Which major city should you visit if you love Japanese food?

Table of contents

Introduction

So, you love Japanese food so much that you can eat it for every meal. And you would like to visit a major city for a nice vacation and would like you indulge in your favorite cuisine but don't know where to go. This is the problem I aim to solve with my report by analyzing the density of Japanese restaurants in some major US cities. While not important to some, this report will help the tourist that would like to visit a popular destination in the US, that has a huge appetite for sushi and the like.

Data

I will use the FourSquare API to collect data about locations of Japanese restaurants in 5 major US cities which are: New York, San Francisco, Las Vegas, Seattle and Chicago. These are some of the most popular US cities to visit according to Trip Advisor and I believe that these cities have some of the best Japanese food the United States can offer. I will use the latitude and longitude of the cities within an API request to pull back the name, address and geospatial data on Japanese restaurants in the area. We will use this data to determine which city would be best to visit and reserve a hotel room.

Methodology

We will be using statistics to solve the majority of our problem. We will obtain the mean location of the Japanese restaurants and then take the average of the distance of the venues to the mean coordinates in order to illustrate the density.

We will then plot our results on a map so we can visualize our data in order to make the results clear.

Machine learning doesn't appear to be required for this particular issue as knowing the density of the restaurants should suffice to reach our answer.

Analysis

Before we analyze any data, we will need to import the libraries required for the analysis. We will need the following libraries:

```python
import numpy as np # library to handle data in a vectorized manner

import pandas as pd # library for data analsysis
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import json # library to handle JSON files

from geopy.geocoders import Nominatim # convert an address into latitude and longitude values

import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into a pandas dataframe

# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors

import folium # map rendering library
```

*Now **we will make a request to the FourSquare API and store the data in a json file** and normalize our results in order map them out. We will limit the results to the top 100 venues, in our case, Japanese restaurants, in each city.*

Folium is used to plot the results we have for our Japanese restaurant geospatial data on a map of their respective cities.

