

TheAnalyticsTeam

# **Sprocket Central Pty Ltd**

Data analytics approach

Brian Ikkanda

# Agenda

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

# Introduction

## Recommend Top 1000 New Customers to Target

### Context and Goal

- Sprocket Central Pty Ltd is a bicycle company that specializes in high-quality bikes and accessible cycling accessories.
- The marketing team at Sprocket Central is looking to boost business by analyzing their existing customer datasets (demographics, addresses, and transactions).
- The overall goal is to recommend the top 1000 customers that Sprocket Central should target to drive higher value for the company.

### Approach

- Data exploration
- Model Development
- Interpretation/Recommendation



# Data Exploration

## Data Quality Assessment and Data Cleaning

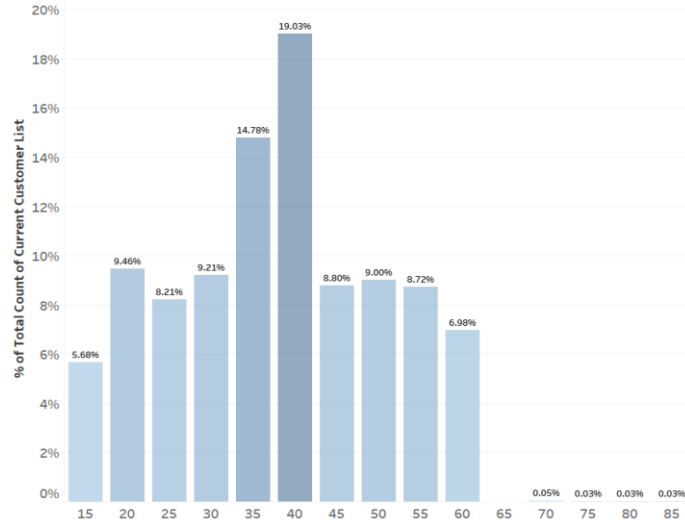
	Customer Demographics	Customer Address	Transaction Data
Accuracy <i>Correct values</i>	DOB: inaccuracies	Additional customer ID's (4001-4003)	
Completeness <i>Data fields with values</i>	Missing last_name, DOB, job_title, tenure values	Missing customer ID's (3,10,22,23)	Missing customer_id, online_order, brand, product_line, product_class, product_size, product_first_sold_date values
Consistency <i>Values free from contradiction</i>	Gender: inconsistent terms	State: inconsistent terms	
Currency <i>Values up to date</i>	Deceased customers		
Relevancy <i>Data items with value meta-data</i>	-Default column not relevant -Age: field to be added		-Cancelled status transactions -Profit: field to be added
Validity <i>Data containing allowable values</i>			Standard_cost: variations in data types
Uniqueness <i>Records that are duplicated</i>	No duplicated values	No duplicated values	No duplicated values

# Data Exploration

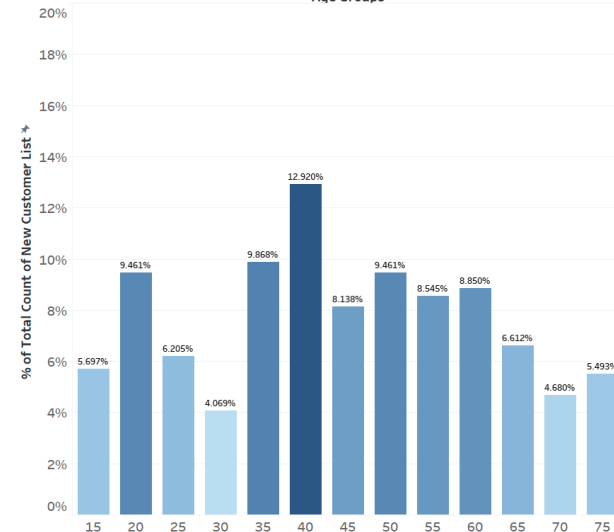
## Age Distributions

- For both sets of data, the 35-45 age range hold the largest percentage of customers.
- The 20-25 age group holds a significant percentage.
- For both distributions, those that are younger than 20 are a small percentage of the customers.
- The new customer data shows a smaller percentage of customers in the 30-35 range and also has a significant group of customers older than 60 years old.

