Revealing Insights into Customer Behavior and Segmentation

*Final Project of Business Intelligence Bootcamp Batch 9
by* **Nabila Ijtihatil Ismail**

Table of Contents

PROJECT BACKGROUND

The project aims to gain a comprehensive understanding of customer behavior, identify sales trends, and develop customer segmentation strategies. By implementing this project, we aim to optimize marketing strategies, enhance customer satisfaction, and boost overall sales performance.

BUSINESS PROBLEM

Currently, competition in the retail and fashion business is getting tighter, as proven by the increasing number of similar businesses. One way to face this growing competition is to recognize customer behavior. By understanding customer behavior, we can create marketing strategies that are right on target so that can increasing overall sales

TOOLS

DATASET

Consumer Behavior and Shopping Habits Dataset:

DATA OVERVIEW

	Customer ID	Age	Gender	Item Purchased	Product Category	Purchase Amount (USD)	Location	Review Rating	Subscription Status	Shipping Type	Promo Code Used	Previous Purchases	Payment Method	Frequency of Purchases	Age Group	Shipping Mode
0	1	55	Female	Blouse	Clothing	53	Kentucky	3.1	Yes	Express	Yes	14	Venmo	Fortnightly	51-60	Online
1	2	19	Female	Sweater	Clothing	64	Maine	3.1	Yes	Express	Yes	2	Cash	Fortnightly	10-20	Online
2	3	50	Female	Jeans	Clothing	73	Massachusetts	3.1	Yes	Free Shipping	Yes	23	Credit Card	Weekly	41-50	Online
3	4	21	Female	Sandals	Footwear	90	Rhode Island	3.5	Yes	Next Day Air	Yes	49	PayPal	Weekly	21-30	Online
4	5	45	Female	Blouse	Clothing	49	Oregon	2.7	Yes	Free Shipping	Yes	31	PayPal	Annually	41-50	Online
5	6	46	Female	Sneakers	Footwear	20	Wyoming	2.9	Yes	Standard	Yes	14	Venmo	Weekly	41-50	Online
6	7	63	Female	Shirt	Clothing	85	Montana	3.2	Yes	Free Shipping	Yes	49	Cash	Quarterly	61-70	Online
7	8	27	Female	Shorts	Clothing	34	Louisiana	3.2	Yes	Free Shipping	Yes	19	Credit Card	Weekly	21-30	Online
8	9	26	Female	Coat	Outerwear	97	West Virginia	2.6	Yes	Express	Yes	8	Venmo	Annually	21-30	Online
9	10	57	Female	Handbag	Accessories	31	Missouri	4.8	Yes	2-Day Shipping	Yes	4	Cash	Quarterly	51-60	Online

- The dataset provides a detailed overview of consumer preferences and purchasing behaviors
- Contains of 3900 rows and 16 columns
- This dataset has no missing value and no duplicated data

see the dataset **here**!

DATA PREPROCESSING

1. **Handling Abnormal Data**

There were indications of input errors in customer gender because many male customers bought women's clothing, such as dresses and skirts, with the number of purchases being twice as much as that of female customers. Therefore, it was necessary to change the gender from male to female and vice versa.

2. **Create new columns such as Age Group and Shipping Mode.**

- For the age group column, divide customers into 6 groups
- For the shipping mode column, group shipping types into online and offline. In the offline category, only store pickup is included, while other shipping types are classified as online.

