

TheAnalyticsTeam

# Sprocket Central Pty Ltd

Data analytics approach

By,  
Sugandh Bansal

# Agenda

1. Introduction
2. Data Cleaning and Exploration
3. Data Insights and Visualizations
4. Conclusion and Recommendation

## Introduction

### **Analysing report of 3 datasets (customer demographic, customer address and transactions) from Sprockets to find high-value customers among 1000 new customer's data**

Sprockets Central Pvt. Ltd. specializes in high quality bikes and accessible cycling accessories to riders. The marketing team of the company is looking for a report to identify high-value customers from their 1000 new customers dataset.

To get the result, we used some analysing tools to find the customer segmentation, which can bring value to the company. For that we did:

- Age and gender based analysis for customers
- Purchasing record of bike in past 3 years
- State analysis on the base of transaction
- Job industry distribution based on new and old customers
- Customer segmentation on the basis of transaction and profit margin
- RFM modelling

## Data cleaning

1. First, we identified anomalies and rectified them for better analysis to the dataset. This involved:

- \* Cleaning inconsistent and irrelevant values and corrupted data, like redundant columns (default, unnamed), blank records etc.
- \* Updated unformatted data, like gender segmentation format, state defining format
- \* Adding a new column for calculations and analysis, like, customer's age by using their provided DOB and profit margin by using list price and standard cost, and also age range to segment the age
- \* Filtered results for age below 100 year age to get the data for present condition

2. After cleaning the data we merged the 3 given datasets ('Demographics', 'Customer's address' and 'Transactions') on basis of Customer ID to form a complete dataset into one dataset containing all the necessary information

3. Checked for any duplicity or repetition in the datasets

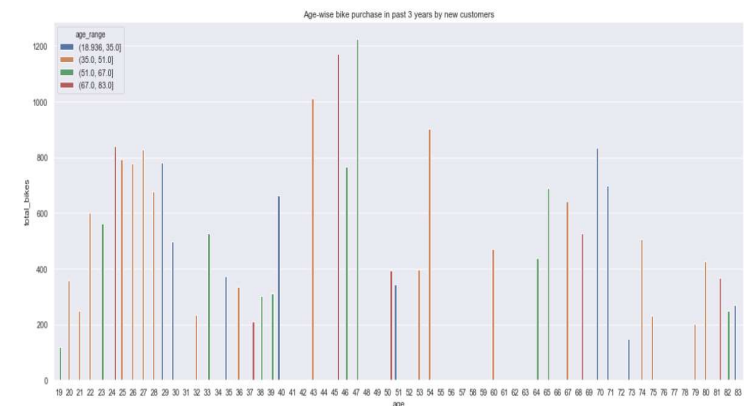
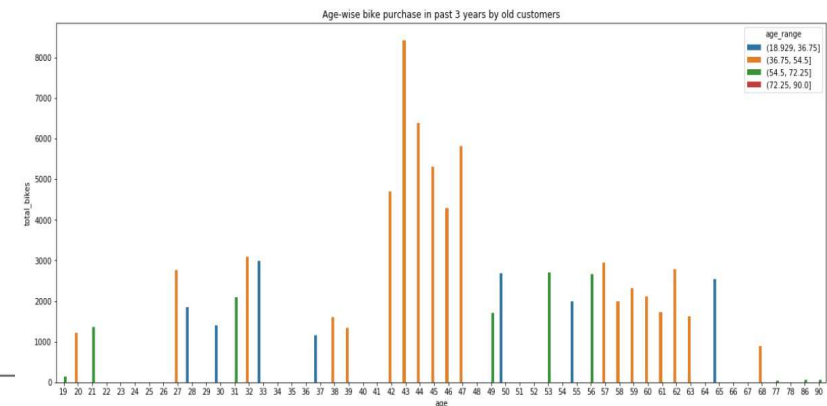
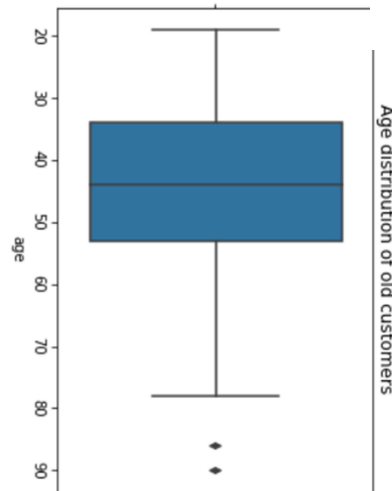
# Data Exploration

