TheAnalyticsTeam

Sprocket Central Pty Ltd

Data analytics approach

KPMG Data Analytics Team – Claudia Wang

Agenda

1. Introduction

2. Data Exploration

3. Model Development

Introduction

This project aims to analyze the existing customer dataset to discover insights for customer trends and behavior and make recommendations on 1,000 potential customers to maximize values.

Three datasets for existing customers:
- Customer demographic;
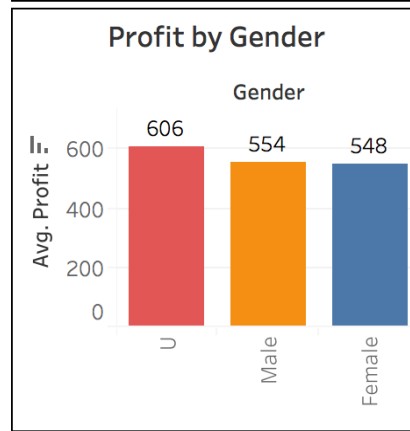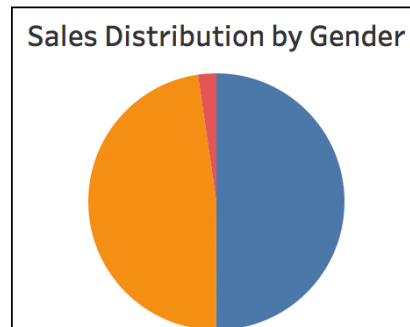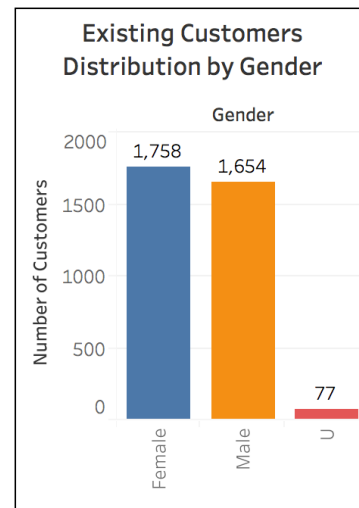- Customer address;
- Transactions.

One dataset for new customers:
- 1,000 new customers

Data Exploration

1. Gender.

- Quantity: Female > Male > U;
- Most sales generated by Female;
- Although Female made most sales, the average profit yielded by them is lowest.

Sales Distribution by Gender

Existing Customers Distribution by Gender

Profit by Gender
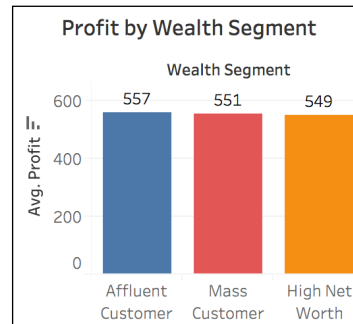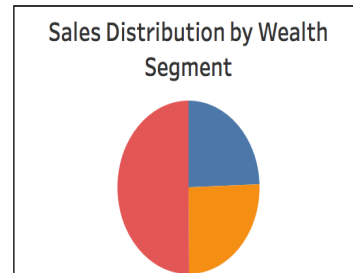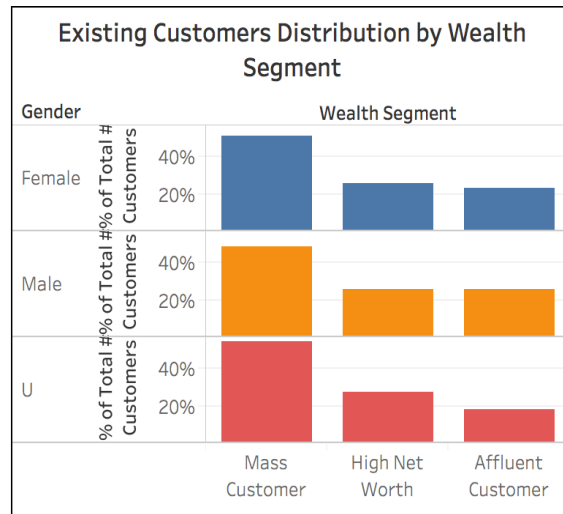
Data Exploration

2. Age.

- Divided age by 6 groups;
- Slight difference in age distribution for Female and Male;
- More sales generated by age '36 to 45' and '46 to 55';
- Customers who are 25 and under made the highest average profit.

