TheAnalyticsTeam

Sprocket Central Pty Ltd

Data analytics approach

Atefeh Gholipour

Agenda

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

Introduction

Targeting high value customers based on customer demographics and attributes

- Sprocket Central Pty Ltd is a medium size bikes & cycling accessories organisation.

- They provided KPMG with 3 datasets including Customer Demographic, Customer Addresses and Transactions data in the past 3 months.

- They ask KPMG team to effectively analyse the datasets to help Sprocket Central Pty Ltd grow its business.

- The goal is finding which of the1000 new customers should be targeted to drive the most value for the organization.

- It is achieved by following 3 phases as follows - Data Exploration, Model Development and Interpretation.

Data Quality Assessment

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

- In first phase, based on Standard Data Quality Dimension, all tables are cleaned and prepared for analysis.

|	Accuracy	Completeness	Consistency	Currency	Relevancy	Validity
Customer demographic	• DOB (have outlier)	• job_title (missing value) • Costumer_id	• gender	• deceased_ indicator (N)	• default	
Customer address		• Costumer_id	• state		• order_status (cancelled order)	
Transaction		• Costumer_id • Online_order • brand				• list_price (format) • product_first_ sold_date (format)

Age Distribution

- In both Od and New customer the majority of customers are aged between 40-49.

- There is a drop in new customer aged 30-39.

- There is a significant increase in new customers in the 70-79 age group.
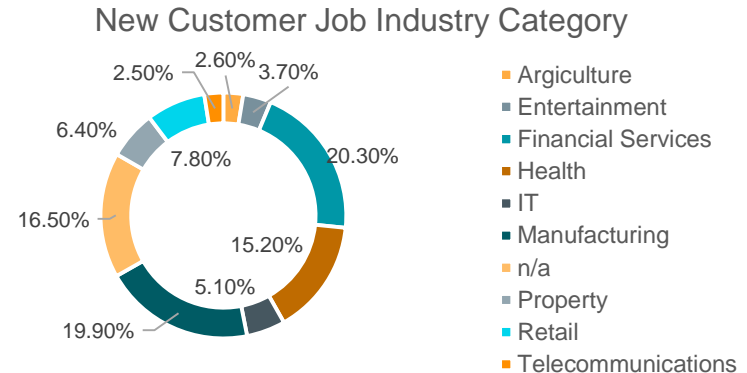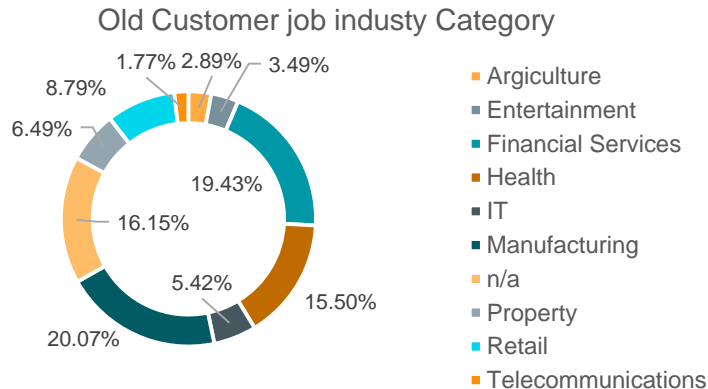

Old Customer Age Distribution
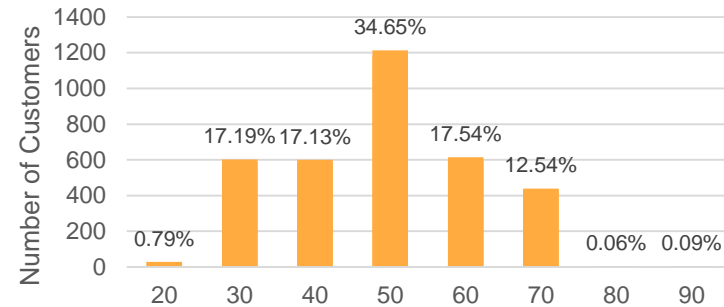
New Customer Age Distribution
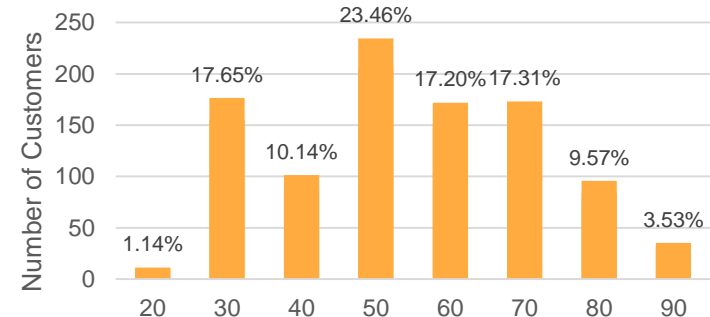
Data Exploration

Job Industry Distribution

- Around 20% of customers (both Old and New) are working in manufacturing and financial services.

