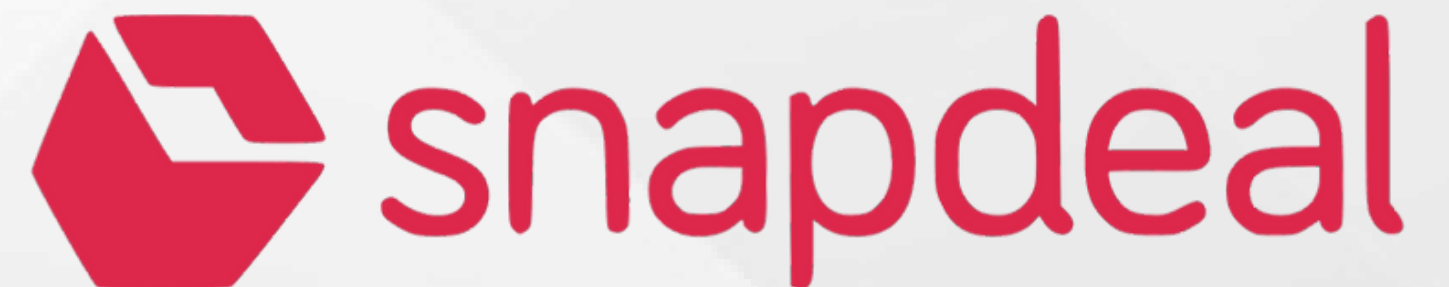# CUSTOMER SEGMENATION ANALYSIS

INSIGHTS AND STRATEGIES

by - Ishant Cane

snapdeal

INTRODUCTION

- The objective of this analysis is to group the customer base into different categories according to their interaction with the platform and purchase patterns.

- By identifying key characteristics of each segment, the business can personalize its marketing efforts, improve customer retention, and encourage repeat customers.

- The customers were divided into segments using the K-Means clustering algorithm, and useful insights were determined by analyzing the resulting clusters.

●SUMMARY

- Customer segmentation in retail business is vital to the success of modern enterprises. The sudden change in consumer expectations has led to the rise of targeted personalization. Customer segmentation is the proven method of ensuring that the business remains profitable despite intense competition in an unpredictable business environment.

- Customer segmentation separates customers into groups based on common characteristics, such as spending habits, gender, similar interests, and geographic locations.

- Customer segmentation is of uttermost importance in E-commerce business because it helps in building stronger customer relationships, improve product offerings, and accurately guides efficient marketing campaigns.

- In this project, we will be using K-means clustering algorithm to identify different customer segments which classify different customers into different categories.

CUSTOMER SEGMENTATION

CUSTOMER

Visiting

LOYAL

HIGH VALUE

LOW VALUE

HIGH POTENTIAL

Spending

Segmentation type

K-MEANS CLUSTERING

K-means clustering is an unsupervised machine learning algorithm used to partition observations into clusters, where each observation belongs to the cluster with the nearest mean (centroid). The algorithm aims to minimize the sum of squared distances between data points and their assigned cluster centroids.
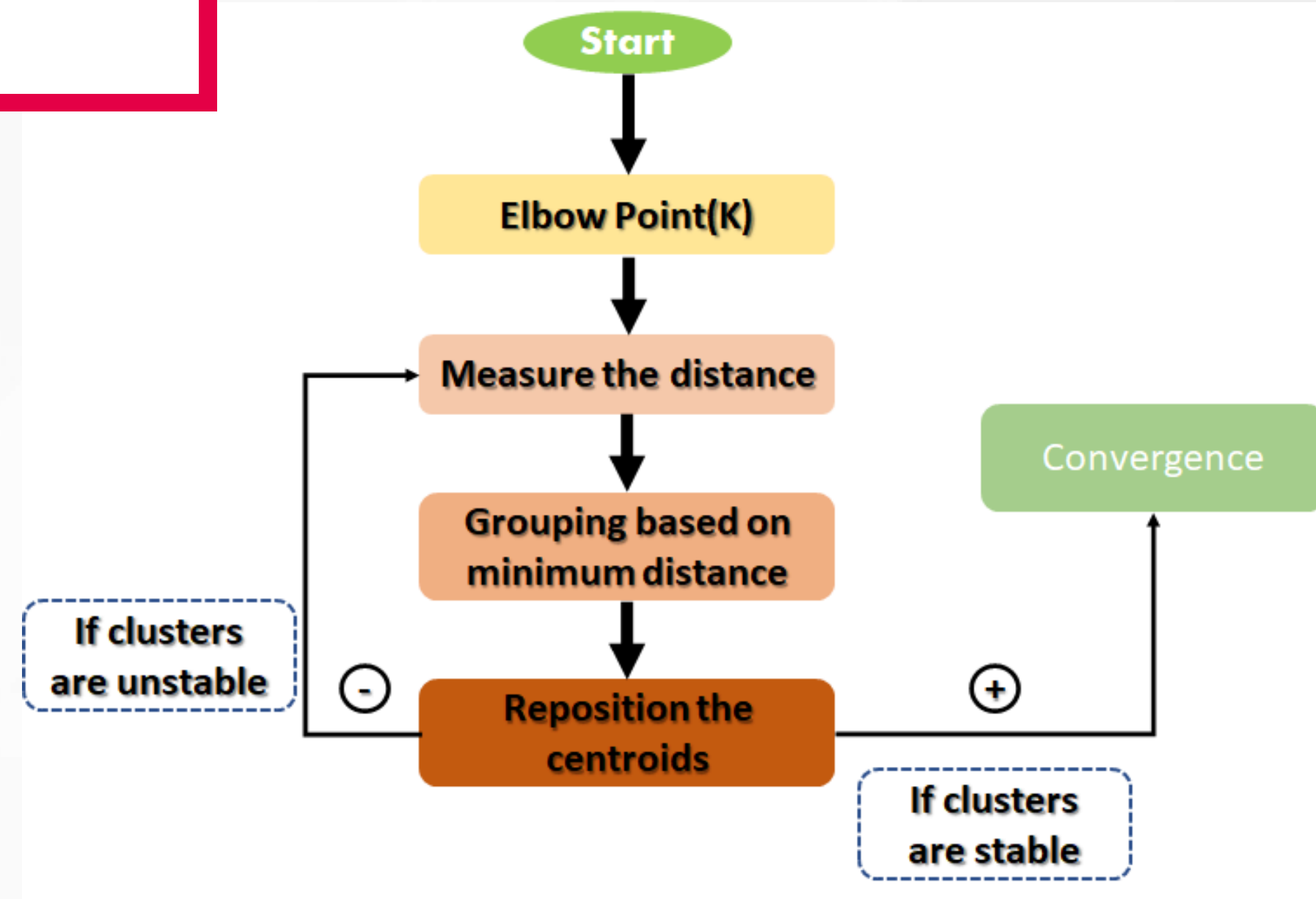
DATASET OVERVIEW

