DATA SCIENCE PORTFOLIO

Iska Okta Fauziah

Hello!

My name is Iska Okta Fauziah

A detailed-oriented fresh graduate. Passionate about data science. Seeking to looking for data-driven experience as data analyst and data scientist.

EDUCATION

Universitas Hasanuddin
Makassar

2018-2022
Bachelor of English Literature

Dibimbing.id
South Jakarta(Remote)

2022-2023
Data Science Bootcamp

August 2023 - now

360DigiTMG
Data Science (Virtual Internship Batch 14)
- ● Obtaining materials regarding data science, include; Cross Industry Standar Process for Machine Learning with Quality Assurance, SQL, and python. Furthermore, the material will be implemented into a practical project case at the culmination of the internship period.

Jan - March 2022

Al-Qad Institute
Online English Private Tutor
- ● Prepared and arranged material into power point.
- ● Taught students via Google Meeting on Saturday and Sunday.
- ● Evaluated student homework and given feedback regarding the homework.

July - Dec 2021

Penerbit BRIN (Badan Riset dan Inovasi Nasional)
Social Media Specialist & Article Writing Division (Virtual Internship)
- ● Writing caption for social media post on instagram, twitter, and tiktok.
- ● Become a coordinator of essay writing.
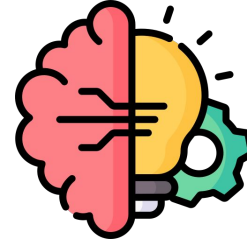- ● Writing an essay and a book review and it is published on website of Penerbit BRIN.

WORKING EXPERIENCES
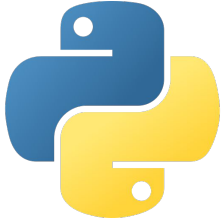
PYTHON
CODING

DATA
VISUALIZATION

MACHINE
LEARNING

SQL
DATABASE

TABLEAU

LOOKER STUDIO
(GOOGLE DATA
STUDIO)

SKILLS

PROJECTS

1 **Flight Delay Prediction**
Link Project

2 **Customer Segmentation Using K-Means**
Link_Project

1

Flight Delay Prediction

Objective:
To predict whether the flight will be delayed or not.

Analysis:
There are some features that affect the delay, namely day of departure, time, length of flight using machine learning classification for prediction of flight delay.

Result:
By using KNN algorithm can predict flight delay by increasing the profit of the airline IDR31 Billion.

2 Customer Segmentation Using K-Means

Objective:

To know the segmentation of customer based on their behaviors and to make the best strategy for each cluster.

Analysis:

- Most of customers from United Kingdom.
- After adding total price column, it have high correlation with quantity column.

Result:

- Customer Need Attention → Recency < 150 days ago, Frequency < 1000 times.
- Customer at Risk → Recency >= 140 days ago, Frequency < 500 times.
- Customer Champion → Recency < 50 days ago, Frequency > 4000 times.
- Customer Loyal → Recency < 100 days ago, Frequency 500 - 3000 times.

CERTIFICATION

CERTIFICATE
OF COMPLETION
Student ID: 8217
Certificate no: 201029DS180607O11
THIS CERTIFICATE IS PROUDLY PRESENTED TO

Iska Okta Fauziah

FOR SUCCESSFULLY COMPLETING
Data Science Bootcamp
HELD ON: 17 Desember 2022 – 11 Juni 2023

ZAKY MUHAMMAD SYAH
CEO DIBIMBING.ID

VEGI FATURRAHMAN
HEAD OF BOOTCAMP

Data Science Bootcamp, 2023
Dibimbing.id

G Great Learning

CERTIFICATE OF COMPLETION

Presented to
Iska Okta Fauziah

For successfully completing a free online course
Data Visualisation in Tableau

Provided by
Great Learning Academy
(On April 2023)

Data Visualization in Tableau, 2023
Great Learning Academy

G Great Learning from **BYJU'S**

CERTIFICATE OF COMPLETION

Presented to
Iska Okta Fauziah

For successfully completing a free online course
Data Science Foundations

Provided by
Great Learning Academy
(On January 2023)

Data Science Foundation, 2023
Great Learning Academy

CERTIFICATION

Intro to Data Analytics, 2022
Revou

Introduction to Data Science with Python, 2022
DQLab
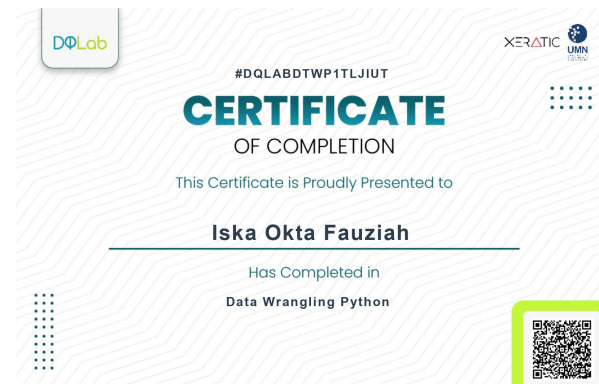
Data Wrangling Python, 2022
DQLab

CUSTOMER
SEGMENTATION
Using K-Means Clustering

OUTLINE

1

Background and Objective

2

Exploratory Data Analysis

3

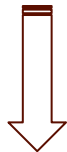**Data Pre-Processing**

4

Modelling And Analysis

5

Recommendations

1 **Background and Objective**

Retail industry is becoming more competitive

service to customers must be improved to retain our customer

Creating better relationship with customers

Increased company profits

Background

How to Know and Provide the Best Services to Customers?

How to Know and Provide the Best Services to Customers?
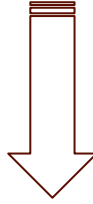
Doing customer segmentation.
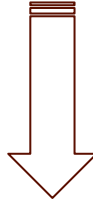
Customer segmentation is the process of grouping customers or consumers in a specific market into smaller segments based on shared characteristics, behaviors, or preferences.

To help businesses and organizations better **understand their customers**, formulate **appropriate marketing and communication strategies**, and ultimately **enhance the products or services they offer**.

Objective

541909 Rows   Numerical (4 features)

8 Features   Categorical (4 features)

This is a transactional dataset which contains all the transactions occurring between 01/12/2010 - 09/12/2011

Link Dataset

- **InvoiceNo**
- **StockCode**
- **Description**
- **Quantity**
- **InvoiceDate**
- **UnitPrice**
- **CustomerID**
- **Country**

2 Exploratory Data Analysis

Top 5 Countries with The Highest Customers
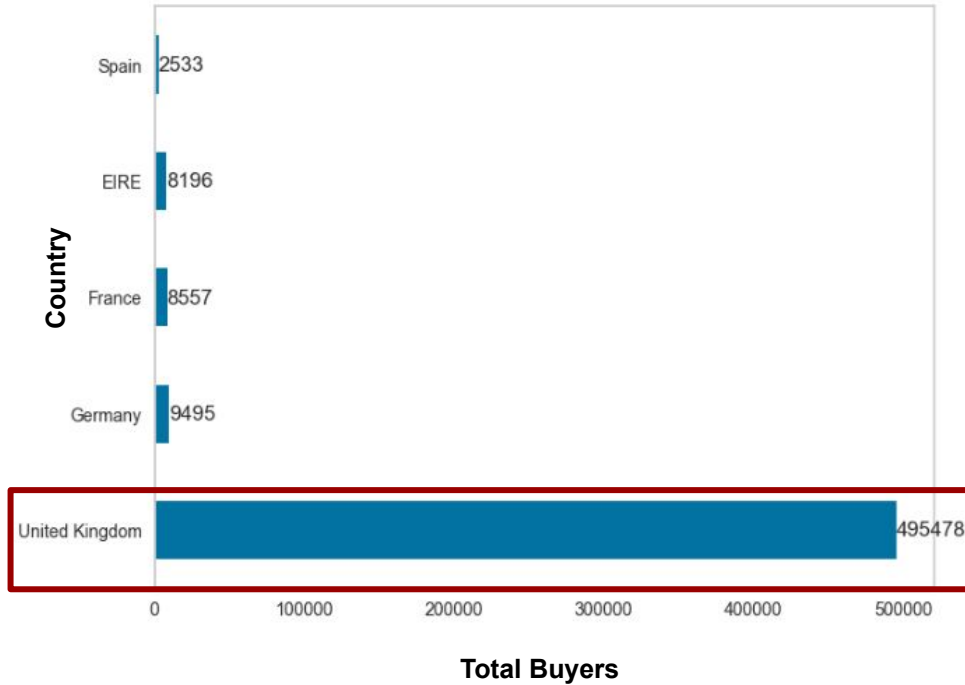
United Kingdom has the biggest buyer with a total amount of **495478** followed by Germany, France, EIRE, & Spain

