

TheAnalyticsTeam

Sprocket Central Pty Ltd

Data analytics approach

[Division Name] - [Engagement Manager], [Senior Consultant], [Junior Consultant]

Agenda

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

Introduction

IDENTIFY AND RECOMMEND TOP 1000 CUSTOMER FOR FURTHER MARKETING TARGETING

Outline of Problem

- The Marketing team from Sprocket Central Pty Ltd is looking to boost business by analyzing their existing customer dataset and predict the customer trend and behavior.
- There are 3 datasets given by the company
- We are going to do the Data Analysis through 3 phases of Data Exploration, Model Development and Interpretation, and recommend 1000 potential customers to the company

Content of this Analysis

- ✓ Sales per Brand based on Product Line
- ✓ New and Old customer Age Distribution
- ✓ New and Old customer distribution based on job industry
- ✓ Number of customer in each state
- ✓ Percentage of New and Old customer Bike Purchased by Gender
- ✓ Number of Car Owned and Not Owned in each state
- ✓ Age, Wealth and Profit Analysis
- ✓ RFM Analysis and Customer Profile Distribution

Data Exploration

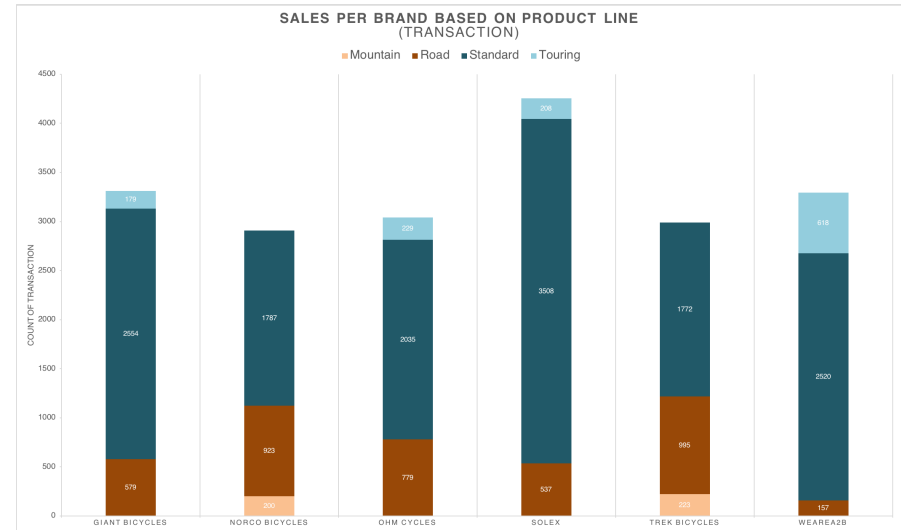
DATA QUALITY ASSESSMENT AND MITIGATION

	Customer Demographic	Customer Address	Transaction Data
Accuracy	<ul style="list-style-type: none">• DOB: inaccurate• Age: missing		<ul style="list-style-type: none">• Profit: missing
Completeness	<ul style="list-style-type: none">• Customer ID: incomplete• Last name: blanks• Job title, industry: blanks	<ul style="list-style-type: none">• Customer ID: incomplete	<ul style="list-style-type: none">• Customer ID: incomplete• Online order: blanks• Brand: blanks
Consistency	<ul style="list-style-type: none">• Gender: inconsistency	<ul style="list-style-type: none">• State: inconsistency	
Currency	<ul style="list-style-type: none">• Deceased customer: some are deceased		
Orderliness		<ul style="list-style-type: none">• Postcode: format general to number• Property value: format to number	<ul style="list-style-type: none">• List price: format number to currency• Product first sold: format number to date
Relevancy	<ul style="list-style-type: none">• Default column: hided		<ul style="list-style-type: none">• Order status: Cancelled filtered out