**Current Customer Data**



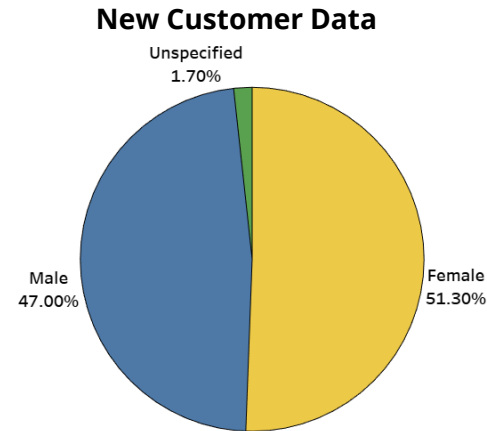
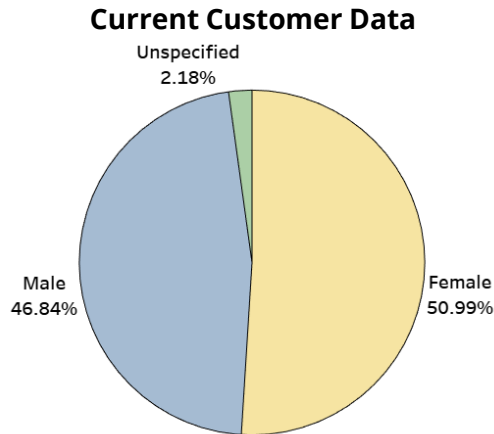
**New Customer Data**



# Data Exploration

## Bike Related Purchases Within the Past 3 Years by Gender

- Females make up the slight majority with regards to bike related purchases
- And that trend continues with the new customer data.

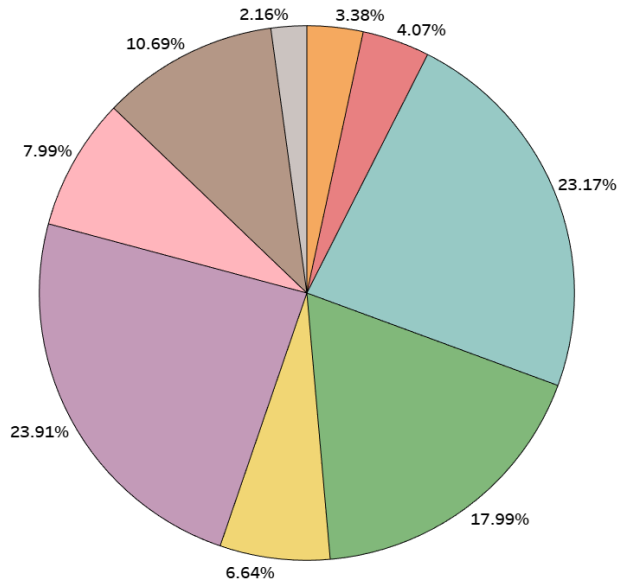


# Data Exploration

## Job Industry Distribution

- Financial Services and Manufacturing make up a large percentage of both current and new customers.
- Agriculture and Telecommunications make up the smallest percentage of both current and new customers.

