# How to Get Target Customers By Analyzing Historical Offer?

A data-based approach using Starbucks Simulated Offer Data.

## Introduction

Customers could react totally different towards certain offers from sellers. Proper marketing strategies can help raise sales tremendously. However, if you provide every member with a discount offer, it could even cut your profits since part of your consumers, most likely the high-income loyal group, never check coupon before spending money in your store. Therefore, getting know your customers and selecting discount-driven ones tend to be highly crucial.

How can we get target customer group? How can we decide on whether a transaction was motivated by a coupon? How does the target group look like? Driven by those questions, I started a data science project using simulated historical offer data sets provided by Starbucks.

There’re 3 different data sets in total:

* portfolio.json - containing offer ids and meta data about each offer (duration, type, etc.)
* profile.json - demographic data for each customer
* transcript.json - records for transactions, offers received, offers viewed, and offers completed

A screenshot of a cell phone

Description automatically generated A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

## PART I: Data Cleaning, Transformation and Combination

After checking null value for 3 data sets, I found there’re only missing values in Profile but those values, basically demographic information, don’t affect my logic, so I just ignored them. Then, to make data cleaner, I changed column name, remapped customer and offer ID and separate dictionary value to generate multiple columns.

Cleaning and transforming data in a short time, it’s time to integrate 3 data sets and combine every offer transaction information. The result data frame includes demographic, profile and each transaction from 2015 to 2019.

## PART II: Tease Out Non-Targeted People

Because the objective of this project was to get a consumer group responding best to certain offer types, it’s important to tease out ignorer since there transaction history could be considered in my analysis.

To realize that, I firstly selected customers who had both offers completed and transaction-only orders because only those customers could be non-targets. Then, I self-designed and calculated an offer-interact index as shown below:

The smaller the offer-interact index is, the less interaction there is between offers provided and customers. To tease out ignorer, I only selected users whose index is larger than (mean – 1\*std).

## PART III: Created K-means Clustering model to get target user group

K-means Clustering model assists me in separating all users to groups based on the variables related to my subject. In this case, I care most about user-offer interaction, so I used corresponding factors, acquired the best K and applied K-means clustering to get answers.

Based on my objective, my variables in training data include offer completed rate, offer view, offer received (without viewing and using) rate for 3 different offer type provided by Starbucks. Then, I used Elbow Method to get a best K as shown below:

A close up of a device

Description automatically generated

Although the trend is not highly apparent, I selected K=8 since there’s still a little slower trend compared to K<8.

Lastly, I used Sklearn package to create and fit a K means model and get 8 groups as a result.

A screenshot of a social media post

Description automatically generated

As a result, Group 1 responds best to Bogo and Discount offer and Group 4 is the next. Additionally, I’ve written every step into a data function. To get the best group, I can just input K and offer type to acquire customer profiles motivated best.

## Conclusion

In this project, I basically processed ETL to prepare for analysis, filtered data by personal assumption and applied unsupervised machine learning model to get our target group. There might be a more precise way to acquire offer ignorer and this could be my next step. Customer segmentation could be highly vital especially for giant sellers since they are able to provide multiple solutions to give their customers the best purchasing experience.

Hence, have you ever found your target users?