POLICE STATIONS VS DONUT SHOPS

Introduction

This project aims to show the relationship between the locations of police stations and food joints in order to determine the popularity of donut shops near the stations. Also, the stations will be clustered to determine a similarity between them.

Three cities take part in this project: Chicago, New York City, and Los Angeles.

Even though the reason for this subject was pure curiosity, this project can be used as the first step in marketing research, for example, to find out what is the most popular brand of a donut shop near police stations.

Data

Chicago: https://www.chicago.gov/city/en/depts/cpd/dataset/police_stations.html

New York City: https://www1.nyc.gov/site/nypd/bureaus/patrol/precincts-landing.page

Los Angeles: http://www.lapdonline.org/our_communities/content_basic_view/6279

FourSquare API (https://foursquare.com/city-guide) was used to get information about food joints for each police station.

OpenStreetMap Nominatim API (https://nominatim.org/release-docs/develop/api/Overview/) was used to get coordinates for the cities and each police station.

Data Analysis

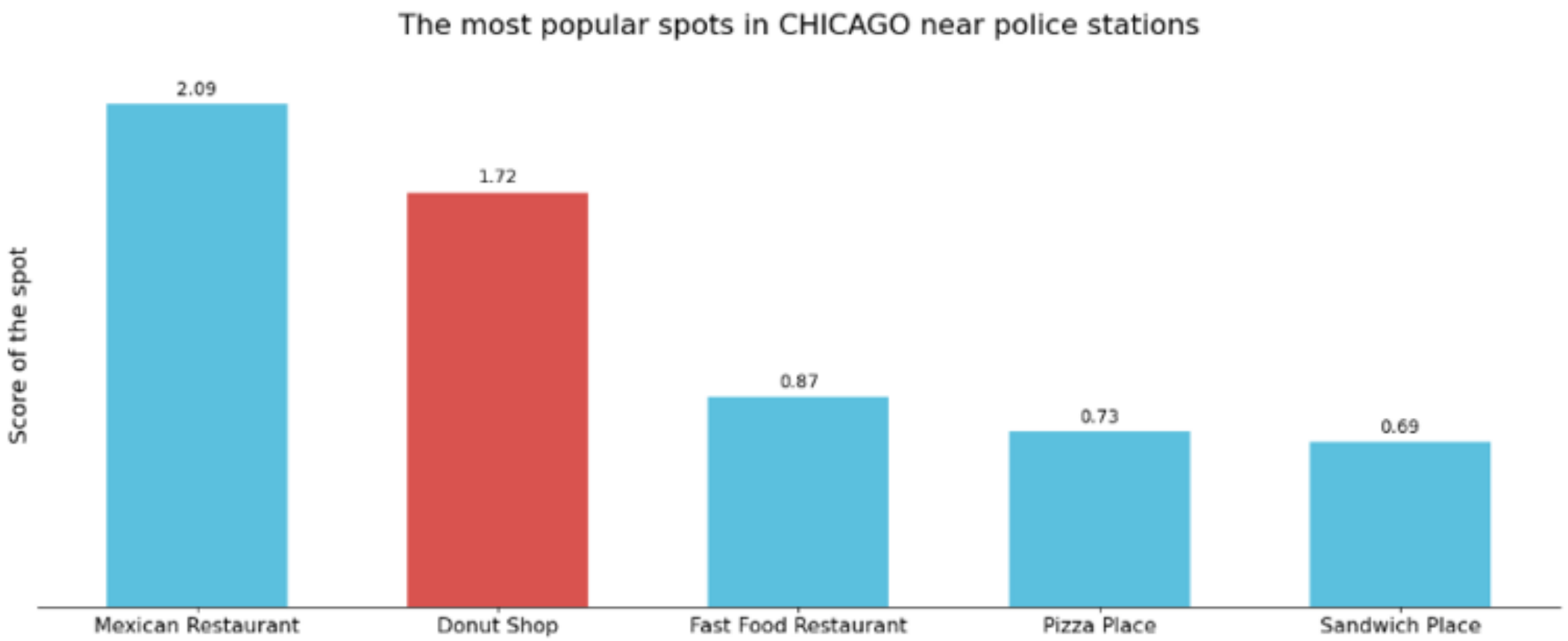
In the second step, I calculated a score for each category of food joints for each police station.

$$Score = \frac{Number\ of\ food\ joints\ from\ the\ same\ category}{Median\ of\ the\ distance\ for\ the\ category\ to\ police\ station}$$

Using the sum of this score for each category, I sorted the data to find the most popular categories near police stations for the city. This answered the main question of this project.

