Capstone Proyect - The Battle of the Neighborhoods - Week 1

Applied Data Science Capstone IBM/Coursera

1. Introduction: Business Problem

Problem Background:

We focus our business problem on the North American states, as they are much larger and we have much more information. Because they are such large populations, apart from the people who live there, they attract a lot of tourism. People who travel and want to know everything about it. And when you travel, it means that you have to eat and you have to look for places to eat, and we are going to look for specific places to eat. But we leave out the typical ones like hamburgers and pizzerias and in this case we will focus on Mexican food.

We have a highly competitive market and more with the theme of food, since wherever you go, you have to eat, and if we go out of the usual, as we said before, pizzas and hamburgers, we will analyze the market for Mexican food, which has many people like, because they like spicy.

Problem Description:

A place that offers food, either a restaurant, or a fast food place, or any place that resembles it, where the food is prepared and served together with the drink, to be taken in the same place in the case of a restaurant or to be taken away if it is a fast food place, which offers you both options, or to be taken in the same place or to be taken away.

We are going to focus on the cities that have a larger population like New York, to see in which of them we can find more Mexican food places. In all these cities we find a great gastronomic culture where a great amount of international cuisines are included, but in this case, going a little bit out of the ordinary, we focus on Mexican food.

At the time we are, there are many tourists, who like this type of food more and they make their trips thinking about the possibility of being able to taste this type of food, outside their native land and for the simple fact that it is their favorite, and they want to stay, near a place that offers the type of food they want to eat. That's why we focus our study on tourists who want to eat that kind of food when they go on a trip to those cities.

With this, we are going to offer a possibility to strategically plan a trip so that wherever you stay, you have a chance to eat the kind of food you like.

Target Audience

We will recommend our tourists the best location to stay depending on the cities they visit, to ensure that they can eat the kind of food they like, which in this case will be Mexican food.

Success Criteria

To be successful in our project, we must make a good recommendation to tourists, because if they are looking for a specific type of food, what they want is precisely that, to be able to choose their hotel according to the areas where they have a greater number of restaurants or Mexican food establishments.

2. Data

Depending on the problem we want to solve, the factors that will influence our decision are

- The number of restaurants in the neighborhood where you will be staying during your vacation
- The number and distance of restaurants there are.
- The distance from the place you are staying to the city centre.

The tool we decided to use is the Foursquare API. We start from what is considered the center of the city that we are going to visit and focus there we will see the restaurants or Mexican food places that are there and the distance to them.

The following data sources will be needed to extract/generate the required information

- We will start from the centers of the candidate areas that will be obtained using Foursquare API.
- We will also obtain the number of restaurants or Mexican food places and their location.

The cities on which we will focus our study are the cities with the largest populations, extracted from the following page:

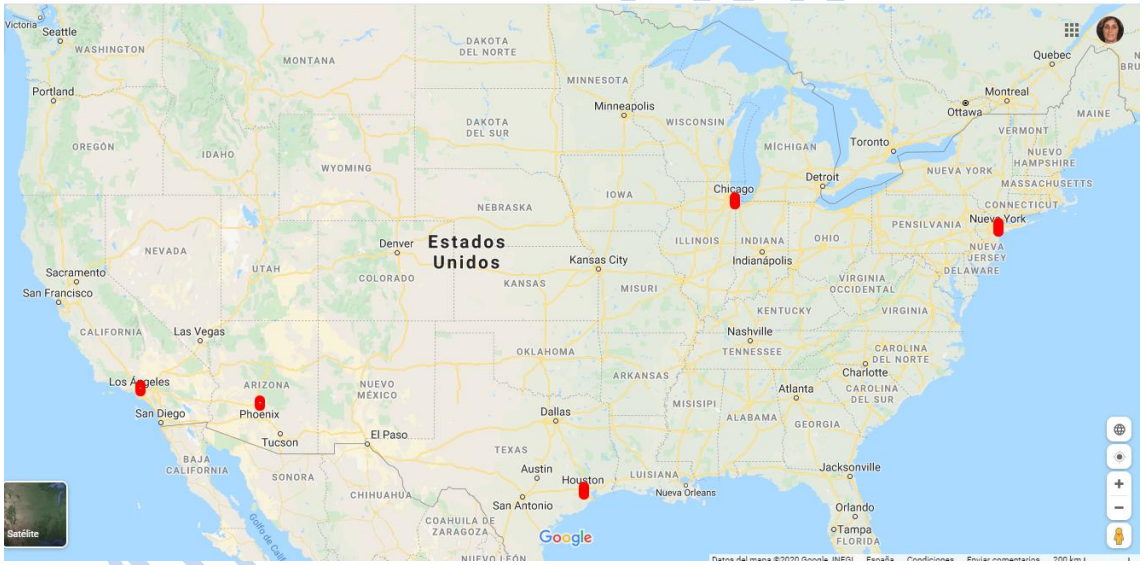
http://worldpopulationreview.com/us-cities/

Keeping the top five

The 200 Largest Cities in the United States by Population 2020

What are the largest cities in the United States? The US city with the biggest population is New York City. Here is a list of the top ten most populated cities in the US:

1. New York City, NY (Population: 8,601,186)
2. Los Angeles, CA (Population: 4,057,841)
3. Chicago, IL (Population: 2,679,044)
4. Houston, TX (Population: 2,359,480)
5. Phoenix, AZ (Population: 1,711,356)
6. Philadelphia, PA (Population: 1,576,596)
7. San Antonio, TX (Population: 1,565,929)
8. San Diego, CA (Population: 1,453,775)
9. Dallas, TX (Population: 1,379,735)
10. San Jose, CA (Population: 1,033,519)



I'm showing an extract of the data we worked with from the query on the Foursquare API.