Data Exploration

SALES PER BRAND BASED ON PRODUCT LINES

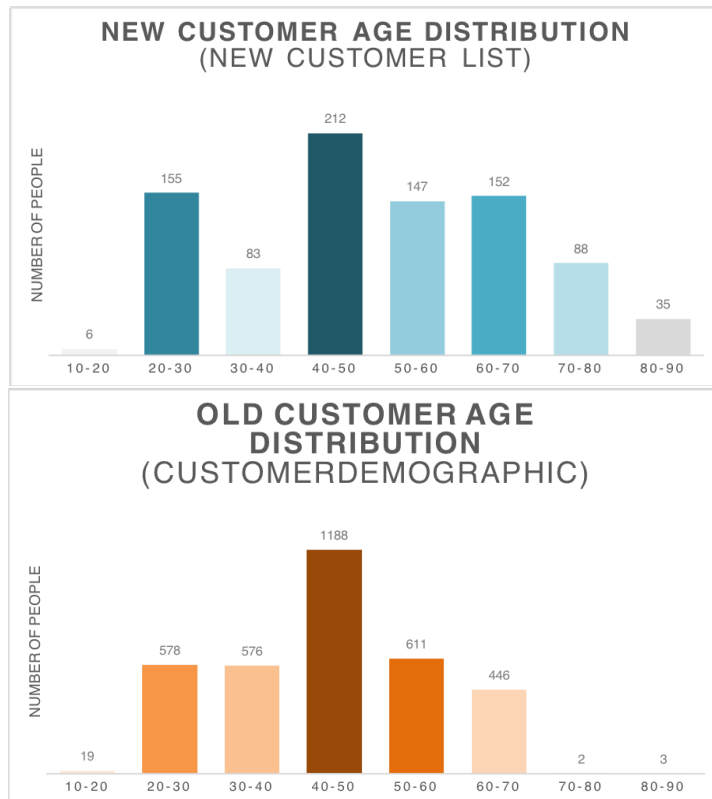
- From the observation, the 'Standard line' has reached the most product sold for all the brand, followed by 'Road line'.
- The least sold product line are 'Mountain' and 'Touring'.
- The Most popular brand among the customer is 'Solex'.
- Numerically, the other Bike Brands distributed fairly good amount to the sales with a few hundred bikes differences.



Data Exploration

NEW AND OLD CUSTOMER AGE DISTRIBUTIONS

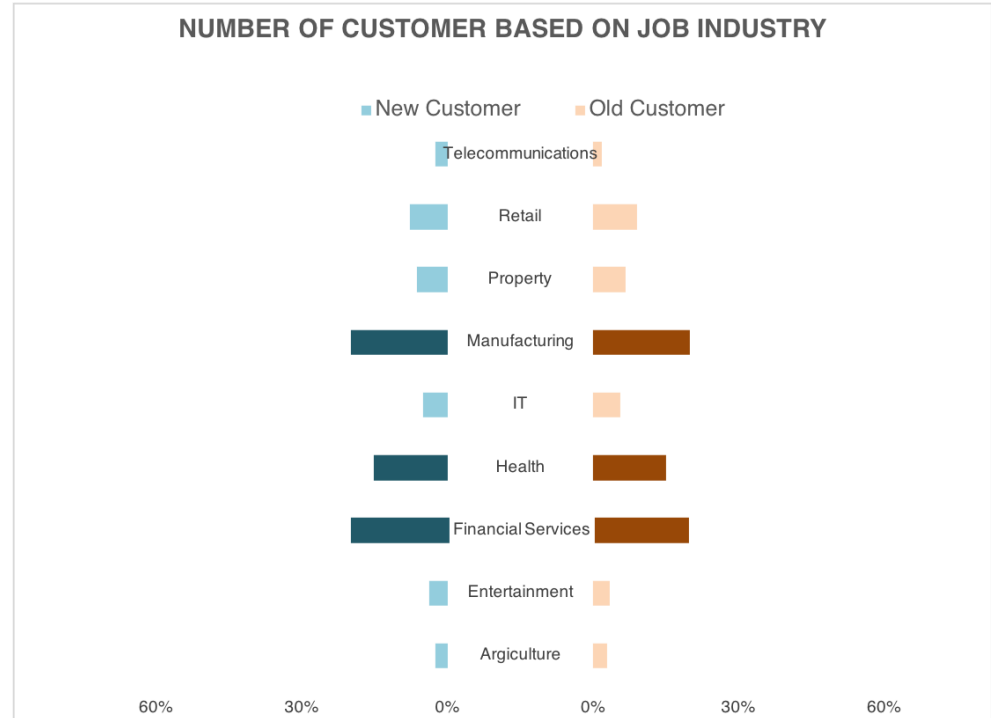
- The largest number of customer are aged between '40-50' for both New and Old customer.
- The lowest age group are '10-20' and '80-90' for both groups, with New group having more customer at the aged between '70-80' compared to the Old group.
- There is a steep drop of customers in the age of '30-40' in the New group.



