PROJECT – 1 Report
Dynamic Customer Segmentation

SUBMITTED IN THE PARTIAL FULFILLMENT REQUIREMENT

FOR THE AWARD OF DEGREE OF

Bachelor of Technology

(COMPUTER SCIENCE and ENGINEERING)

SUBMITTED BY –

Group 1

GAGAN SINGHAL (220461)

AYUSHI SINGH (220462)

YOGESH SHARMA (220465)

ANUSHA SINGH (220641)

UNDER THE SUPERVISION OF

DR. ATUL MISHRA

SCHOOL OF ENGINEERING AND TECHNOLOGY

BML MUNJAL UNIVERSITY™

FROM HERE TO THE WORLD

BML MUNJAL UNIVERSITY

Gurugram, Haryana - 122413

Dec 2023

CANDIDATE'S DECLARATION

I hereby certify that I have worked on project entitled, "**Dynamic Customer Segmentation**", in partial fulfillment of requirements for the award of Degree of **Bachelor of Technology** in Computer Science and Engineering at **BML Munjal University** is an authentic record of our own work carried out during a period from August 2023 to December 2023 under the supervision of Dr. Atul Mishra.

Gagan Singhal (220461)
Ayushi Singh (220462)
Yogesh Sharma (220465)
Anusha Singh (220641)

Date –
Place – BML Munjal University

This is to certify that the above statement made by the candidate is correct to the bestof my knowledge.

Dr. Atul Mishra
Assistant Professor

ABSTRACT

This project introduces a strong data-driven methodology for improving customer understanding, segmentation, and revenue forecasting. Four related datasets are carefully cleaned and integrated as the primary focus, guaranteeing data quality and consistency for further analysis. We examine the subtleties of consumer behavior, uncovering patterns and insights by utilizing the potent RFM analysis (Recency, Frequency, Monetary Value).

In order to overcome the difficulties presented by interconnected datasets, our approach is combining data cleaning, RFM analysis, and sophisticated modeling techniques. The comprehensive methodology's results revealing three distinct segments with significantly different purchase patterns and revenue potential, each with its own potential for tailored product recommendations, focused marketing campaigns, and resource optimization.

The project has a broad impact since it gives businesses the tools, they need to make wise decisions that will increase revenue and optimize customer engagement. This work advances customer-centric business strategies and offers practical insights for immediate implementation by identifying three distinct customer groups with greatly different purchase patterns and revenue potential.

This project offers a data-driven foundation for strategic decision-making and sustainable growth, acting as a road map for companies looking to fully utilize their customer data.

ACKNOWLEDGEMENT

I am highly grateful to **Dr. Atul Mishra, Professor**, BML Munjal University, Gurugram, for providing supervision to carry out the seminar/case study from August-December 2023.

Dr. Atul has provided great help in carrying out our work and is acknowledged with reverential thanks. Without the wise counsel and able guidance, it would have been impossible to complete the training in this manner.

I would like to express thanks profusely to thank **Dr. Atul Mishra**, for stimulating us time to time. I would also like to thank entire team of BML Munjal University. I would also thank our team who devoted their valuable time and helped me in all possible waystowards successful completion.

Gagan Singhal
Ayushi Singh
Yogesh Sharma
Anusha Singh

LIST OF FIGURES

TABLE OF CONTENTS

Chapter 01
Introduction to Organization

Sprocket Central Pty Ltd specializes in high-quality bikes and user-friendly cycling accessories. Their marketing team desires to increase sales by analyzing their existing customer dataset to identify customer trends and behavior. Customer demographics, transactions, a new customer list, and customer address were all labeled datasets.

KPMG Analytics is the data and analytics arm of KPMG, one of the Big Four professional services firms. They offer a wide range of data and analytics services to help businesses gain insights from their data, improve decision-making, and achieve their goals.

Sprocket Central Pty Ltd has approached KPMG's Team for business advice. Sprocket Central Pty Ltd is interested in learning more about KPMG's Analytics, Information & Modelling team's expertise.

Sprocket Central Pty Ltd, in particular, requires assistance with its customer and transaction data. The company has a large dataset containing information about its customers, but its team is unsure how to effectively analyze it in order to help optimize its marketing strategy.

Chapter 02
Introduction to Project

2.1 Overview

Sprocket Central's data analytics project, fueled by RFM analysis and K-means clustering, holds immense potential for optimizing marketing strategies. To unlock this potential, building and training a robust machine learning model using the Random Forest Regressor and the Elbow Method is crucial.

The first step involves cleaning and validating the extensive three dataset. This includes scrubbing for missing values, inconsistencies, and formatting issues. Techniques like dropping records with null values, a thorough examination for duplicate records was performed and outlier removal ensure data quality and pave the way for accurate analysis.

The Recency, Frequency, and Monetary (RFM) analysis was employed to categorize customers into distinct segments, revealing patterns that can guide marketing and business strategies. The dataset was further enriched by incorporating customer age groups and developing a detailed customer tag based on RFM scores. The resulting dataset was exported for further exploratory data analysis (EDA), paving the way for a comprehensive understanding of customer segments. Finally, the distribution of customers across different segments offers a comprehensive overview, emphasizing the prevalence of certain customer categories.

Next comes feature engineering. We created new features based on existing ones, capturing hidden relationships and enhancing model performance. Scaling, normalization, and feature interaction analysis are valuable tools in this process. Finally, we split the prepped data into training and testing sets, ensuring the model learns from unseen data during evaluation.

Now, we unleash the power of the Random Forest Regressor. We choose key hyperparameters like the number of trees, maximum depth, and minimum leaf size. Experimenting with different combinations using grid search or random search helps identify the optimal settings that minimize the chosen error metric, such as mean squared error. With the chosen hyperparameters, the model trains on the training data, learning the intricate patterns within it.

Subsequently, we employed clustering techniques to categorize customers into distinct segments based on their relevant characteristics. This segmentation provides a nuanced view of customer groups, enabling targeted strategies for customer engagement and marketing efforts.
With the optimal number of clusters identified, we can assign each data point to its corresponding cluster, segmenting customers based on their characteristics and behavior. This segmentation forms the basis for targeted marketing strategies. We can analyze the behavior and preferences within each segment, tailoring marketing campaigns to resonate deeply with each group.

2.2 Existing System

The current approach of the project lacks the depth required for optimal results. To enhance the analysis, we should apply the advanced techniques used in the models developed in previous phases. In the Data Exploration phase, we can leverage the insights gained from the Random Forest Regressor model, extending beyond conventional exploratory data analysis. By incorporating feature importance's from the model, the intern can prioritize the most influential variables for a more nuanced understanding of customer behavior.

Moving into Model Development, we should apply the dynamic data transformations and ensemble learning techniques used in the Random Forest model to enhance the predictive accuracy. This holistic approach connects the model outputs to actionable marketing strategies, fostering a more comprehensive and impactful solution for driving value in targeted marketing efforts.

2.3 User Requirement Analysis

In our initial review of the datasets provided by Sprocket Central Pty Ltd for the customer targeting project, we have identified several data quality issues that warrant attention for a more robust analysis.
Firstly, we observed missing values across key fields, such as 'Age' and 'Gender' in the customer demographic dataset. Additionally, inconsistencies in data types, duplicate entries in the transaction's dataset, and the presence of outliers in numerical fields were noted. Addressing these issues is imperative to ensure the accuracy and reliability of the subsequent analysis.
Moving forward, our team is committed to resolving these data quality concerns in preparation for the three phases outlined in our project approach: Data Exploration, Model Development, and Interpretation. Our objective is to collaboratively refine and enhance the datasets to extract meaningful insights, ultimately facilitating an effective customer targeting strategy.

2.4 Feasibility Study

The feasibility study for the customer targeting project at Sprocket Central Pty Ltd will concentrate on ensuring the technical and operational viability of the proposed approach across three pivotal phases: Data Exploration, Model Development, and Interpretation. Technically, we will assess the existing infrastructure's readiness for advanced analytics and machine learning applications, ensuring seamless integration during the Data Exploration phase. Operationally, our focus will be on aligning the project with the day-to-day operations of the marketing department, evaluating workflow implications, and identifying training needs for staff during Model Development.

Chapter 03
Literature Review

3.1 Comparison

The current project approach acknowledges a need for depth and optimization. In comparison with existing models and exploratory data analysis, there are notable similarities and areas for alignment. Both emphasize the utilization of advanced techniques, such as the Random Forest Regressor model, to extract meaningful insights during the Data Exploration phase. Our exploratory data analysis likely aligns with the suggestion to extend beyond conventional methods, seeking deeper insights.

Moreover, the recommendation to incorporate feature importance from the Random Forest Regressor model closely aligns with a data-driven approach that aims to prioritize influential variables. This reinforces the importance of leveraging key insights gained from your models to guide further exploration.

In addition to the outlined approaches, we are incorporating RFM (Recency, Frequency, Monetary) analysis to further enhance customer segmentation and decipher purchase patterns. This additional technique adds depth to our analysis, allowing us to understand not only the historical spending behavior of customers but also their recency and frequency of transactions. By combining RFM analysis with the proposed regression and clustering models, we aim to capture a comprehensive view of customer segments, considering both their historical purchasing patterns and predictive features.

The project will involve developing three distinct models. The first model, a regression model, leverages a broad spectrum of features from all available tables to predict total revenue or customer count of purchases. Building on insights from this regression model, the second model involves clustering, focusing on high-importance features identified in the initial regression model. This clustering model aims to uncover distinct customer segments based on shared characteristics, facilitating more targeted marketing efforts.

Furthermore, the third model is designed to predict total revenue or customer count of purchases specifically using features available in the new customer list. This targeted regression model serves the strategic purpose of prioritizing customer targeting for the newly acquired customer base. Through the integration of RFM analysis and the delineated modeling strategies, we collectively cover customer segmentation and purchase pattern analysis, fostering a more robust understanding of customer behavior for effective marketing optimization.

3.2 Objectives of the Project

The project's primary objective is to optimize targeted marketing strategies for Sprocket Central Pty Ltd through a data-driven approach. Specifically, the project aims to achieve the following:

3.2.1. Enhanced Revenue Prediction: Develop a robust regression model leveraging a comprehensive set of features from diverse datasets to accurately predict total revenue or customer count of purchases. This model will serve as a foundational tool for understanding and forecasting the overall financial impact.

3.2.2. Customer Segmentation and Clustering: Utilize clustering techniques, informed by feature importance from the regression model, to identify distinct customer segments. This will enable a more granular understanding of customer behavior and preferences, facilitating targeted marketing efforts tailored to specific segments.

3.2.3. New Customer Prioritization: Develop a targeted regression model using features exclusive to the new customer list. This model will prioritize targeting strategies for the newly acquired customer base, ensuring focused efforts on high-potential customers.

3.2.4. Integration of RFM Analysis: Incorporate RFM (Recency, Frequency, Monetary) analysis to augment customer segmentation and illuminate purchase patterns. The integration of RFM will provide valuable insights into historical customer behavior, enhancing the overall understanding of the customer base.

3.2.5. Actionable Insights for Marketing Optimization: Connect the outputs of the models to actionable marketing dashboard. The ultimate goal is to provide Sprocket Central Pty Ltd with clear, data-driven insights for optimizing marketing initiatives, resource allocation, and customer engagement.

These defined objectives collectively form the foundation for a comprehensive and impactful project, aligning with Sprocket Central Pty Ltd's strategic goals of enhancing customer targeting and maximizing revenue.

Chapter 04
Exploratory Data Analysis

4.1 Dataset

4.1.1 CustomerDemographic

This dataset encompasses a record of past three years, encompassing a wide array of customer particulars, transaction types, professional backgrounds, car ownership status, gender, and various other pertinent attributes. The dataset is rich in information and can be instrumental for in-depth analyses.

4.1.2 Transactions

This dataset contains comprehensive transaction information, including payment type, order status, product brands, purchase costs, and other relevant details, linked to each customer via their unique customer ID and transaction ID. It is the biggest and major analytics value holder.

4.1.3 NewCustomerList

This dataset contains information on new customers. It includes essential attributes like names, gender, bike-related purchase history, date of birth, job details, wealth segment, car ownership, tenure, address, property valuation, rank and value information.

4.1.4 CustomerAddress

This dataset includes customer identifiers (customer_id), address details, postal codes, state information, country location, and property valuation data. The dataset's geographic and property-related attributes make it particularly suitable for conducting customer segmentation analysis.