Top 10 products with the Most Transactions
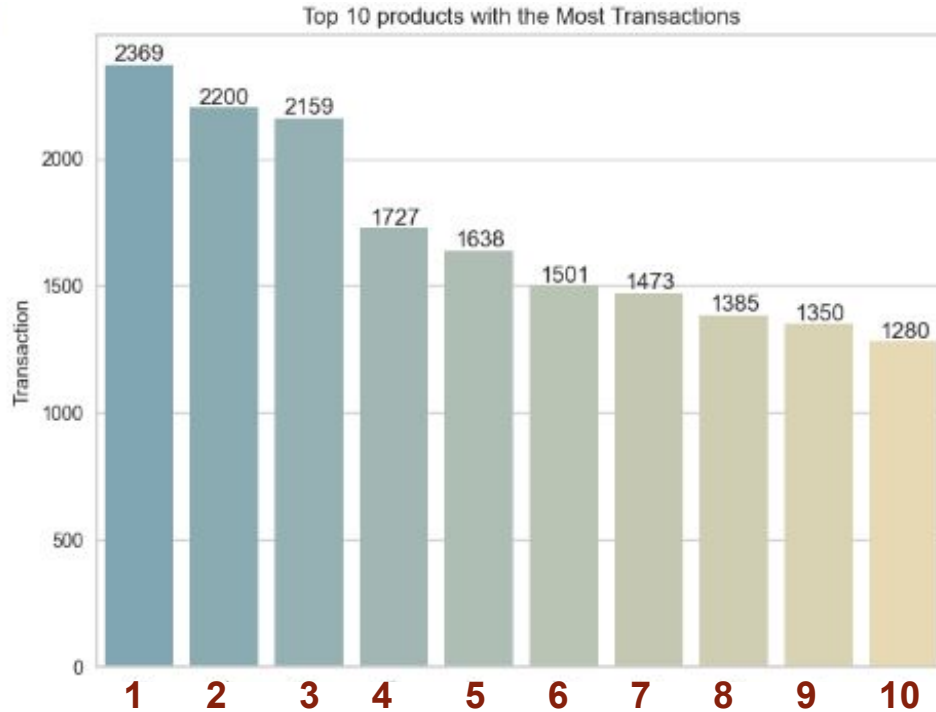
Top 10 Products with the Most Transactions

1. WHITE HANGING HEART T-LIGHT HOLDER
2. REGENCY CAKESTAND 3
3. UMBO BAG RED RETROSPOT
4. PARTY BUNTING
5. LUNCH BAG RED RETROSPOT
6. ASSORTED COLOUR BIRD ORNAMENT
7. SET OF 3 CAKE TINS PANTRY DESIGN
8. PACK OF 72 RETROSPOT CAKE CASES
9. LUNCH BAG BLACK SKULL
10. NATURAL SLATE HEART CHALKBOARD

Correlation Heatmap

Quantity and **total_price** has high correlation.

It can be assumed that **the more items purchased, the higher the price**.

Note: In EDA, it added one column 'total_price'

3 **Data Pre-Processing**

STEP1 — Data Cleaning

STEP2 — Filtering Data (3 Countries UK, Germany, & France)

STEP3 — Creating Dataset RFM for Modelling

STEP4 — Data Scalling
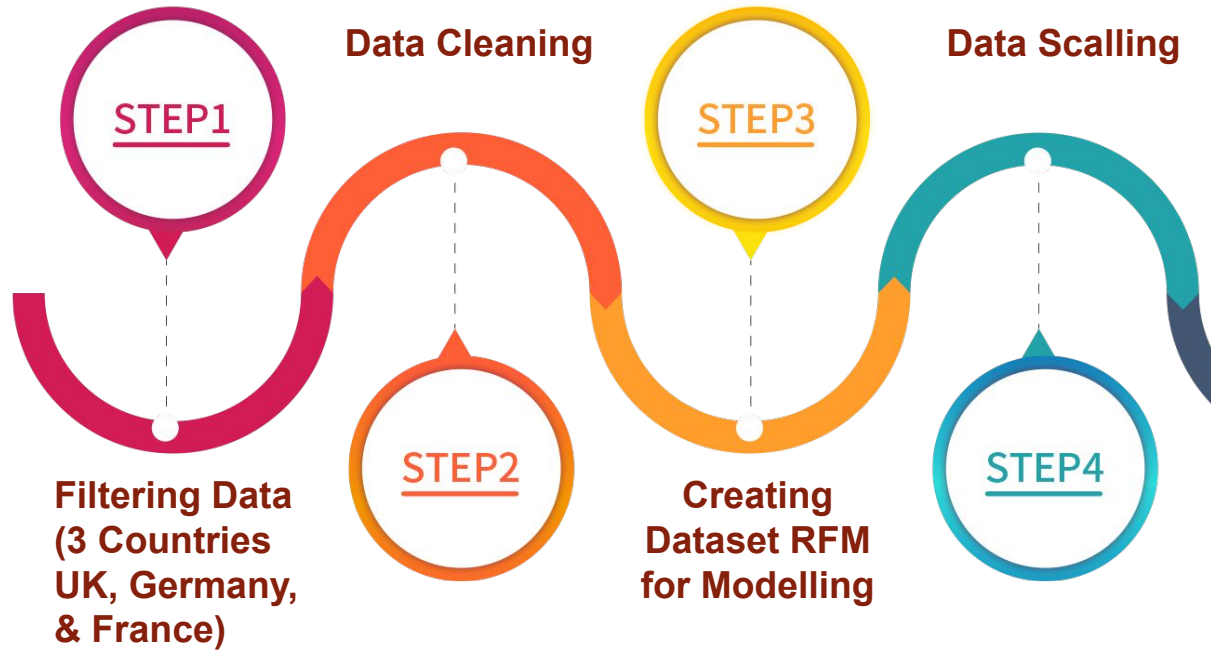
What is RFM?

RFM stands for :

Recency (R): How recently customers have made their purchases.

Frequency (F): How often customers have made their purchases.

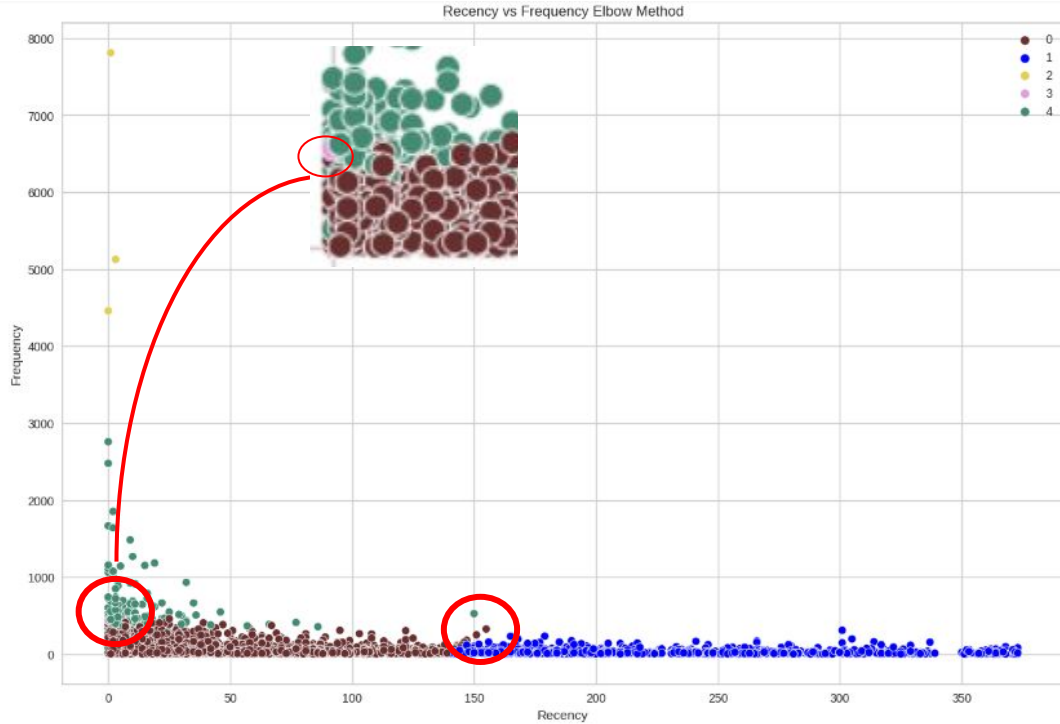**Monetary**(M): How much money customers have paid for their purchases.

