Centers of coffee

Report

Business Problem

An imaginary global coffee company is interested in entering a new city (Rome, Italy), they have decided to open coffee shops to compete with all the local vendors. They want to open the new shops as close as possible to the competition in order to have the cheapest prices in the area and run the competition out of business. They have come to me to ask for a geographical analysis, to find the perfect place for their shops. The coffee shop industry is worth over 100 billion dollars and is growing fast.

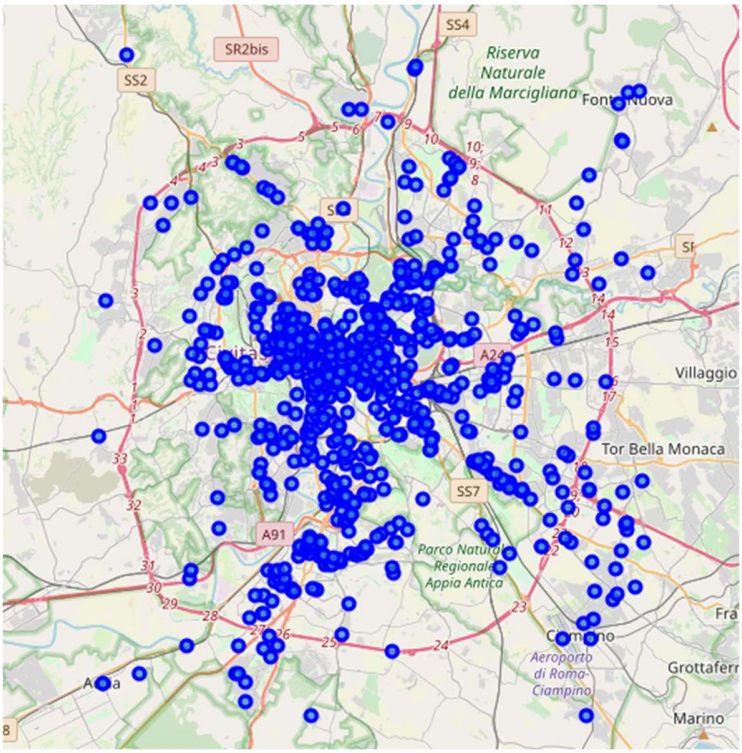
In this project I will use k-means to find the locations in Rome that minimize the distance to all coffee shops of Rome, Italy.

Data

The data will be obtained by the foursquare API, I'll be looking for geolocation data of all coffee shops and café's in Rome, and I will be applying k-means algorithms to this data in order to find the locations that minimize the distance to all coffee shops.

The data will be obtained by the foursquare API, I'll be looking for geolocation data of all coffee shops and café's in Rome, and I will be applying k-means algorithms to this data in order to find the locations that minimize the distance to all coffee shops and café's.

The gathered data looks like this:



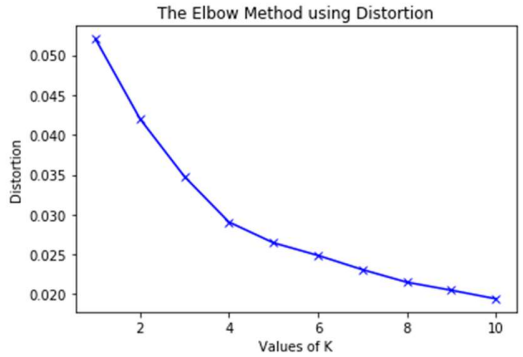
We can see that there are a lot of coffee shops and café's in Rome. 801 to be precise.

Methodology

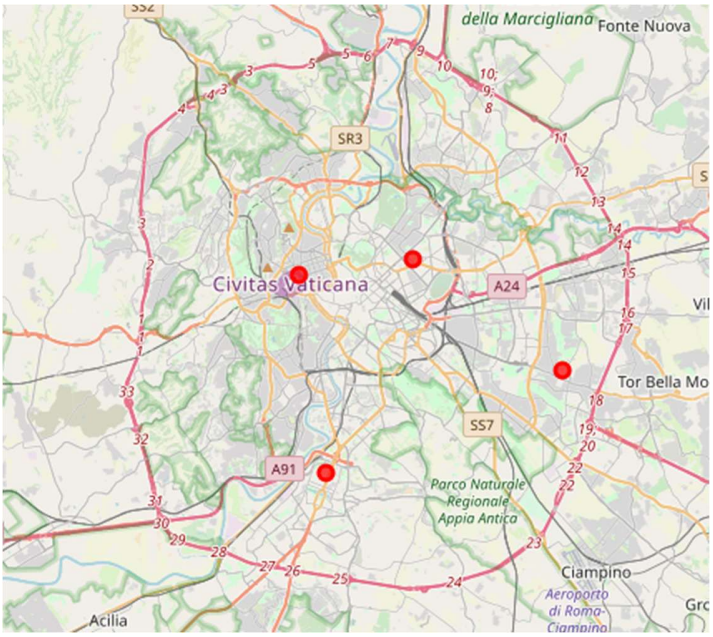
I will use K-means to determine the best location for the coffee shops, first I will use the elbow method too determine how many coffee shops should the company open. Then I will find the clusters and the centroids which will determine the location of the new coffee shops.

