The Battle of Neighbourhoods

Introduction:

More than 10 billion donuts are made each year in the United States alone. Originating with the Dutch, these deep-fried cakes enjoy a rich history and a growing fan base. Specialty doughnut shops may offer wacky creations like peanut butter and sriracha, but most people are happy to indulge in the regular old glazed variety.

Some of the cities in the US with the best donuts and have a large number of donut shops are New York, NY, Chicago, IL, Boston, MA and Seattle, WA. These cities are known for having some of the best donut shops in the US.

Problem:

The problem to be solved is to analyse and find the donut shops' locations in the major US cities and find the best place to have a donut and which city has the closest donut shops and the most number of donut shops in the city.

Data Section:

The Foursquare API will be used to collect data about locations of donut shops in 4 major US cities which are: New York, NY, Chicago, IL, Boston, MA and Seattle, WA. These are one of the most populated US cities and I am hopeful that they will contain the best donut shops in the US.

Methodology:

The main objective here is to asses which city would have the highest Donut Shop density.The Four Square API was used through the venues channel. The near query was used to get venues in the cities. The CategoryID was used to set it to show only Donut Shops. An Example of my requests:

https://api.foursquare.com/v2/venues/explore?&client_id=&client_secret=&v=20180605&New York, NY&limit=100&categoryId=4bf58dd8d48988d148941735

That 4bf58dd8d48988d148941735is the Id of the Donut Shop Category. Also, Foursquare limits us to maximum of 100 venues per query. This request was repeated for the 4 studied cities and got their top 100 venues. The name and

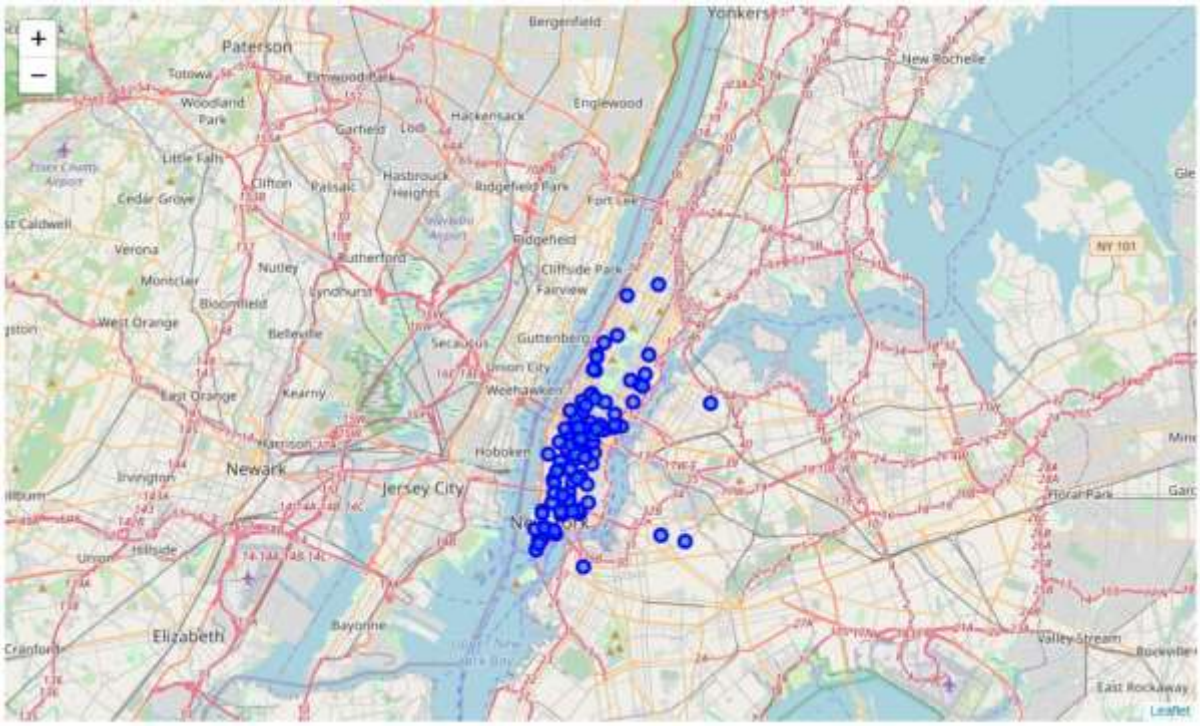
coordinate data was saved from the result and plotted on the map for visual inspection.