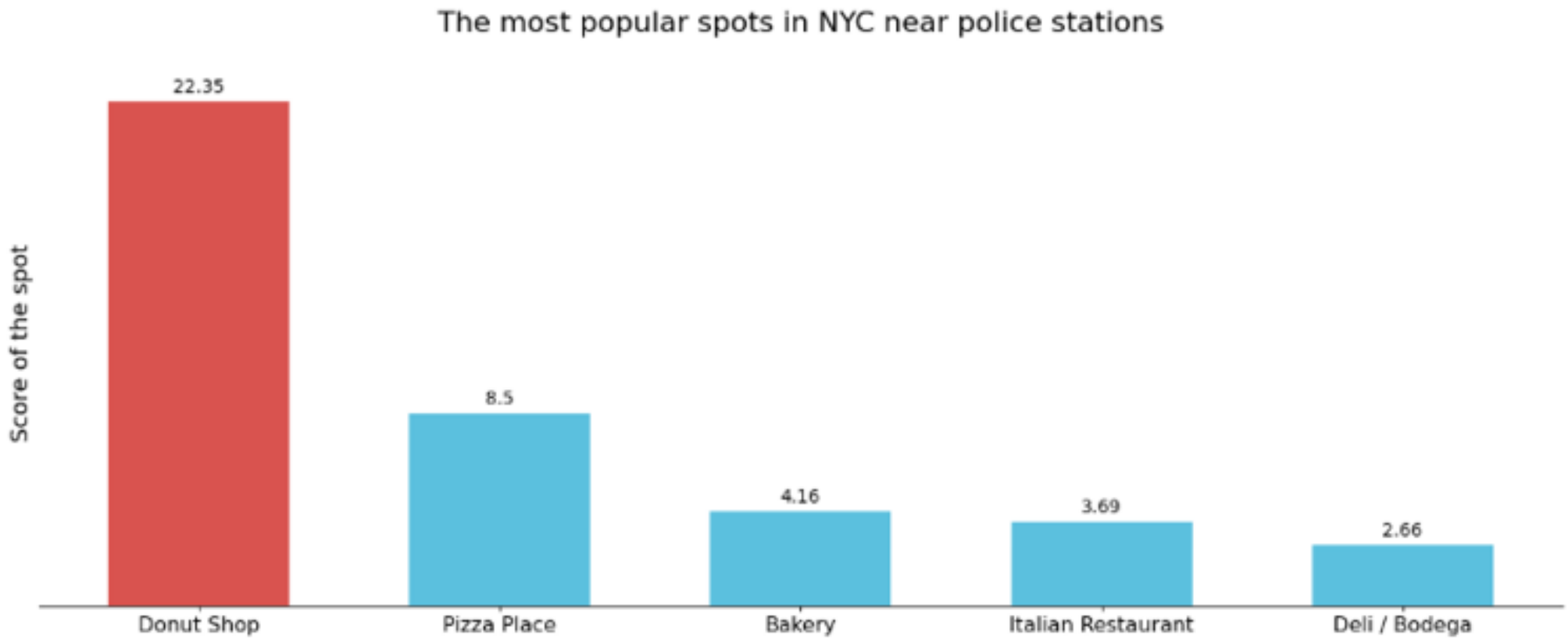
The most popular food joints near police stations for the city

