

Introduction

Let's say you have never been to the US and you want to have coffee while you are there. So, you want to go to a place with a high density of coffee places around you. The problem we aim to solve is to analyze the coffee shops locations in the major US cities and find the best place for our tourist so that he can have a good coffee. Our main target are tourists with a taste of coffee.

Data section

I used the FourSquare API to collect data about locations of coffee shops in 5 major US cities which are: New York, NY, San Francisco, CA, Jersey City, NJ, Boston, MA and Chicago, IL. These are one of the most populated US cities randomly picked and I am hopeful that they will contain the best coffee shops in the US.

Methodology

My main target here is to assess which city would have the highest coffee shop density. I used the FourSquare API through the venues channel. I used the near query to get venues in the cities. Also, I use the Category ID to set it to show only coffee shops.

An Example of my requests:

```
LIMIT = 100 # Maximum is 100
cities = ["New York, NY", 'Chicago, IL', 'San Francisco, CA', 'Jersey City, NJ', 'Boston, MA']
results = {}
for city in cities:
    url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&near={}&limit={}&categoryId={}'.format(
        CLIENT_ID,
```

```
CLIENT_SECRET,  
VERSION,  
city,  
LIMIT,  
"4bf58dd8d48988d1e0931735") #COFFEE SHOP CATEGORY  
results[city] = requests.get(url).json()
```

That **4bf58dd8d48988d1e0931735** is the Id of the coffee shops Category. Also, Foursquare limits us to maximum of 100 venues per query.

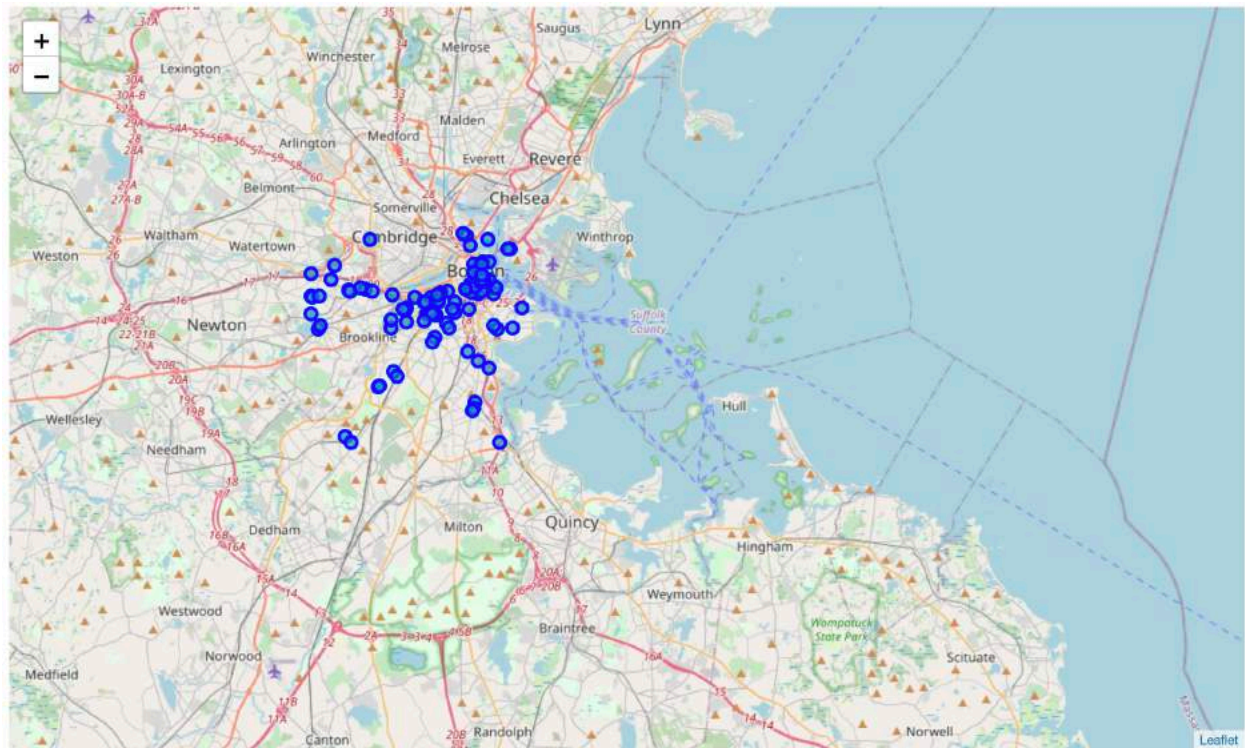
Moreover, I repeated this request for the 5 studied cities and got their top 100 venues. I saved the name and coordinate data only from the result and plotted them on the map for visual inspection.

