

TheAnalyticsTeam

Sprocket Central Pty Ltd

Data analytics approach

Amarachi Ugagbe

AGENDA

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

INTRODUCTION

Identification and Recommendation of top 1000 target customers from given datasets.

Problem Outline:

- Sprocket Central Pty Ltd is a medium sized organization that specializes in sales of bike and cycling accessories.
- They are interested in getting data-driven insights that will boost their business growth.
- With the provision of data sets, they are to analyze and recommend the top 1000 customers who should be targeted to drive higher business values for the company.

Data Analysis Content:

- **New and Old Customer Age distributions**
- Bike related purchases over the last 3 years by gender
- Wealth segmentation by age category
- Job Industry distributions
- Number of cars owned and not owned by state
- RFM analysis and customer classification.

This analysis will go through some phases as: Data Exploration; Model Development and Interpretation.

DATA EXPLORATION

Data Quality Assessment

Basic Data Quality Issues

- Accurate: Correct Values are imputed
- Complete: Data fields are with values
- Consistent: Values are free from contradictions
- Current: Data Values are up to date
- Valid: Values imputed are the right and allowed values
- Relevant: Data items are with value meta-data
- Unique: Data sets are not duplicated.

Table to show summary of work done

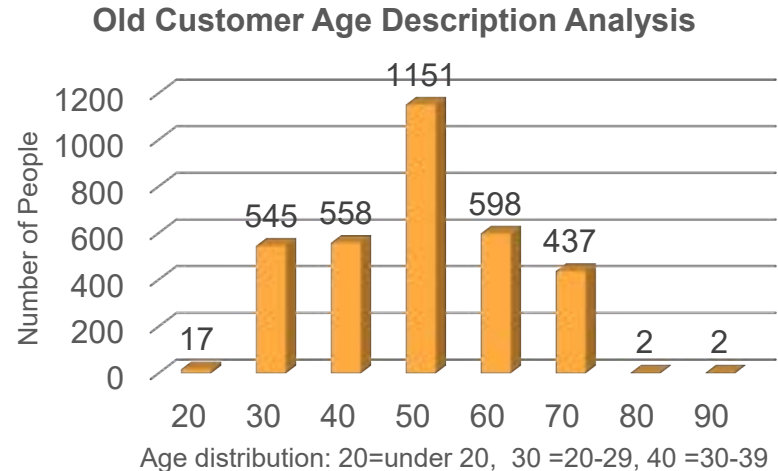
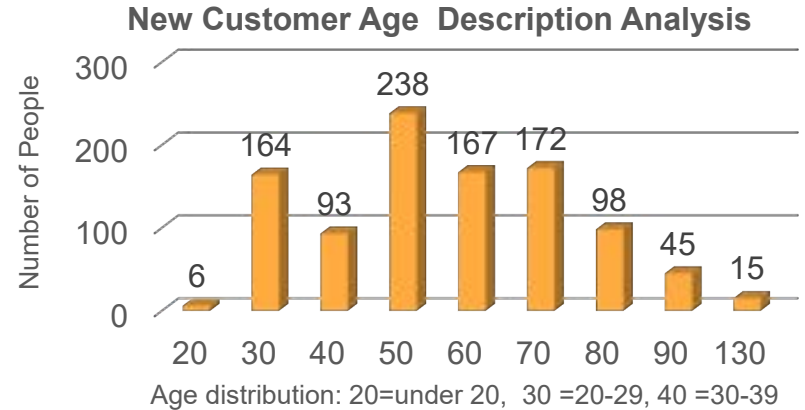
	Accurate	Complete	Consistent	Current	Relevant	Valid
Customer Demographic	Last Name Date of Birth	Job title had blank spaces, Customer ID incomplete	Gender was inconsistent	Filtered out deceased column	Default column was deleted	
Customer Address		Customer ID incomplete, online order column had blanks	States inconsistent			
Transactions	Profit column is missing,	Customer ID is incomplete, Brand is black			Cancelled Order Status is missing	List price format

An email has already been sent with respect to this Data Quality.