4.2 Exploratory Data Analysis and Visualisations

4.2.1 Data Cleaning

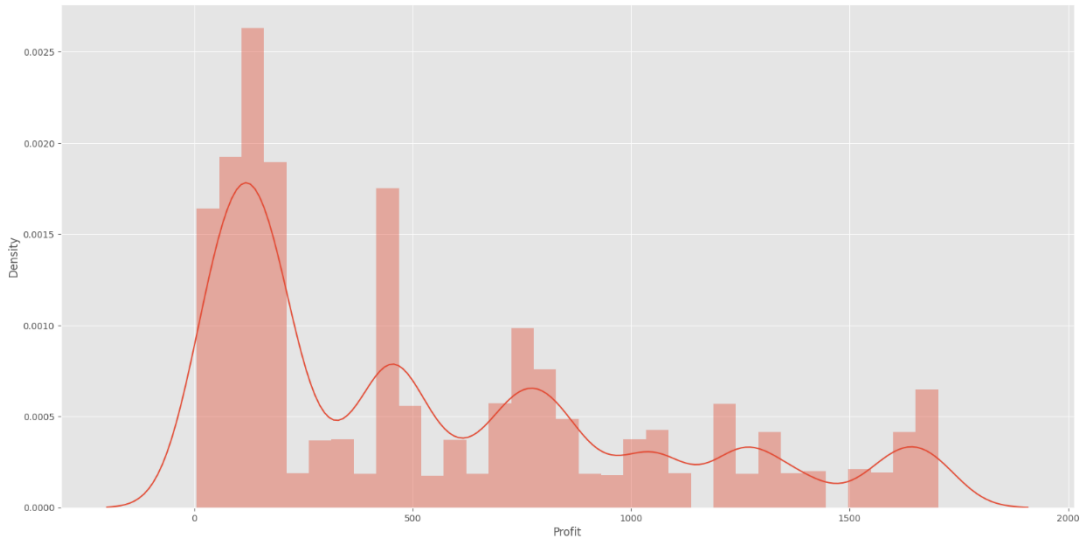

Figure 4.2.1 A

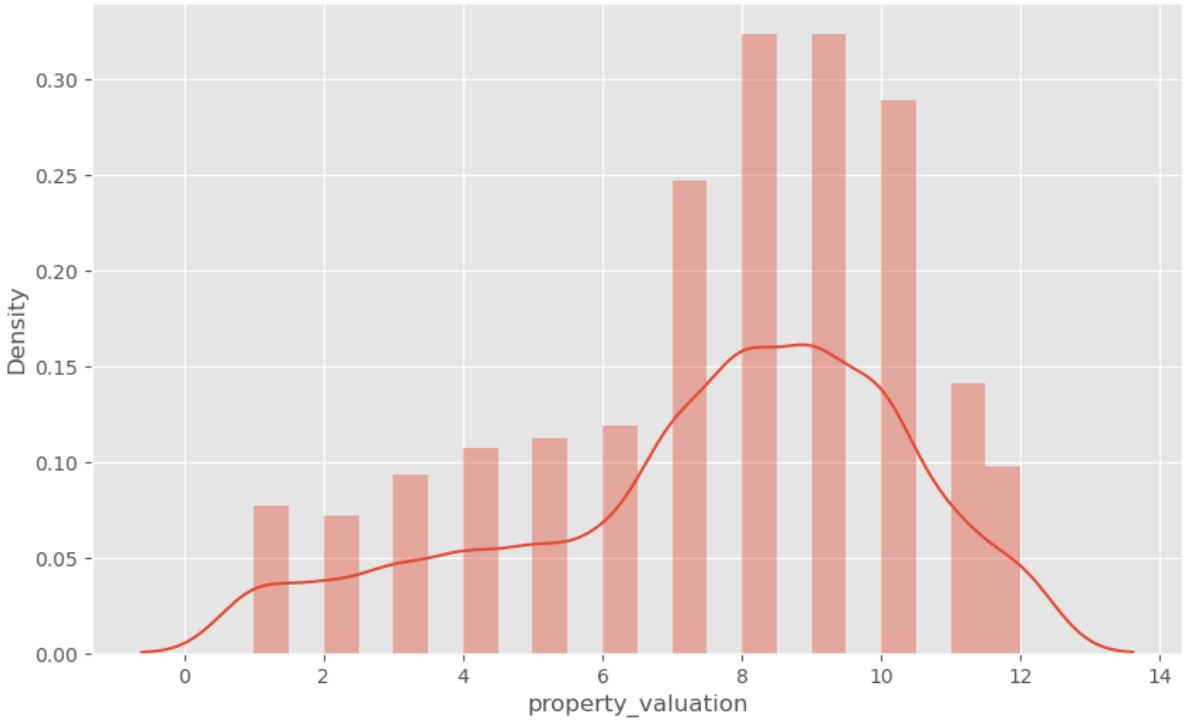

Figure 4.2.1 B

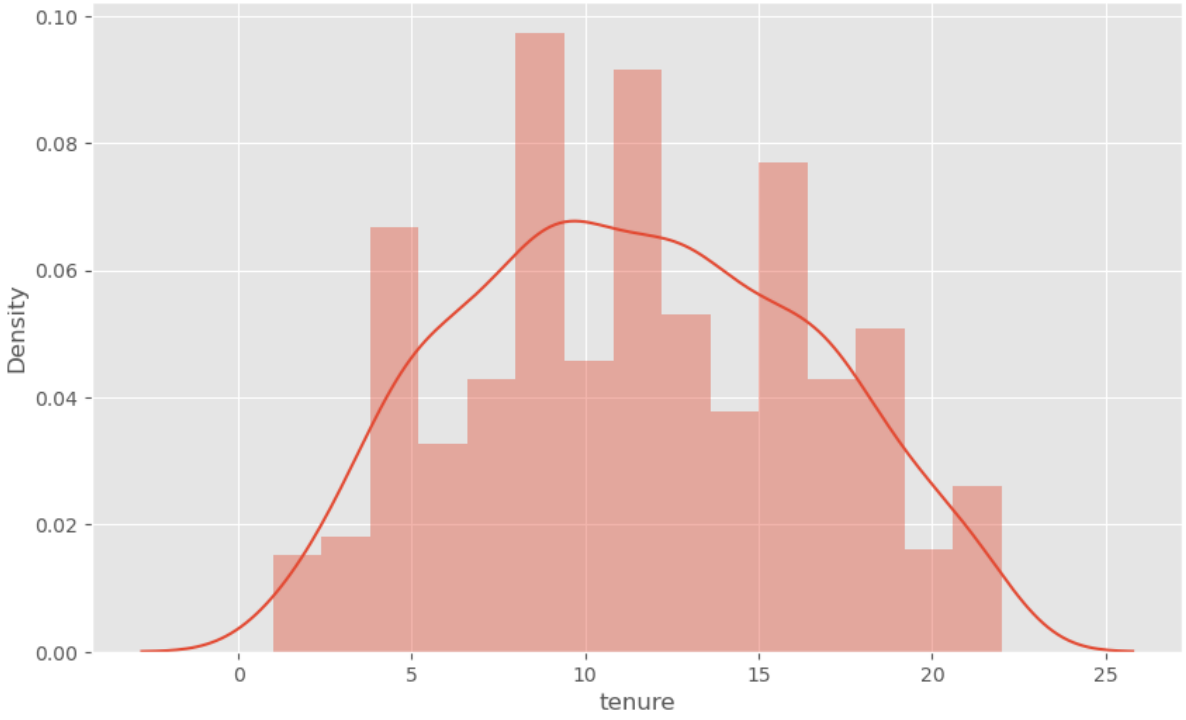

Figure 4.2.1 C

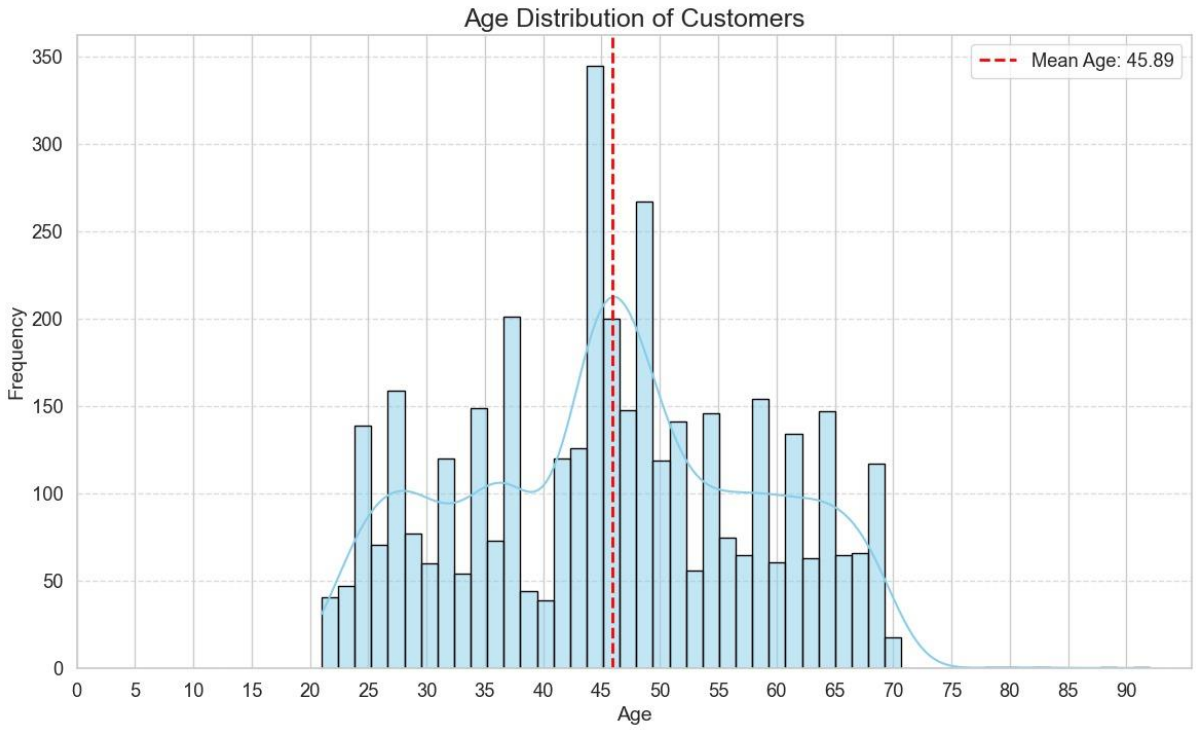

Figure 4.2.1 D

4.2.2 RFM Analysis

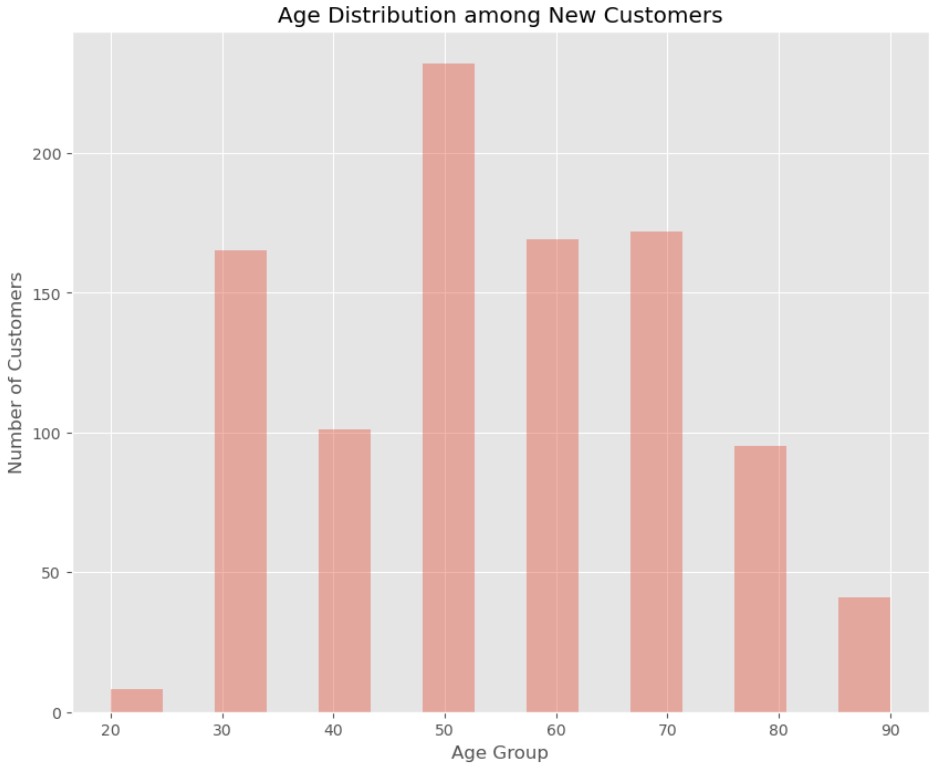

Figure 4.2.2 A Here 20 defines as age under 20, 30 defines as age bracket between 20-29