4 **Modelling And Analysis**

Elbow Method

Distortion Score Elbow for KMeans Clustering
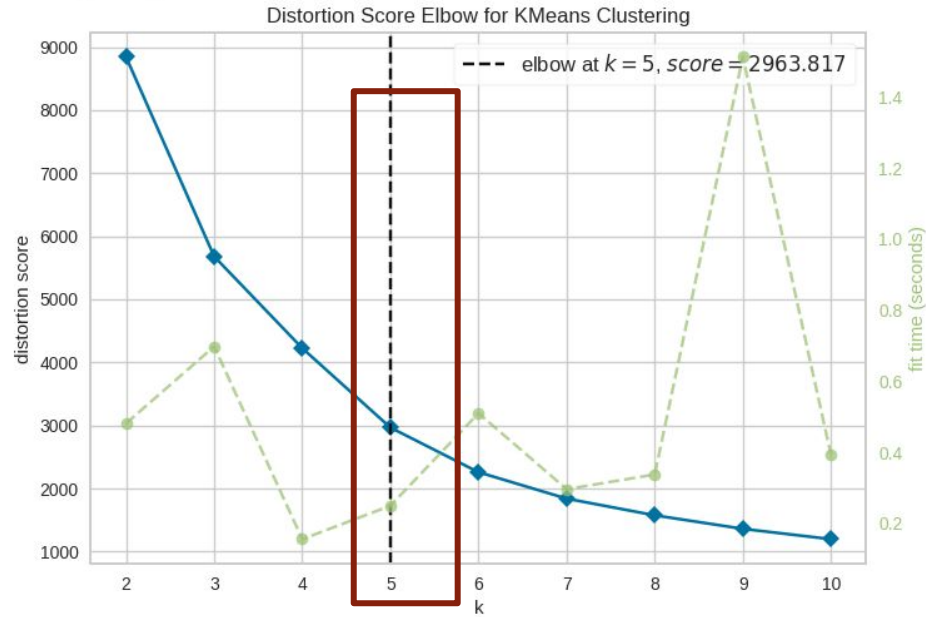
According to the graph, the best K value for K-Means Clustering is **K = 5**.

Visualization Clustering Recency & Frequency using Scatter Plot

Cluster 0 ●
Cluster 1 ●
Cluster 2 ●
Cluster 3 ●
Cluster 4 ●

This diagram shows the distribution of customer data which is divided into five clusters according to the K-Means Clustering algorithm.
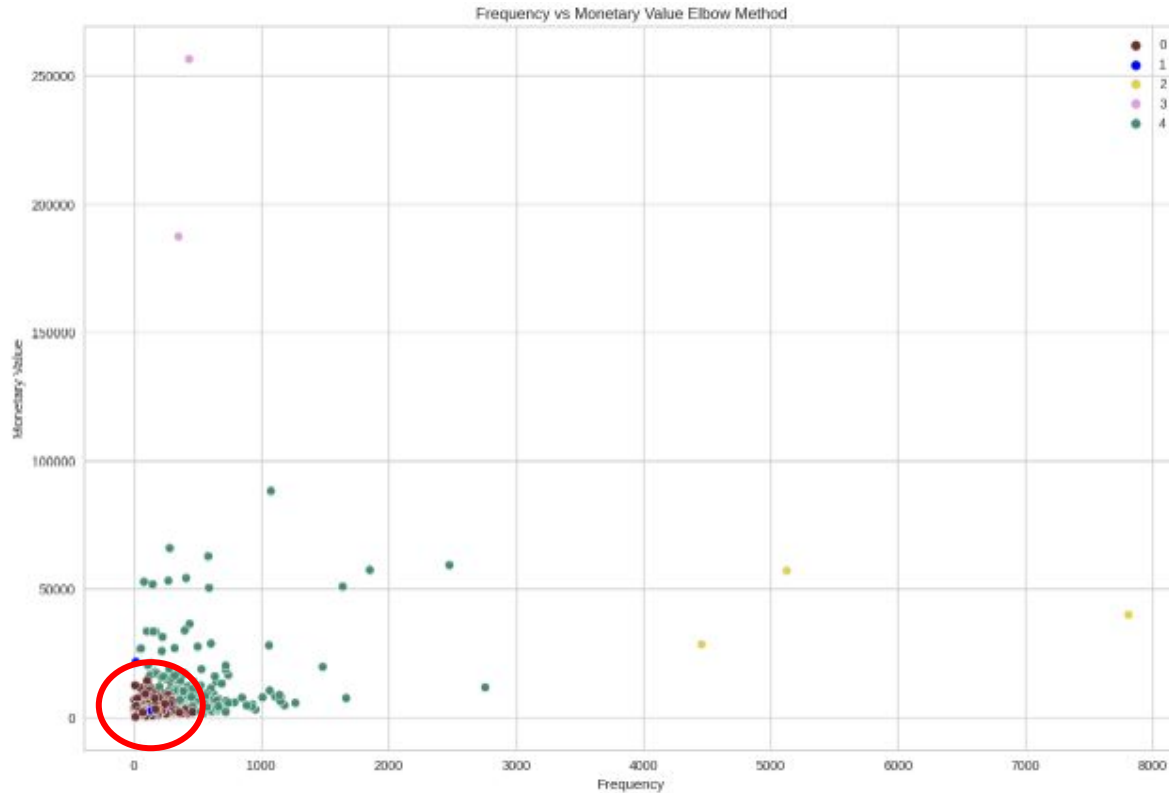
Visualization Clustering Recency & Monetary Value using Scatter Plot
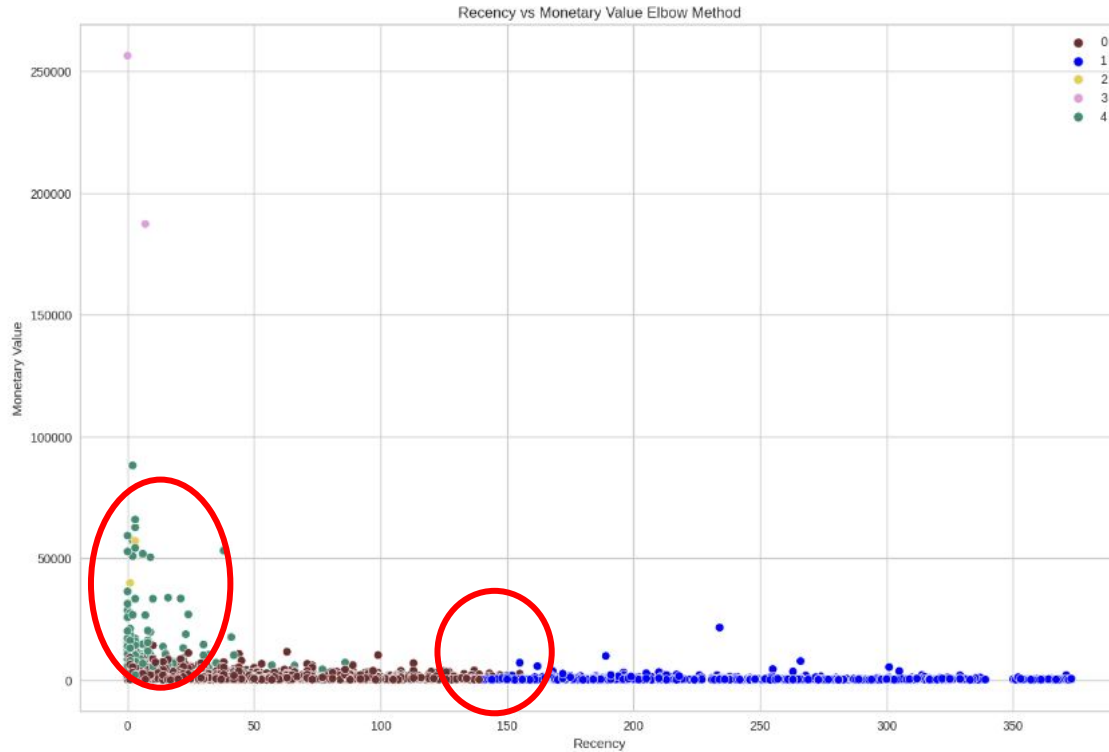
Recency vs Monetary Value Elbow Method

Cluster 0 ●
Cluster 1 ●
Cluster 2 ●
Cluster 3 ●
Cluster 4 ●

This diagram shows the distribution of customer data which is divided into five clusters according to the K-Means Clustering algorithm.