Here are the results prior to plotting the top 100 Japanese restaurants in each city:

```
Total number of Japanese restaurants in New York, NY = 328

Total number of Japanese restaurants in Chicago, IL = 260

Total number of Japanese restaurants in San Francisco, CA = 270

Total number of Japanese restaurants in Las Vegas, NV = 262

Total number of Japanese restaurants in Seattle, WA = 237
```
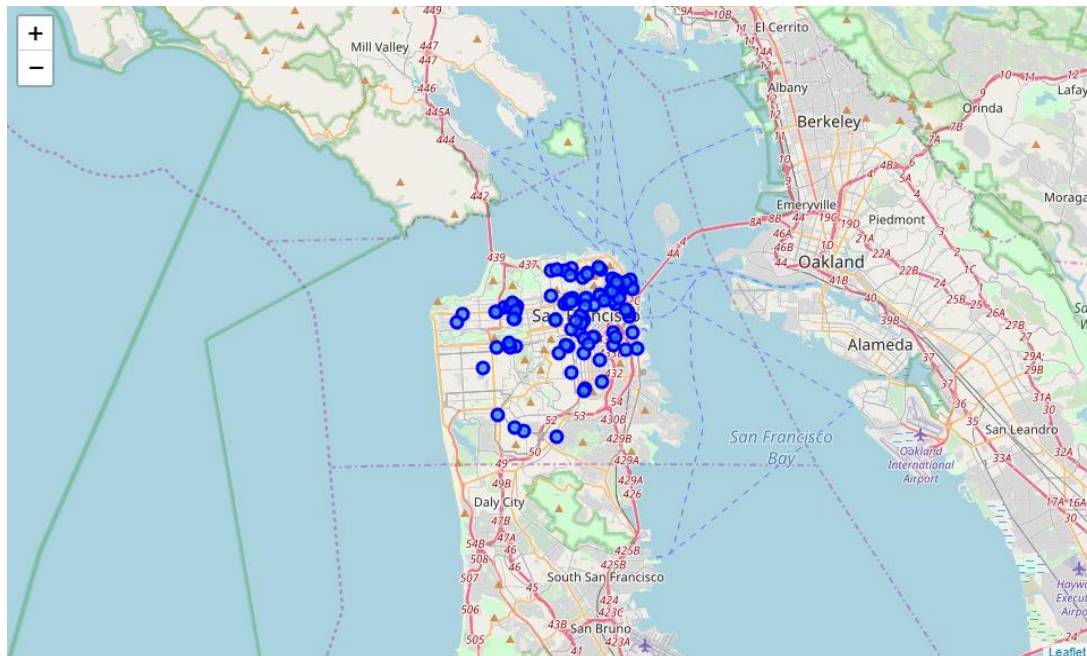
Now let's take a look at the maps of each city to visualize our data.
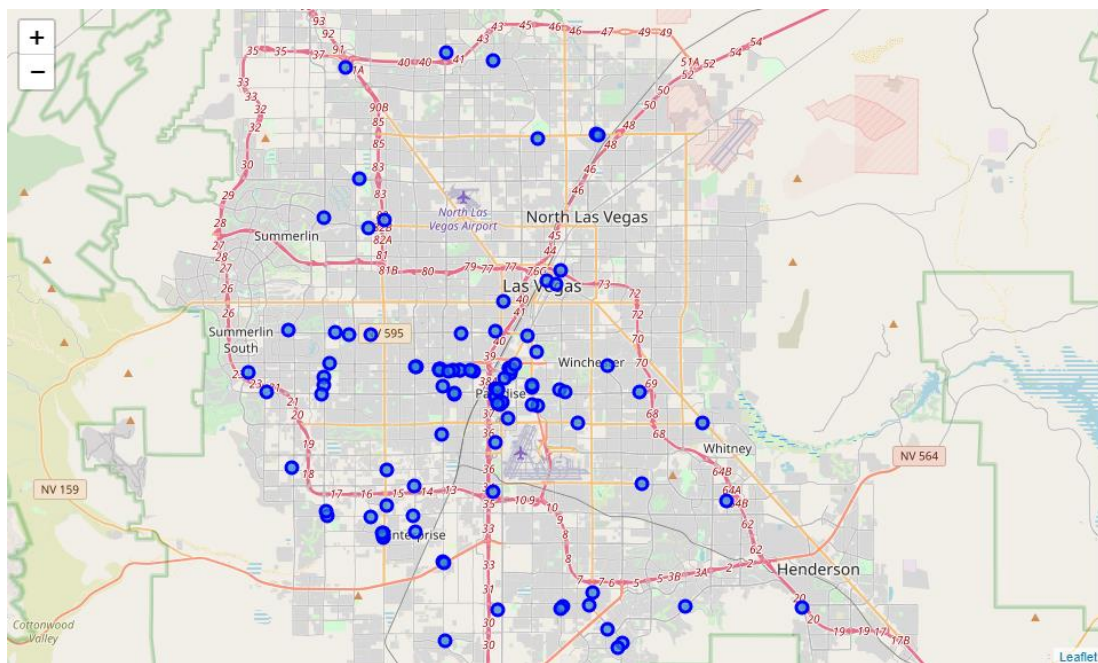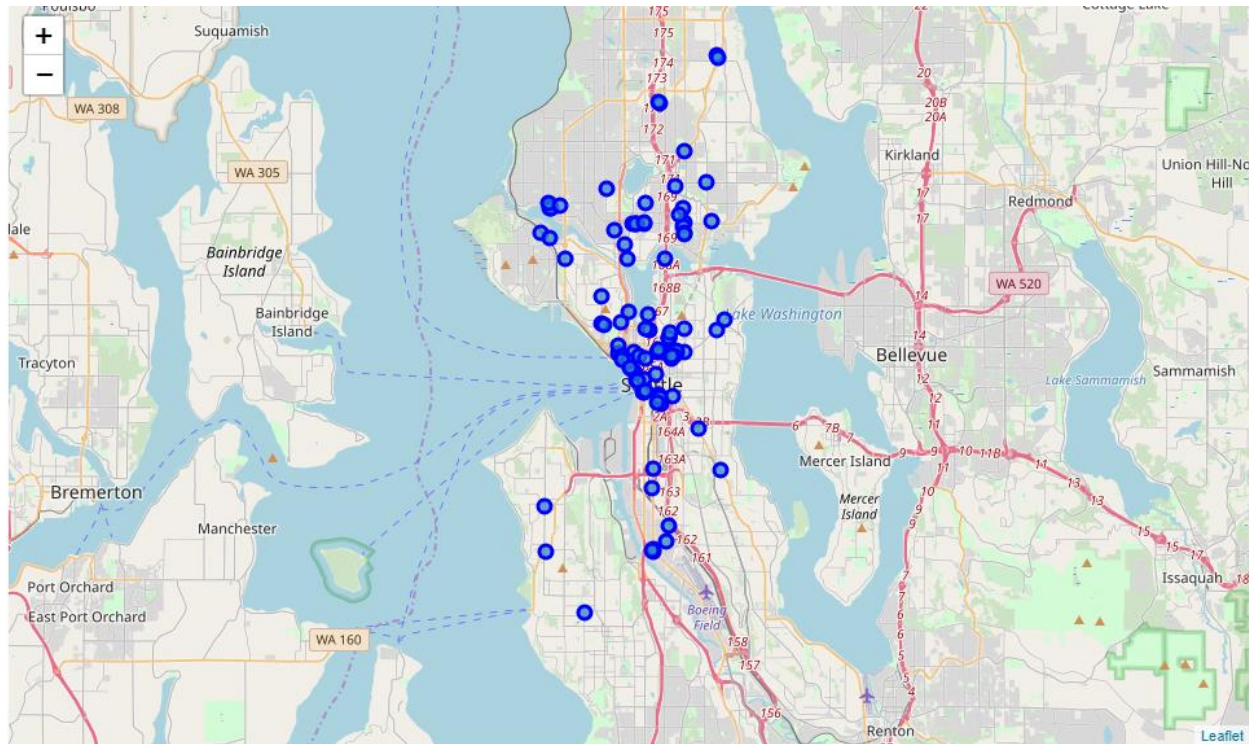
Map of New York

