Predict Customer Personality to Boost Marketing Campaign by Using Machine Learning

Helmy Naufal Aziz

Overview

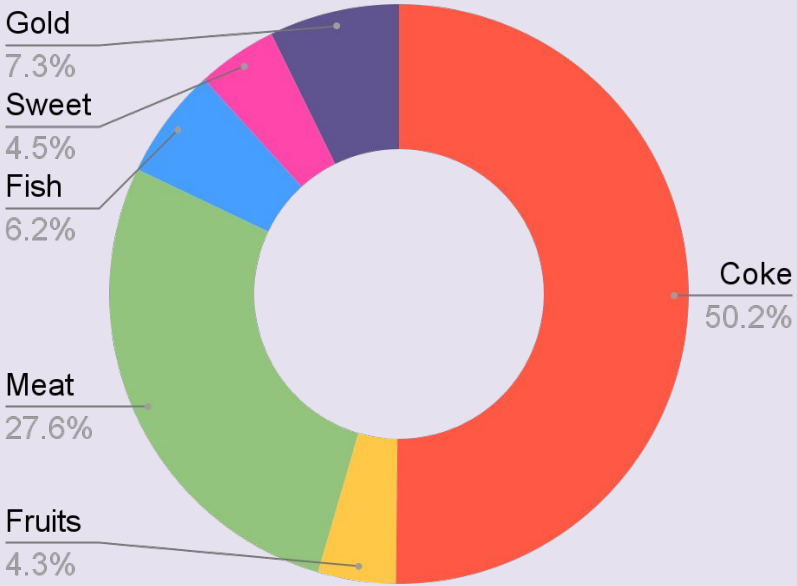
A company can grow rapidly when they know how their customer behaves, so they can provide better service and benefits to the customer that have the potential to be loyal customers. By processing historical marketing campaign data, a company can improve its performance and target the right customers so they can make transactions on the company's platform. From this data insight, we can focus on creating a cluster prediction model to make it easier for companies to make decisions.

01

Data Exploratory

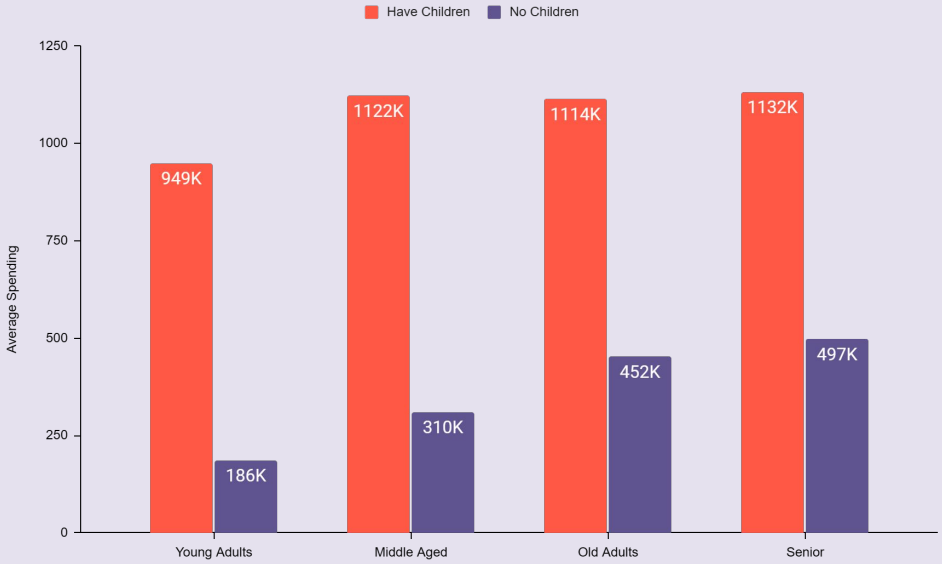
What insights do we get from the data?

