

TheAnalyticsTeam

Sprocket Central Pty Ltd

Data analytics approach

Gaurav Singh

Agenda

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

Introduction

Tools

- **Microsoft Excel**
- **Tableau**



Introduction

Recommend Top 1000 Customers from the given Data

- Sprocket Central Pty Ltd specialises in high-quality bikes and cycling accessories.
- Their marketing team is looking to boost business by analysing their existing customer dataset.
- Using the existing 3 datasets, recommend 1000 new customers that Sprocket Central should be targeted to drive the most value for the organisation.

Data Exploration

Data Quality Assessment/Data Cleaning

	Accuracy	Completeness	Consistency	Currency	Relevancy	Validity
Customer Demographic	DOB: inaccurate, Age: missing	Job_title: blanks customer_id: missing	Gender: inconsistency	Deceased customers: filter out	Default: deleted	
Customer Address		Customer_id: missing	States: inconsistency			
Transactions	Profit: missing	Customer_id: missing, online_order: blanks, brands: blanks	Gender: inconsistency	List_price, Standard_cost, profit		List price: format, Product_first_sold_date: format

Standard Data Quality Dimensions

Correct Values

Accuracy

Data Fields with Values

Completeness

Values Free from Contradiction

Consistency

Values up to Date

Currency

Data Items with Value Meta-data

Relevancy

Data Containing Allowable Values

Validity

Records that are Duplicated

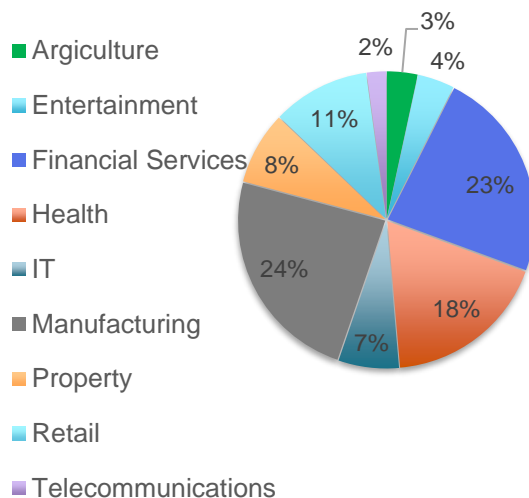
Uniqueness

Data Exploration

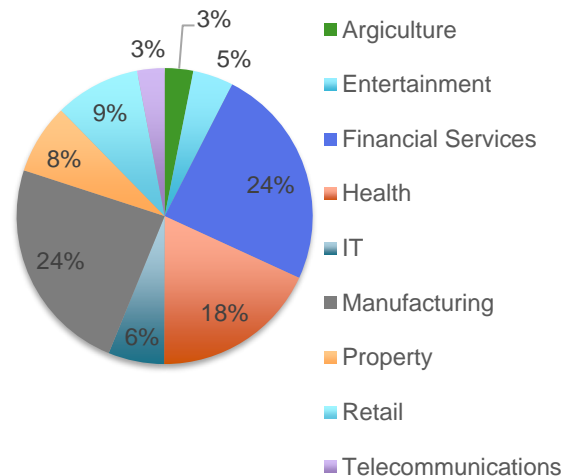
Job Industry Distribution

- 20% of new customers are in Manufacturing and Financial Services.
- The least number of customers i.e. 3% are in Agriculture and Telecommunications.
- Similar pattern in existing customers, 24% and 23% are in Manufacturing and Finance Services.