DATA ANALYSIS & INSIGHT

Customer Distribution by Gender

Male 32.0%

68.0% Female
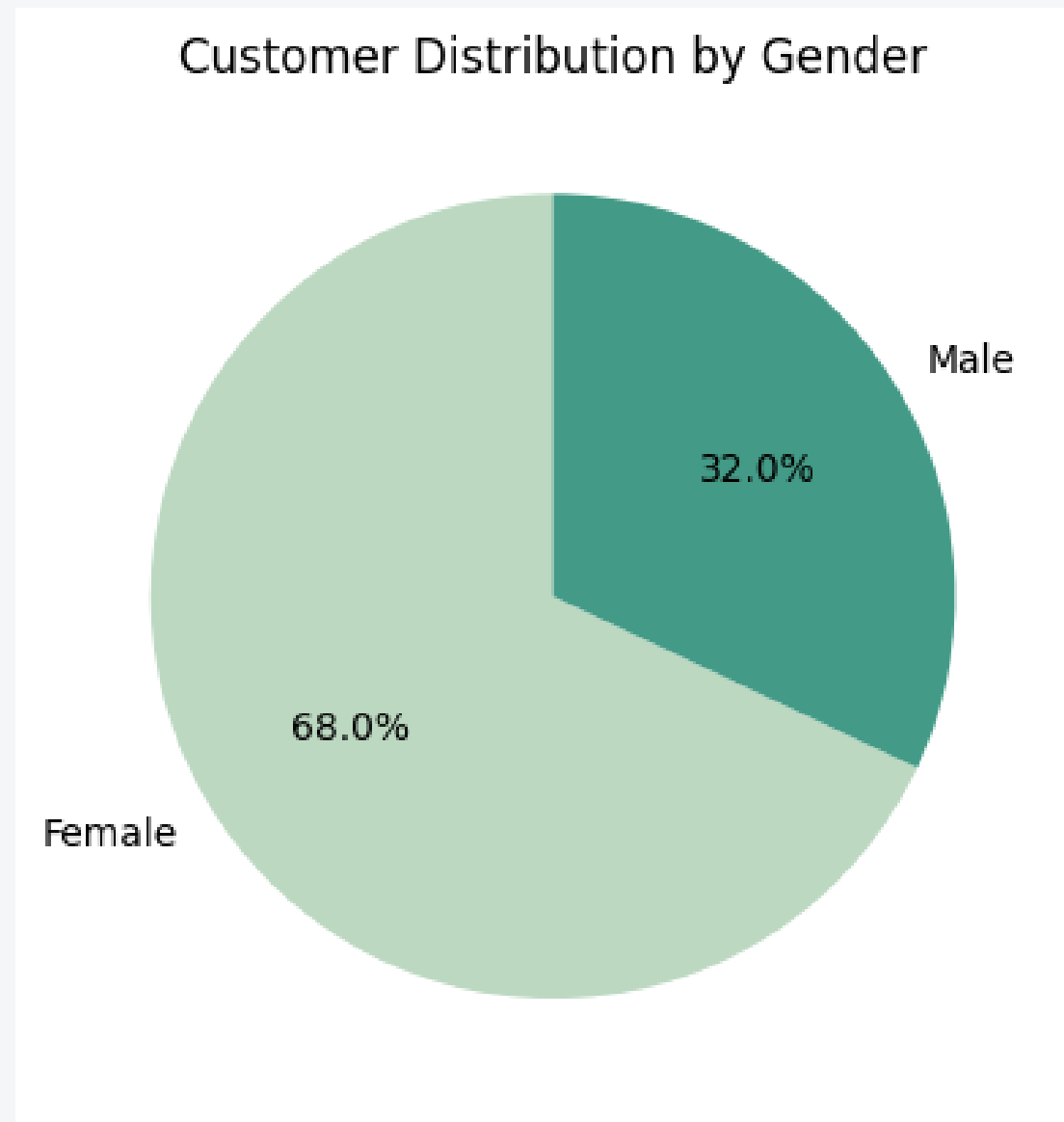
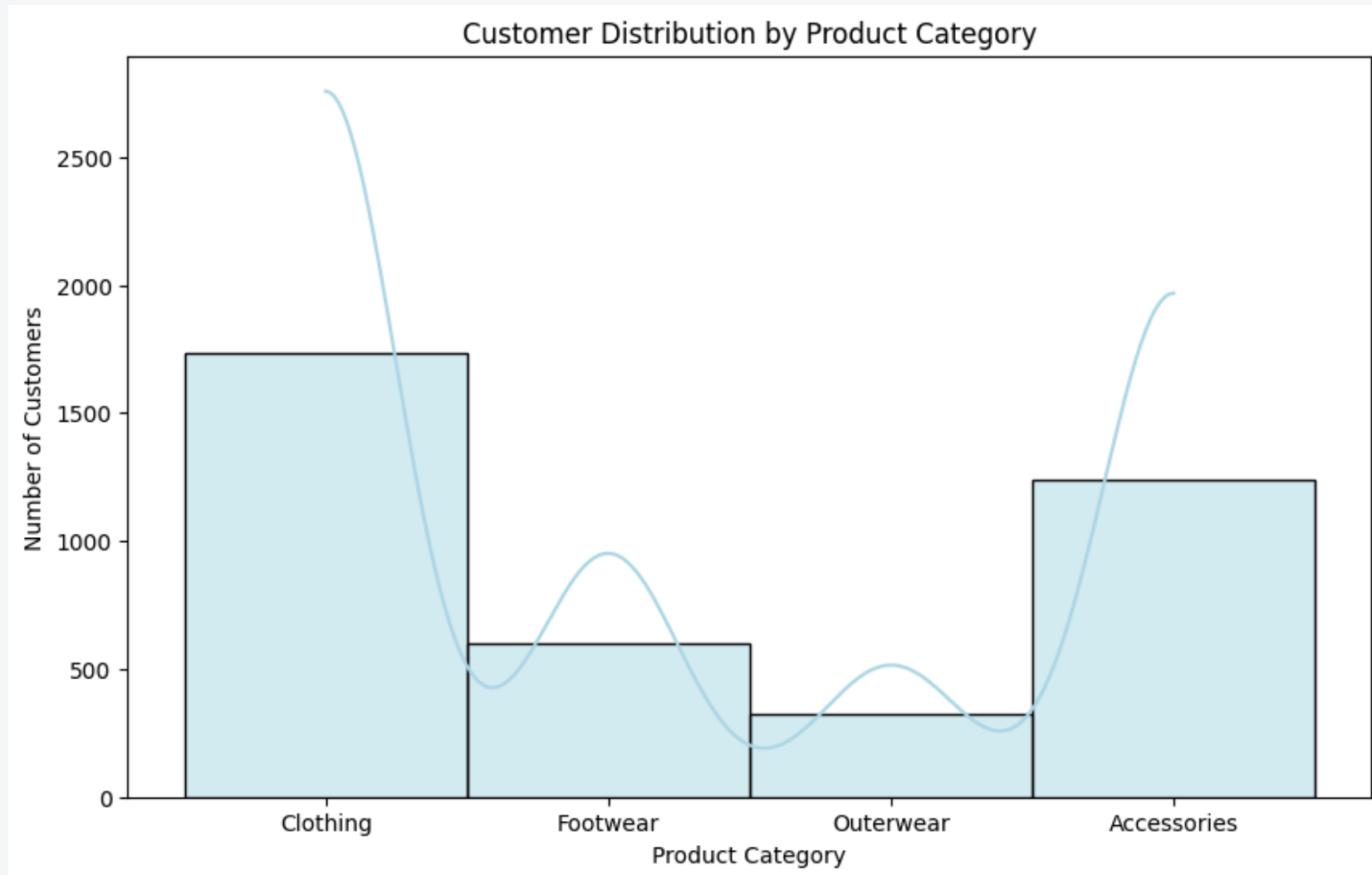**Dominated by female customers** as much as 68% or 2,652 people