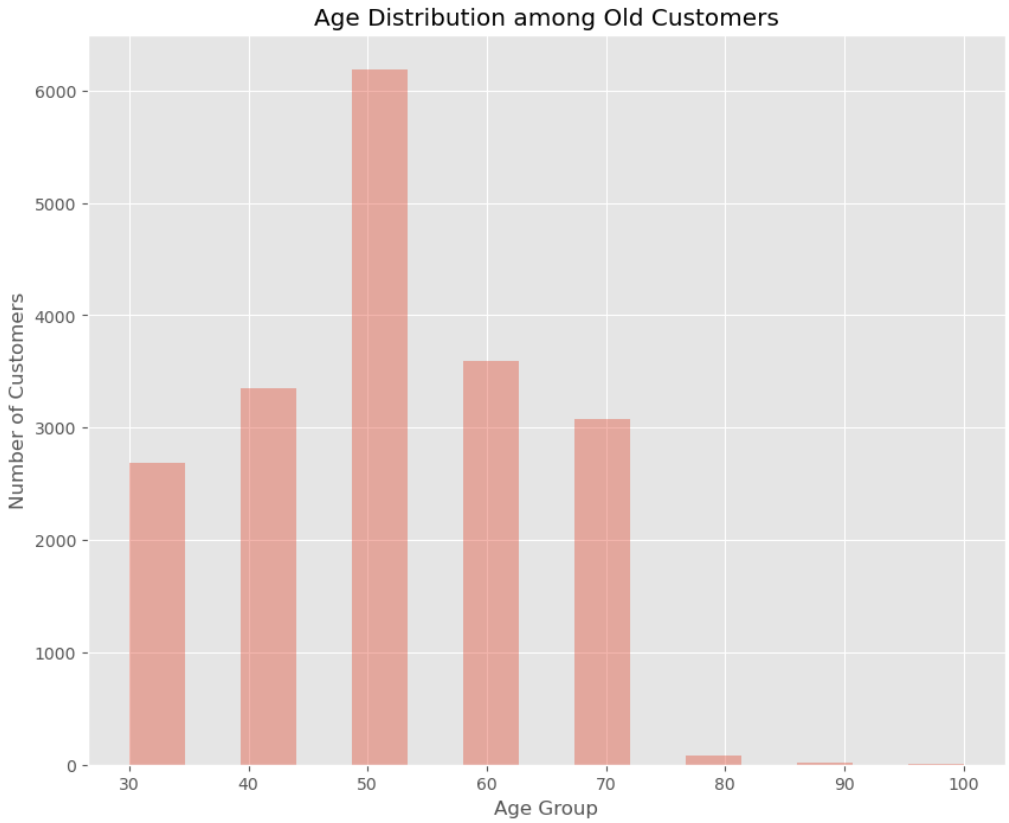

Figure 4.2.2 B Here 20 defines as age under 20, 30 defines as age bracket between 20-29

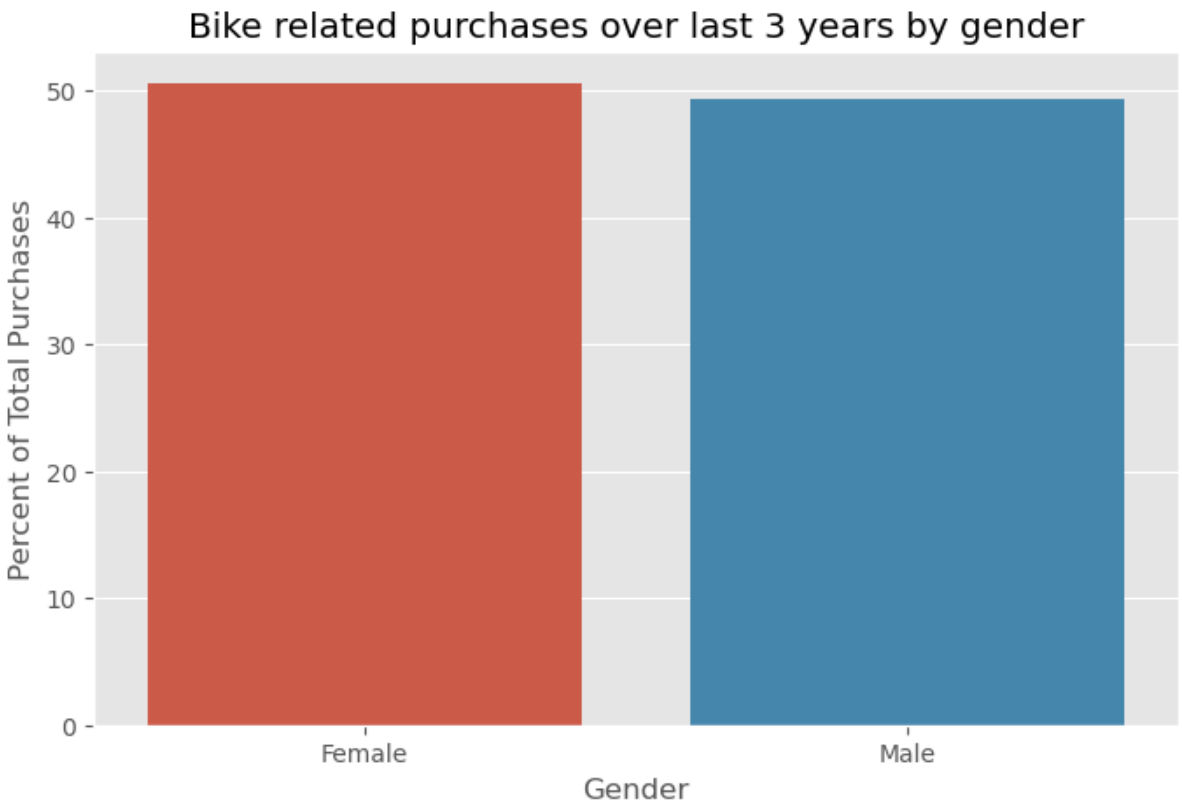

Figure 4.2.2 C Here Female encompasses 50.6% purchases and 49.4% purchases

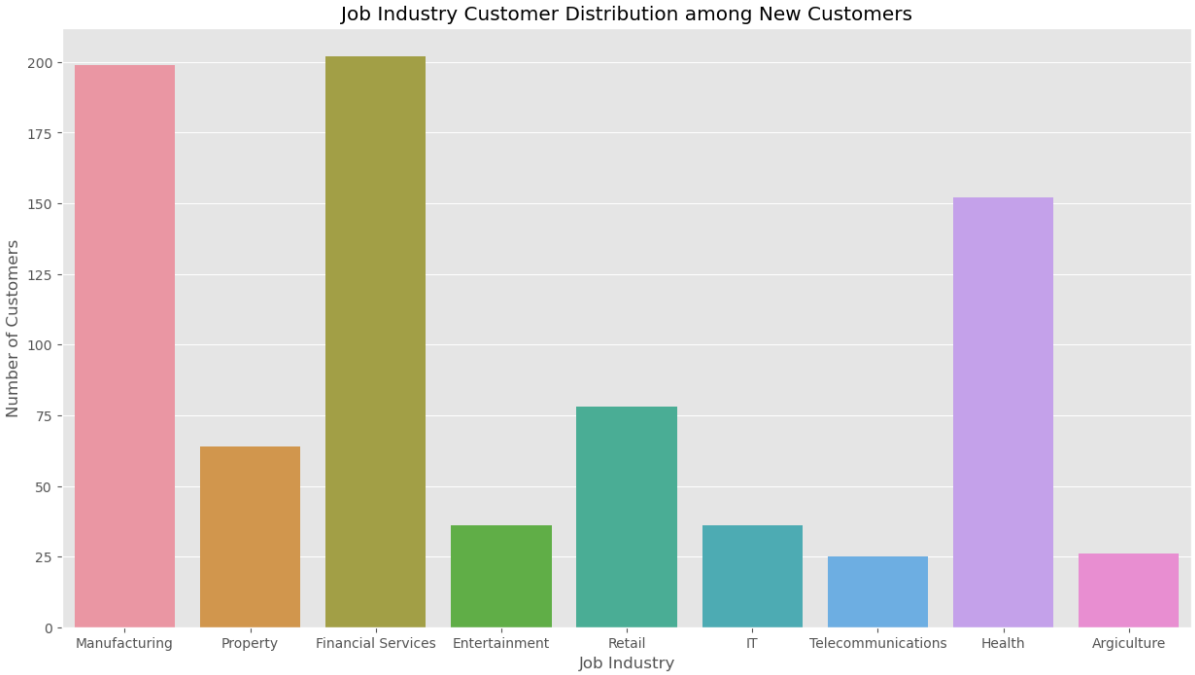

Figure 4.2.2 D

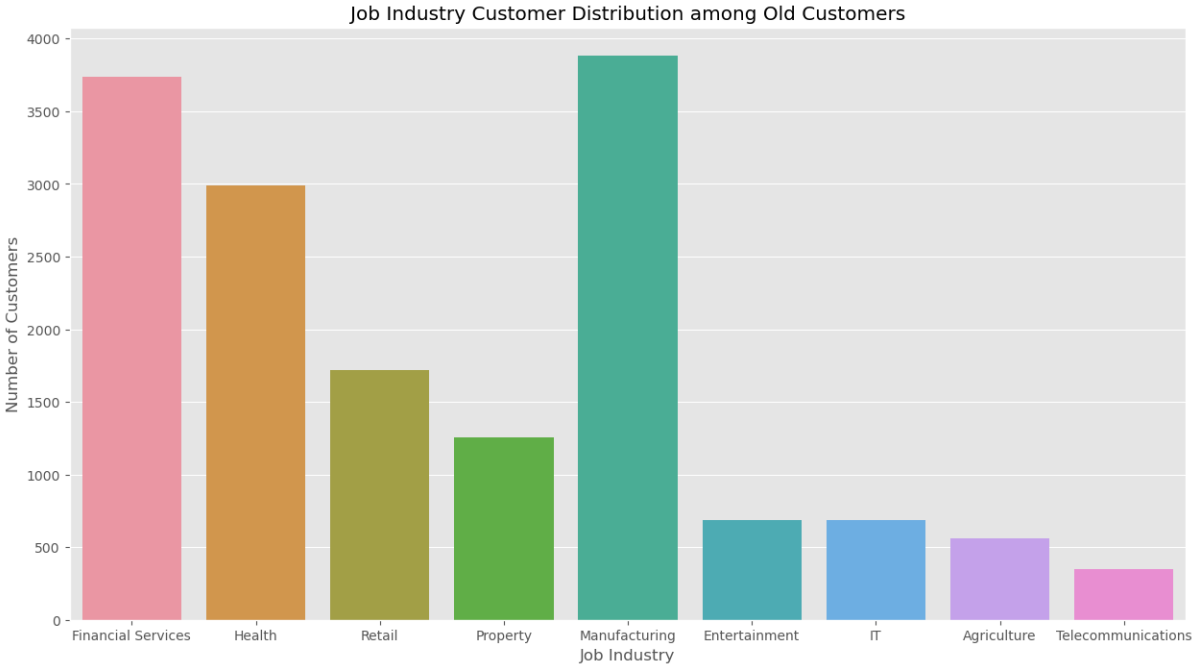

Figure 4.2.2 E

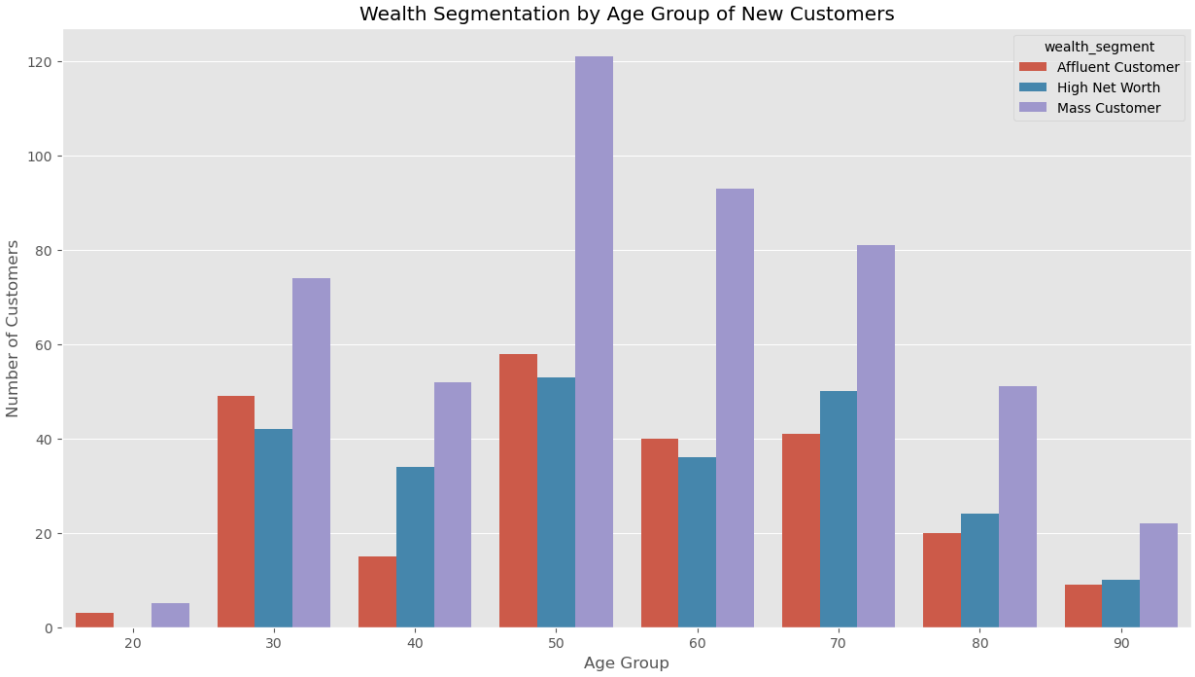

Figure 4.2.2 F In all the age groups the greatest number of customers are 'Mass Customers'. The 2nd largest customer base being the 'High Net Worth' group. In the age group 40-49 the 'Affluent Customer' group outperforms 'High Net Worth' group

