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 to Target from Datasets

Outline of Problem

- Sprocket Central is a company that specializes in high-quality bikes and cycling accessories.
- Their marketing team is looking to boost business sales by analyzing provided datasets.
- Using the 3 datasets provided the aim is to analyze and recommend 1000 customers that Sprocket Central should target to drive higher value for the company.

Contents of Data Analysis

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

Data Exploration

Data Quality Assessment and 'Clean Up'

Key Issues for Data Quality Assessment

- Accuracy: Correct Values
- Completeness: Data Fields with Values
- Consistency: Values Free from Contradiction
- Currency: Values up to Date
- Relevancy: Data items with Value Meta-data
- Validity: Data Containing Allowable Values
- Uniqueness: Records that are Duplicated

Summary Table

| Datasets	Accuracy	Completeness	Consistency	Currency	Relevancy	Validity
Customer Demographic	DOB: inaccurate Age: missing	Job title: blanks Customer id: incomplete	Gender: inconsistency	Deceased customers: filter out	Default column: delete	
Customer Address		Customer id: incomplete	States: inconsistency			
Transactions	Profit: missing	Customer id: incomplete Online Order: blanks Brand: blanks			Order status: filter out cancelled	List price: format Product sold date: format

Data Exploration

'New' and 'Old' Customer Age Distributions

- Most customers are aged between 40-49 in 'New'. In 'Old' majority of customers are aged between 40-49 also.
- The lowest age groups are under 20 and 80+ for both 'New' and 'Old' customer lists.
- The 'New' customer list suggests 40-49 and 60-69 are most populated.
- The 'Old' customer list suggest 20-69.
- There is a steep drop of customers in the 30-39 age group in 'New'.
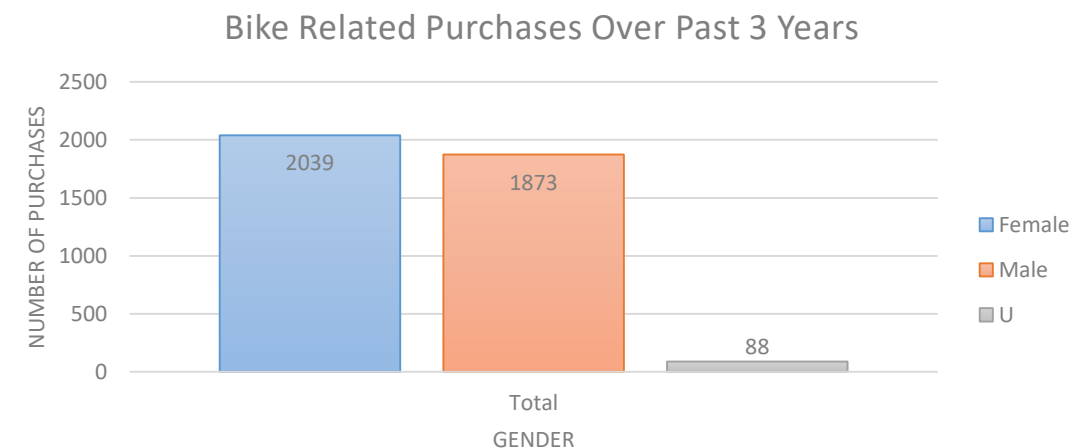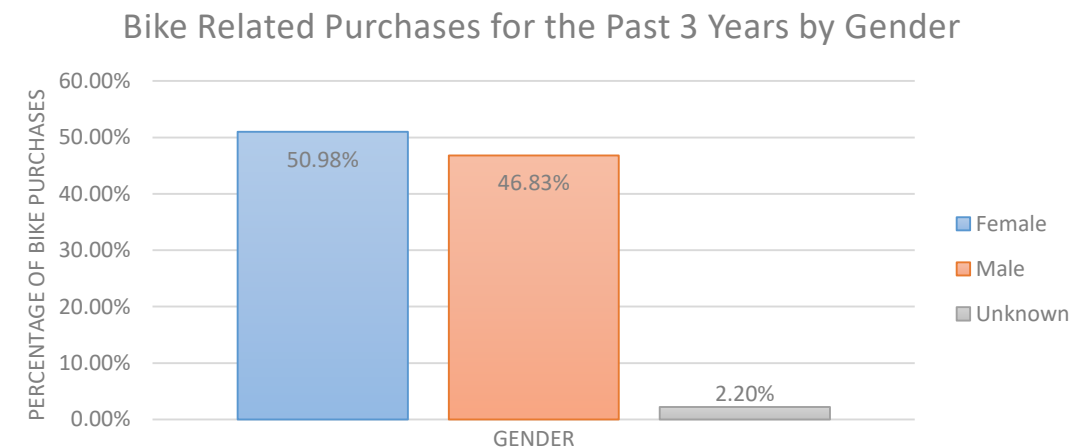