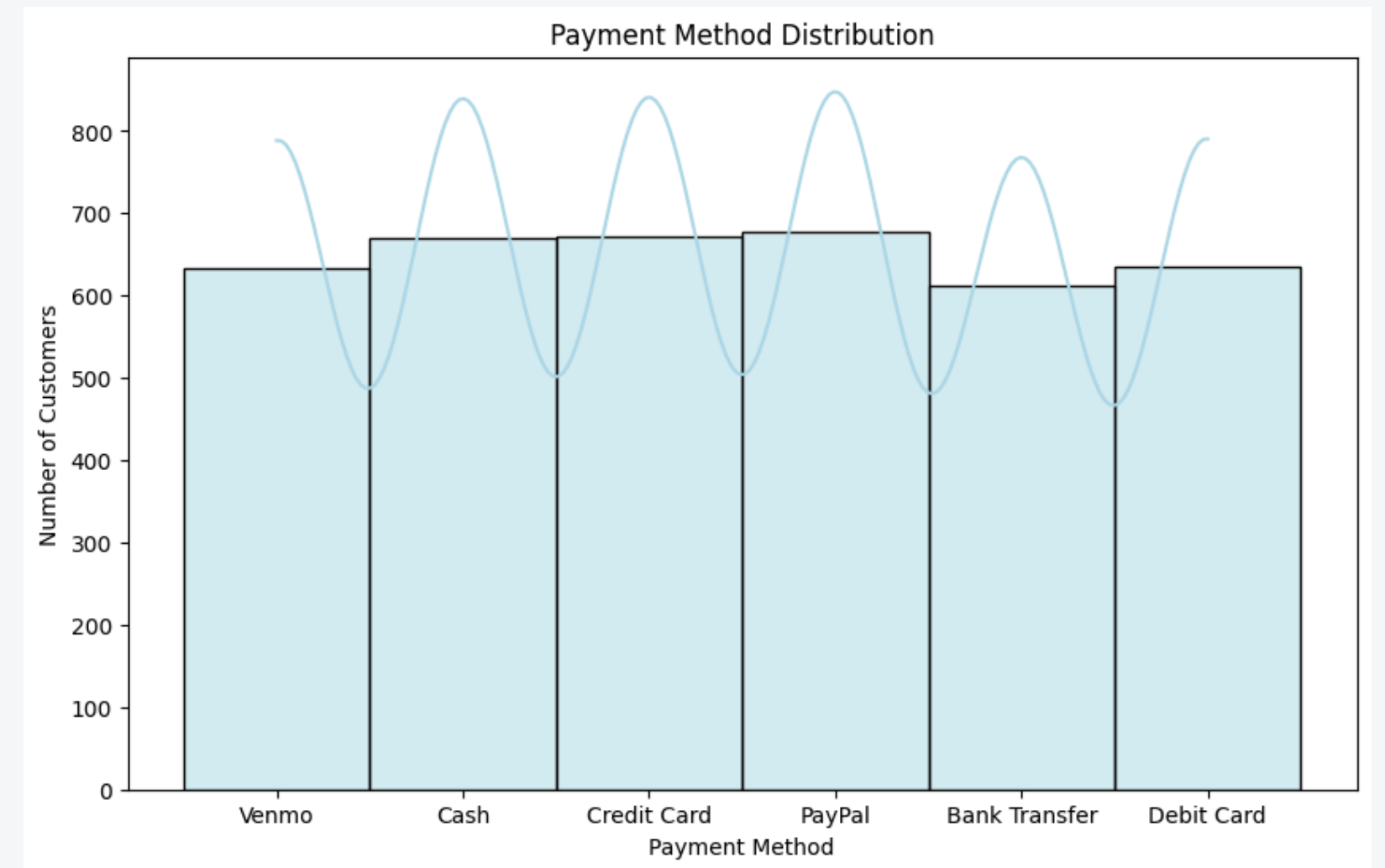
DATA ANALYSIS & INSIGHT
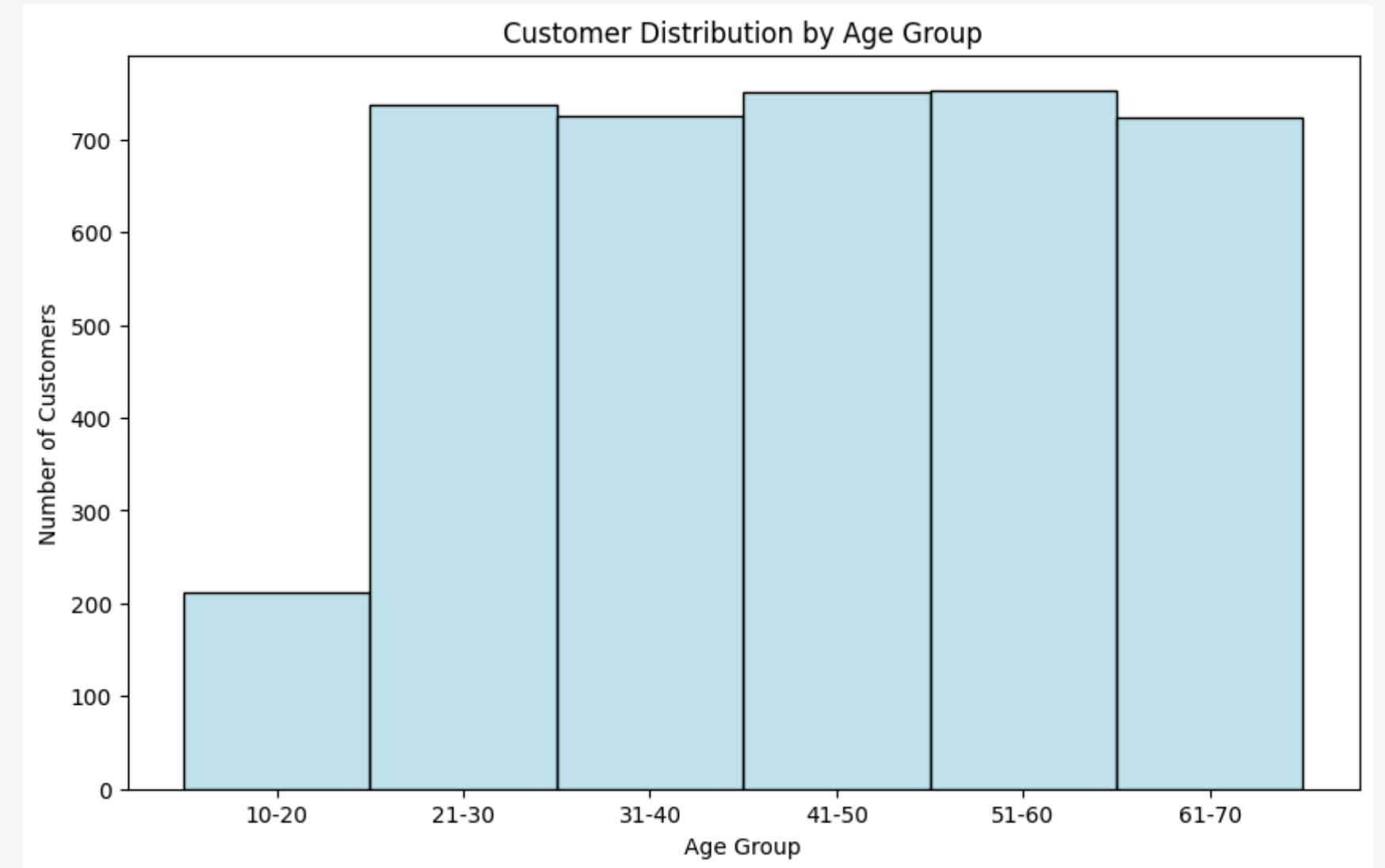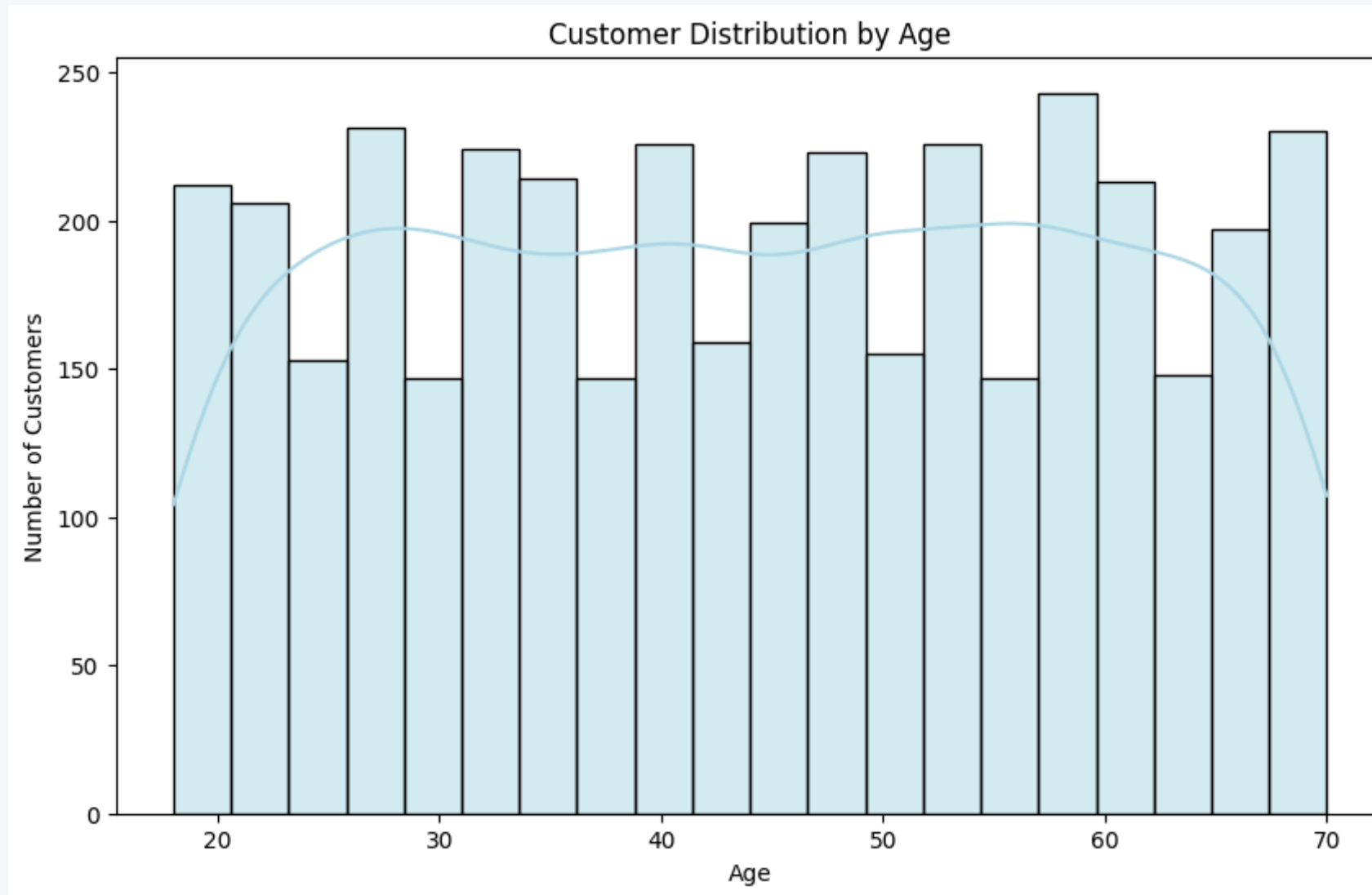
The **average age of customers is 44 years**, with the youngest being 18 years old and the oldest being 70 years old.

