



# Sprocket Central Pty Ltd

Data Analytics Approach

---

THE ANALYTICS TEAM

# Contents

---

1. Introduction
2. Data Exploration
3. RFM Model Development
4. Interpretation

# Introduction

---

- 1000 potential customers do not have prior transaction history.
- Generate useful customer insights
- Optimize resource allocation for targeted marketing.

# Data Exploration

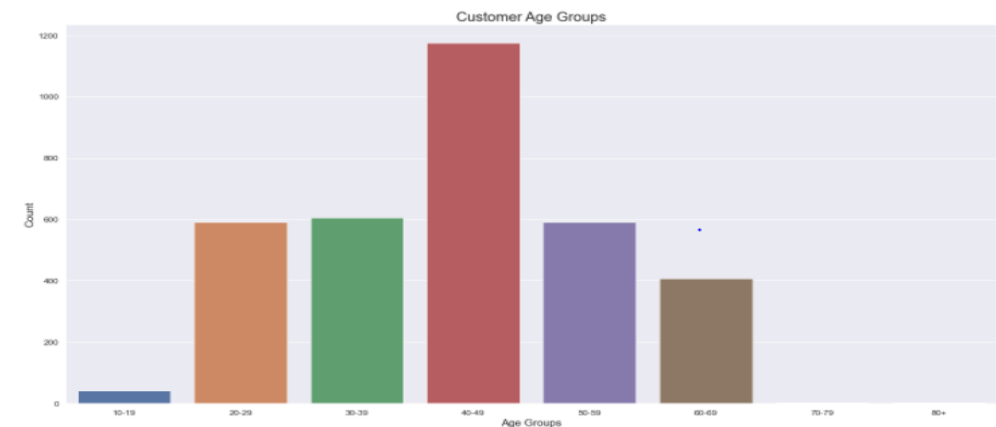
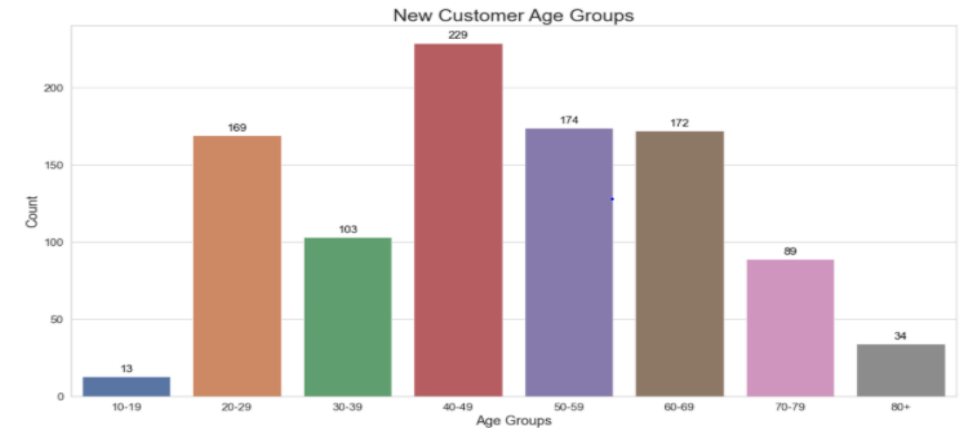
---

## **Data Cleaning and Quality Assessment**

- Accuracy: Correct Values
- Completeness: Data Fields with Values
- Consistency: Value Free From Contradiction
- Currency: Values Up to Date
- Relevancy: Items with Value Meta-data
- Validity: Data Containing Allowable Values
- Uniqueness: Duplicated Records

# Data Exploration – Customer Age Distribution

- Most customers age around 40 to 49 years old
- The least popular groups are 10 to 19 and 70 years old and above.
- Popular shoppers tend to be between 20
- to 69 years old.