- The smallest percent of customers (both Old and New) are working in agriculture and telecommunications sectors.

Old Customer job industy Category

1.77% 2.89% 3.49%
8.79%
6.49%
19.43%
16.15%
5.42%
15.50%
20.07%

- Argiculture
- Entertainment
- Financial Services
- Health
- IT
- Manufacturing
- n/a
- Property
- Retail
- Telecommunications

New Customer Job Industry Category

2.50% 2.60% 3.70%
6.40%
7.80%
20.30%
16.50%
15.20%
5.10%
19.90%

- Argiculture
- Entertainment
- Financial Services
- Health
- IT
- Manufacturing
- n/a
- Property
- Retail
- Telecommunications

Number of customers who owns car in different state

- The number of New customers in NSW who have not a car are higher than the number who own a car.

- QLD has the most difference between New customers who car owned and not owned.

- In VIC two groups are almost equal.
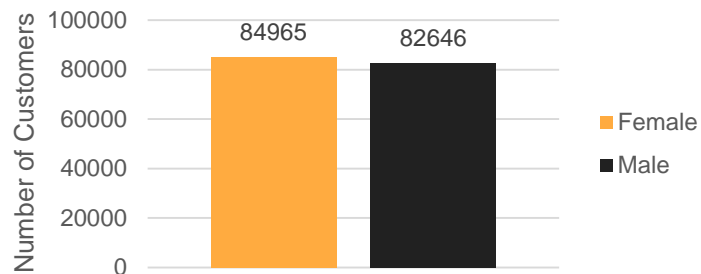
Old Customers Owned Car

New Customers Owned Car
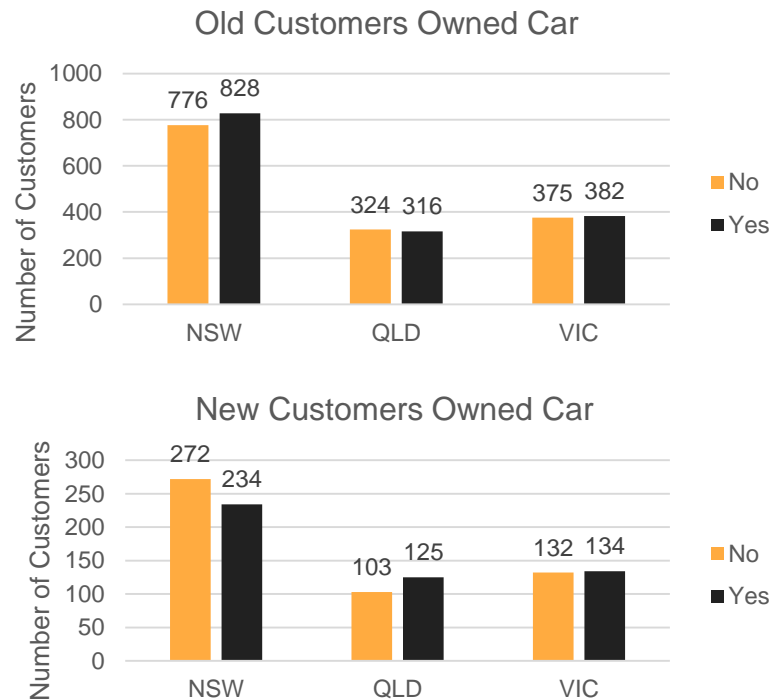
Past 3 years purchase behavior by gender
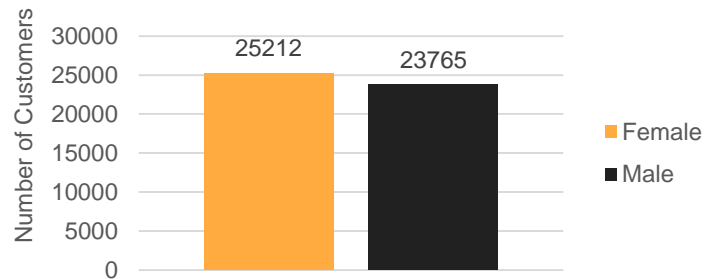
- Females customers made purchase a little more than males customers.

