



Shopping Trend Analysis

A Project Report submitted in partial fulfilment of the requirements of

AICTE Internship on AI: Transformative Learning

with

TechSaksham — A joint CSR initiative of Microsoft & SAP

By

Krish, mrkrishdhiman@gmail.com

Under The Guidance of

Aditya Prashant

Master Trainer, Edunet Foundation

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my mentors and faculty members at TechSaksham Team, whose invaluable guidance and support have been instrumental in the successful completion of this project. I am particularly grateful to Pavan Sumohana for their expertise and insights that helped me navigate through various challenges during the project. I also wish to thank my peers for their encouragement and constructive feedback, which motivated me to strive for excellence. Lastly, I acknowledge the contribution of various resources and tools that enabled the analysis and visualization of data, making this project possible.

ABSTRACT

The analysis of shopping trend datasets offers crucial insights into consumer behaviour, enabling businesses to make data-driven decisions to enhance marketing strategies, inventory management, and customer satisfaction. This study examines a dataset that encompasses various attributes, including product categories, purchase frequency, seasonal trends, and subscription statuses, with the aim of identifying key patterns and correlations. By leveraging the Python Pandas library, the dataset is cleaned, transformed, and analysed to uncover actionable insights. Seasonal trends are evaluated to understand how consumer preferences shift throughout the year, while the subscription status column is used to analyse its impact on purchase behaviour and loyalty. Statistical measures and visualisations are employed to highlight the most popular product categories, peak shopping periods, and variations in spending habits among different customer groups. Additionally, advanced techniques such as clustering and regression analysis are explored to predict future shopping trends and identify potential target segments. The results provide actionable recommendations for businesses, including personalised marketing campaigns, optimised inventory planning, and tailored subscription offers. This study demonstrates the power of data analysis in understanding and leveraging shopping trends, offering a foundation for businesses to stay competitive in dynamic markets while improving customer experiences.

CONTENT

Introduction.....	Page 1
Acknowledgement.....	Page 2
Abstract.....	Page 3
Content.....	Page 4
Figures.....	Page 5
Problem 1.....	Page 6
Problem 2.....	Page 7
Problem 3.....	Page 8
Problem 4.....	Page 9
Problem 5.....	Page 10
Problem 6.....	Page 11
Problem 7.....	Page 12
Problem 8.....	Page 13
Problem 9.....	Page 14
Problem 10.....	Page 15
Problem 11.....	Page 15
Problem 12.....	Page 16
Problem 13.....	Page 17
Problem 14.....	Page 18
Problem 15.....	Page 19
Problem 16.....	Page 20
Problem 17.....	Page 21
Problem 18.....	Page 22
Problem 19.....	Page 23

FIGURES

Figure No	Figure Caption	Page
Figure 1	Age Distribution	6
Figure 2	Average Purchase Amount By Product Category	7
Figure 3	Proportion of Purchases by Gender	8
Figure 4	Most Commonly Purchased Item in Each Category	9
Figure 5	Customer Spending By Season	10
Figure 6	Average Rating By Product Category	11
Figure 7	Total Spending: Subscribed vs Non-Subscribed Customers	12
Figure 8	Most popular Payment Method Across Customers	13
Figure 9	Total Spending: Promo Code Users vs Non-Users	14
Figure 10	Purchase Frequency Across Age Groups	15
Figure 11	Preferred Shipping Type for Different Product Categories	16
Figure 12	Average Purchase Amount with vs Without Discount	17
Figure 13	Product colour Popularity Distribution	18
Figure 14	Distribution Of Previous Purchases	19
Figure 15	Average Purchase Amount by Review Ratings	20
Figure 16	Product Category Purchases by Age Group	22
Figure 17	Average Purchase Amount By Gender	23

CHAPTER 1

SOLUTION TO PROBLEMS

Problem Statement 1 :- What is the overall distribution of customer ages in the dataset?

Solution

:-

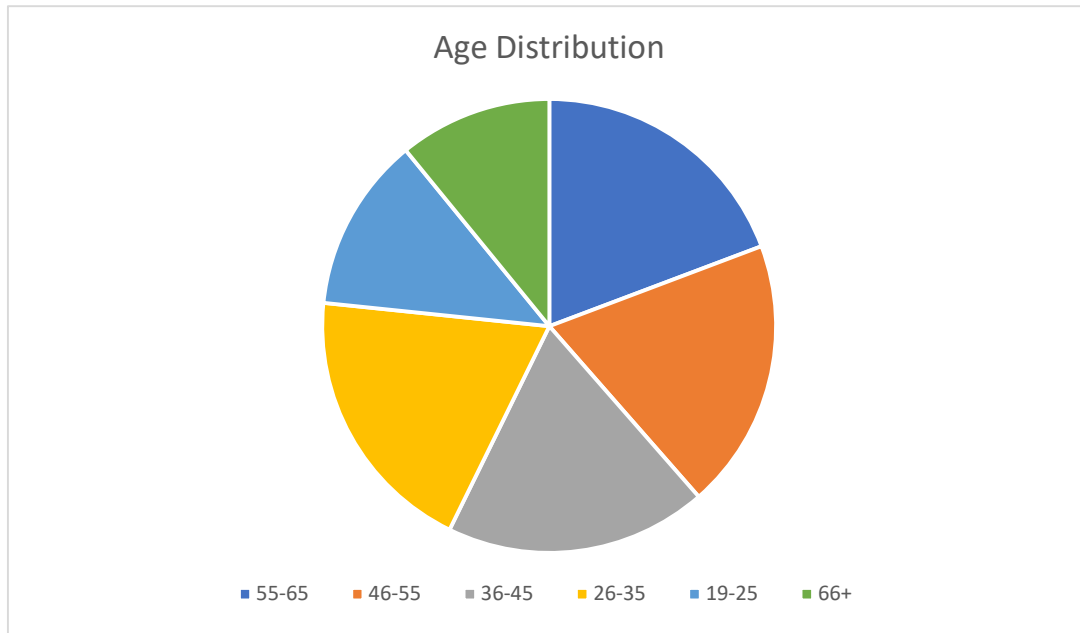


Figure 1 Age Distribution

Problem Statement 2 :- How does the average purchase amount vary across different product categories?