◆ Source: Online Retail dataset

◆ Contains 541,909 transactions

◆ Period: Dec 2010 – Dec 2011

◆ Data fields:

- InvoiceNo – Transaction ID
- StockCode – Product code
- Description – Product description
- Quantity – Units purchased
- InvoiceDate – Date & time of transaction
- UnitPrice – Price per unit (£)
- CustomerID – Unique customer identifier
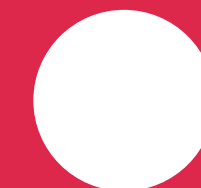- Country – Customer's country

For this analysis, I have used the **RFM approach** to create clusters. It include creating new features from the existing features as:
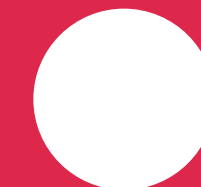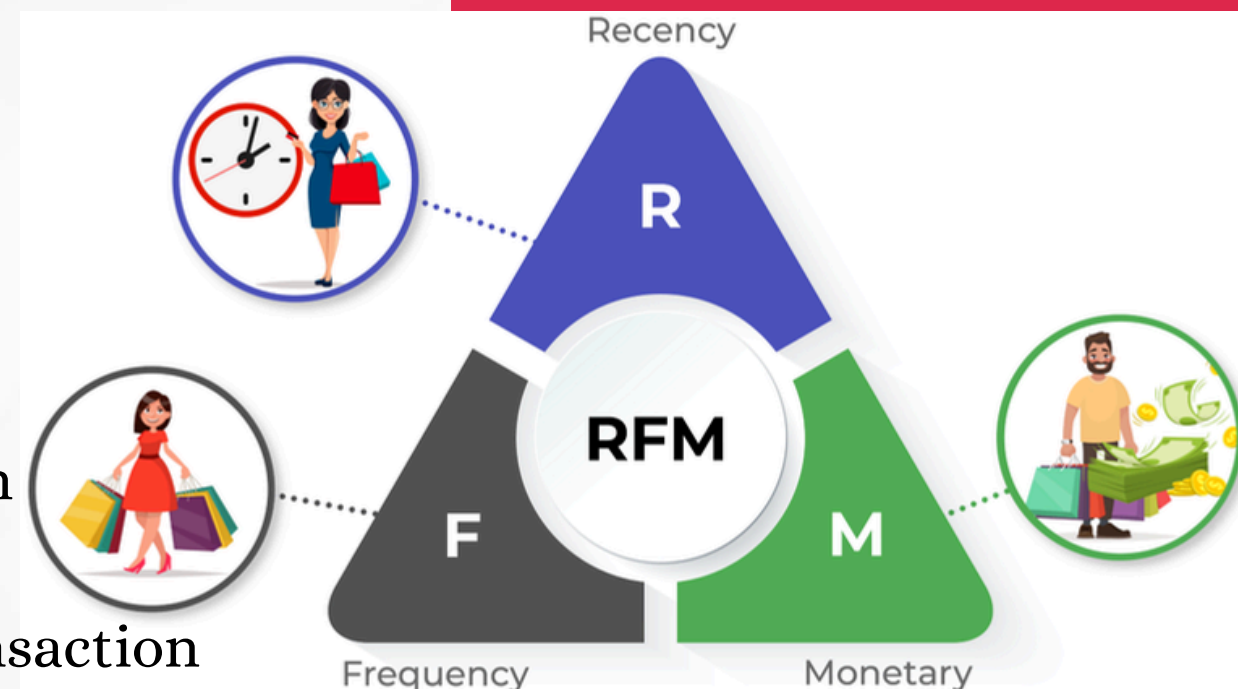
Recency: How many days has passed since the last purchase by a customer
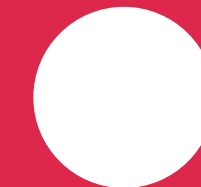
Frequency: How many times customer had made purchase in a year

Monetary: Total amount spent by each customer

STEPS TAKEN IN THIS PROCESS

The following steps has been taken in performing of this analysis:

Preprocessing

- The raw customer profile data was processed by removing data with missing values and handling outliers using the Interquartile Range (**IQR**) method.
- Creating new features.
- Then, data was standardize using **StandardScaler.**

