Market Segmentation from Customer Behavior

Drew Pickett

Introduction

The idea is to develop a customer segmentation to define a marketing strategy for a credit card company. The Dataset being used summarizes the usage behavior of about nine thousand active credit card holders during the last six months. The file is at a customer level with eighteen behavioral variables or features. Finding the optimal clustering model for the customer segmentation will be the final goal.

Problems

- Does the Data set need feature engineering? Preprocessing?

- How many clusters work best for this Data set

- Which clustering model works best?

- How can we interpret the data the clusters make?

- How can we visualize the data to make a marketing strategy?

Methodology

First, we had to go through the Data and do some feature engineering. This dataset provided some missing values in some of its features, so we must iterate through the Data and replace those values with the mean value of each feature. The Data also had some outliers, so we need to find and delete the outliers. Finally we can scale and normalize the Data so the clustering model can easily interpret the Data.

Now we can use the elbow method to find the correct number of clusters to group the data into. For this instance six clusters seemed to be the optimal choice because the inflection point of the cluster score is at exactly six.

From here we can now implement the K-means Clustering model with the parameters of having six. In order to interpret the Clusters we need to print them out into histograms and perform analysis on what they could mean.

Results

Dataset

|	CUST_ID	BALANCE	BALANCE_FREQUENCY	PURCHASES	ONEOFF_PURCHASES	INSTALLMENTS_PURCHASES	CASH_ADVANCE
0	C10001	40.900749	0.818182	95.40	0.00	95.4	0.000000
1	C10002	3202.467416	0.909091	0.00	0.00	0.0	6442.945483
2	C10003	2495.148862	1.000000	773.17	773.17	0.0	0.000000
3	C10004	1666.670542	0.636364	1499.00	1499.00	0.0	205.788017
4	C10005	817.714335	1.000000	16.00	16.00	0.0	0.000000