Map of Chicago

Map of San Francisco

Map of Las Vegas

Map of Seattle

Now that we can visualize the density of Japanese restaurants in each city, we can see that New York appears to be the best option due to the high number of Japanese restaurants within proximity to each other. This will make vacation planning easy as our eatery of choice will not inhibit the ability to search for the best hotel that will suit the price point for any tourist.

However, let's make sure we are correct by getting an accurate measure of the restaurant density in each city and then visualize them by plotting them on a map.

We will be using statistics to solve this portion of our problem. We will obtain the mean location of the Japanese restaurants and then take the average of the distance of the venues to the mean coordinates in order to illustrate the density.

Let's take a look at the code that will do this for us:

```python
maps = {}
for city in cities:
    city_lat = np.mean([results[city]['response']['geocode']['geometry']['
bounds']['ne']['lat'],
                        results[city]['response']['geocode']['geometry']['
bounds']['sw']['lat']])
    city_lng = np.mean([results[city]['response']['geocode']['geometry']['
bounds']['ne']['lng'],
                        results[city]['response']['geocode']['geometry']['
bounds']['sw']['lng']])
    maps[city] = folium.Map(location=[city_lat, city_lng], zoom_start=11)
```

```python
    venues_mean_coor = [df_venues[city]['Lat'].mean(), df_venues[city]['Ln
g'].mean()]
    # add markers to map
    for lat, lng, label in zip(df_venues[city]['Lat'], df_venues[city]['Ln
g'], df_venues[city]['Name']):
        label = folium.Popup(label, parse_html=True)
        folium.CircleMarker(
            [lat, lng],
            radius=5,
            popup=label,
            color='blue',
            fill=True,
            fill_color='#3186cc',
            fill_opacity=0.7,
            parse_html=False).add_to(maps[city])
        folium.PolyLine([venues_mean_coor, [lat, lng]], color="red", weigh
t=1.5, opacity=0.5).add_to(maps[city])

    label = folium.Popup("Mean Co-ordinate", parse_html=True)
    folium.CircleMarker(
        venues_mean_coor,