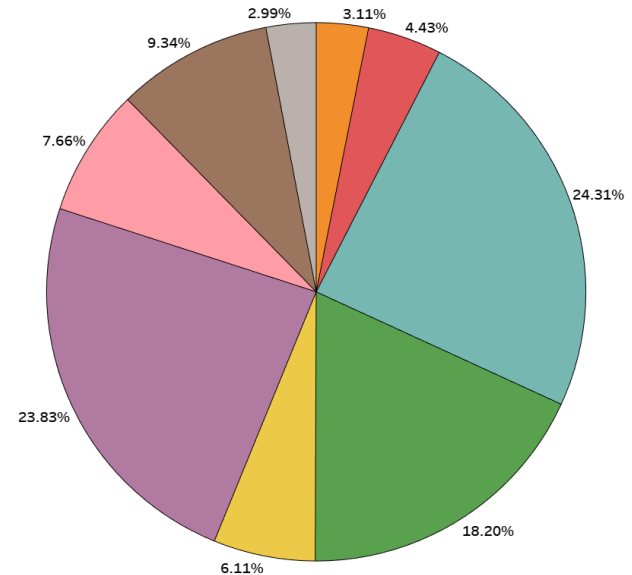
**Current Customer Data**



**Job Industry Category**



**New Customer Data**

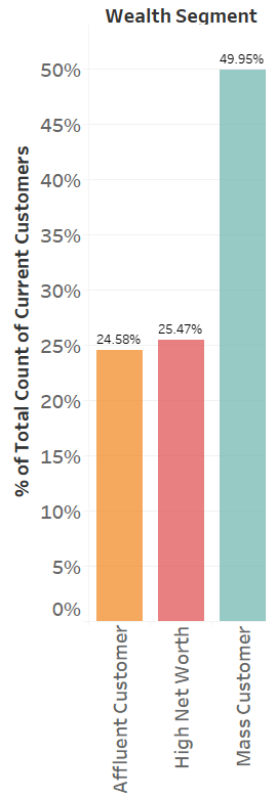


# Data Exploration

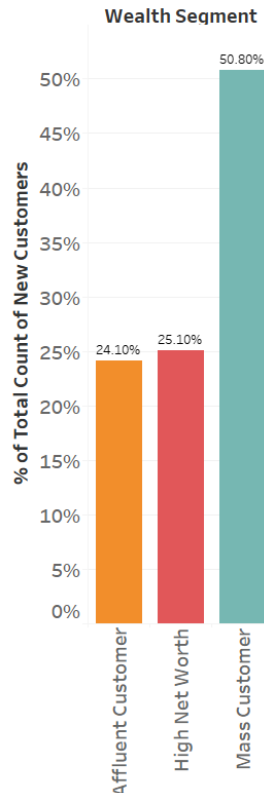
## Wealth Segmentation

- A significant number of customers are classified as 'Mass Customer' almost doubling the 'Affluent' and 'High Net Worth' customers.
- Customers labelled 'High Net Worth' are only slightly more than 'Affluent Customers'

Current Customer Data



New Customer Data



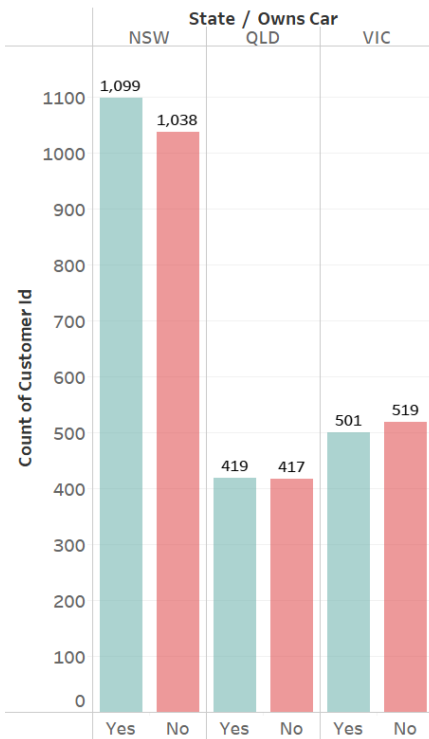


# Data Exploration

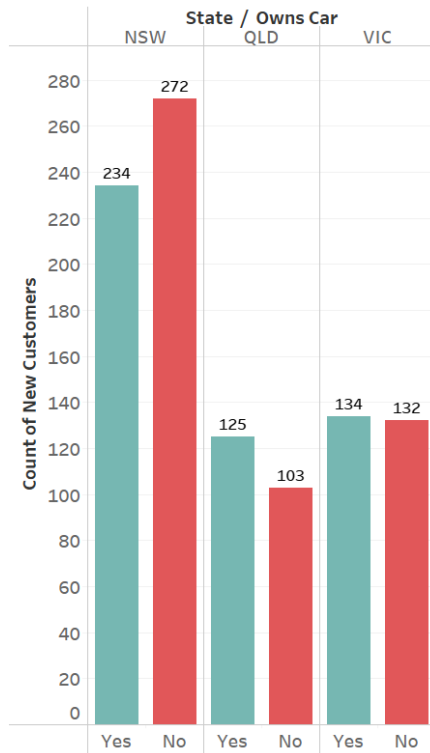
## Number of Cars Owned by State

- Amongst the data of current customers, there is a larger number of customers that owned cars at NSW. NSW seems to have a higher number of people from which data was collected.
- The new customer data shows that there are more customers WITHOUT cars in NSW.
- Both QLD and VIC seem to have more customers with cars in the New Customer data set.

