TheAnalyticsTeam

Sprocket Central Pty Ltd

Data analytics approach

Created by Tram Nguyen

Agenda

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

Introduction

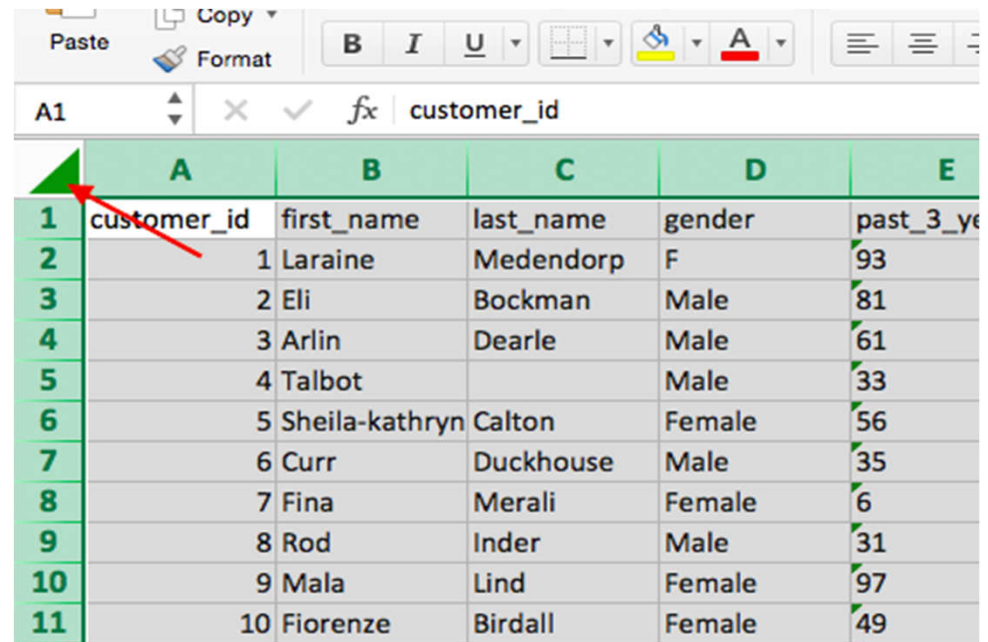**Module task 1**   **Module task 2**   **Module Task 3**

Task1 : identify all data quality issues.

- **Data incompleteness**
- **Data inconsistency**
- **Data Inaccuracy**
- **Data not coming from the same source**

Recommendations are attached in email to clients.

Data Exploration

Module task 2 : Exploration into customer demographics, geographics, psychology, and behaviour. **Links**
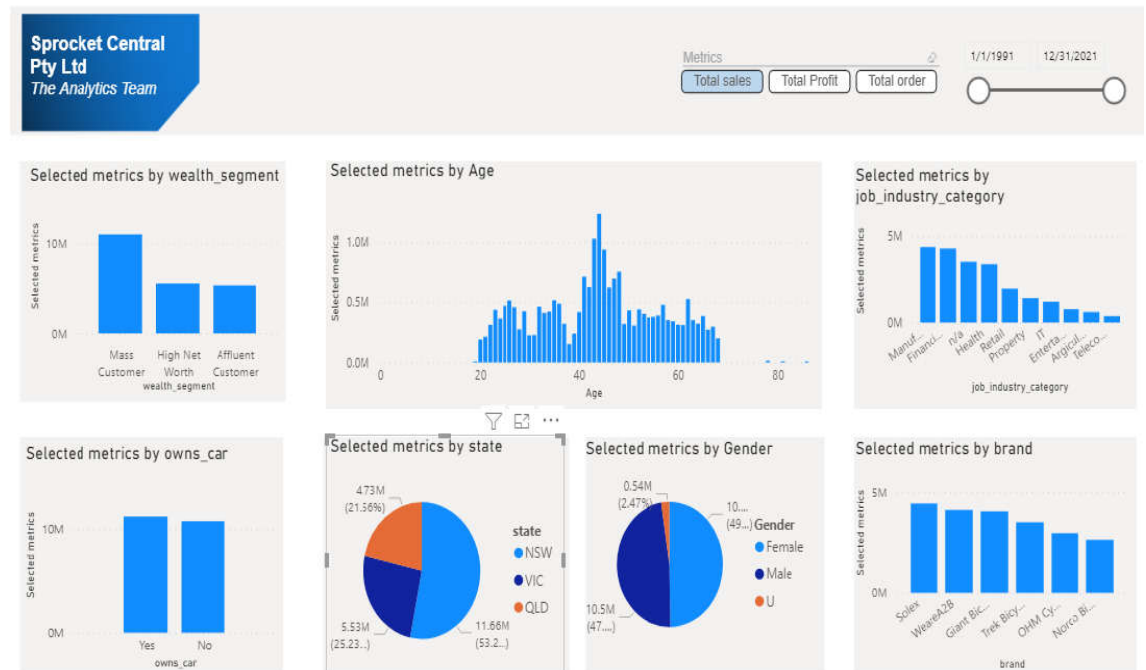
Demographics metrics:
- Age: 20-70 focus group: 40-50 (highest frequency)
- Wealth segment: Mass customer occupies the largest proportion of sales and profit of 10 million and 5.5 million.
- Gender: almost the same distribution between male and female customers.
- Job industry: Customer working in manufacturing, financial service and health sector generate the highest sales and profit (75%),
- Car ownership: 50% of customer owning car.

