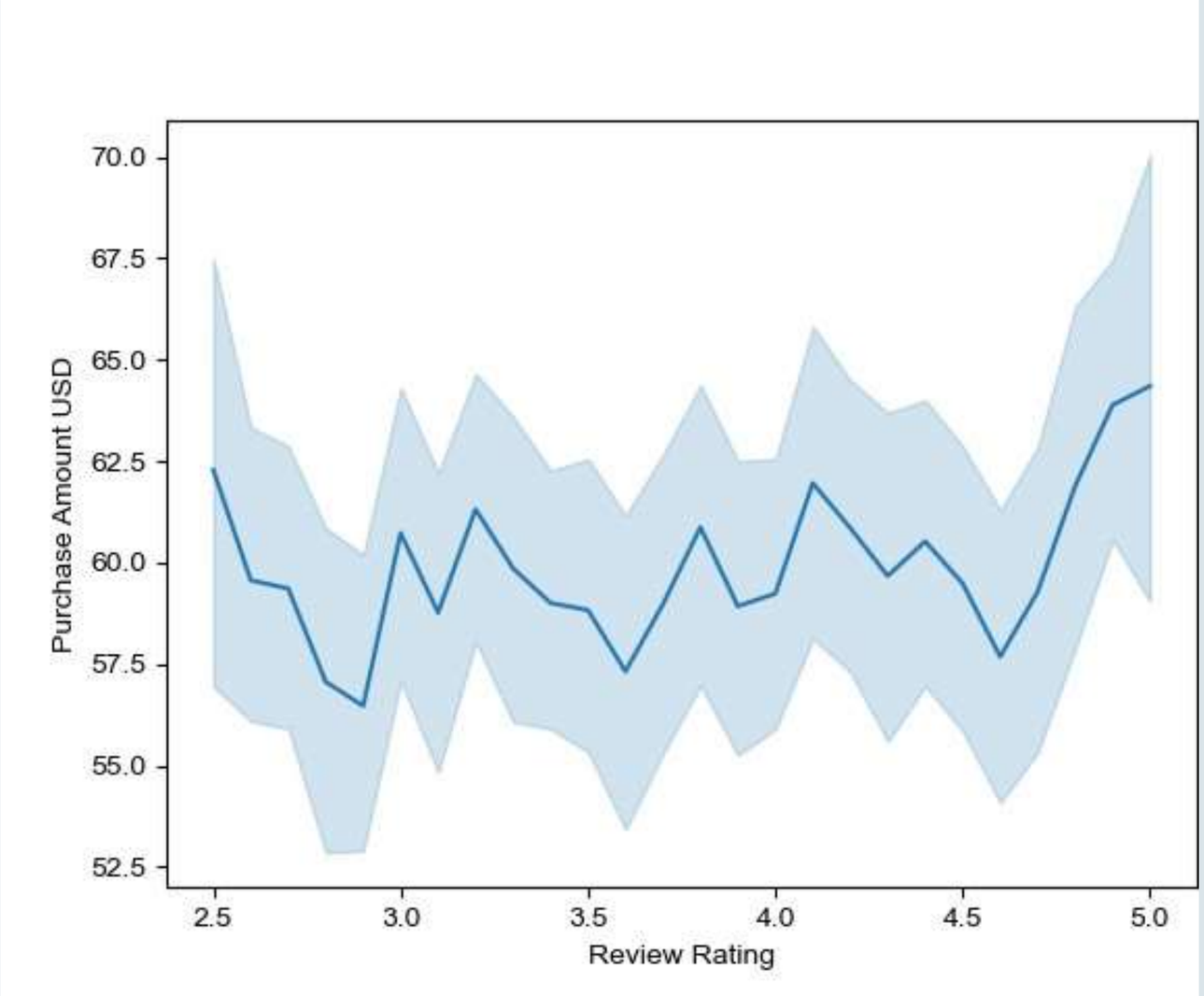


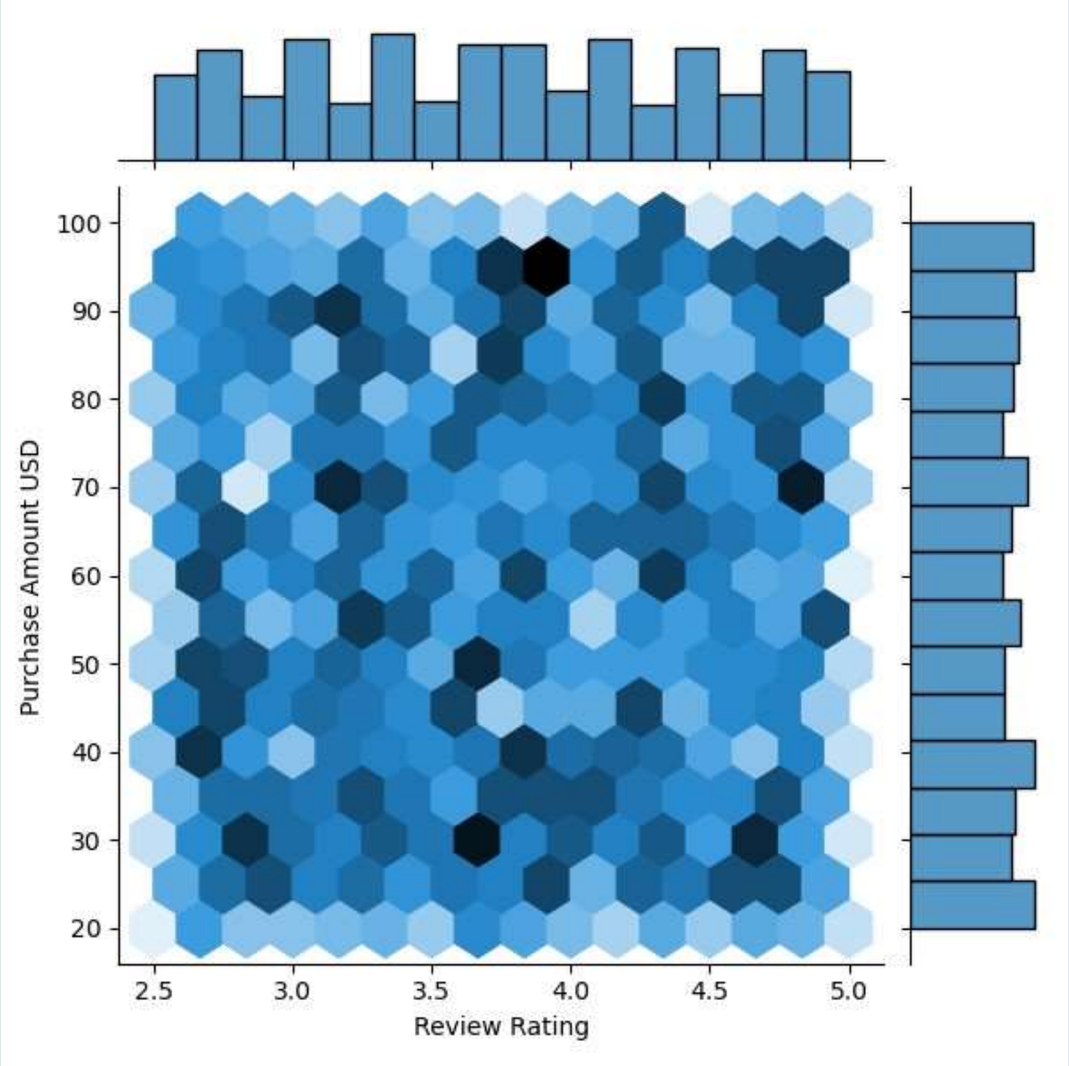
MCA



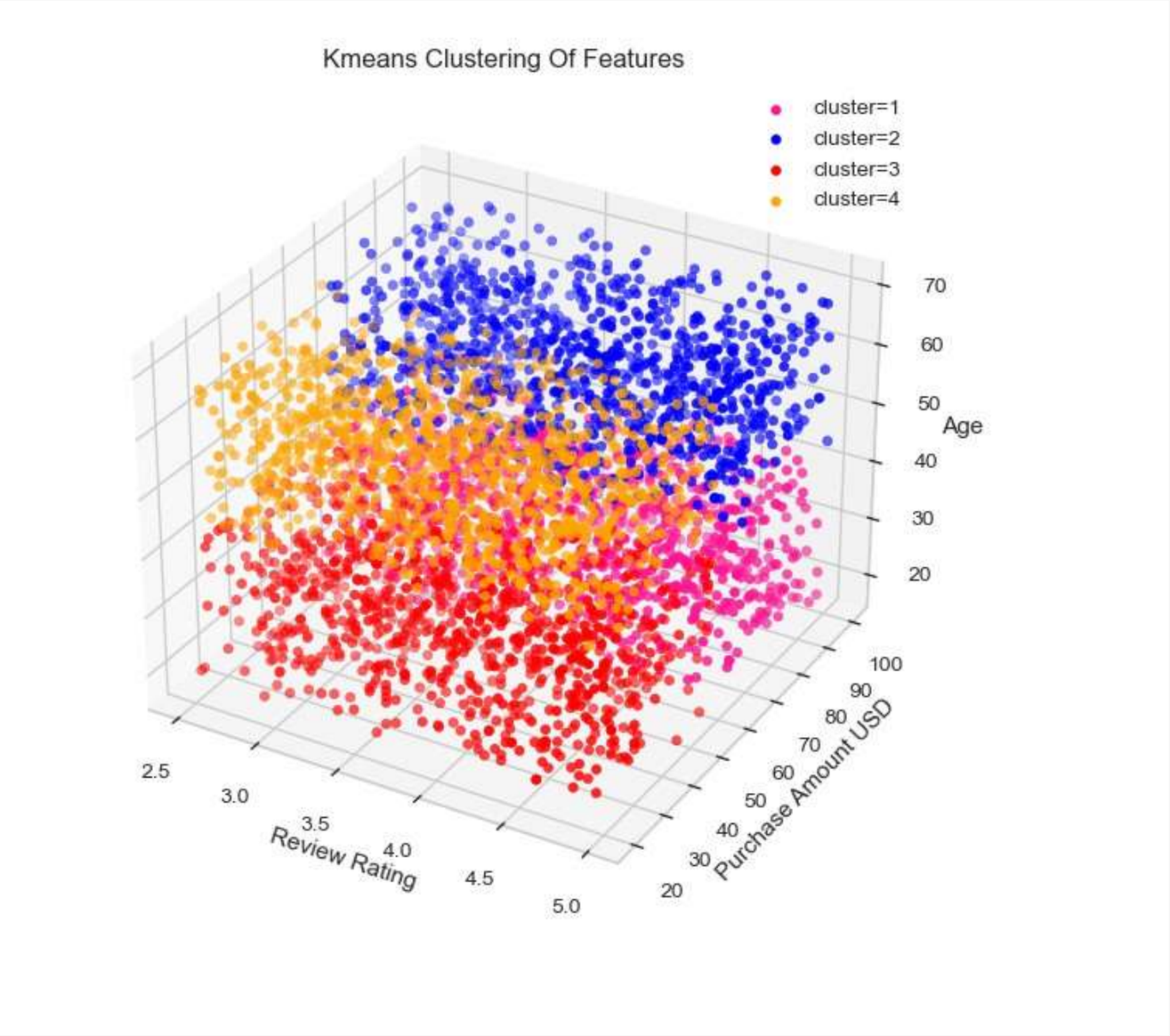
PCA

"Review Rating" and “Purchase Amount USD" relation





Customer Segmentation



One of possible segmentations

- **4 clusters with differences between “Age” (18-40, 40-70) and “Purchase Amount USD”(20-60, 60-100)**

Thank you

By Vanshika Rawat