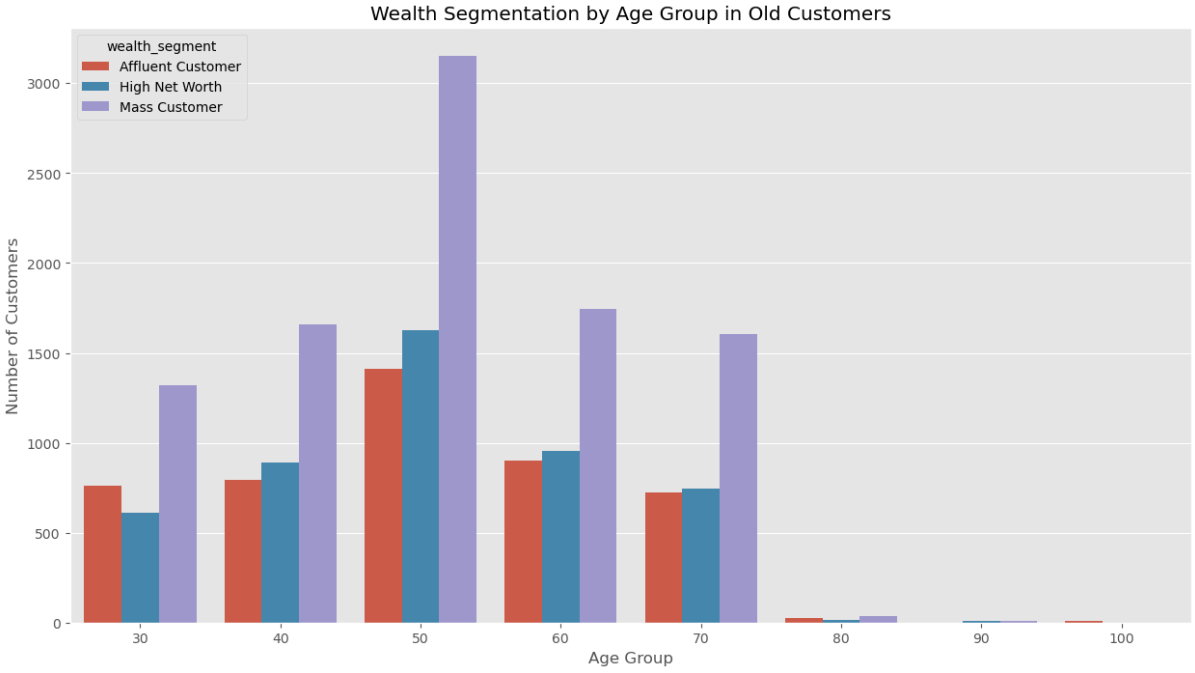

Figure 4.2.2 G Similar trend (like that of New Customers) is observed among Old Customers

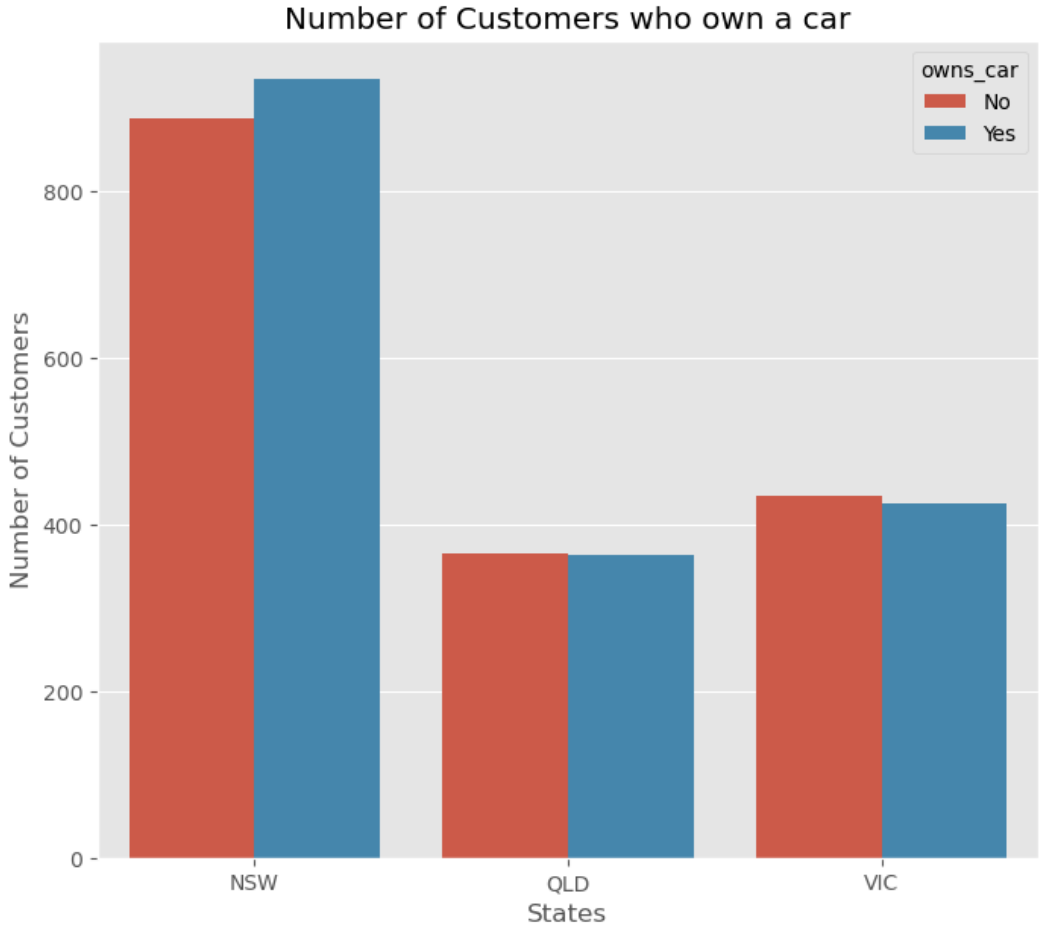

Figure 4.2.2 H NSW has the largest number of people that do not own a car. It seems that a higher amount of data is collected from NSW compared to other states. In QLD the distribution between customers having a car or not having is even. In Victoria the number is split evenly. Both the numbers are significantly lower than that of NSW.

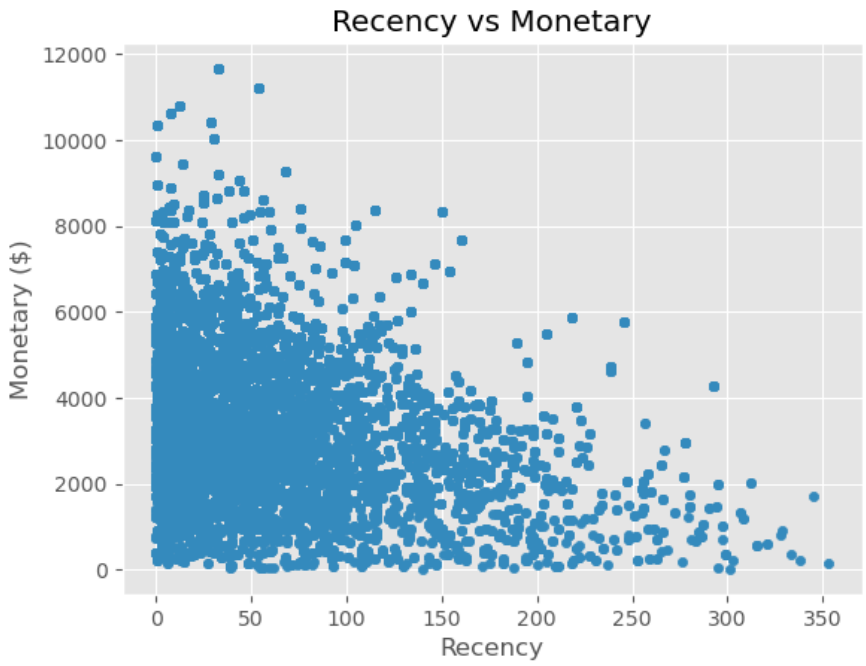

Figure 4.2.2 I

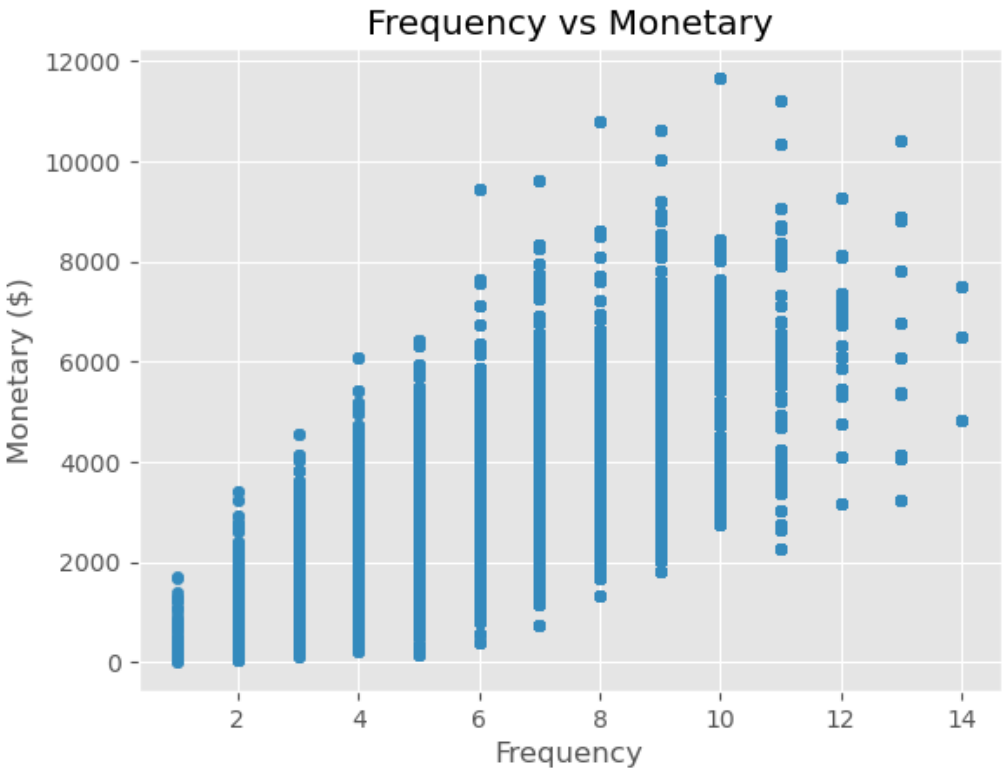
Frequency vs Monetary
Monetary ($)
Frequency
12000
10000
8000
6000
4000
2000
0
2
4
6
8
10
12
14

