

TheAnalyticsTeam

# Sprocket Central Pty Ltd

Data analytics approach

Analytics, Information and Modelling Team – Justine Picar

# Agenda

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

# Introduction

## TARGETING HIGH VALUE CUSTOMERS FOR SPROCKET CENTRAL PTY LTD.

Sprocket has been a long-standing KPMG client specializing in high-quality bikes and accessible cycling accessories.

### Goal:

Determine customer patterns and behaviors of 1,000 new clients and properly allocate budget and resources based on our findings for the appropriate market segments using the following data from Sprocket's marketing team:

- Customer Demographics
- Customer Addresses
- Customer Transactions from Jan-Dec 2017

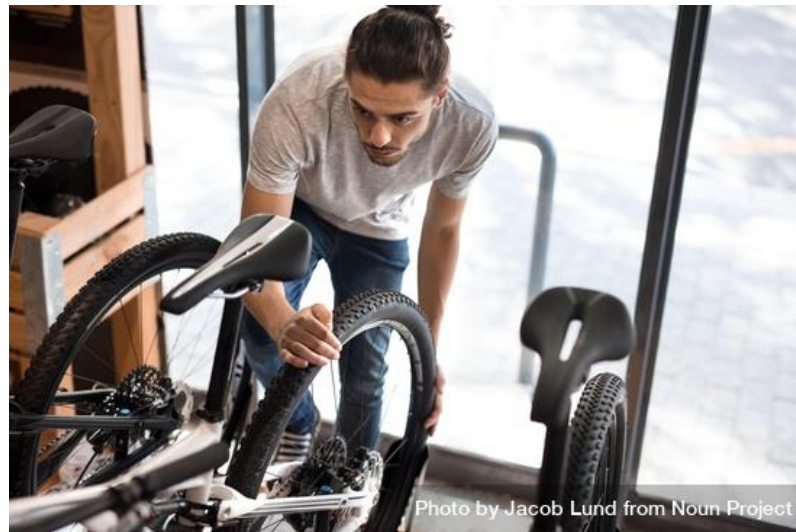


Photo by Jacob Lund from Noun Project

# About the Dataset

## Customer Demographics

- 4000 customers
- 12 usable features
- There are missing values, duplicates, and formatting issues

## Customer Addresses

- 3999 addresses
- 6 usable features
- There are some formatting issues

## Transactions

- 20,000 data points
- 13 usable features
- There are missing values and formatting issues

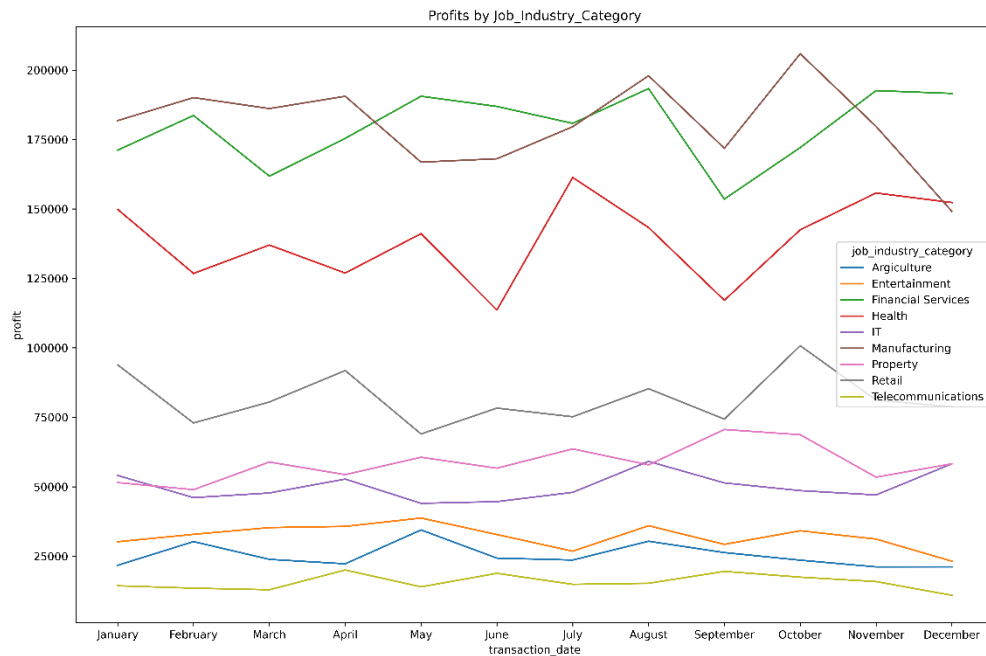
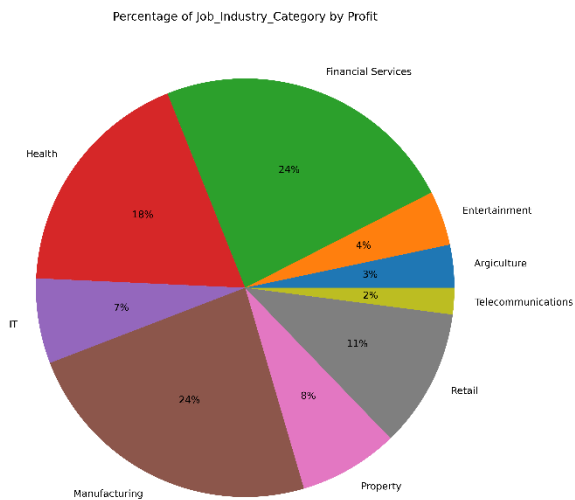
## Final Merged Dataset