Current Customer Data



New Customer Data



# Model Development

## Recency, Frequency, Monetary (RFM) Analysis

- Target customers can be identified using RFM analysis
- Recency, frequency, and monetary categories give insight into the level of engagement of customers.



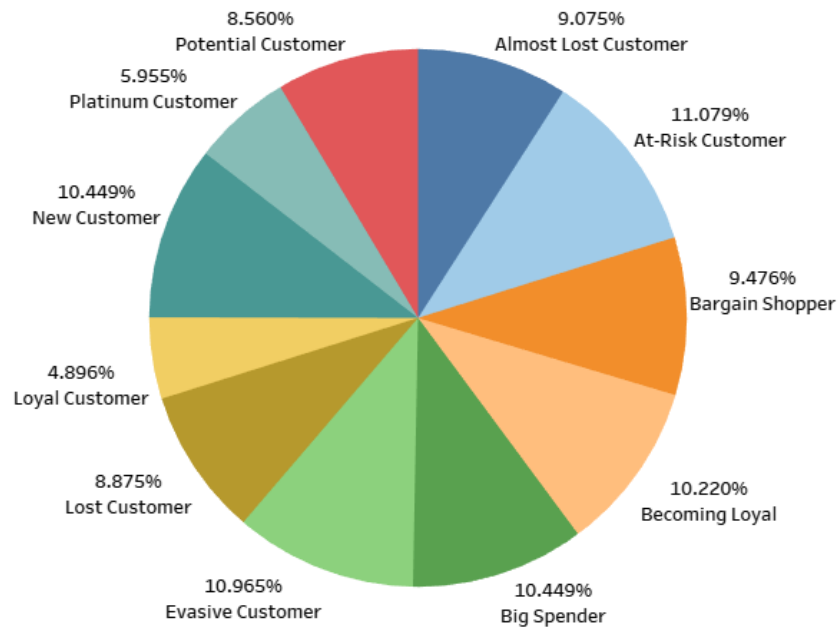
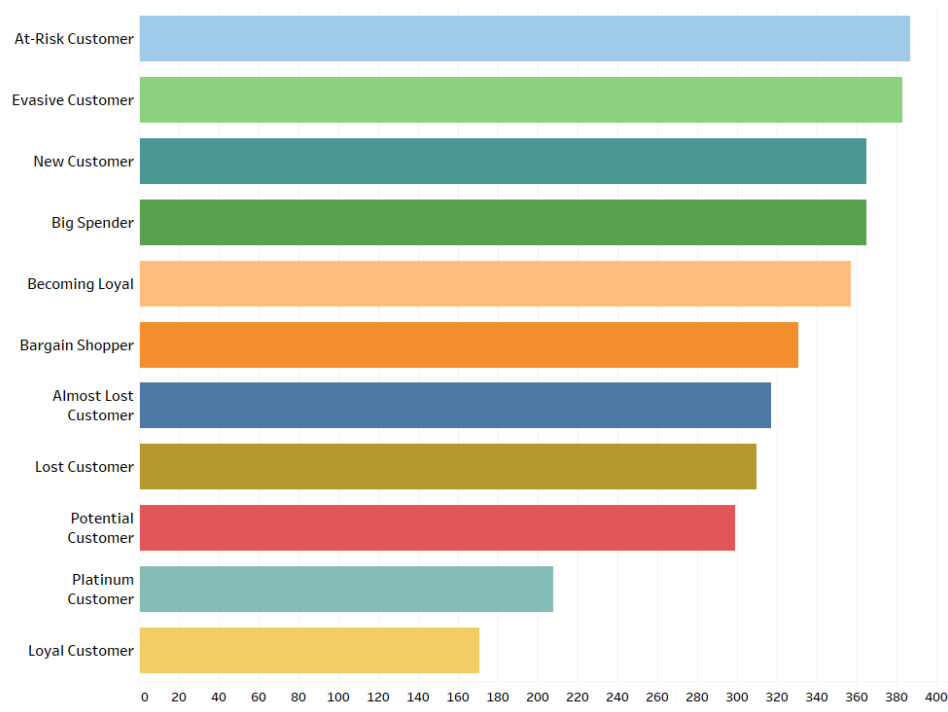
# Model Development