Figure 4.2.2 J

4.2.3 Model Deployment & Further Analysis

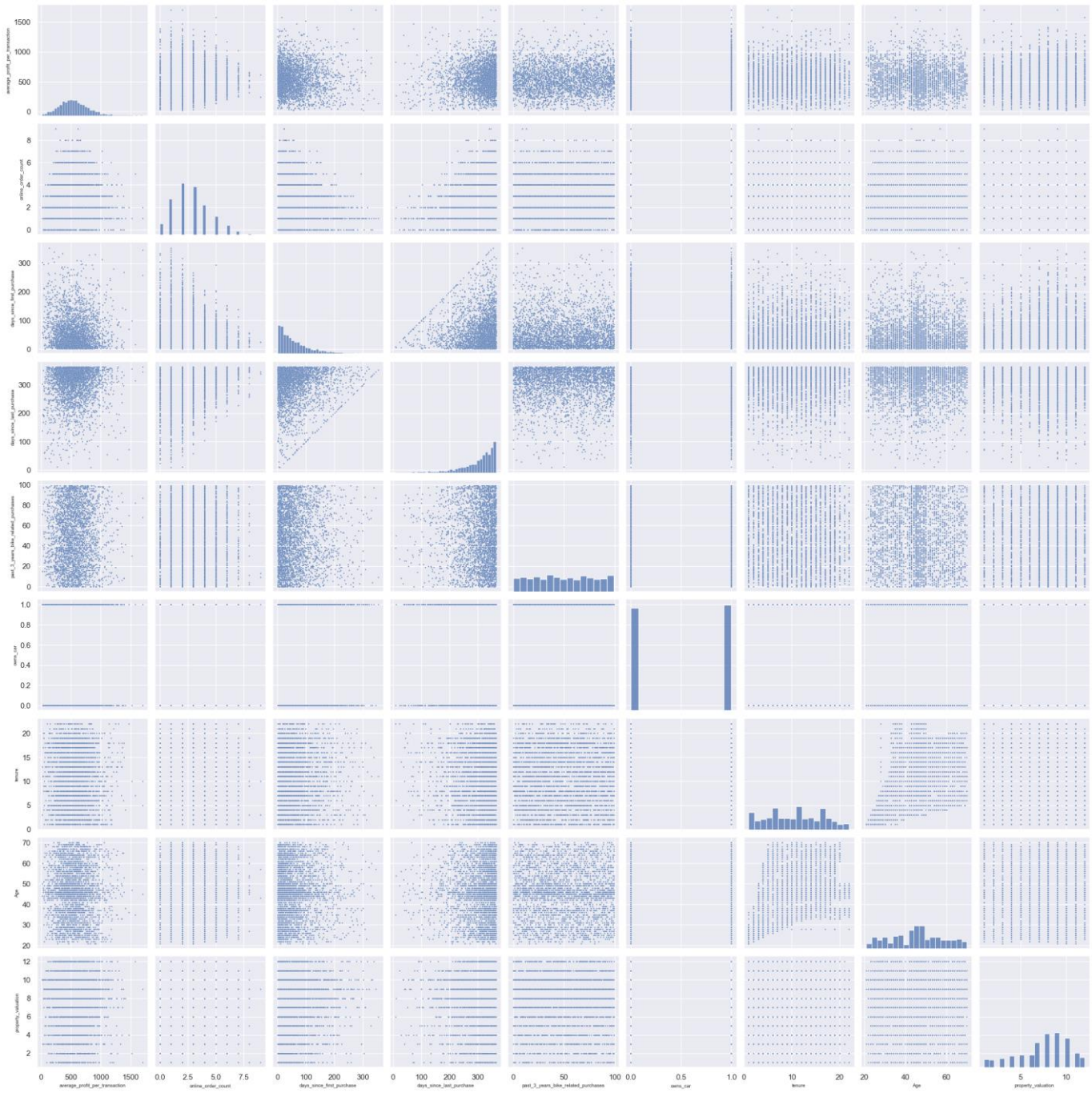

Figure 4.2.3 A

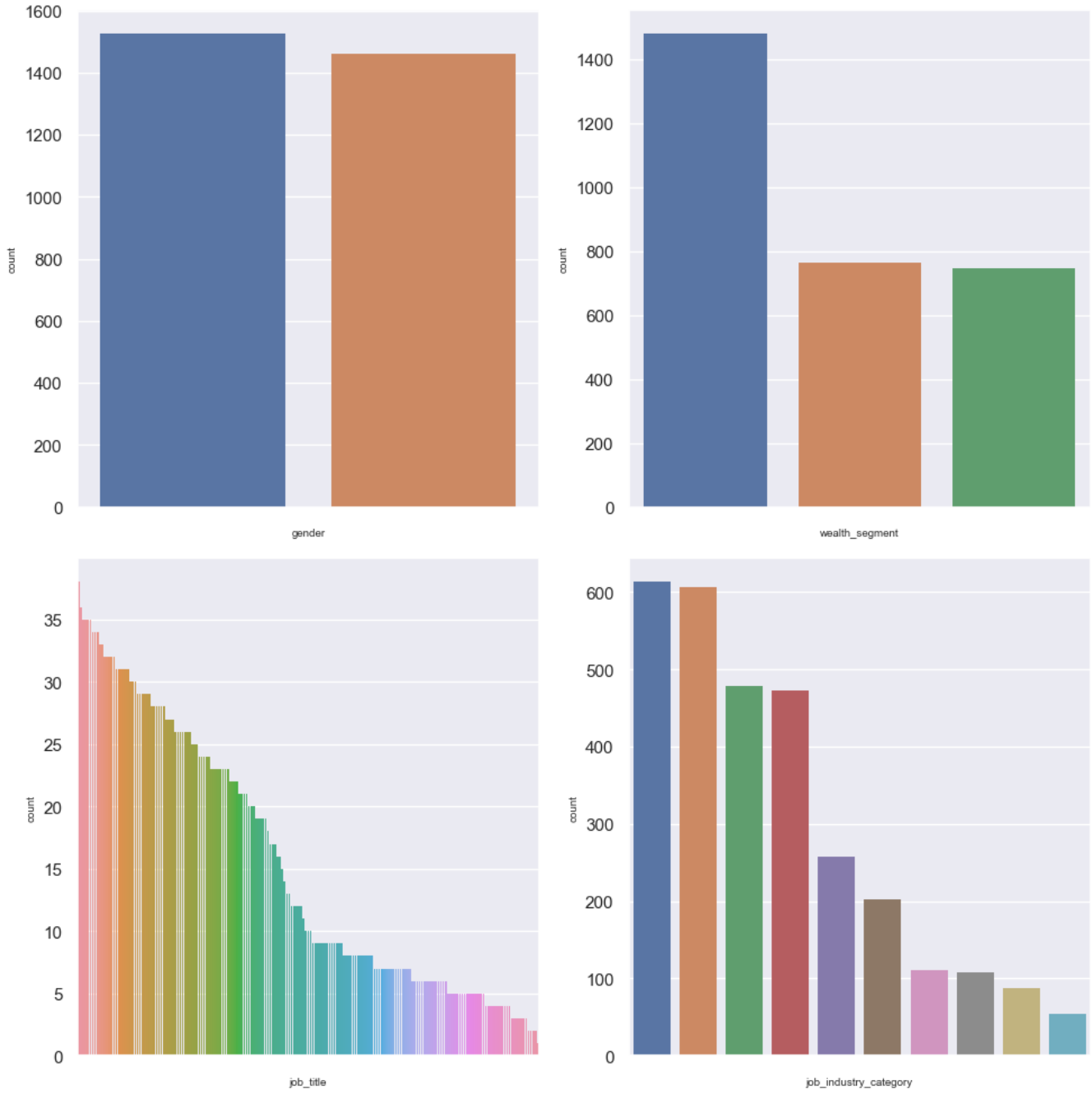
1600
1400
1200
1000
800
600
400
200
0
count
gender
1400
1200
1000
800
600
400
200
0
count
wealth_segment
35
30
25
20
15
10
5
0
count
job_title
600
500
400
300
200
100
0
count
job_industry_category