Data Exploration

NEW AND OLD CUSTOMER DISTRIBUTION BASED ON JOB INDUSTRY

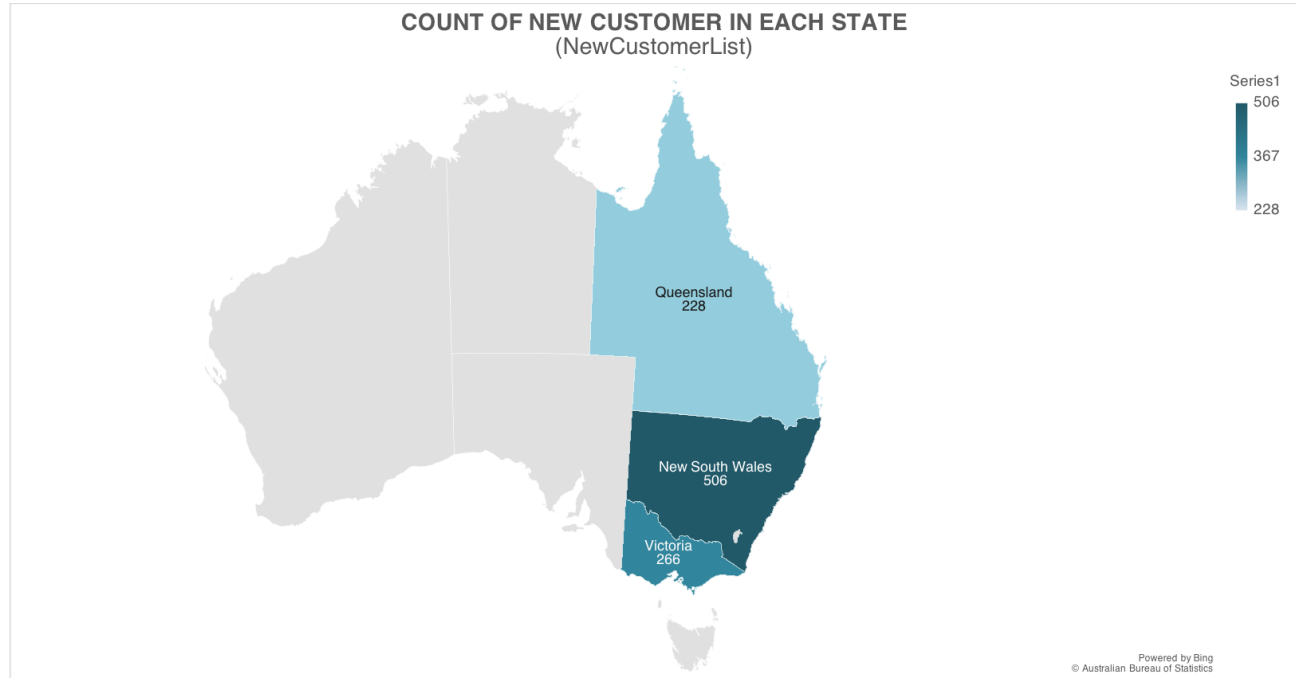
- From both datasets, we see almost the same pattern regarding the job industry distribution among customer.
- Most of the customer from both 'New' and 'Old' groups are belong to 'Financial Services', 'Health' and 'Manufacturing Sector'.
- All other industries stand only a small portion of customer.



