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
5. Recommendation

Introduction

Sprocket Central Pty Ltd

- **Background:** a long-standing KPMG client who specializes in high-quality bikes and accessible cycling accessories to riders
- **Request:** use existing customer datasets to determine customer trends and behavior
- **Goal:** recommend which of 1000 new customers should be targeted to drive the most value for our client
- **Data Analysis Approach:** use following customer analysis and segmentation models to carefully assess existing customers in different dimensions
    - RFM (Recency, Frequency and Monetary value) Segmentation Analysis
    - Demographic Segmentation (Age, Gender, Wealth Status)
    - Geographic Segmentation (States)

Data Exploration

Data Quality Assessment and Data Integrity Improvement

- **Factors that reduce data integrity:**

  3 Datasets not in sync: customers in 'Transactions' and 'CustomerAddress' datasets but not found in 'CustomerDemographic' dataset (merge 3 datasets to generate a dataset that only contain customers who exist across 3 datasets)

  Completeness: 'Transactions' and 'CustomerDemographic' datasets contain missing values for certain columns (if those records do not create any material difference to the later analysis, they should be excluded from the datasets)

  Inconsistency:

  for certain columns of all 3 datasets, there are multiple representations of the same value;

  'Transactions' and 'CustomerDemographic' datasets have inconsistent datatypes for values of the same column

Model Development

Recency, Frequency, and Monetary Value Segmentation Analysis

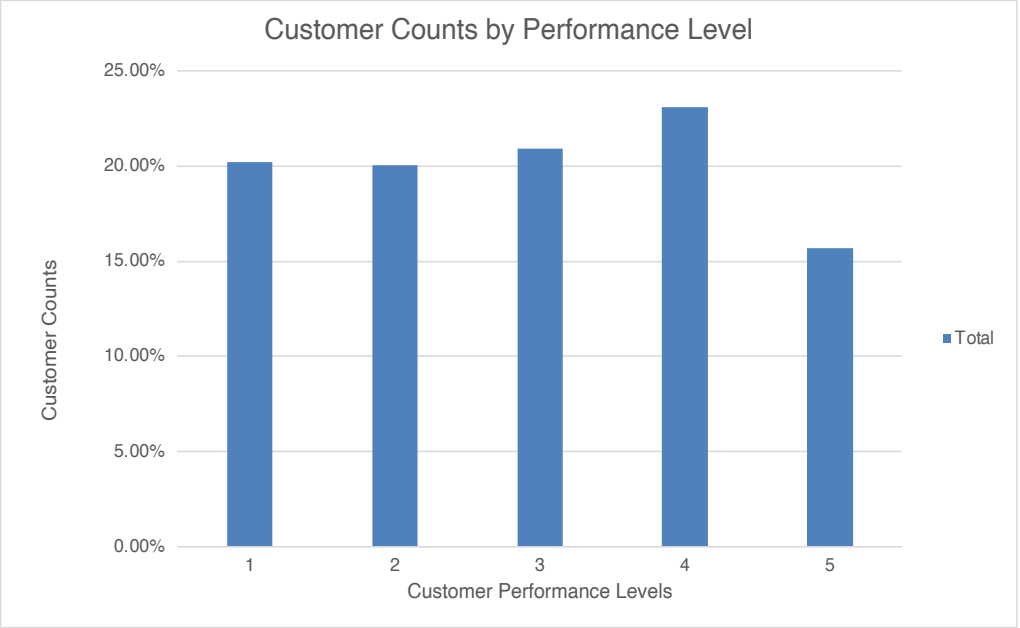
- Recency: How many day(s) is the last order date of an existing customer away from 12/30/2017?
- Frequency: How many transactions each existing customer had placed before 12/30/2017?
- Monetary: How much total profit had been generated by each existing customer before 12/30/2017?

- Categorize existing customers on 'Transactions' table into 5 performance levels based on their R score, F score and M score calculations by using 20 percentile, 40 percentile, 60 percentile, 80 percentile and 100 percentile.
- RFM Value = Weighted %* R score + Weighted % * F score + Weighted % * M score
- Use RFM Value to determine the performance level of an existing customer
- 5 as being the highest level while 1 as being the lowest
- Perform analysis on those level 5 customers for more detailed insights of customer characteristics in order to make marketing target decision.