SOLUTION :-

Average Purchase Amount Distribution by Product Category

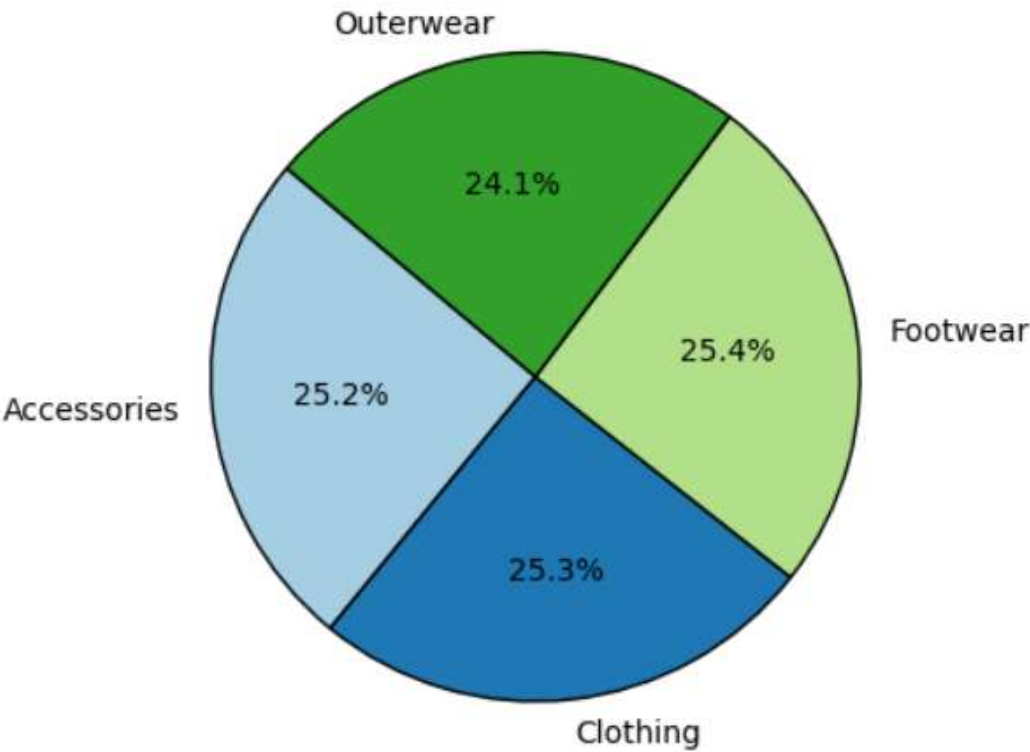


FIGURE 2 :- Graph of Average Purchase Amount Distribution by Product Category

Category	Average Purchase Amount
Accessories	59.838710
Clothing	60.025331
Footwear	60.255426
Outerwear	57.172840

Problem Statement 3 :- Which gender has the highest number of purchases?

Solution :-

Proportion of Purchases by Gender

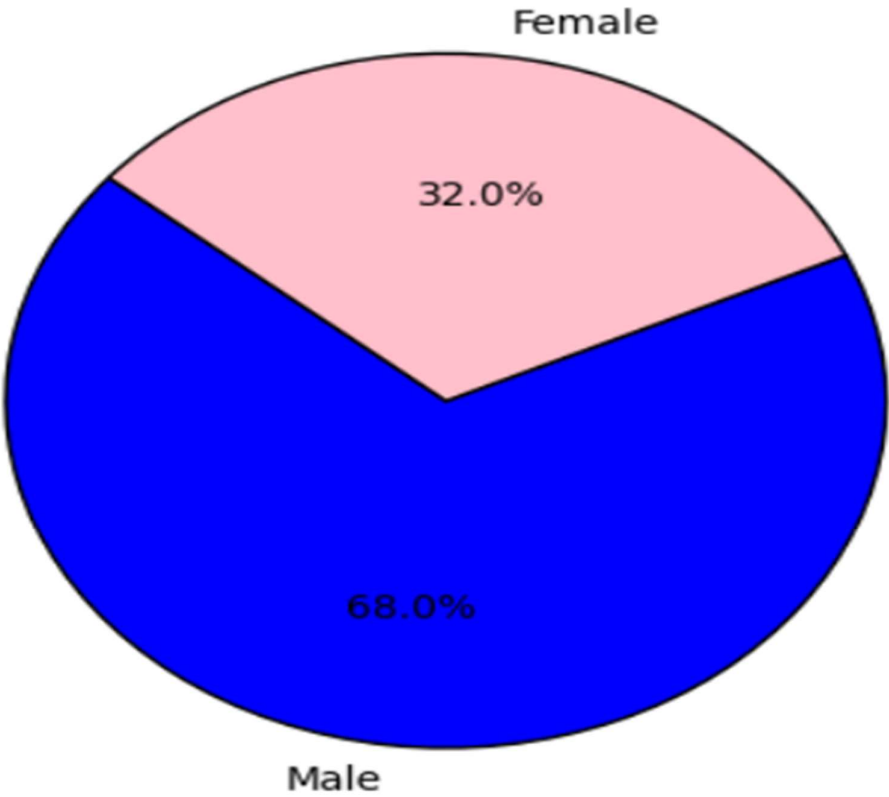


FIGURE 3 :- Proportion of Gender in Total Purchases

Gender	Proportion in Total Purchase
Male	2652
Female	1248

Problem Statement 4 :- What are the most commonly purchased items in each category?

Solution :-

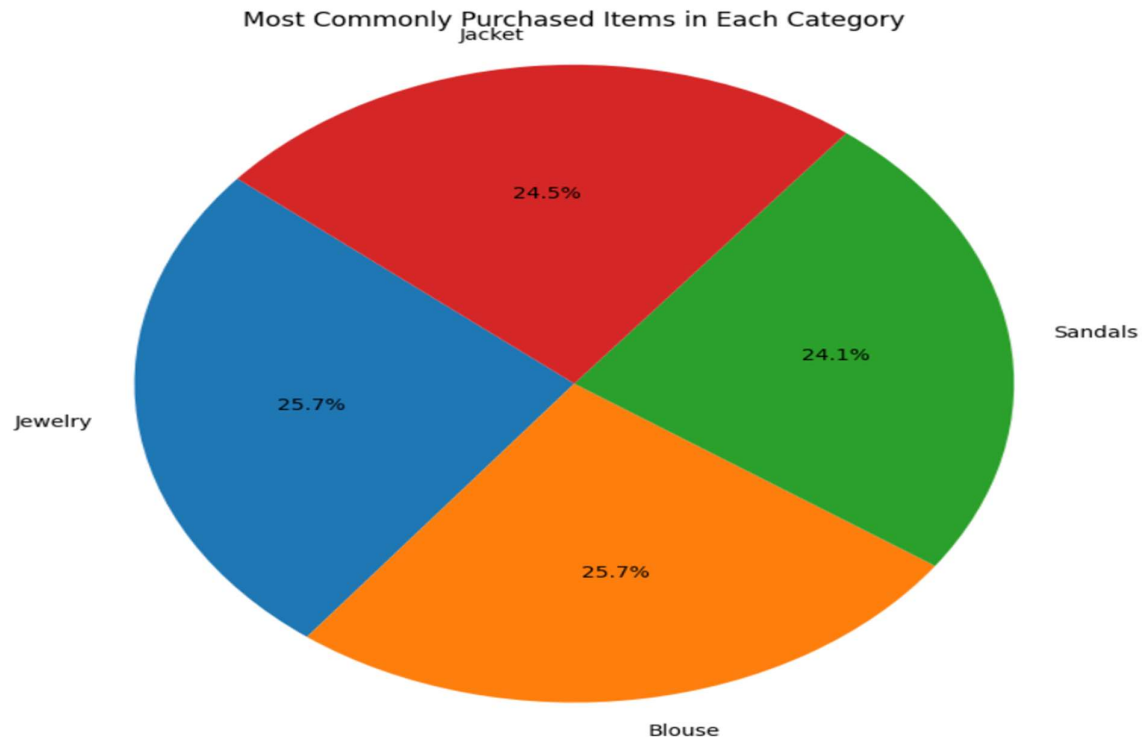


Figure 4 Most commonly Purchased Item in each category

S.no	Category	Item Name	No. Of Item
1	Accessories	Jewellery	171
2	Accessories	Belt	161
3	Accessories	Sun Glasses	161
4	Accessories	Scarf	157
5	Accessories	Hat	154
6	Accessories	Hand Bag	153
7	Accessories	Backpack	143
8	Accessories	Gloves	140
9	Clothing	Blouse	171
10	Clothing	Pant	171
11	Clothing	Shirt	169
12	Clothing	Dress	166
13	Clothing	Sweater	164
14	Clothing	Socks	159
15	Clothing	Skirt	158
16	Clothing	Shorts	157
17	Clothing	Hoodie	151
18	Clothing	T Shirt	147
19	Clothing	Jeans	124
20	Footwear	Sandals	160
21	Footwear	Shoes	150
22	Footwear	Sneakers	145
23	Footwear	Boots	144
24	Outerwear	Jacket	163
25	Outerwear	Coat	161

Problem Statement 5 :- Are there any specific seasons or months where customer spending is significantly higher?