Clustering

The optimal number of clusters came out to be **4** using **Elbow method.** Thus 4 clusters were created using **K-means** algorithm.
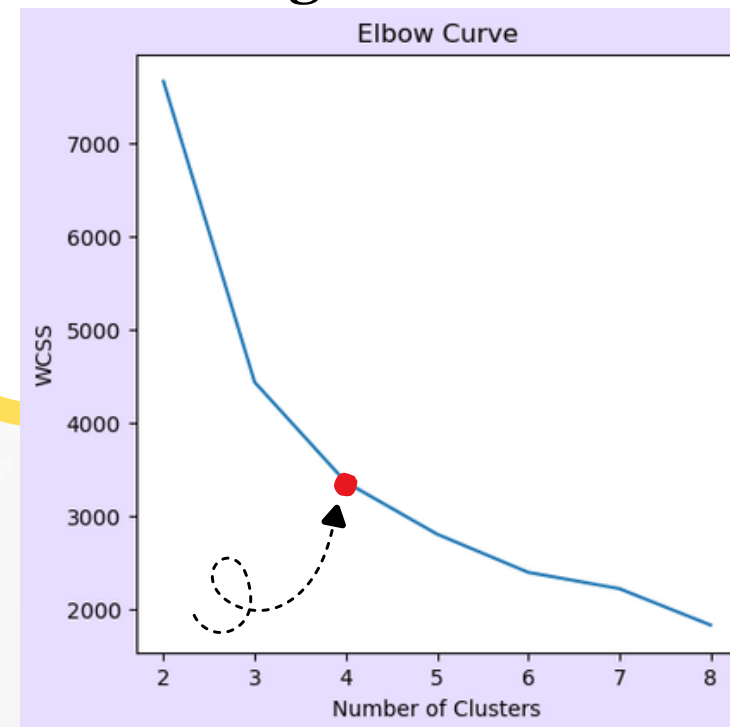

Analysis

All key variables such as purchase **frequency**, **monetary** value, and **recency** were analyzed for each cluster to identify customer behaviors and trends.

CLUSTERS DISTRIBUTION

By Applying Elbow curve method, we got 4 clusters : **cluster 1, cluster 2, cluster 3,** and **cluster 4.**

◆ **Cluster 0** contributes the highest with **53.60%** of customers.

◆ **Cluster 1** contributes **17.47%** of customers.

◆ **Cluster 2** contributes **24%** of customers

◆ **Cluster 3** contributes the lowest **4.73%** of customers.


Cluster Distribution
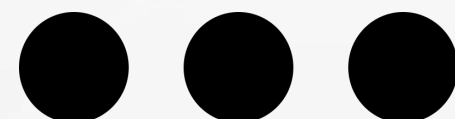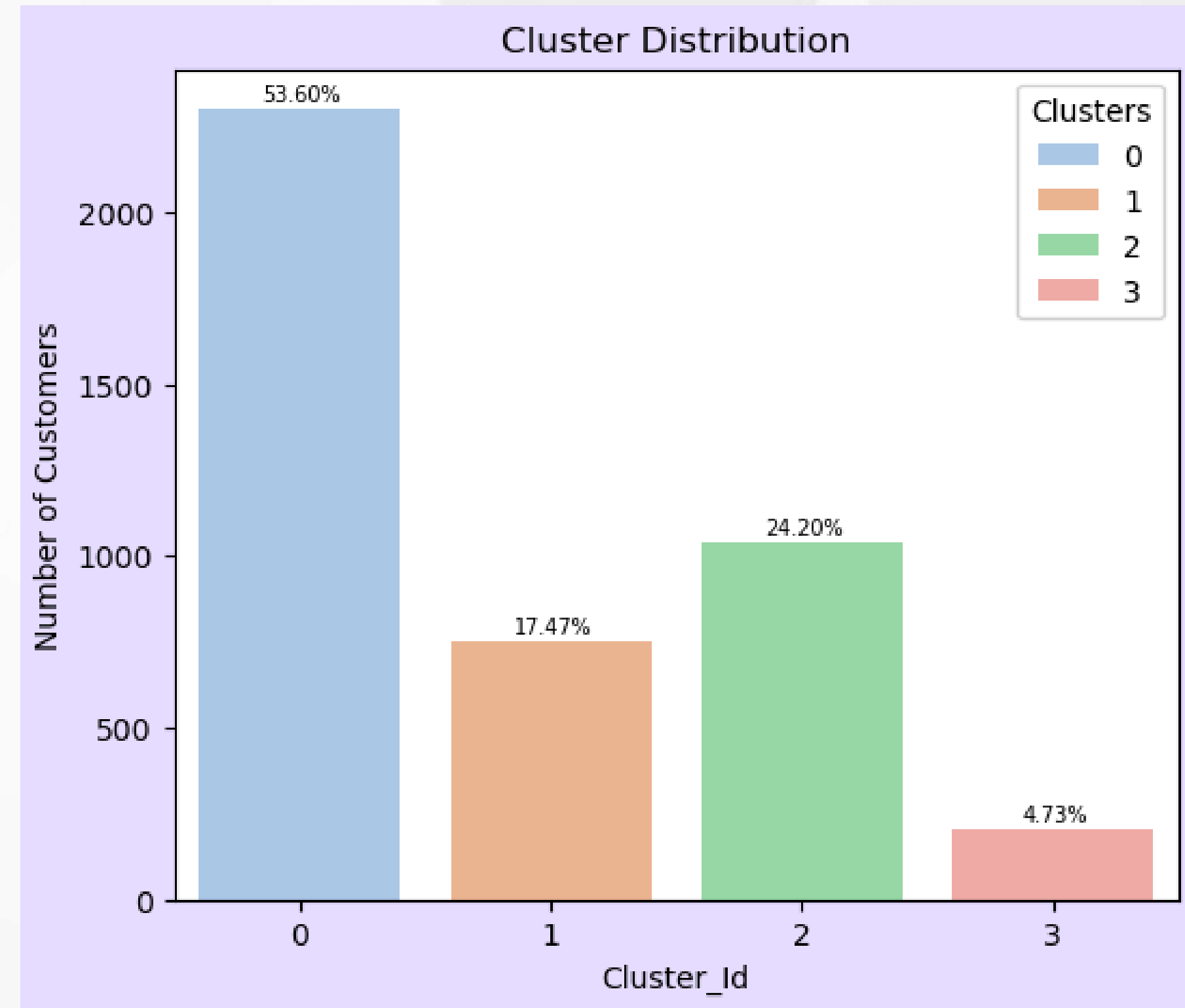
WHO SPEND THE MOST?