DATA EXPLORATION

New and Old Customer Age distributions

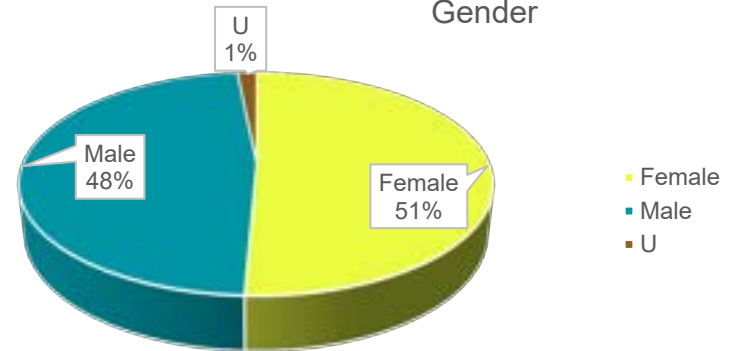
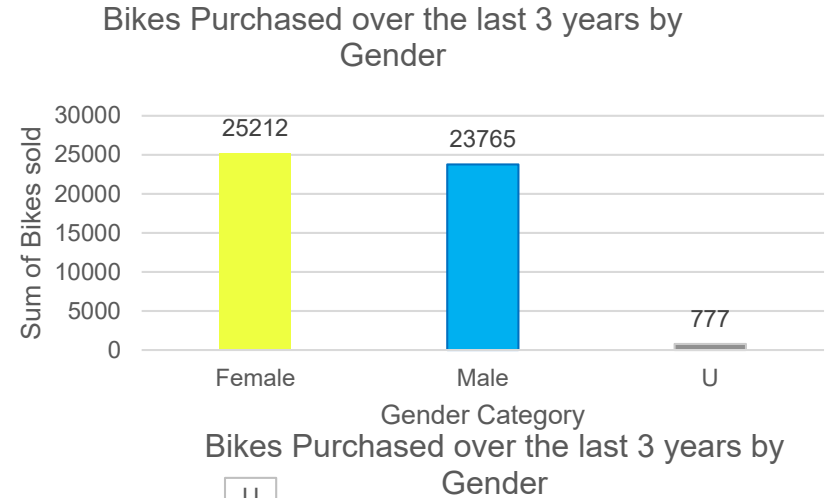
- In the New Customer Age Distributions, most customers are aged between 40-49 years old. Also in the old customer age distribution, most customers are aged between 40-49 years old.
- For New customers, least age groups are under 20 years old. While for Old Customers the least age grade is over 80 years and also under 20 years.
- New Customer list, the most populated age grades are: 20-29, 40 -69 years.
- Old Customer list, the most populated age grades are: 20-69 years old.
- There is a sharp decline of customers in the new customer list in the 30-39 age category.



DATA EXPLORATION

Bike related purchases over the last 3 years by gender.

- For the last three years, about 51% of bike related purchases were made by Females in comparison to 48% of purchase by Males. About 1% of the purchases were made by unknown gender.
- Females purchased 1447 bikes more than males.
- Generally, females have the majority of bike related sales.