Solution :-

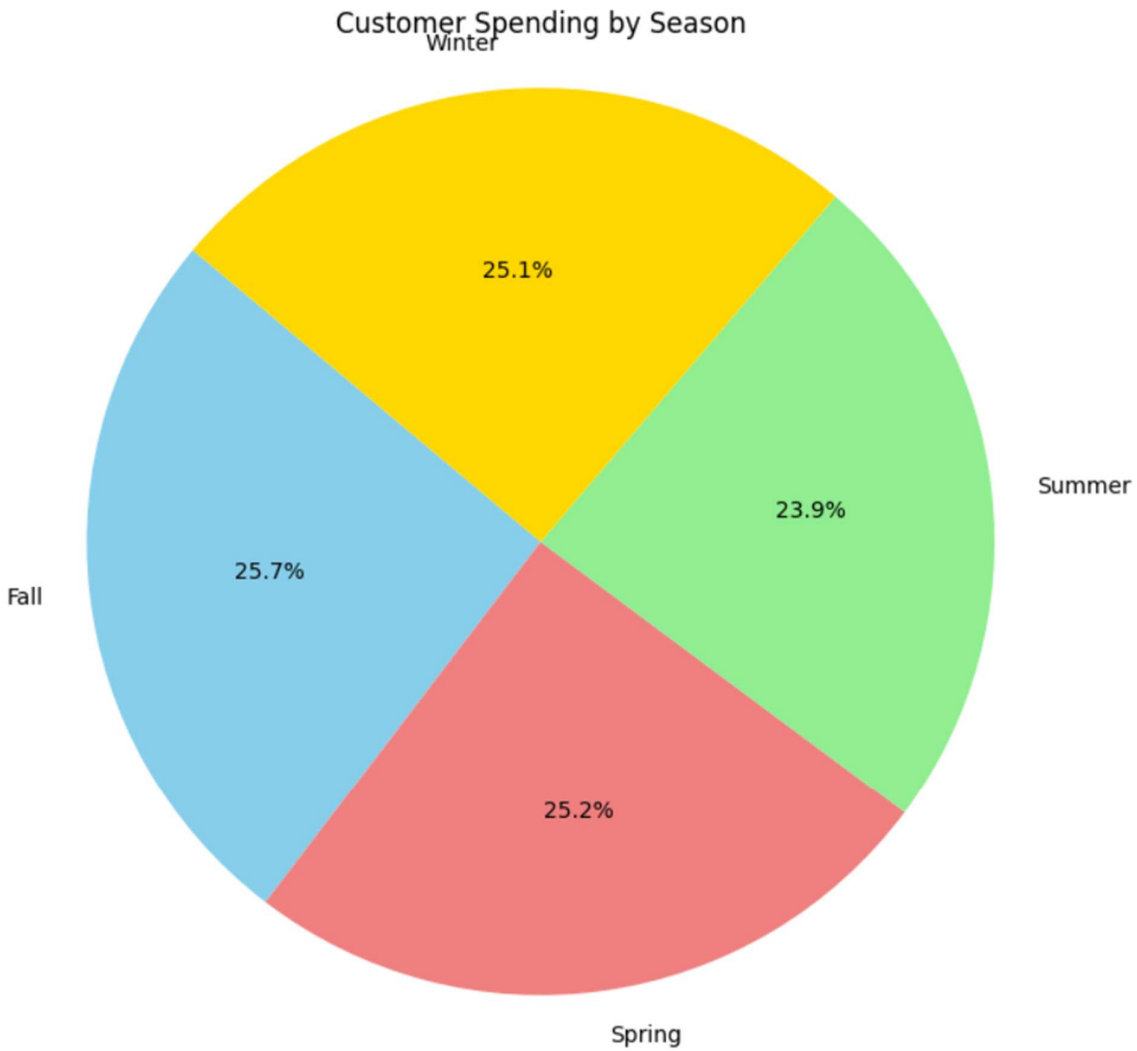


Figure 5 Customer Spending By Season

Fall	60018
Spring	58679
Summer	55777
Winter	58607

Problem Statement 6 :- What is the average rating given by customers for each product category?

Solution :-

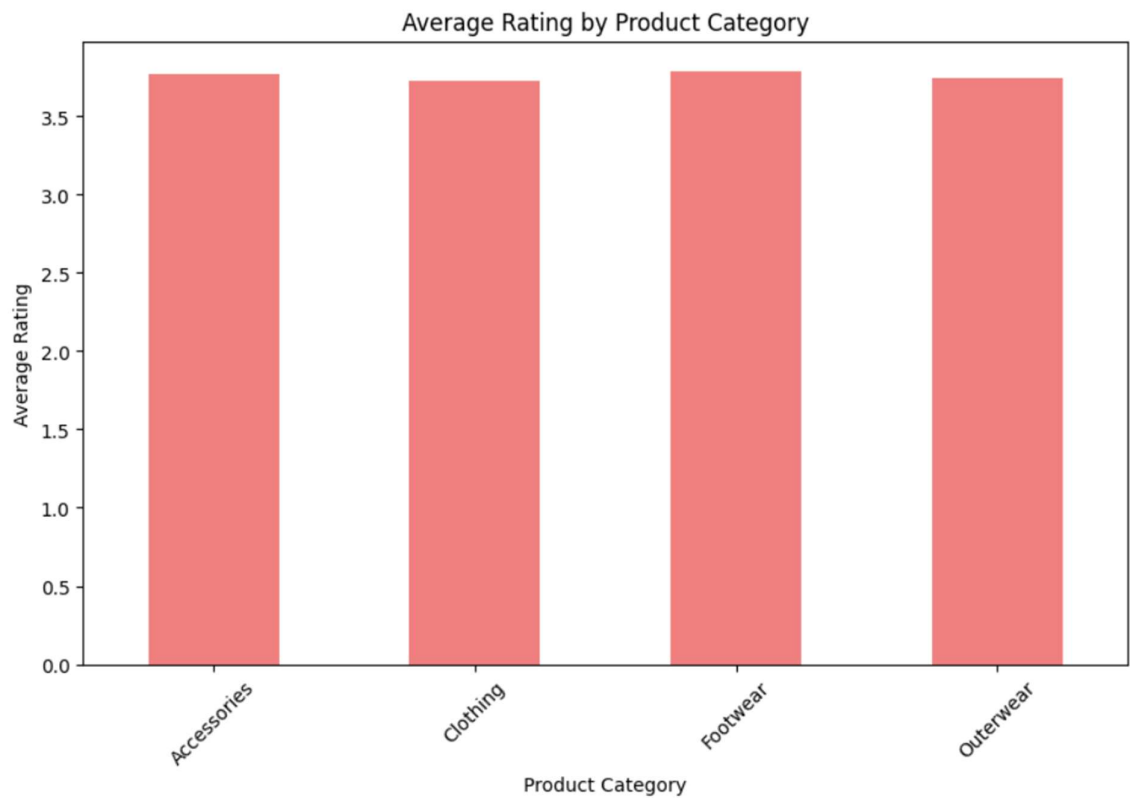


Figure 6 Average Rating by Product Category

Accessories	3.768629
Clothing	3.723143
Footwear	3.790651
Outerwear	3.746914

Problem Statement 7 :- Are there any notable differences in purchase behavior between subscribed and non-subscribed customers?