- While customers in **cluster 1** contributes the most, customers in **cluster 3** spends highest averagely.

- Customers in **cluster 0** have low average spends but they contribute highly.

- Customers in **cluster 2** have lowest average spends and they also contribute lowest.
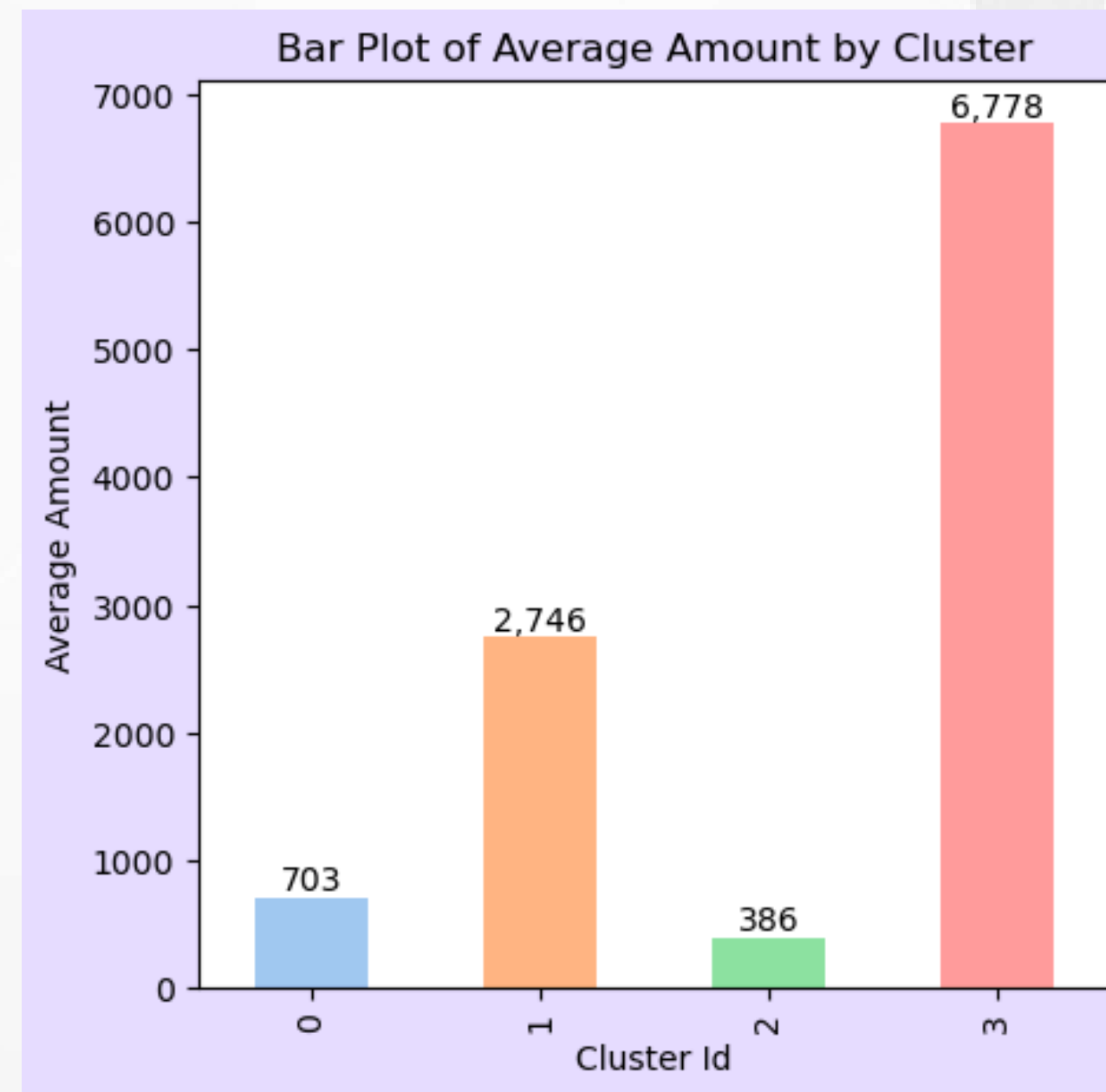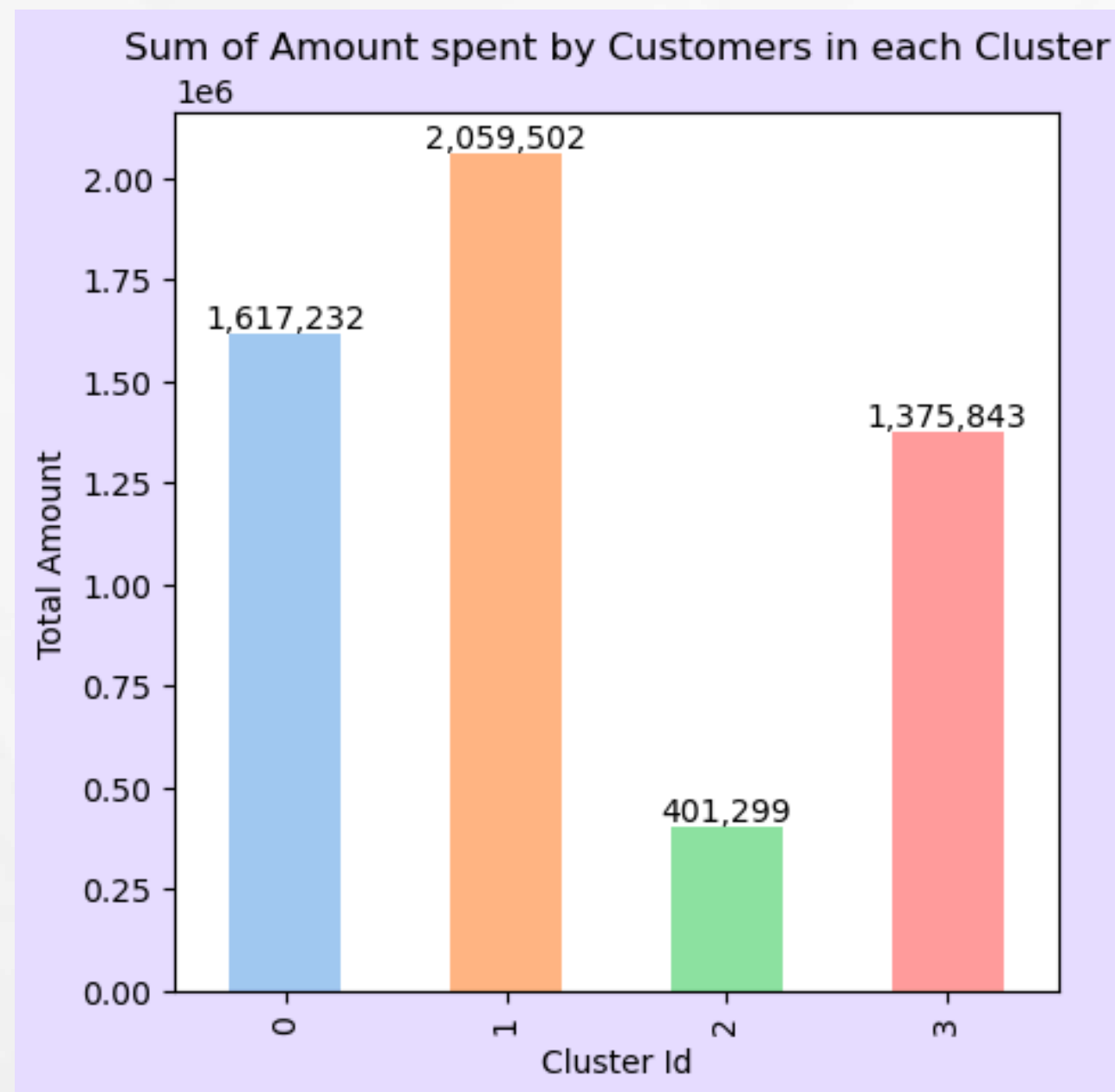
WHO BUYS THE MOST?