Figure 4.2.3 B

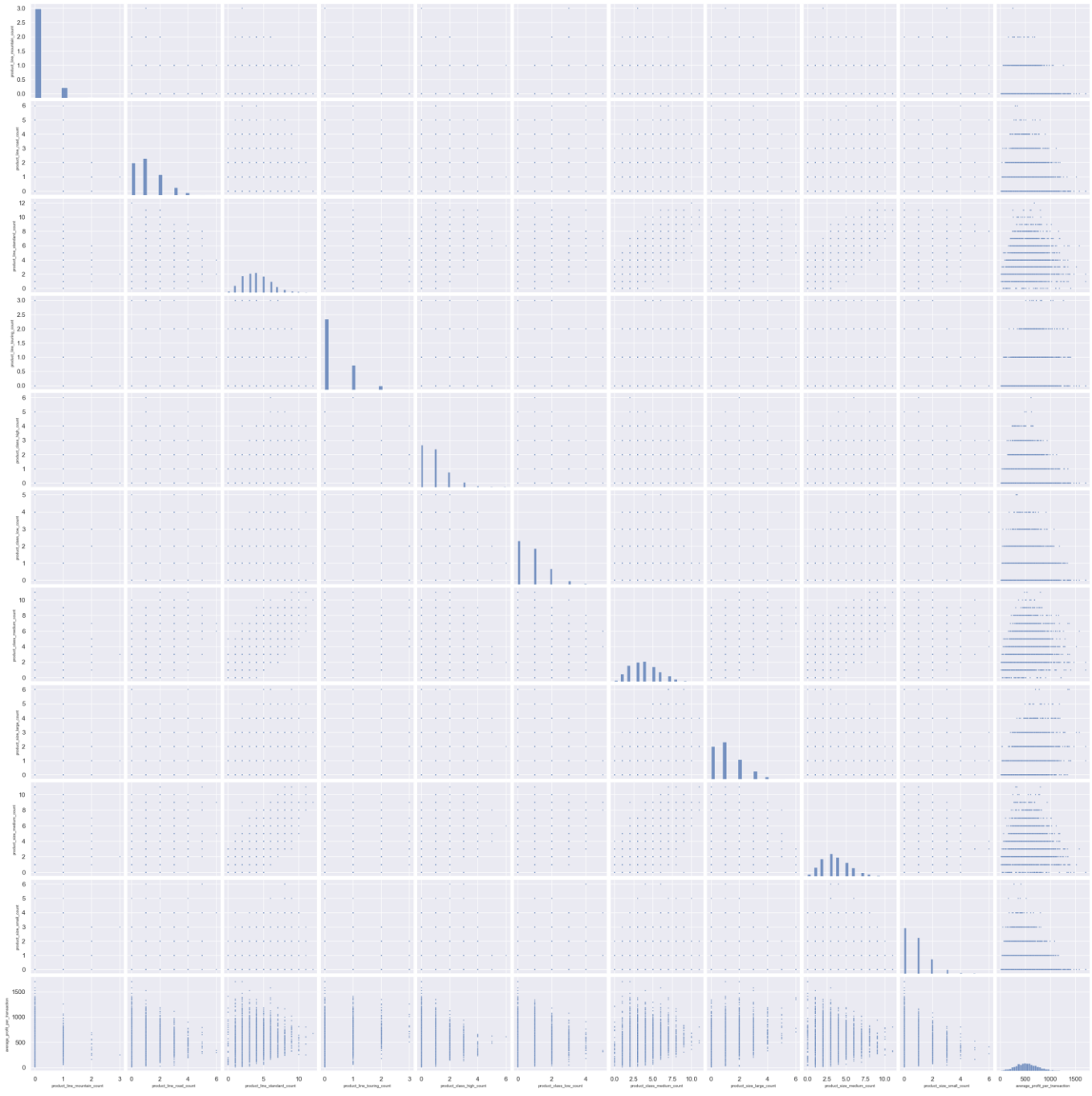

Figure 4.2.3 C

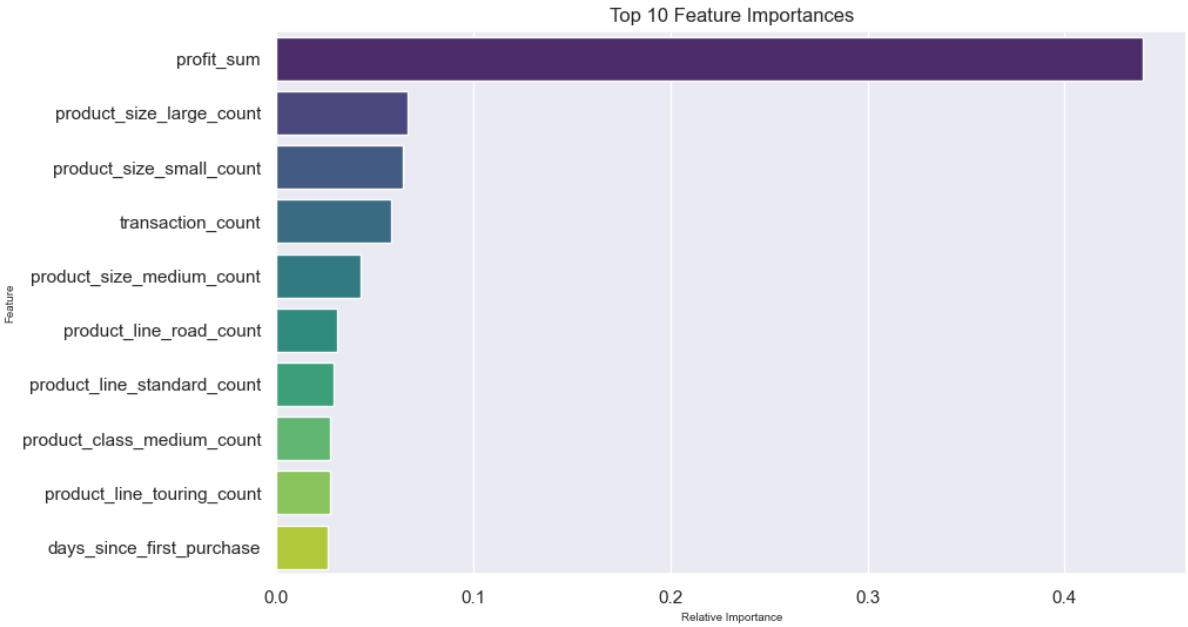

Figure 4.2.3 D

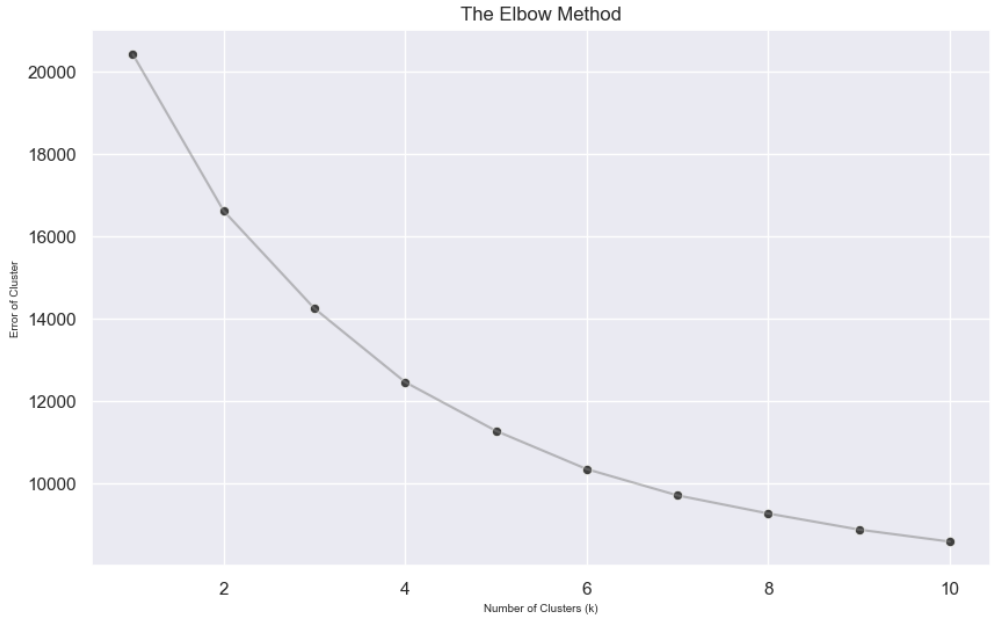

Figure 4.2.3 E

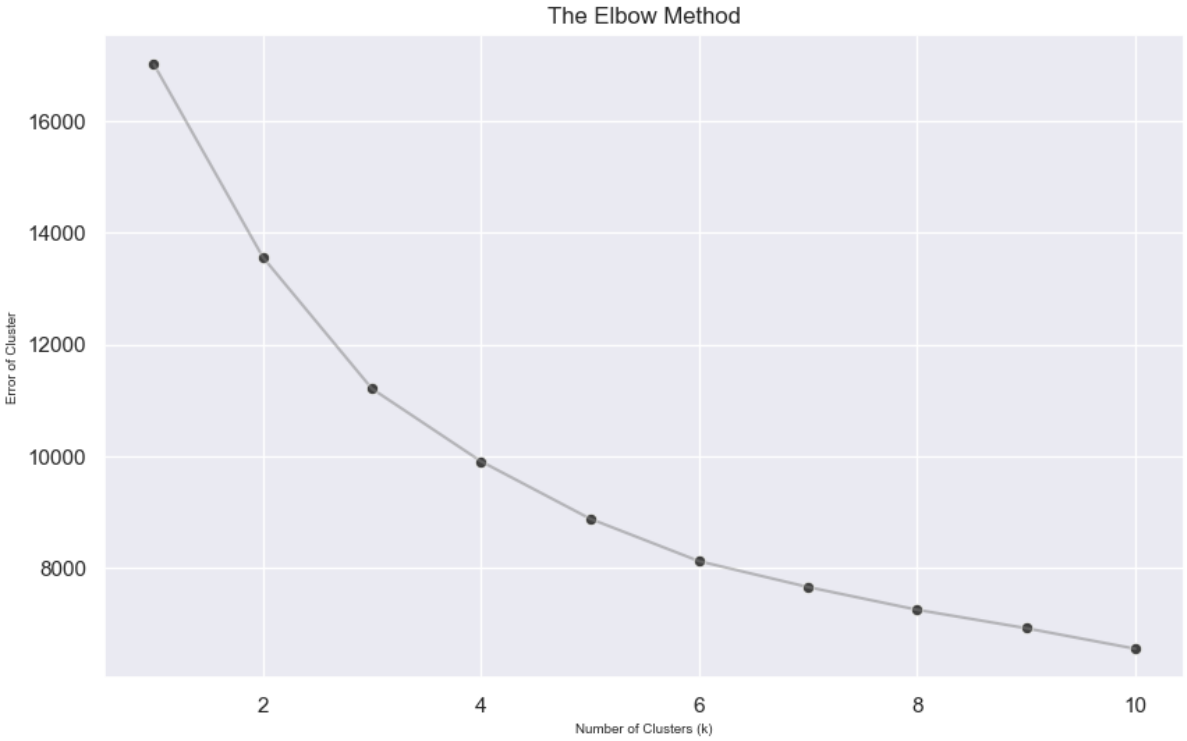

Figure 4.2.3 F

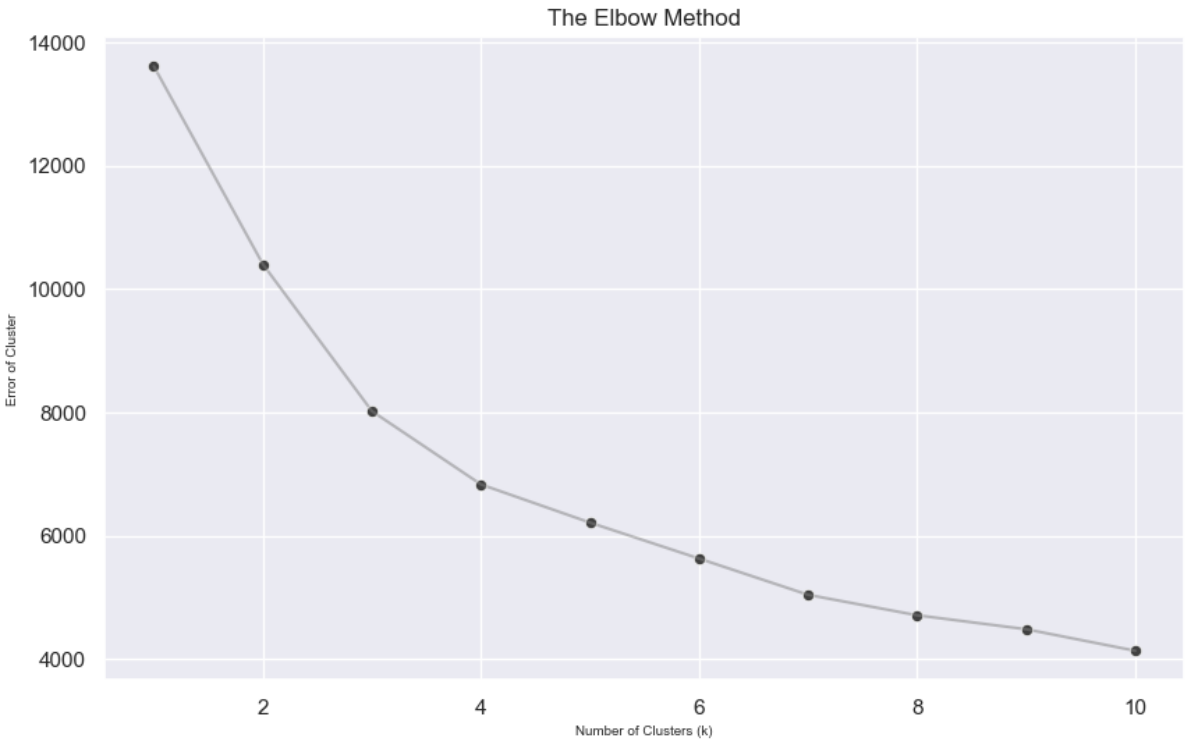

Figure 4.2.3 G

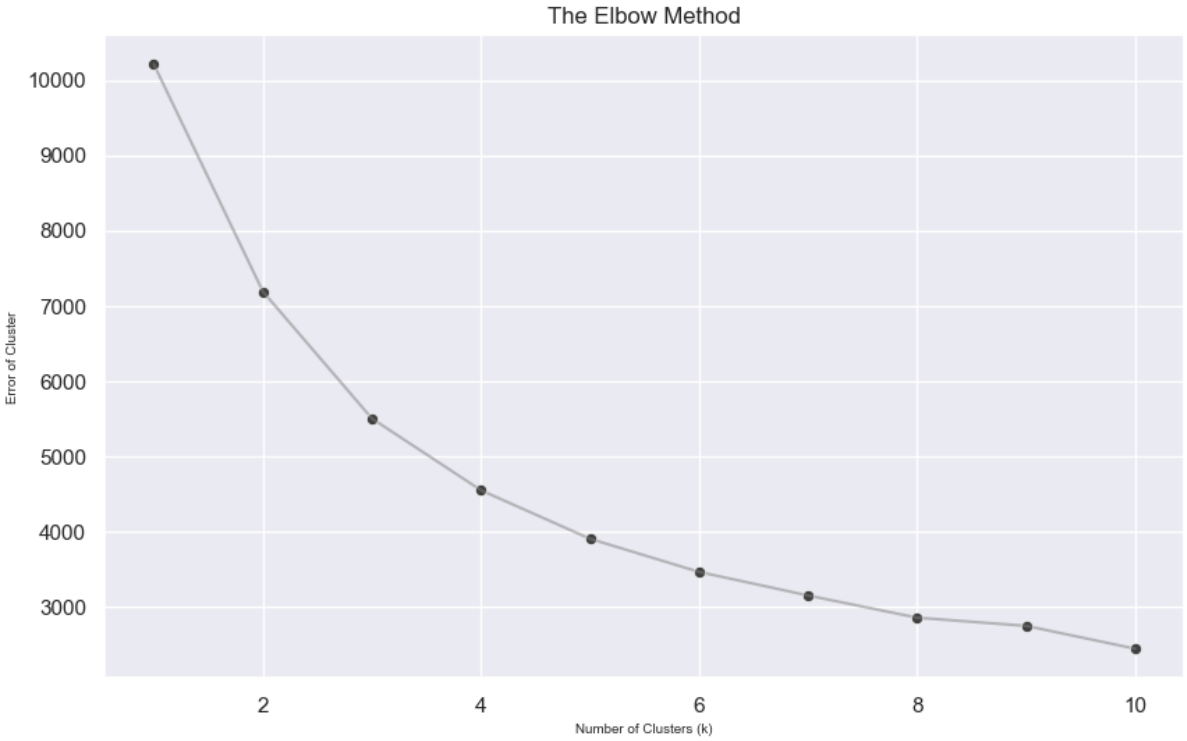

Figure 4.2.3 H

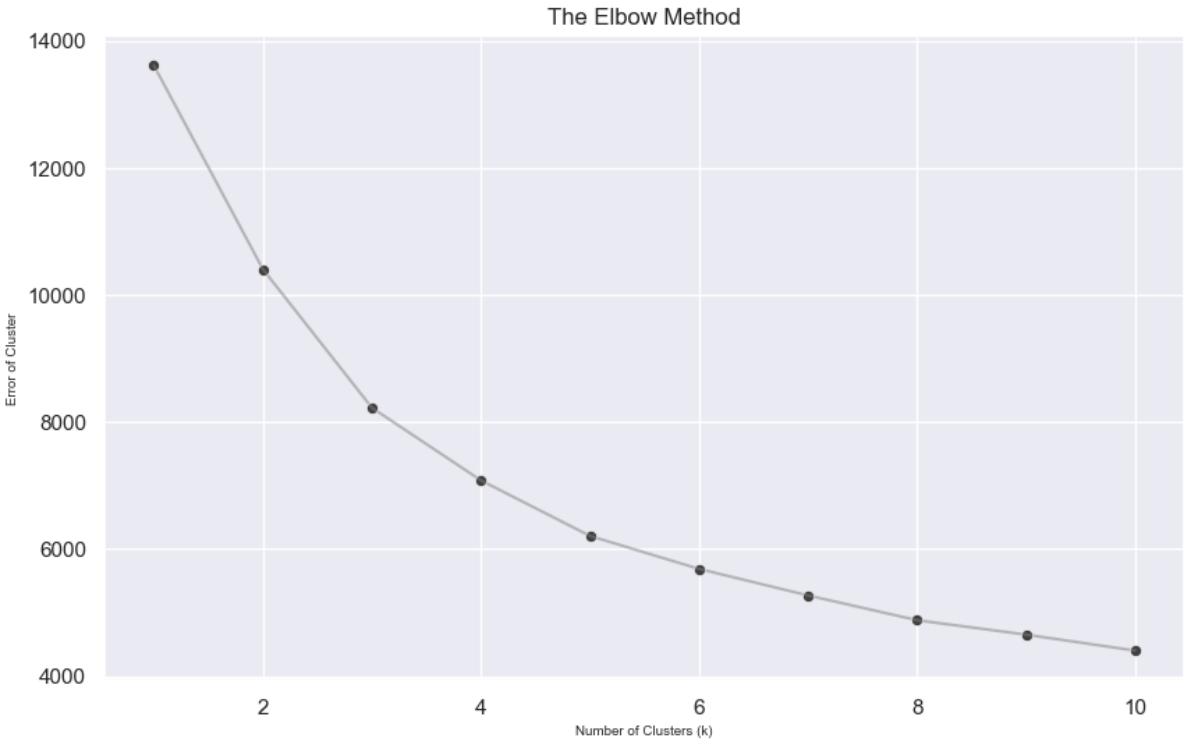

Figure 4.2.3 I

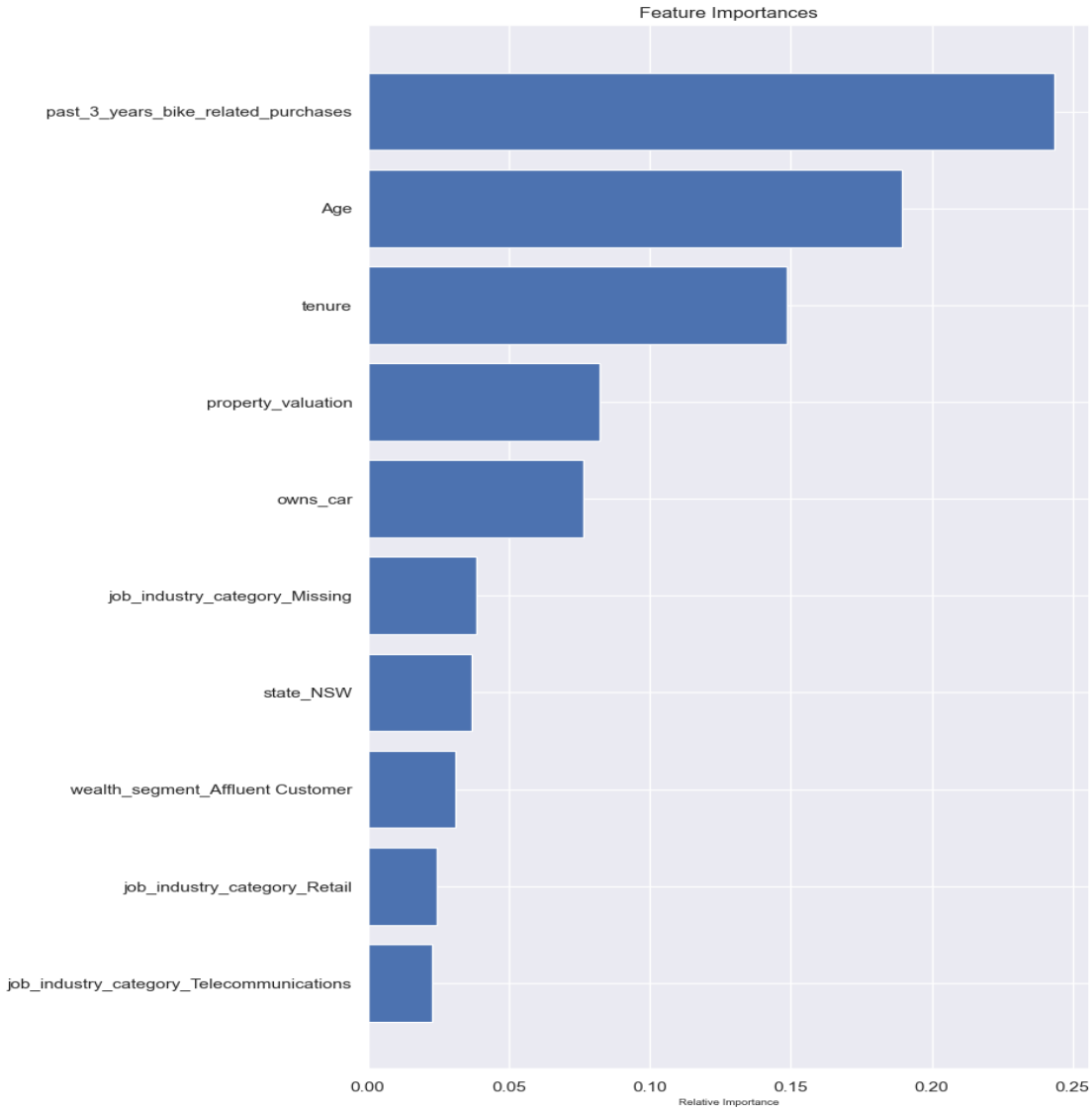

Figure 4.2.3 J

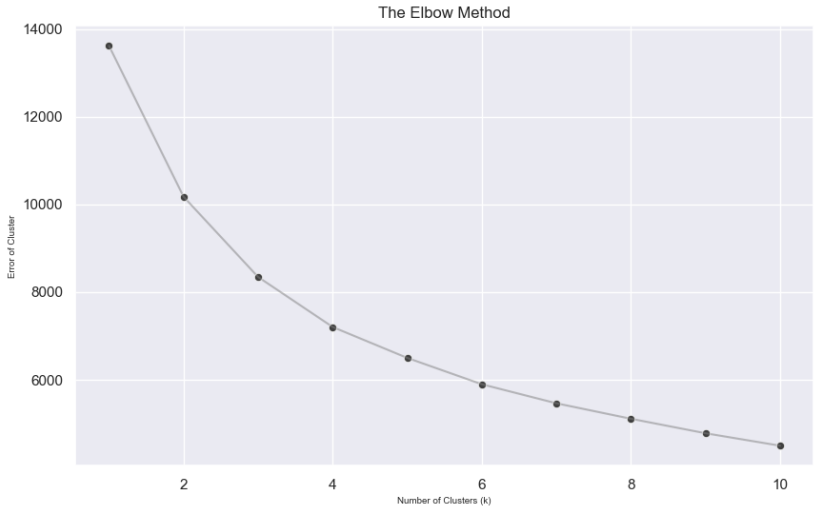

Figure 4.2.3 K

Chapter 05
Methodology

5.1 Introduction to Tools

In this project, python programming language will be employed to drive effective data analysis, model development, and the optimization of targeted marketing efforts at Sprocket Central Pty Ltd. The project will predominantly leverage:

5.1.1 Python is a powerful and widely used programming language in data science. Its extensive library ecosystem, which includes Pandas, NumPy, and Scikit-learn, makes it useful for data manipulation, statistical analysis, and machine learning model development. Python's readability and ease of use enhance the efficiency of the project's coding and analytical processes.

5.1.2 Jupyter Notebooks provide an interactive and collaborative environment for executing code, visualizing data, and documenting the entire analytical process.

5.1.3 The implementation of machine learning libraries such as scikit-learn underscores the application of predictive modeling techniques, with the RandomForestRegressor facilitating accurate regression predictions for tasks like revenue forecasting. Additionally, the utilization of the KMeans algorithm highlights the use of machine learning for clustering, a method employed to categorize data points into distinct groups, fostering enhanced customer segmentation.

5.1.4 Visualization tools like Matplotlib, Seaborn, and Plotly will be employed to create compelling visualizations that aid in conveying complex patterns and trends discovered during data exploration and analysis.

All the libraries used are relevant to the tasks being performed. Let's review the libraries and their functions:

-pandas (import pandas as pd):
Purpose: Used for data manipulation and analysis. It provides data structures such as DataFrame that are essential for handling and processing tabular data.

-numpy (import numpy as np):
Purpose: Used for numerical operations and array manipulations. It is often used in conjunction with pandas for efficient data handling.

-matplotlib (import matplotlib.pyplot as plt):
Purpose: A popular data visualization library. It can generate static, interactive, and animated plots.

-seaborn (import seaborn as sns):
Purpose: Seaborn, which is built on top of matplotlib, provides a high-level interface for creating informative and appealing statistical graphics. It is frequently used in data visualization.

-os (import os):
Purpose: Provides a way to interact with the operating system, and in this case, is used for file and directory operations.