Solution :-

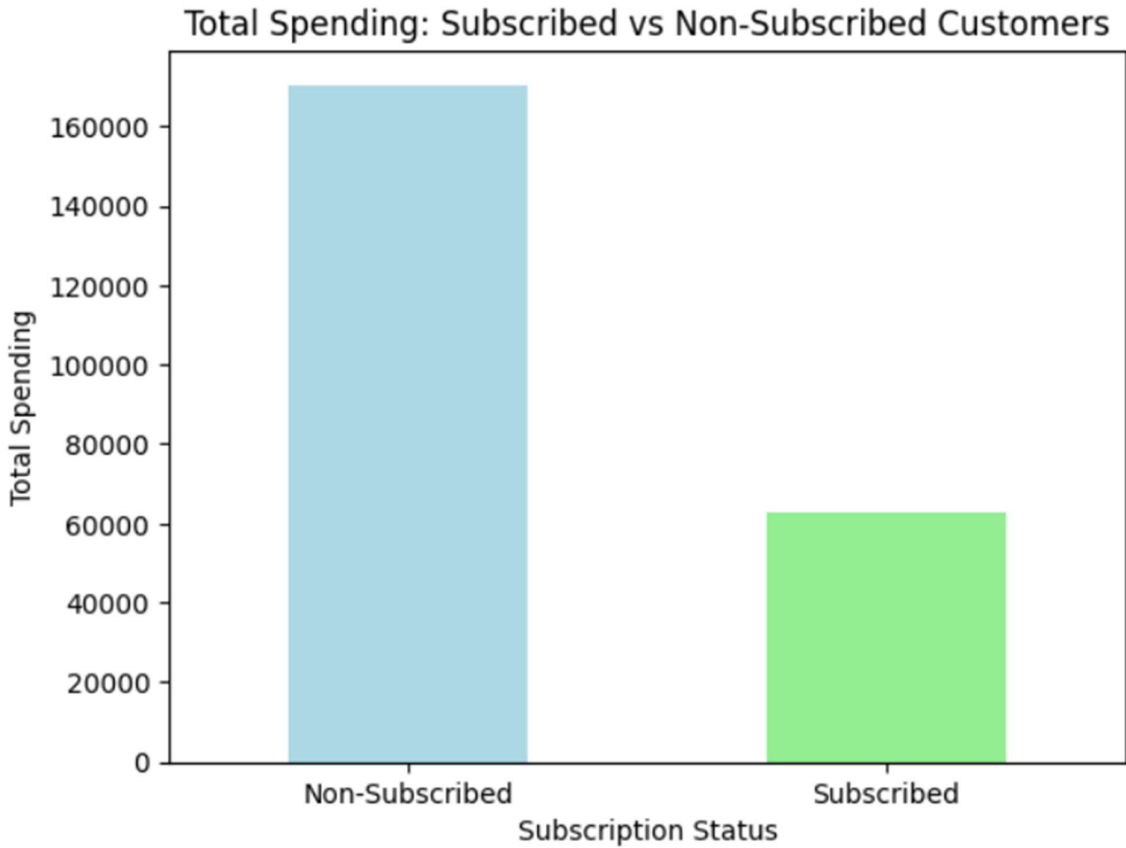


Figure 7 Total Spending Subscribed vs non Subscribed

Subscription Status	Number Of Customers
No	170436
Yes	62645

Problem Statement 8 :- Which payment method is the most popular among customers?

Solution :-

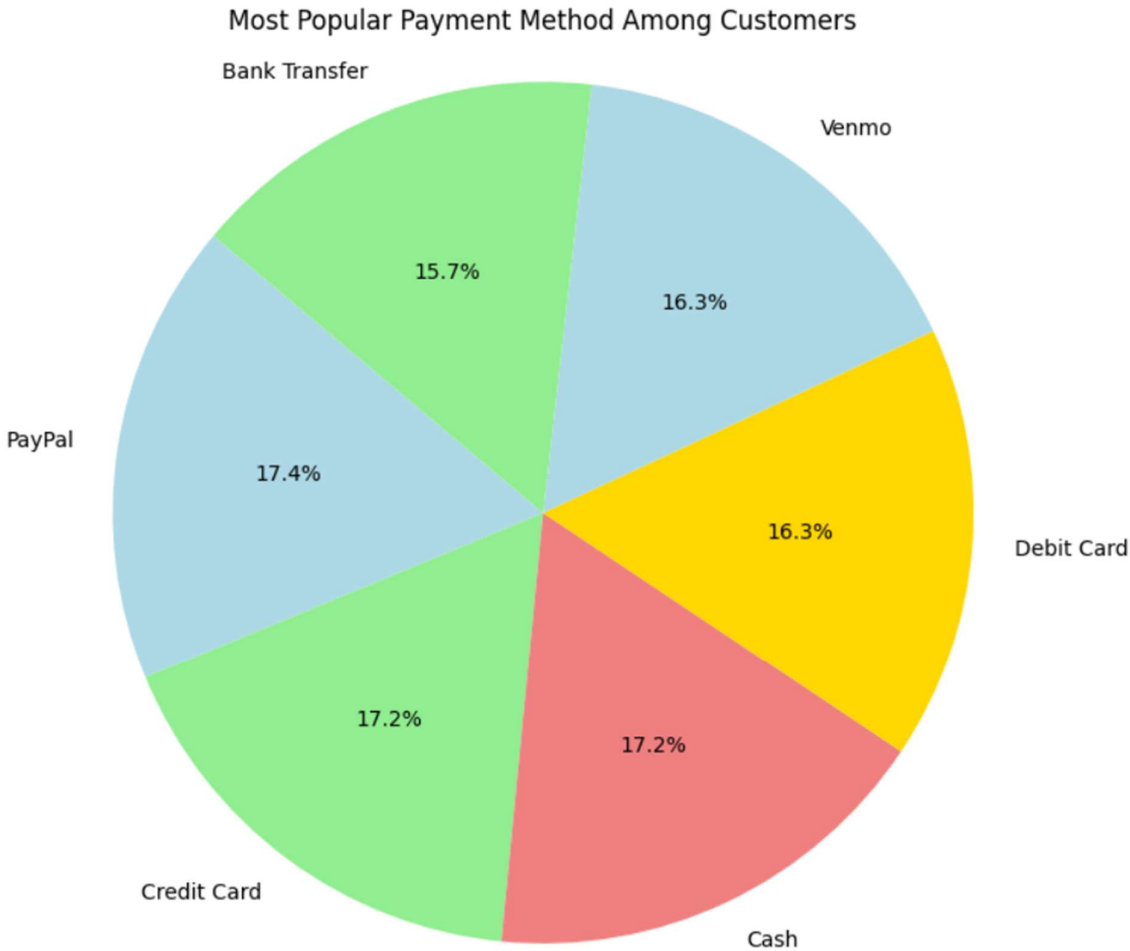


Figure 8 Most Popular Payment Method Among Customers

Payment Method	No Of Customer Used
Paypal	677
Credit Card	671
Cash	670
Debit Card	636
Venom	634
Bank Transfer	612

Problem Statement 9 :- Do customers who use promo codes tend to spend more than those who don't?