Existing Customers Distribution by Age Group

Sales Distribution by Age

Profit by Age Group
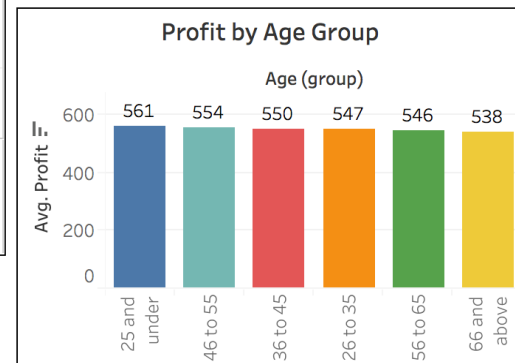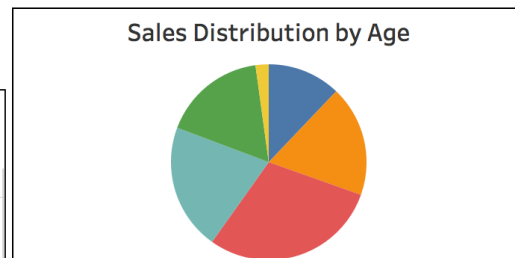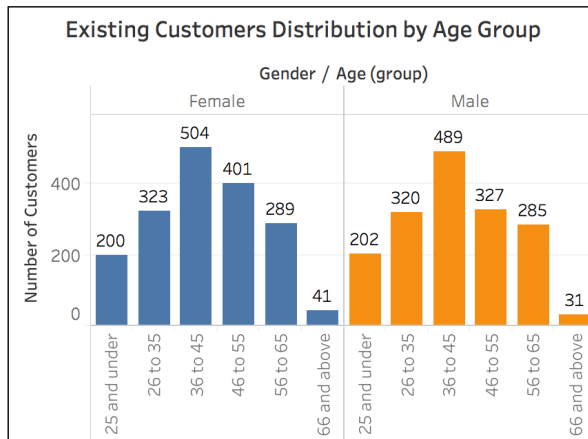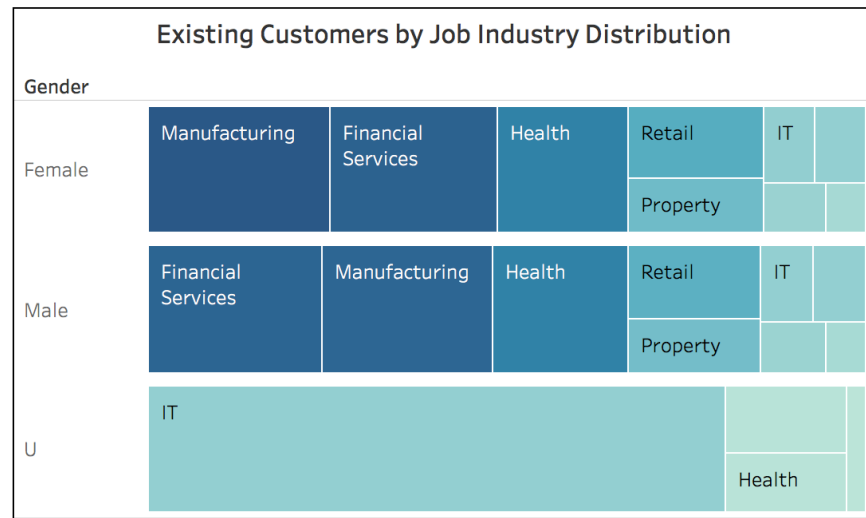
Data Exploration

3.Wealth Segment.

- Most sales yielded by 'Mass Customer';
- Sales from 'High Net Worth' and 'Affluent Customer' were similar;
- 'Affluent customer' made most average profit.

Existing Customers Distribution by Wealth Segment

Sales Distribution by Wealth Segment

Profit by Wealth Segment

Data Exploration

4. Job Industry.

- The job industry distribution for Female and Male is similar, but that for U is different;
- Most U customers work in IT industry.