        radius=10,
        popup=label,
        color='red',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(maps[city])

# Let's see what our results are without ploting them
    print(city)
    print("Mean Distance from Mean coordinates")
    print(np.mean(np.apply_along_axis(lambda x: np.linalg.norm(x - venues_
mean_coor),1,df_venues[city][['Lat','Lng']].values)))
```
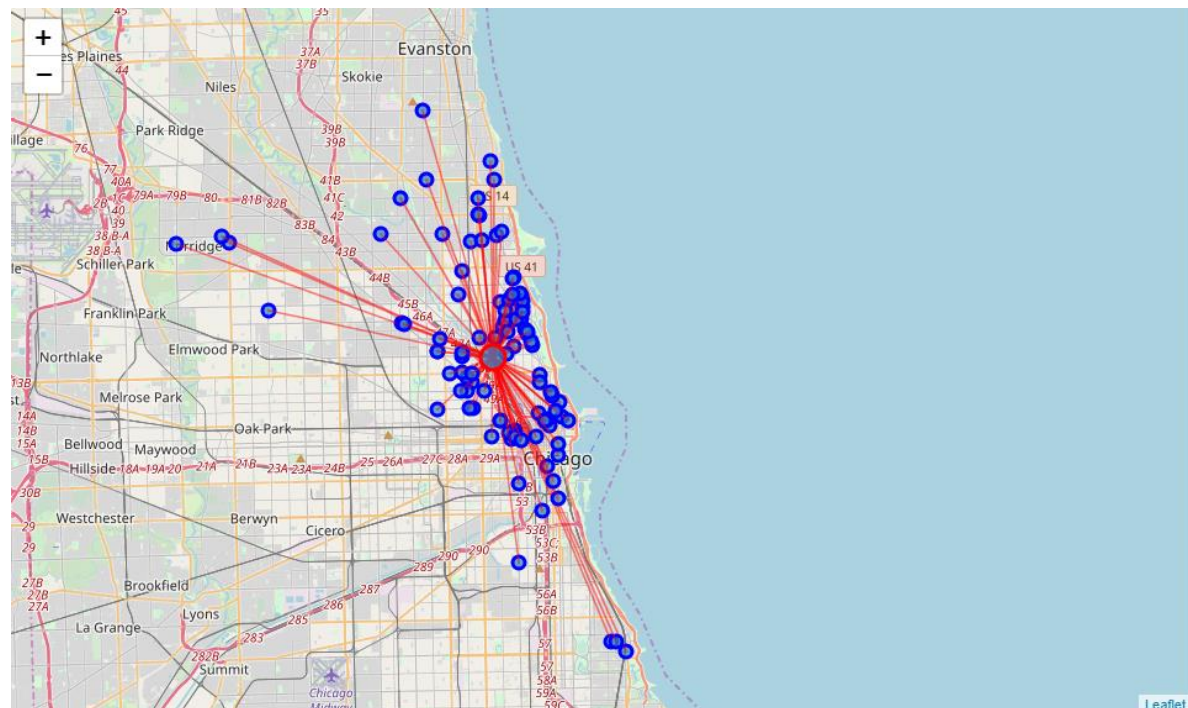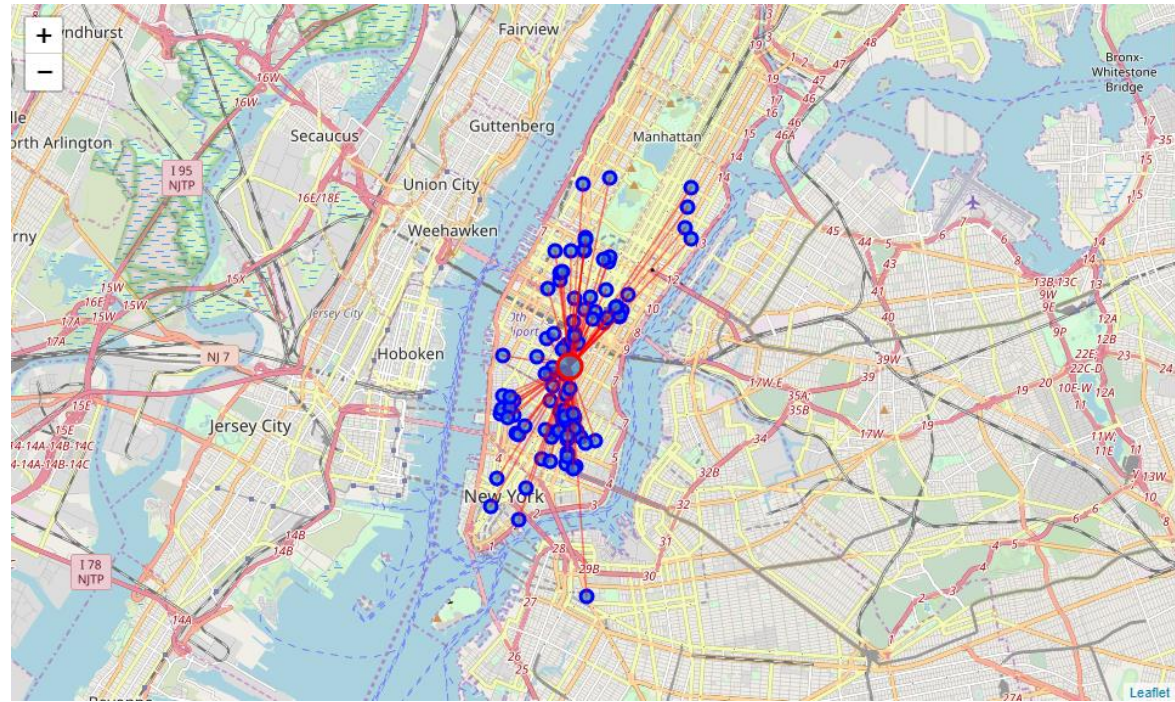
The result of this code allows us to map the true density of Japanese restaurants in each city and will print the following:

```
New York, NY
Mean Distance from Mean coordinates
0.01809379278802192
Chicago, IL
Mean Distance from Mean coordinates
0.04497019707422612
San Francisco, CA
Mean Distance from Mean coordinates
0.02742247138198929
Las Vegas, NV
Mean Distance from Mean coordinates
0.0719585335788342
Seattle, WA
Mean Distance from Mean coordinates
0.037010855726453955
```
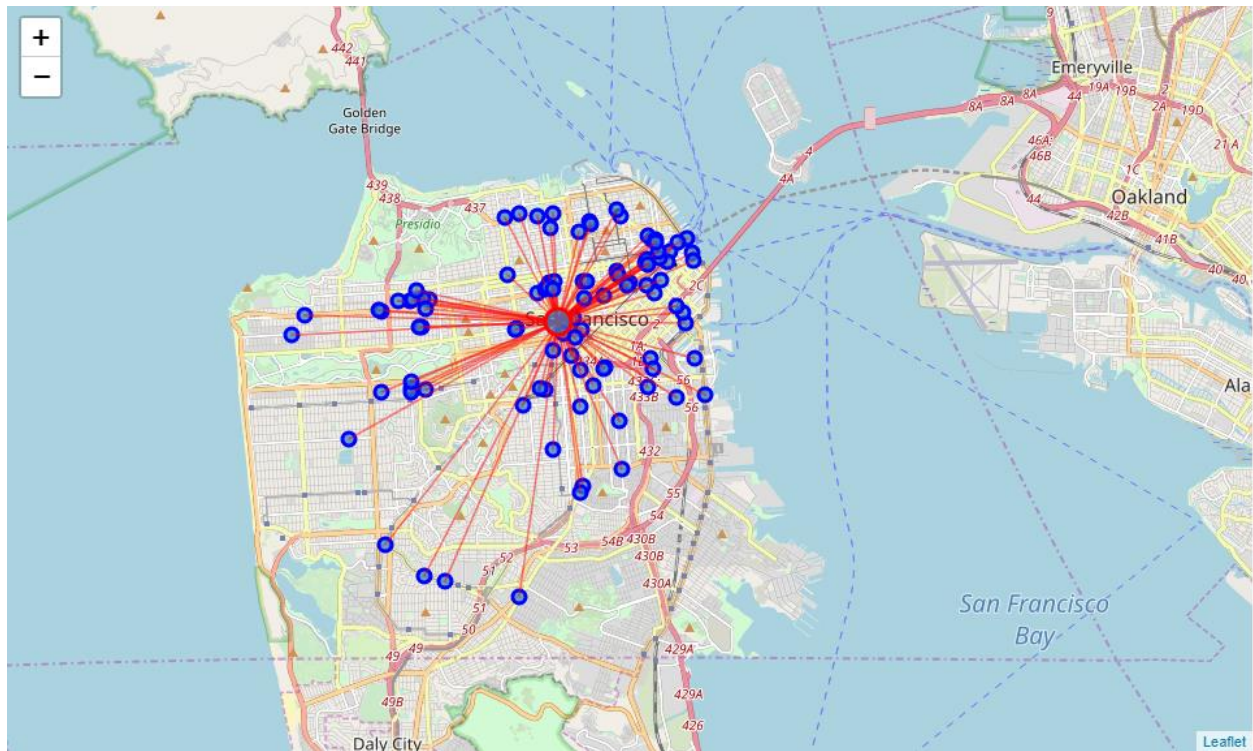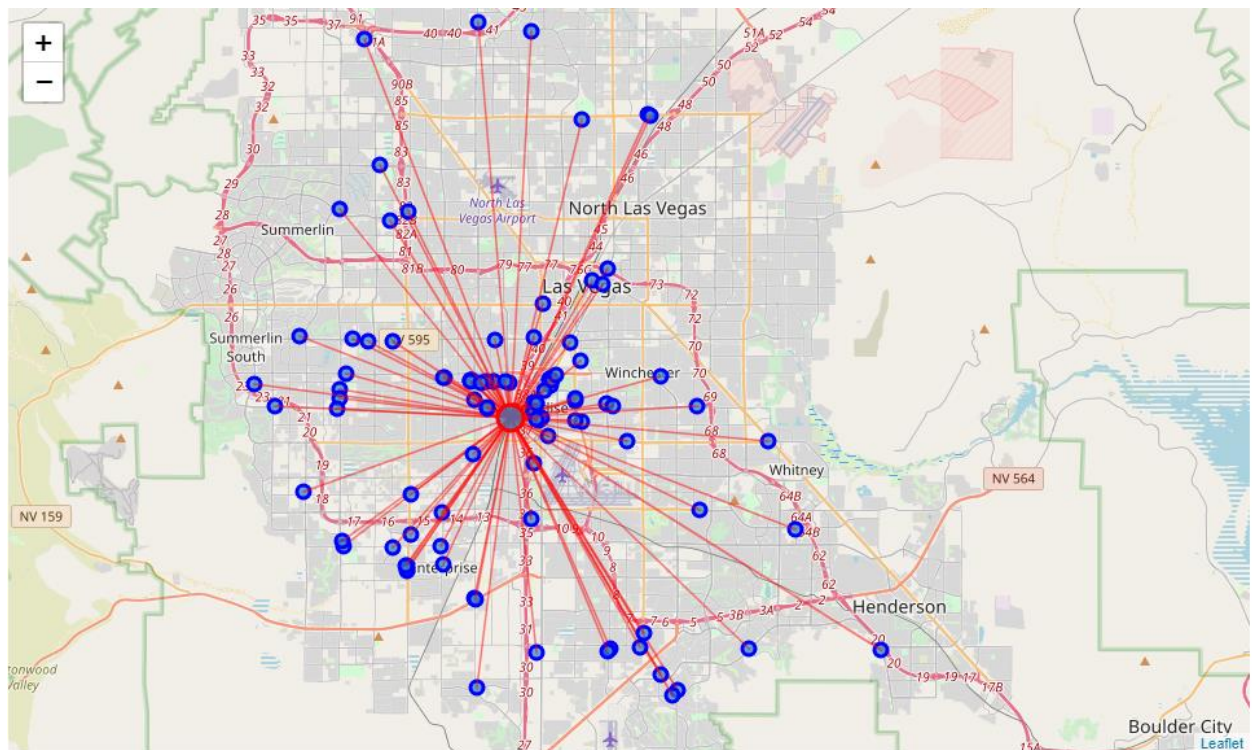
Now we can see the new maps we created that confirms our preliminary findings.

