TheAnalyticsTeam

Sprocket Central Pty Ltd

Data analytics approach

KPMG Analytics Division- Jaynil Pandya

Agenda

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

Introduction

We are provided with Customer Dataset which gives insights on type, age, demographic, and profit of customers. We have identified target customer base among the new customers.

Outline of the problem-
- Sprocket Central is a company that specializes in high quality bike and accessories.
- The Marketing Team is looking to boost sales.
- To target 1000 new customers that will bring the highest value to the business.

Approach to solve the problem:
- Identifying the age group that purchases frequently.
- Identifying the demographic for largest customer base.
- Classifying target customer based on different criteria.
- Recency, Frequency, and Monetary Analysis.

Data Exploration

Issues discovered and addressed while performing Data Quality Assessment.

| Data Quality Parameters	Issues	Mitigation
Accuracy	• Outliers in DOB column in multiple sheets. • Spelling error in 'job_category' column in *CustomerDemographic.*	• Create a new age column wherever DOB is mentioned for better analysis. • Find and replace the error with correct spelling.
Completeness	• Null values in many columns. • Default column in *CustomerDemographic* is filled with irrelevant values.	• Filter out and assign *NA* to null-values in columns like address, profession, etc… • Deleting the column for better analysis at further stages.
Consistency	• Gender attribute in multiple sheets is not consistent. • Some data fields have full state names instead of short forms.	• Edit values using find and replace. • For states, VLOOKUP can also be used.

Data Exploration

Currency	• **The 'Y' indicator in deceased_indicator is irrelevant.**	• **Discard the records with 'Y' as deceased_indicator.**
Relevancy	• Redundant column of rank in NewCustomerList	• Keep only one rank column (Advisable to keep the one which formulates the rank).
Validity	• In *transactions,* product_first_sold_date is of wrong data type. • In *transactions,* list_price and standard_cost is not of currency type.	• Ensure all the values have relevant data types. This will be helpful while fetching data for model building.

Data Exploration: Age-wise Distribution
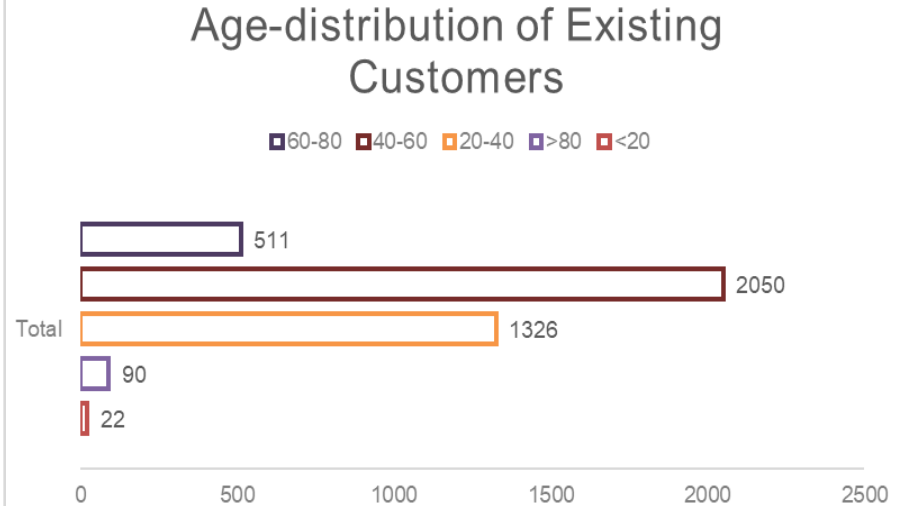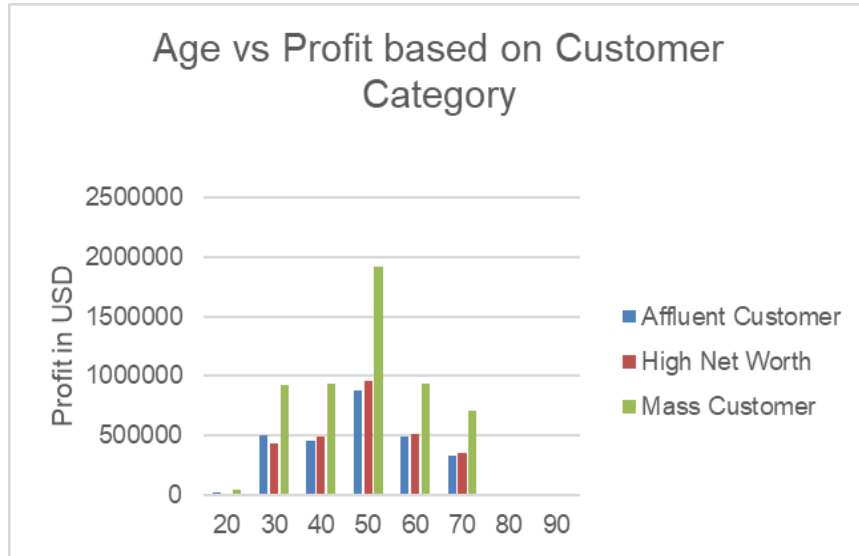
- It is evident from the above graphs that majority of the existing customers are of the age group 40-60 and 20-40. It is conclusive to say that the average consume age is between **40-50.**

- The age group of 50 is also the most profitable group for the company. Hence, it should be a target group among new customers.