- 20500 data points
- Reformatted values, and identified 5 additional customers from transactions and addresses
- Regrouped by features like tenure and age for more general granularity

# Data Exploration

## Profit by Job Industry

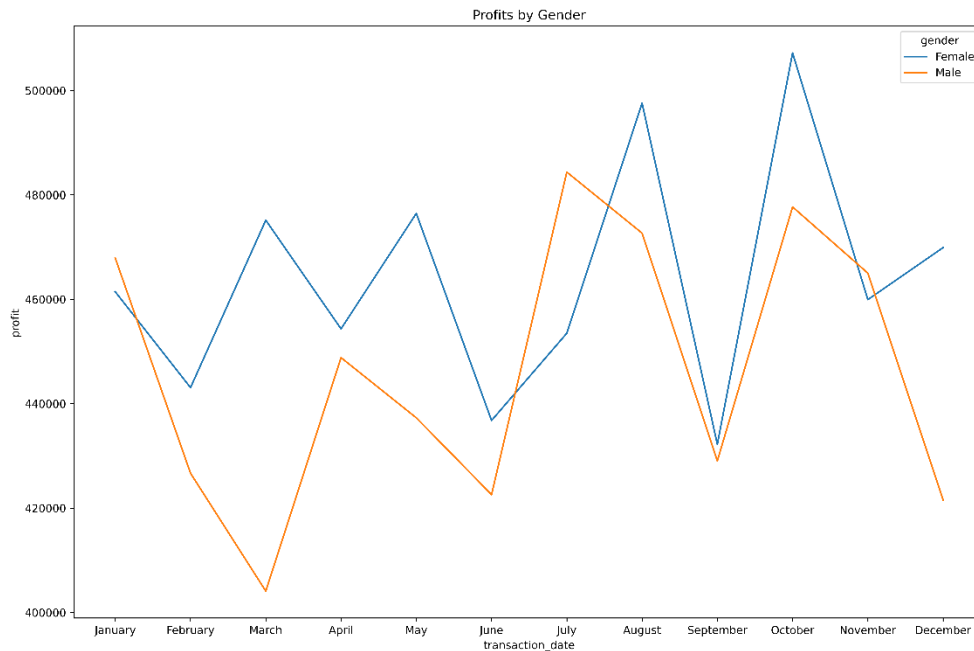
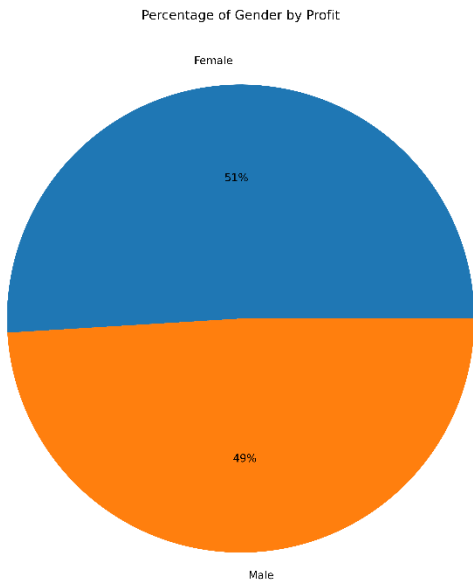
Manufacturing, Financial Services, and Health make up a large proportion of customers and profit