Existing Customers



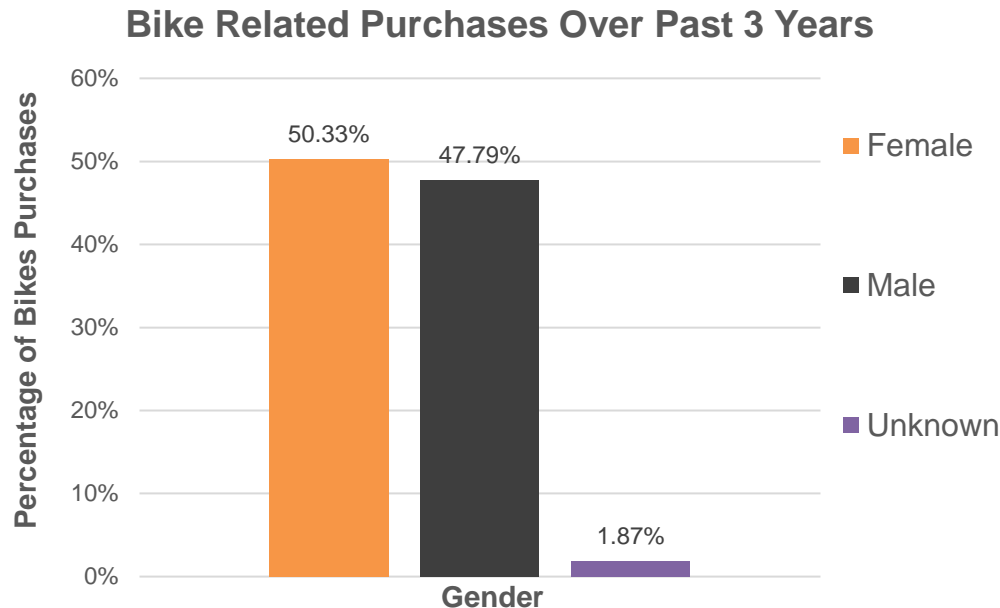
New Customers



Data Exploration

Bikes Related Purchases Over Last 3 Years

- Over the last 3 years, about 50% of bikes related purchases were made by females to 48% of purchases made by males.
- 2% purchases made by unknown gender.
- Females make up majority of bike related sales.

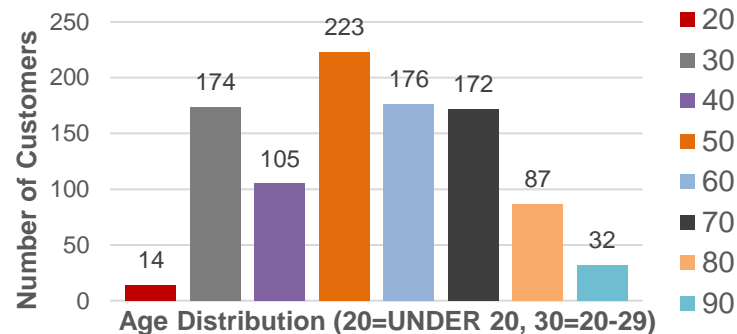


Data Exploration

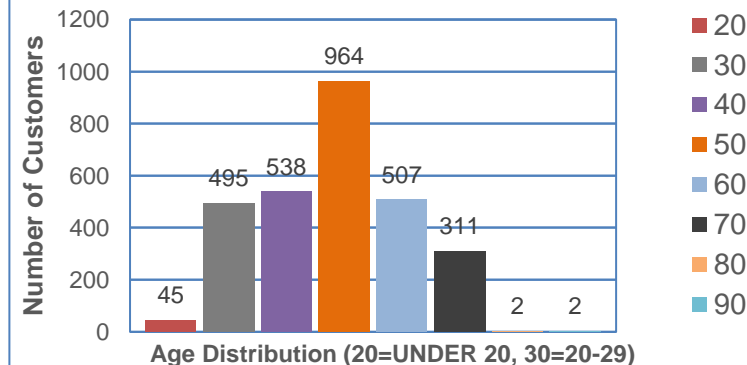
Customer Age Distribution

- Most 'New and Existing Customers' are aged between 40-49.
- The 'New', suggests age group 30, 50-70 are most populated.
- The 'Existing', suggests age group 30-60.
- The lowest age group in new customers, are younger i.e. under 20 and in existing, between 70-89.