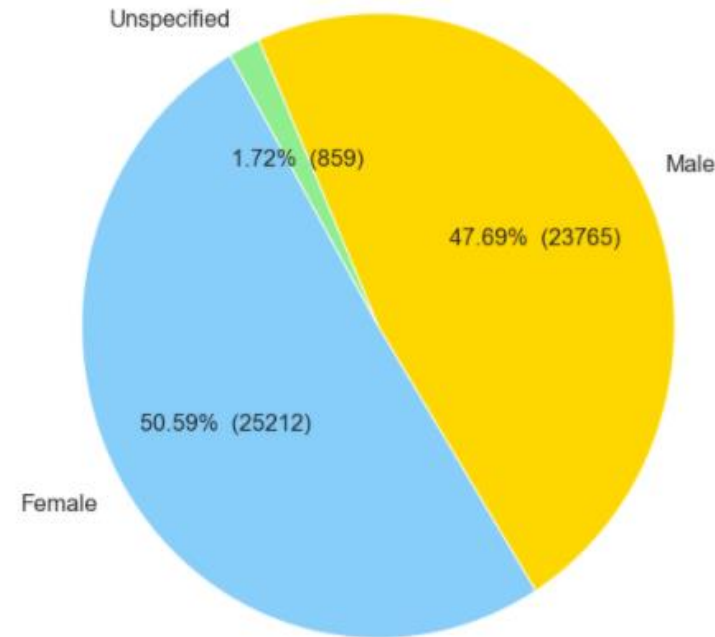


# Data Exploration – Purchases in the Past 3 Years By Gender

---

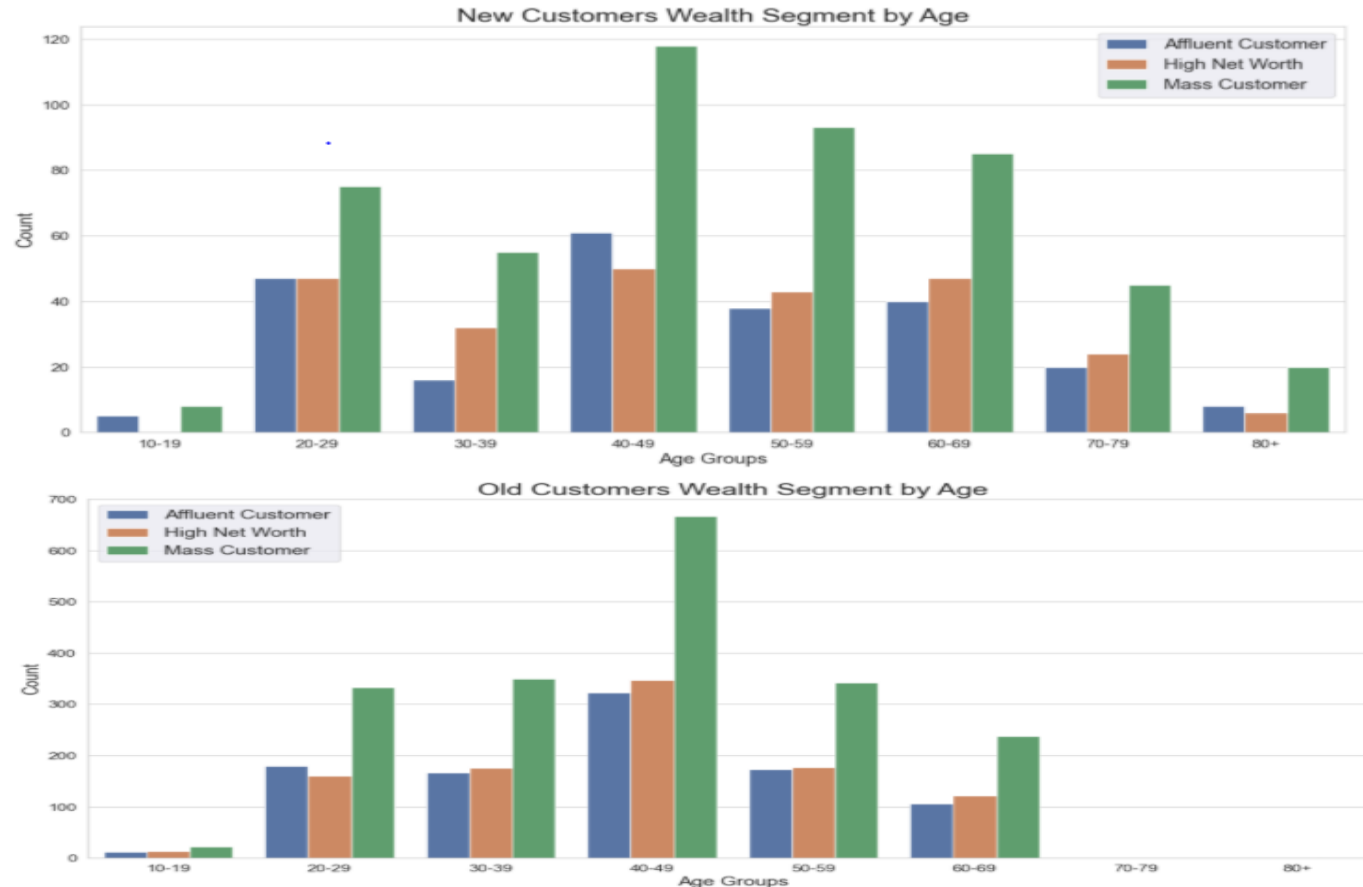
- Customer gender demographic evenly spread between male and females.
- Females make up majority of the bike purchases in the past three years