Data Exploration

NUMBER OF CUSTOMER IN EACH STATE

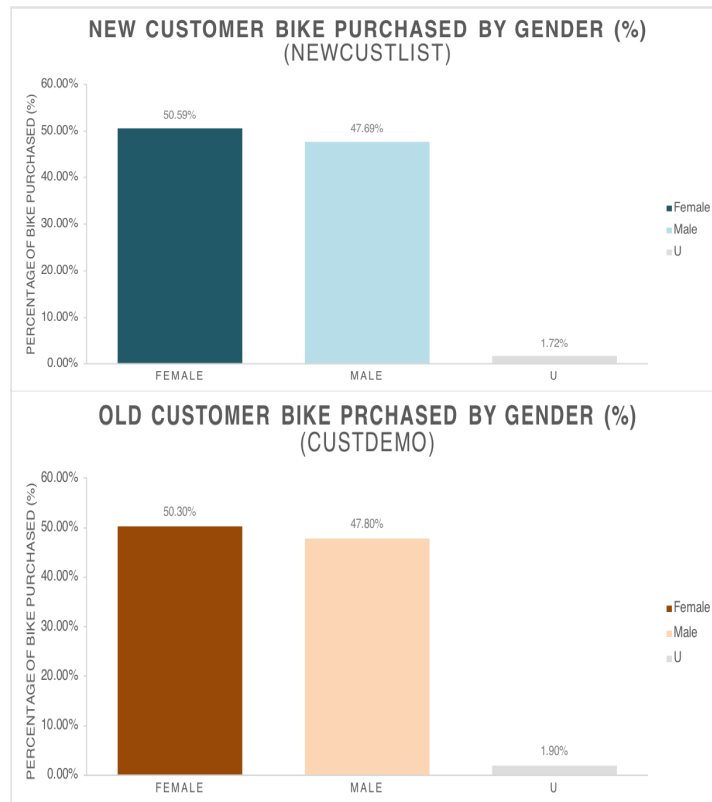
- The new customers belong to the three states of Australia, Queensland having 228 potential customers, Victoria having 266 customers and New South Wales having 506 customers.
- Result in NSW ranking the most potential state for customer targeting.



Data Exploration

NEW AND OLD CUSTOMER BIKE PURCHASED BY GENDER

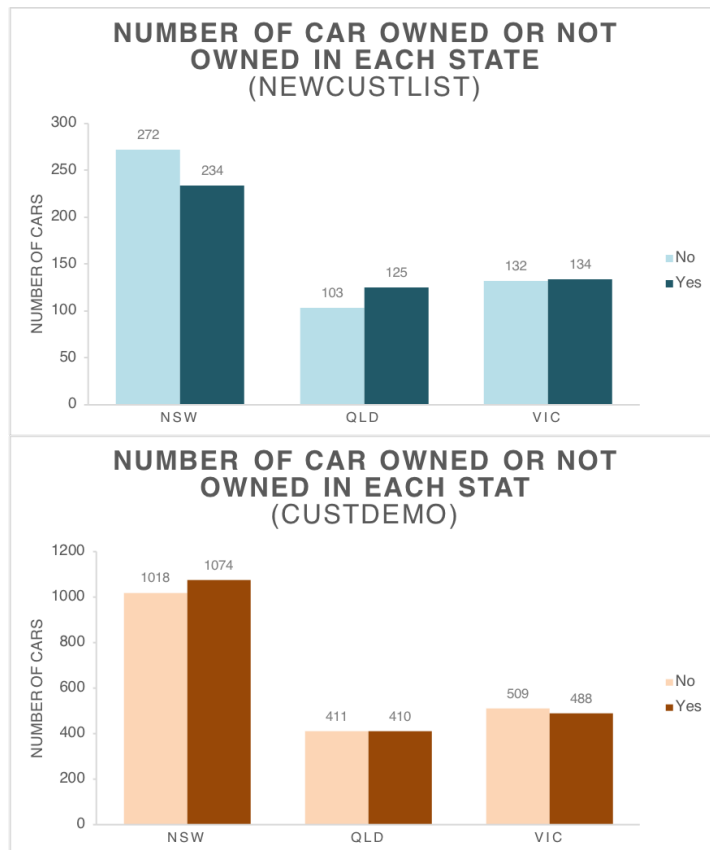
- Almost the same pattern demonstrated in the graphs of bike purchased by Gender, with about 50% were 'Females', and around 48% were 'Males', while about the other 2% were made by 'Unknown' gender.
- Numerically, 'Females' purchased almost 5000 bike more than 'Males' for Old customer dataset. (Appendix 1)
- While from the New customer dataset, 'Females' made around 1500 more bike purchased than 'Males'. (Appendix 2)