◆ **Cluster 3** has the highest frequency and the highest ordered quantity

◆ **Cluster 0** has the moderate frequency and the moderate ordered quantity

◆ **Cluster 1** has the low frequency and the low ordered quantity

◆ **Cluster 2** has the lowest frequency and very low ordered quantity

Bar Plot of Average Frequency by Cluster

Bar Plot of Average Quantity by Cluster
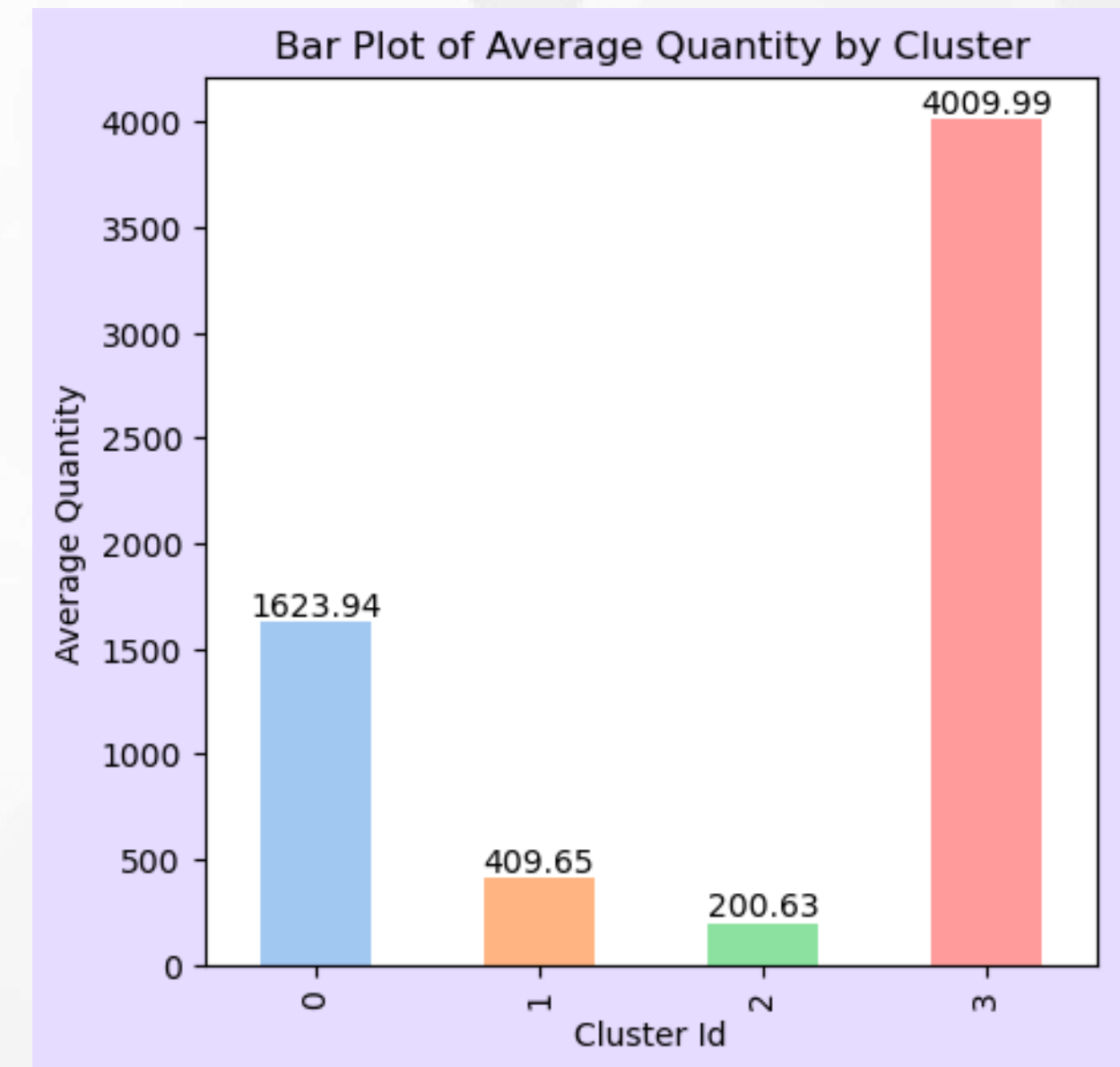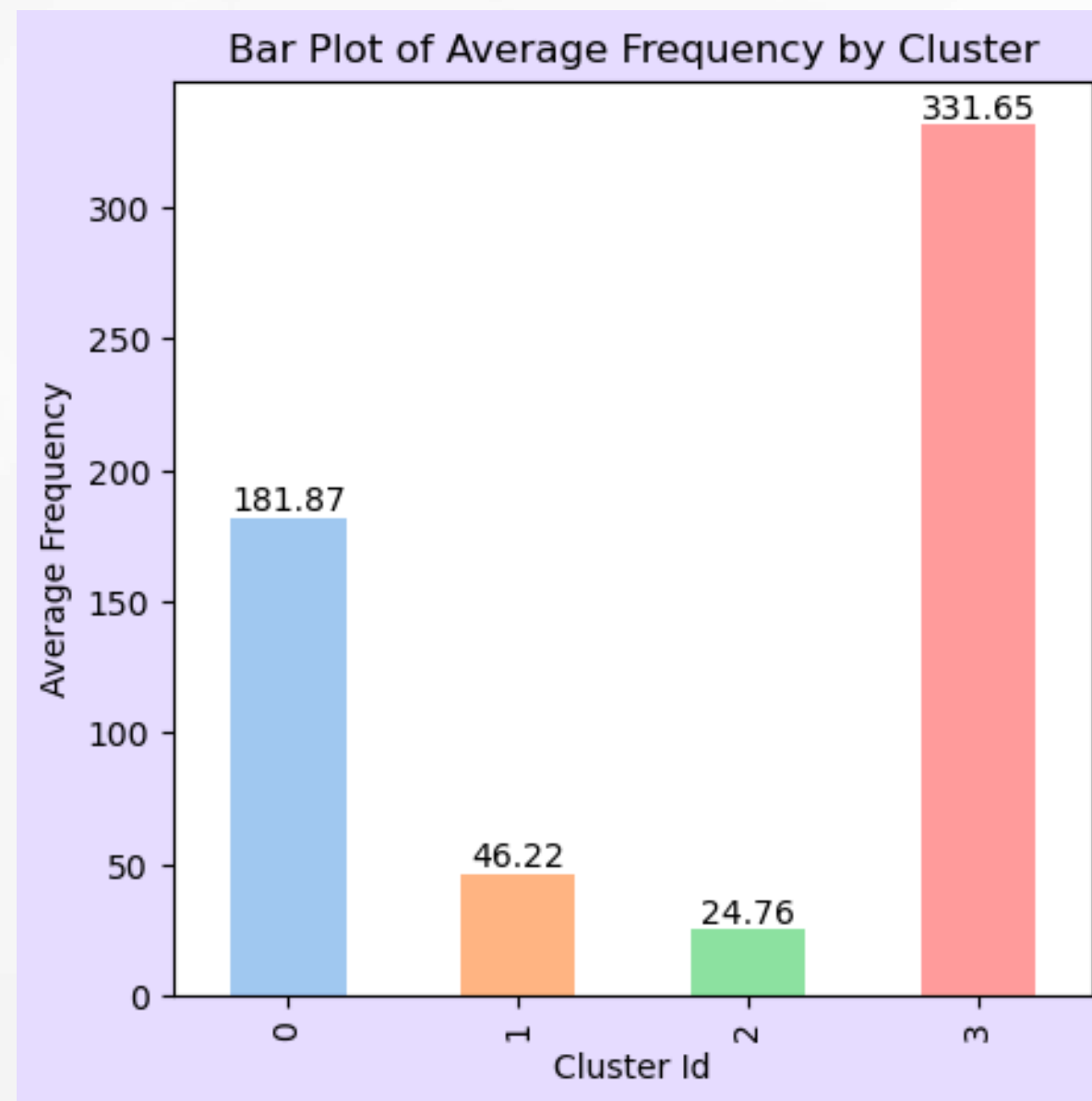
WHO SHOPS RECENTLY?

◆ **Cluster 3** shops the most recently, while **cluster 1** and **2** also shopping recently

◆ **Cluster 2** has a very high average recency, means these customers has probably stopped shopping from us.