Solution :-

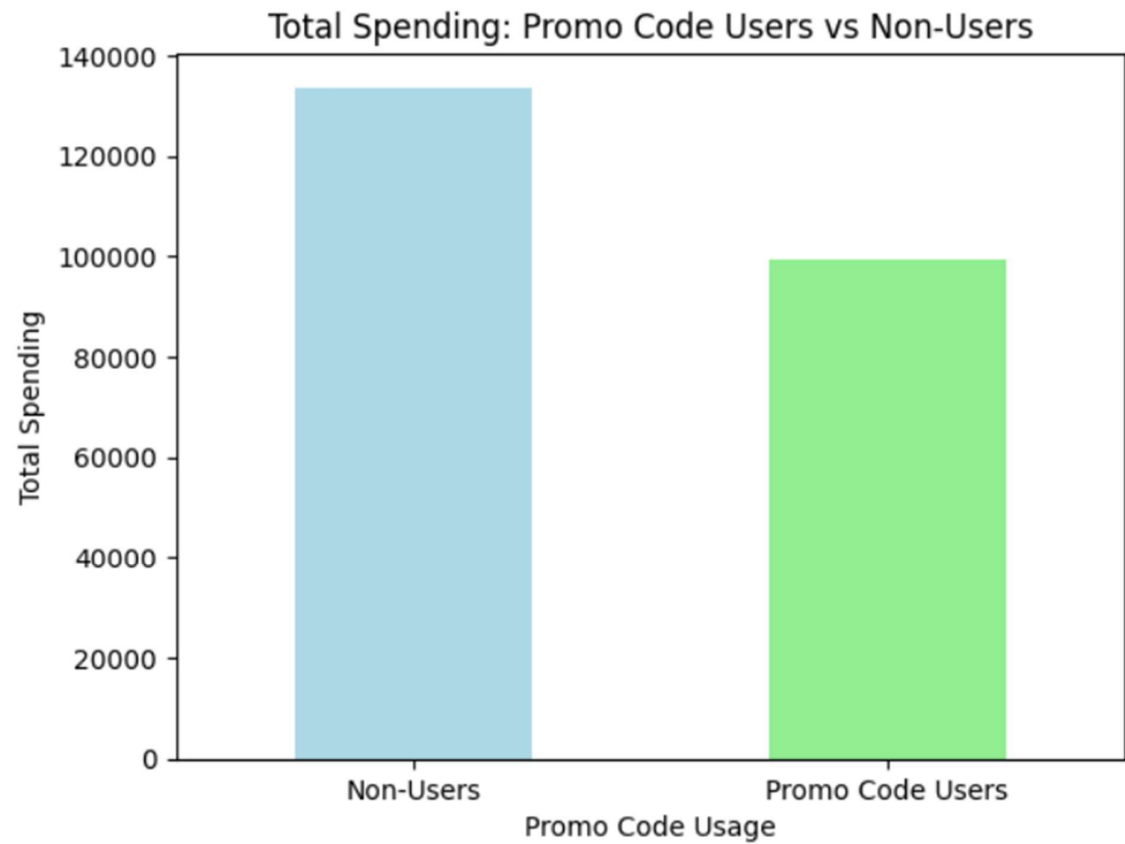


Figure 9 Total Spending Promo Code Users vs Non Users

Promo Code Used	Number of Customers
Yes	99411
No	133670

Problem Statement 10 :- How does the frequency of purchases vary across different age groups?

Solution :-

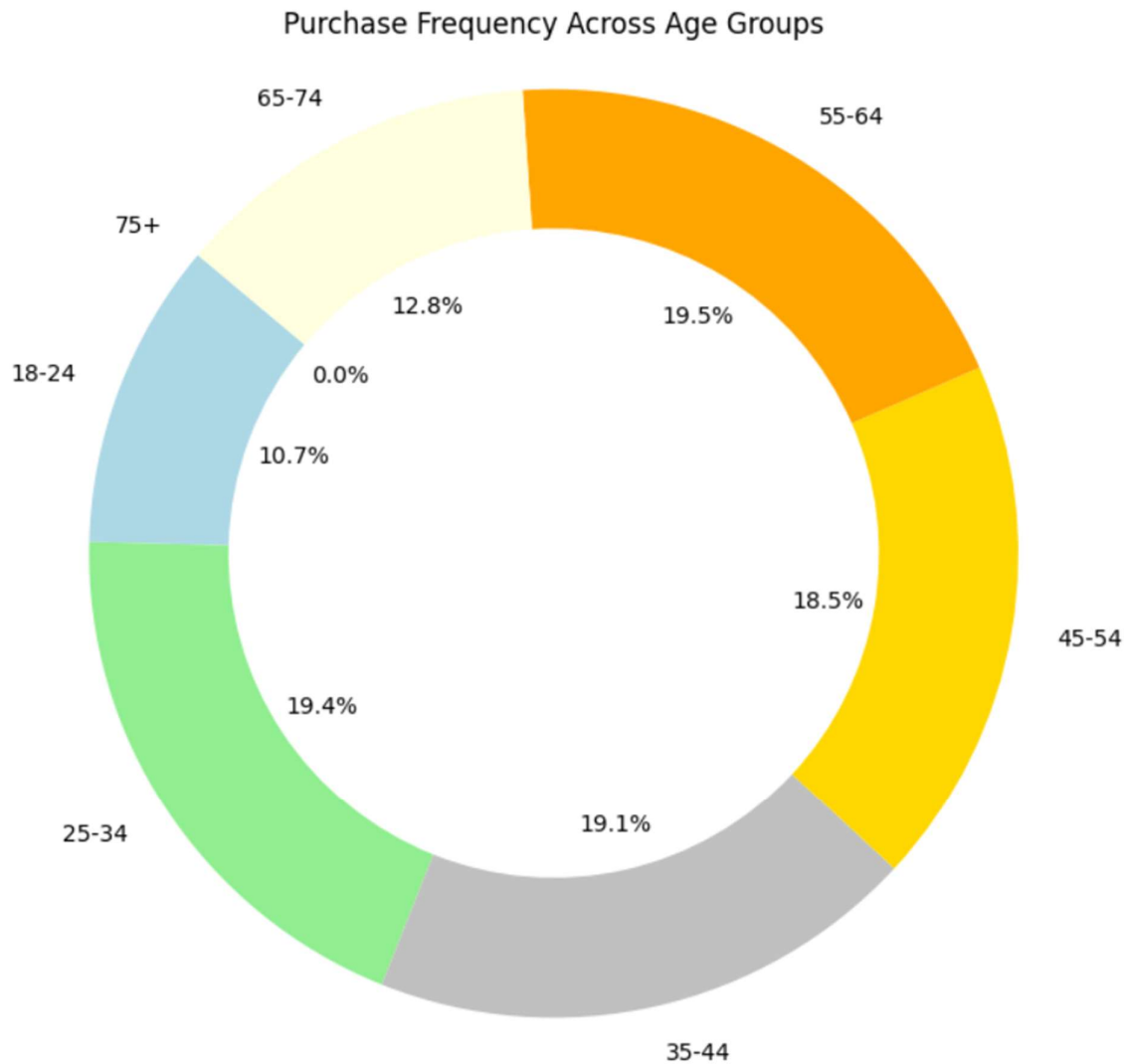


Figure 10 Purchase Frequency Across Age Groups

Age Group	Purchase
18-24	418
25-34	755
35-44	746
45-54	720
55-64	761
65-74	500
75+	0

Problem Statement 11 :- Are there any correlations between the size of the product and the purchase amount?

Solution :- Yes The Correlation between product size and purchase amount: -0.02245572497375391

Problem Statement 12 :- Which shipping type is preferred by customers for different product categories?