Data Exploration

NUMBER OF CAR OWNED OR NOT OWNED IN EACH STATE

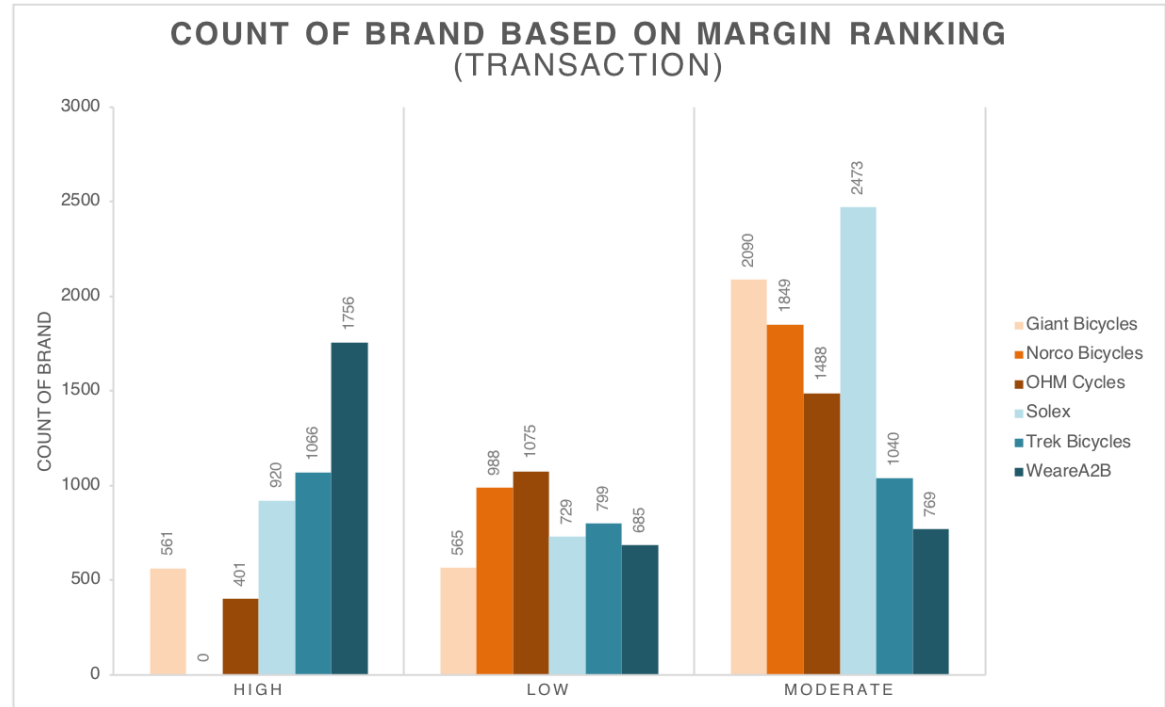
- The distributions from VIC and QLD are spread almost evenly in both customer groups with just a few number differences.
- With the dataset we got, mostly are from NSW, people from this state seems to decreased their car owning making the car owned rate falling lower than those who do not owned in the New customer group.