## Age distribution :

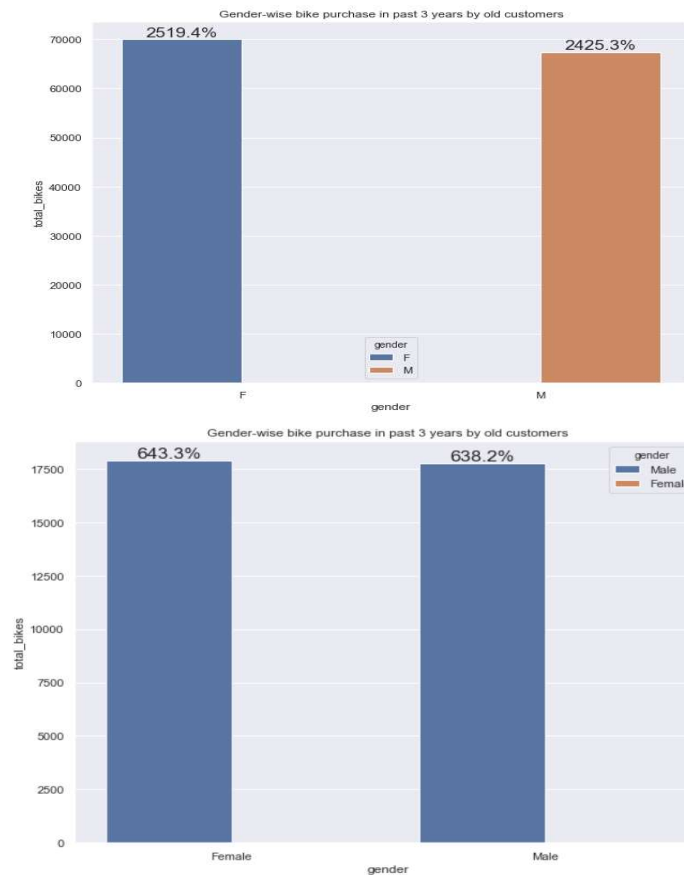
We can see that the age distribution in the 3 datasets is between 20 to 90 years, though mostly customers from the age of 42 to 47 seemed most active.

According to the dataset, the study says that the customers of the age group of 42 to 47 have been recorded to have made the most purchases with Sprockets in last 3 years. Thus, this should be a targeted age group to consider for ferocious marketing for the marketing team.

We can see in the second figure, in new customer dataset that age group of 42 to 54 has more potential customer than any other age group.



# Data Exploration



## Gender-wise analysis :

Bike purchase record analysis is on the base of the gender. The female population is higher in the dataset. Approximately more than 50% of the bikes were bought by females, comparatively with the male customers accounting for approximately 48% of the purchases. Female customers having higher bike purchase records. We can see in the figure, that the number of females are more than males in the customer datasets, so hypothetically, we can say that females use bikes more than males.

Thus, I would recommend to make up majority of the sales offers should be effectively targeted to females. In second figure, in new customer dataset also females are more in purchasing than males, but the figures have very marginal difference.

