Best Zoo in the United States

Data Science Capstone – The Battle of Neighborhoods

Introduction

Let us say you have never been to the United States and you want to visit the best zoos while you are there. Therefore, you want to go to a place with a high density of zoos around you. The problem we aim to solve is to analyze the zoos' locations in the major US cities and find the best place for our tourist so that he can have the best zoo experience. Our main target are tourist families.

Data section

I will use the Foursquare API to collect data about zoo locations 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 and I am hopeful that they will contain the best zoo in the US.

Methodology

My main target here is to asses which city would have the highest zoo density. I used the Foursquare API through the venues channel. I used the near query to get venues in the cities. In addition, I use the Category ID to set it to only show zoos. An Example of my requests:

```
https://api.foursquare.com/v2/venues/explore?&client_id=&client_secret=&v=20180605&New York, NY&limit=100&categoryId=4bf58dd8d48988d1ca941735
```

That 4bf58dd8d48988d1ca941735 is the Id of the zoo category. In addition, Foursquare limits us to maximum of 100 venues per query.

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

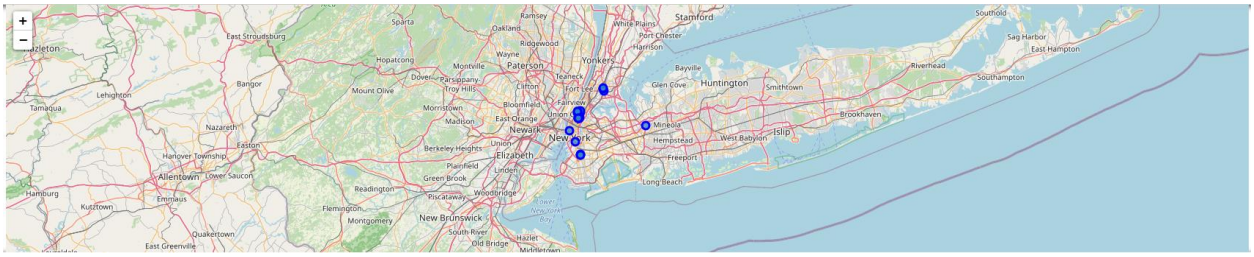
Next, to get an indicator of the density of zoos, I calculated a center coordinate of the venues to get the mean longitude and latitude values. Then I calculated the mean of the

Euclidean distance from each venue to the mean coordinates. That was my indicator; mean distance to the mean coordinate.

Results

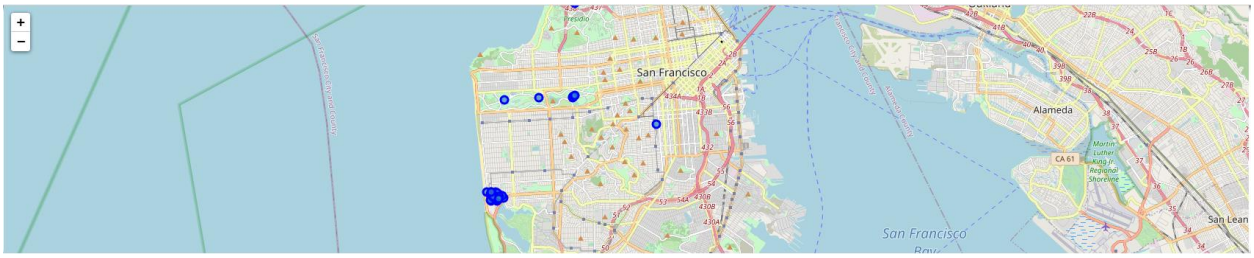
For the first visual inspection, we can see that they all have multiple zoos. The following visualizations are pictures of the geoplot generated with folium:

New York:

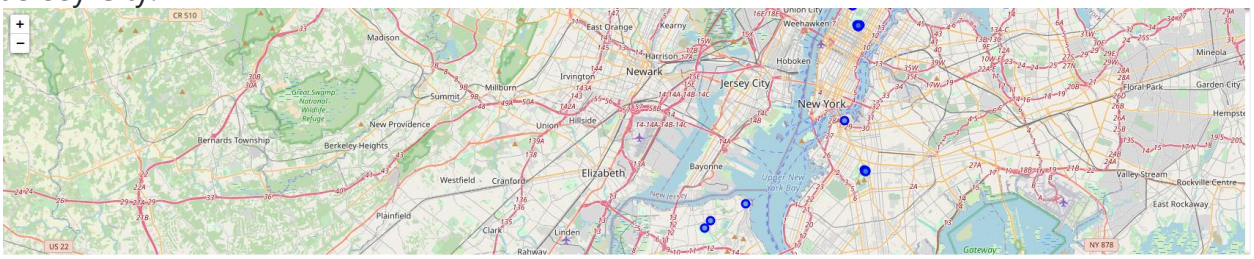


Chicago:

San Francisco:



Jersey City:

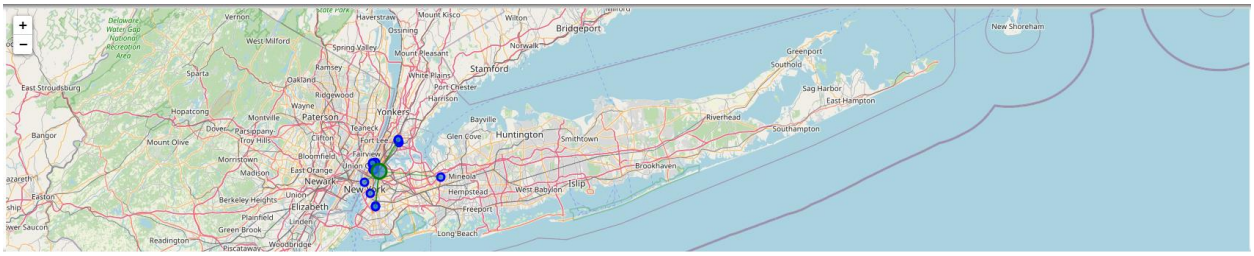


Boston:

Upon First inspection we see that San Francisco, and Illinois are the most densely cities. In the next phase, we will calculate the mean coordinate and the mean distance to mean coordinate. We represent the mean coordinate with a big green circle and distances with green lines.

New York:

```
Mean Distance from Mean coordinates
0.05351486288380877
```

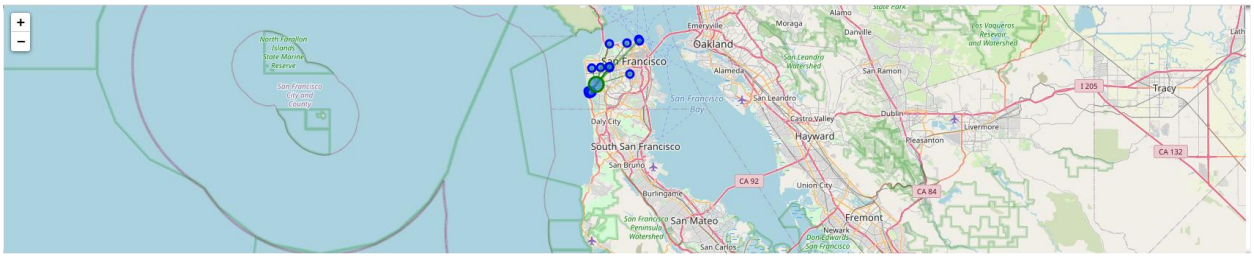


Chicago:

```
Mean Distance from Mean coordinates
0.16471304536264
```

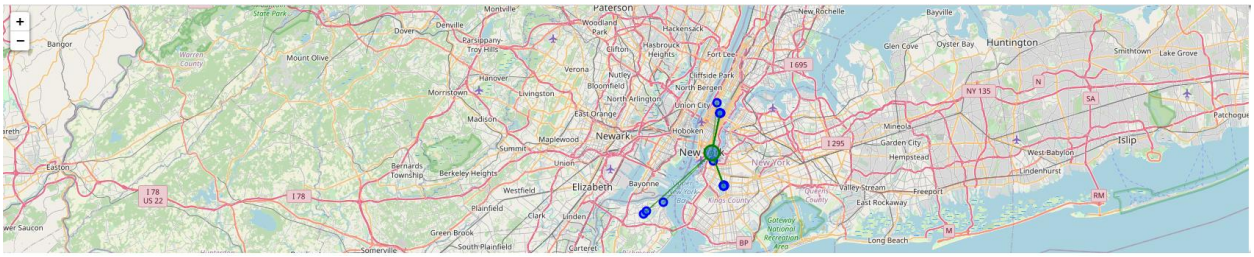
San Francisco:

```
Mean Distance from Mean coordinates
0.025891823899077602
```



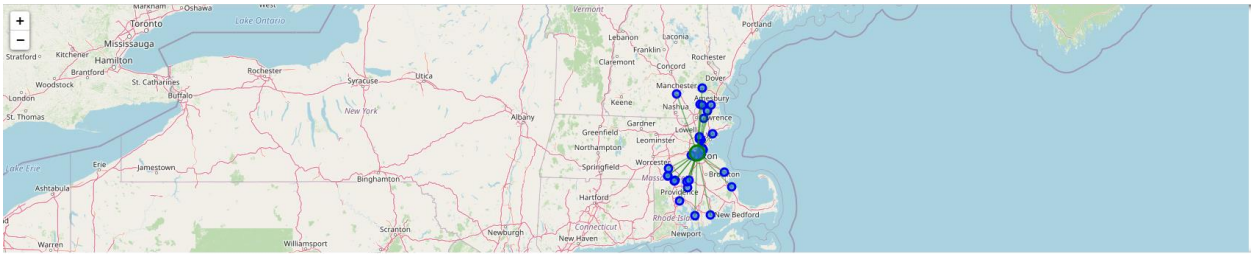
Jersey City:

```
Mean Distance from Mean coordinates
0.0648934717472367
```



Boston:

```
Mean Distance from Mean coordinates
0.2994649827539902
```



Discussion

One thing I noticed in the figure is that there is a cluster of zoos in California that I thought might be the same zoo in San Francisco, CA. If this were the case, the best state to go to for zoos could be Chicago, IL.

One consideration to do further work on is to move the location of the Foursquare API query until we get all the zoos in each city and do the calculations again.

Conclusion

We can conclude that San Francisco, California is the best, most dense area to visit in the US for zoo entertainment.