New Customer Age Distribution
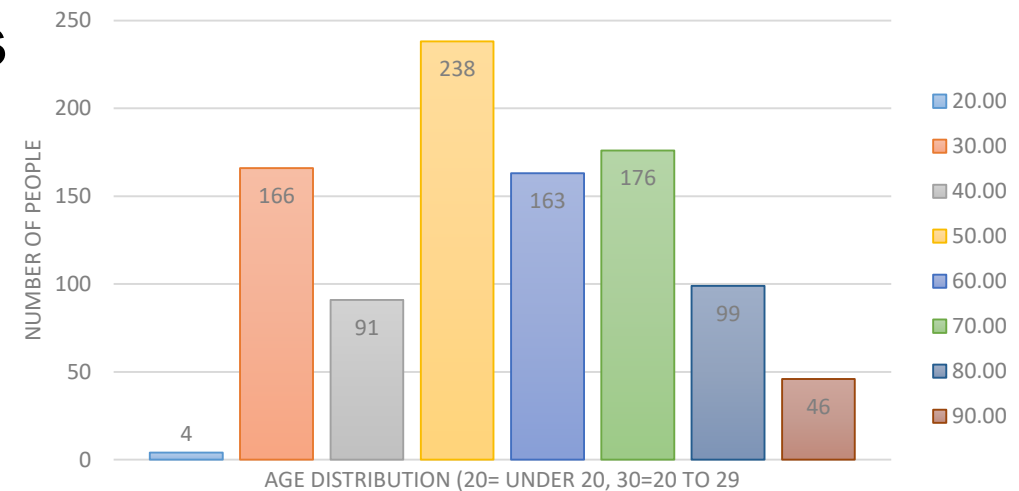
Old Customer Age Distribution
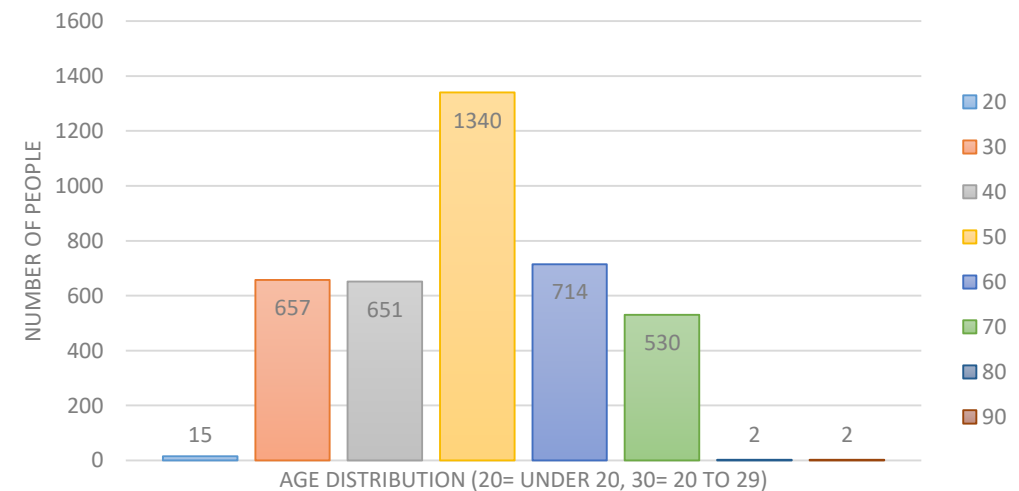
Data Exploration

Bike Related Purchases Over Last 3 Years by Gender

- Over the last three years about 50% of bike related purchases were made by females to 47% of purchases made by males. Approximately 2% were made by unknown gender.
- Females make up majority of bike related sales.