Revenue Stream



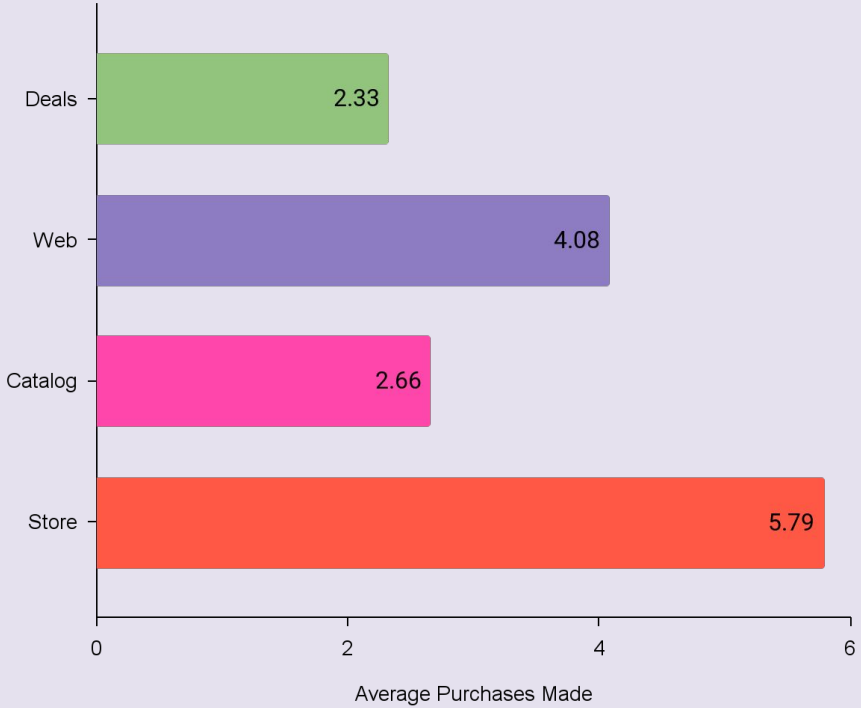
More than 50% of the company revenue comes from Coke product while Fruits and Sweet products have the lowest revenue. We can focus on these two lowest products for the next campaign, so we can gain more revenue from these products.

Impact of having children on the total spending in each age segmentation



Customers who have no children will spend more than those who have children, and the young adults customers have the lowest spending than other age categories. We can focus on targeting young adults customer to spend more on our products.

Customer Purchase Method



Most customers made purchases directly in the store, with an average of more than 5 times purchases made by each customer. It is probably because the customer can ensure the product before purchasing it. Meanwhile purchase made with discount and catalogue become the lowest method.

02

Preprocessing

Do the data clean enough to be processed?

Preprocessing

Data Cleaning

Clean data from missing values and duplicates data

Feature Selection

Remove unnecessary features

Feature Encoding

Transform the categorical values into numerical ones

Feature Transformation

Transform the data to the normal distribution

03

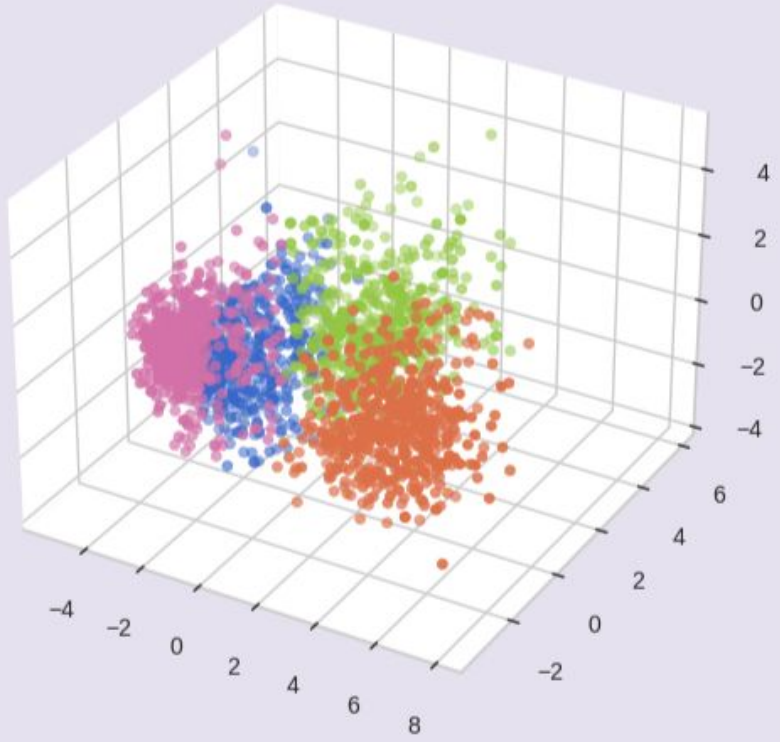
Data Modeling

How is the model performance?