New Customer Total Bike Related Purchases For The Past 3 Years By Gender



# Data Exploration – Wealth Segmentation By Age

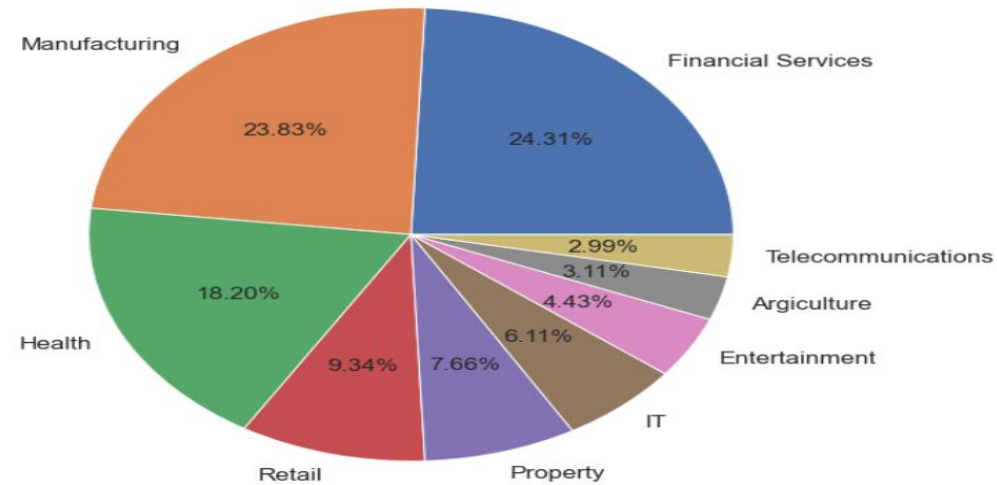
- Most customers are labelled as Mass Customer.
- More customers labelled as High Net Worth than Affluent by a small margin.



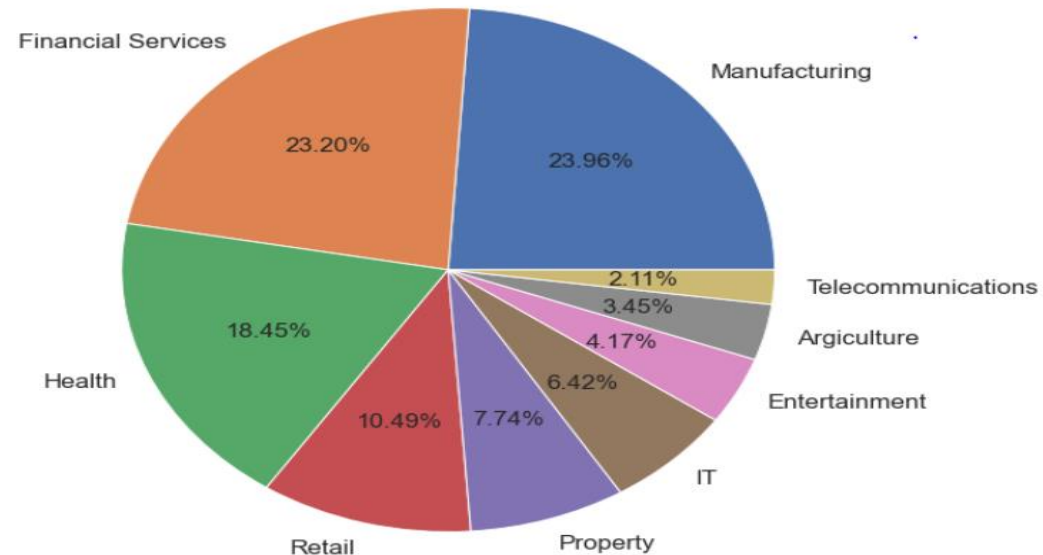
# Data Exploration – Job Industry Distribution