# Data Exploration

## Yearly Total Profit by Gender

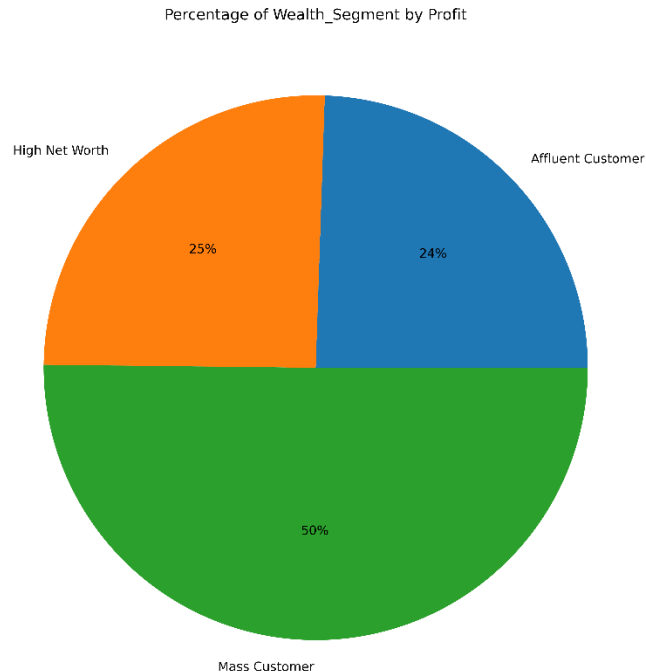
Women spend slightly more than men. However, we should still consider both market segments.



# Data Exploration

## Wealth Segment

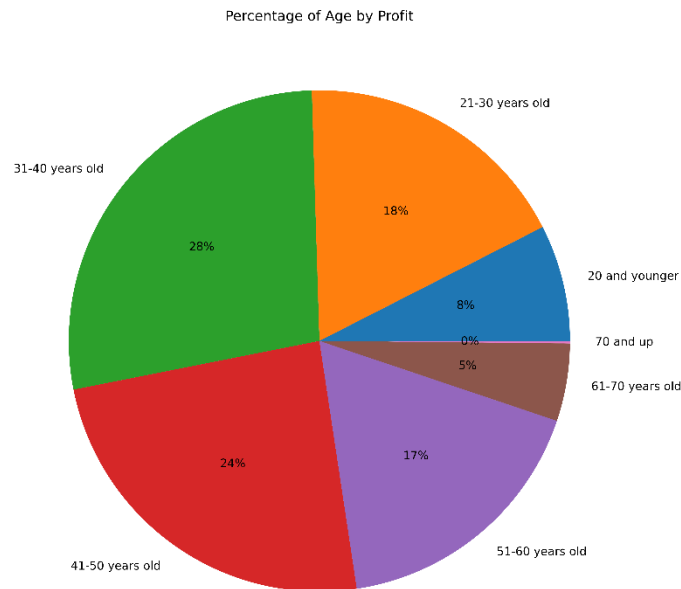
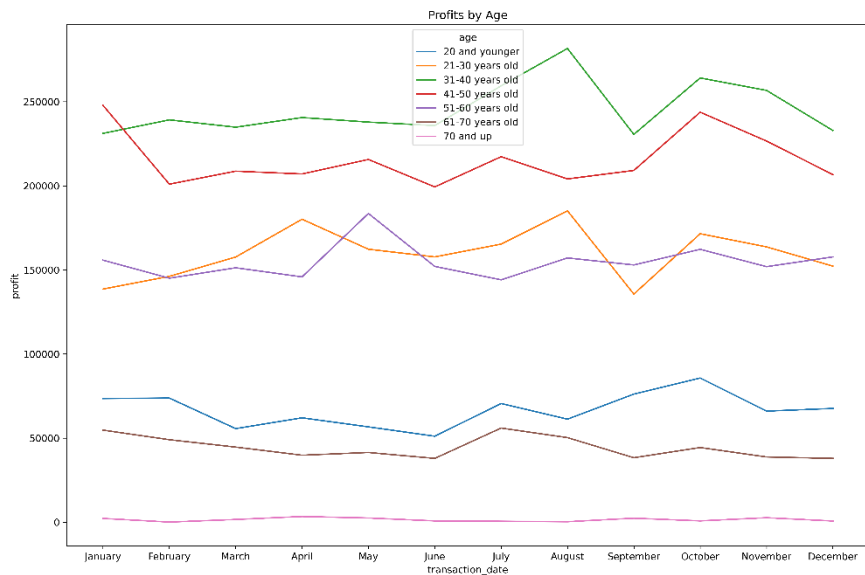
Mass Customers make up 50% of total profits in 2017