Solution :-

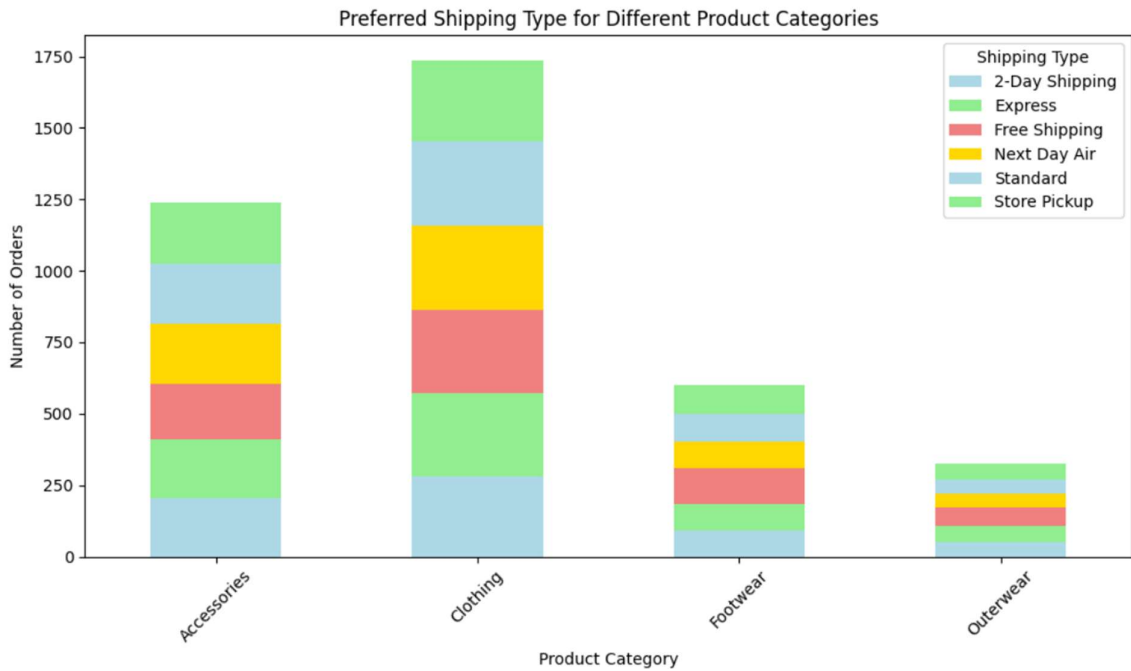


Figure 11 Preferred Shipping Type for Different Product Categories

Category	2-Days Shipping	Exprees	Free Shipping	Next Day	Air Standard	Store Pickup
Accessories	206	203	195	211	208	217
Clothing	281	290	294	293	297	282
Foot Wear	90	96	122	93	100	98
Outer Wear	50	57	64	51	49	53

Problem Statement 13 :- How does the presence of a discount affect the purchase decision of customers?

Solution :- Discount Applied :-



Figure 12 Average Purchase Amount with vs Without Discount

Yes	59.279070
No	60.130454

Problem 14 :- Are there any specific colors that are more popular among customers?

Solution :-

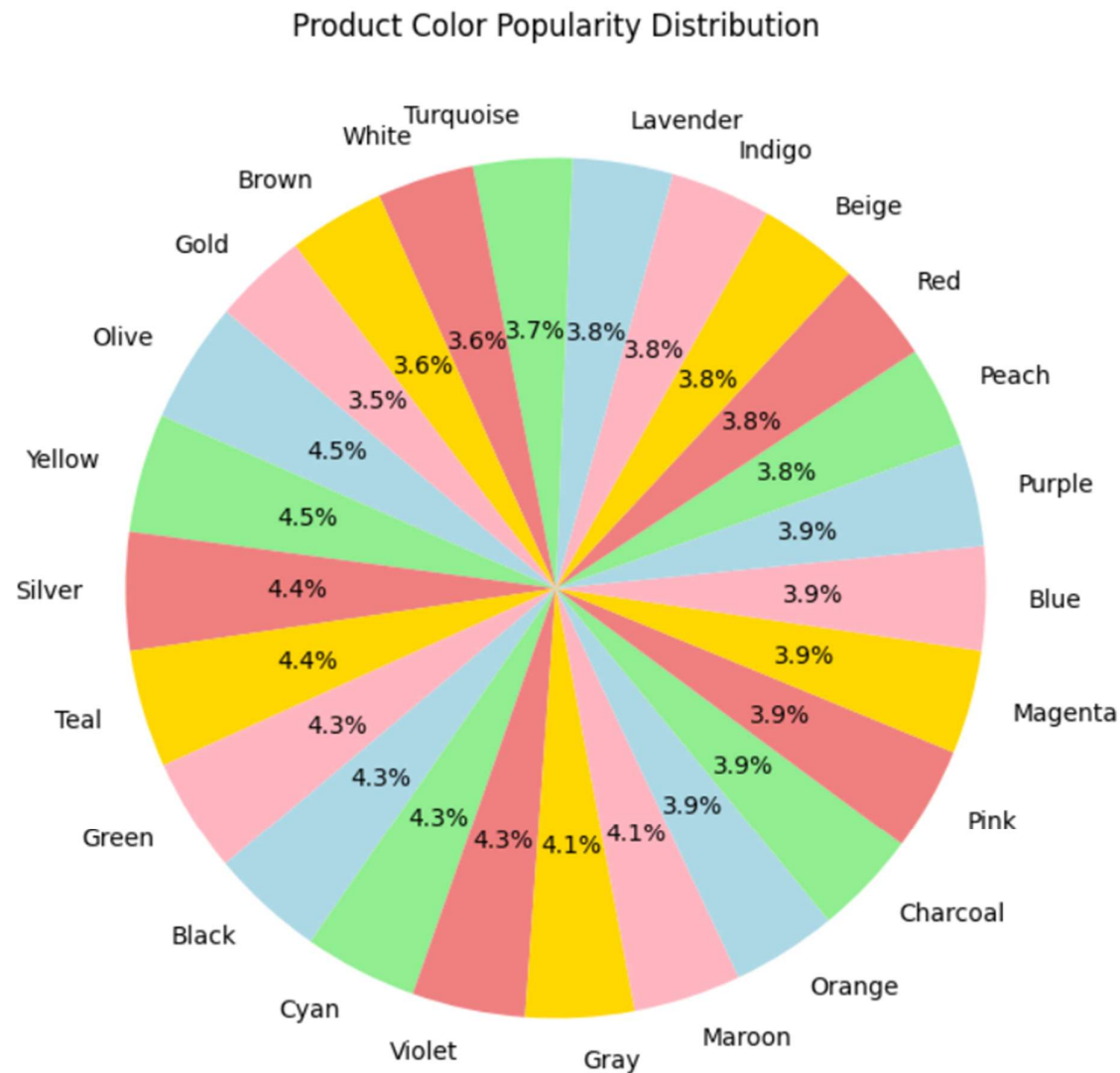


Figure 13 Product Colour Popularity Distribution

Colour	Purchases	Colour	Purchases	Colour	Purchases
Olive	177	Grey	159	Red	148
Yellow	174	Maroon	158	Beige	147
Silver	173	Orange	154	Indigo	147
Teal	172	Charcoal	153	Lavendar	147
Green	169	Pink	152	Turquoise	145
Black	167	Magenta	152	White	142
Cyan	166	Purple	151	Brown	141
Violent	166	Peach	149	Gold	138

Problem 15 :- What is the average number of previous purchases made by customers?