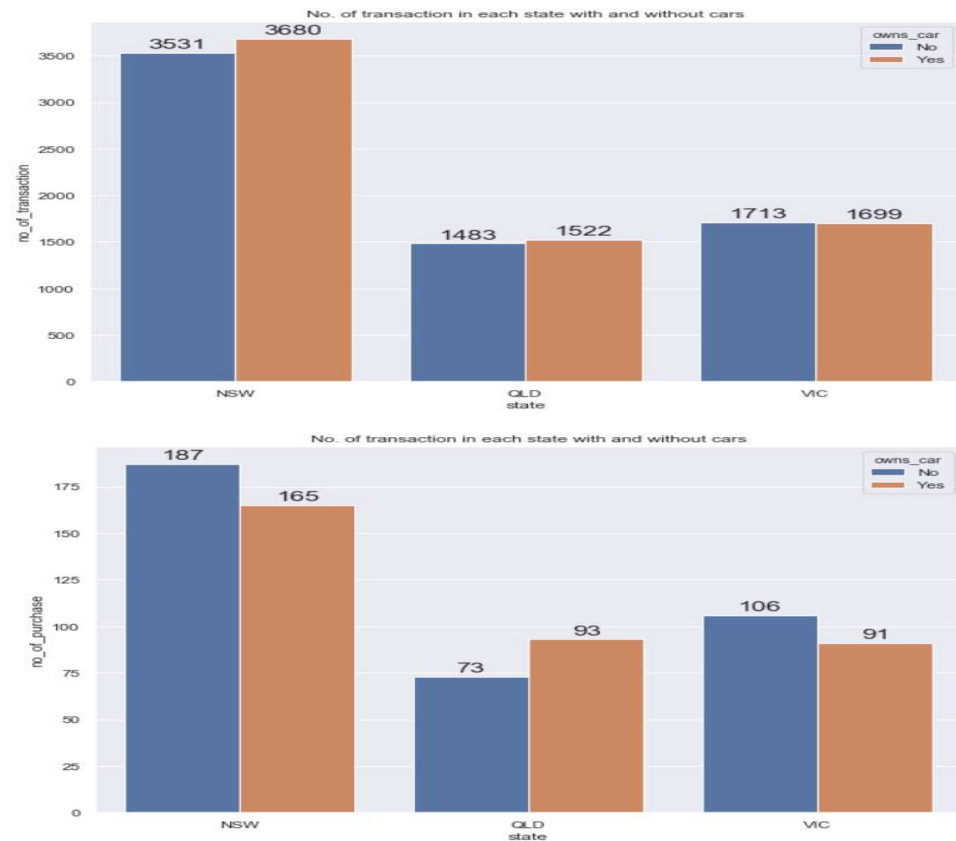
## Data Exploration

### Number of transaction:

Here in this figure, we can see that though the customers use cars, they still bought bikes in more numbers than to those who did not have cars, specially in NSW (New South wales). The NSW had the highest number of customers than the other states and QLD had the lowest number of customers.

Thus, the marketing team should focus people, who live in the New South Wales state and owns a car, as they will get potential more buyers from there, comparatively to another state. We can see in second figure that the new customer dataset also indicates potential customers in NSW.

These customers can be targeted with special schemes for encouragement There's still a very large population in all 3 states that are not yet enrolled as Customers of Sprockets.

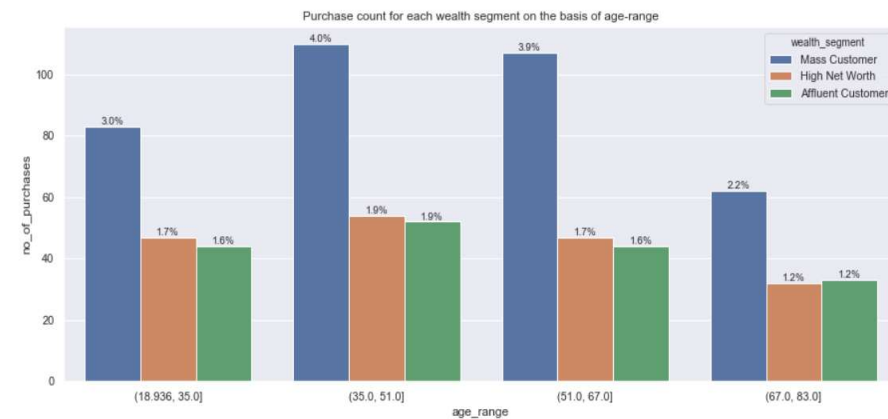
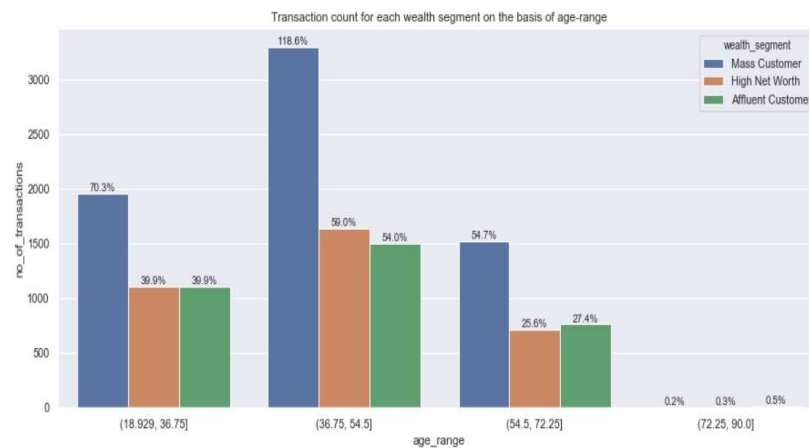