- Top three job industries that customers work in are financial services, manufacturing, and health
- Least popular job industries are telecommunications, agriculture, entertainment, and IT

New Customer Job Industry Category Distribution



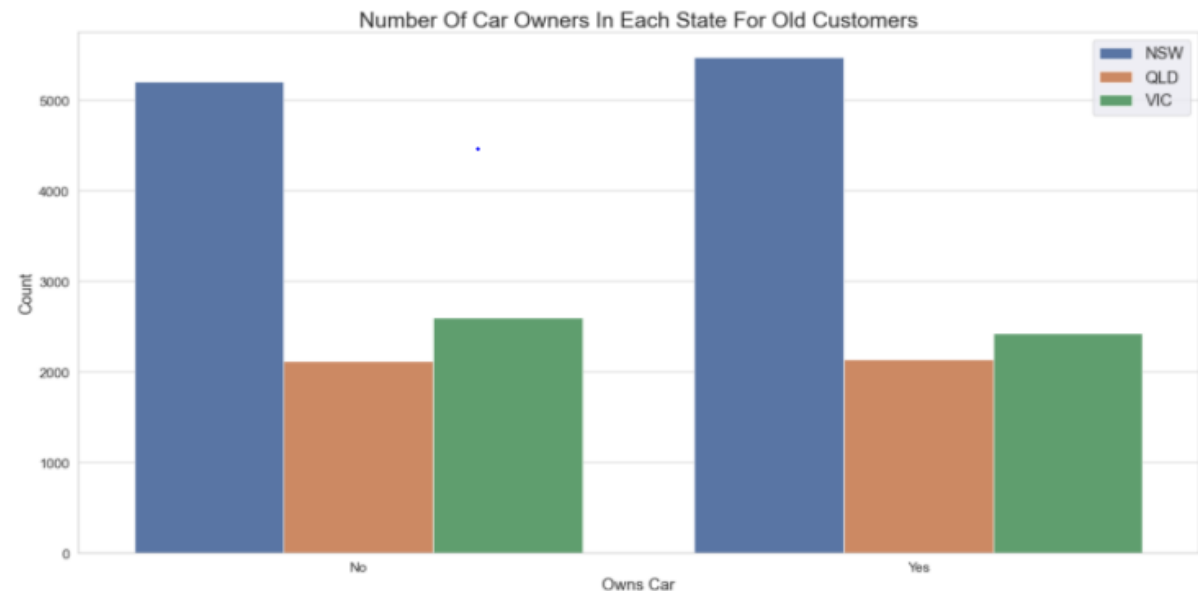
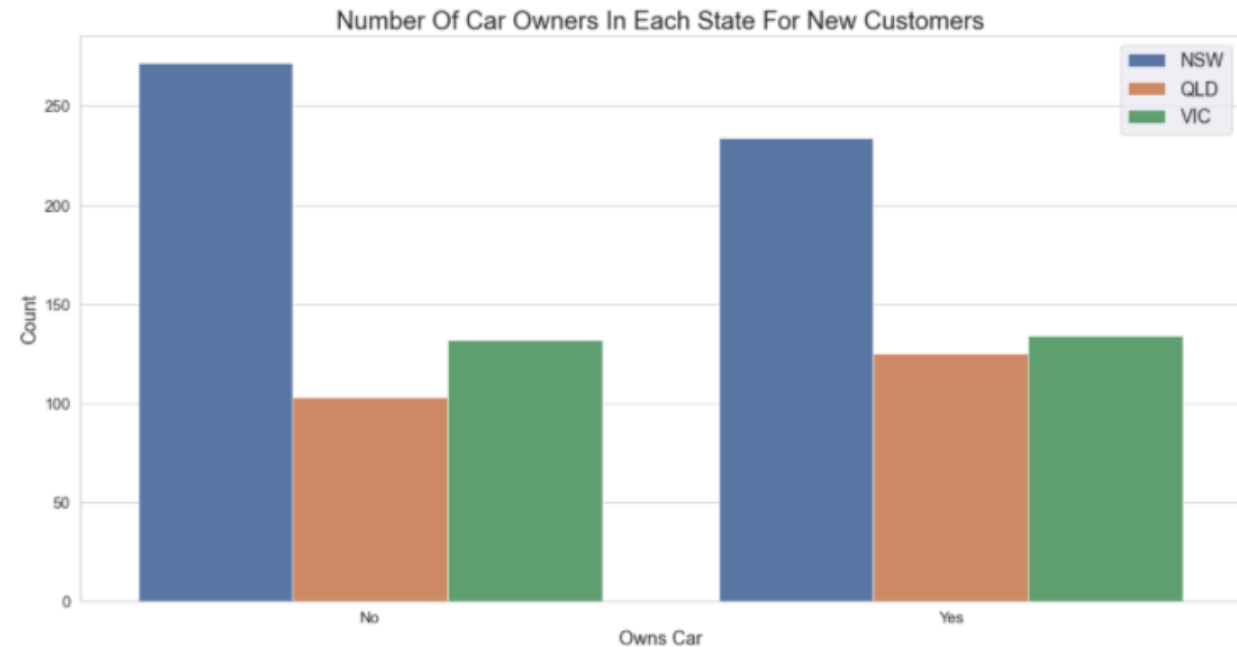
Old Customer Job Industry Category Distribution