Old Customer Gender Distribution
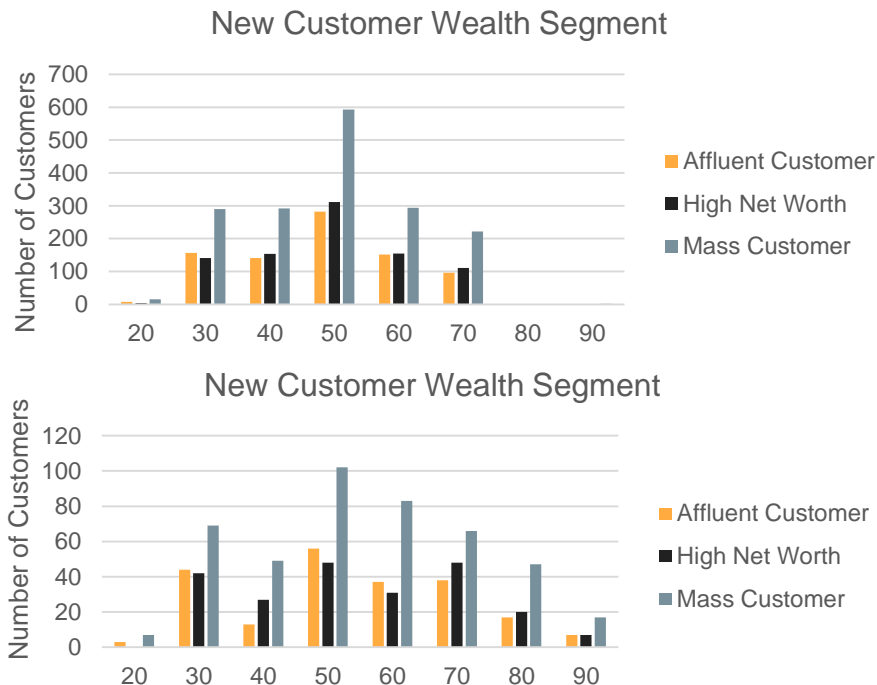
New Customer Gender Distribution

Past 3 years purchase behavior by gender

- In all age groups, Mass customers formed the largest number of customers.

- Affluent customers are in aged 40-49 in both New and Old list.

New Customer Wealth Segment

New Customer Wealth Segment

RFM Analysis

- RFM (Recency, Frequency, and Monetary) is a method used for analysing customer value and used to determine which customer should be targeted.

- Maximum and Minimum of R_Score, F_Score and M_Score for each customer labels.

Row Labels	Max of R_Score	Min of R_Score	Max of F_Score	Min of F_Score	Max of M_Score	Min of M_Score
1. Champion	4	4	4	3	4	2
2. Potential Loyalists	4	4	3	2	4	1
3. Recent Users	4	3	4	1	4	1
4. Can't Lose Them	3	3	4	2	4	1
5. Need Attention	3	3	2	1	4	1
6. Loyal Customers	2	2	4	2	4	1
7. Price Sensitive	2	2	2	1	3	1
8. About to Sleep	2	1	4	1	4	1
9. Hibernating	1	1	2	1	4	1
10. Lost Customers	1	1	1	1	1	1

