

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

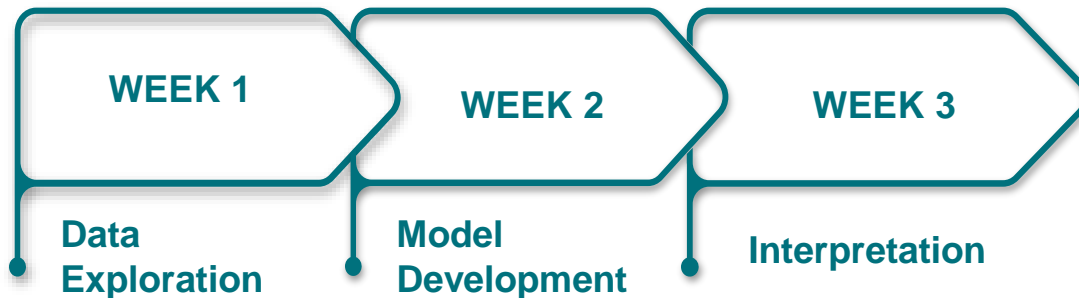
Introduction

Analyzing Customer & Transaction data

- Sprocket Central Pty Ltd needs help with its customer and transactions data.
- Plan to boost business:



- 3 week scope:



Data Exploration : Week 1

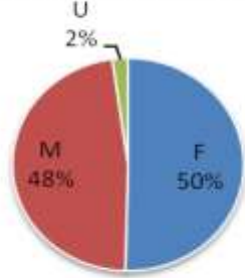
Data Quality Assessment identified following issues:

DATASETS	ACCURACY	COMPLETENESS	RELAVANCY	VALIDITY	CONSISTENCY
Customer Demographic	<ul style="list-style-type: none">•DOB column: Inaccurate values•Columns for age and age group to be added to dataset	6 columns i.e 46% columns has null values	Default Column: Irrelevant values		Gender column: Inconsistent values
Customer Address					States column: Inconsistent values
Transaction	<ul style="list-style-type: none">•Profit column to be added to the dataset	7 columns i.e 53% columns has null values		Data type incorrect for product sold column	

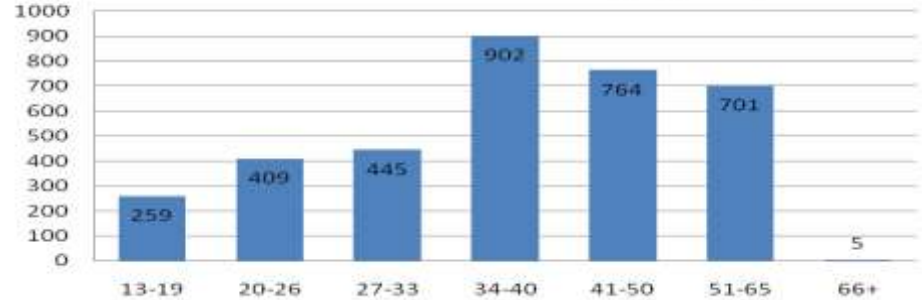
Additionally Sprocket Central Pty Ltd has given us a new list of 1000 potential customers with their demographics and attributes. However, these customers do not have prior transaction history with the organization

Data Exploration : Customer Distribution based on gender, age group, state and car ownership

Distribution based on gender



Distribution based on age group



Distribution based on state and car ownership



Insights:

Number of female customers is slightly more than males

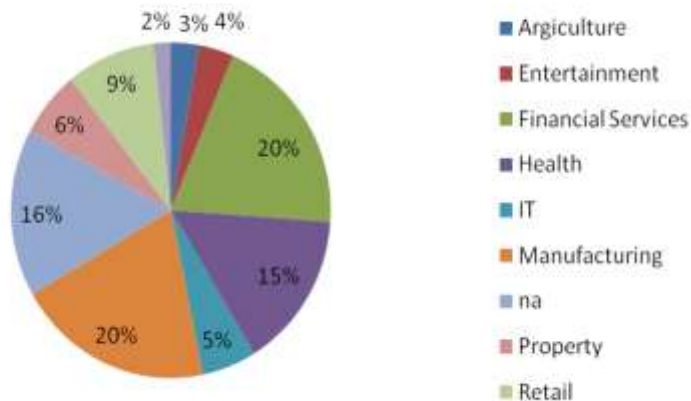
Age group 34-40 has high number of customers

NSW has higher number of customers

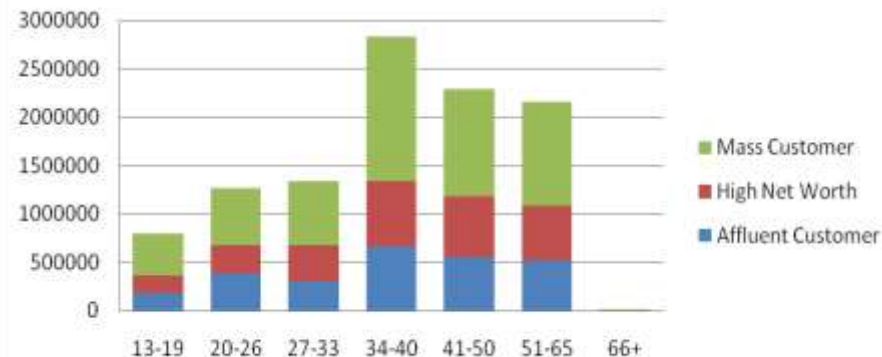
All the states have almost equal ratio of customers who Own and don't own cars

Data Exploration : Profit Distribution based on job industry, age group and wealth

Profit Distribution based on Industry



Profit Distributed based on age group and wealth



Insights:

Manufacturing, Finance and health industry shows higher profits

Profits is higher in mass customers in the age group of 34-40

Model Development: Week 2

- Determine a hypothesis based on business questions to check whether it is actually true or not.
- Identify the trends and the customer segment that has higher value
- Perform RFM analysis
- Train Machine Learning Models on the train data set and measure the accuracy of each model to choose one to deploy on the new Customer dataset.
- Deploy the chosen model on New Customers dataset and extract resulting recommended customers dataset

Interpretation: Week 3

Visualization and presentation of findings

Prepare a dashboard that will help visualize important trends, data distributions, business insights.

The dashboard will have filters to choose between brands, job type, months and states.

Appendix