# Data Exploration

## Total Profit by Age Group

The largest market is between 30-50 years old at combined total profit of ~\$500k in 2017

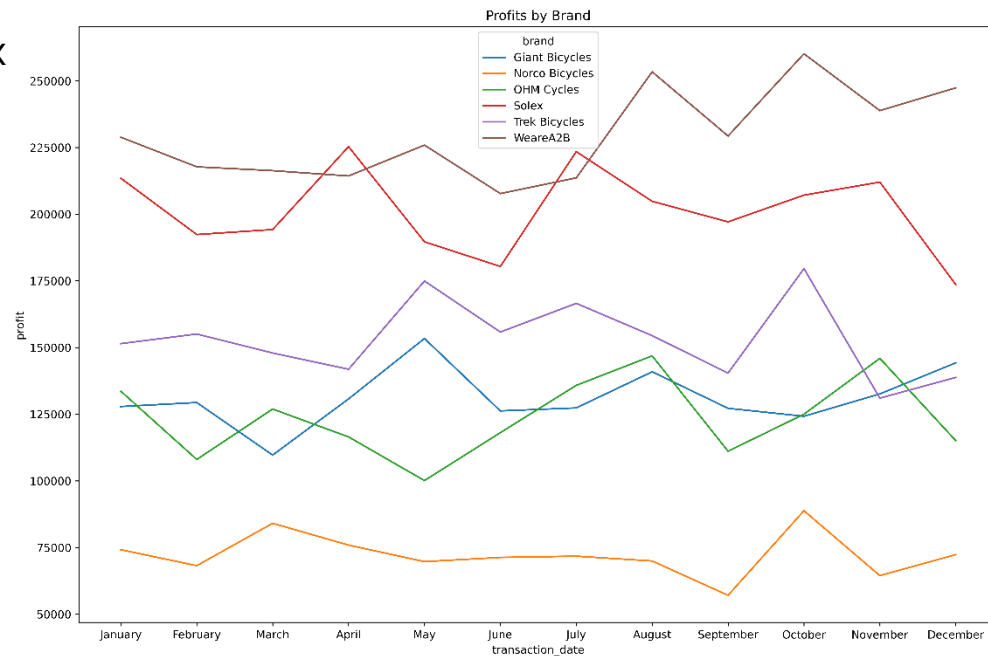
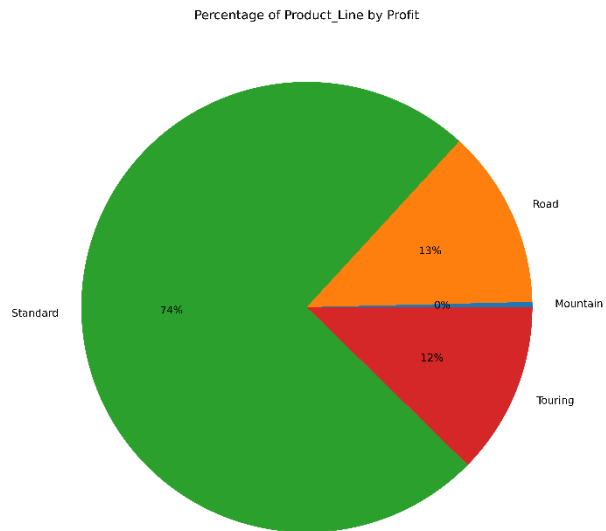




# Data Exploration

## Total Profit by Brand & Product Line

- Profitable brands: WEareA2B and Solex
- Profitable product line: Standard Bikes



# Model Development

## Tasks

- Feature engineering and Pre-processing
- Dropping irrelevant features, filling in missing values, and transforming the data
- Models will include Random Forest, Categorical Boosting, and KNN to predict which customers are most likely to make purchases
- We will determine which demographics are most important to focus our marketing strategy on

# Interpretation

## Marketing Strategy

- Most purchases are made later in the year around the spring time when most people are outdoors in Australia; Sprocket should begin preparing for this season and uptick in bike purchases
- Sprocket should focus on selling and increasing their inventory on Standard bikes from WeareA2B and Solex during this time period
- Based on total profits from 2017, Sprocket should focus on the following demographics: Customers in their 40s, Mass consumers, and customers that work in the Manufacturing and Financial industry

# Appendix

# Appendix

You can find the cleaned dataset and the full analysis on [GitHub](#).