## Customer Definition List

Rank	Profile	RFM Value	Description
1	Platinum Customer	444	Most recent purchases, purchases often, highest spending
2	Loyal Customer	442	Most recent, buys often, low amount spent
3	Becoming Loyal	422	Most recent, bought more than once, low amount spent
4	Big Spender	344	Recently purchased, very frequent buyer, high amount spent
5	Potential Customer	331	Recently purchased, frequent purchases, spent small amount
6	New Customer	311	Recently purchased, very low frequency, very low amount spent
7	At-risk Customer	233	Low recency, fairly frequent buyer, moderate amount spent
8	Bargain Shopper	221	Low recency, low frequency, very small amount spent
9	Almost Lost Customer	133	Very low recency, fairly frequent buyer, but high amount spent
10	Evasive Customer	121	Very low recency, low frequency, very small amount spent
11	Lost Customer	111	Very low RFM

# Model Development

## Distribution of Customers



# Interpretation

## Top 1000 Customers

Rank	Profile	Description	No. of Customers	Cumulative Count	Selection
1	Platinum Customer	Most recent purchases, purchases often, highest spending	208	208	208
2	Loyal Customer	Most recent, buys often, low amount spent	171	379	171
3	Becoming Loyal	Most recent, bought more than once, low amount spent	357	736	357
4	Big Spender	Recently purchased, very frequent buyer, high amount spent	365	1101	264
5	Potential Customer	Recently purchased, frequent purchases, spent small amount	299	1400	0
6	New Customer	Recently purchased, very low frequency, very low amount spent	365	1765	0
7	At-risk Customer	Low recency, fairly frequent buyer, moderate amount spent	387	2152	0
8	Bargain Shopper	Low recency, low frequency, very small amount spent	331	2483	0
9	Almost Lost Customer	Very low recency, fairly frequent buyer, but high amount spent	317	2800	0
10	Evasive Customer	Very low recency, low frequency, very small amount spent	383	3183	0
11	Lost Customer	Very low RFM	310	3493	0

# Interpretation

## Top 1000 Customers

Rank	Profile	Description	No. of Customers	Cumulative Count	Selection
1	Platinum Customer	Most recent purchases, purchases often, highest spending	208	208	208
2	Loyal Customer	Most recent, buys often, low amount spent	171	379	171
3	Becoming Loyal	Most recent, bought more than once, low amount spent	357	736	357
4	Big Spender	Recently purchased, very frequent buyer, high amount spent	365	1101	264
5	Potential Customer	Recently purchased, frequent purchases, spent small amount	299	1400	0
6	New Customer	Recently purchased, very low frequency, very low amount spent	365	1765	0
7	At-risk Customer	Low recency, fairly frequent buyer, moderate amount spent	387	2152	0
8	Bargain Shopper	Low recency, low frequency, very small amount spent	331	2483	0
9	Almost Lost Customer	Very low recency, fairly frequent buyer, but high amount spent	317	2800	0
10	Evasive Customer	Very low recency, low frequency, very small amount spent	383	3183	0
11	Lost Customer	Very low RFM	310	3493	0

# Interpretation

## Top 1000 Customers

Rank	Profile	Description	No. of Customers	Cumulative Count	Selection
1	Platinum Customer	Most recent purchases, purchases often, highest spending	208	208	208
2	Loyal Customer	Most recent, buys often, low amount spent	171	379	171
3	Becoming Loyal	Most recent, bought more than once, low amount spent	357	736	357
4	Big Spender	Recently purchased, very frequent buyer, high amount spent	365	1101	264

- The goal is to identify the top 1000 customers to target to drive higher value for the company
- These target customers will be the most recent customers (since they are purchasing the most current stock of merchandise). But these customers should also buy fairly frequently and spend the most amount of money to be considered high value customers.