Model Development

COUNT OF BRANDS BASED ON THREE LEVELS OF MARGIN RANKING

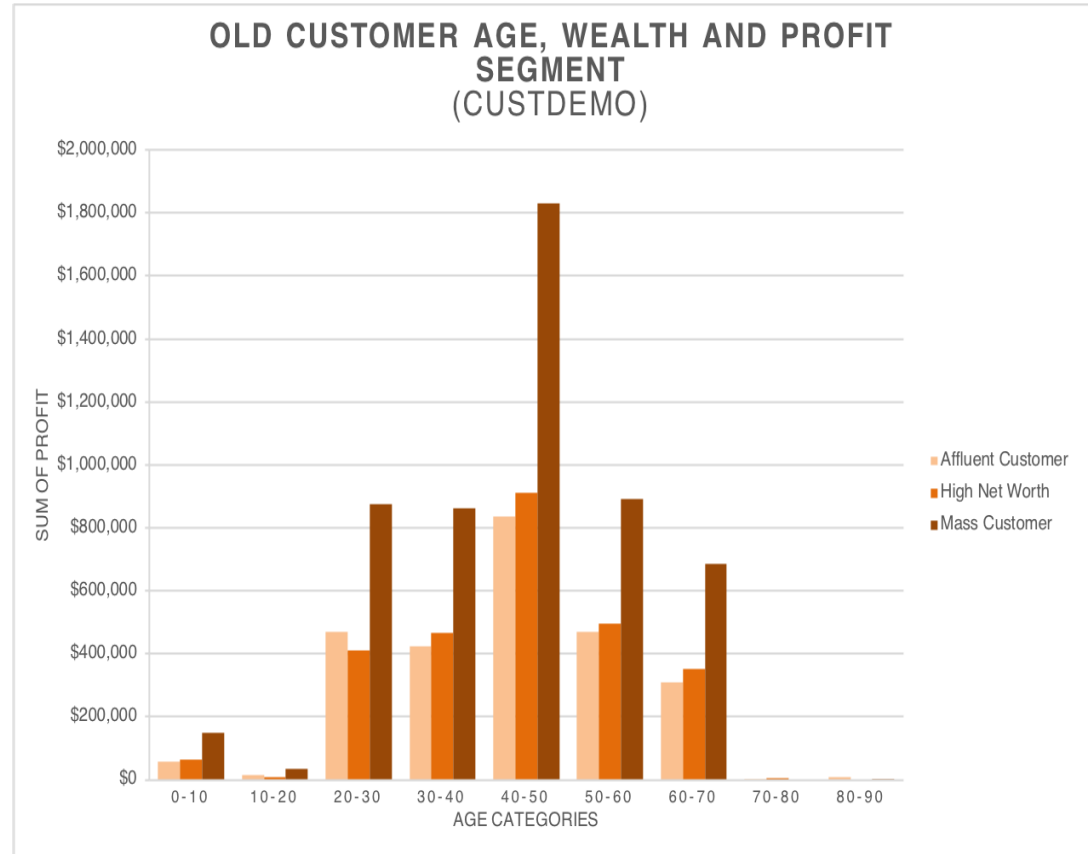
- The company made mostly 'Moderate' margin based on the profit that made from selling all brands of bike.
- Within those of 'Moderate' margin, the brand 'Solex' contributed the largest proportion.
- The 'High' margin come at second place where the company generated most High margin profit from the brand 'WeareA2B'.
- Hence, while trying to increasing the sale of 'Solex', the company should also focus on Increase the sales of 'WeareA2B' as it generated most the high profit.



Model Development

AGE, WEALTH AND PROFIT ANALYSIS

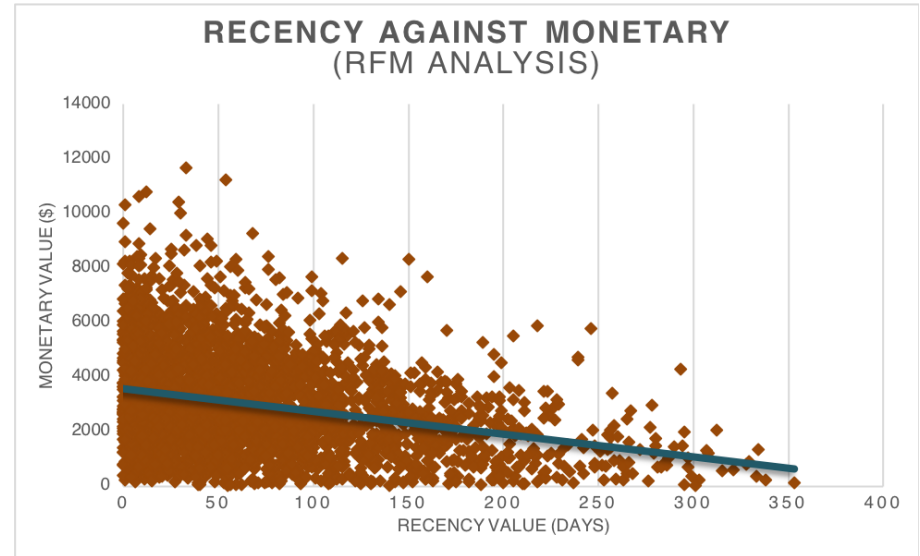
- Large amount of the Transaction happened to be the 'Mass Customer'.
- Customer Aged between '40-50' has the most profit distribution to the company's sales.
- The profit followed by the other age groups of '20-40' and '50-70' with also having 'Mass Customer' as the largest proportion.