Bike Related Purchases for the Past 3 Years by Gender

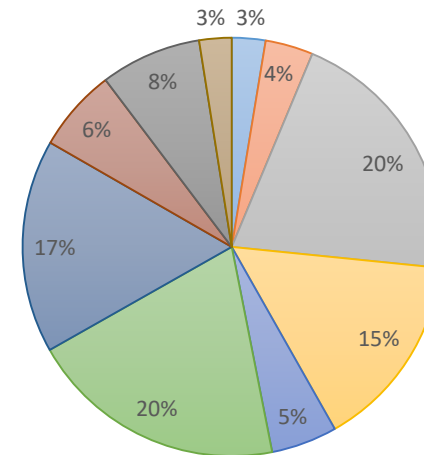Bike Related Purchases Over Past 3 Years
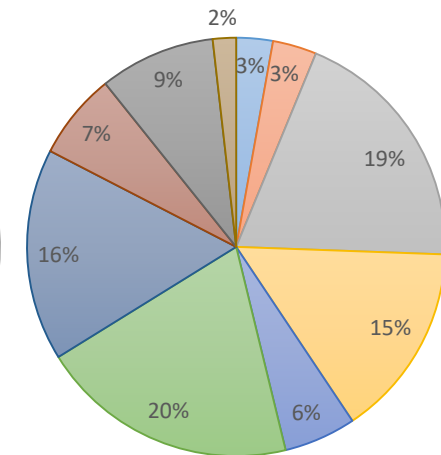
Data Exploration

Job Industry Distribution

- 20% of 'New' Customers are in Manufacturing and Financial Services.
- The smallest number of customers are in Agriculture and Telecommunications at 3%
- Similar pattern in 'Old' customer list, at 20% and 19% in Manufacturing and Financial Services respectively.

'New' Customer List

'Old' Customer List

Data Exploration

## Wealth Segmentation by Age Category

- In all age categories the largest number of customers are classified as 'Mass Customer'
- The next category is the 'High Net Worth' customers.
- The 'Affluent Customer' can outperforms the 'High Net Worth' customer in the 40-49 age group.

New Customer Wealth Segment by Age

NUMBER OF PEOPLE IN EACH AGE CATEGORY

	20	30	40	50	60	70	80	90
■ Mass Customer	1	77	47	123	91	84	52	24
■ High Net Worth		41	33	55	30	52	26	12
■ Affluent Customer	3	48	11	60	42	40	21	10

Old Customer Wealth Segment by Age

NUMBER OF PEOPLE IN EACH AGE CATEGORY

	20	30	40	50	60	70	80	90
■ Mass Customer	9	277	276	569	310	213	2	1
■ High Net Worth	2	151	157	292	141	107		
■ Affluent Customer	2	137	125	294	151	121		1

Numbers of Cars Owned and Not Owned by State

- NSW has the largest amount of people that **do not** own a car. NSW seems to have a higher number of people from which data was collected.
- Victoria is also split quite evenly. But both numbers are significantly lower than those of NSW.
- QLD has a relatively high number of customers that own a car.

Number of Cars Owned in Each State
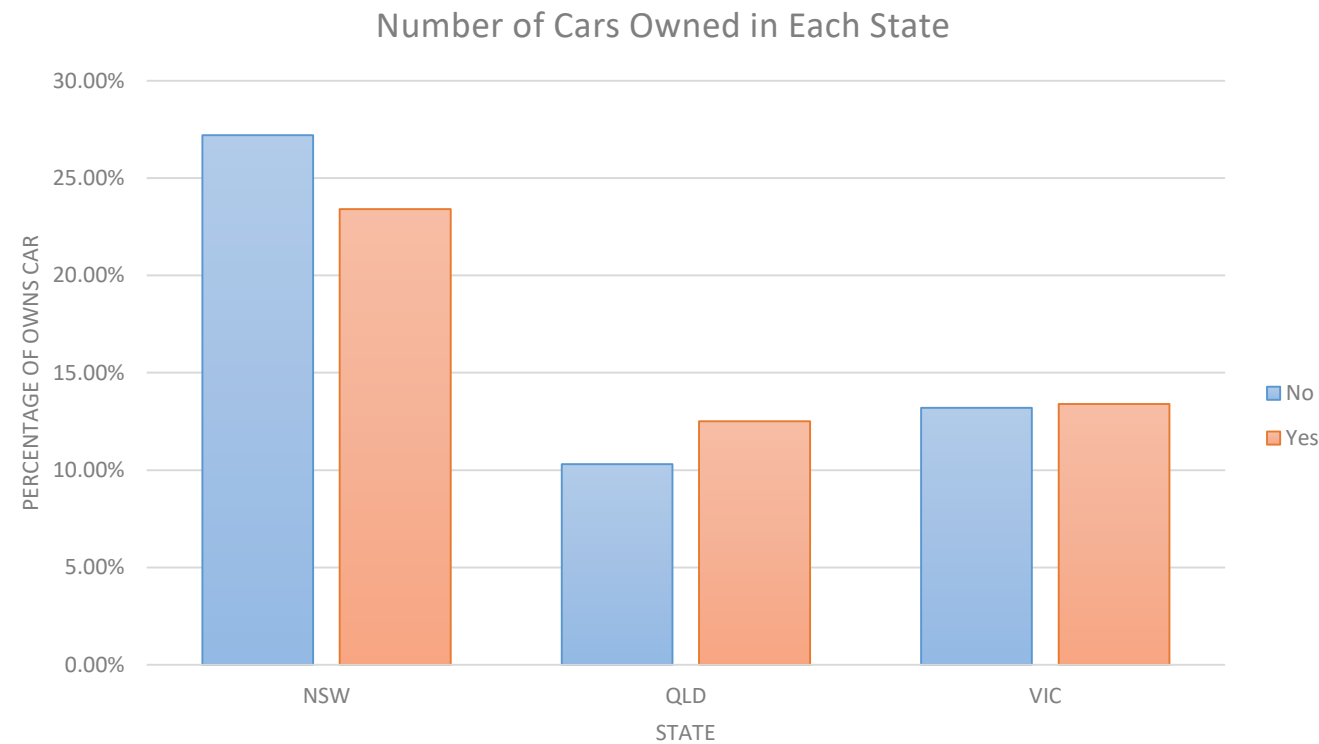
Model Development