Most customers come from the 51–60 year age group, followed by the 41–50 year age group.
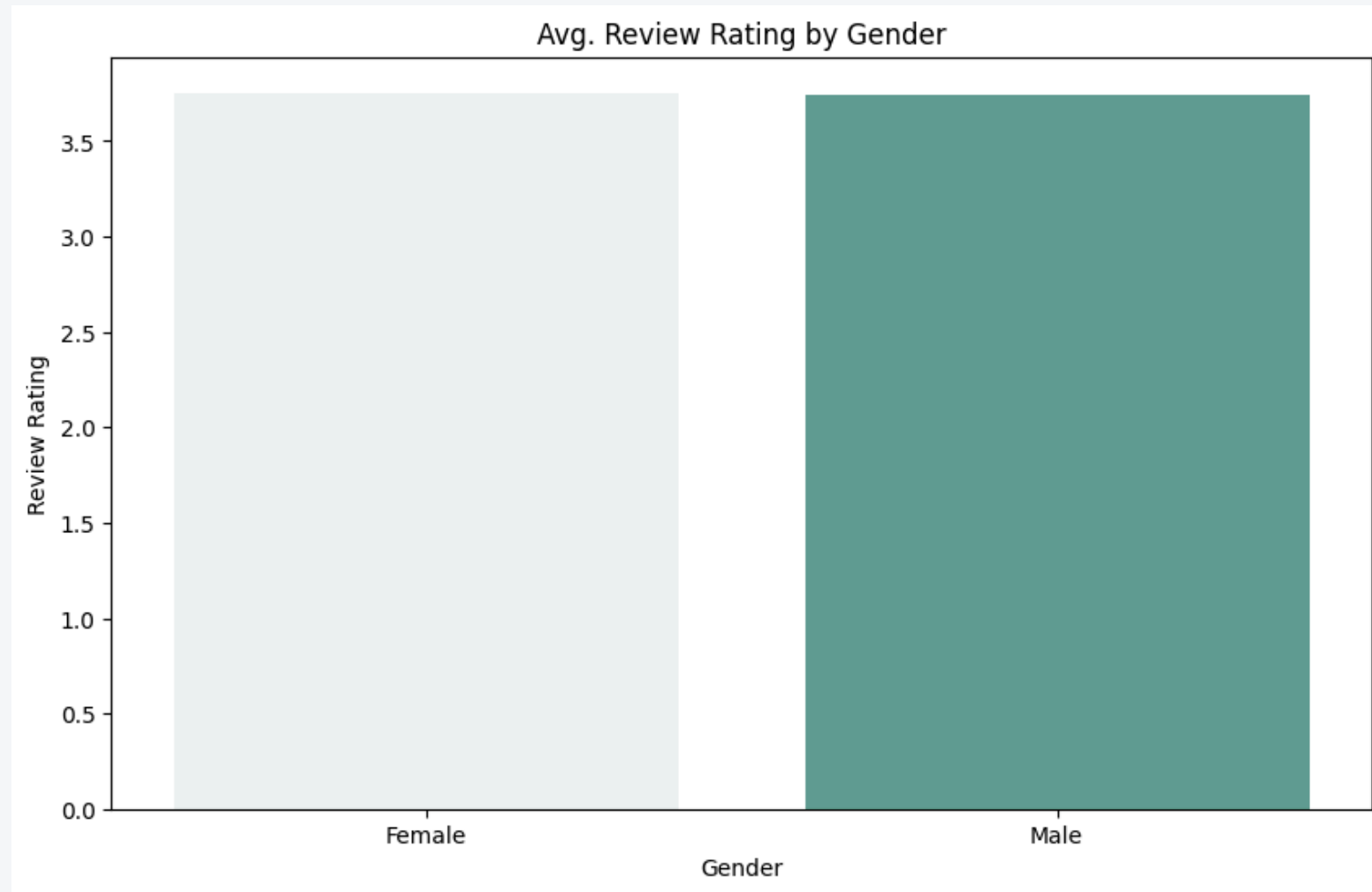
DATA ANALYSIS & INSIGHT

