Problem: Which major city has the greatest choice of pizza places in the US?

Introduction:

If someone who has never been to the US wants to try a new pizza everyday in a particular city, how would they go about finding the city with the wide variety of choices for pizza available in a short distance? The aim is to analyse Pizza stores' locations in major US cities and figure out which city has the widest variety of pizza chains on offer. Importantly, the different pizza places should not be spread out too far as that would make it difficult for our tourist to sample them if they are located very far from each other

Data section:

FourSquare API will be used to collect data about locations of Pizza stores in US cities of New York, NY, San Francisco, CA, Jersey City, NJ, Boston, MA and Chicag, IL.

Methodology:

Main target here is to assess which city has the highest density of pizza stores. FourSquare API through venues channel will be used. Near query will be used to get venues in the cities. Also CategoryID will be used to look at only pizza places Once their top 100 venues are recognised, the names and coordinate data from the result will be used to plot them on a map.

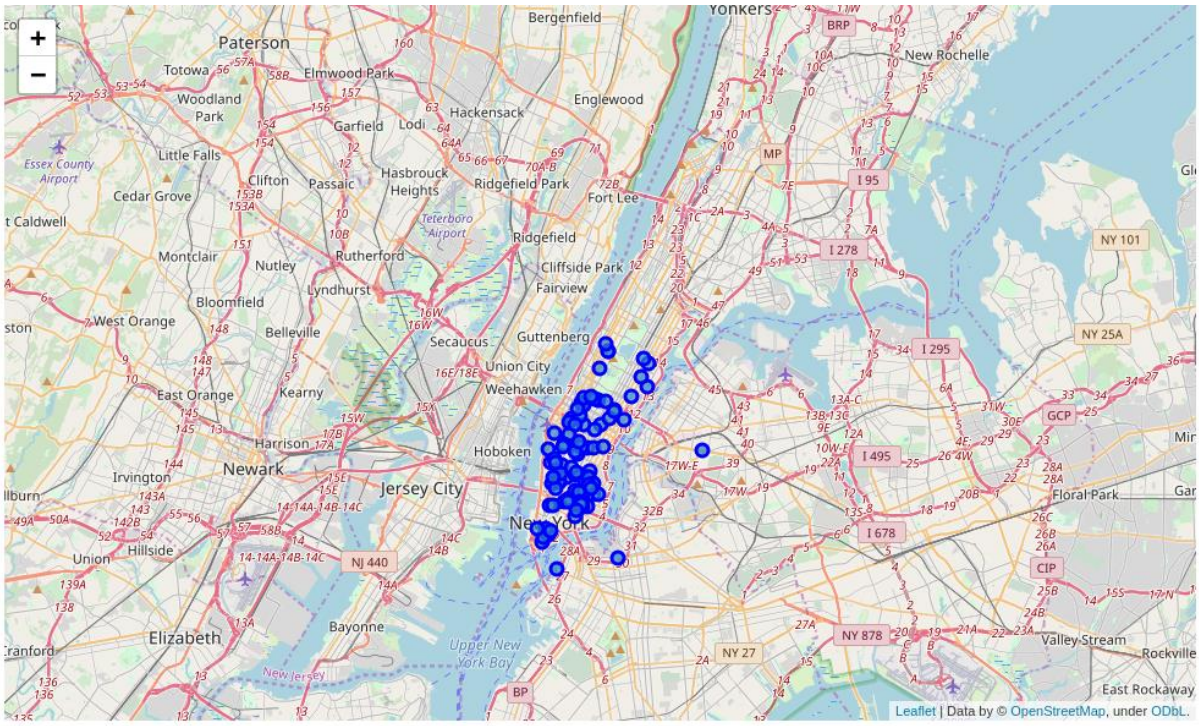
After this, to get an indication of density of pizza places, the centre coordinate of venues will be calculated to extract the mean longitude and latitude values. Post that, mean Euclidean distance from each venue to the mean coordinates will be calculated.

This should give an idea about the city with the widest variety of pizza places at an accessible distance.