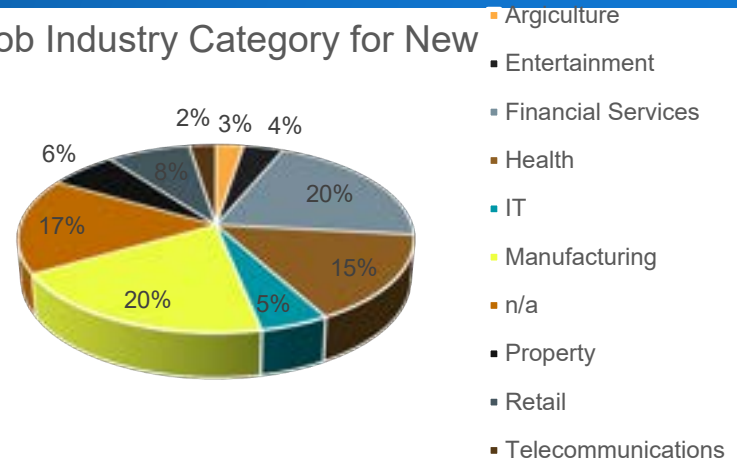


DATA EXPLORATION

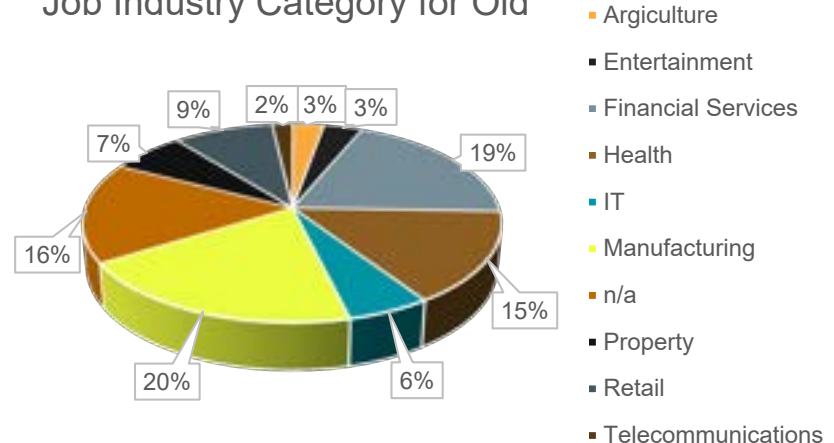
Job Industry Distribution

- For the New Customer list, The customers from Manufacturing and Financial Services gave the highest at 20% each.
- For the Old Customer list, The customers from Manufacturing and Financial Services gave the highest at 20% and 19% respectively.
- The smallest number of customers are in Agriculture and Telecommunications category in both Old and New Customers, at 2% each.

Job Industry Category for New



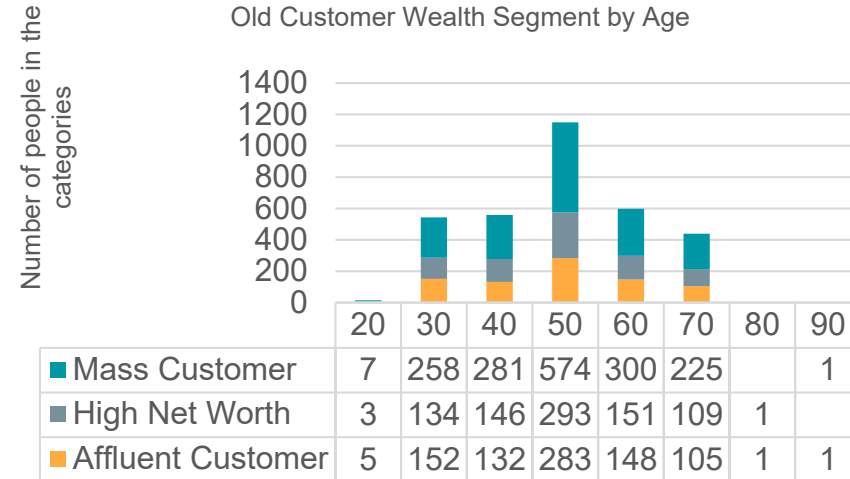
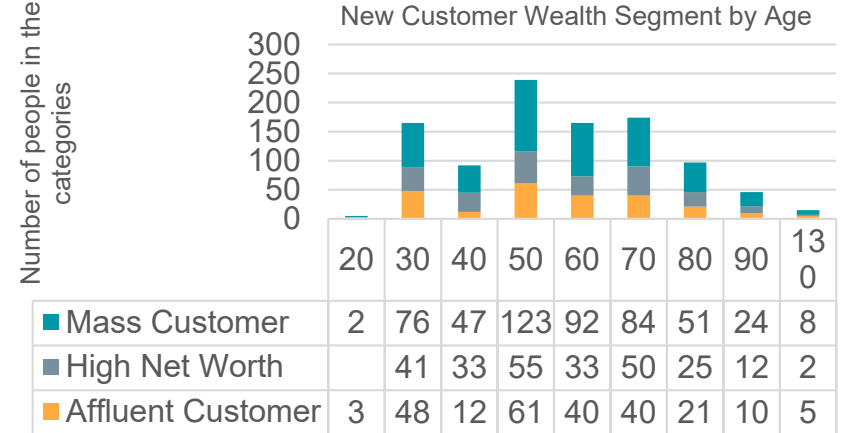
Job Industry Category for Old