Bar Plot of Average Recency by Cluster
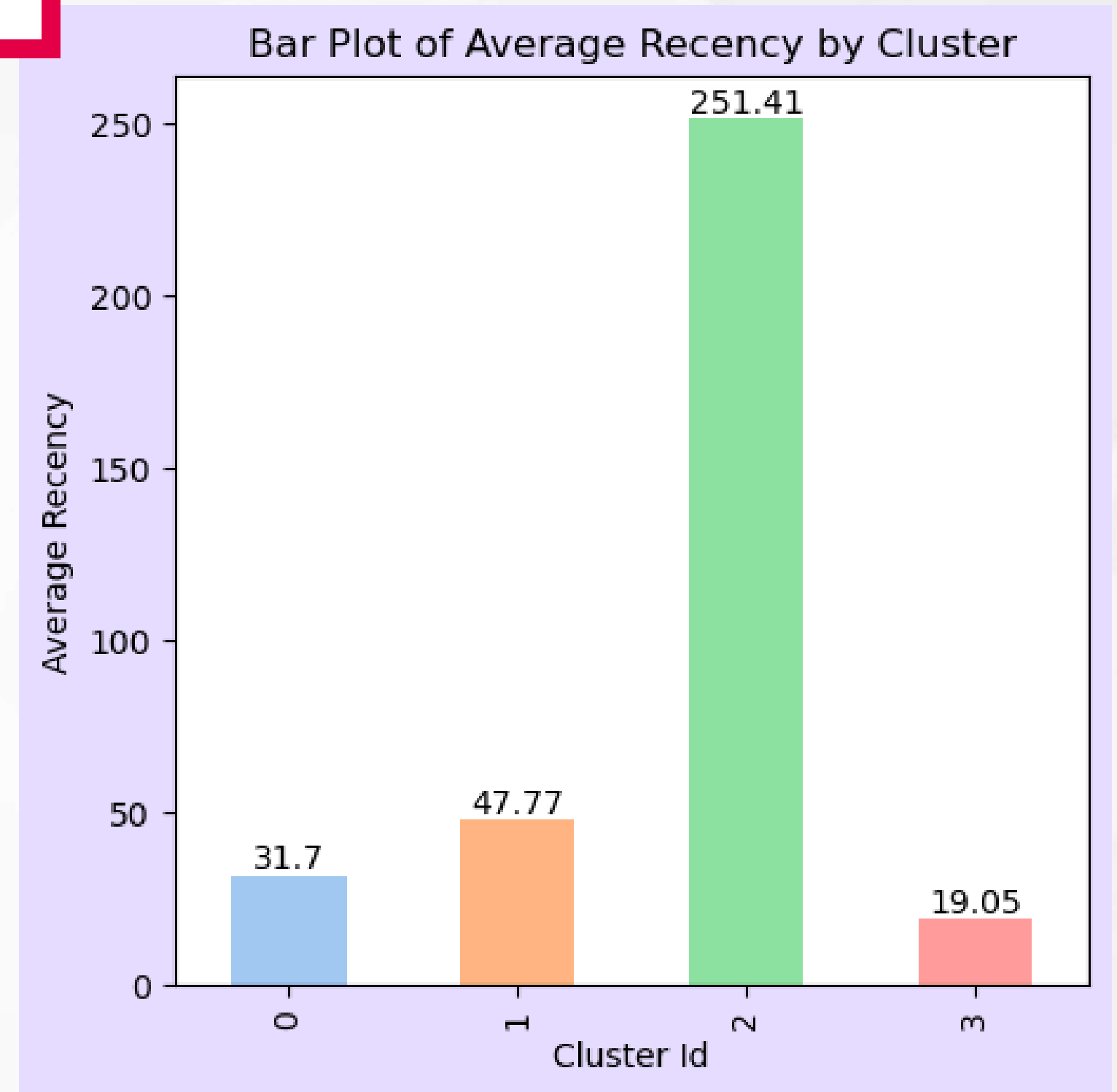
Customer Profile

Cluster 0: Less Active Customers

Key Characterstics

- Largest number of buyers.
- Spend in low amount.
- Less frequent
- Order in moderate quantity.

Customer Behaviour

These customers are occasional shoppers who exhibit low loyalty and engagement.
This segment could include new customers who have only made one or two purchases or occasional shoppers. They contribute little to revenue.

How to engage them more?

- We can provide them first purchase or entry level discounts to get them started.
- Use personalized promotions (discounts, coupons, bundled offers) to increase spending.
- Send reminder emails or product recommendations to increase engagement.

Customer Profile

Cluster 1: Loyal & Engaged Customers

Key Characterstics

- Spend in descent amount.
- Moderate to high Frequent.
- Make recent purchase.
- Order in low to moderate quantity.

Customer Behaviour

This segment represents loyal customers who regularly made purchase. They buy often, spend a decent amount, and shop recently but order in low quantity.
They add good value to the company.

How can we make them spend more?

- We can introduce discount on high amount purchase
- Encourage their loyalty by offering personalized deals or exclusive recommendations based on their preferences.
- Reward their loyalty with points, cashback, or exclusive deals.
- Giving them early access to new launches or sales.

Customer Profile

Cluster 2: At Risk or Churned Customers

Key Characterstics

- Least Spenders
- Very low frequency.
- Had long time since making any purchase..
- Order very less quantity.

Customer Behaviour

This group is the least engaged and may have stopped being a customer. They have the highest recency and least ordered quantity and also spend least.

How to bring them back?

- We can conduct welcome back campaign.
- Giving them special discounts or free deliveries to get them back.
- Remind them of new and big launches.
- But don't overspend on these since they had already churned.

Customer Profile

Cluster 3: The VIPs and Most Valuable

Key Characterstics

- Spends high amount
- Have very high frequency.
- Have low recency means makes regular purchase.
- Order in high quantity.

Customer Behaviour

This segment represents the most valuable customers, who are highly engaged with frequent purchases and a high lifetime value.