# Data Exploration – Car Ownership By State

- Most customers from New South Wales don't own a vehicle
- Consider to allocate more stocks and resources on the NSW region.

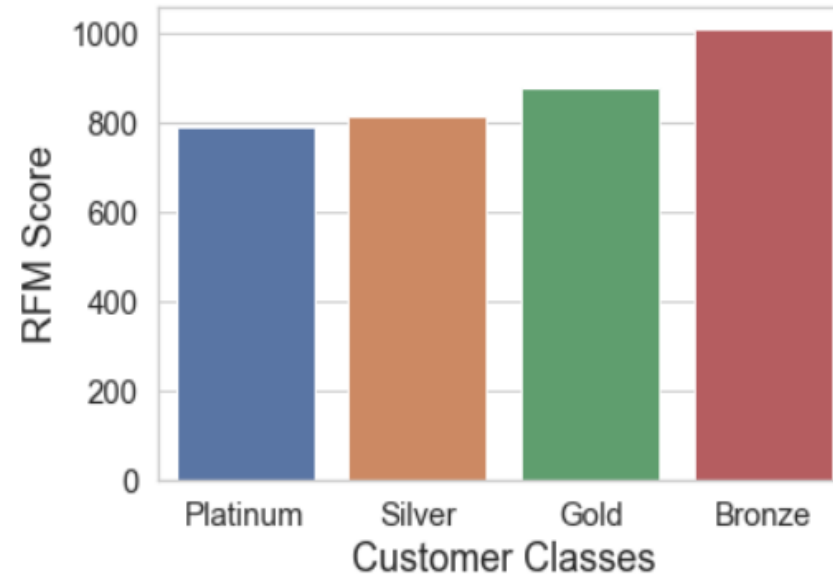


# Model Development – RFM Analysis

- RFM stands for Recency, Frequency and Monetary.
- Frequency and Monetary affects customer lifetime value.
- Recency affects customer retention
- Classified into four groups by their RFM score from highest to lowest respectively: Platinum, Gold, Silver, and Bronze