DATA EXPLORATION

Wealth Segmentation by Age Category

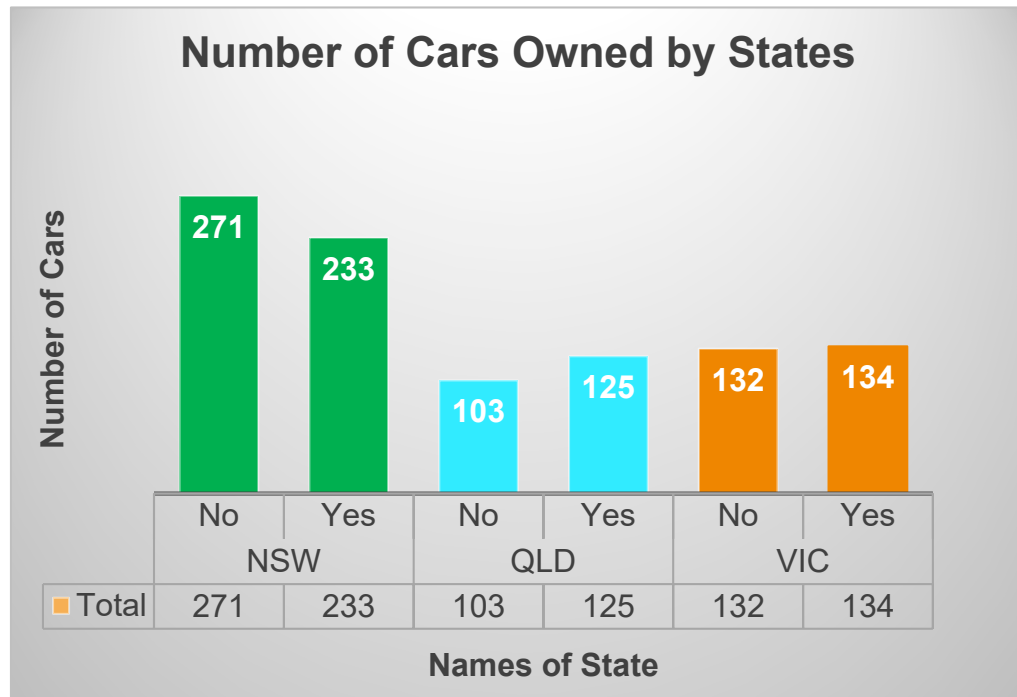
- In all the Age Categories both New and Old Customers, the highest number of customers are under the “Mass Customer” Segment.
- The next Segment with second highest customer is the “High Net Worth”.
- The Affluent Customers outperformed the High net Worth customers in the ages of 40-59 years.