The **most purchased products are from the clothing category**, while the least purchased are from the outerwear category.
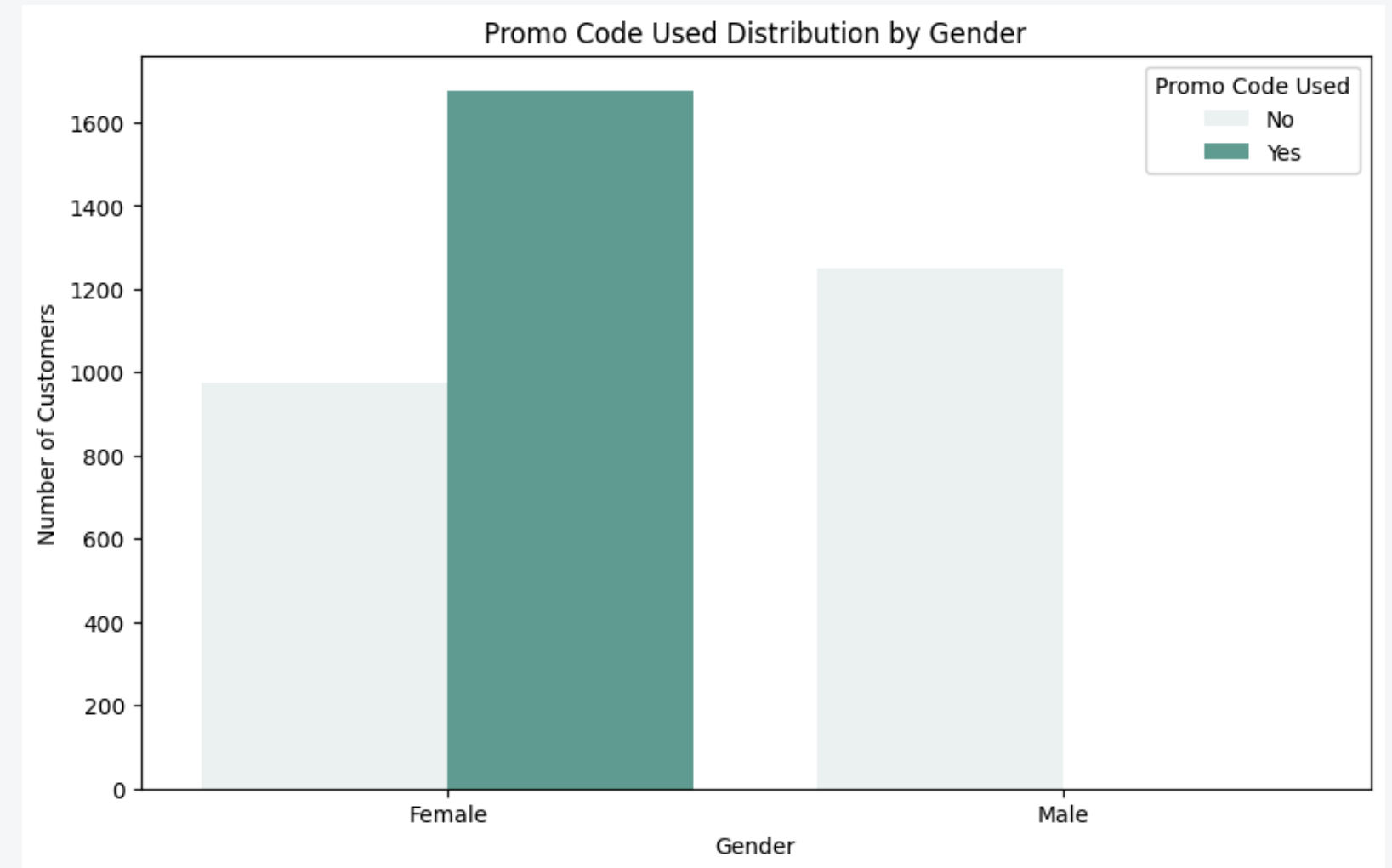
The **most used payment method is PayPal**, while the least used is bank transfer.

DATA ANALYSIS & INSIGHT

Avg. Review Rating by Gender

Promo Code Used Distribution by Gender

Female customers tend to give higher review ratings than male customers.

More female customers use promotions than those who don't, while no male customers use promotions.