## Data Exploration

### Wealth – segment basis :

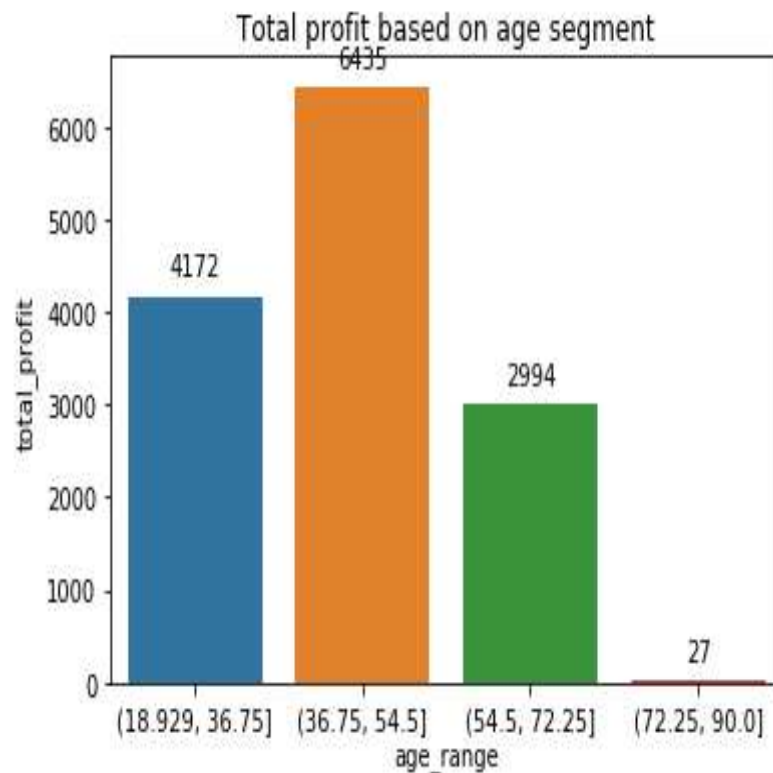
Here in this figure, we can see in all age groups, most customers of Sprockets falls under the Mass Customer wealth segment. Even the most transactions are made by the mass customer segment, specifically who comes in the age range of 37 to 55.

Thus, the marketing team should focus the segment and age range, who comes under mass customer segment and age range 37 to 55. In the new customer dataset also the mass customer segment is highest with the age range of 35 to 67.





## Data Exploration



### Total profit margin basis:

Here in this figure, we can see that the customers from age group 36 to 55 gave company most profit. They came back or purchased more products than any other age group.

Thus, on the basis of profit analysis, marketing team should give attention to the age group 36 to 55 and plan to make offers for them.

