Clustering Customers Analysis

By Maulana Shendi

https://www.linkedin.com/in/maulana-shendi/

About Data

Dataset about a supermarket that wants to do a clustering analysis of its consumers to find out the characteristics of consumers

- Dataset : https://www.kaggle.com/imakash3011/customer-personality-analysis
- Notebook : https://github.com/mpsmaul/my-project/tree/master/Customer%20Personality

Goals

- The model is expected to predict the characteristics of the buying habits of consumer products.

- Models can reduce marketing costs and increase marketing ROI

- Reduce 30% marketing cost

Model Introduction

- The model uses machine learning k means clustering

- There are 4 clusters obtained from the model

- The data is scaled with a standard scaler

- The previous model was carried out dimensional reduction with PCA

Alfa Cyan

- People size 826 (37%)
- Average age 55 years (median)
- Average income 68774 US Dollars
- Love all product in our market (Wine, Fruits, Meat, Fish, Sweet Product, and Gold product)
- Always visit the store to shop, not infrequently also shop through the web and catalogs

Beta Red

- People size 324 (14.4%)
- Average age 55 years (median)
- Average income 51550 US Dollars
- Love all Wine products, and low spending
- Always visit the store to shop and often visit the store website
- Happy to follow the company's campaign

Teta Orange

- People size 112 (5%)
- Average age 47 years (median)
- Average income 82092 US Dollars
- Love all products in our market (Wine, Fruits, Meat, Fish, Sweet Products, and Gold products) and usually buy in large quantities compared to other consumers
- Always visit the store to shop, not infrequently also shop with the web and catalogs. Relatively rarely visit the web when compared to other consumers
- Always enthusiastic about the company's campaigns

Delta Purple

- People size 976 (43.6%)
- Average age 48 years (median)
- Average income 33729 US Dollars
- Low spending compared to other consumers
- Always visit the store to shop, not infrequently also shop through the web and catalogs
- Rarely or never participates in company campaigns

Conclusion

- Companies can maximize services at offline and web stores because overall visitors prefer to shop through these two platforms

- For alpha cyan and teta orange, all products in the store can be offered

- For beta red, only wine products can be offered

- According to data, the most effective campaigns are campaign 4 and campaign 6.

Guesstimate Method for analysis potential advantages of the Supermarket

ROI = Profit – Marketing cost / Marketing cost

Before using clustering
ROI = 500.000 – 200.000 / 200.000 = 1.5 %

After using clustering
ROI = 500.000 – 140.000 / 140.000 = 2.5 %

Thank you and who do you think is the impostor !?!?