Visualization Clustering Frequency & Monetary Value using Scatter Plot

Cluster 0 ●
Cluster 1 ●
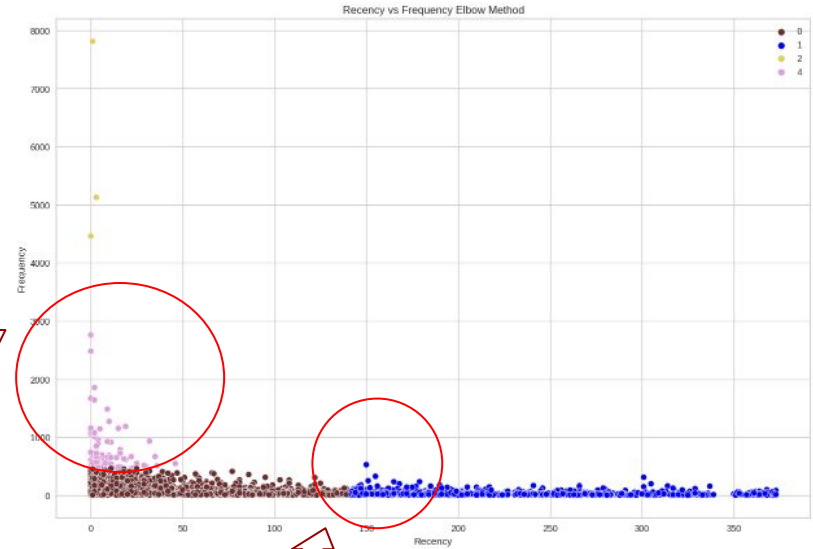Cluster 2 ●
Cluster 3 ●
Cluster 4 ●

This diagram shows the distribution of customer data which is divided into five clusters according to the K-Means Clustering algorithm.

After seeing the distribution of clustering using scatter plot, it can be seen that there are some range of distribution of clustering is not clear. Therefore, we need to reduce the cluster.
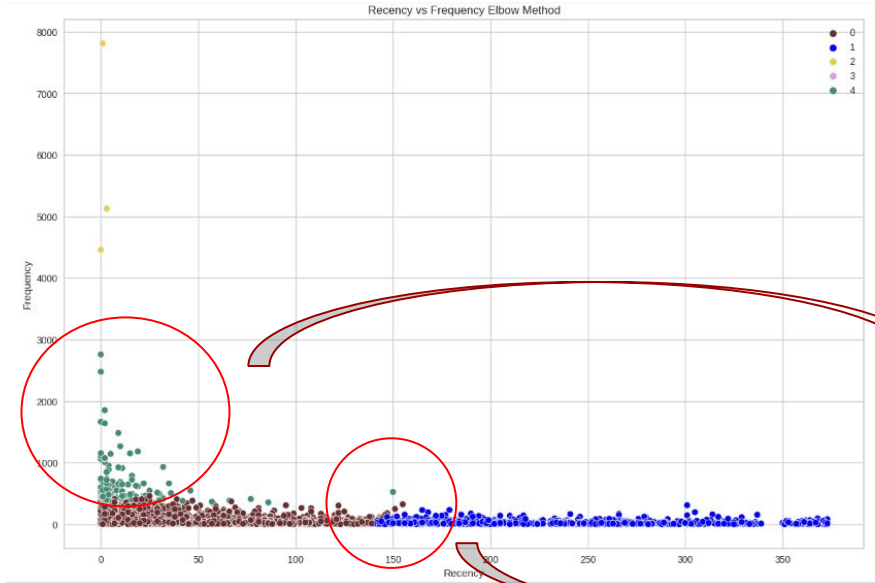
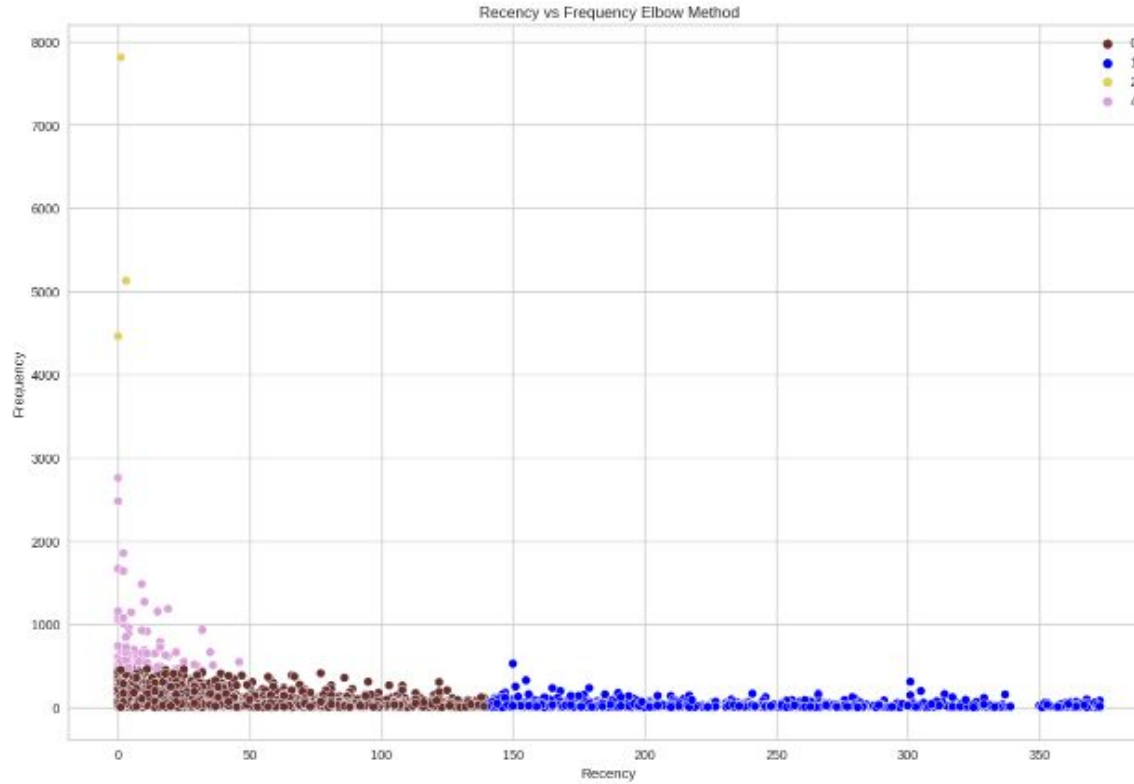**Before Reducing**