Interpretation

Findings of RFM Analysis

- Almost ¼ of the existing customers are considered level 4 based on RFM value generated from 50% weighted of recency score, 30% weighted of frequency score and 20% weighted of monetary score
- Only 15.7% of existing customers has highest performance

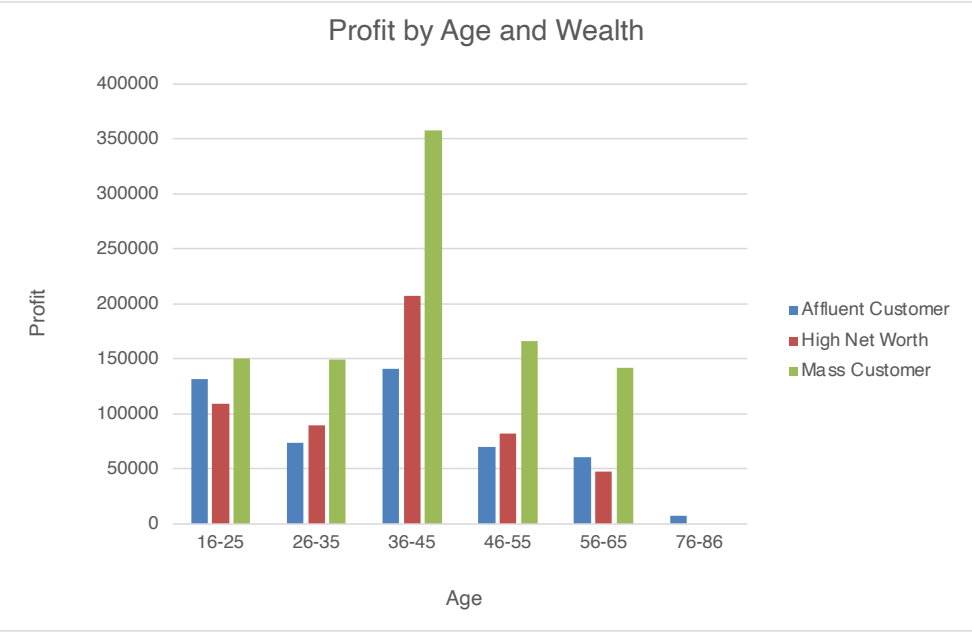


Demographic Segmentation – Age, Wealth and Profit

- Which age cluster and wealth status of level 5 customers had generated the most profit for the organization?
- Calculate age of each existing customer by subtracting their date of birth from 12/30/2017
- Each age cluster contains ages within 10 years

Interpretation

Findings of Age and Wealth Segmentation



- Most of profit are generated from customers who are at average wealth and age range from 36 to 45 years old across different wealth_segment

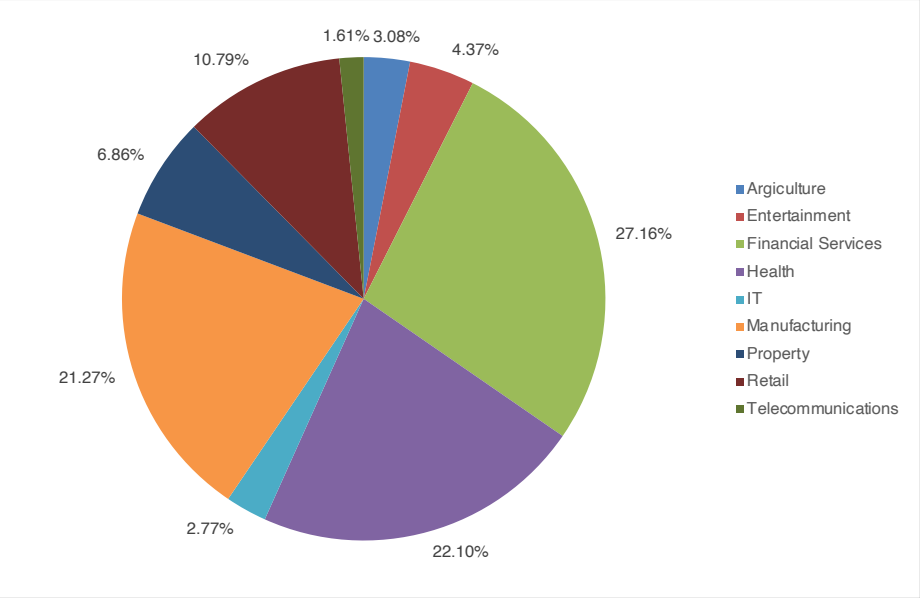
Demographic Segmentation – Customer Job Industries and Profit