Next, to get an indicator of the density of Donut Shops, a center coordinate of the venues was calculated to get the mean longitude and latitude values. Then I calculated the mean of the Euclidean distance from each venue to the mean coordinates. That was the indicator; mean distance to the mean coordinate.

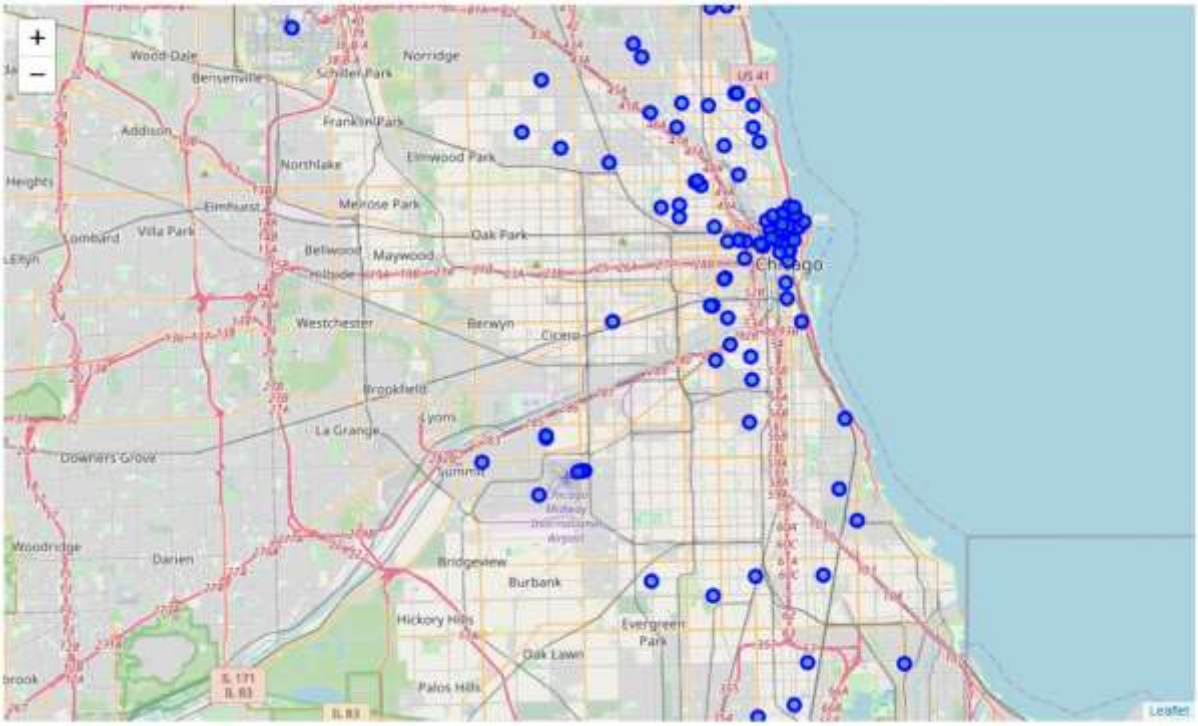
Results:

For our initial visual inspection, we see that all the 4 cities have multiple donut shops and often more than Foursquare would like to supply us. The following maps here are the pictures of the geoplot generated with the folium library:

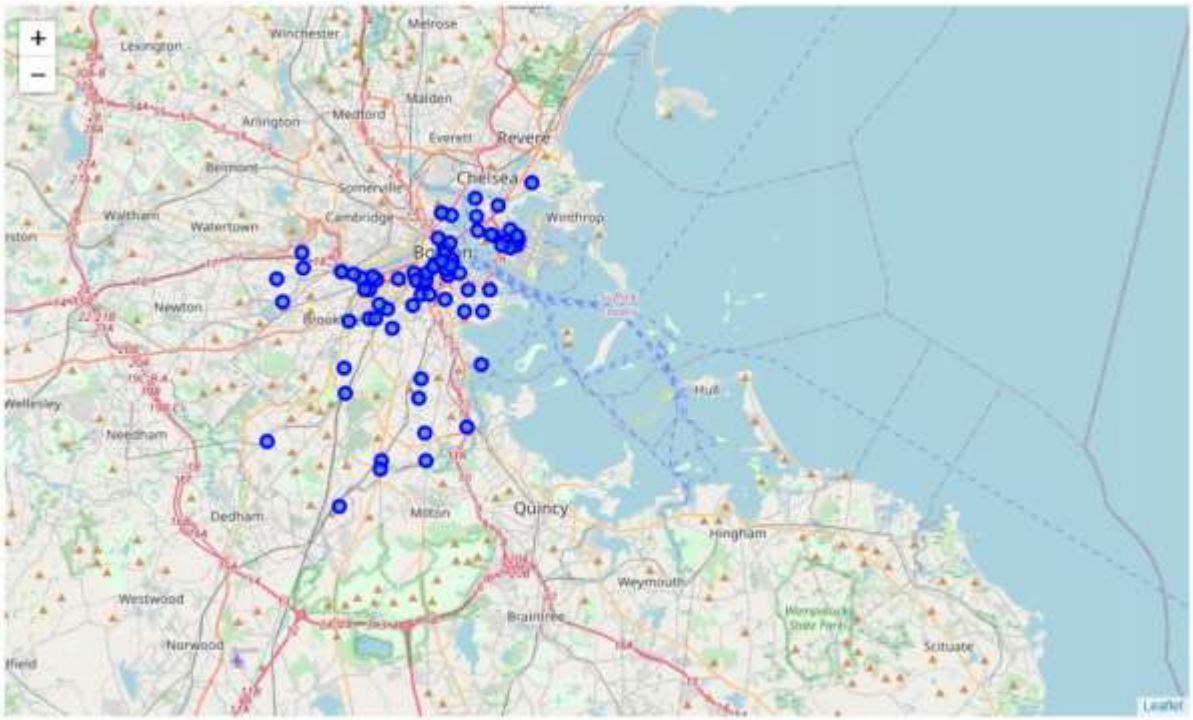
New York:



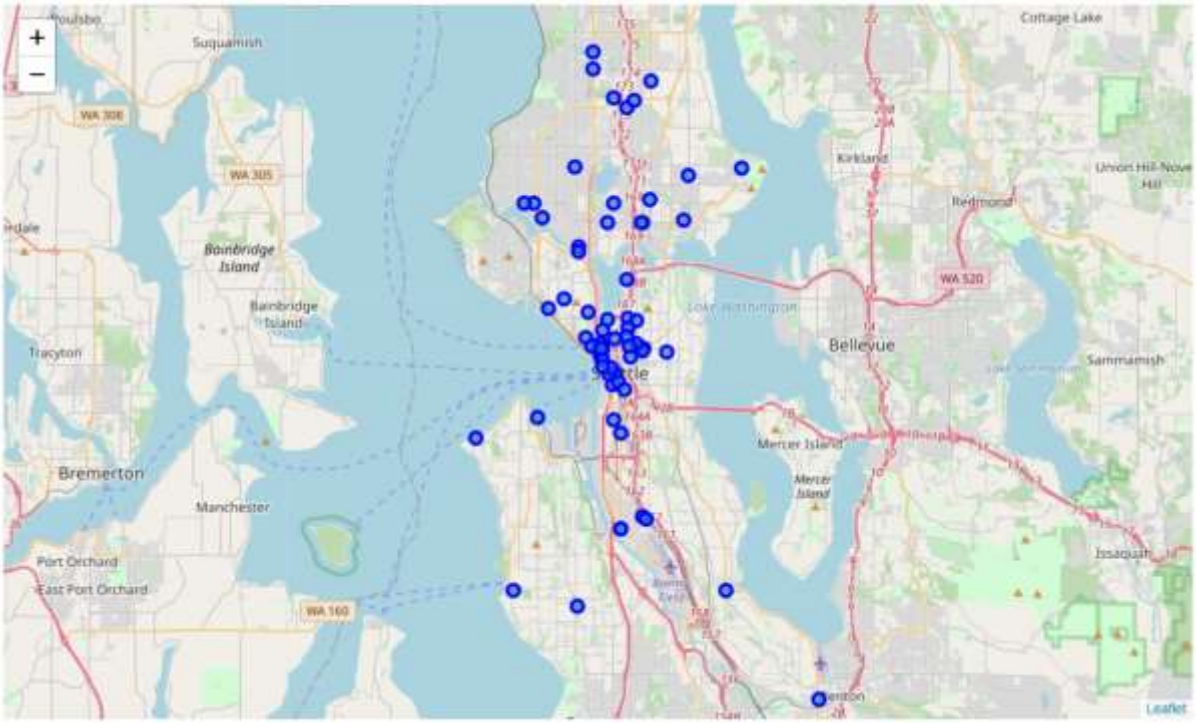
Chicago:



Boston:

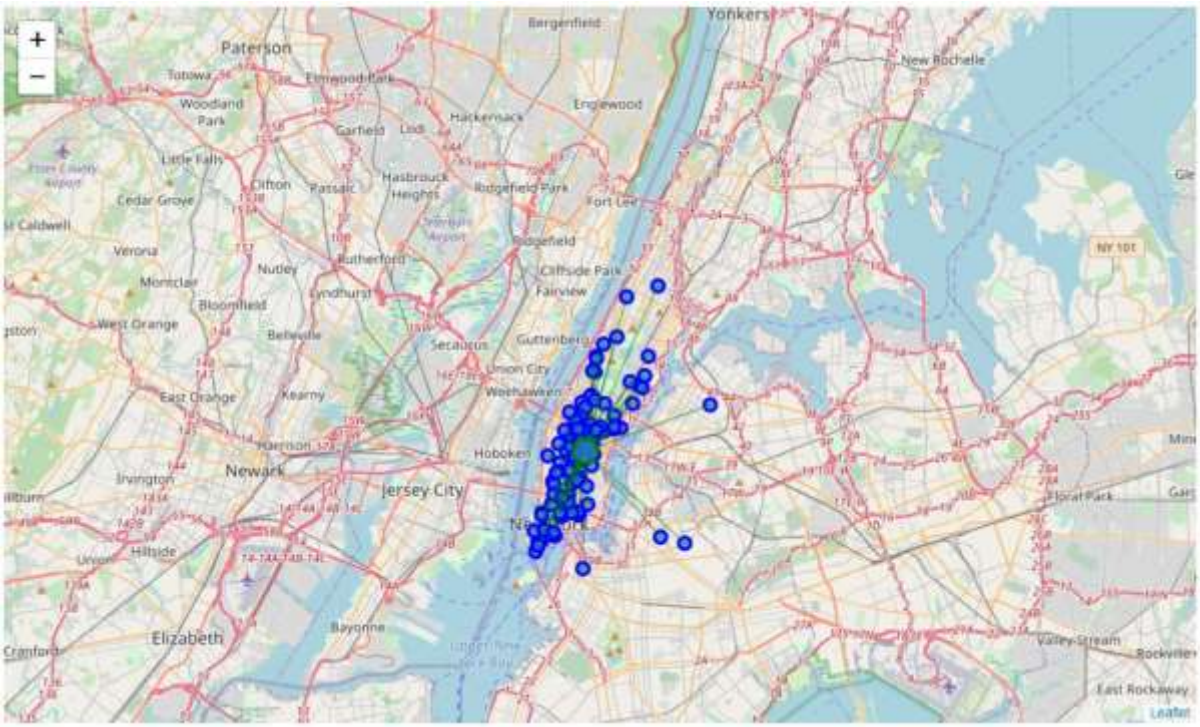


Seattle:

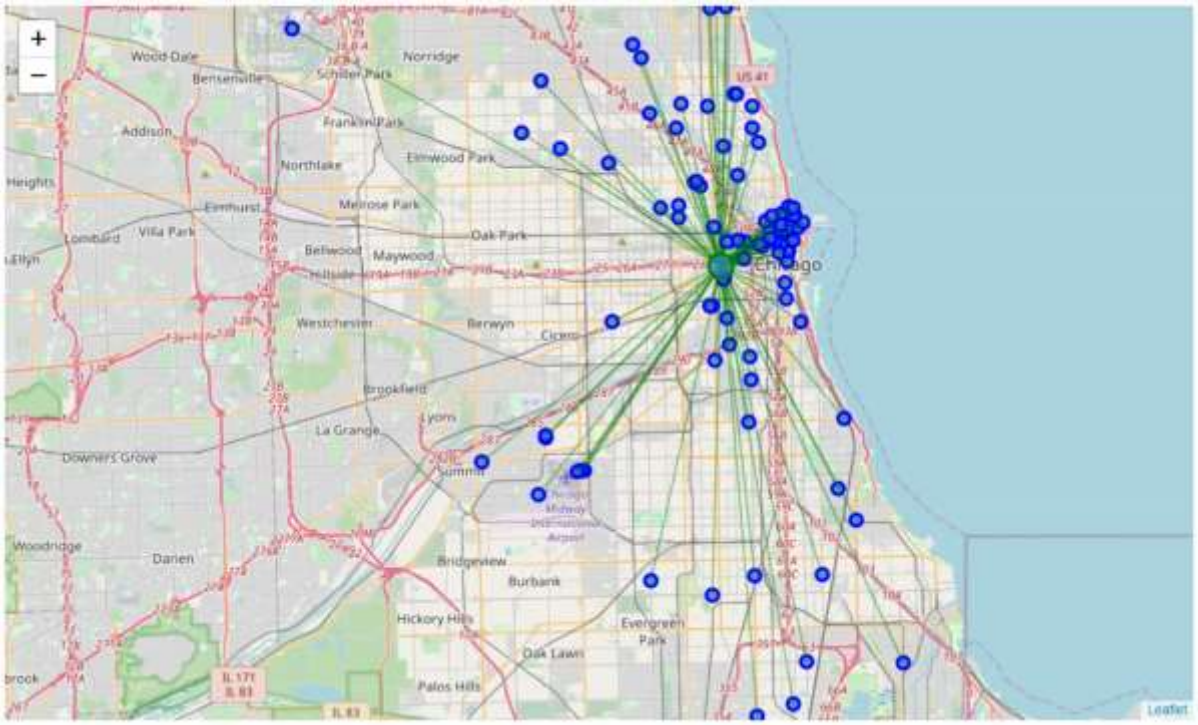


On making the first inspection we see that New York is the most dense city with donut shops. In the next step of the analysis we Calculate the Mean coordinate and the mean distance to mean coordinate(MDMC). We represent the mean coordinate with a green circle and distances with green lines.