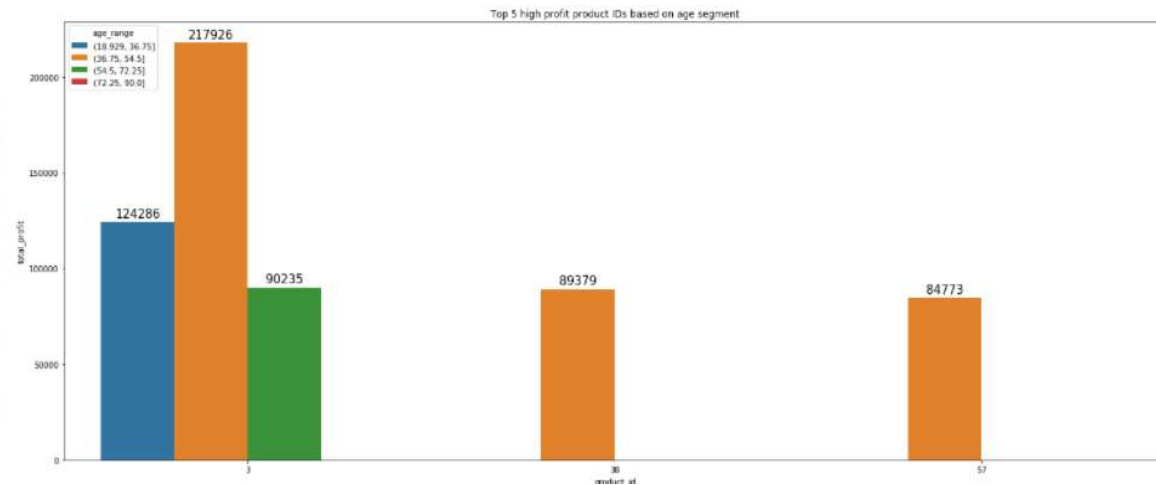
## Data Exploration

### Top five purchasing products on the base of age range:

Here in this figure, we can see that the customers from different age groups bought mostly product with the id-3. That means it is the most popular product among all. After product id 3, it is product id- 38 and 57 that had high on demand and gave company good profit.

Thus, marketing team should focus on these products, while making any offer or to attract more new customers.

	age_range	product_id	total_profit
104	(36.75, 54.5]	3	217926.40
3	(18.929, 36.75]	3	124286.15
205	(54.5, 72.25]	3	90235.15
139	(36.75, 54.5]	38	89378.52
158	(36.75, 54.5]	57	84773.00

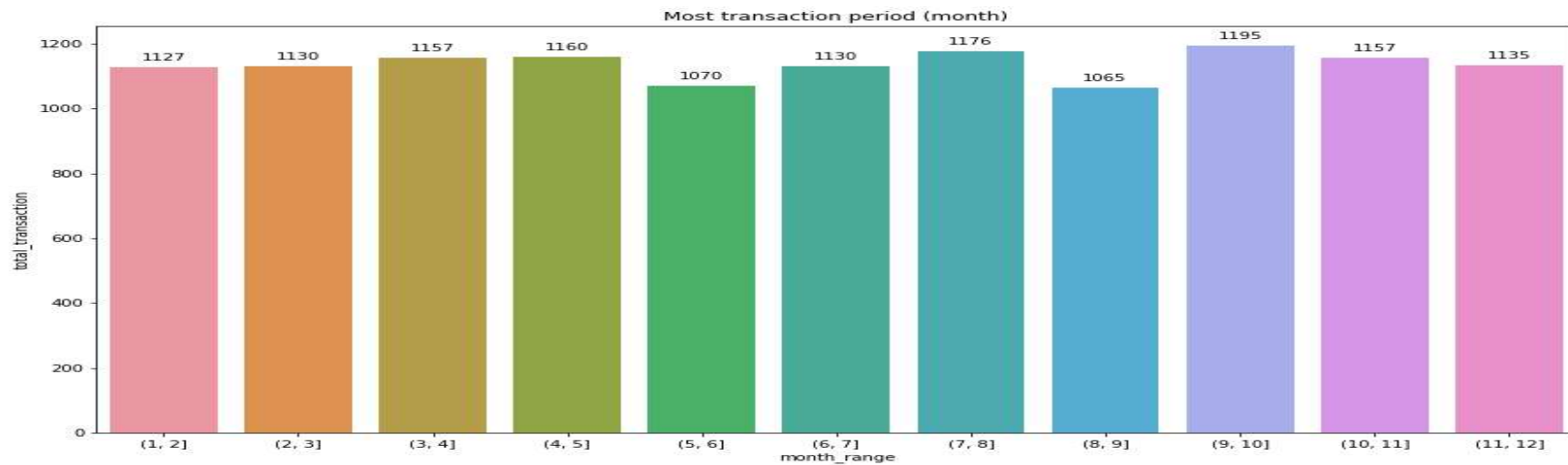


## Data Exploration

### Most transaction period (month):

Here in this figure, we can see that the customers made the most transactions in the 9 and 10 month, that means in September and October, comparatively to another months in the year.