Model Development

RFM ANALYSIS – SCATTER-PLOT (1)

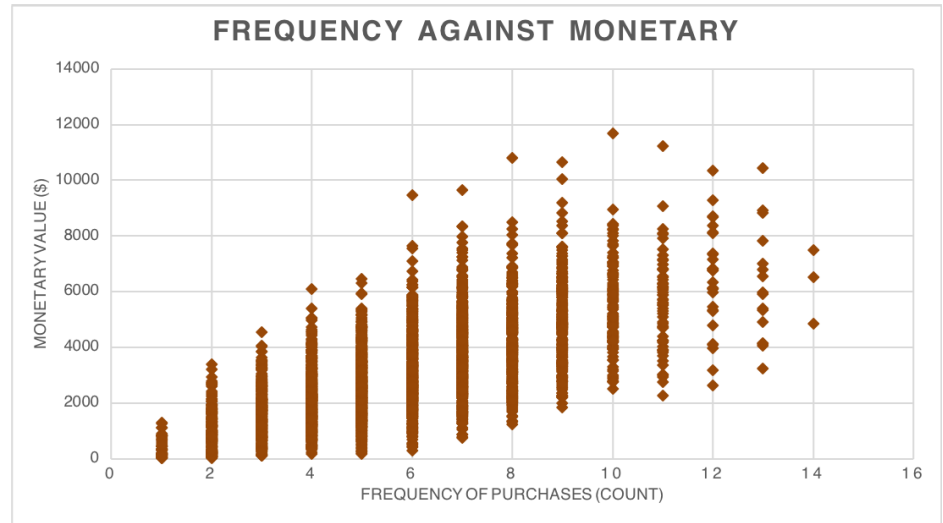
- The scatter-plot of Recency against Monetary has demonstrated that the customer who made purchase more recently have generated more revenue to the company, compare to the customer made purchase long ago.
- Customer from recent past at '50 to 100' days also generated a moderate amount of revenue.
- There are a few high revenues which are generated from the recent customer as well.
- From the pattern, customer from the past tend to spend less on bike purchases compare to recent days and vice versa when the profit keep increasing based on recency.



Model Development

RFM ANALYSIS – SCATTER-PLOT (2)

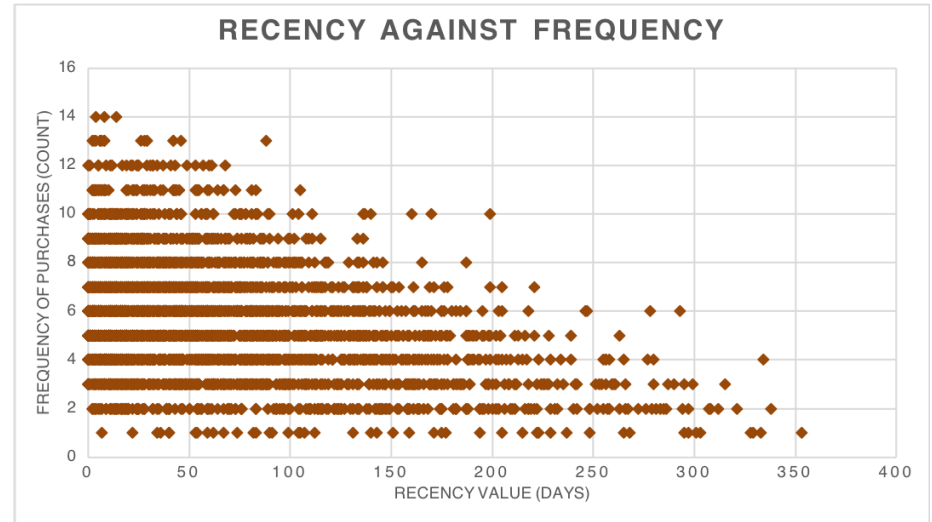
- There is a positive relationship between Frequency and Monetary, which result in more profit being generated when customer made more frequent purchases.