DATA EXPLORATION

Car Owners Analysis

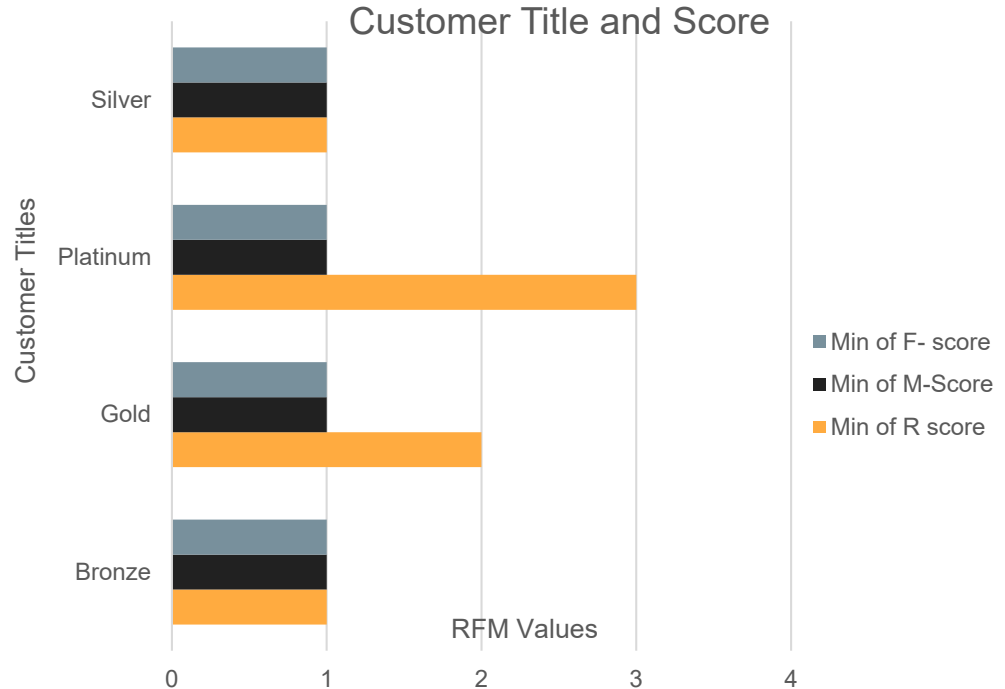
- NSW has the highest number of people who do not own a car. It also has the highest number of people that perhaps visited the company and data was collected.
- VIC has an almost evenly shared number of those that own a car and those that do not own a car.
- QLD is the state that has the highest number of customers who are car owners.



MODEL DEVELOPMENT

RFM Analysis and Customer Segmentation

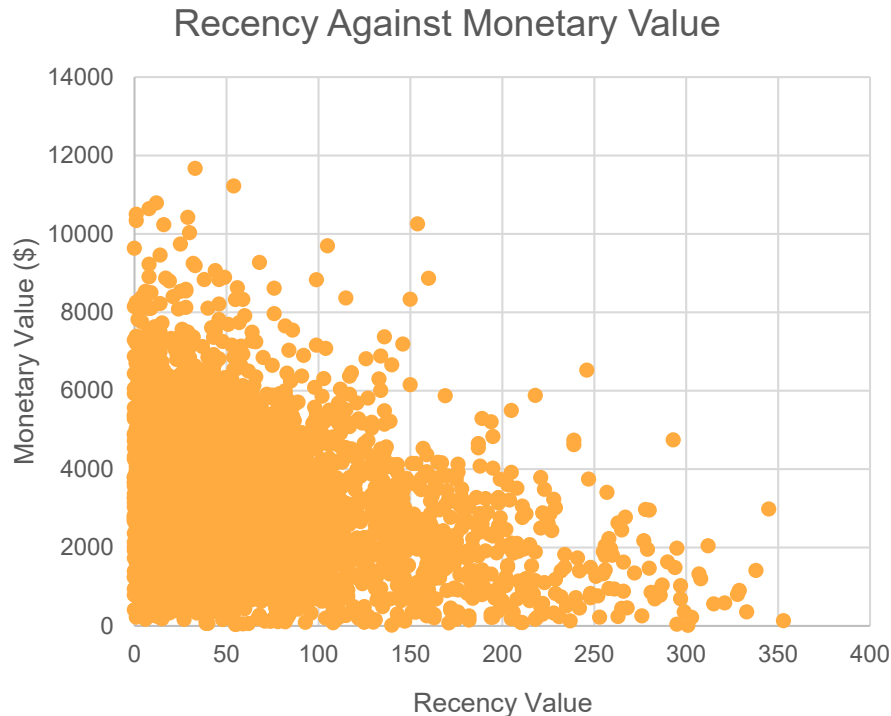
- To identify the best customers and perform targeted marketing campaigns, companies used a marketing technique called **RFM Analysis**, which ranks and groups customers based on the **Recency**, **Frequency** and **Monetary** total of their recent transactions.
- This RFM Model shows the customers that has high level of engagements with the business with respect to how recent their transaction is, how frequent they make the transactions and how much was spent in that given period.