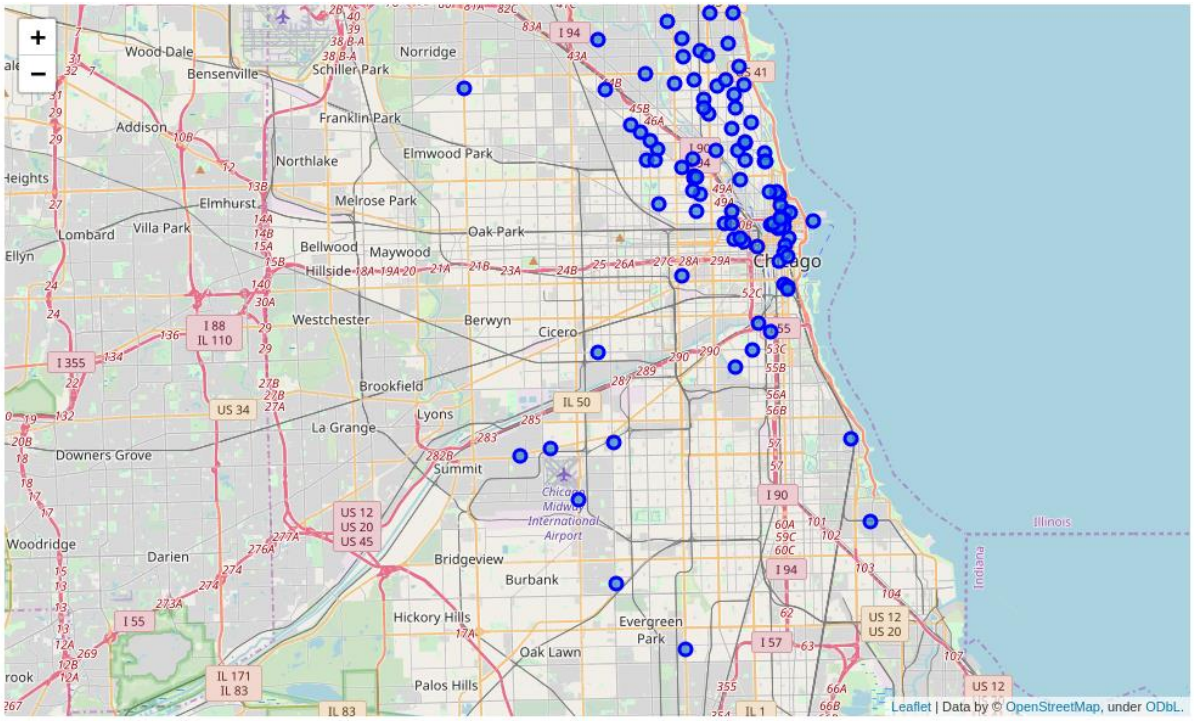
Results

From the initial visual inspection, one can see that all cities analysed here have multiple pizza places. The following here are the pictures of the geoplot generated with folium:

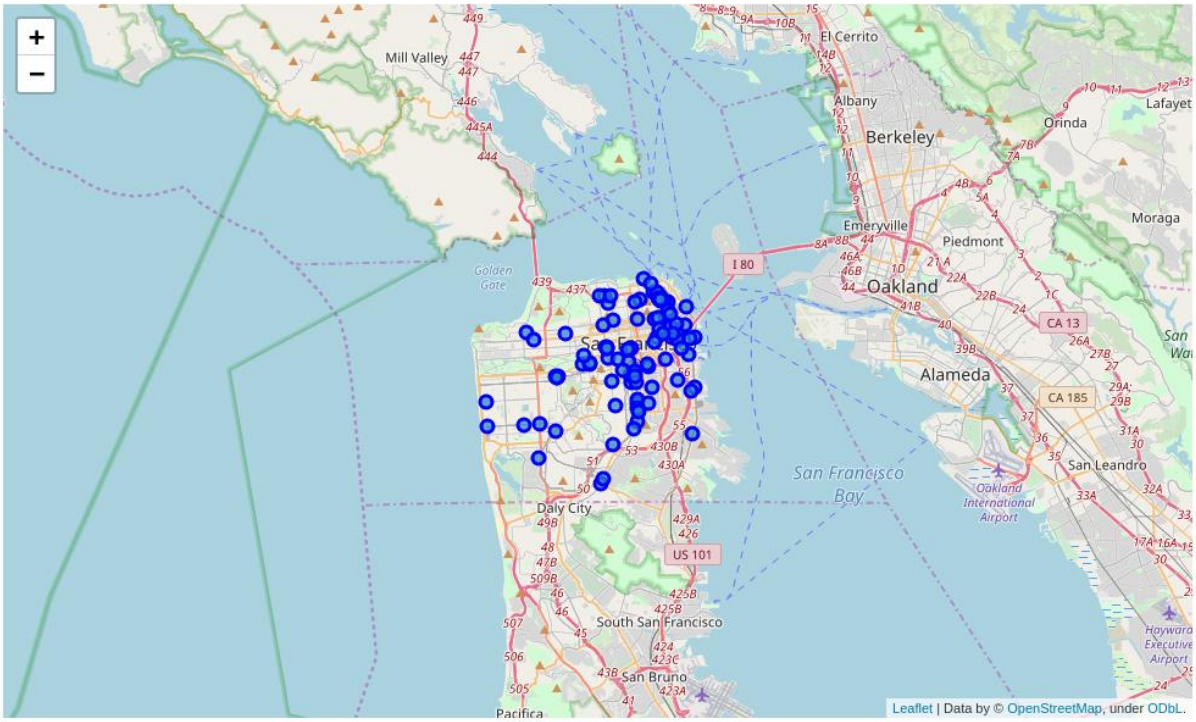
New York:



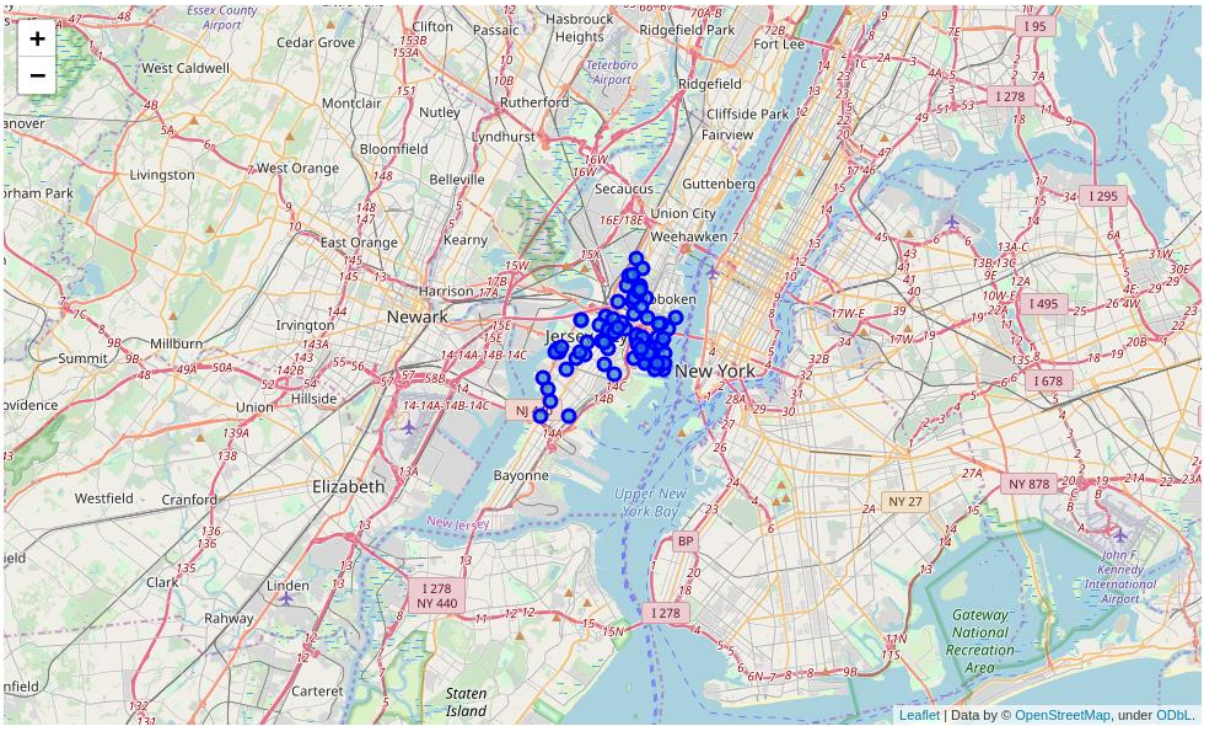
Chicago:



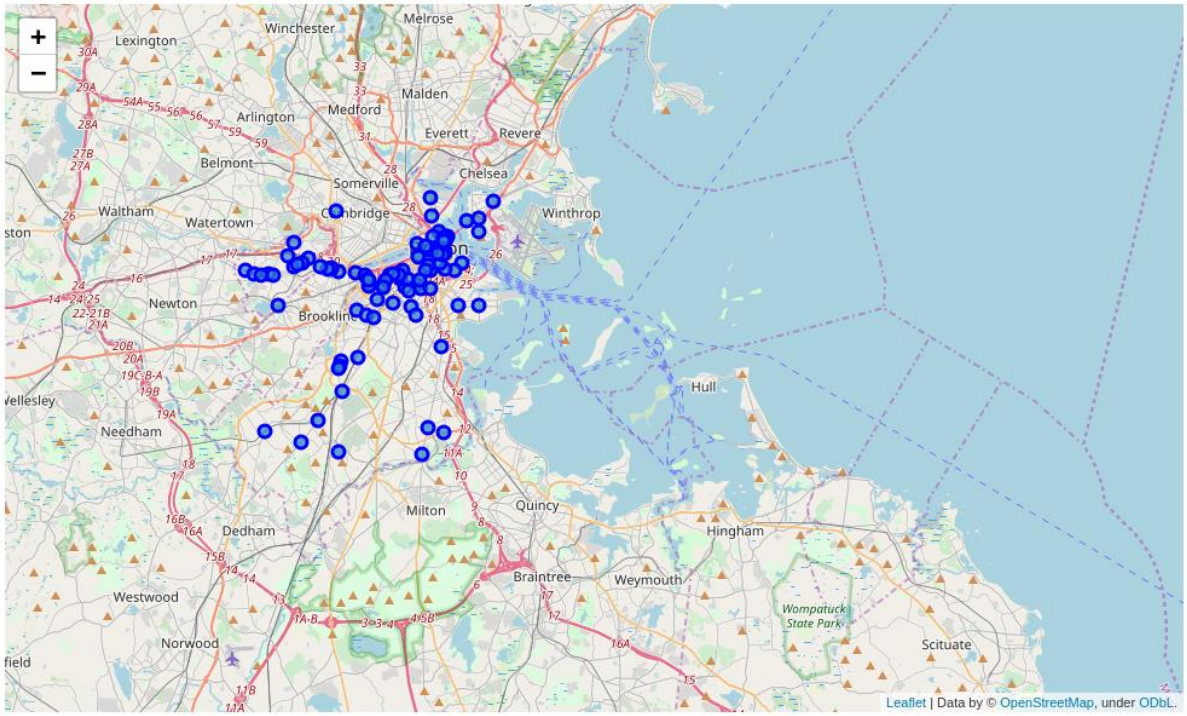
San Francisco



Jersey city:



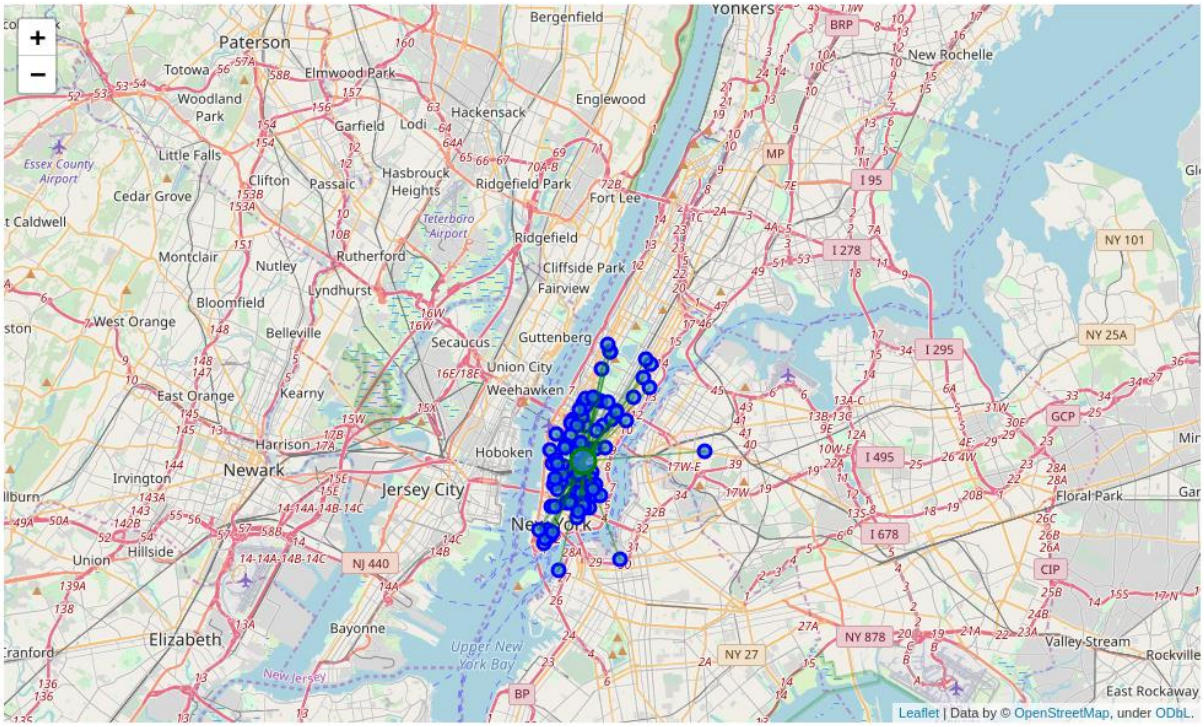
Boston:



Upon first inspection we see that New York, Jersey City and San Francisco are the cities with the highest density of pizza places. In the next phase the mean coordinate and the mean distance to mean coordinate (MDMC) was calculated. The mean coordinate with a big green circle and distances with green lines are represented.

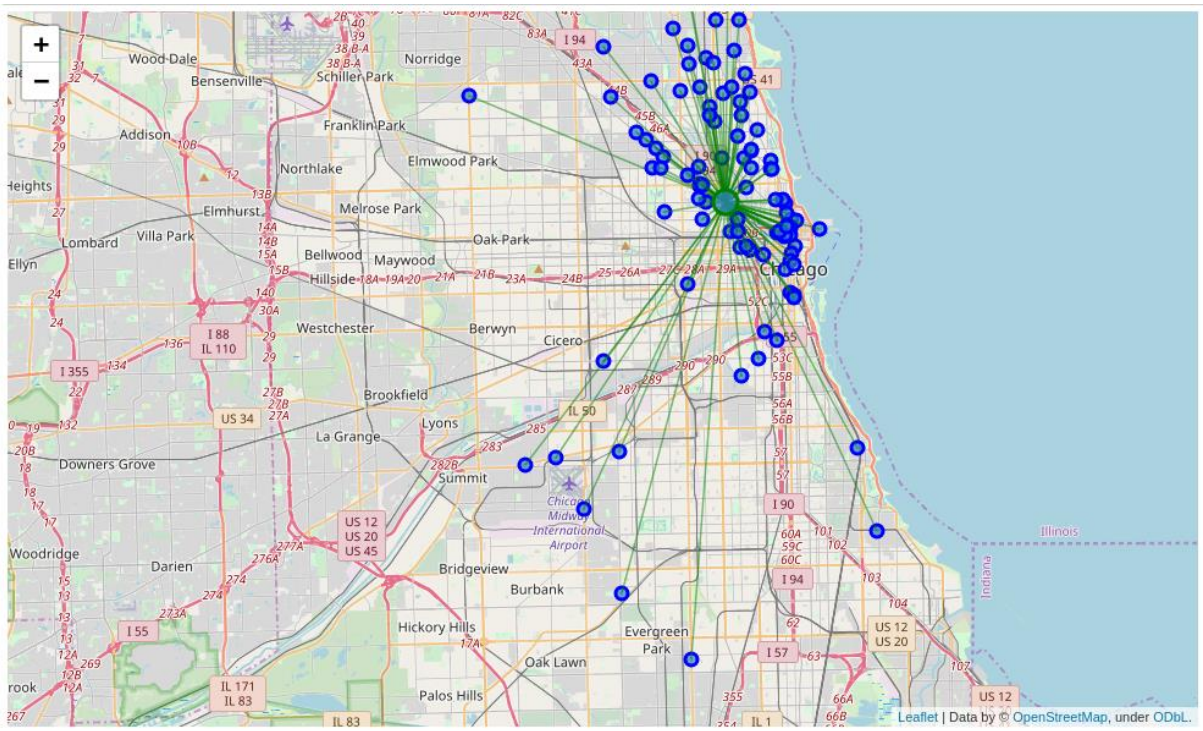
New York:

MDMC: 0.021556



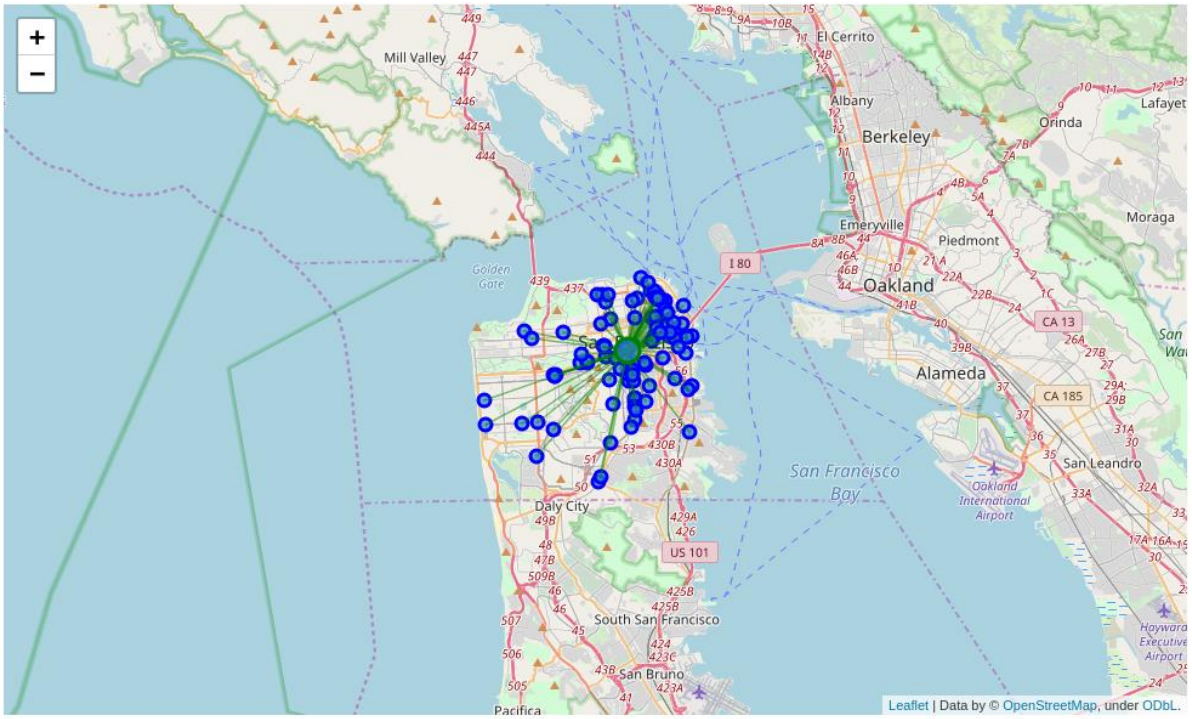
Chicago:

MDMC: 0.052805



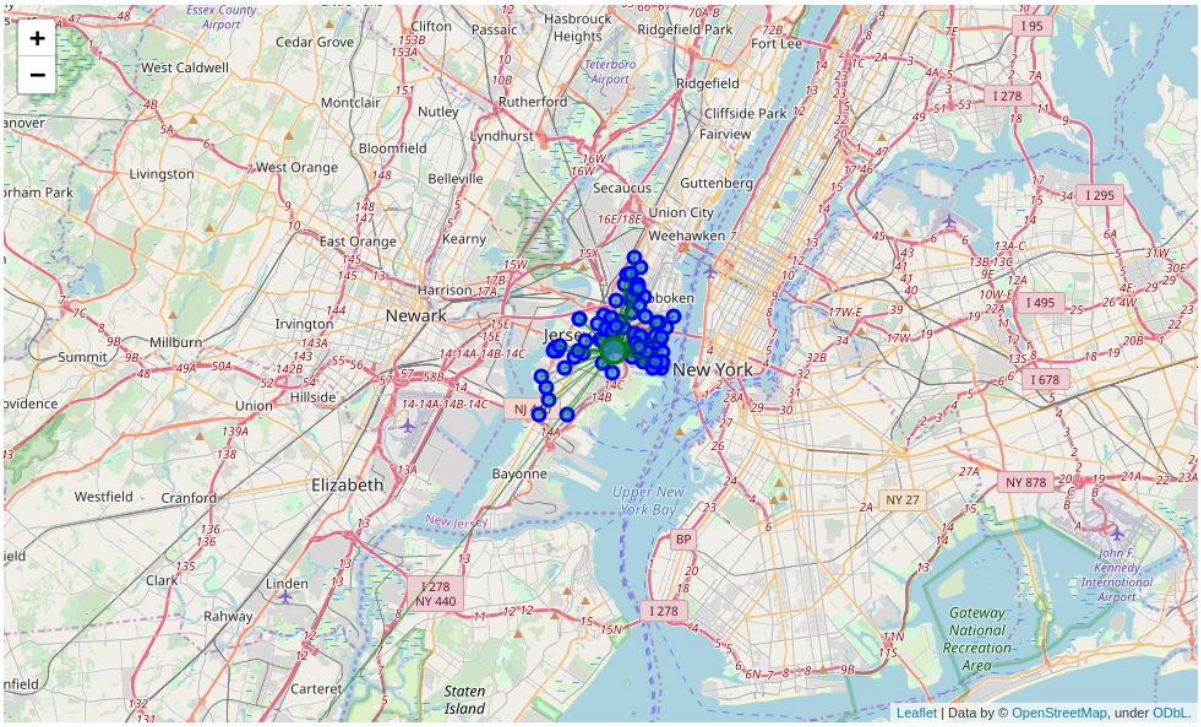
San Francisco:

MDMC: 0.028633



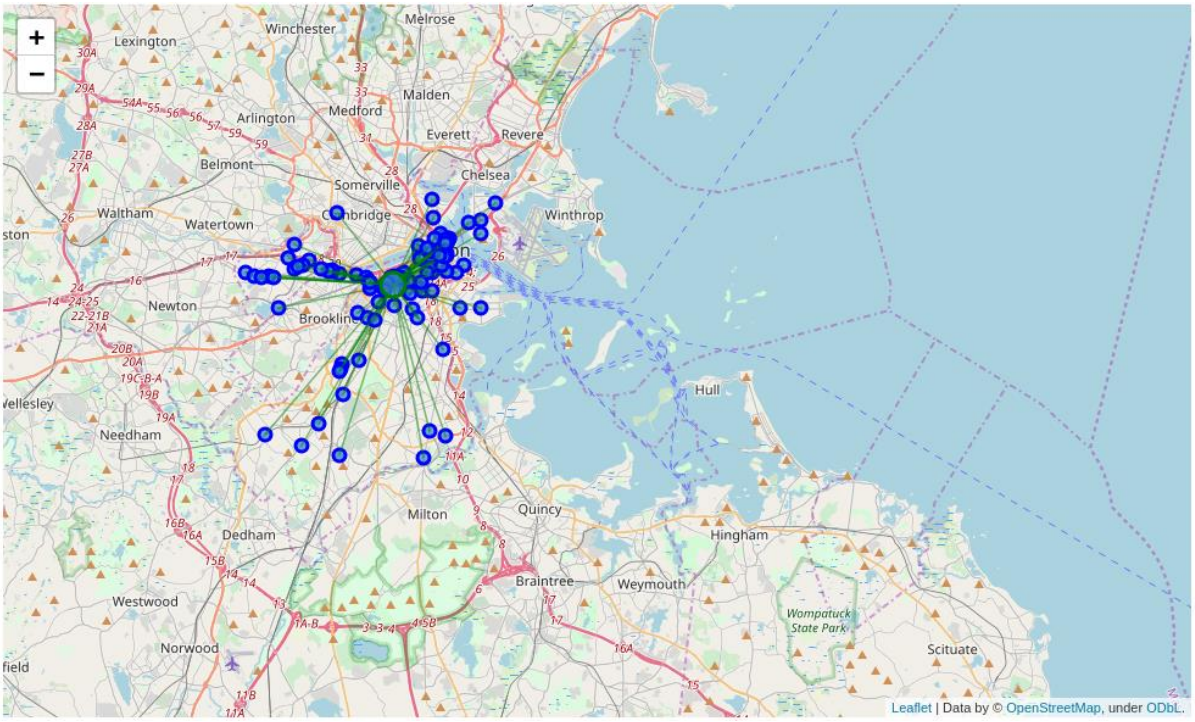
Jersey City:

MDMC: 0.029950



Boston:

MDMC: 0.035126



Therefore the results are :

1. New York
2. San Francisco
3. Jersey City
4. Boston
5. Chicago

Conclusion:

New York has the highest number of pizza places in short distance from each other, therefore the tourist can book a hotel room near the mean coordinate and enjoy the wide variety of pizza places at his disposal.