New Customer Age Distribution



Existing Customer Age Distribution

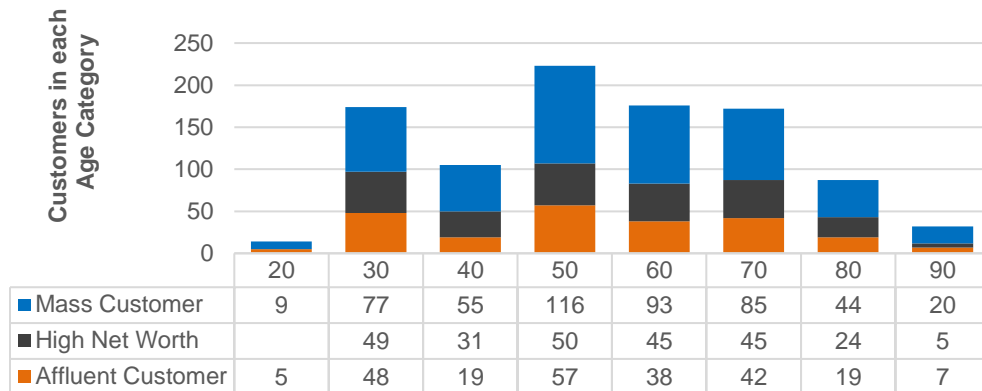


Data Exploration

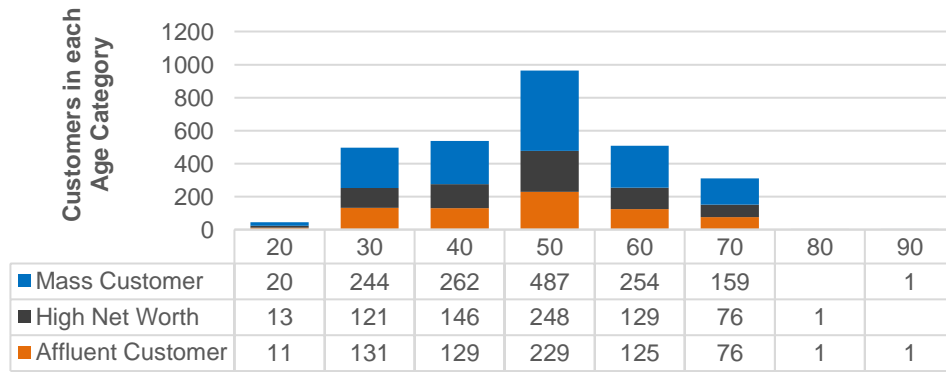
Wealth Segmentation by age

- Mass Customer = Most number of customers
 - High Net Worth = holding assets greater than given amount.
 - Affluent Customer = Wealthy Customers
-
- The Affluent Customers are more than High Net Worth customers in the 50 age category.

New Customer Wealth Segmentation by Age Category



Existing Customer Wealth Segmentation by Age Category

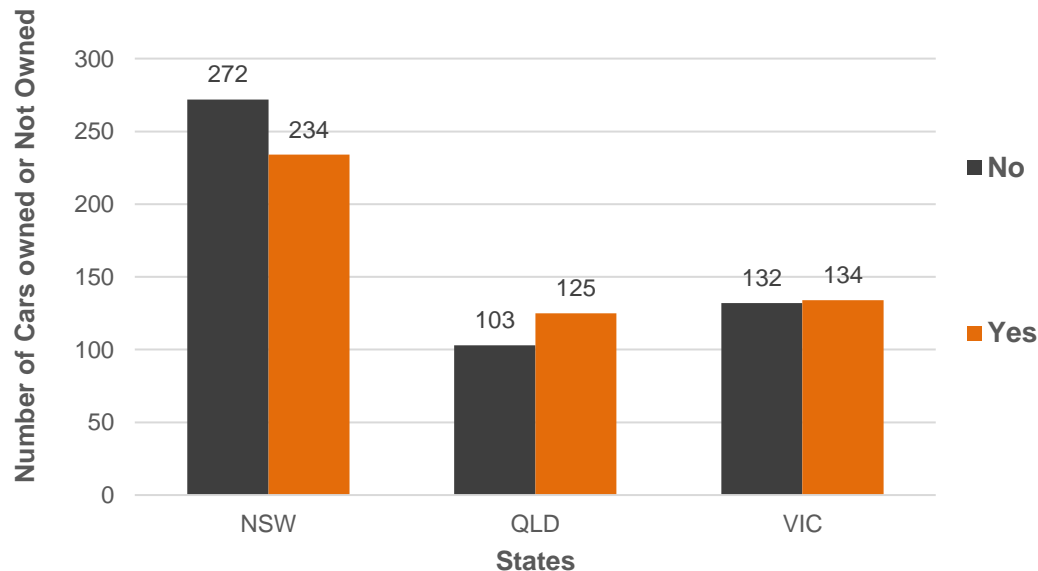


Data Exploration

Number of Cars owned and Not owned.