Chicago



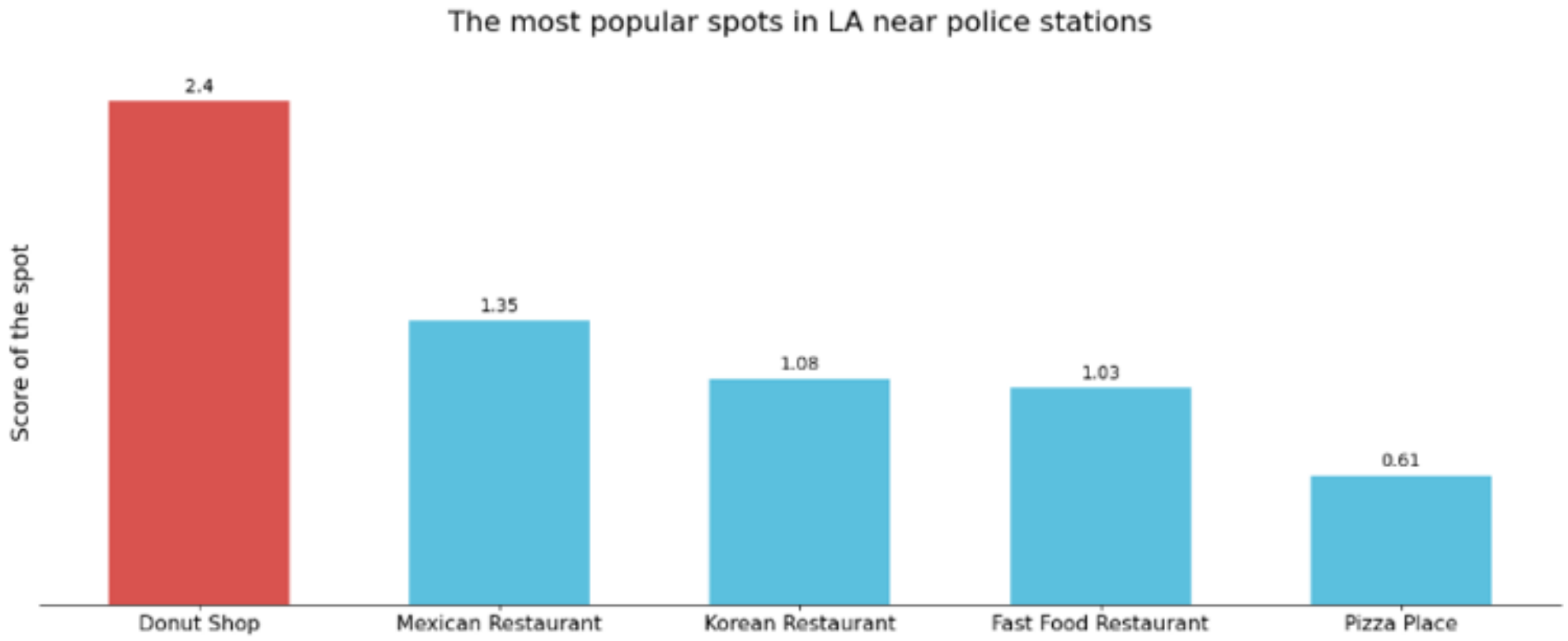
The most popular food joints near police stations for the city

New York City



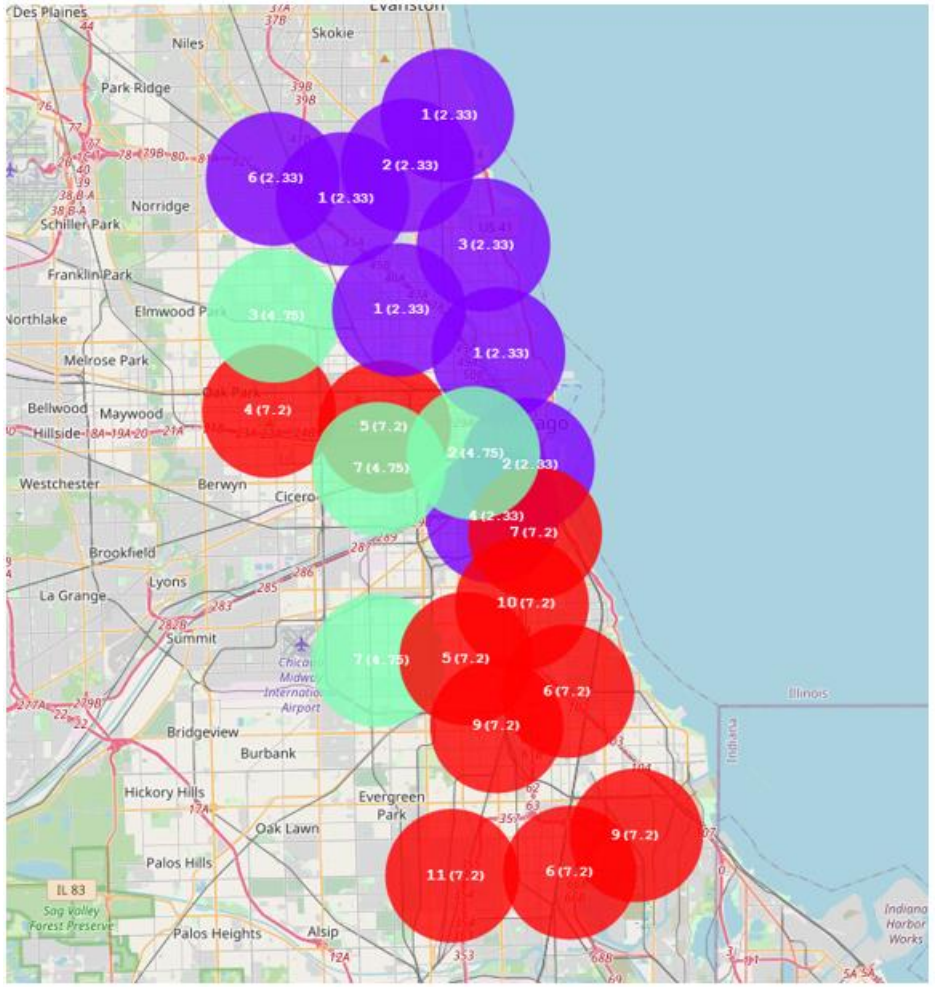
The most popular food joints near police stations for the city

