**Background**

Before visiting, people want to know which places are hotspots in the city. It is the same for media which wants to broadcast the best places around. The united states has the best places in the world, including the best pizzerias. Anyone around the world would visit the united states in search of hotspots pizza stores. Even media giants head to the states to explore the best pizzerias there is. Some medias do not have a lot of budget, they want to visit one city to discover the pizzerias.

**Problem**

The problem at hand, is a media company from Italy which wants to visit the United States of America to showcase the best pizza spots in the country. They want to do it with little cost as possible, so they want to visit one city which might have the most, best and closest pizza places as possible. They want to choose between the five biggest cities in the country: New York, San Francisco, Jersey City, Boston and Chicago city.

They do not want to move around a lot, so they want to book one hotel in the city so that they could have access to the best places in the city. It is expensive to move across the country exploring pizzerias and transporting filming equipment is not easy.

**DESCRIPTION OF THE DATA**

The pizza stores data in the five major cities will be collected through the FourSquare API. The data will show the total number of pizzerias on FourSquare and map the distance from the hotel booked by the media company. It will narrow down the choices of pizzerias that they could visit based on the distance as they want to save travel costs. They are hopeful that this will be enough to help them get to the best pizzerias in the country with little cost as possible. Apart from recommending the closest pizzerias from their hotel, it will show a map of the locations of the pizzerias.

**Cleaning**

In the data there would be outliers, like a pizzeria in New York, which closer to Jersey City CBD than to New York, such will be dropped out of the data as they decrease the score of the city.

**exploratory data analysis**

I used the folium and Geopy to do the visualization of the cities that I was exploring.

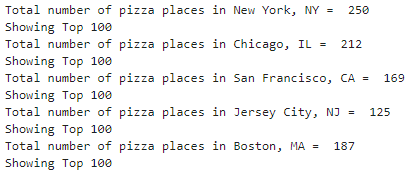
**inferential statistical**

There was not any inferential apart from the mean distance calculated from the hotel to the pizzeria

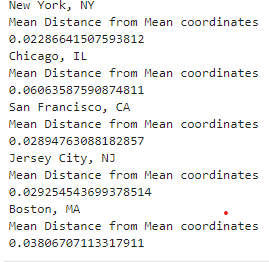
**RESULTS**

THE FOLLOWING ARE THE RESULTS AND VISUALIZATIONS OBTAINED:

The total number of pizzerias on FourSquare are shown below, with only the top 100 selected. The results clearly show that the most pizzerias are in New York, followed by Chicago, Boston, San Francisco and finally Jersey City. Most pizzerias do not mean the closest to the hotel, that needs to be taken into consideration.

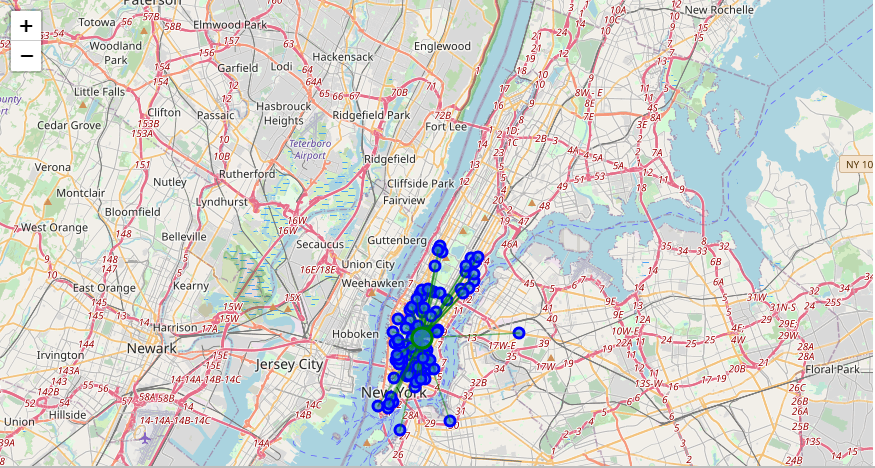
****

The next figure visualizes the mean distance of the pizzerias from the hotel chosen. Most hotels are closer to the hotel in New York city.