Geographics metrics:
- Most of sales derived from New South Wales (53%), Victoria(25%), Queensland(21%)

Data Exploration

Module task 2 : Exploration into customer demographics, geographics, psychology, and behaviour.

Psychology metrics:
- Customers prefer buying products belonging to normal brands Solex, WeareA2B, Giant Bicycle. Other special purposed brands are demanded by a smaller group of customers.
- Standard product line, medium product class and medium product size are the main product attributes when it comes to sales and profit.
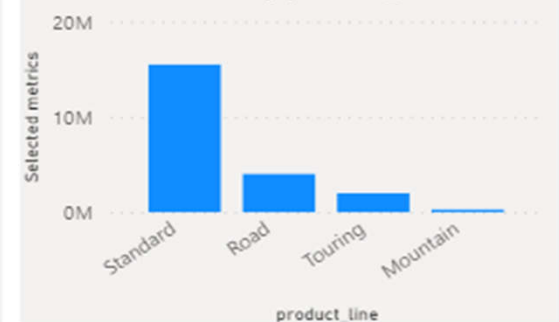
Behaviour metrics:
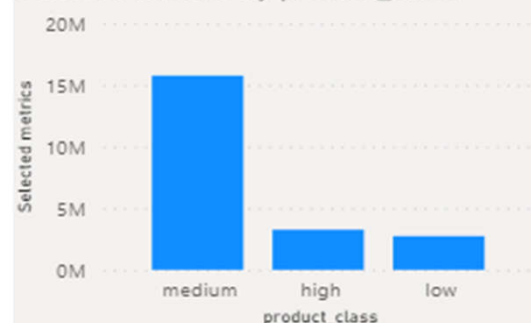- There is a balance distribution between online and offline channels.
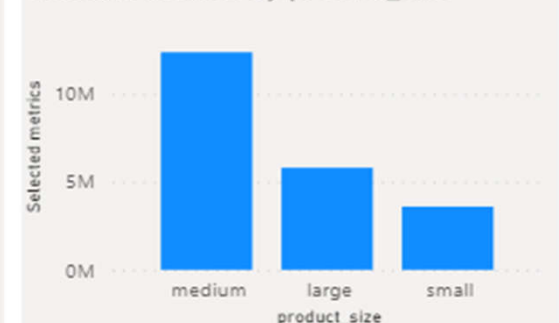

Selected metrics by brand
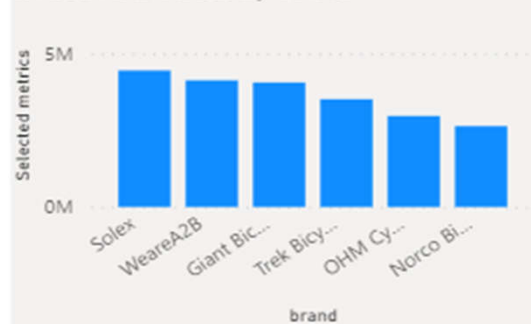
Selected metrics by product_line

Selected metrics by product_class

Selected metrics by product_size

Model Development

Module Task 3: Business model using Recency, Frequency, Monetary Value (RFM) **Links**