Solution :- Average number of previous purchases: 25.35153846153846

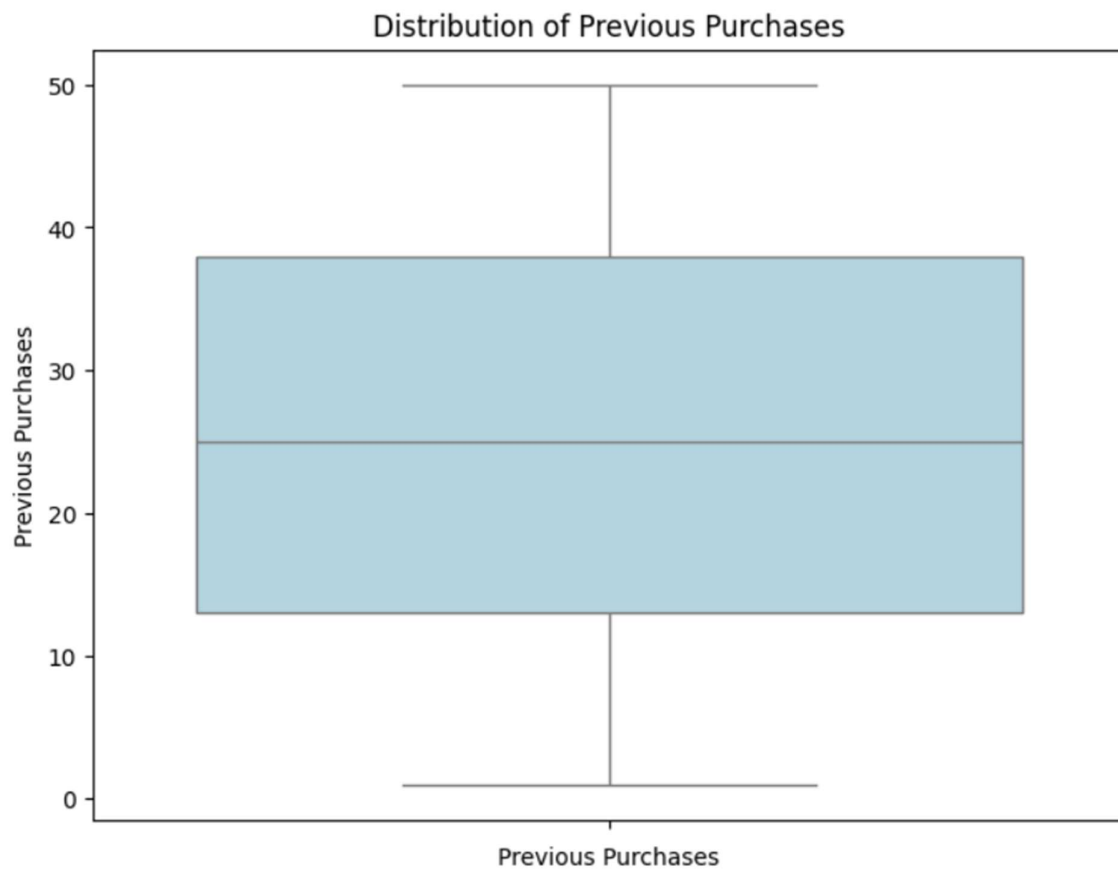


Figure 14 Distribution of Previous Purchases

Problem Statement 16 :- How does the purchase amount differ based on the review ratings given by customers?

Solution :-

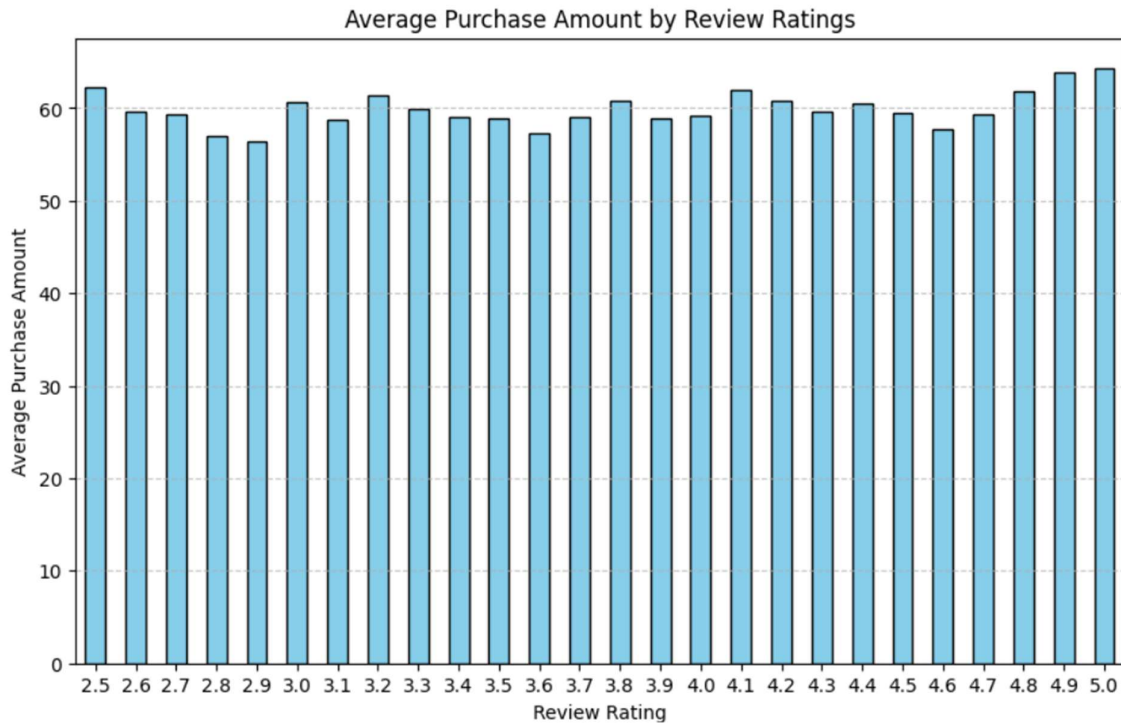


Figure 15 Average Purchase Amount by Review Ratings

2.5	62.287879
2.6	59.566038
2.7	59.363636
2.8	57.066176
2.9	56.470588
3.0	60.728395
3.1	58.770701
3.2	61.315789
3.3	59.861842
3.4	59.005495
3.5	58.833333
3.6	57.322148
3.7	58.974359
3.8	60.873239
3.9	58.926380
4.0	59.237569
4.1	61.959459
4.2	60.853801
4.3	59.673469
4.4	60.525316
4.5	59.489209
4.6	57.683908
4.7	59.283784
4.8	61.881944
4.9	63.885542
5.0	64.352941

Problem Statement 17 :- Are there any noticeable differences in purchase behavior between different locations?