So, the marketing team should focus on this time period to making any offer or to attract more new customers.

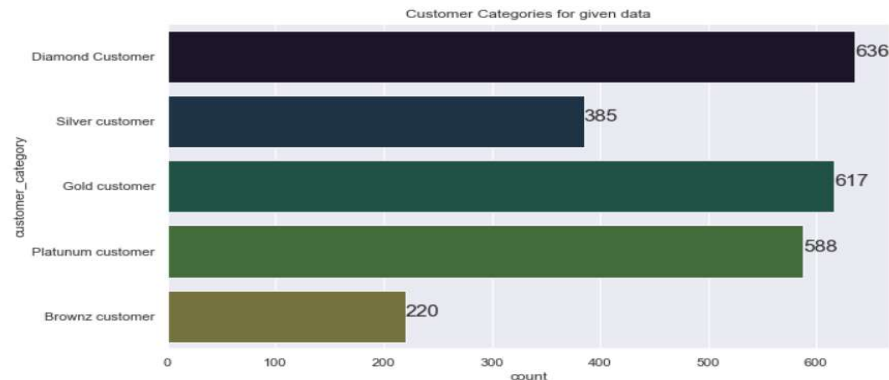


## Model development

### RFM Analysis and Modelling for Customer Segmentation:

RFM stands for Recency, Frequency and Monetary value, each of which corresponds to a key customer trait. These RFM metrics are important indicators of a customer's behaviour because frequency and monetary value affects a customer's lifetime value. Recency affects retention, a measure of engagement. This method is one of the most popular, easy to use segmentation method for analysing customers trend and behaviour.

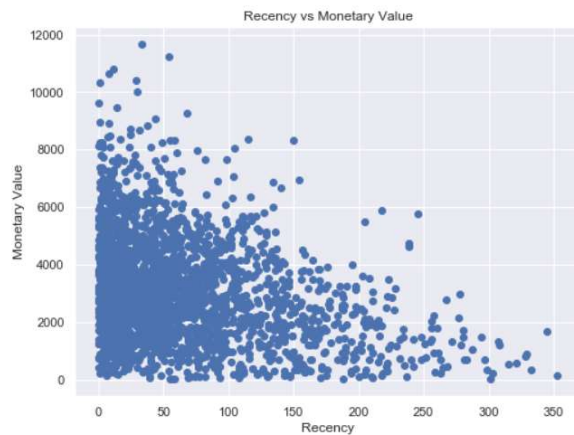
We will be using this model to cluster customers into separate classes on basis of the 3 KPI's used in this method – Recency (how recently has the customer made a purchase), Frequency (how many purchases/products has the customer made/bought) and Monetary value (Total profit from a customer's transaction(s)).



# Interpretation

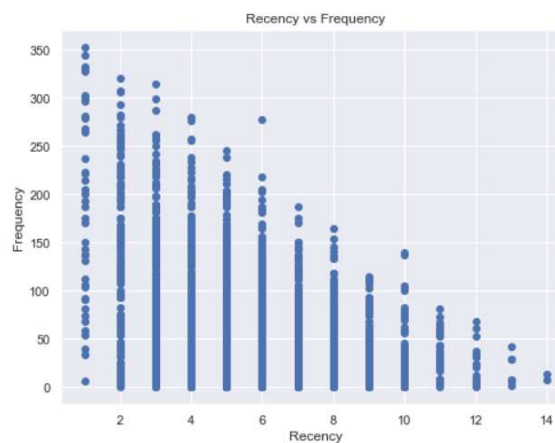
## Recency Vs Monetary Value:

From the relation between Recency and Monetary value, the Monetary Value is more for more Recent items i.e. Customers who have purchased more recently have generated more revenue/profit.