Data Exploration: Demographic Distribution

State-wise age group distribution

Type of Customer State-wise
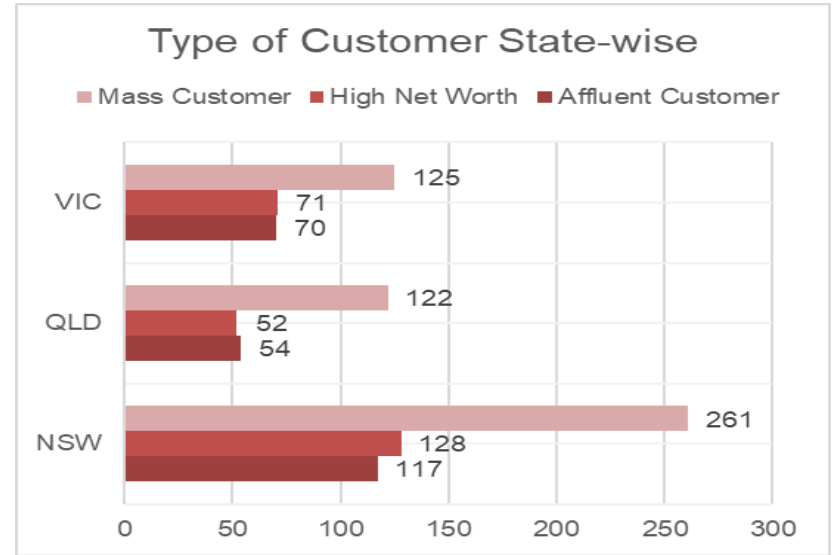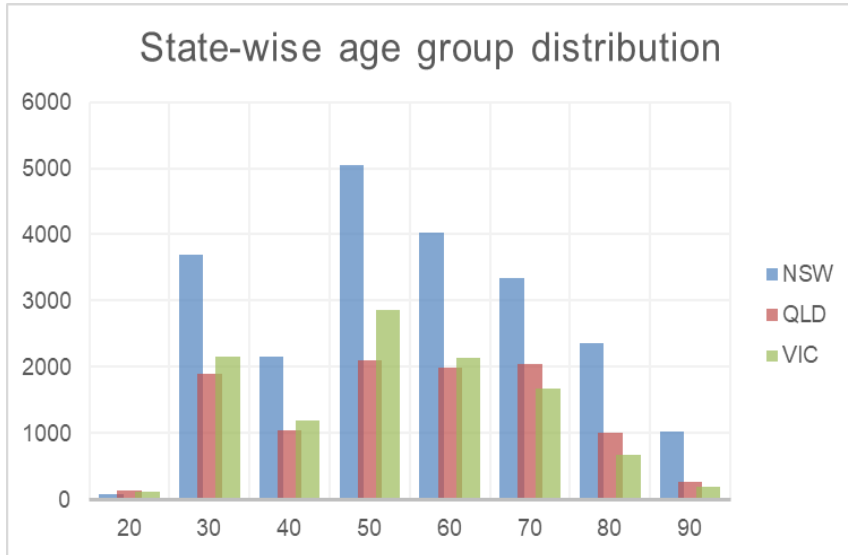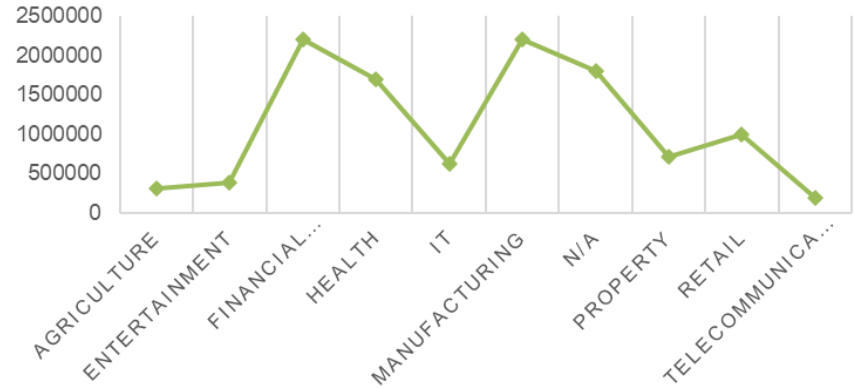
- **New South Wales (NSW)** has the highest number of customers followed by Victoria (VIC). NSW has almost **2x** the customers in every category. Thus, these two states should be of the highest priority.
- An interesting observation from second graph, most customers from every state are **Mass Customers**, hence we need to target new customers who could be classified as mass customers.