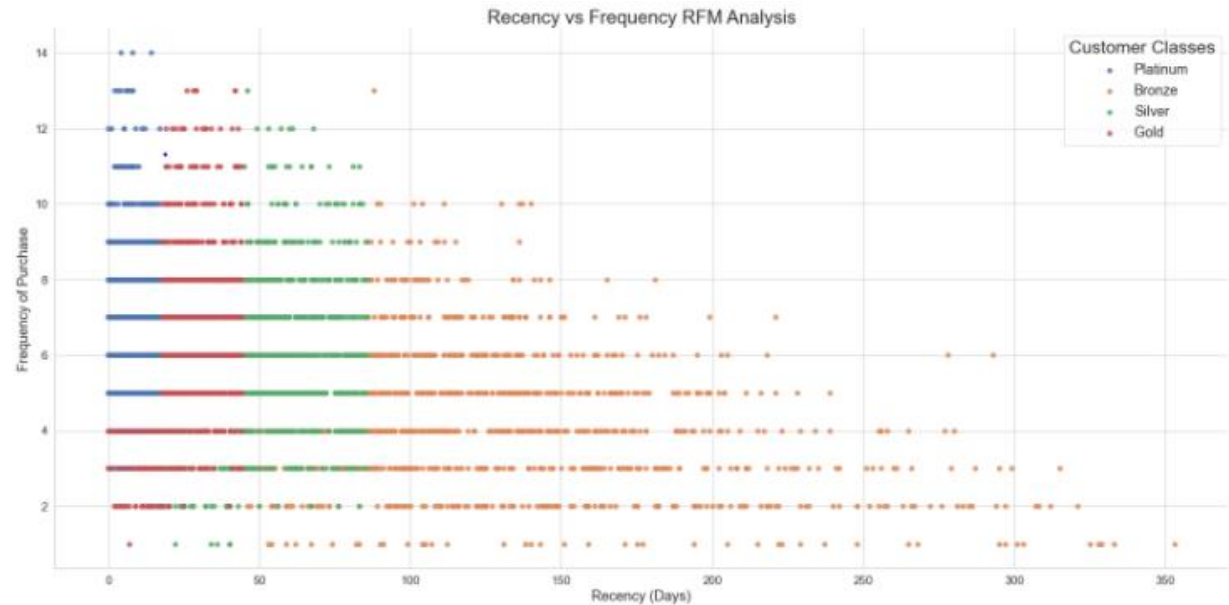
customer_title		title_description	count
Platinum		bought more often and spent the most	791
Gold		bought relatively often and spent large amount...	878
Silver		average purchase frequency and average spendin...	813
Bronze		very low recency, frequency and money spent	1009