Model Development

Customer segmentation based on RFM values

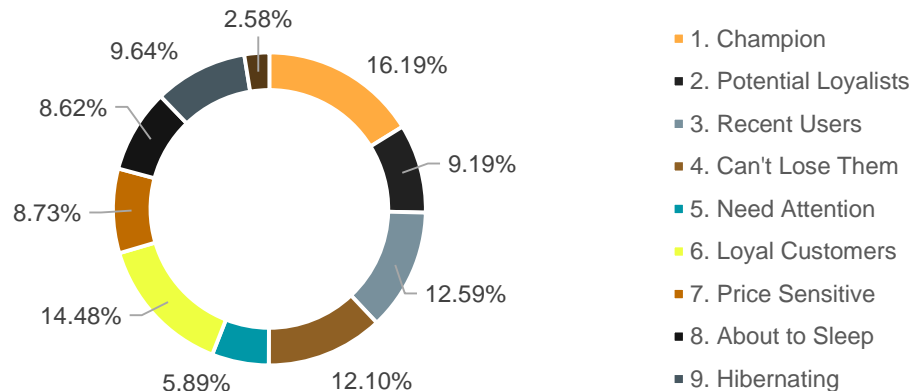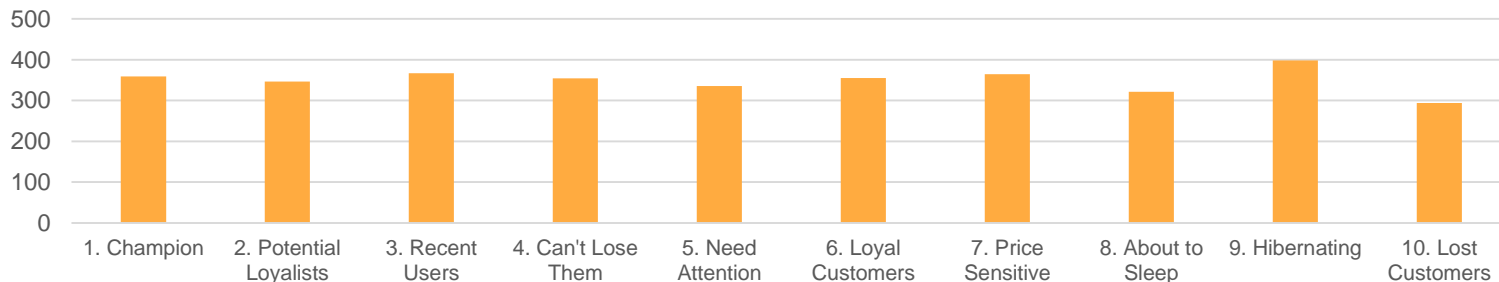| Label	Characteristics
Champion	Dream customer. Bought recently, Buy frequently, and buys big. Create positive word of mouth for the brand amongst their circles. Contributes a disproportionate amount of revenues to the business, and hence must be treated with care.
Potential Loyalists	High spending customers, who needs to feel valued in order to become a loyalist of the brand. Significant contributor to the business revenue, with potential to contribute further.
Recent Users	Have recently started using the brand, need to be cultivated by the brand to become long-term brand loyalists. Important segment which needs to be monitored constantly to prevent drop-outs after a few recent purchases.
Can't Lose Them	High value customers. As the label suggests, this segment requires continuous attention, as these customers are regular users of the brand at some risk of switching, and need to feel valued, to continue using it.
Need Attention	This segment constitutes users who are considering stopping use of the brand due to various reasons. In order to prevent this, it is necessary to scrutinize their behaviour and create further personalisation in communication to ensure increase in brand trust and loyalty.
Loyal Customers	This segment is the bread and butter of every business, as customer is this segment have a clear and positive view of the brand. These customers are happy with the product and service provided, and are unlikely to switch to alternatives. Important to make these customers feel valued.
Price Sensitive	These customers are always on the lookout for price-led promotions. Their loyalty and purchase is not defined by brand interactions, but by changes in the price that they have to pay. Segment which can be easily leveraged through offers and discounts.
About to Sleep	Existing users who haven't made purchases in a long time, and are on the verge of being lost customers. The business needs to make a last push to ensure customer retention, and create the most relevant and pertinent messaging possible in order to do so.
Hibernating	Existing users of the brand, who have not bought recently. These customers need to be induced to make a purchase as soon as possible, or else the brand will lost out on the trust and relevance it has created previously.
Lost Customers	These customers have stopped using brand altogether, and have decided to use alternatives instead. Very difficult to re-engage with this segment, as they already have a negative view of the brand or a more trustworthy alternative.

Customers contribution in profit

- Champion customers has the most contribution and Lost customer has the least contribution in profit.

Old Customers Profit, Total Profit = $11,144,921.72

Legend:
- 1. Champion
- 2. Potential Loyalists
- 3. Recent Users
- 4. Can't Lose Them
- 5. Need Attention
- 6. Loyal Customers
- 7. Price Sensitive
- 8. About to Sleep
- 9. Hibernating

Percentages: 16.19%, 9.19%, 12.59%, 12.10%, 5.89%, 14.48%, 8.73%, 8.62%, 9.64%, 2.58%

Old Customer Segmentation

Categories: 1. Champion, 2. Potential Loyalists, 3. Recent Users, 4. Can't Lose Them, 5. Need Attention, 6. Loyal Customers, 7. Price Sensitive, 8. About to Sleep, 9. Hibernating, 10. Lost Customers

Interpretation

Customer label distribution

- The number of customer in each segment is given a table.

- The customers who bought recently, bough very frequently in the past and tend to spend more than others are highlighted who should be targeted.

Row Labels	Number of customers
1. Champion	359
2. Potential Loyalists	346
3. Recent Users	367
4. Can't Lose Them	354
5. Need Attention	335
6. Loyal Customers	355
7. Price Sensitive	364
8. About to Sleep	321
9. Hibernating	398
10. Lost Customers	294

Appendix