Best neighborhood in Mazatlan for a Butcher shop

By Ivan Alarcon

1. Intruduction

1.1 Background

Mazatlan is a coastal city in Sinaloa Mexico in which seafood is mainly consumed, but which has begun to have a demographic and cultural growth, in which the consumption of meat in BBQs has increased, and the demand for better quality cut increases equally.

1.2 Problem

The client wants to take advantage of this demand and open a butcher shop with fine cuts and better quality of exported meat. The client wants to see which are the best neighborhoods in which he can open his store and that have a greater demand for these products since the preference for Seafood for even greater in most of the city.

2. Data acquisition

2.1 Data source

This project will rely on public data from internet and Foursquare. For this project we download the information of Postal code from Mazatlan, Mexico. Getting the geo position of the neighborhoods and getting all the information of the restaurants that are in those areas.

2.2 Data cleaning

As we get the information of the post code from Mazatlan city all that we did were to delate some columns and NA rows, after that we get all the latitude and longitude of the neighborhoods, with this last information we download all the busines that are in those neighborhoods and merge them in one file.

3. Methodology

The goal of this project is to group together the similar neighborhood in the city of Mazatlan to find the best place to open a buster shop.

3.1 Categorize

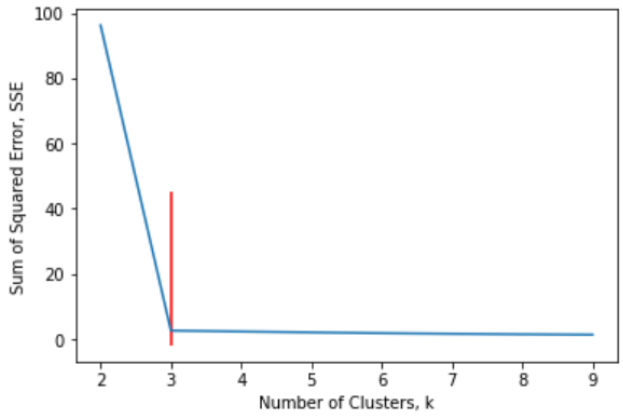
Onces when we have all the information of the restaurants in the neighborhoods we proceed to extrac the information the category of each restaurants and create a new padas frame this information will be useful to see in which neighborhood is more popular the steakhouse or BBQ places it will mean that most of the person from Mazatlan frequent that area for a steak.

3.2 Clustering

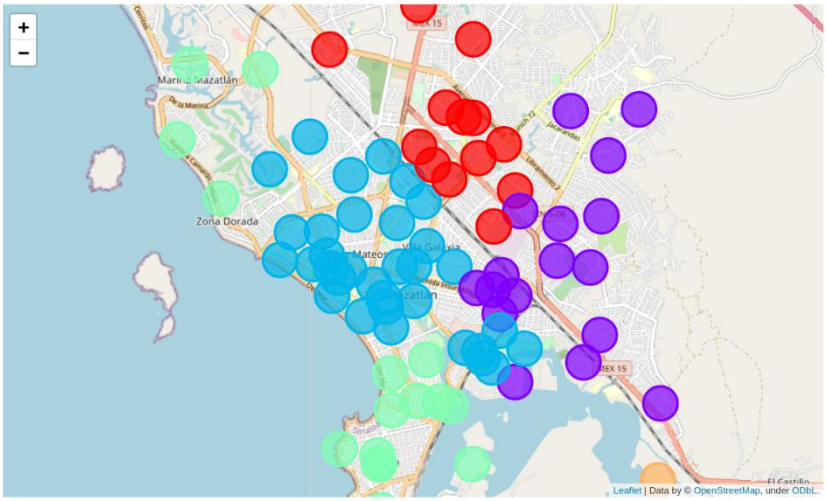
For this part we will cluster the neighborhoods for the city of Mazatlan with the information that we have collect from above. We analyze each of the neighborhoods to se how frequents are each of the categories in each of them, after that we take which ones were the most 10 common venues of each neighborhood.

3.3 Best k mean

The elbow method is used to solve the problem of selecting k. Interestingly, the elbow method is not perfect either but it gives significant insight that is perhaps not top optimal but sub-optimal to choosing the optimal number of clusters by fitting the model with a range of values for k.
The approach for this is to run the k-means clustering for a range of value k and for each value of k, the **Sum of the Squared Errors (SSE)** is calculated., calculate sum of squared errors (SSE). When this is done, a plot of k and the corresponding SSEs are then made. At the elbow (just like arm), that is where the optimal value of k is. And that will be the number of clusters to be used. The whole idea is to have minimum SSE.



With this information we start to clustering by neighborhoods using the k-means to cluster the neighborhood into 5 clusters.



3.4 Analyzing each cluster

In this part we analyze each of the cluster to see which is the best, to open the butcher shop.

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4	Mazatlan	0	Seafood Restaurant	Taco Place	Burger Joint	Mexican Restaurant	Convenience Store	Beach	Pizza Place	Juice Bar	Sushi Restaurant	Restaurant
6	Mazatlan	0	Seafood Restaurant	Taco Place	Mexican Restaurant	Burger Joint	Beach	Sushi Restaurant	Pizza Place	Hotel	Flea Market	Juice Bar
7	Mazatlan	0	Seafood Restaurant	Taco Place	Mexican Restaurant	Burger Joint	Beach	Sushi Restaurant	Pizza Place	Hotel	Juice Bar	Restaurant
8	Mazatlan	0	Seafood Restaurant	Taco Place	Mexican Restaurant	Burger Joint	Sushi Restaurant	Pizza Place	Restaurant	Coffee Shop	Beach	Convenience Store
9	Mazatlan	0	Seafood Restaurant	Taco Place	Mexican Restaurant	Sushi Restaurant	Restaurant	Burger Joint	Fast Food Restaurant	Pizza Place	Beach	Steakhouse
10	Mazatlan	0	Seafood Restaurant	Taco Place	Mexican Restaurant	Convenience Store	Pizza Place	Burger Joint	Hotel	Flea Market	Juice Bar	Sushi Restaurant

4. Results

The following are the highlights of the 5 clusters above:

1. We can see that Seafood and Taco place are the most populars plases in most of the neighborhood
2. We can see that neighborhood 1 and 2 are the ones that betwent their most commons have a Steakhouse or a BBQs which means that in these areas are more popular this meals and could be the main options for the Butcher shop

5. Conclusions

Cluster 1 and 2 were selected because in these the Steakhouses and BBQs are more popular and it is the kind of clients and areas we are looking for where they prefer these foods, it is also important to complement this analysis with data from another platform such as Uber eats to verify that these products are purchased more in these neighborhoods.