Solution :- Yes, There are noticeable differences in purchase behavior as Stated Through this Table

S.NO	Location	Total Purchase Amount	Avg Purchase Amount	NO Of Purchases
1	Alabama	5261	59.11	89
2	Alaska	4867	67.59	72
3	Arizona	4326	66.55	65
4	Arkansas	4828	61.11	79
5	California	5605	59.00	95
6	Colorado	4222	56.29	75
7	Connecticut	4226	54.17	78
8	Delaware	4758	55.35	86
9	Florida	3798	55.85	68
10	Georgia	4645	58.79	79
11	Hawaii	3752	57.72	65
12	Idaho	5587	60.07	93
13	Illinois	5617	61.05	92
14	Indiana	4655	58.92	79
15	Iowa	3752	60.88	69
16	Kansas	5587	54.55	63
17	Kentucky	5617	55.72	79
18	Louisiana	4655	57.71	84
19	Maine	4201	56.98	77
20	Maryland	3437	55.75	86
21	Massachusetts	4402	60.88	72
22	Michigan	4848	62.09	73
23	Minnesota	4388	56.55	88
24	Mississippi	4795	61.03	80
25	Missouri	4384	57.91	81
26	Montana	4533	60.25	96
27	Nebraska	4977	59.44	87
28	Nevada	4883	63.37	87
29	New Hampshire	4691	59.42	71
30	New Jersey	5784	56.74	67
31	New Mexico	5172	61.90	81
32	New York	5514	60.42	87
33	North Carolina	4219	60.79	78
34	North Dakota	3802	62.89	83
35	Ohio	5014	60.37	77
36	Oklahoma	5257	58.34	75
37	Oregon	4742	57.33	74
38	Pennsylvania	5220	66.56	74
39	Rhode Island	4649	61.44	63
40	South Carolina	4376	58.40	76
41	South Dakota	4243	60.51	70
42	Tennessee	4926	61.97	77
43	Texas	3871	61.19	77
44	Utah	4443	62.57	71
45	Vermont	4860	57.17	85
46	Virginia	4842	62.88	77
47	Washington	4623	63.32	73
48	West Virginia	5174	63.87	81
49	Wisconsin	4196	55.94	75
50	Wyoming	4309	60.69	71

Problem Statement 18 :- Is there a relationship between customer age and the category of products they purchase?

Solution :- The Below Figure Shows The Relationship.

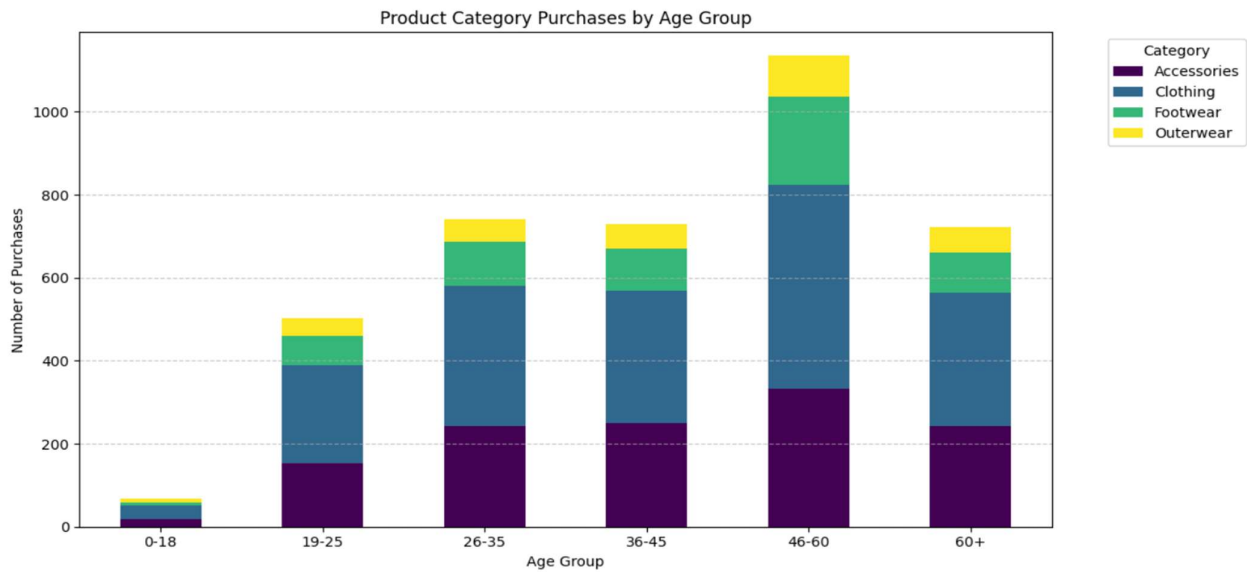


Figure 16 Product Category Purchases by Age Group

Age Group	Accessories	Clothing	Footwear	Outerwear
0-18	18	33	8	10
19-25	154	236	71	41
26-35	243	337	108	54
36-45	249	319	102	59
46-60	333	492	212	98
60+	243	320	98	62

Problem Statement 19 :- How does the average purchase amount differ between male and female customers?

Solution :- The average Purchase Amount differ are Shown Below By the Use of graph

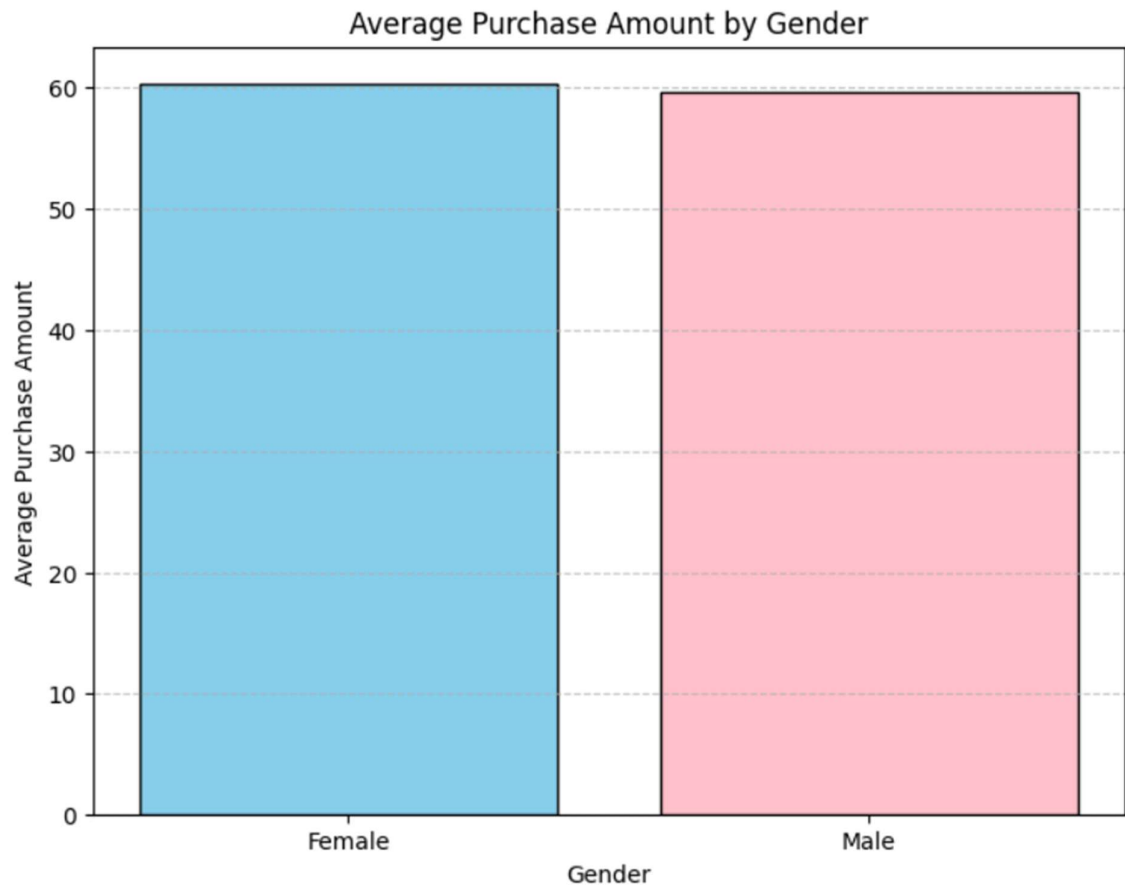


Figure 17 Average Purchase Amount by Gender

Gender	Percentage
Male	59.536199
Female	60.249199