

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

This is a presentation to describe the process of prospecting new customers for SPRocket Central Pty. Ltd.

This is a part of KPMG Australia Data Analytics Virtual Experience Program. A given situation for this case study is that, Sprocket Central Pty Ltd, a long-standing KPMG client who specialises in high-quality bikes and cycling accessories plans to expand the customer base. SPRocket task KPMG data analyst team to identify new prospect customers within a given customer list. To do that SPRocket would supply additional dataset to built a model to identify a prospect.

For reference, see the Forage project [page](#)

GOAL

Deliver a presentation and data dashboard to SPRocket Marketing Team

The Process

Here is the process we used to extract insight from the dataset.

1. Data Acquiring
2. Data Cleaning
3. Data Validation
4. EDA and Feature Engineering
5. Analyze the data
6. Deliver the insight presentation and Dashboard

Data Exploration

Some key insights from the EDA with 20k transactions from 2017

- Top 3 profitable brands are
 - WeareA2B 25% of total profit
 - Solex 22% of total profit
 - Trex Bicycle 17% of total profit
- Most profitable and best selling size is “medium”
- Best selling Product line is “Standard” bike
- Offline/Online sales are almost equal in both value and transaction number.
- No cyclical sales pattern throughout the year

Data Exploration (Continued)

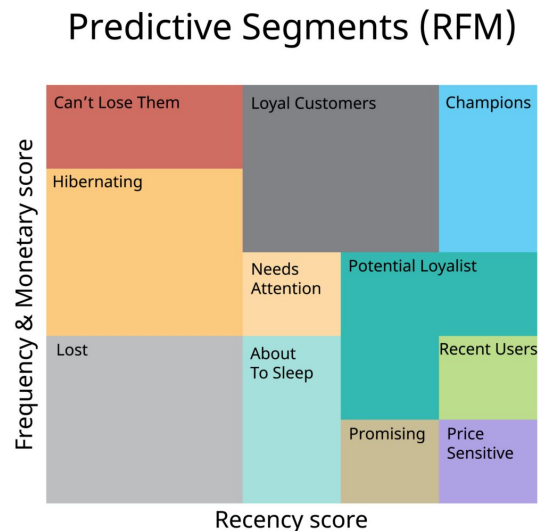
Some key insights from the EDA with 20k transactions from 2017

- Customer segment that contribute the most to the sales and profit is “Mass Customer”
 - With over 50% sales and profit contribution
- More than 50% of customers live in NSW, Australia
- Majority of the customers are in their 40's
- Top 3 customer's job industry are
 - Financial Service
 - Manufacturing
 - Health

Model Development

I plan to implement RFM model to rank the prospect base on a predicted profitability of a prospect.

- Customers will be rank using RFM model
 - Recency of purchases
 - Frequency of purchases
 - Money spent during the purchases
- RFM will act as a guideline to determine the segmentation of all prospects.
- Will utilize only Recency and Monetary score for model's simplicity



Data Exploration (Continued)

Some key insights from RFM model

- 24.5% of profit came from '22' customers
 - bottom 40% of recency score
 - bottom 40% of monetary score
 - median purchasing amount of '22' customer is \$456.45 per transaction
 - in fact, '22' customers made up to 29% of all customer
- 34% of profit came from '33' and '43' customers combined
 - bottom 80% of recency score
 - bottom 60% of monetary score
 - median purchasing amount of '33' is \$583.27 and of '43' is \$445.21
 - '33' and '43' made up to 23.1% of all customer

Data Exploration (Continued)

Some key insights from RFM model

- We suggest SPRocket Central should focus on these 3 groups of customer since there is a significant chunk of customer that participate in both good amount of sales and business activities.
 - 22
 - 32
 - 33
 - 43
- Other group, such as '44', '21' or '54', either have too small members or too small basket size.

RM_score_eng	SUM of profit_eng	SUM of profit	MEDIAN COUNTUNIQUE of customer_id
22	24.58%	\$2,715,536.61	\$456.45
32	16.20%	\$1,789,366.3	\$309.81
21	5.38%	\$593,924.18	\$167.21
33	17.02%	\$1,879,580.9	\$583.27
43	16.92%	\$1,869,196.91	\$445.21
11	2.07%	\$228,143.46	\$356.50
44	7.20%	\$794,914.98	\$690.49
42	2.50%	\$275,858.33	\$167.21
54	3.55%	\$392,300.60	\$448.68
53	1.47%	\$162,039.56	\$189.28
23	1.24%	\$136,479.61	\$812.44
31	0.27%	\$30,135.49	\$104.24
55	0.63%	\$69,749.65	\$542.87
12	0.21%	\$22,799.67	\$1,666.4
34	0.44%	\$48,522.34	\$1,544.6
45	0.34%	\$37,656.76	\$903.11
Grand Total	100.00%	\$11,046,205.4	\$445.21

Appendix