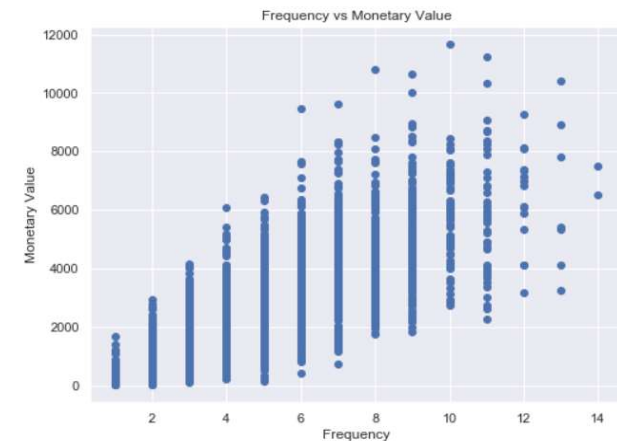
## Frequency Vs Monetary Value:

Relation between Frequency and Monetary value is quite linear, i.e. customers who have purchased more number of items, have provided more profit to Sprockets. Thus Customers classified as Platinum and Potential Loyalists who visit frequently correlate with increased profit for the business.



## Recency Vs Frequency:

There is sort of an inverse relationship between Recency and Frequency and rightfully so, since the customers who had visited long back have a high Recency value, and since they haven't returned in a long time, they have a low frequency value associated. Customers who visited more recently (0-50 days ago), have a higher frequency possibility.

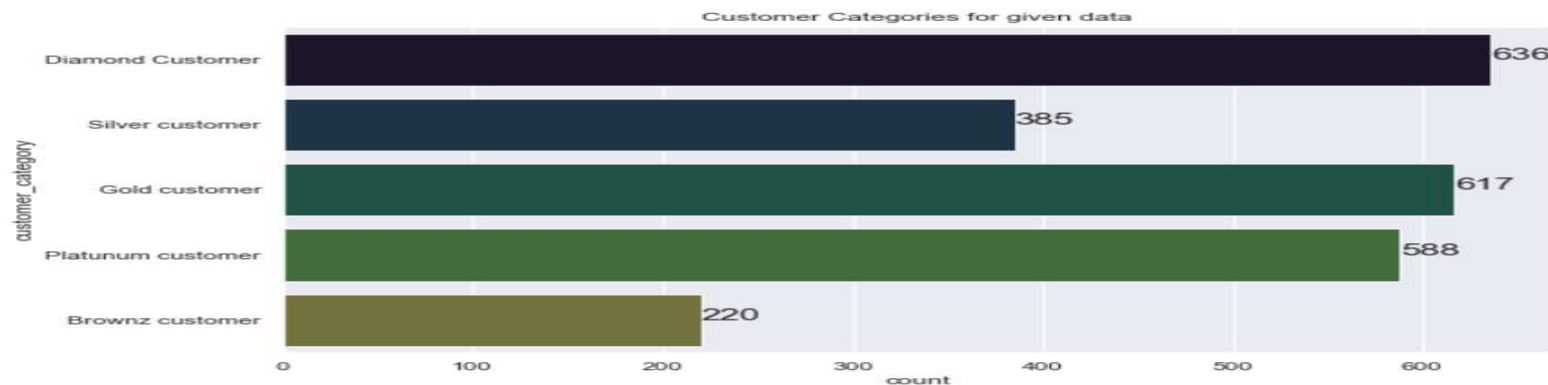


# Interpretation

## Customer category division:

According to RFM Model, all customers are segmented into 5 different categories –

1. Diamond customer: These are the customer, who brought profit to the company. These are the repeating and most recent and frequent customers, who made maximum transaction with the company.
2. Silver customer: These customers are also recent but not that much frequent as the diamond. But they also came back and made many transactions and brought profit to the company.
3. Gold customers: They are moderate customers, who once in a while comes back but no regular.
4. Platinum customer: These are the customers who bought good amount but didn't make any transaction recently or frequently.
5. Brownze customer: They are the ones, who made order but never came back.



## Conclusion

In the end, according to the analysis and report, the most valuable customers falls under age group 36-55 and from the job industry finance, health and manufacturing. In the state NSW, the customers has chances to respond on any offer or come back to make transaction.

I hope this evaluation and analysis report will help the company to make maximum profit and bring the valuable input to get more potential customers.

**Note:** The data and information in this document is reflective of a hypothetical situation and client. This document is to be used for KPMG Virtual Internship purposes only.

# Thank you