- Determine which customer job industry are tend to create more profit for the organization
- Group level 5 customers by different job industries
- Level 5 customers are from 9 main job industries: Agriculture; Entertainment; Financial Services; Health; IT; Manufacturing; Property; Retail; Telecommunications
- Use a pie chart to demonstrate how many percentage of all level 5 customers are from each of these job industries

Interpretation

Findings of Job Industries Segmentation

- Top 3 job industries of level 5 customers are Financial Services, Health, Manufacturing



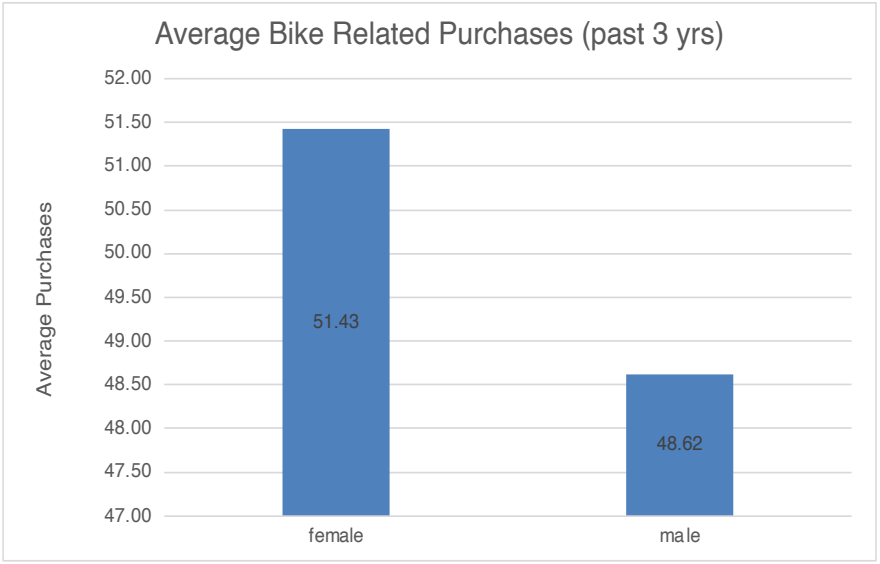
Model Development

Demographic Segmentation – Gender and Bike Related Purchases

- Determine if gender plays a role in bike related purchases from last 3 years
- If bike related purchase is a gender-neutral behavior, then both genders should have similar average purchases from the past 3 years
- Classified all level 5 customers with transactions into 2 gender groups
- Add all purchase numbers up from each distinct female customer in the past 3 years and divided by total distinct female customers, will result in average purchases from female customers in the past 3 years
- Add all purchase numbers up from each distinct male customer in the past 3 years and divided by total distinct male customers, will result in average purchases from male customers in the past 3 years