DATA ANALYSIS & INSIGHT

	Age Group	Freq. of Purchases	Purchase Amount	Previous Purchases	Promo Code Used
Subscriber	41–50 yo, 51–60 yo	Annually, Weekly	21–40 USD	31–40 x	Yes
Non-Subscriber	51–60 yo, 21–30 yo	Quarterly	21–40 USD	21–30 x	No

CUSTOMER SEGMENTATION
## with K-Means Clustering

The Elbow Method
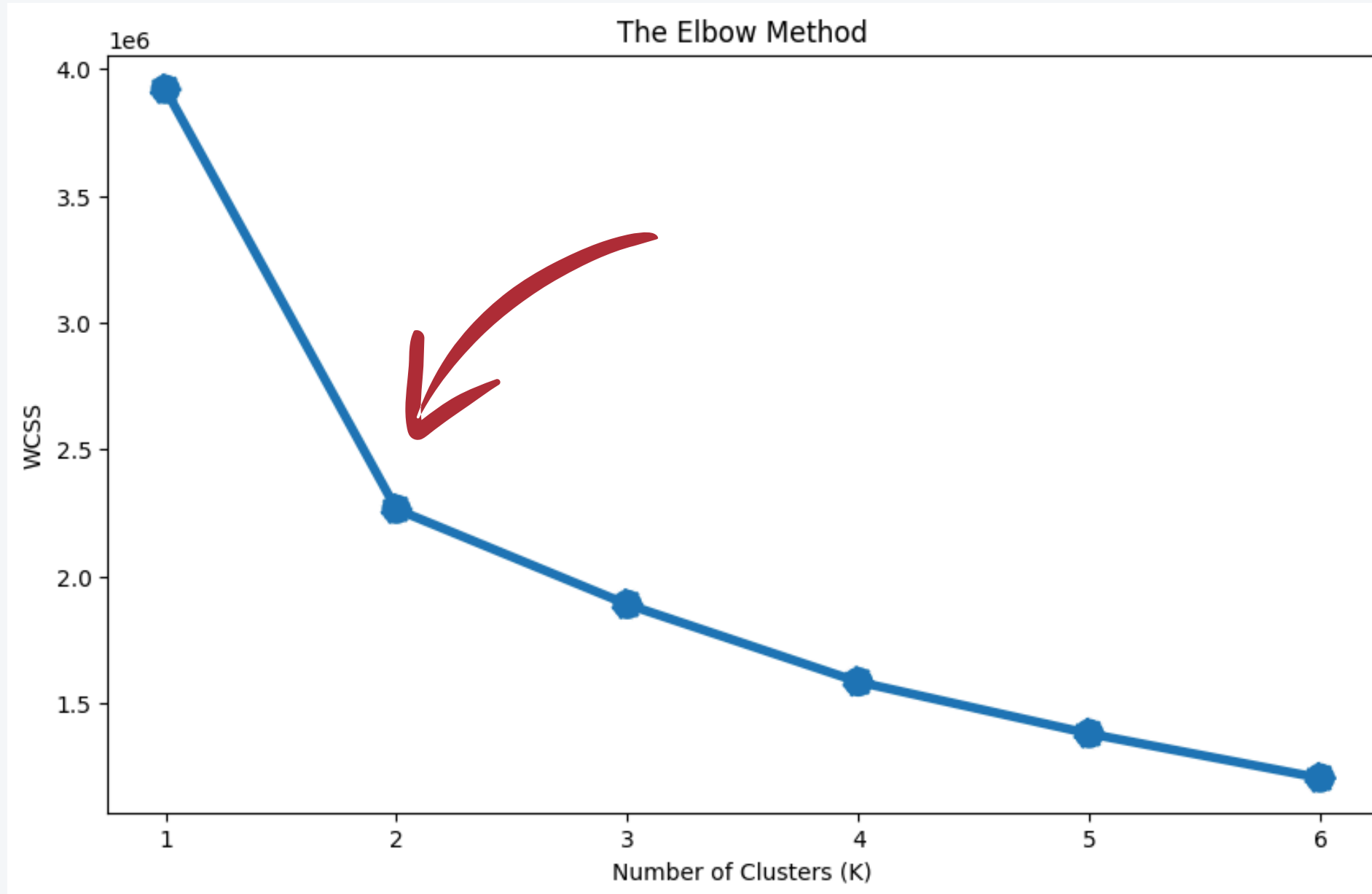
```
Iteration: 1 with inertia value: 3919258.173333335
Iteration: 2 with inertia value: 2263773.100029841
Iteration: 3 with inertia value: 1888236.8477324147
Iteration: 4 with inertia value: 1583006.1291083808
Iteration: 5 with inertia value: 1378868.3097649105
Iteration: 6 with inertia value: 1203838.4879756665
```

From the elbow method, it can be seen that the point that forms the elbow is when K=2 and the inertia value after K=2 begins to decrease linearly. So **the optimal value for this data is K = 2.** For further analysis, we're going to use K=2.