- According to chart, large number of people lives in New South Wales (NSW).
- NSW has largest amount of people that do not own a car.
- Victoria (VIC), is evenly split and significantly lower than NSW.
- In QLD, most number of customers own a car.

Number of Cars Owned per State

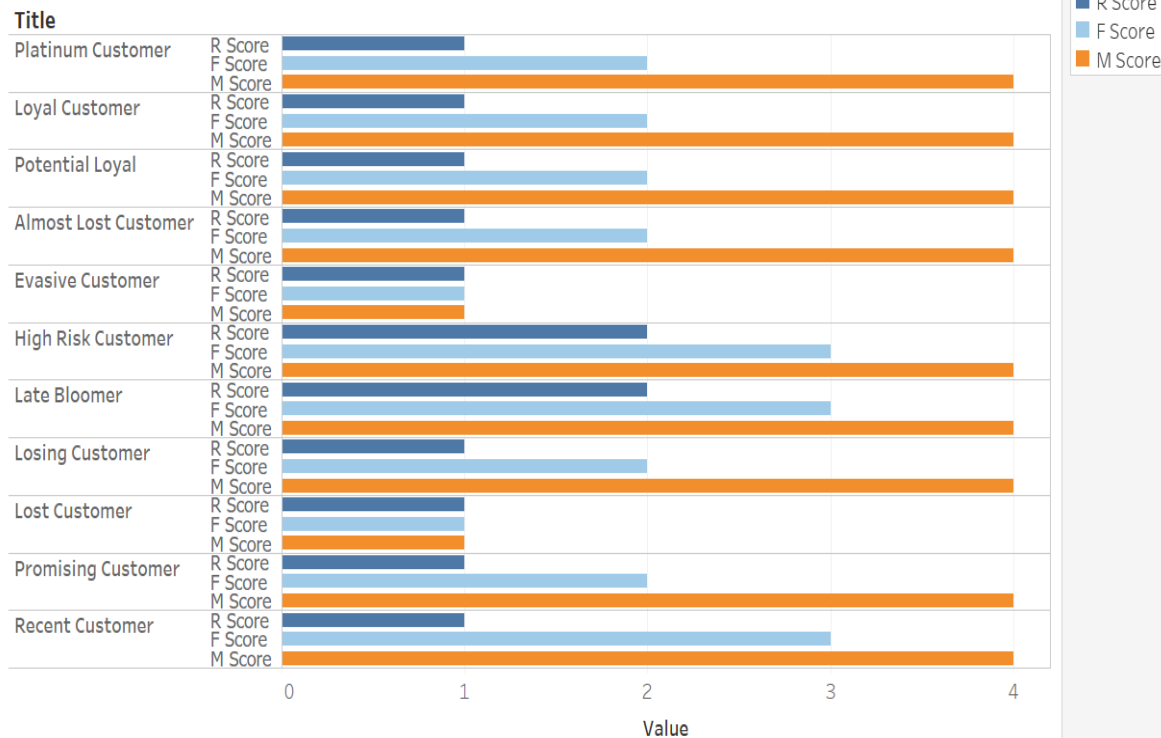


Model Development

Customer Classification and RFM Analysis

- RFM Analysis is used to determine which customers should target to increase its revenue and values.
- RFM stands for Recency, Frequency and Monetary value
- Recency – visits or the customer activity
- Frequency – The frequency of customer transactions
- Monetary – The intentions to spends or purchasing power of customers.

Customer Title and RFM Score(1-4)