-datetime (import datetime as datetime):
Purpose: Used for working with dates and times. It is employed for manipulating and formatting date-related information in the data.

-IPython (from IPython.display import Image):
Purpose: Specific to Jupyter Notebooks, it allows displaying images within the notebook.

-pickle (import pickle):
Purpose: Used for serializing and deserializing Python objects. In this code, it is used to save and load machine learning models.

-sklearn (scikit-learn) libraries:
RandomForestRegressor, train_test_split, GridSearchCV, KFold, mean_absolute_error, StandardScaler, MinMaxScaler, KMeans
Purpose: scikit-learn is a machine learning library that provides tools for data mining and data analysis. The specific modules used here are for building and evaluating machine learning models, scaling features, and performing clustering.

-plotly (import plotly and import plotly.express as px):
Purpose: Used for creating interactive plots and visualizations. Plotly express is a high-level interface for creating various types of plots.

5.2 Constraints

5.2.1 Data Quality and Availability: The project's success is contingent on the quality and availability of the provided datasets. Incomplete or inconsistent data may limit the accuracy of models and insights.

5.2.2 Time Constraint: The project's completion may be subject to time constraints, affecting the depth of analysis or the complexity of implemented models. Tight deadlines could impact the thoroughness of exploration.

5.2.3 Business Understanding: Working on business related project despite being from technical background make it complex in finding useful variable in dataset.

Chapter 06
Results

6.1 RFM Analysis Results

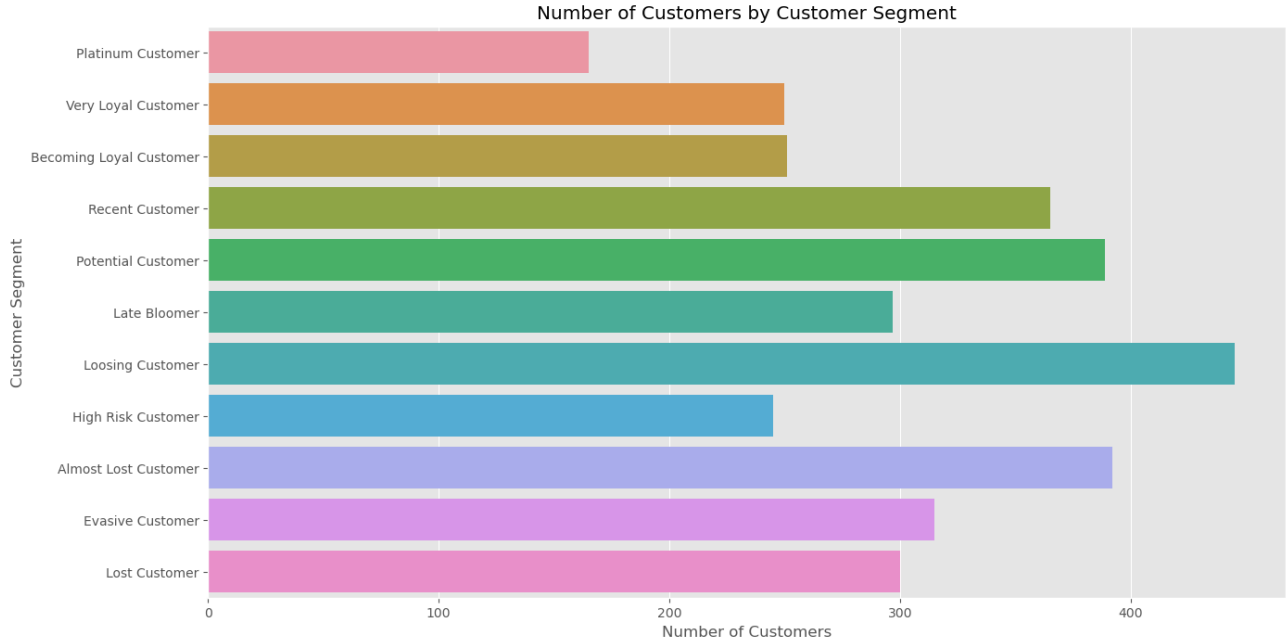

Figure 6.1 A bar graph representing data from an RFM analysis for customer segmentation. The graph displays various customer segments and their respective counts.

There are 12 different customer segments displayed in distinct colors, including “Platinum Customer”, “Very Loyal Customer”, “Becoming Loyal Customer”, “Recent Customer”, “Potential Customer”, “Late Bloomer”, “Loosing Customer”, “High Risk Customer”, “Almost Lost Customer”, “Evasive Customer” and “Lost Customer”. Each segment has a corresponding-colored bar that represents the number of customers in that segment. The x-axis is labeled “Number of Customers” and it ranges from 0 to over 400. The y-axis labels represent the different customer segments. Bars vary in length indicating the number of customers in each segment; for instance, “Lost Customers” have the highest count while “Platinum Customers” have one of the lowest.

6.2 Random Forest Regressor Results

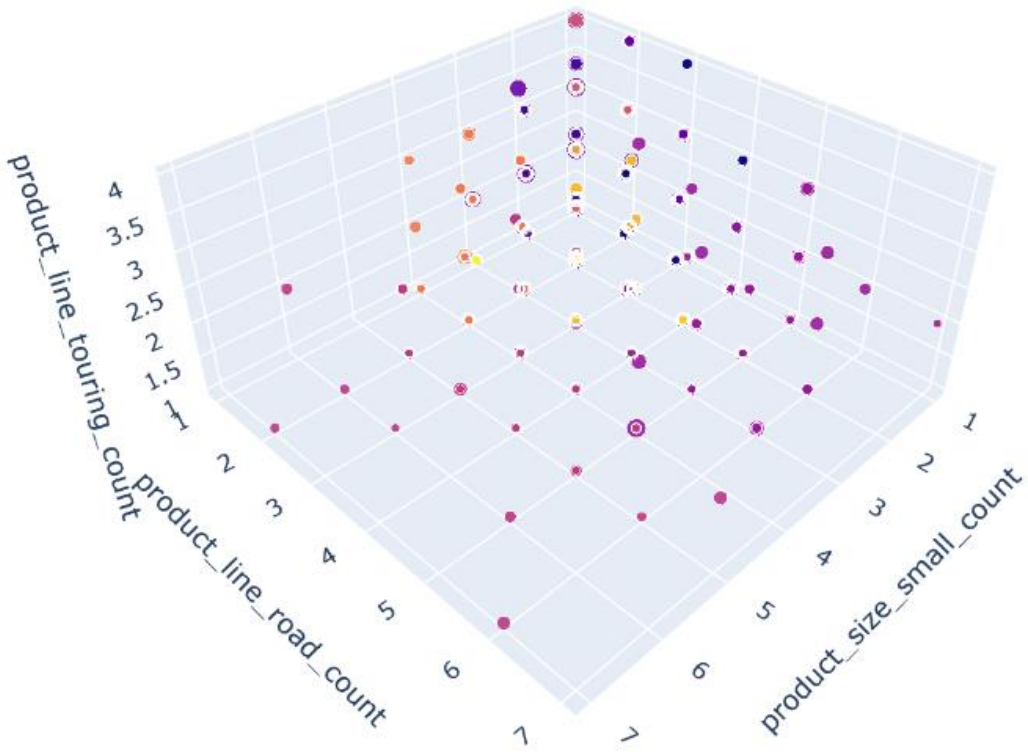

Figure 6.2.1 The 3D plot visualizes the relationship between three variables: 'product_size_small_count', 'product_line_road_count', and 'product_line_touring_count'. The plot shows data points scattered across the plot, indicating varied correlations between these three variables.