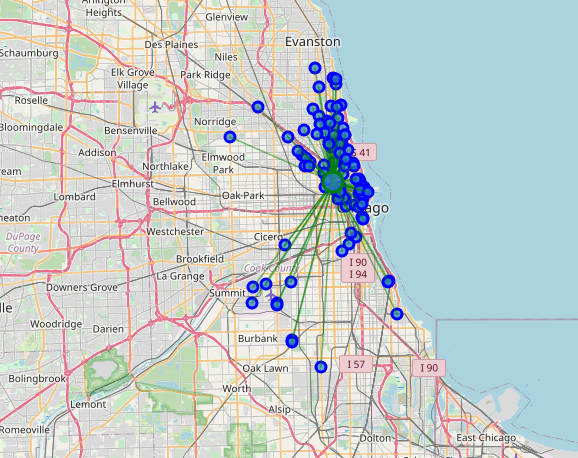
****

**Following are the map visualizations of the pizzerias with reference to the hotel.**

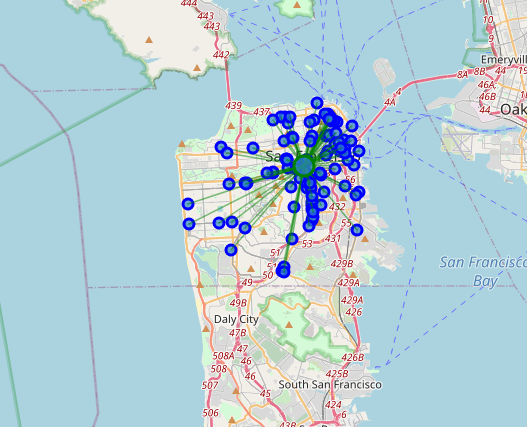
In Jersey City: The pizzerias are not very far from the hotel, with few outliers.

****

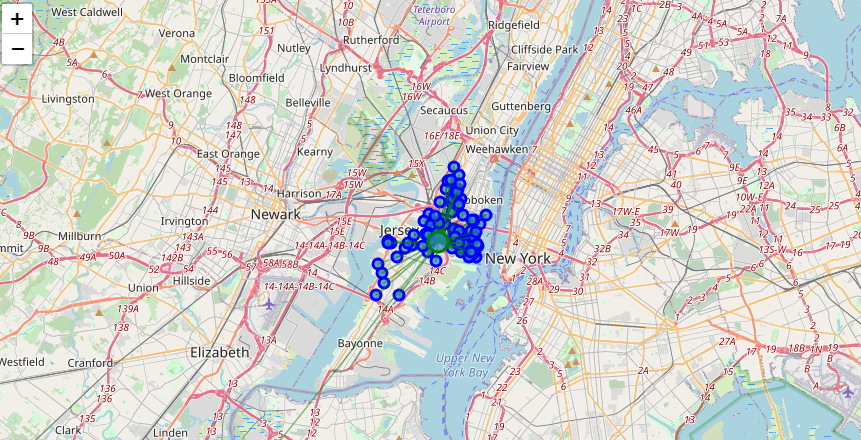
In Chicago City: There are several outliers on the map.

****

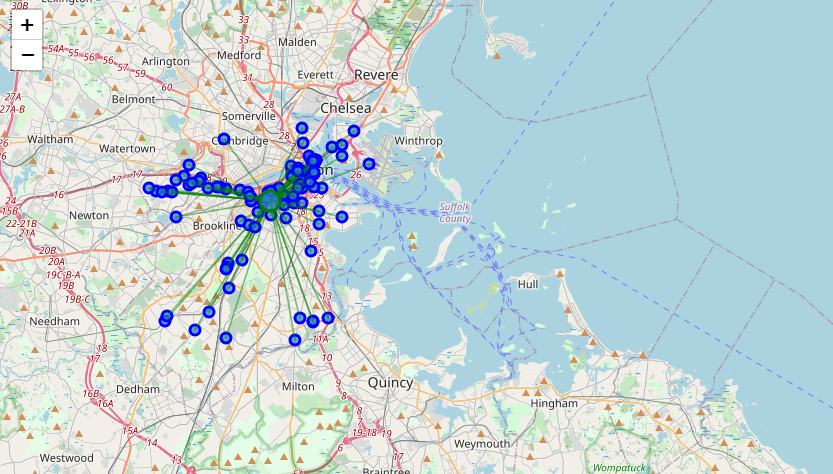
In San Francisco: The stores are a bit scattered

****

In New York: Clustered with a few outliers.

****

Boston: There is a general dispersion.

****

**Discussion**

**Cities which have a lot of parks, recreational areas, monuments and other social areas takes up space in the CBD and spreads the pizzerias apart.**

**CONCLUSION**

After visualization of the pizzerias, Jersey City in New Jersey emerged as the winner, with the lowest mean distance from the hotel.

****