MODEL DEVELOPMENT

Scatter-Plot based on RFM Analysis

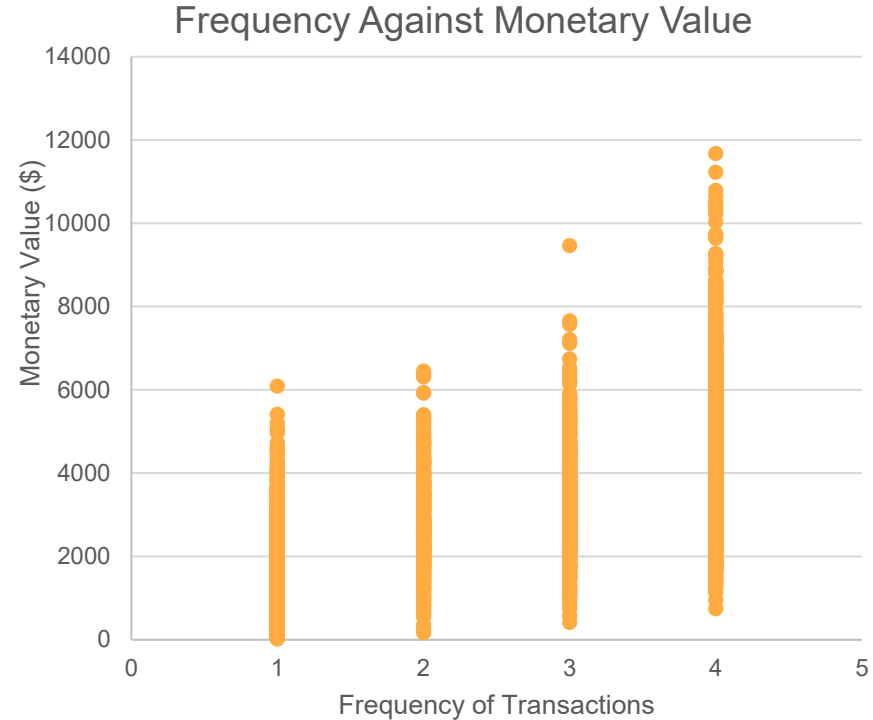
- This chart shows that the customers that are recent in their purchase made more or high monetary impact on the business, than those that visited a long time ago.
- Those that visited from 0-100 days made more monetary impact (generated more revenue) on the business/company.
- Those that visited more than 200 days ago generated low revenue for the business/company.