CUSTOMER SEGMENTATION
visualization with Radar Chart

Compare Cluster
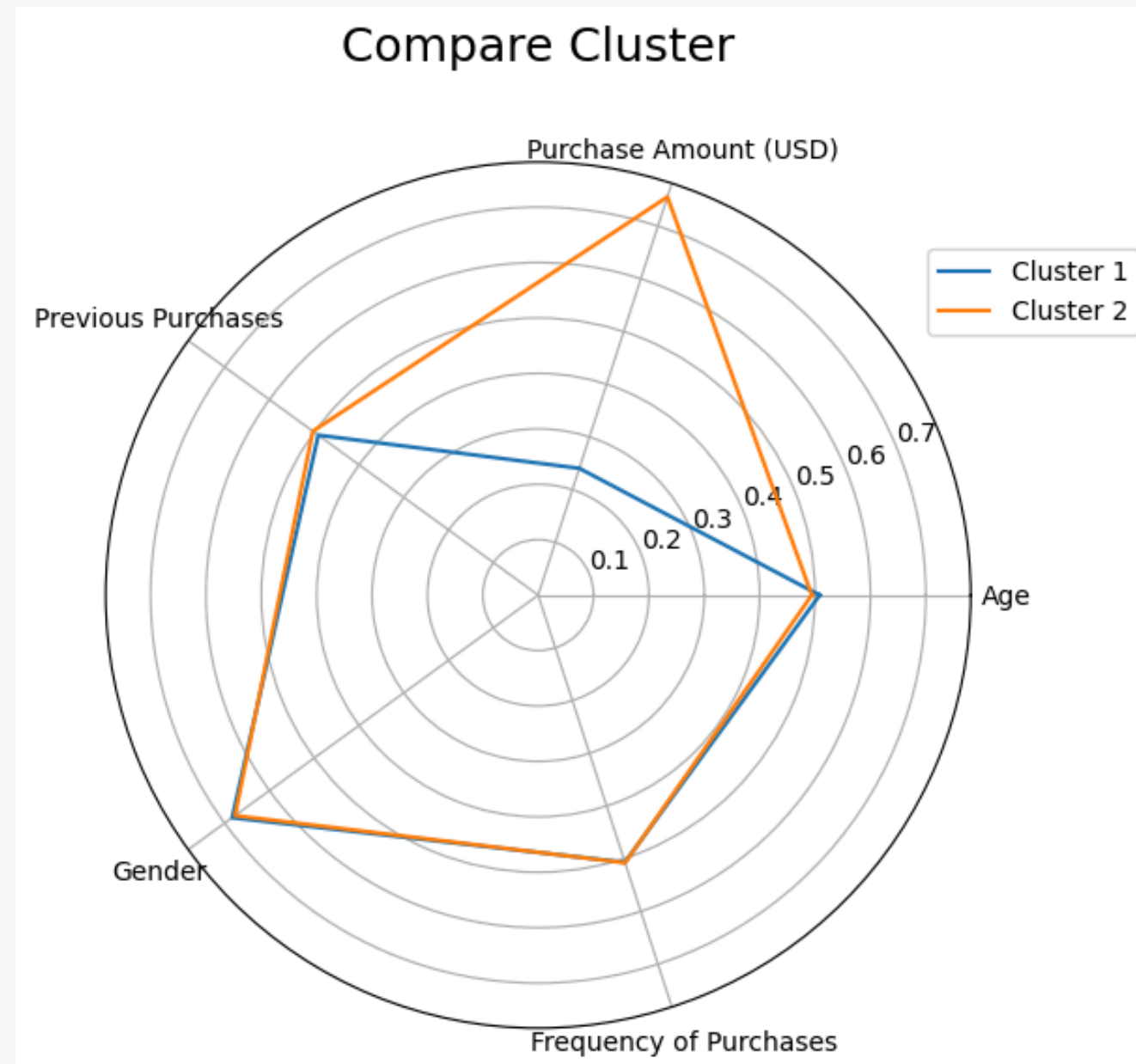
Based on the results, it turns out that the 4 variables, namely **previous purchases, age, gender, and frequency of purchases from both clusters have similarities.** The only difference is in the purchase amount, with **Cluster 2 spending more on purchasing goods than Cluster 1.**

RECOMMENDATION

1. **Marketing Strategies** :
- cluster 1 : discount, special deals on frequently purchased items and bundled offers
- cluster 2 : loyalty programs and exclusive offers

2. **Boost Total Order**: identify the factors causing high total orders in the top 1 of each product category and seasonality, then apply these insights to parts with low total orders

SALES DASHBOARD

Total Customer	Total Product	Total Order	Total Customer Location
3,900	**25**	**3,900**	**50**

Select Customer Age: All
Select Product Category: All
Select Location: All

Customer Count by Gender

Female 32%
Male 68%

Customer by Subscription Status

Yes 27%
No 73%
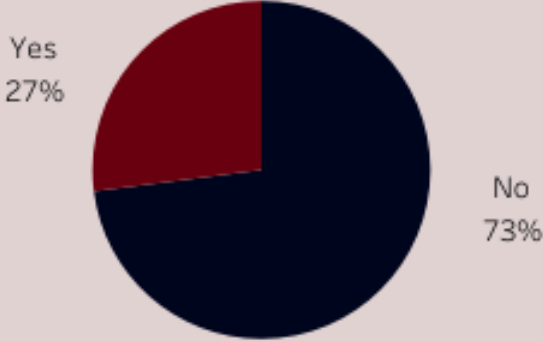
Customer by Promo Used

Yes 43%
No 57%
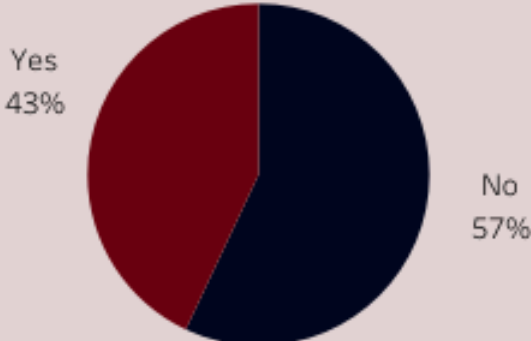
Customer by Discount Applied

Yes 43%
No 57%

Total Order by Product

Accessories — Clothing — Footwear — Oute..

Jewelry, Sunglasses, Belt, Scarf, Hat, Handbag, Backpack, Gloves, Pants, Blouse, Shirt, Dress, Sweater, Socks, Skirt, Shorts, Hoodie, T-shirt, Jeans, Sandals, Shoes, Sneakers, Boots, Jacket, Coat