Data Exploration: Customer Job Industry Analysis

Industry-wise Sales Distribution

- Agriculture
- Entertainment
- Financial Services
- Health
- IT
- Manufacturing
- Property
- Retail
- Telecommunications
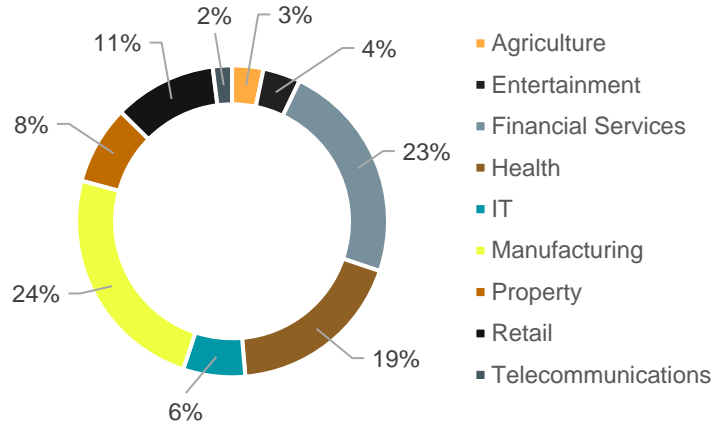
PROFIT VS CUSTOMER JOB INDUSTRY

- Customers involved in jobs related to **Manufacturing** and **Financial Services** have generated solid sales and highest profits for the company.
- They have jointly contributed to **$4.4 million** in profits. **Health sector** has had a generous contribution of **$1.75 million** in profits.
- Health industry employees are also a good source of revenue for the company.
- **Target: Manufacturing, Financial Services, and Health industries.**

Data Exploration: Miscellaneous Observations

Gender-wise sales distribution: New Customers

(chart showing values: Female 25212, Male 23765, U 859)

State-wise Car Ownership

(stacked bar chart — NSW: No 272, Yes 234; QLD: No 103, Yes 125; VIC: No 132, Yes 134)