Clustering

We will classify the customer by creating a cluster with the following steps:

1. Perform dimensionality reduction using **Principal Component Analysis (PCA)**
2. Determine the number of the cluster using **Elbow Method**
3. Label the customer cluster using **K-Means Clustering**

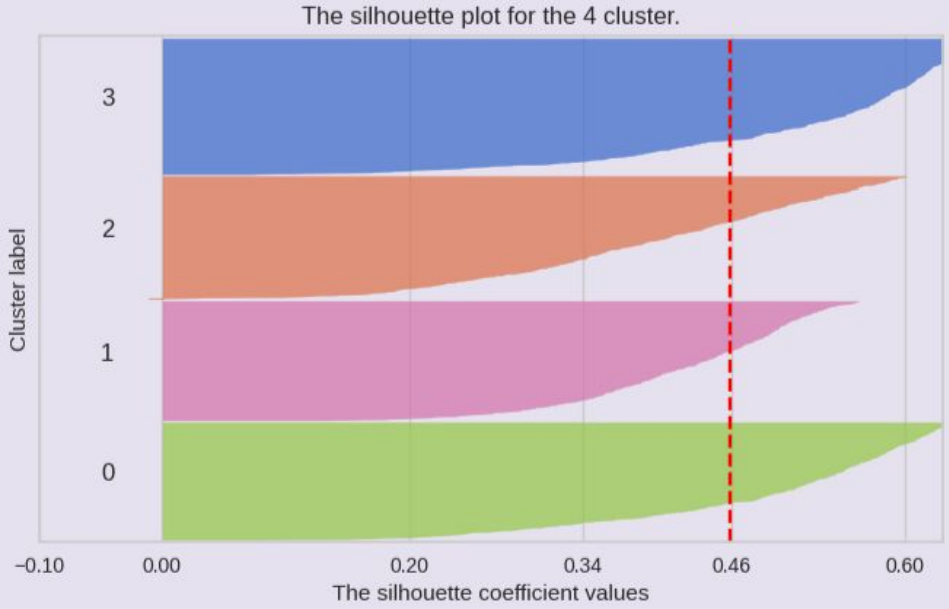


Evaluate Model

We use **silhouette score** to evaluate the model performance. From the graph we can find out several things:

- The average silhouette score is 0.46
- All the clusters are above the average silhouette score
- The cluster size are fairly distributed

We can conclude the model is good.



04

Interprating Cluster

What is the profile of each cluster?

Cluster Profile

After doing several data exploratory, we can conclude the profile of each cluster, as follows:

High Spender

- Customers are spread in all age ranges
- Most of the customers have no children
- Frequently make a purchase in the store and from the catalog
- Least group that use coupon with only 1 coupon used each month
- Least visited the website with only 2 visits a month
- Have a high income with 50 to 100 million per year
- Have high spending with average spending of 1,39 million

Moderate Spender

- Dominated by customers aged 31 (Young Adults) to 80 years (Senior)
- Most customers are parents with 1 - 2 children, with a slight majority of teenagers.
- Frequently visit the website with 5 visits per month
- Frequently make a purchase in the store and from the website with an average of 3 coupons used each month
- Have a moderate income of 30 to 80 million per year
- Have moderate spending with average spending of 867.7K

Cluster Profile

After doing several data exploratory, we can conclude the profile of each cluster, as follows:

Low Spender

- Dominated by customers aged 45 (Middle Aged) to 80 (Senior)
- Mostly a parent with 1 - 3 children, with a slight majority of teenagers
- Frequently make a purchase using coupons with more than 2 coupons used each month
- Frequently visit the website with 6 visits a month
- Have a moderate income of 30 to 70 million per year
- Have low spending with average spending of 156.2K

Risk of Churn

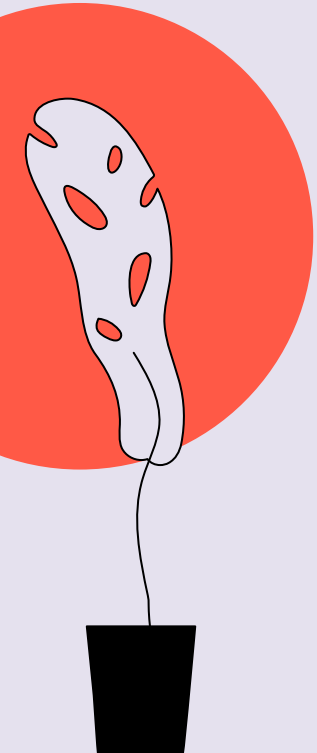
- Dominated by customers aged 27 (Young Adults) to 60 years (Old Adults)
- Have 1 - 2 children, with a slight majority of kids
- Most frequently visited the website with almost 7 visits a month
- Quite often use coupons with almost 2 coupons used each month
- Have low income with less than 50 million per year
- Have low spending, with average spending of only 94.2K

05

Recommendation

What action should be taken to solve the problem?

Recommendation



01 Focus on improving service to retain the High Spender customers so they are satisfied with our services and have a pleasant experience.

02 Give a personalized recommendation to increase the transaction of Moderate Spender customers. Optimize coupons used for this group, as this group uses more coupons than other groups.

03 Needs further analysis to increase the conversion rate for Low Spender and Risk of Churn customers, as these groups frequently visit the website but have low transactions.