After Reducing

Recency & Frequency

**Visualization
After
Reducing Cluster**

Cluster 0 ●
Cluster 1 ●
Cluster 2 ●
Cluster 4 ●

Analyzing Clustering Based on Recency & Frequency
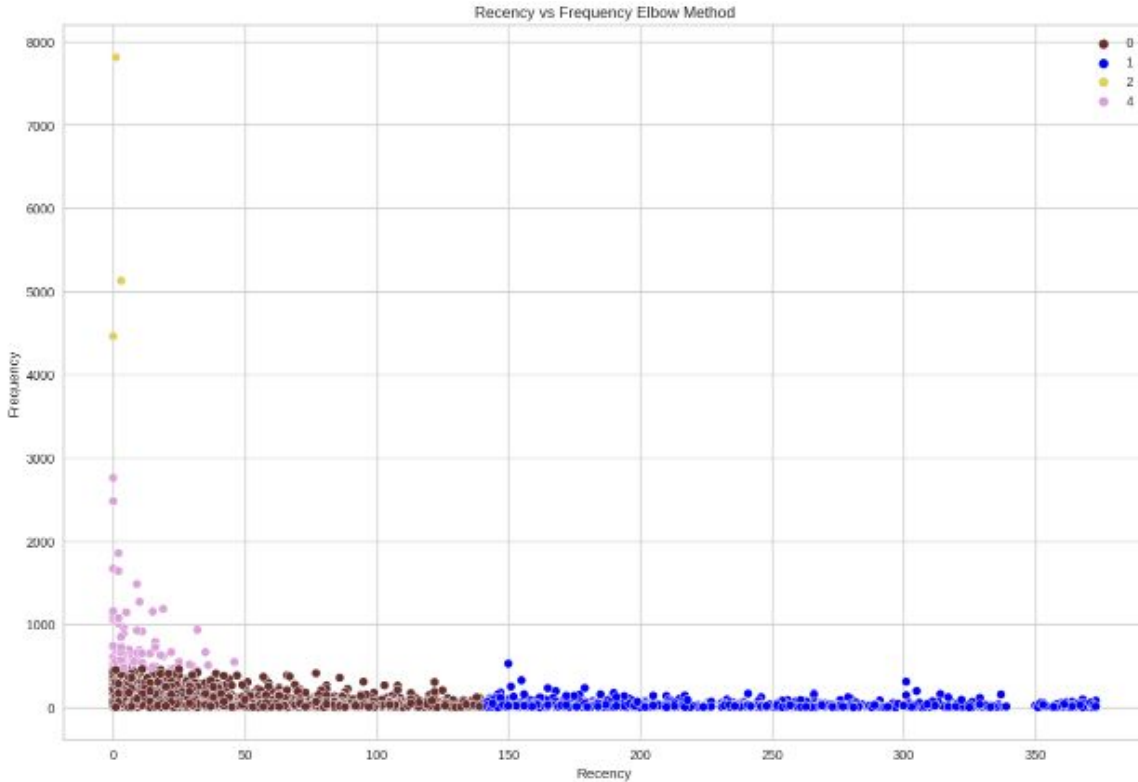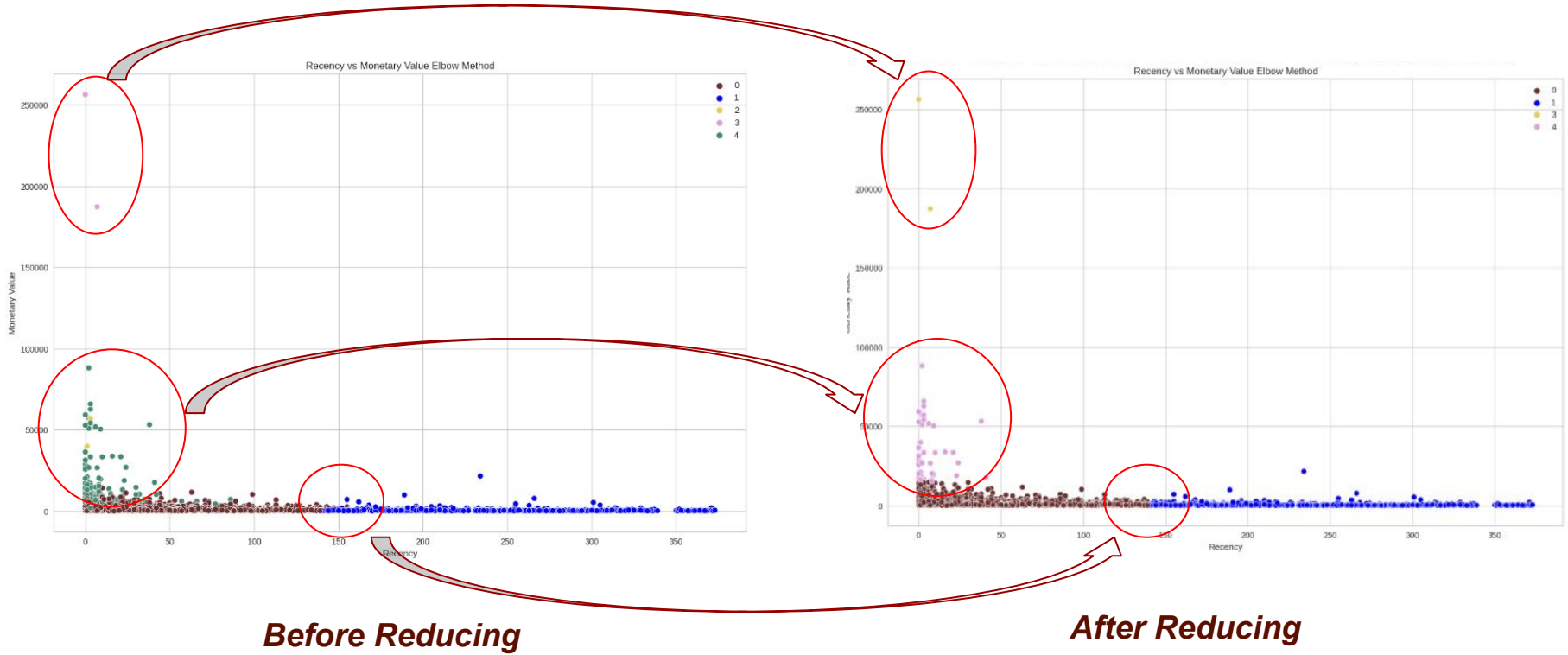

Recency vs Frequency Elbow Method

- **Cluster 0 (Customer Need Attention)**
  for customers in this group, most of them have a recency less than 150 days and the transaction frequency is not too high.
- **Cluster 1 (At Risk)**
  customers who have made transactions more than or equal to 140 days ago, and the frequency of transactions is also very small.
- **Cluster 2 (Champion)**
  customers who made their last transaction less than 50 days and have the highest transaction frequency
- **Cluster 4 (Loyal)**
  most of the customers in this group have the last transaction less than 100 days and the transaction frequency is quite high.
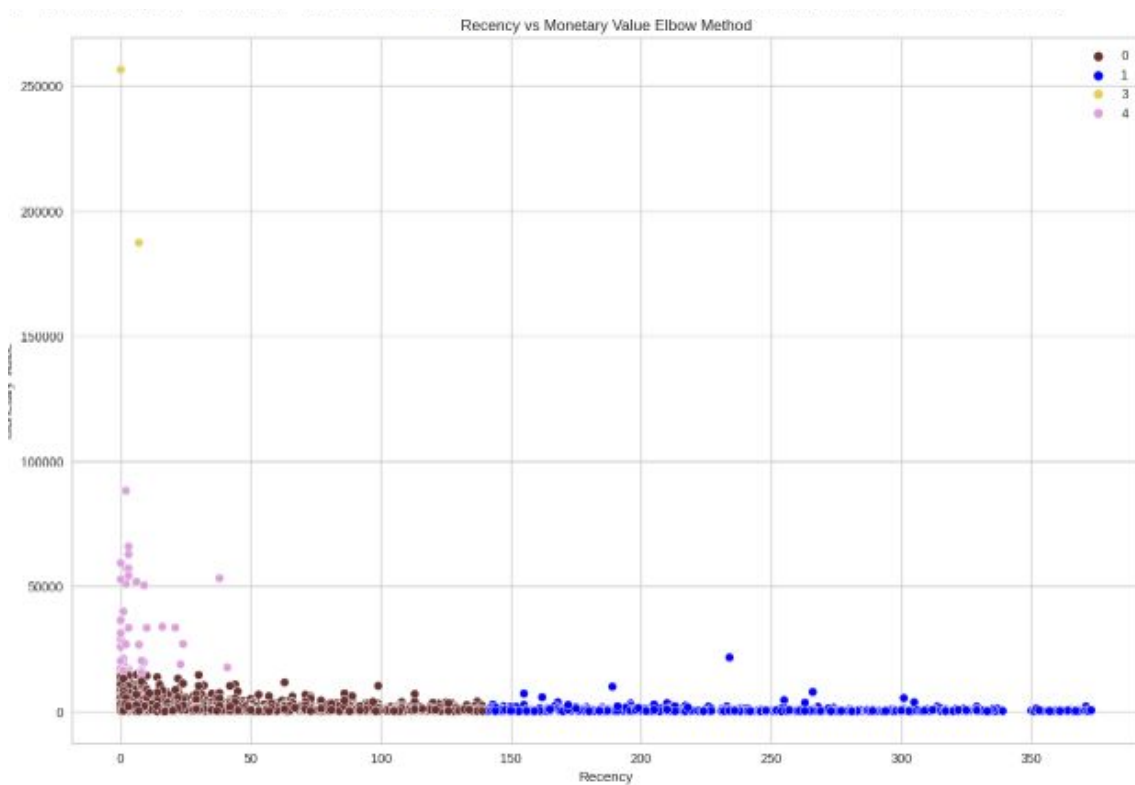
Percentage of Each Cluster for Recency & Frequency
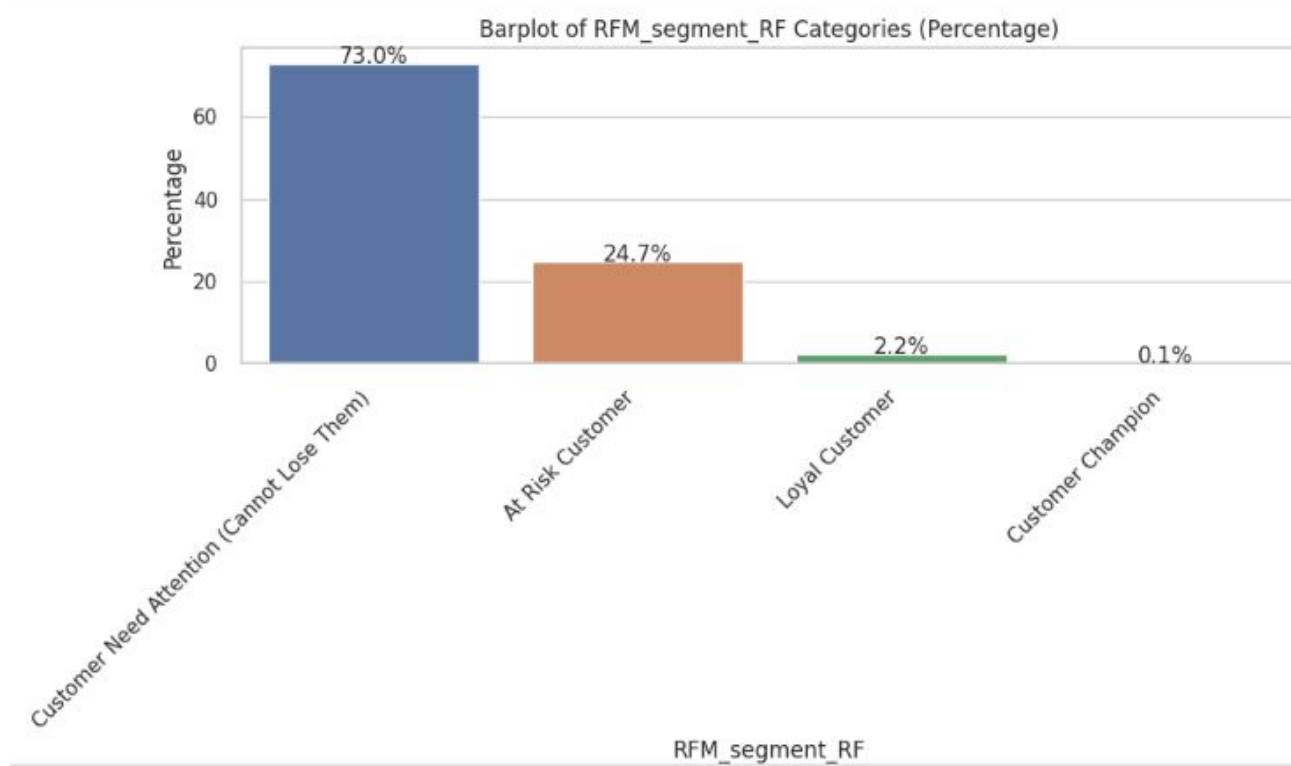
Barplot of RFM_segment_RF Categories (Percentage)