The x-axis is labeled "product_size_small_count", ranging from approximately 1 to 8. The y-axis is labeled "product_line_road_count", ranging from approximately 1 to 17. The z-axis is labeled "product_line_touring_count", ranging from approximately 2 to 5. The size and color of the markers are determined by the values of 'product_size_large_count' and 'cluster', respectively. The larger the value of 'product_size_large_count', the larger the marker size. The color of the markers is determined by the value of 'cluster'.

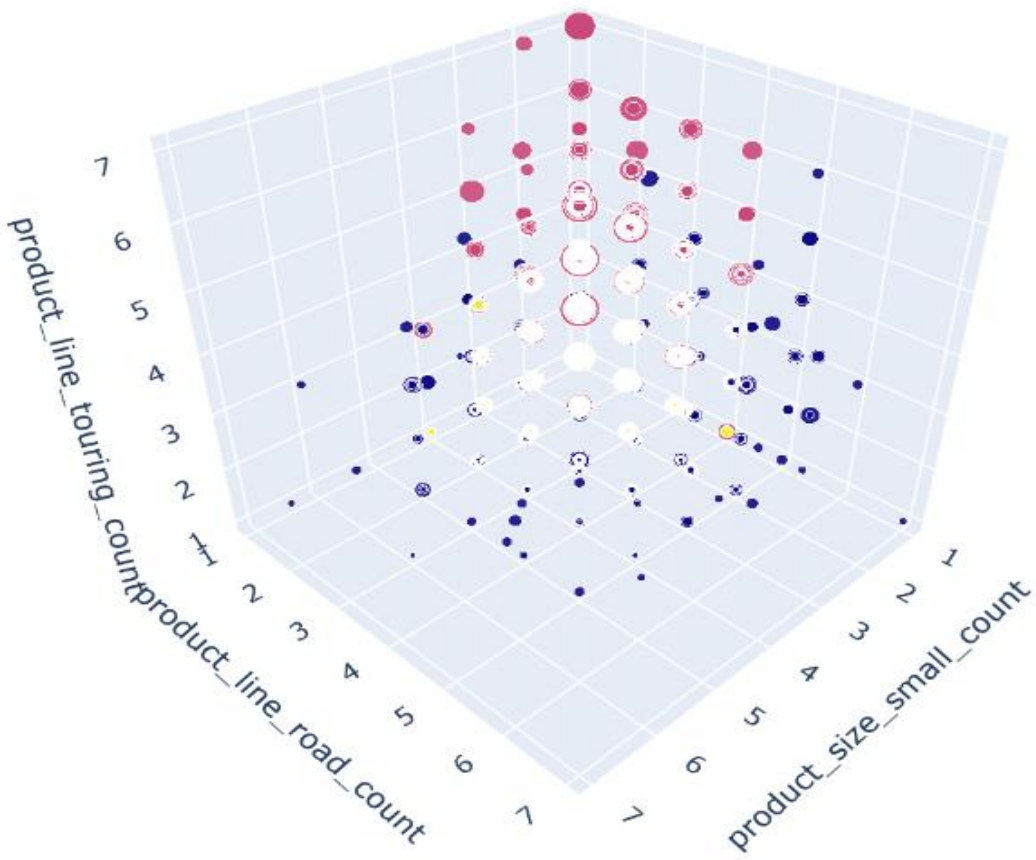

Figure 6.2.2 The 3D scatter plot visualizes data points in three dimensions, where each axis represents a different feature: ‘product_size_small_count’, ‘product_line_road_count’, and ‘product_size_large_count’. This visualization can be helpful to understand the distribution and relationships between these features and how they contribute to different clusters or groups within your data.

The size of the markers in the plot is determined by the ‘average_profit_per_transaction’ and their color by the cluster they belong to, as identified perhaps by a clustering algorithm like K-Means. The x-axis is labeled “product_size_small_count”, ranging from approximately 1 to 8. The y-axis is labeled “product_line_road_count”, ranging from approximately 1 to 17. The z-axis is labeled “product_size_large_count”, ranging from approximately 1 to 10. The size and color of the markers are determined by the values of ‘average_profit_per_transaction’ and ‘cluster’, respectively. The larger the value of ‘average_profit_per_transaction’, the larger the marker size. The color of the markers is determined by the value of ‘cluster’.

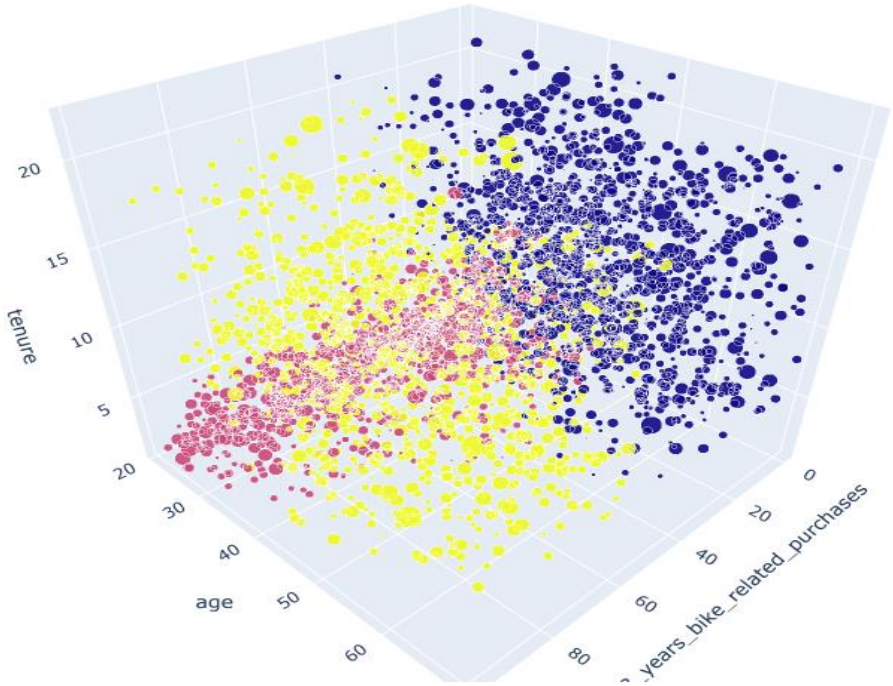

Figure 6.2.3 The 3D plot visualizes data points in three dimensions, where each axis represents a different feature: 'past_3_years_bike_related_purchases', 'age', and 'tenure'.. This visualization is helpful to understand the distribution and relationships between these features and how they contribute to different clusters or groups within your data.

The size of the markers in the plot is determined by the 'average_profit_per_transaction' and their color by the cluster they belong to, as identified by a clustering algorithm like KMeans or hierarchical clustering. The x-axis is labeled "past_3_years_bike_related_purchases", ranging from approximately 0 to 100. The y-axis is labeled "age", ranging from approximately 18 to 90. The z-axis is labeled "tenure", ranging from approximately 0 to 23. The size and color of the markers are determined by the values of 'average_profit_per_transaction' and 'cluster', respectively. The larger the value of 'average_profit_per_transaction', the larger the marker size. The color of the markers is determined by the value of 'cluster'.

Project GitHub link: https://github.com/SKies2003/Dynamic-Customer-Segmentation

Chapter 07

Conclusion and Future Scope

7.1 Conclusion

The integration of data cleaning, RFM analysis, Random Forest Regressor, K-Means clustering, and advanced visualizations forms a robust framework for businesses seeking to optimize their customer segmentation and marketing strategies. Data cleaning lays the foundation by ensuring the accuracy and reliability of the input data, eliminating duplicates, handling missing values, and standardizing formats. This sets the stage for more meaningful and accurate analyses.

RFM analysis provides a simple yet powerful way to categorize customers based on their recency, frequency, and monetary value, offering a foundational understanding of customer behavior. The Random Forest Regressor enhances predictive modeling, allowing businesses to uncover complex relationships between various features and customer behavior. This not only aids in segmentation but also provides insights into the factors influencing customer preferences. K-Means clustering, a machine learning algorithm, further refines segmentation by grouping customers with similar traits, helping businesses tailor their marketing strategies to distinct customer profiles. The combination of these methods allows for a nuanced understanding of customer segments, enabling targeted and personalized marketing approaches.

The incorporation of advanced visualizations enhances the interpretability of results, providing a clear and intuitive representation of complex data patterns. Visualizations can illustrate the relationships identified by Random Forest Regressor, showcase the clusters formed by K-Means, and offer a visual overview of RFM segments. This aids in communication and decision-making within the organization.

In essence, this comprehensive approach empowers businesses to move beyond generic marketing strategies and engage with their customer base in a more personalized manner. By understanding the unique characteristics of different customer segments and the underlying factors influencing their behavior, businesses can implement targeted marketing initiatives, fostering sustained growth, customer satisfaction, and overall success.

7.2 Future Scope

The future scope of customer segmentation, incorporating data cleaning, advanced analytics, and machine learning models, holds considerable promise for businesses looking to stay competitive and responsive to evolving market dynamics. Several trends and possibilities may shape the future of this field:

7.2.1 Advanced Machine Learning Techniques: As machine learning continues to advance, businesses may leverage more sophisticated algorithms and techniques beyond Random Forest Regressor and K-Means clustering. Deep learning models, reinforcement learning, and other cutting-edge approaches could offer deeper insights into customer behavior patterns.

7.2.2 Integration with AI-driven Personalization: The combination of customer segmentation with AI-driven personalization can lead to highly tailored customer experiences. Businesses might use predictive analytics to anticipate individual customer needs and deliver

personalized content, recommendations, and offers in real-time.

7.2.3 Ethical Considerations and Privacy: With increasing concerns about data privacy, businesses will need to navigate ethical considerations when collecting and using customer data. Future advancements may include more transparent and consensual data practices, ensuring that customers feel secure about how their information is utilized.

7.2.4 Predictive Customer Lifetime Value (CLV): Beyond segmentation, businesses may focus on predicting customer lifetime value using advanced analytics. This involves forecasting the potential revenue a customer can generate over their entire relationship with the company, guiding strategic decisions on resource allocation and customer retention efforts.

7.2.5 Customization for Niche Markets: Businesses may increasingly tailor their segmentation models to address niche markets and unique customer segments. This level of granularity allows for hyper-personalized marketing strategies that resonate with specific customer demographics and preferences.

In summary, the future of customer segmentation involves a continued evolution towards more sophisticated, real-time, and ethical approaches. The integration of emerging technologies and a deeper understanding of customer behavior will play a crucial role in helping businesses stay agile and responsive in an ever-changing market landscape.

PRJ-1

ORIGINALITY REPORT

10%	8%	2%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Guru Nanak Dev Engineering College Student Paper	2%
2	Submitted to The Hong Kong Polytechnic University Student Paper	1%
3	ebin.pub Internet Source	1%
4	www.ir.juit.ac.in:8080 Internet Source	1%
5	Guixiang Liao, Yongbo Li, Yongzai Xi, Ning Lu, Shan Wu. "Application of High-Resolution Aeromagnetic and Gamma-ray Spectrometry Surveys for Litho-Structural Mapping in Southwest China", Minerals, 2023 Publication	1%
6	Submitted to University of San Francisco Student Paper	1%
7	dokumen.tips Internet Source	1%

8	Submitted to Deenbandhu Chhotu Ram University of Science and Technology Student Paper	<1 %
9	Submitted to Punjab Technical University Student Paper	<1 %
10	medium.com Internet Source	<1 %
11	Submitted to University of North Texas Student Paper	<1 %
12	grietinfo.in Internet Source	<1 %
13	Submitted to University of Birmingham Student Paper	<1 %
14	Submitted to Glasgow Caledonian University Student Paper	<1 %
15	www.brecw.ac.in Internet Source	<1 %
16	Submitted to HELP UNIVERSITY Student Paper	<1 %
17	ntrs.nasa.gov Internet Source	<1 %
18	s3-eu-west-1.amazonaws.com Internet Source	<1 %
19	dspace.dtu.ac.in:8080	

Internet Source

<1%

Exclude quotes	On	Exclude matches	< 6 words
Exclude bibliography	On		