Interpretation

Findings of Gender and Bike Related Purchases Segmentation



- Female customers are likely to make more bike related purchases than male

Model Development

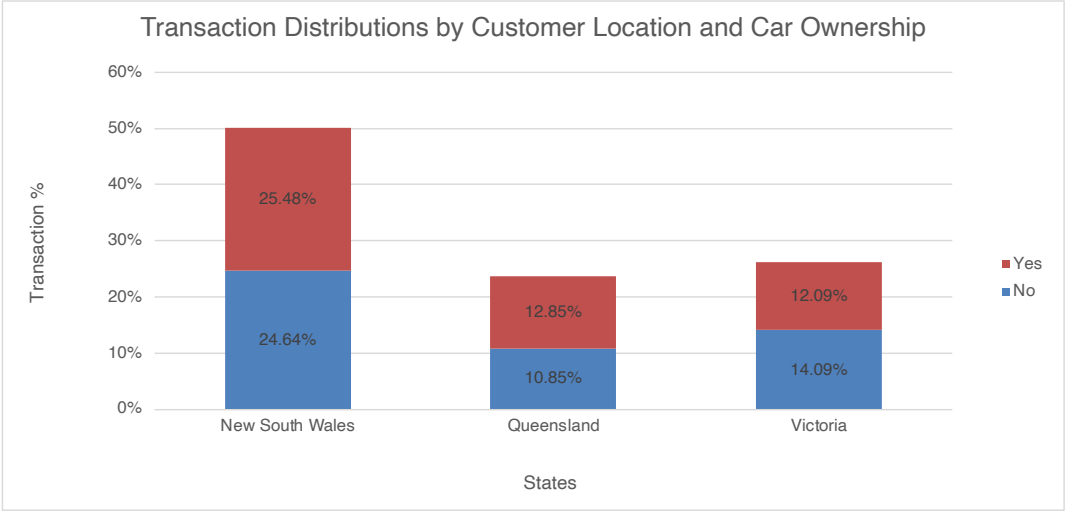
Geographic Segmentation – Car Ownership and States

- Which state tend to have more level 5 customer than others?
- Sprocket Central has level 5 customers mainly from 3 states: New South Wales; Queensland; Victoria
- At the same time, we also want to know if owning a car will make any different on making the purchase decision
- Using stacked bar chart can clearly visualize the behavior of level 5 customers from different state as well as for those who own and not own a car

Interpretation

Findings of Car Ownership and State Segmentation

- Half of level 5 customer transactions are made by who live in New South Wales
- Level 5 customers who own a car in New South Wales have slightly higher transaction rate than who don't own any cars
- It is opposite for the state of Victoria



Recommendation

New targeting customers should have these characteristics:

- Most of the high value customers will be female compared to male
- Aged range from 36 to 45 across wealth_segment
- Working in Financial Services, Manufacturing and Health industries
- Currently live in state of New South Wales and own a car

first_name	last_name	gender	past_3_years_bike_related_purchases	DOB_corrected	age	job_title	job_industry_category	wealth_segment	state	country
Sybilla	MacCart	Female	88	1987-01-15	36.00	Paralegal	Financial Services	Mass Customer	NSW	Australia
Martelle	Tuppeny	Female	52	1981-02-03	42.00	Marketing Assistant	Manufacturing	Mass Customer	NSW	Australia
Patricia	Everix	Female	34	1978-02-19	45.00	Director of Sales	Health	Mass Customer	NSW	Australia
Lacy	Drance	Female	27	1978-02-05	45.00	Graphic Designer	Manufacturing	Affluent Customer	NSW	Australia
Daryl	Pauncefort	Female	12	1979-06-18	44.00	Community Outreach Specialist	Financial Services	Mass Customer	NSW	Australia
Loleta	Aberdalgy	Female	45	1981-02-15	42.00	Occupational Therapist	Health	Mass Customer	NSW	Australia
Kaylyn	Jakaway	Female	45	1980-07-30	43.00	Registered Nurse	Health	Affluent Customer	NSW	Australia
Amabel		Female	71	1981-09-14	42.00	Chief Design Engineer	Financial Services	Mass Customer	NSW	Australia
Ajay	Worham	Female	80	1979-09-30	44.00	Computer Systems Analyst I	Manufacturing	Mass Customer	NSW	Australia
Selle	Casper	Female	98	1978-03-27	45.00	Social Worker	Health	Mass Customer	NSW	Australia
Lynnell	Shoesmith	Female	44	1981-01-29	42.00	Occupational Therapist	Health	High Net Worth	NSW	Australia
Janaye	Eade	Female	23	1984-12-13	39.00	Chief Design Engineer	Health	Affluent Customer	NSW	Australia
Therese	Brotherhood	Female	30	1981-07-28	42.00	Food Chemist	Health	Mass Customer	NSW	Australia
Darlleen	Shalcras	Female	77	1980-09-14	43.00	Health Coach I	Health	Mass Customer	NSW	Australia
Adria	Van den Velde	Female	39	1978-07-10	45.00	Nuclear Power Engineer	Manufacturing	Affluent Customer	NSW	Australia
Joane	Caldes	Female	84	1978-03-17	45.00	Senior Cost Accountant	Financial Services	High Net Worth	NSW	Australia