Before Reducing

After Reducing

Recency & Monetary Value

Recency vs Monetary Value Elbow Method

Visualization After Reducing Cluster

Cluster 0 🔴
Cluster 1 🔵
Cluster 3 🟡
Cluster 4 🔴

Analyzing Clustering Based on Recency & Monetary Value
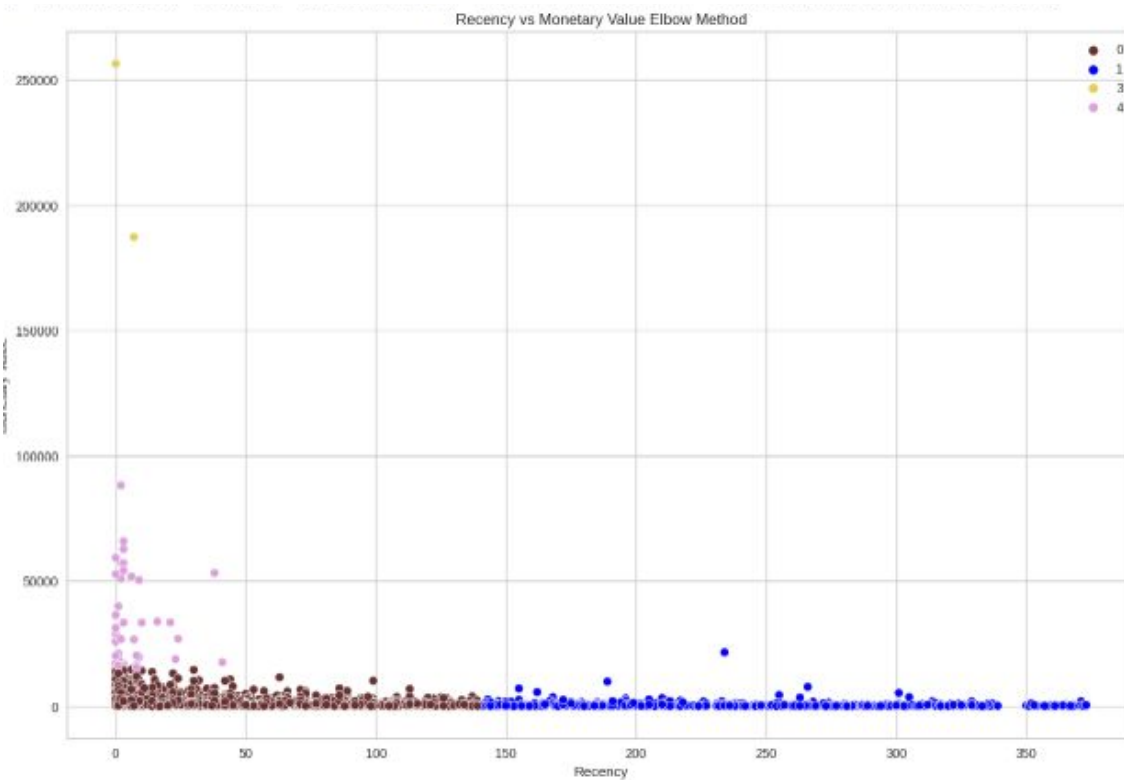

Recency vs Monetary Value Elbow Method

- **Cluster 0 (Customer Need Attention)**
  for customers in this group, most of them have a recency less than 150 days and spent money on the transaction is not too high.
- **Cluster 1 (At Risk)**
  customers who have made last transactions more than or equal to 140 days ago, and very little money spent.
- **Cluster 3 (Champion)**
  customers who made their last transaction less than 50 days and spent a lot of money on the transaction.
- **Cluster 4 (Loyal)**
  most of the customers in this group also have the last transaction less than 50 days and the money spent on the transaction is quite high.