Customer Classes Determined With RFM Analysis



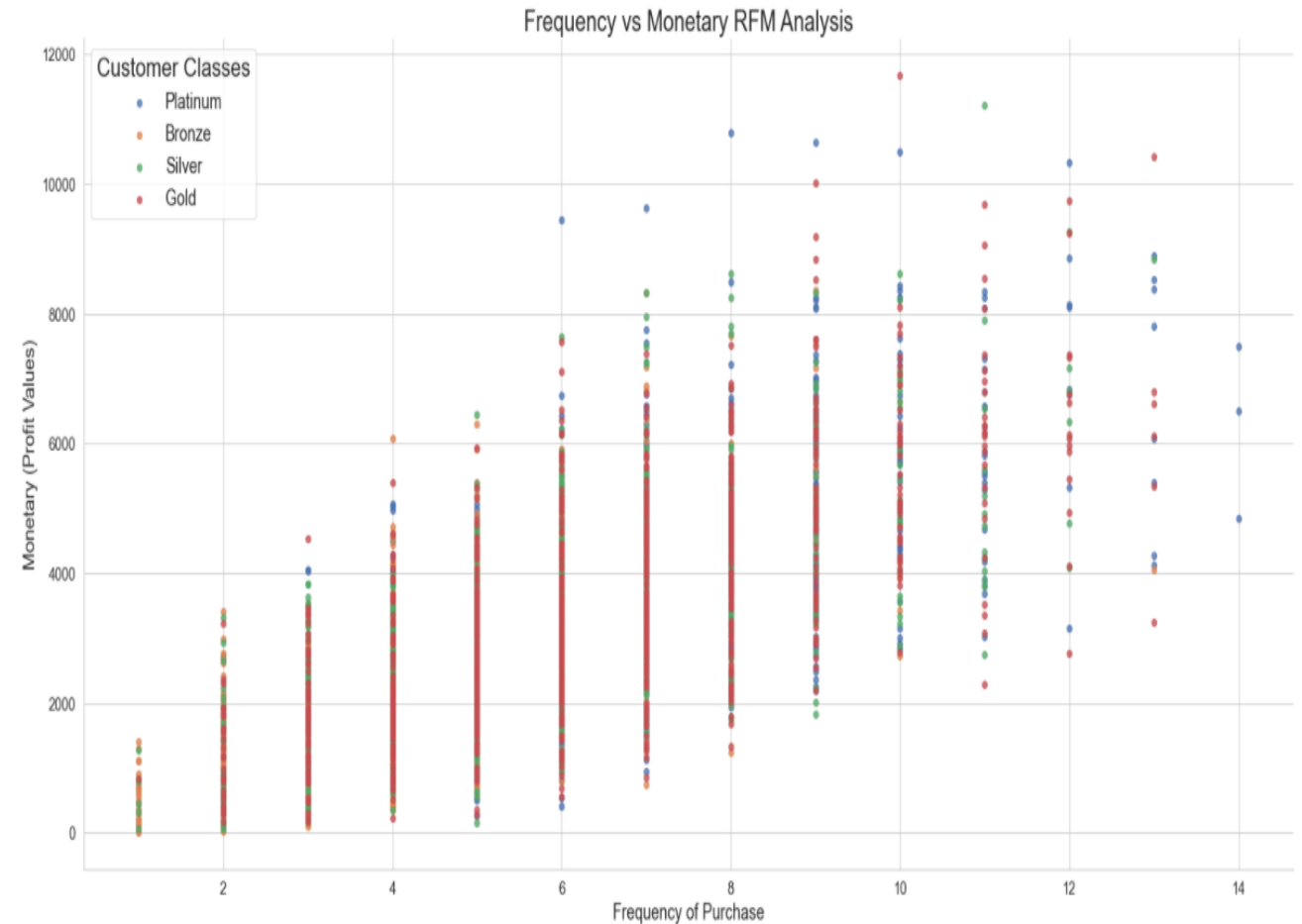
# Model Development – Recency vs Frequency

- Platinum customers bought more recently and frequently than other groups
- Negative correlation between recency and frequency
- The less recent the last purchase is, the less frequent the customer buys



# Model Development - Frequency vs Monetary

- Positive correlation between recency and monetary
- Business should expect higher monetary gains when the more frequently the customer buys



# Interpretation

## – Target Market

- Sprocket should focus on the top – 1000 potential customers based on RFM score and tend to retain most valuable customers.
- Should focus more on the customers within age group 40-49, the customers working for financial service sector, and customers live in NSW do not own car.

RFM_score		customer_title
customer_id		
925	444.0	Platinum
725	444.0	Platinum
793	444.0	Platinum
790	444.0	Platinum
144	444.0	Platinum
...	...	...
2887	344.0	Gold
2474	344.0	Gold
647	344.0	Gold
2755	344.0	Gold
171	344.0	Gold

1000 rows x 2 columns

RFM_score		customer_title
customer_id		
925	444.0	Platinum
725	444.0	Platinum
793	444.0	Platinum
790	444.0	Platinum
144	444.0	Platinum
768	444.0	Platinum
2651	444.0	Platinum
2653	444.0	Platinum
2659	444.0	Platinum
151	444.0	Platinum
739	444.0	Platinum
729	444.0	Platinum
719	444.0	Platinum
2379	444.0	Platinum
2695	444.0	Platinum
2705	444.0	Platinum
3337	444.0	Platinum
695	444.0	Platinum
669	444.0	Platinum
165	444.0	Platinum