Legend: ■ No ■ Yes
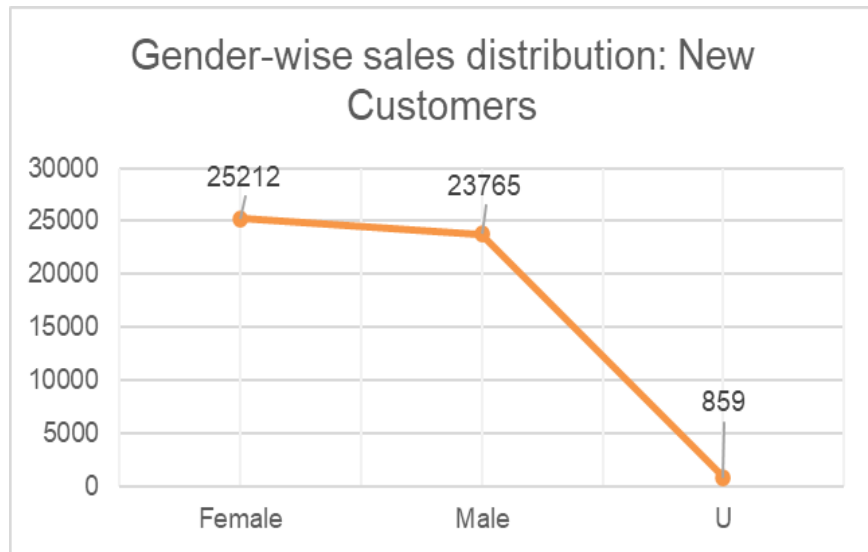
- The male and female numbers are constant while the irregularity with 859 customers as "U" is maintained for consistency of data.
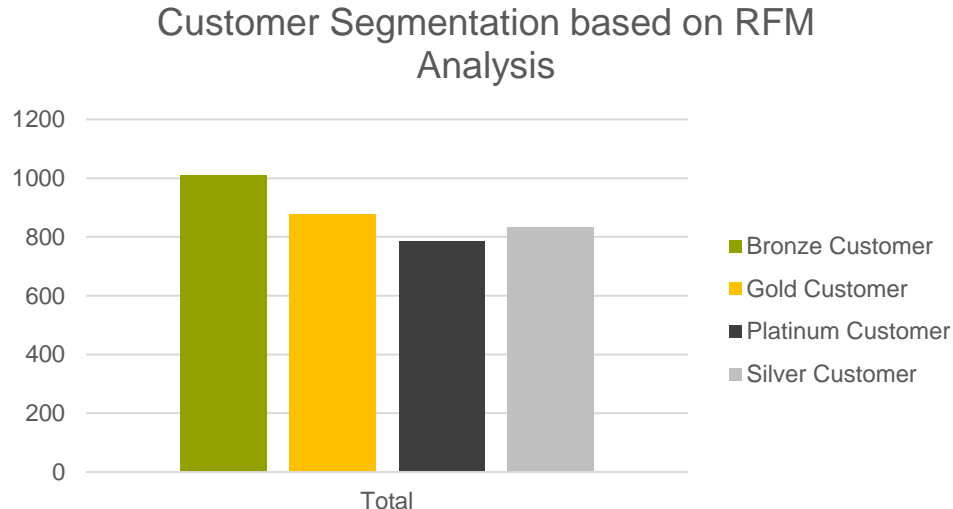- **Female customers** could be prioritised over male.
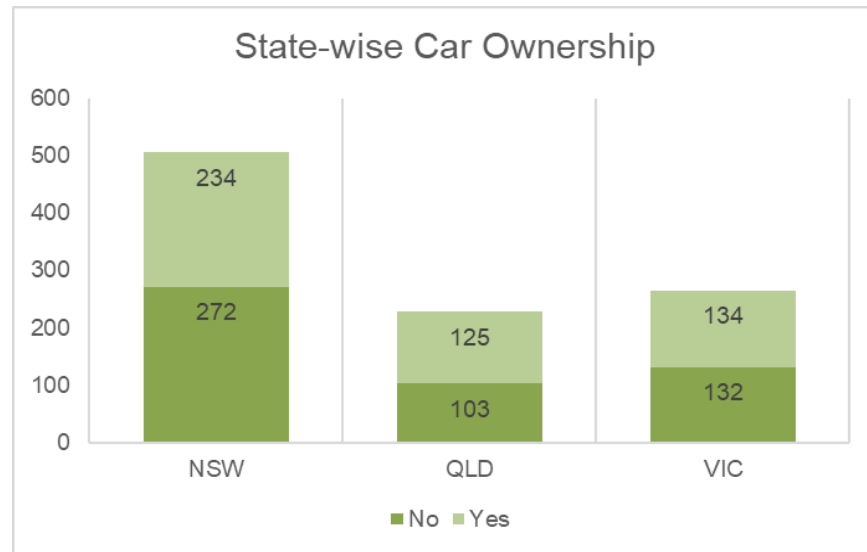
- Another observation to reinforce that **NSW** and **VIC** should be target states.
- NSW is a great opportunity because more than **50% of customers don't own cars.**

Customer Classification: RFM Analysis

- **Recency, Frequency, and Monetary Analysis** is a method of scoring customers based on how frequent, recent, and profitable their purchases were.
- Customers are rated on a score of 1-4, 4 being the highest. The highest score is awarded to the top 25%ile of customers.
- For ease of understanding, on the basis of these scores, the customers are classified as **Bronze, Silver, Gold, or Platinum.**

Customer Segmentation based on RFM
Analysis

Legend:
- Bronze Customer
- Gold Customer
- Platinum Customer
- Silver Customer

Interpretation

No.	Category	Inference
1.	**Platinum (RFM>=411)**	Most loyal, frequent, and high-spending customers. Low advertising needed, best customer support and sales services.
2.	**Gold (RFM>=311)**	They are recent customers with average frequency and who spent a good amount. Offer membership or loyalty programs or recommend related products to upsell them.
3.	**Silver (RFM>=211)**	They are customers who are not frequent shoppers. Start building relationships with these customers by providing onboarding support and special offers to increase their visits.
4.	**Bronze (RFM<211)**	These are customers who have not purchased since a long time, and have not been spending much. They are non-target customers.

Platinum, Gold are customer segments that we need to target and retain. They are the most loyal, frequent, and profit generating customers.

Customer Classification – Targeting High Value Customers

● **These are the high value customers that should be targeted from the new list:**

- Who are currently living in **New South Wales and Victoria.**
- Aged between **40–50.**
- Working in **Financial Service, Manufacturing and Health.**
- Most of the high value customers are female compared to male.

Interpretation

List of few customers that will come under the high-value customer classification

Customer ID	Bike Related Purchases for the last 3 years	Age	Job Industry	Wealth Segment	Owns Cars	State
1842	445	44	Financial Services	Mass Customer	No	New South Wales
2001	168	44	Manufacturing	Mass Customer	Yes	New South Wales
650	486	44	Health	Mass Customer	No	New South Wales
3297	234	46	Manufacturing	Mass Customer	No	Victoria
50	266	41	Manufacturing	Mass Customer	Yes	New South Wales

Thank you!