MODEL DEVELOPMENT

Scatter-Plot based on RFM Analysis

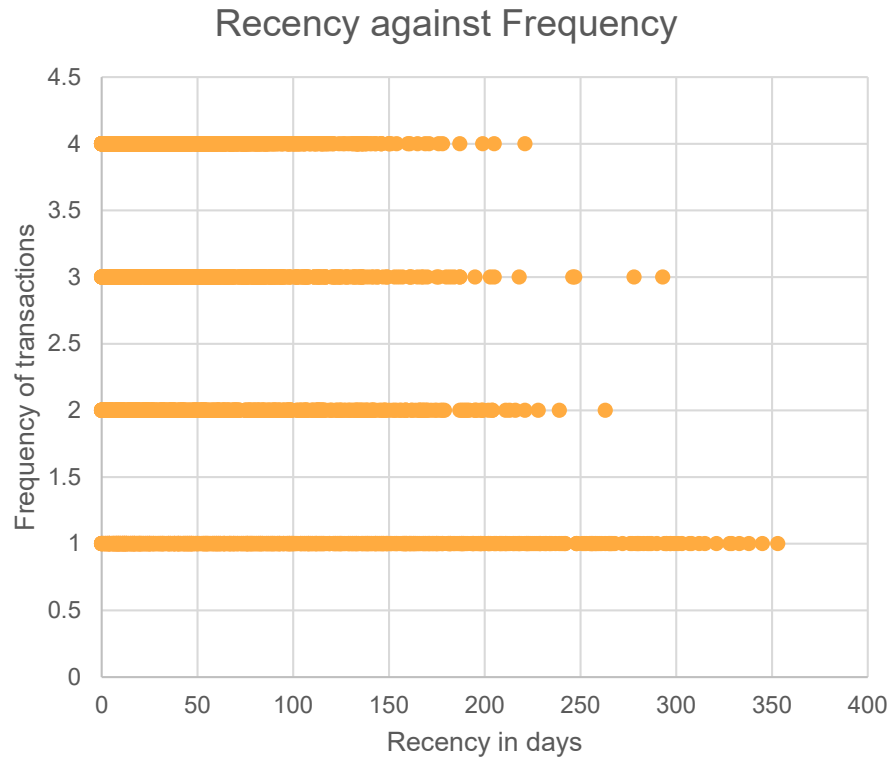
- Platinum and Gold customers visit frequently and this correlates with increased revenue for the business.
- There is a positive relationship between frequency and monetary gain of the business.



MODEL DEVELOPMENT

Scatter-Plot based on RFM Analysis

- Low frequency (0-1) correlated with high recency values (350 days)
- The shorter the number of days of visit, the higher the frequency.
- Very recent customers are also very frequent customers.



MODEL DEVELOPMENT

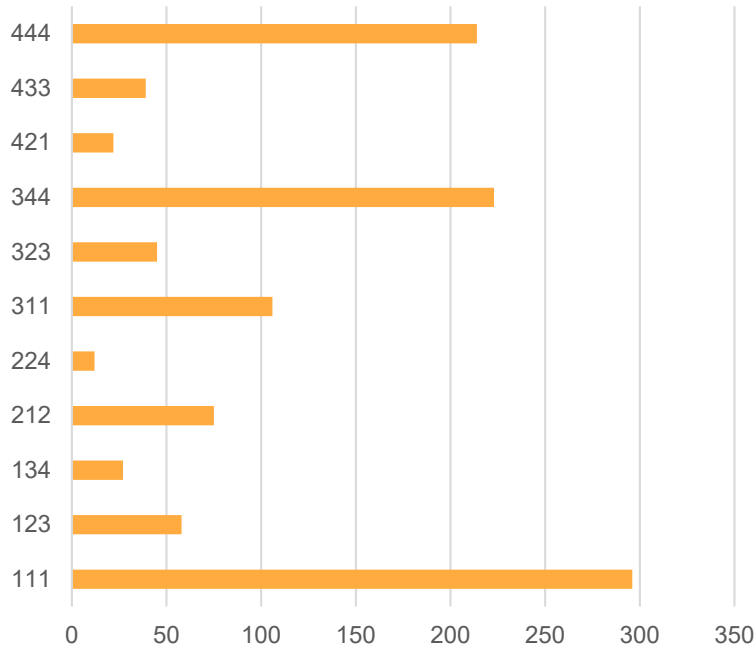
Customer Tittle List as given by assigned RFM Value

Rank	Title	Description	RFM Values
1	Platinum Customer	Most recent buyer, buys often and spends most	444
2	Very Loyal	Most recent buyer, buys often and spends large amount of money	433
3	Becoming loyal	Relatively recent, bought more than once, spends large amount of money	421
4	Recent Customer	Bought recently, not very often, average money spent	344
5	Potential Customer	Bought recently, never bought before, spent small amount	323
6	Late Bloomer	No recent purchase, but RFM value is more than average	311
7	Losing Customer	Purchase was a while ago but below average RFM value.	224
8	High Risk Customer	Purchase was made long time ago, frequency is high , amount spent is high	212
9	Almost Lost Customer	Very low recency, low frequency, but high amount spent	134
10	Evasive Customer	Very low recency, low frequency, small amount spent	123
11	Lost Customer	Very low RFM	111