RFM Analysis and Customer Classification

- RFM analysis is used to determine which customers a business should target to increase its revenue and value.
- The RFM (Recency, Frequency, and Monetary) model shows customers that have displayed high levels of engagement with the business in the three categories mentioned.


Customer Title and Scores

Model Development

Customer Title Definition List with RFM Value Assigned

Rank	Customer Title	Description	RFM Value
1	Platinum Customer	Most recent buy, buys often, most spent	444
2	Very Loyal	Most recent, buys often, spend 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 purchases recently, but RFM value is larger than average	311
7	Losing Customer	Purchases was a while ago, below average RFM value	224
8	High Risk Customer	Purchase was long time age, frequency is quite high, amount spent is high	212
9	Almost Lost Customer	Very low recency, low frequency, but high amount spent	124
10	Evasive Customer	Very low recency, very low frequency, small amount spent	112
11	Lost Customer	Very low RFM	111

Model Development

Customer Title Distribution

Distribution of Customers

Distribution of Customers
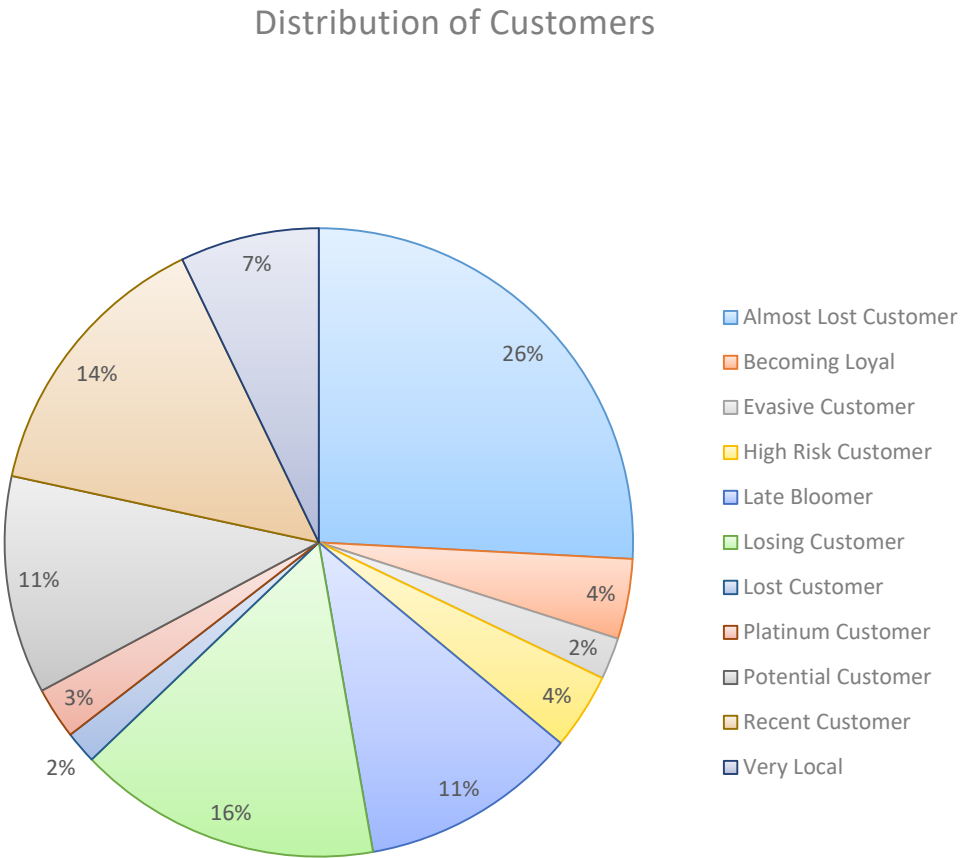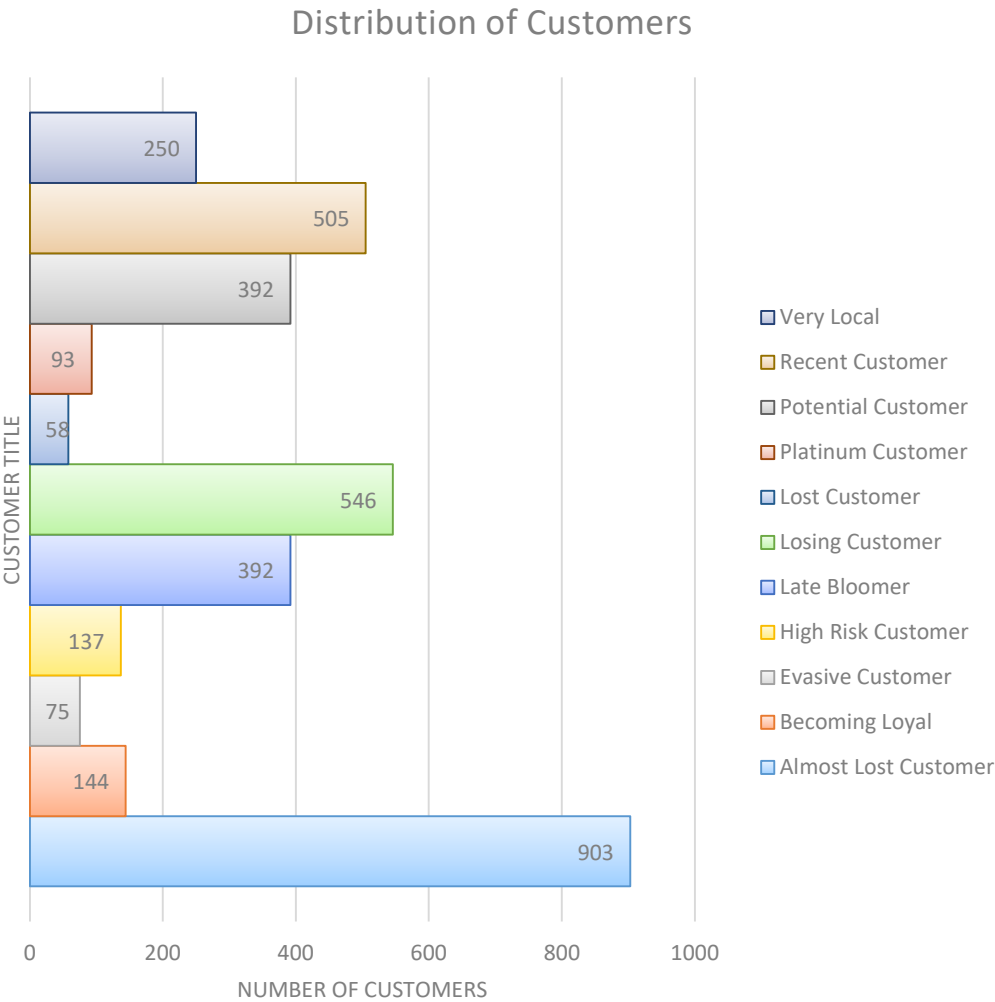
Interpretation

Summary Table of the Top 1000 Customers to Target

Rank	Customer Title	Number of Customers	Cumulative	Customer Selection
1	Platinum Customer	93	93	93
2	Very Loyal	250	343	250
3	Becoming Loyal	144	487	144
4	Recent Customer	505	992	505
5	Potential Customer	392	1384	8
6	Late Bloomer	392	1776	0
7	Losing Customer	546	2322	0
8	High Risk Customer	137	2459	0
9	Almost Lost Customer	903	3362	0
10	Evasive Customer	75	3437	0
11	Lost Customer	58	3495	0

Interpretation

Customer Target and Methodology

Rank	Customer Title	Number of Customers	Cumulative	Customer Selection
1	Platinum Customer	93	93	93
2	Very Loyal	250	343	250
3	Becoming Loyal	144	487	144
4	Recent Customer	505	992	505
5	Potential Customer	392	1384	8

- Filter through the top 1000 customers by assigning the conditions discussed in the table above.

- The 1000 customers discovered would have bought recently, they have bought very frequently in the past and trend to spend more than other customers.

Appendix

Appendix

This is an optional slide where you may place any supporting items.