Percentage of Each Cluster for Recency & Monetary Value
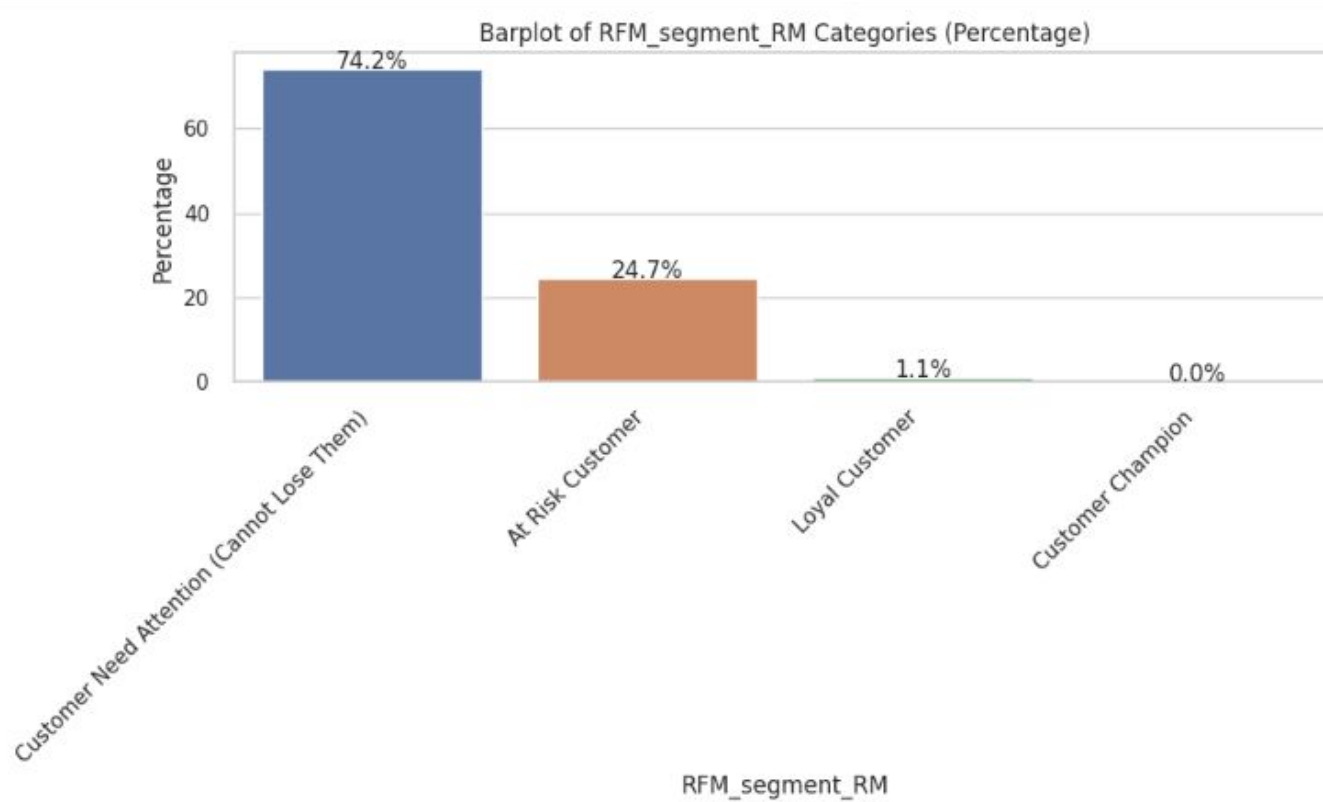
Barplot of RFM_segment_RM Categories (Percentage)
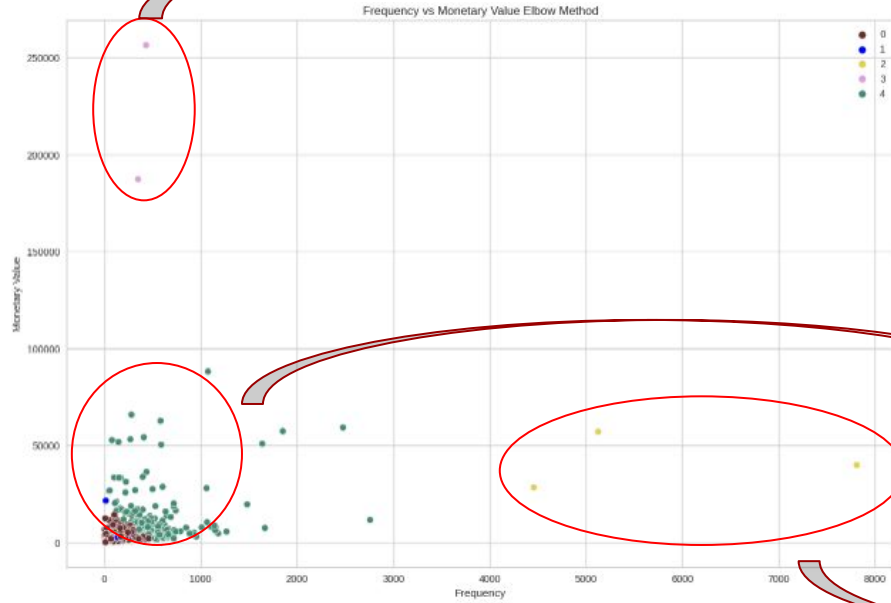
Before Reducing

After Reducing

Frequency & Monetary Value

Frequency vs Monetary Value Elbow Method

Visualization After Reducing Cluster
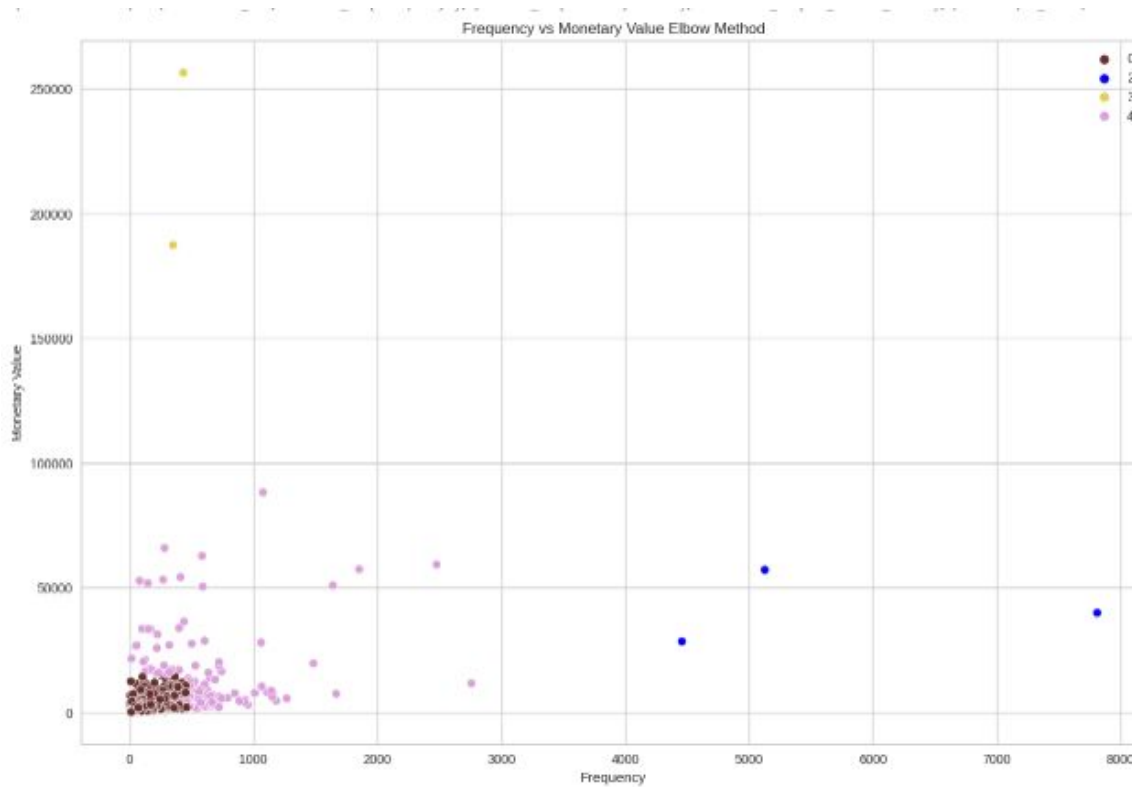
Cluster 0 🟤
Cluster 2 🔵
Cluster 3 🟡
Cluster 4 🔴

Analyzing Clustering Based on Frequency & Monetary Value

Frequency vs Monetary Value Elbow Method
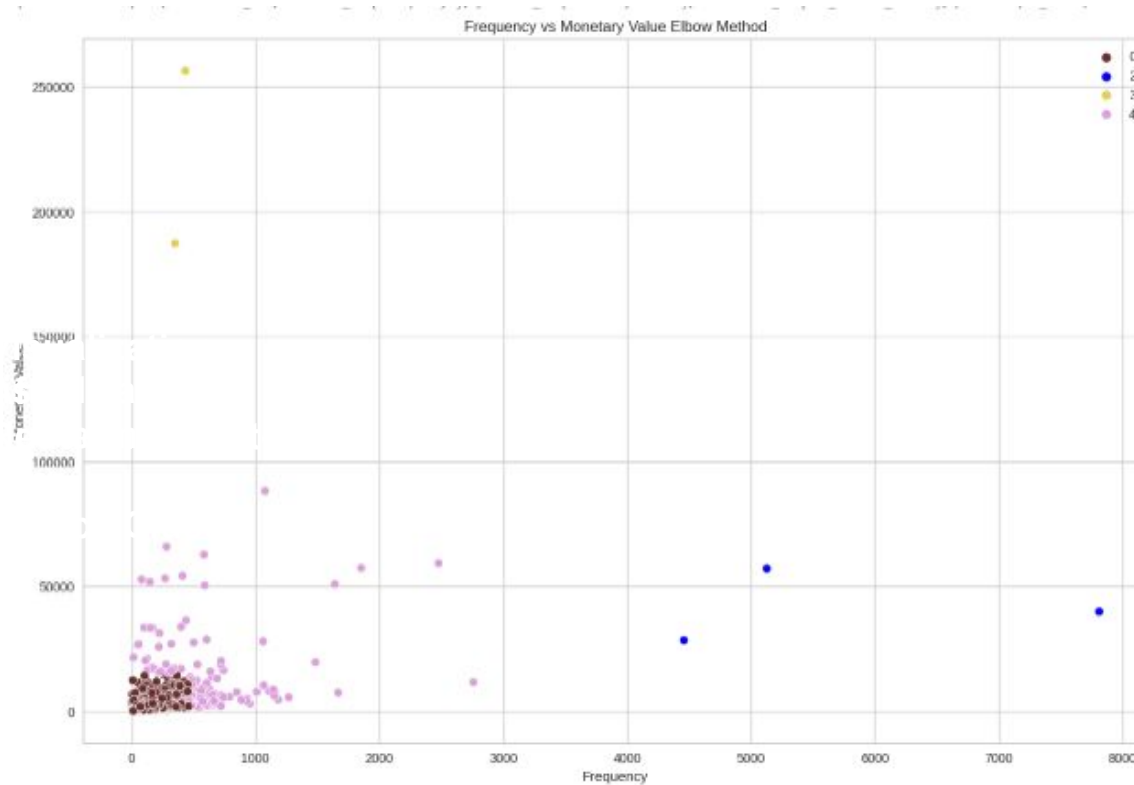
- **Cluster 0 (At Risk)** customers who have low frequency and monetary value.
- **Cluster 2 (Loyal)** customers who have a monetary value of less than $100000, but they are customers who make transactions most often.
- **Cluster 3 (Champion)** Although this customer on this grup have little frequency of transaction, they spend a lot of money.
- **Cluster 4 (Promising)** customers who have frequency of transactions and monetary value is quite high.