Existing Customers by Job Industry Distribution

Data Exploration

## 5. Past 3 Years Bike Related Purchases.

- For Female and Male: there is a slight weak negative relationship between sales and past 3 years bike purchases;
- For U: there is a weak positive relationship between sales and past 3 years bike purchases.
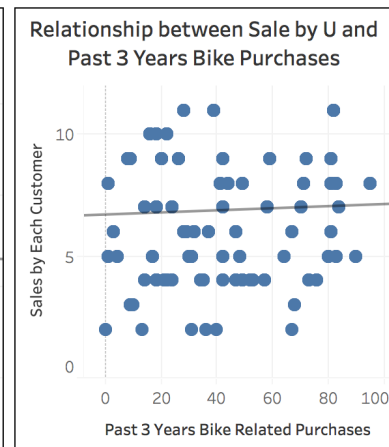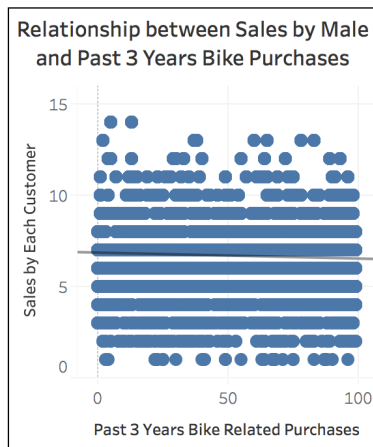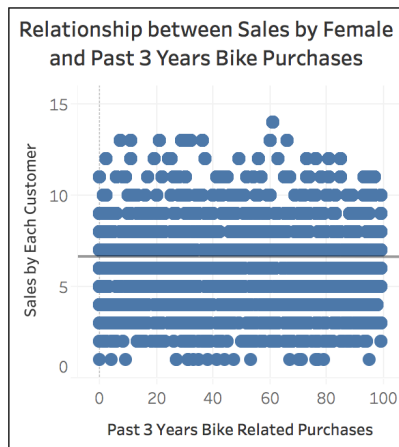
Data Exploration

## 6. Number of Transactions Made by Each Customer.

- For Male and Female, over 30% customers made 5 or 6 transactions during the past year;
- The range of transactions for each U customer is 2 to 11;
- The range of transactions for each Female and Male customer is 1 to 14.

Number of Transactions Made by Each Customer

Number of Transactions by Each Customer
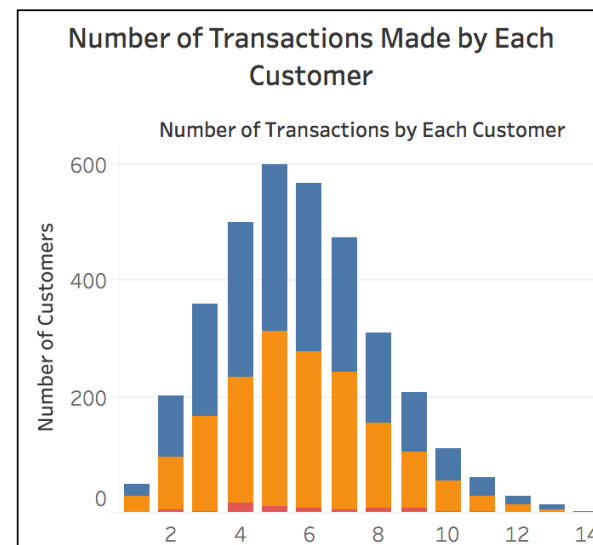
Data Exploration

7. Graphic Distribution.

- Top customers are from Liverpool and Blacktown;
- Liverpool and Blacktown are important for marketing campaigns.

Existing Customers Geographic Distribution
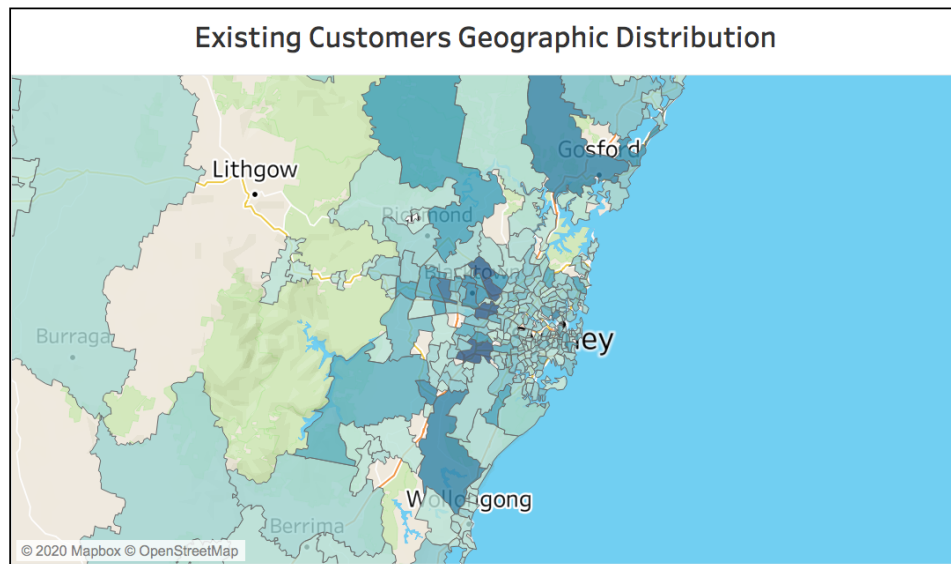
Model Development

Logistics regression.

We will develop logistics regression model to build customer segmentations on potential customers.