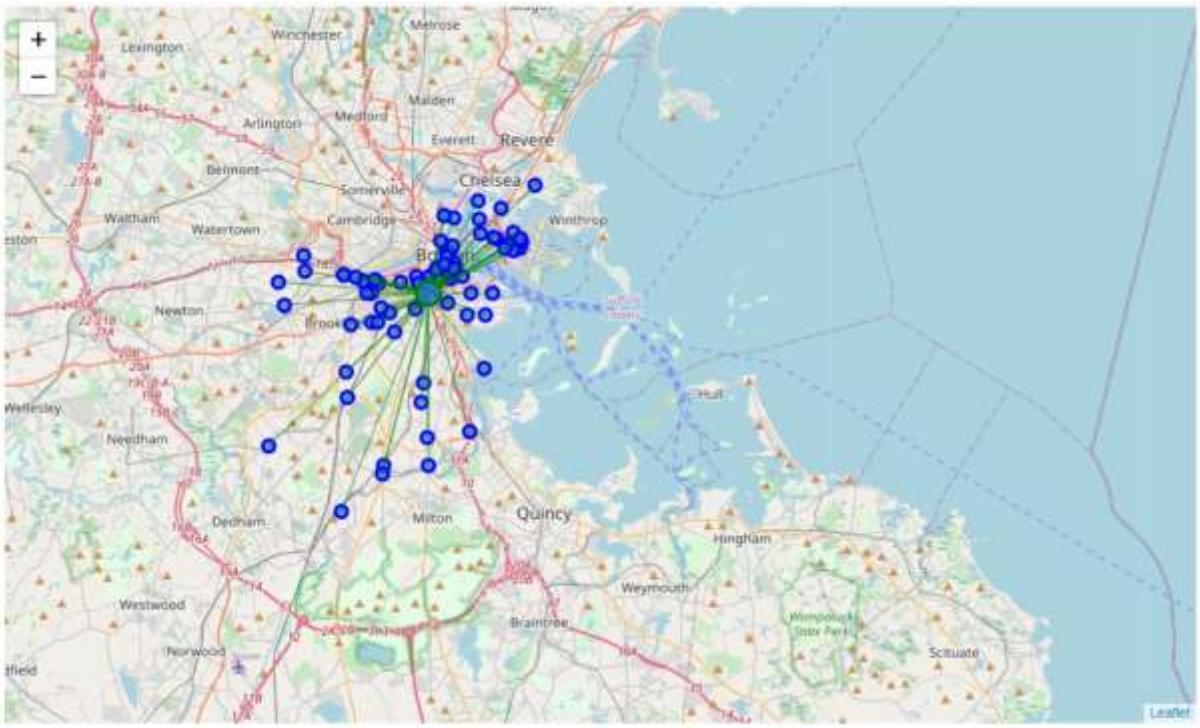
New York: MDMC: 0.0247



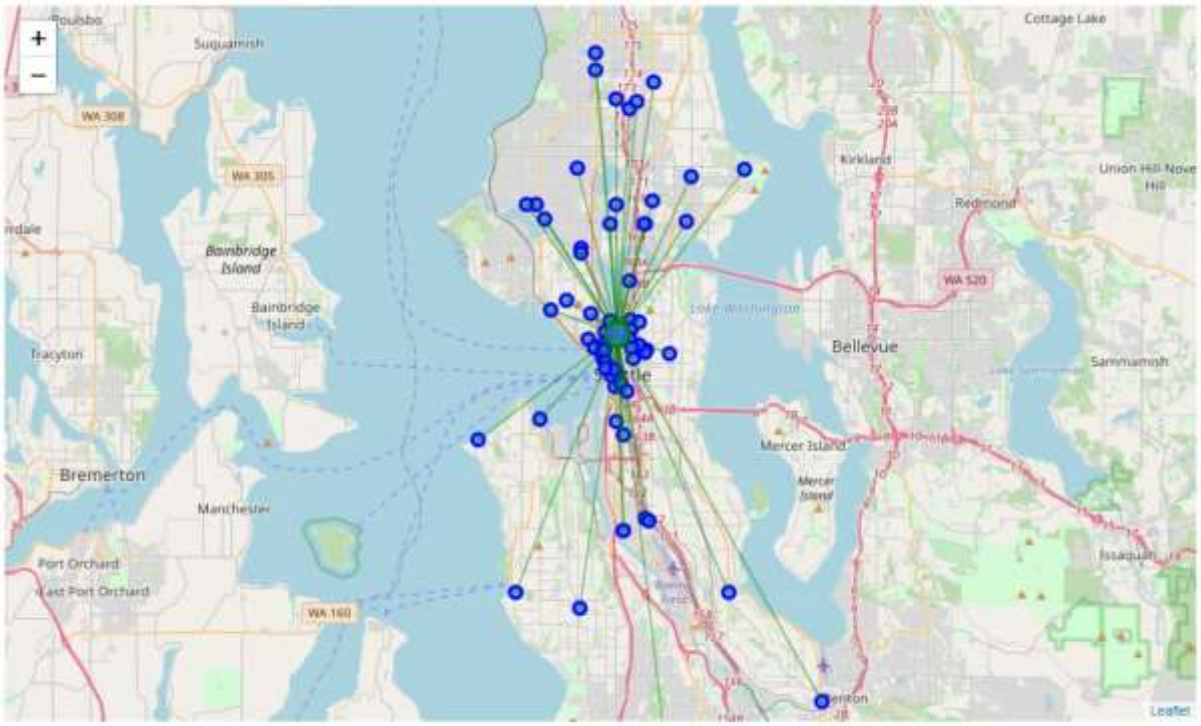
Chicago: MDMC: 0.0695



Boston: MDMC: 0.0377



Seattle: MDMC: 0.0451



The results of the cities are:

1. New York

2. Boston

3. Seattle

4. Chicago

Results:

It is clear that New York is the best option for the city to get a donut from a donut shop and also has many donut shops as options.

The above results can be improved for better results based on the data used to perform the analysis.