Percentage of Each Cluster for Frequency & Monetary Value

Barplot of RFM_segment_FM Categories (Percentage)
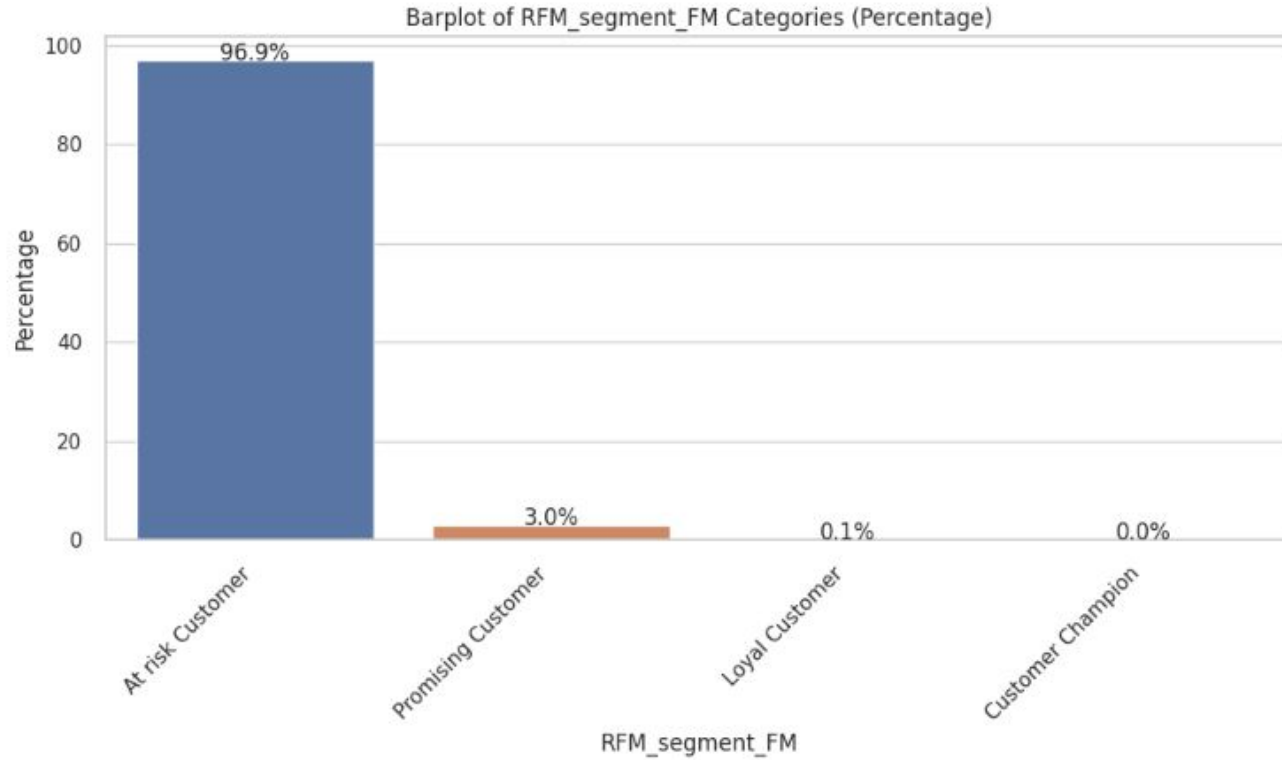
5 **Recommendation**

Recommendation for Clustering Based Recency & Frequency

Customer Need Attention	1. **Personalized Offers**: Provide special discounts or offers based on their transaction history to incentivize them to make more purchases. 2. **Exclusive Content**: Offer exclusive content, such as informative articles or how-to guides related to their past purchases, to keep them engaged with our brand.
Customer at Risk	1. **Reactivation Offers**: Send personalized reactivation offers with enticing discounts or incentives to encourage them to make a purchase. 2. **Win-Back Campaigns**: Create a targeted email or communication series to remind them of the benefits of our products/services and emphasize any new features or improvements.
Customer Champion	1. **VIP Treatment:** Recognize their loyalty with exclusive perks, early access to new products, or priority customer service. 2. **Cross-Selling/Upselling**: Recommend complementary products or upgrades based on their purchase history to increase their transaction value.
Customer Loyal	1. **Frequent Shopper Rewards**: Offer rewards or discounts for reaching certain transaction frequency milestones. 2. **Limited-Time Offers**: Create a sense of urgency with limited-time offers or flash sales to encourage more frequent purchases.

Recommendation for Clustering Based Recency & Monetary Value

Customer Need Attention	1. **Engagement Campaigns**: Send emails or notifications showcasing new products or updates to keep them engaged with your brand. 2. **Feedback Collection**: Reach out to gather feedback on their recent purchases and overall experience to identify areas for improvement.
Customer at Risk	1. **Re-engagement Content**: Send personalized content like product recommendations or success stories to remind them of the benefits of your products/services. 2. **Customer Surveys**: Understand why they haven't been active. Use surveys to gather insights into their reasons for disengagement and tailor your approach accordingly.
Customer Champion	1. **Referral Programs**: Encourage them to refer friends or colleagues, rewarding them for bringing in new business. 2. **Upselling Opportunities**: Present higher-tier products or complementary items they might be interested in based on their past purchases.
Customer Loyal	1. **Loyalty Rewards**: Implement a tiered loyalty program with rewards based on their spending levels, incentivizing even more purchases. 2. **Sneak Peeks**: Give them exclusive previews of upcoming products or services to make them feel valued and engaged.

Recommendation for Clustering Based Frequency & Monetary Value

Customer at Risk	1**. Reactivation Campaigns**: Create targeted reactivation campaigns that emphasize the benefits and value of our products or services. Offer a time-limited discount or promotion to incentivize a new purchase. 2**. Engagement Surveys:** Send out surveys to understand why these customers have low engagement. Gather feedback on their preferences and pain points to tailor your offerings and communications more effectively.
Customer Loyal	1**. Frequent Shopper Discounts**: Offer periodic discounts or rewards specifically for customers who shop often, incentivizing them to spend more during each transaction. 2**. Bundle Deals**: Create bundles of related products or services that encourage customers to spend a bit more to access a better deal.
Customer Champion	1**. Luxury Experience**: Offer a personalized luxury experience for these high-spending customers. Provide concierge services, premium packaging, or expedited shipping to enhance their overall experience. 2. **High-Value Upselling:** Identify premium products or services that align with their spending patterns. Offer tailored upsell recommendations that enhance their purchases and align with their preferences.
Customer Promising	1. **Subscription Models**: Introduce subscription-based models for products or services they frequently purchase to encourage consistent spending. 2. **Customization Options**: Allow these customers to personalize their purchases or experiences, increasing their attachment to our brand.

Thank You!!

Feel Free to Contact me!

https://wa.me/082393617436

iskaoktafauziah03@gmail.com

https://github.com/IskaOktaFauziah03

https://medium.com/@iskaoktafauziah5

www.linkedin.com/in/iska-okta-fauziah