Next, to get an indicator of the density of coffee shops, I calculated a center coordinate of the venues 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 my indicator; mean distance to the mean coordinate.

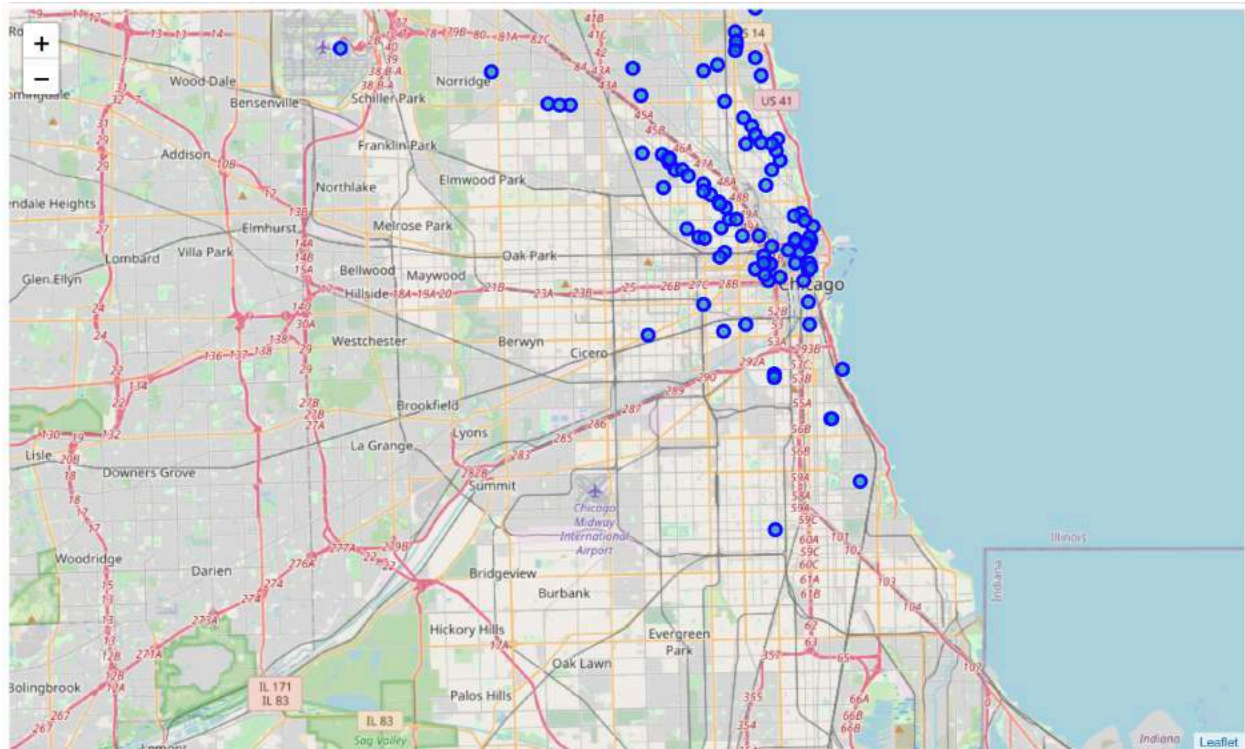
Results

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

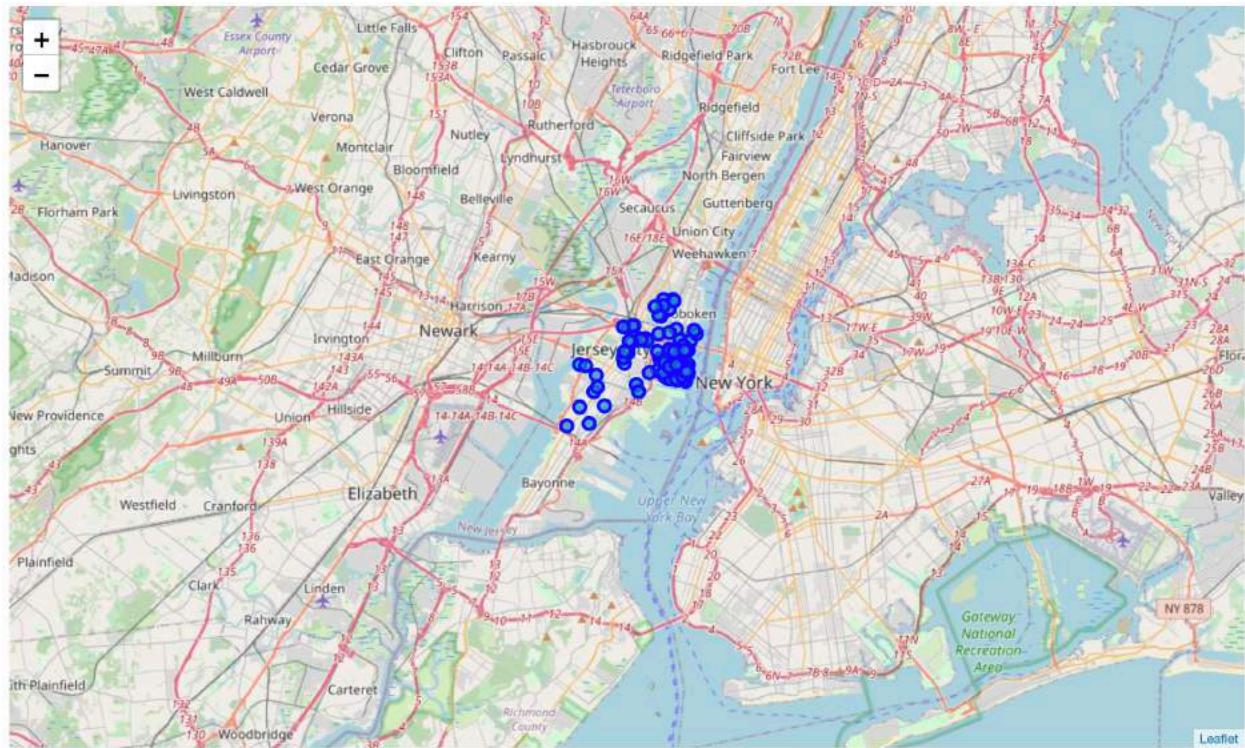
Boston:



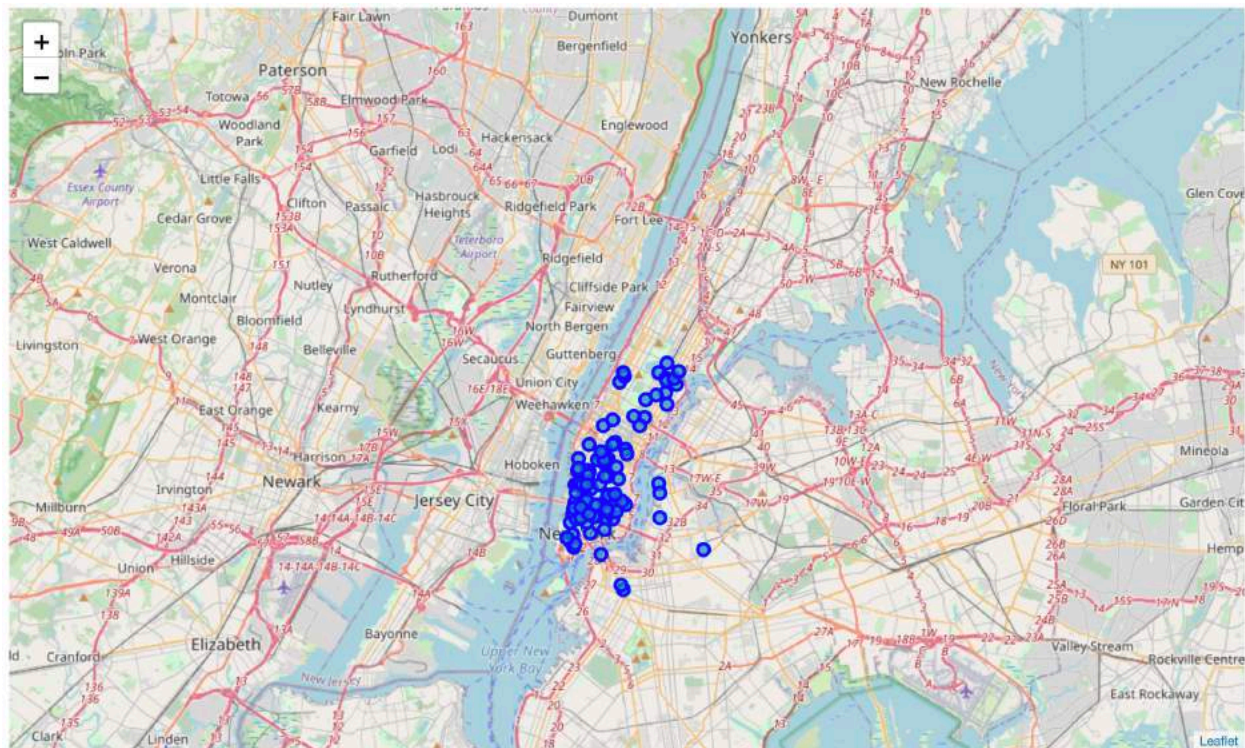
Chicago:



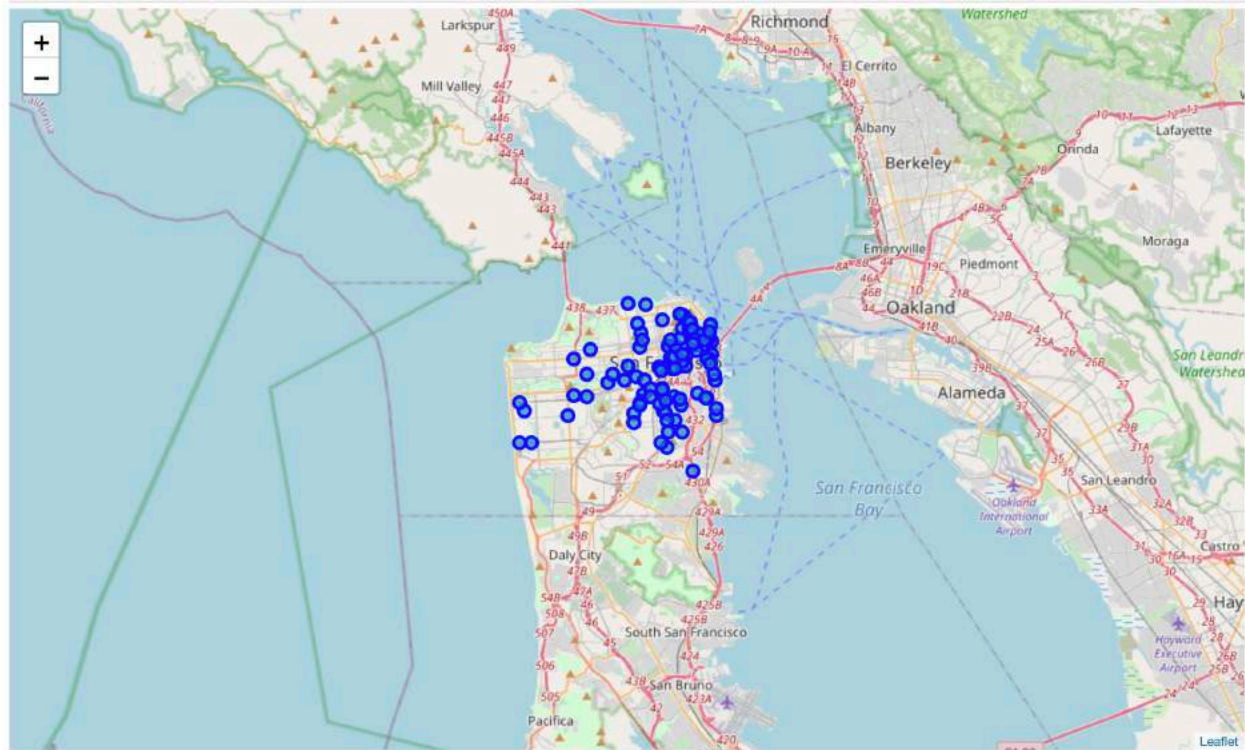
Jersey City



New York:

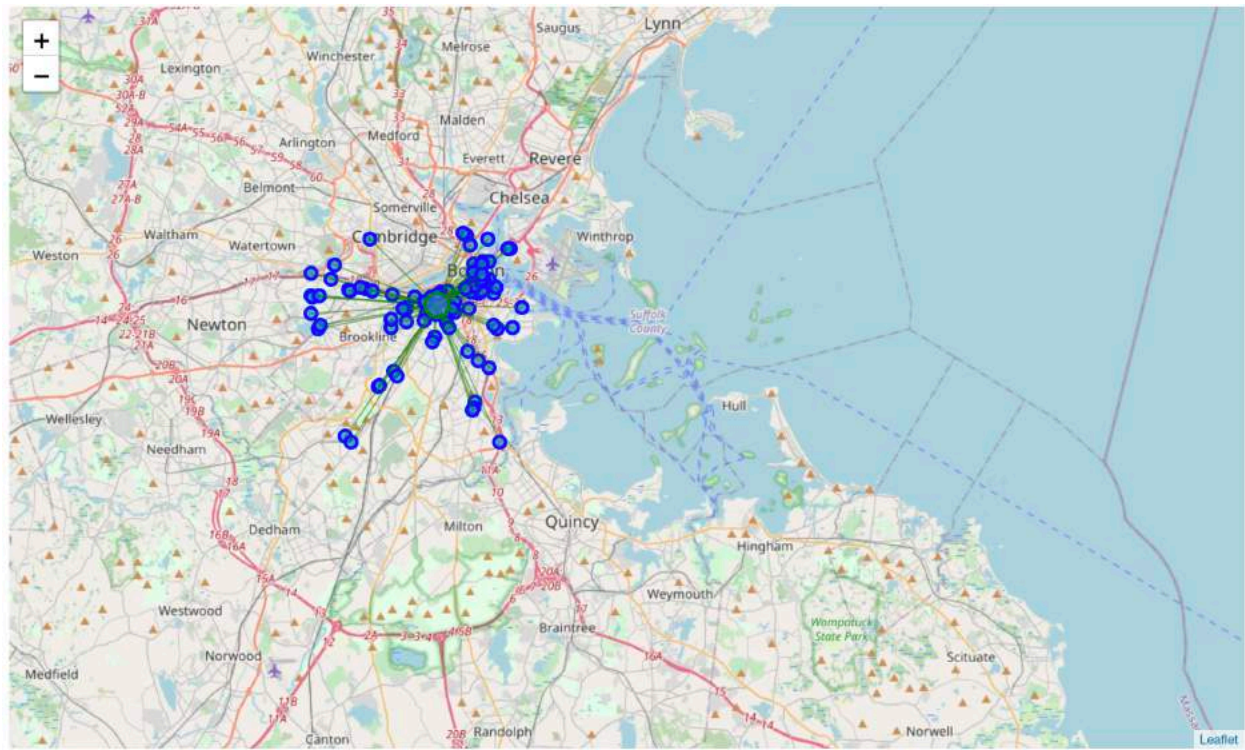


San Francisco:

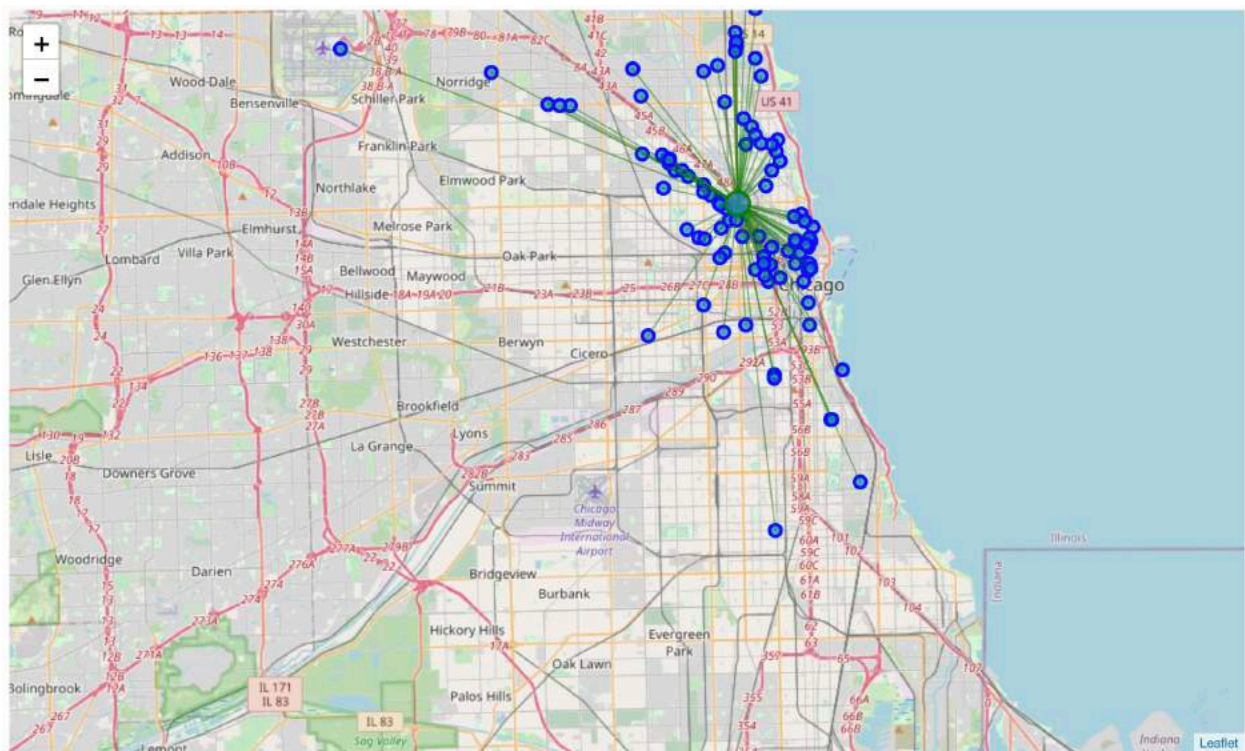


Upon First inspection we see that New York, Jersey City and San Francisco are the most densely cities. In the next phase we Calculate the Mean coordinate and the mean distance to mean coordinate (MDMC). We represent the mean coordinate with a big green circle and distances with green lines.

Boston:



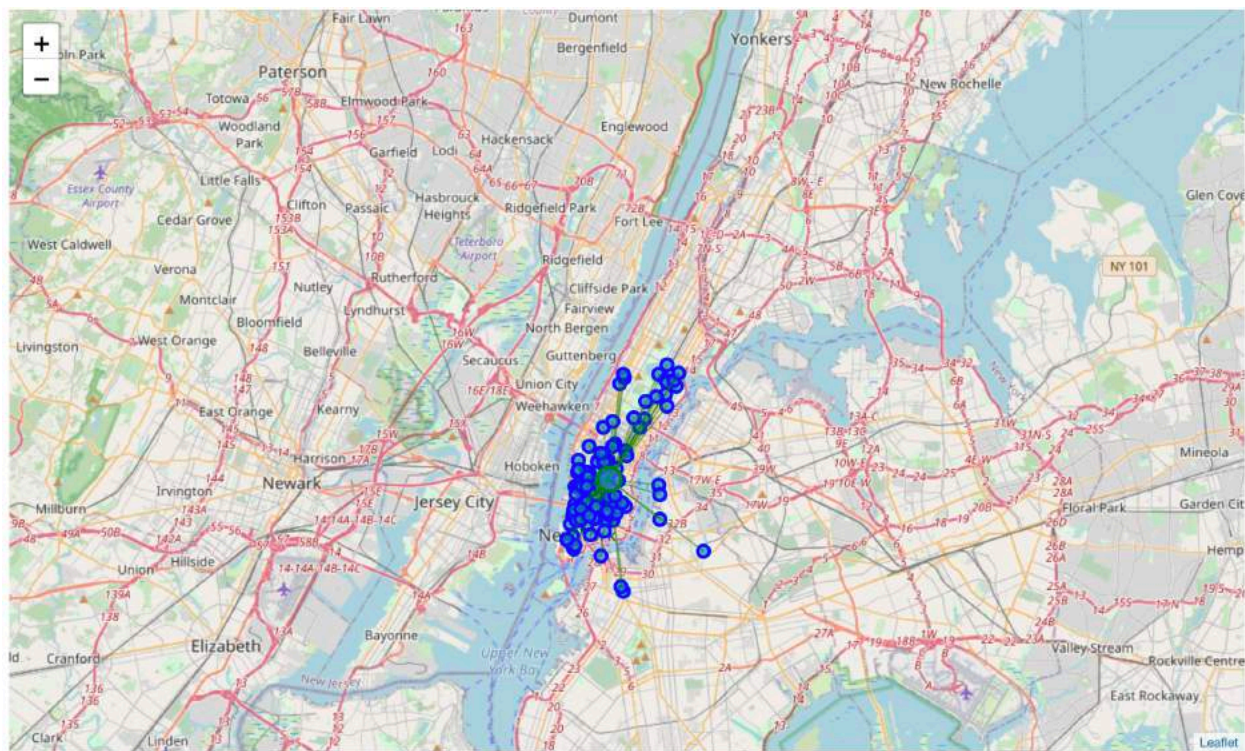
Chicago:



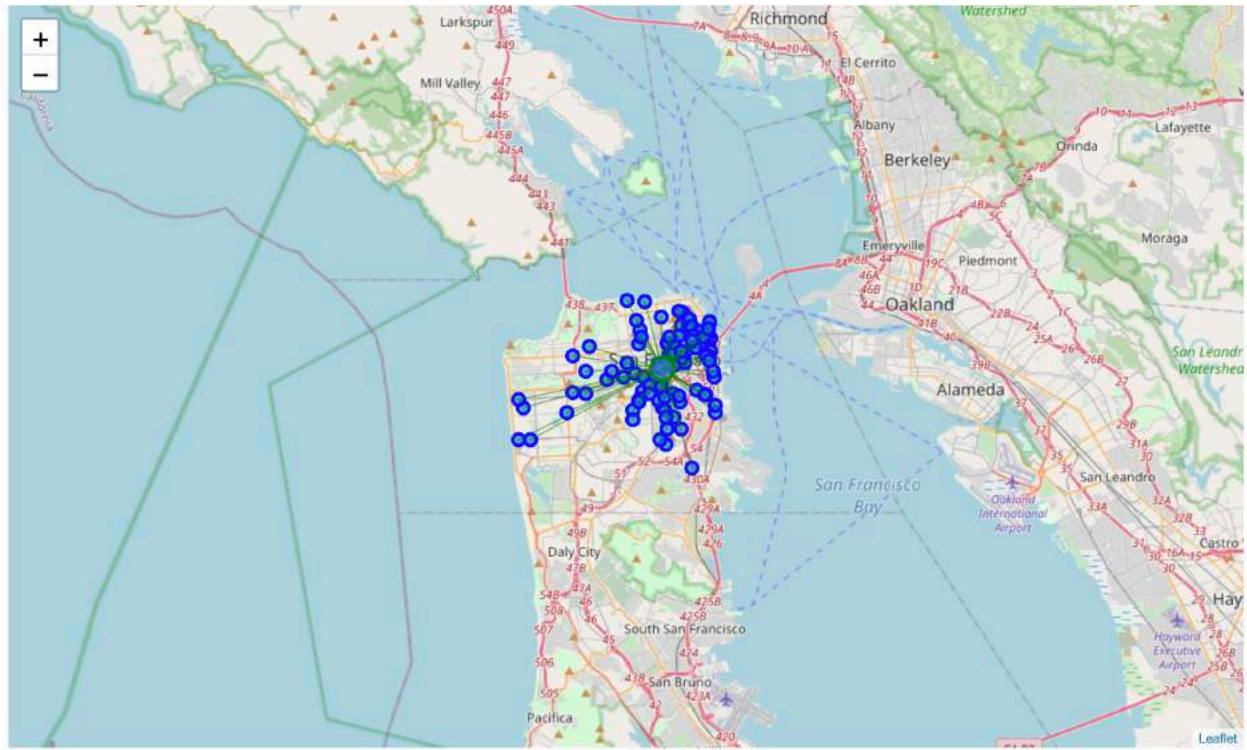
Jersey City:



New York:



San Francisco:



Based on the calculation of the Mean coordinate and the mean distance to mean coordinate (MDMC) our results are:

1. Jersey City, NJ
2. New York, NY
3. San Francisco, CA
4. Boston, MA
5. Chicago, IL

Discussion:

Based on the Mean Distance from Mean Coordinates of the coffee shops, Jersey City, New York, San Francisco, Boston, Chicago are the best places to be surrounded by coffee shops respectively. Jersey City is the best place for a person to be surrounded by coffee

shops from the mean location. and as an add-on New York is the second-best place and which is just a shore away.

Conclusion:

Now there is no doubt that Jersey City area is the best place to try coffee in many coffee shops in the US. Also, if our tourist is done with all the Jersey City coffee shops he can cross to New York and enjoy more coffees.

Also, we would recommend that our tourist to book a hotel close to the mean coordinate so that he is surrounded with many coffee shops around his stay!!