What to do to keep them engage?

- We should treat them specially since they brought highest value and profit to company
- We can give them exclusive perks — like priority support, VIP sales, or personal thank-you notes.
- Since they trust us, we can provide friend or family recommendation discounts to get more such customers through them.
- Do everything to keep them us.

SUGGESTIONS
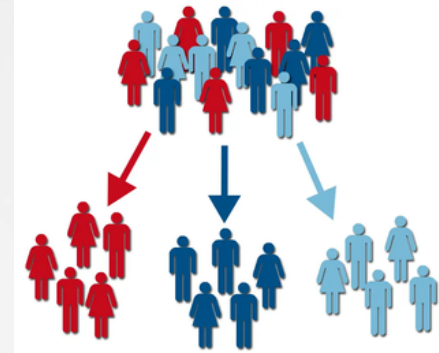
Segment specific market

To maximize engagement, use customize marketing initiatives for each segment.

Increase Revenue

Protect and grow revenue from VIPs & Loyal customers.

Run loyalty programs and special deals for such customers. Feedback Loops and Continuous Improvement:

Personalization and Targeting

Based on past purchases, current engagement levels, and shopping habits, send personalized product recommendations and offers.
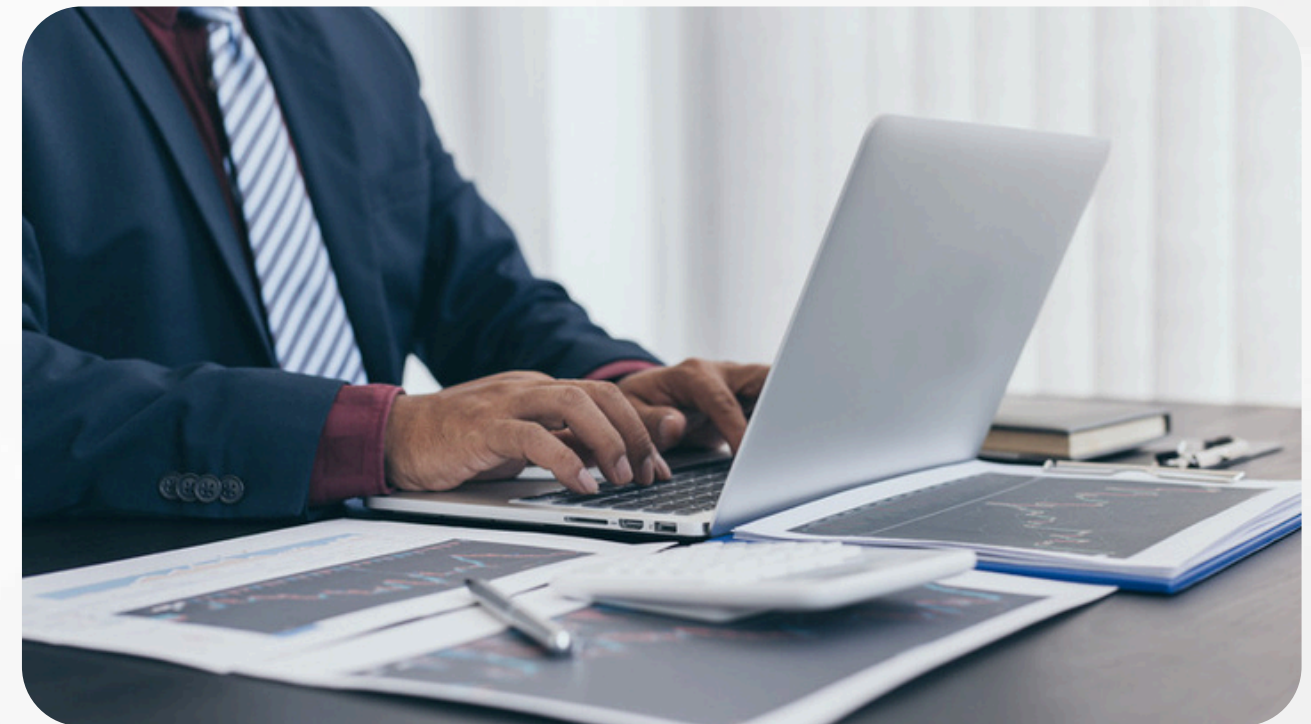
Feedback Loop and Continuous Improvement:

Run targeted feedback campaigns for each cluster. The insights we get will not only confirm profiling but also give clear directions on what actions will actually work.
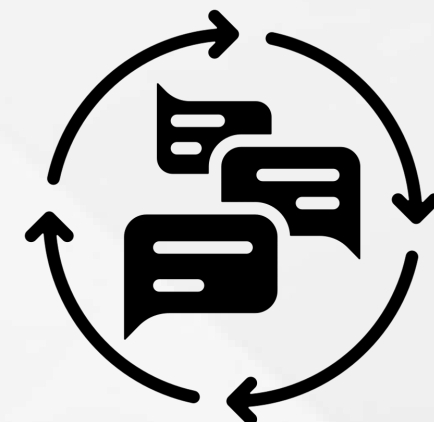
CONCLUSION

In this project, we successfully applied customer segmentation techniques to an online retail dataset consisting of over half a million transactions. After cleaning and preprocessing the data, we engineered meaningful features such as recency, frequency, and monetary value to capture customer purchasing behavior. Using the K-means clustering algorithm, we divided the customer base into distinct groups that reveal important behavioral differences. The results provide actionable insights for the business, such as identifying high-value customers who can be targeted with loyalty programs, occasional buyers who may need promotional offers to increase engagement, and price-sensitive customers who respond better to discounts. Overall, this segmentation empowers the business to tailor its marketing strategies, improve customer satisfaction, and optimize resource allocation, ultimately driving higher profitability and stronger customer relationships.