Total Order by Season

Season	Value
Spring	999
Fall	975
Winter	971
Summer	955
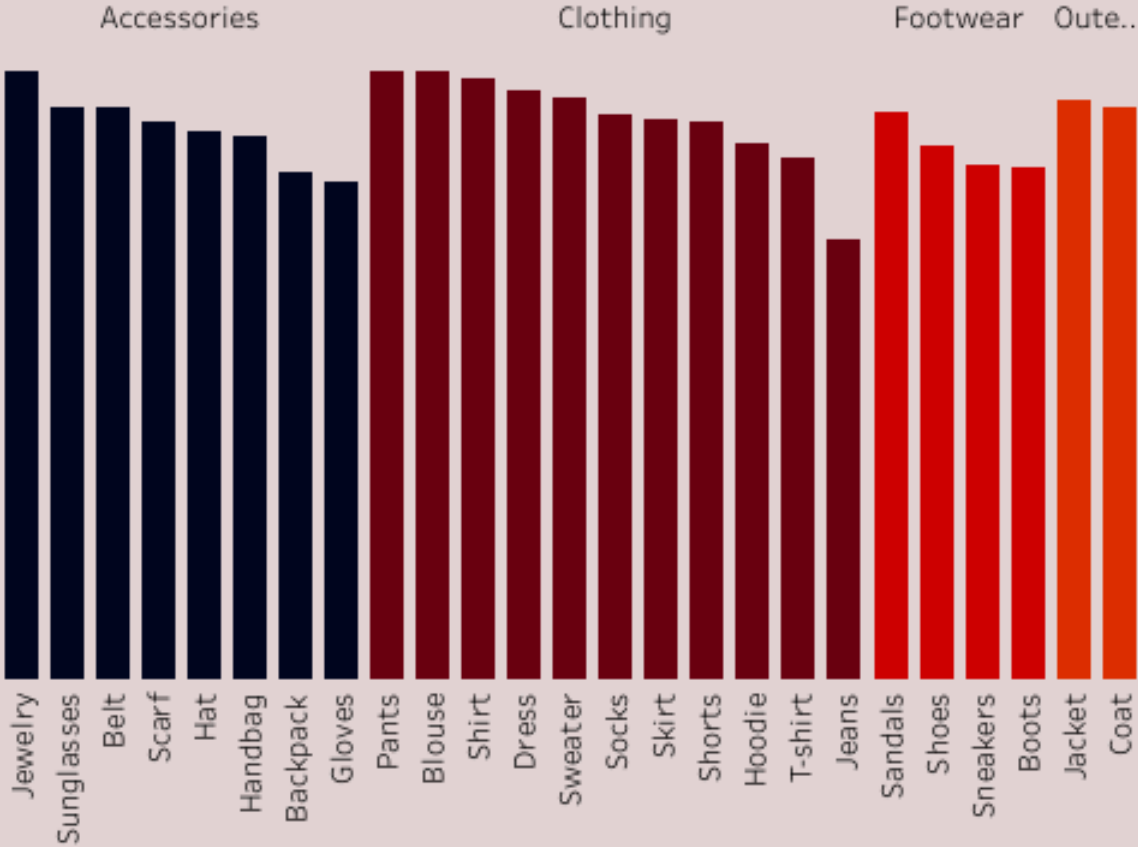
Total Order by Payment Method

Payment Method	Value
PayPal	677
Credit Card	671
Cash	670
Debit Card	636
Venmo	634
Bank Transfer	612

Purchase Amount by Location

Alaska, Canada, Oregon, South Dakota, New Hampshire, Colorado, West Virginia, New Mexico, South Carolina, Florida, Hawaii, Mexico, Colombia

© 2024 Mapbox © OpenStreetMap
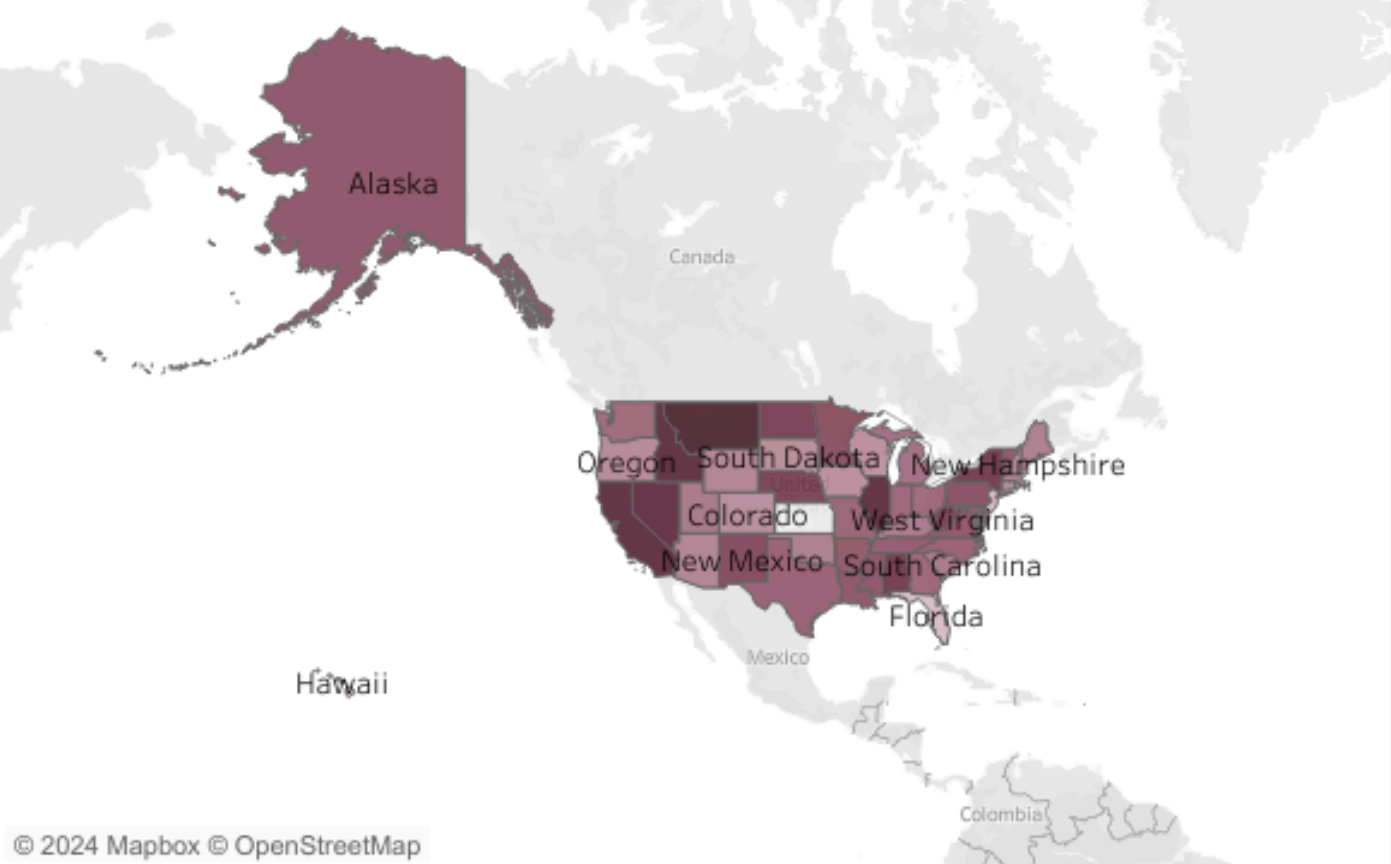
Thanks!

www.linkedin.com/in/nabila-ijtihatil-ismail/

nabila.ijtihatil@gmail.com

appendix

Business Requirement Document

Link Google Colab   Link Tableau Dashboard   Link BRD