Map of New York

Map of Chicago

Map of San Francisco
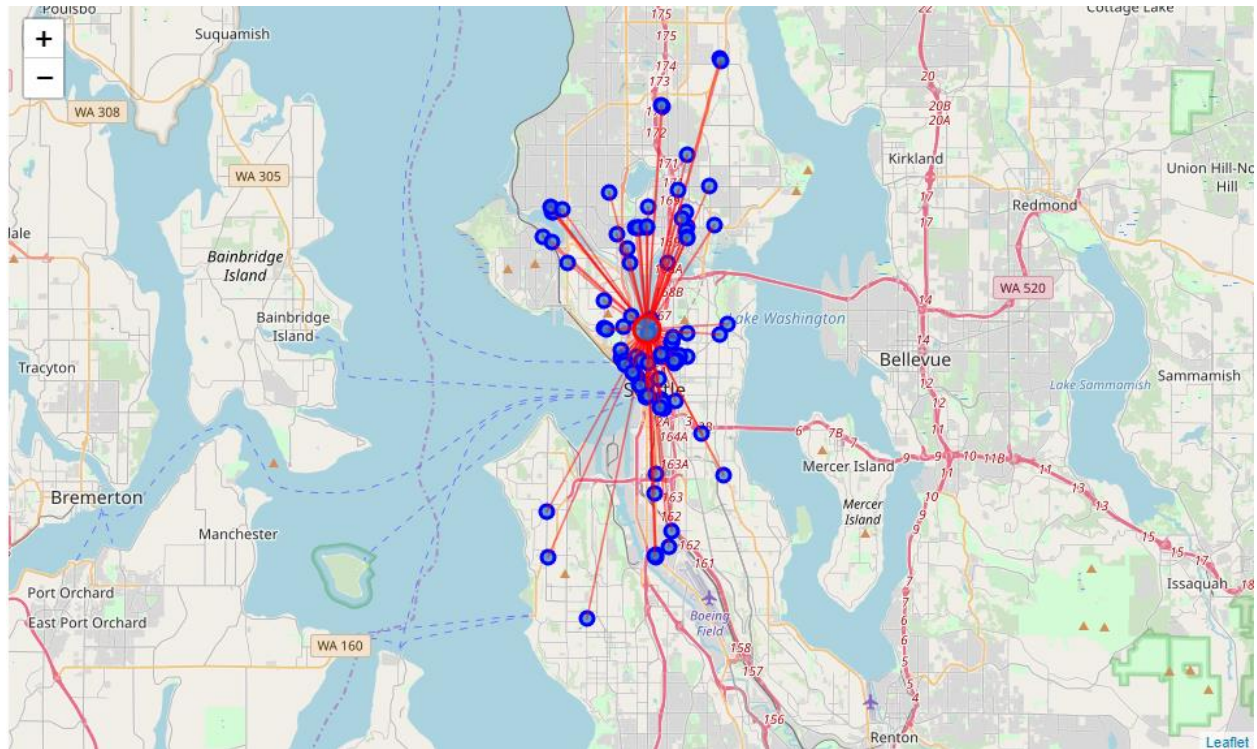

Map of Las Vegas

Map of Seattle

Results and Discussion

So what is the result of our analysis you say? We can see that by figuring the mean distance from the mean coordinates of Japanese restaurants in each city that New York has the highest density of restaurants which appears to make it ideal for a tourist that must have sushi or tempura to survive the day.

The next best city to visit for those who see themselves as connoisseurs of Japanese food would be San Francisco. It has the second highest density of Japanese restaurants.

One thing that both of these great cities have in common when it comes to Japanese cuisine, they both have some of the best ports of the world. They are each world renowned for having the best sea food arrive at their ports each day.

Conclusion

You really can't go wrong with any of these cities as they each have something great to offer outside of Japanese cuisine. However, the clear winner is good ol' New York City.

Resources

- FourSquare API
- TripAdvisor's Top 25 most popular travel destinations in America