Elbow Function
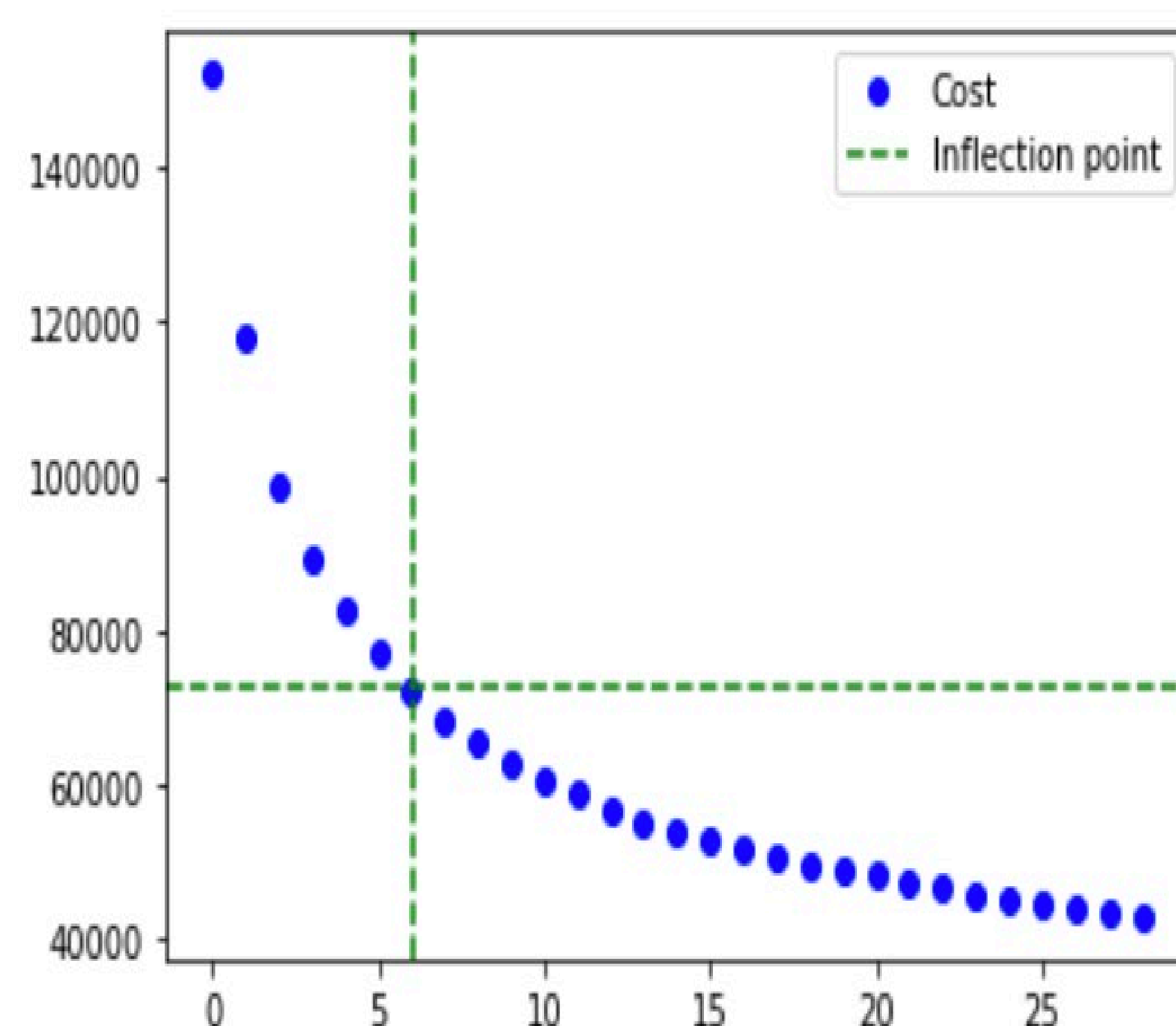
Clusters
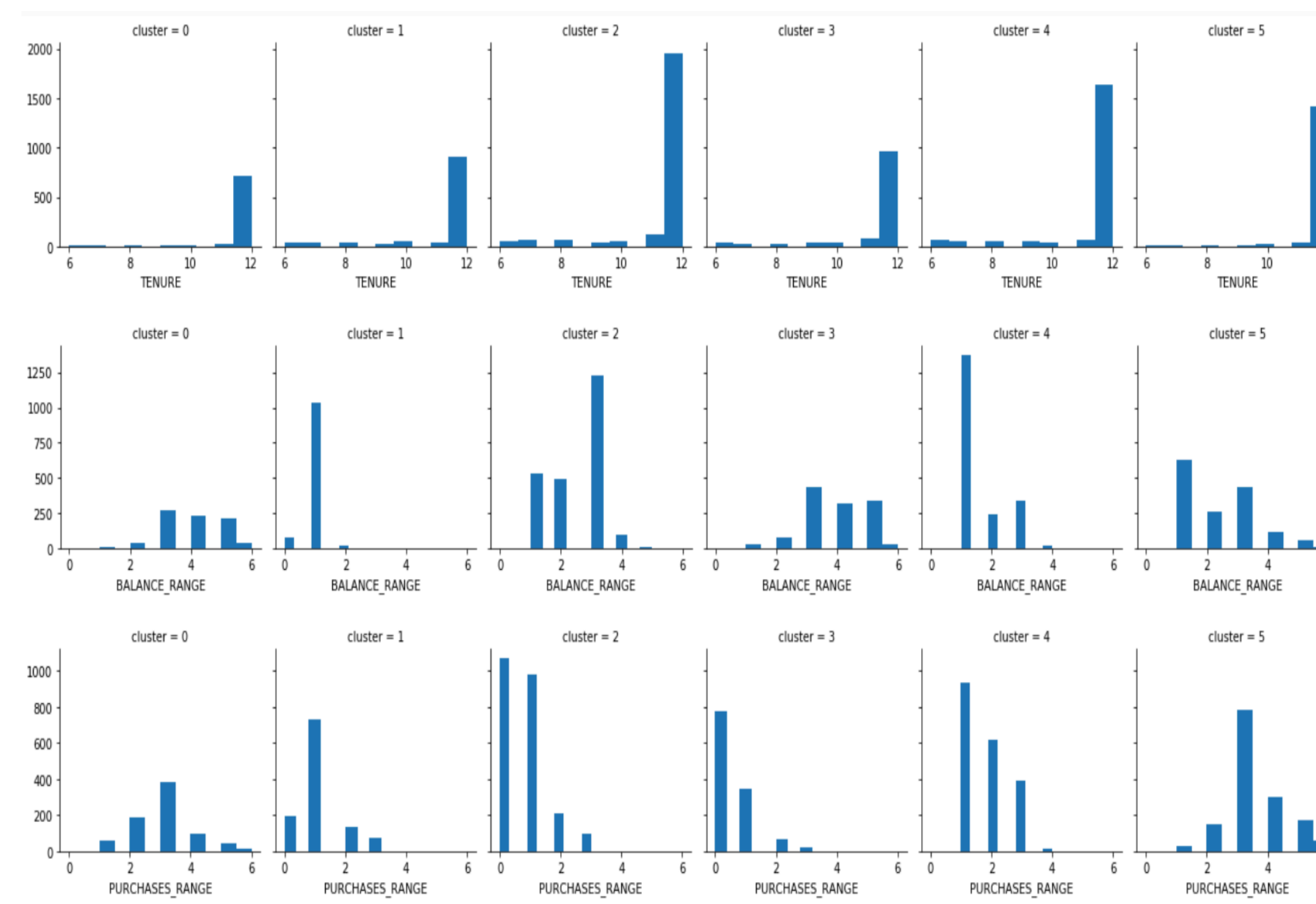
Customer Segmentation from Credit Card Usage.
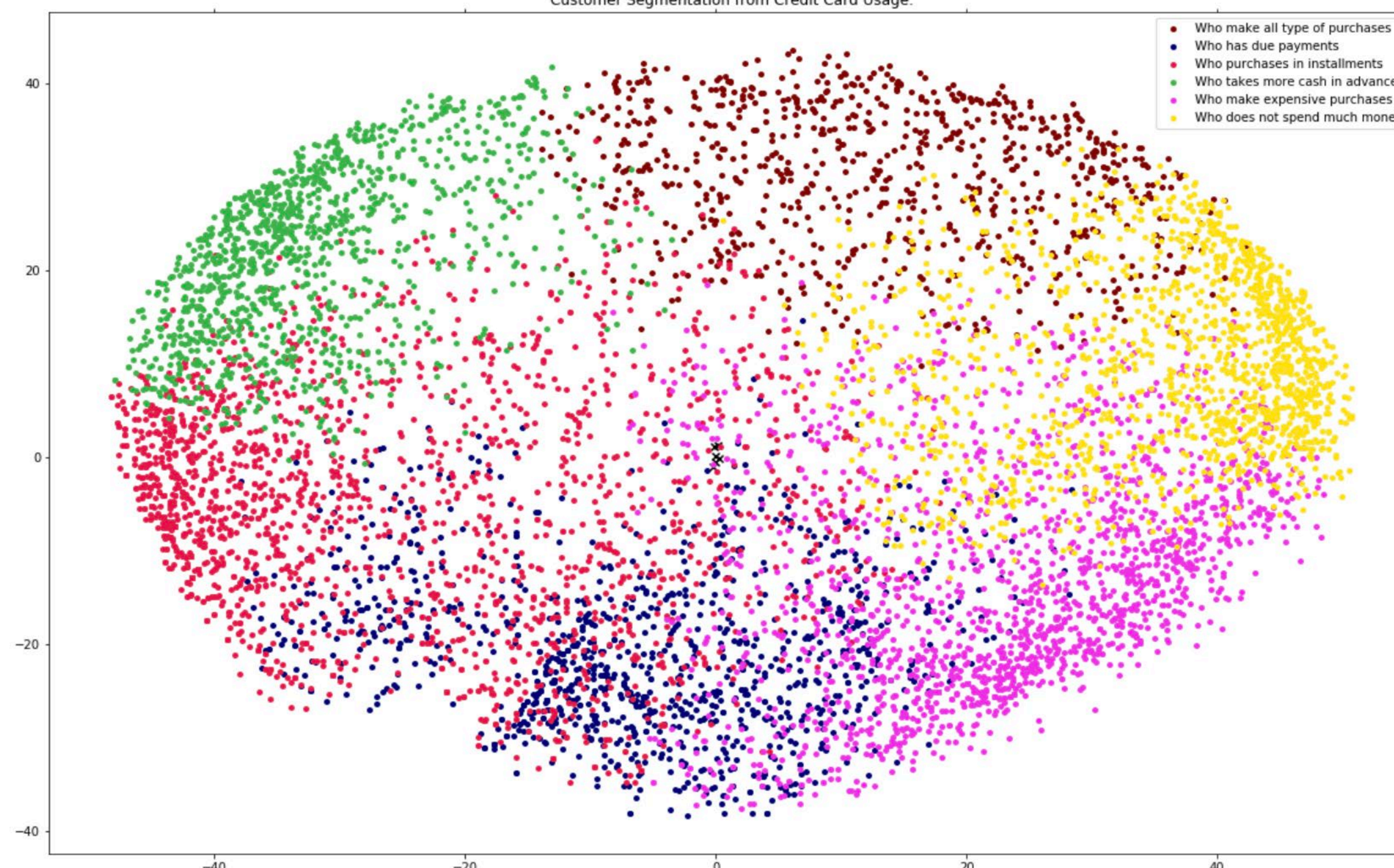
Methodology Continued

Once printing a map of histograms for every cluster and feature we can analyze what each cluster means.

Cluster One: People who will make all types of purchases

Cluster Two: People who have more due payments

Cluster Three: People who pay in installments more

Cluster Four: People who take more cash in advance

Cluster Five: People who make more expensive purchases

Cluster Five: People who rarely spend money

Finally we can finally transform the data into a X, Y plane using the principle analysis component and see what these clusters look like.

Conclusion

From the K-means clustering model we can determine a marketing strategy. It seems that more customers of the credit card make expensive purchases. A good marketing strategy would be an advertising campaign to target these customers. The company can allow customers to extend their card limit to make more expensive purchases or show an advertisement where a customer making an expensive purchases is a simple and easy task to do. We could target every type of customer and make a promotion deal or advertisement for them and their needs and send it to the customer using the emailing or mailing list. Market segmentation is a great way to understand your customers and boost your profits. Without market segmentation the company will have to assume what its customers need and wants. Every company that would like to maximize profits and understand its customer base should implement some sort of market segmentation.