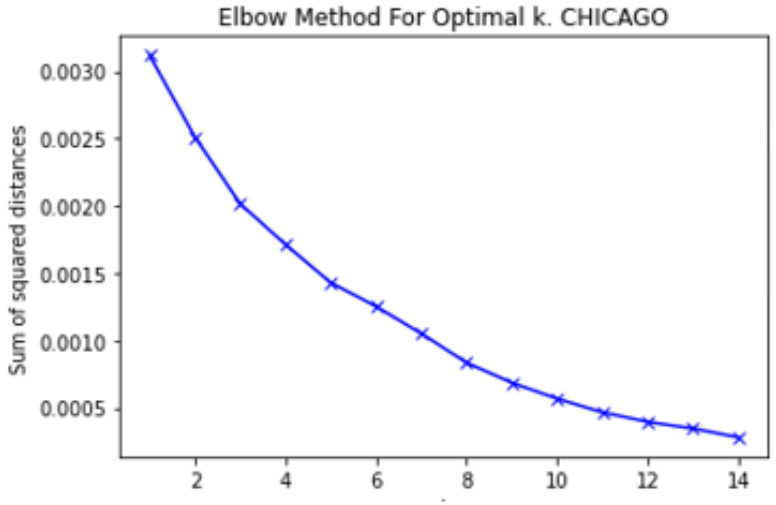
Los Angeles



Clustering

Chicago

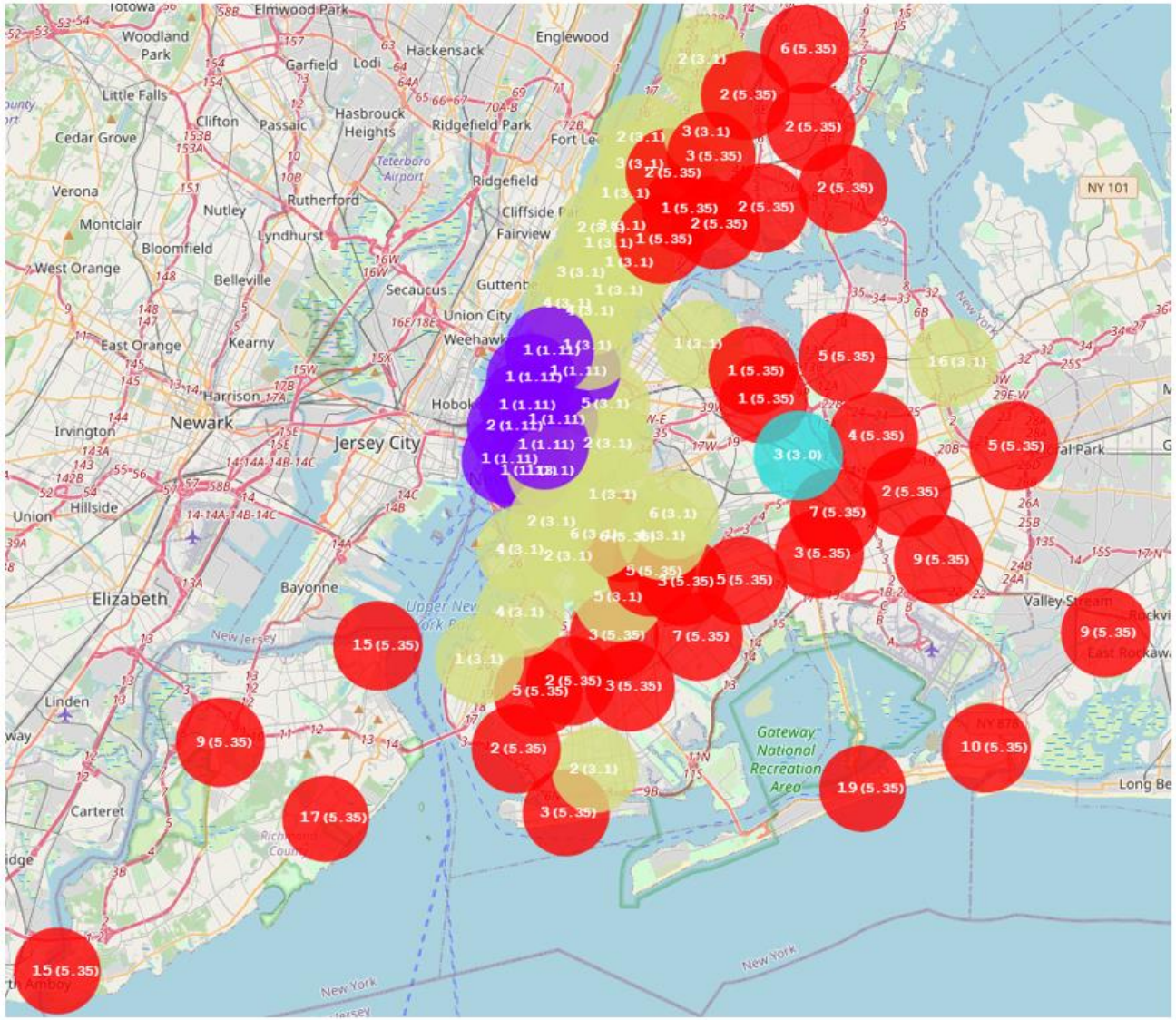


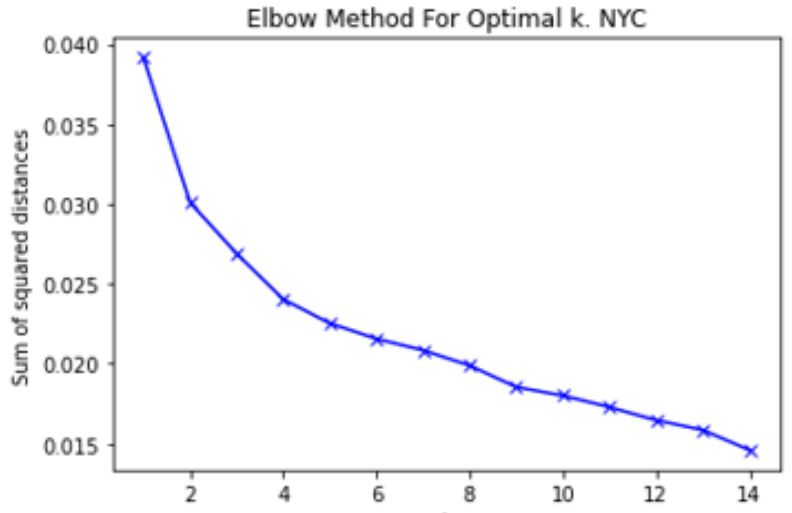


Best K is 3

Clustering

New York City

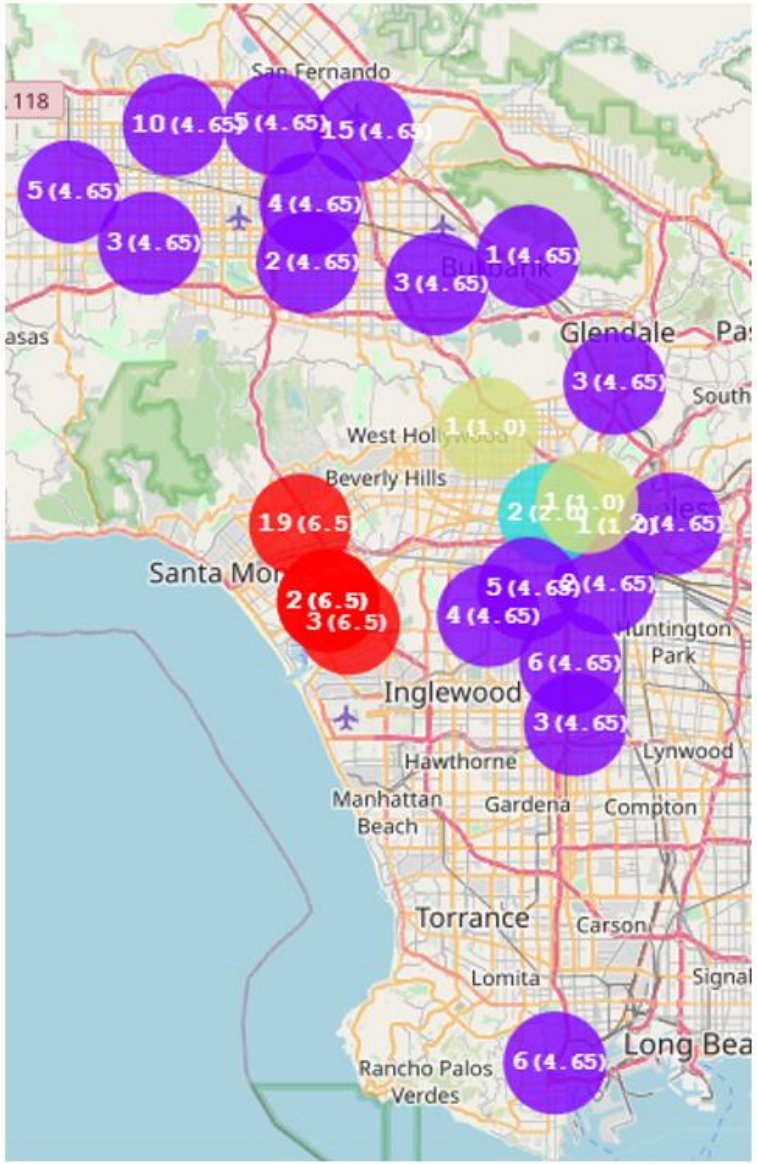


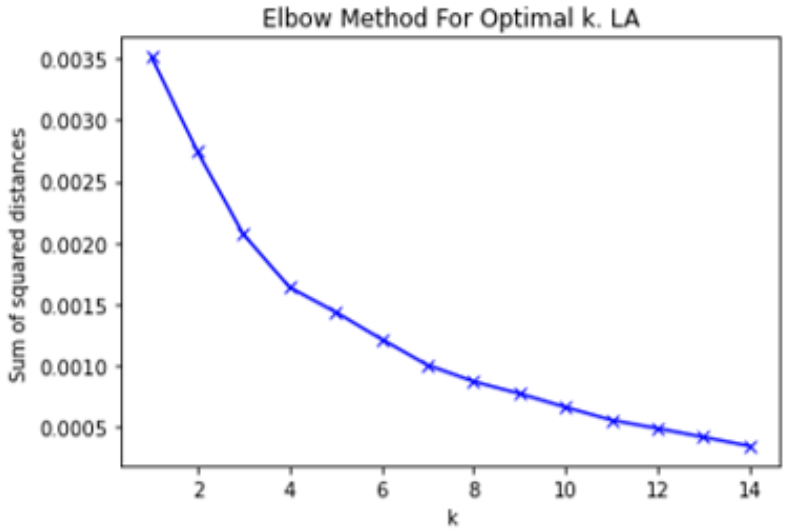


Best K is 4

Clustering

Los Angeles





Best K is 4

Results

The analysis shows that donuts shops are popular near police stations.

The clustering, however, shows that more extensive analysis should be made, especially in New York City and Los Angeles since clusters, in general, are not very different and have outliers

The relationship was defined, but analysis showed that there are steps in this project that could be improved:

- List of the food joints
    - It makes sense to increase the radius around each police station for more objective evaluation.
    - Get the full list of food joints within the radius regardless of popularity defined by the API.
- Make a more extensive analysis of the outliers in clustering.

This concludes the results of our project, but there are further steps that can be performed, for example:

- To find the most popular brands of donut shops near police stations.
- To find the most popular brands of donut shops within the cities and compare the result with donut shops near the stations.