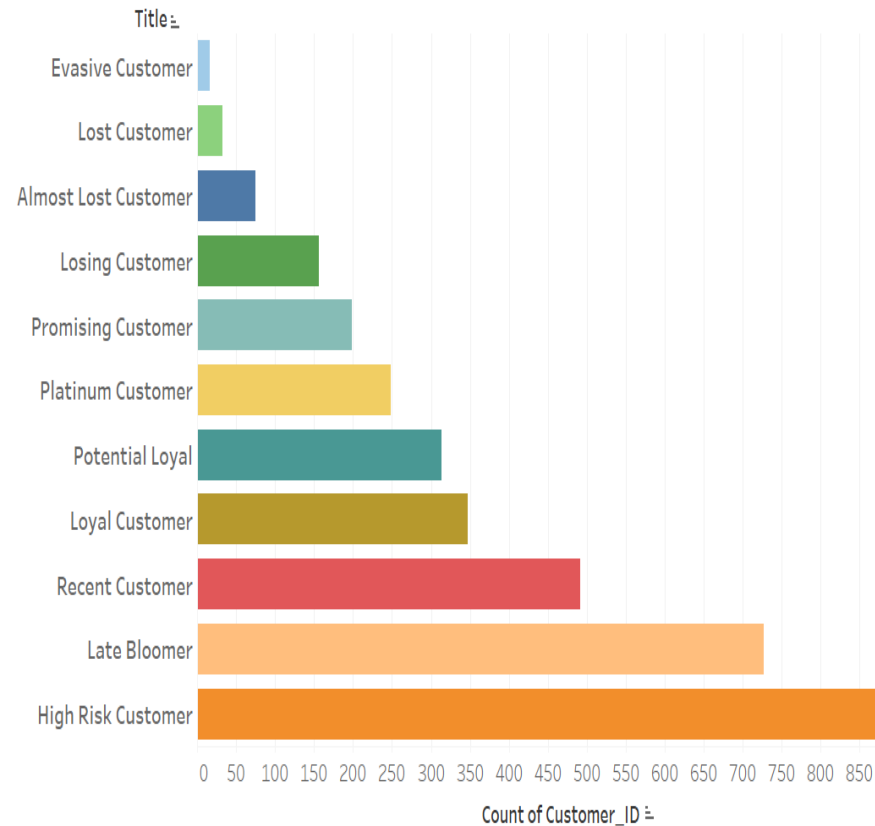
Model Development

Customer Title List with RFM values description

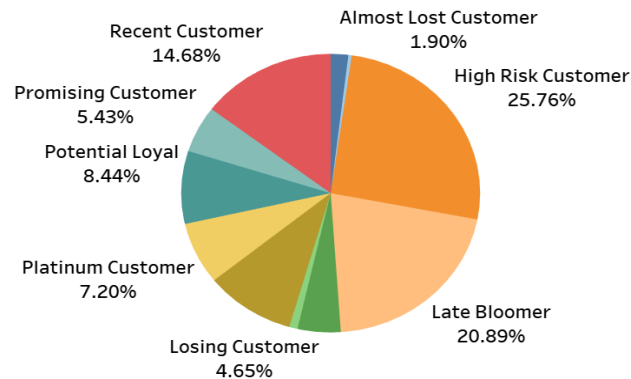
Rank	Customer Title	Description	RFM Value
1	Platinum Customer	Most Recent buy, buys often, most spent	444
2	Loyal Customer	Spend good money often. Responsive to promotions.	433
3	Potential Loyal	Relatively recent, bought more than once, spend large amount of money.	421
4	Recent Customer	Bought most recently, but not often.	344
5	Promising Customer	Recent shoppers, but haven't spent much.	323
6	Late Bloomer	No purchase recently, but RFM is larger than average.	311
7	Losing Customer	Below average recency, frequency and monetary values..	224
8	High Risk Customer	Spent high and purchased often. But long time ago. Need to bring them back!	212
9	Almost Lost Customer	Made biggest purchases, and often. But haven't returned for a long time.	124
10	Evasive Customer	Last purchase was long back, low spenders and low number of orders.	112
11	Lost Customer	Lowest recency, frequency and monetary scores.	111

Model Development

Customer Title Distribution



Customer Title Distribution



Interpretation

Identify Top 1000 Customers to Target

Rank	Customer Title	Number of Customers	Customer Selection
1	Platinum Customer	248	248
2	Loyal Customer	347	347
3	Potential Loyal	314	314
4	Recent Customer	491	91
5	Promising Customer	198	0
6	Late Bloomer	727	0
7	Losing Customer	156	0
8	High Risk Customer	888	0
9	Almost Lost Customer	75	0
10	Evasive Customer	16	0
11	Lost Customer	33	0

Interpretation

Customer Target

Rank	Customer Title	Number of Customers	Customer Selection
1	Platinum Customer	248	248
2	Loyal Customer	347	347
3	Potential Loyal	314	314
4	Recent Customer	491	91

- Identified the top 1000 customers by analysing the data.
- The 1000 customers discovered would have bought recently, frequently in the past and spend more compare to other customers.

Appendix