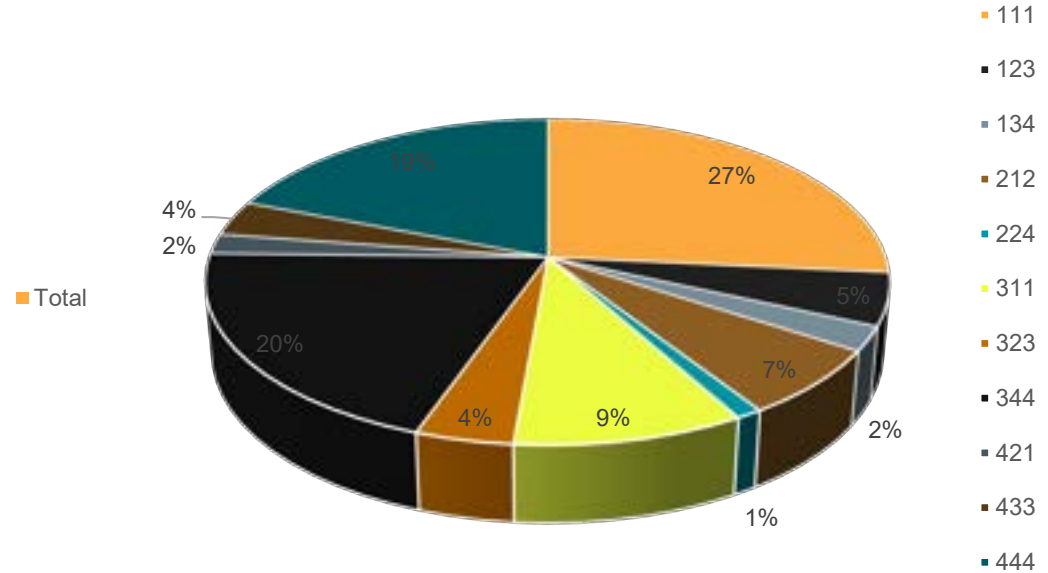
MODEL DEVELOPMENT

Customer Tittle Distribution by assigned RFM Value

Total



Total



INTERPRETATION

Summary Table of top 1000 Customers to Target

Rank	Title	Description	Number of Customers	Cumulative	Customer Selection
1	Platinum Customer	Most recent buyer, buys often and spends most	431	431	431
2	Very Loyal	Most recent buyer, buys often and spends large amount of money	243	674	243
3	Becoming loyal	Relatively recent, bought more than once, spends large amount of money	417	1091	417
4	Recent Customer	Bought recently, not very often, average money spent	360	1451	0
5	Potential Customer	Bought recently, never bought before, spent small amount	291	1742	0
6	Late Bloomer	No recent purchase, but RFM value is more than average	528	2270	0
7	Losing Customer	Purchase was a while ago but below average RFM value.	202	2472	0
8	High Risk Customer	Purchase was made long time ago, frequency is high , amount spent is high	301	2773	0
9	Almost Lost Customer	Very low recency, low frequency, but high amount spent	170	2943	0
10	Evasive Customer	Very low recency, low frequency, small amount spent	140	3083	0
11	Lost Customer	Very low RFM	411	3494	0

INTERPRETATION

Customers Target and Methodology

Rank	Title	Description	Number of Customers	Cumulative	Customer Selection
1	Platinum Customer	Most recent buyer, buys often and spends most	431	431	431
2	Very Loyal	Most recent buyer, buys often and spends large amount of money	243	674	243
3	Becoming loyal	Relatively recent, bought more than once, spends large amount of money	417	1091	326

The Customers that make up the top 1000 are those that have bought recently, buys frequently and spends much money while buying.

Thank You.