Results

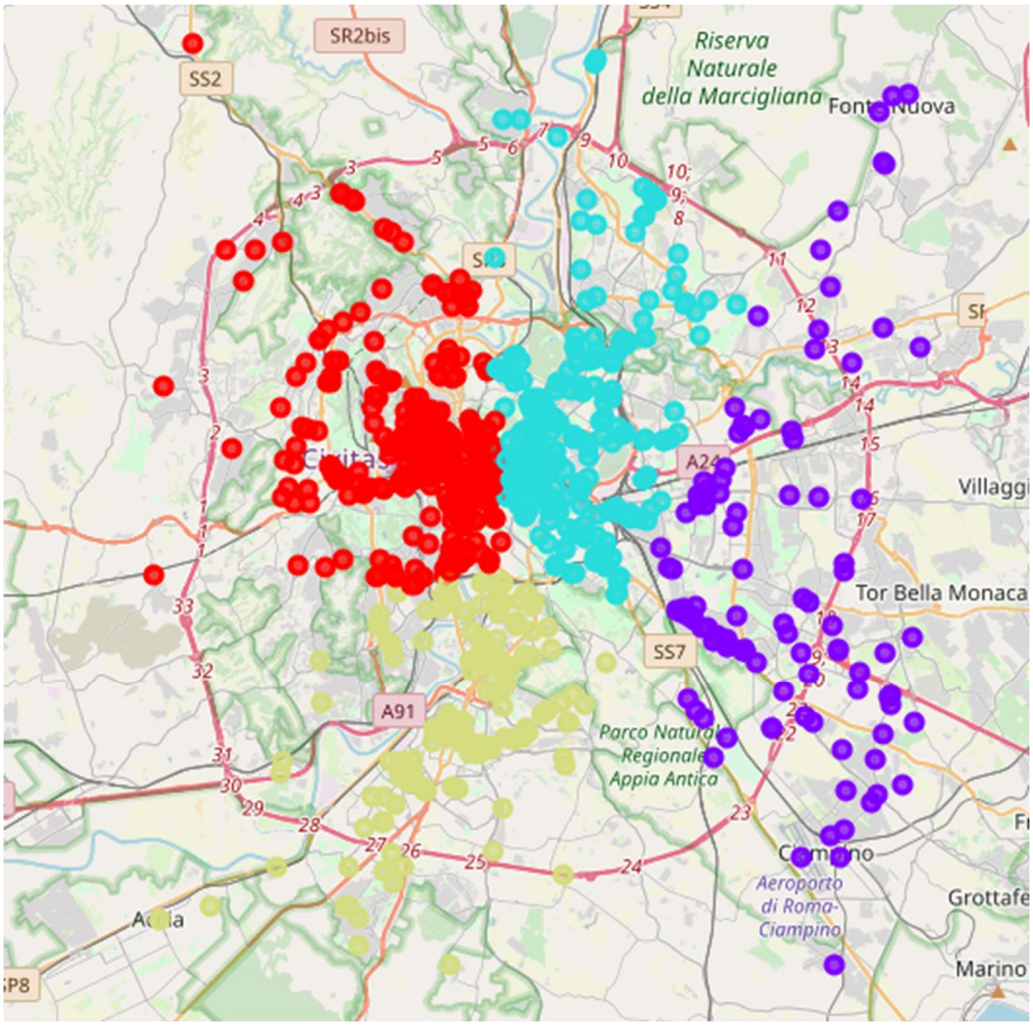
By the elbow method I've determined that 4 coffee shops will be the best number for the company too minimize distance between the new coffee shops to all existing coffee shops and to not waist cost by opening too many coffee shops:



Now I used k-means too find the best locations for the new coffee shops, here they are on a map:



And here are the clusters:



Discussion

This project can be expanded too any location-based industry and any city, I think it is a very useful tool to have. Rome has a lot of coffee shops, so it makes for a good example. It can also be applied with weights for multiple types of shops, for example close to coffee shops but far away from ice cream shops.

Conclusion

I used the foursquare API and I have found the four centers of coffee in Rome.