To calculate **recency**, we need to find out most recent purchase date of each customer and see how many days they are inactive for.
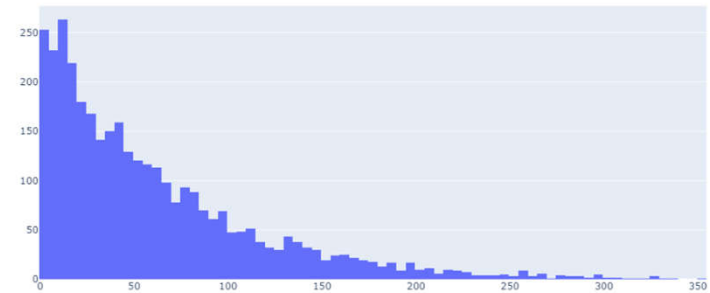
To create **frequency** clusters, we need to find total number orders for each customer.

Finally, we cluster them based on revenue (**monetary**) for each customer.

Recency

Monetary Value
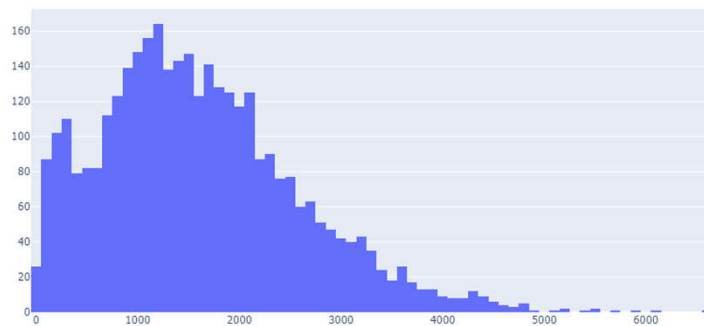
Frequency

Model Development

Module Task 3: Overall Segmentation [Links](#)

•**Low Value**: Customers who are less active than others, not very frequent buyer/visitor and generates very low - zero - maybe negative revenue.

•**Mid Value**: In the middle of everything. Often using our platform (but not as much as our High Values), fairly frequent and generates moderate revenue.

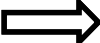•**High Value**: The group we don't want to lose. High Revenue, Frequency and low Inactivity.
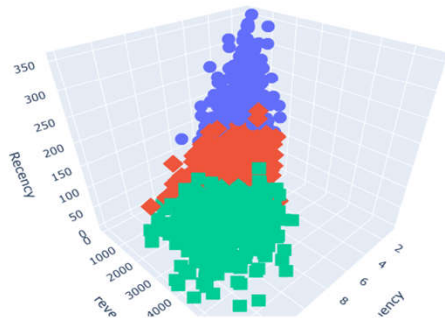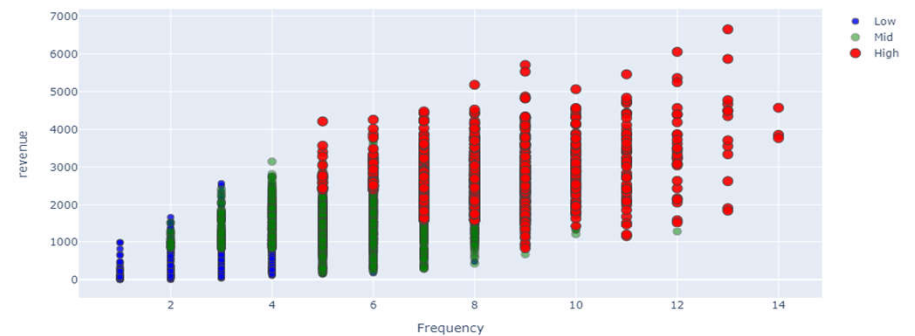
Interpretation

Module Task 3: Overall Segmentation [Links]

Cluster 1: low value customers having overall score from 0-2.
- Made last purchase from 100 days on earlier (more than 50 days inactive).
- Total revenue per user less than $800.
- Total purchase of less than 4.
- ⟹ Recommendation: Increase frequency.

Cluster 2: mid value customers having overall score from 3-6.
- Made last purchase within 40 and 70 days ago.
- Total revenue per user between $800 and $2000.
- Total purchase between 4 and 6.5.
- ⟹ Recommendation: Improve retention + Increase Recency.

Cluster 3: high value customers having overall score from 7-10.
- Made last purchase less than 40 days ago.
- Total revenue per user more than $2200.
- Total purchase more than 7.
- ⟹ Recommendation: Improve retention.

Appendix