Model Development

RFM ANALYSIS – SCATTER-PLOT (3)

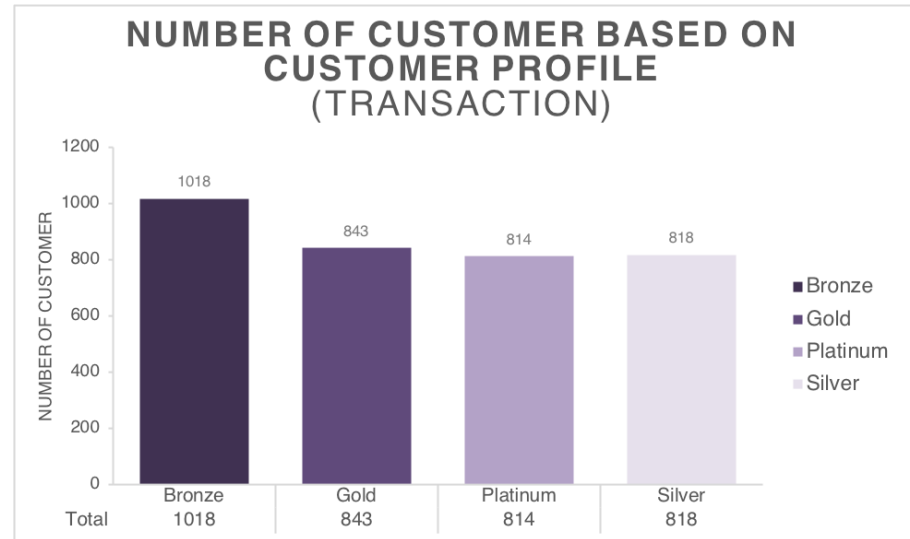
- Customers that have made purchases more recently at the latest '0-50' days have higher chance of making more frequent purchased.
- There is a negative relationship between Recency and Frequency, meaning the customer who made purchases the most recently are also the more frequently purchased customer.



Model Development

RFM ANALYSIS - CUSTOMER PROFILE DISTRIBUTION

- RFM Analysis is used to determine the most valuable customers a business should target to maximize its revenue and value.
- The RFM Analysis conducted from three values, Recency, Frequency and Monetary, which demonstrated how loyal and active the customer is in engaging with the business.
- Based on the dataset of Product Transaction, we can tell that the greatest number of customer are still at Bronze, however, the other rank of profile has quite the similar number which mean the company still has quite a great number of loyal and active customer.



Interpretation

CUSTOMER CLASSIFICATION – TARGETING HIGH VALUE CUSTOMERS

From the above Analysis and Segments, there are a few patterns which we can use to group the customer with the high value to the company, including:

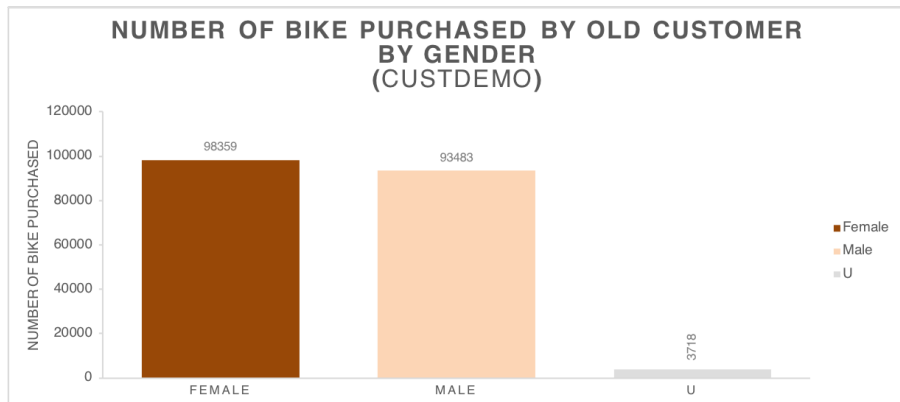
- Aged between '40-50' with 'Mass' wealth segment
- Who are currently living in 'NSW', followed by 'VIC'
- Most of those will be 'Females'
- Those working in 'Financial', 'Health' and 'Manufacturing' industry sector.

first_name	past_3_years_bike_rel	age	job_industry_c	owns_car	tenure
Melba	38	45	Health	No	4
Winnifred	83	45	Financial Services	No	14
Gale	59	44	Financial Services	Yes	17
Martelle	52	41	Manufacturing	No	9
Chanda	15	50	Manufacturing	Yes	13
Patricia	34	44	Health	No	19
Daryl	12	42	Financial Services	Yes	12
Sunny	90	47	Financial Services	No	11

Appendix

Appendices

Appendix 1



Appendix 2