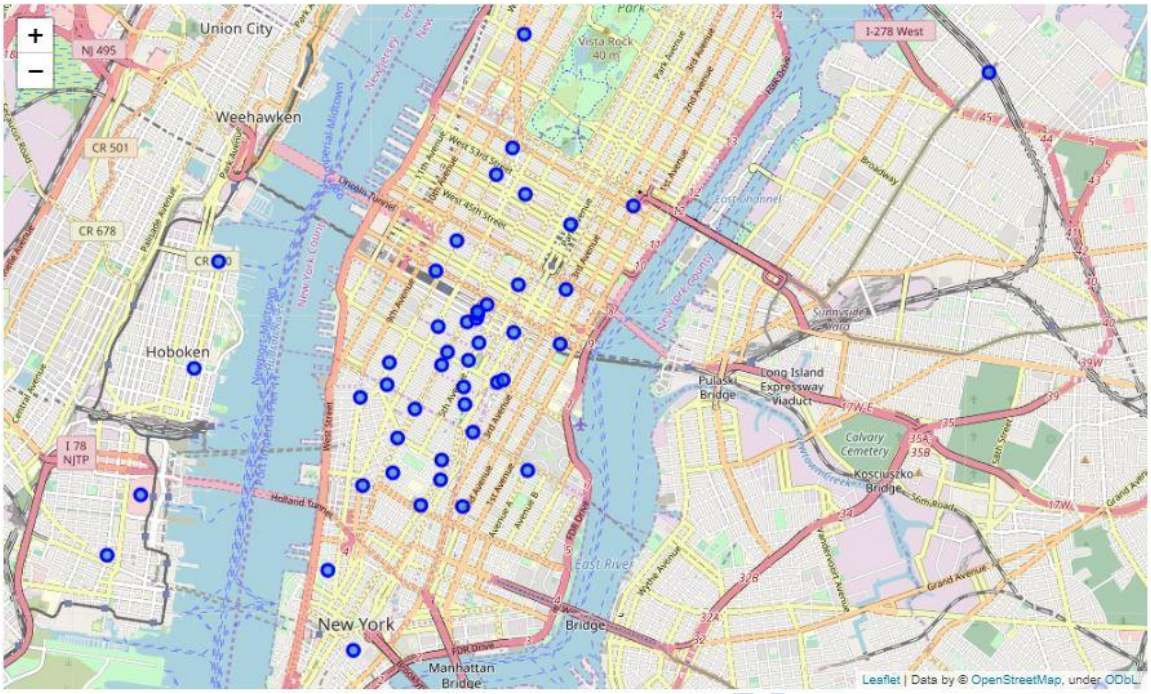
```
In [4]: results
        executed in 600ms, finished 14:21:22 2020-03-07

Out[4]: {'New York, NY': {'meta': {'code': 200,
            'requestId': '5e639ffa1a4b0a001c23aa78'},
           'response': {'venues': [{'id': '4aa44e08f964a5204f4620e3',
              'name': 'Chipotle Mexican Grill',
              'location': {'address': '350 5th Ave Lbby 6',
               'crossStreet': '34th St. between Fifth & Sixth Aves.',
               'lat': 40.748524,
               'lng': -73.9866615,
               'labeledLatLngs': [{'label': 'display',
                 'lat': 40.748524,
                 'lng': -73.9866615}],
               'postalCode': '10118',
               'cc': 'US',
               'city': 'New York',
               'state': 'NY',
               'country': 'United States',
               'formattedAddress': ['350 5th Ave Lbby 6 (34th St. between Fifth & Sixth Aves.)',
                'New York, NY 10118',
                'United States']},
              'categories': [{'id': '4bf58dd8d48988d1c1941735',
                'name': 'Mexican Restaurant',
                'pluralName': 'Mexican Restaurants',
                'shortName': 'Mexican',
                'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v2/food/mexican_',
                 'suffix': '.png'},
                'primary': True}],
              'referralId': 'v-1583587226',
              'hasPerk': False},
             {'id': '4a7a0832f964a5203ce81fe3',
              'name': 'Chipotle Mexican Grill',
              'location': {'address': '464 Park Ave S',
               'crossStreet': 'Park Ave between 31 & 32 Streets',
               'lat': 40.745464745574424,
```

3. Methodology

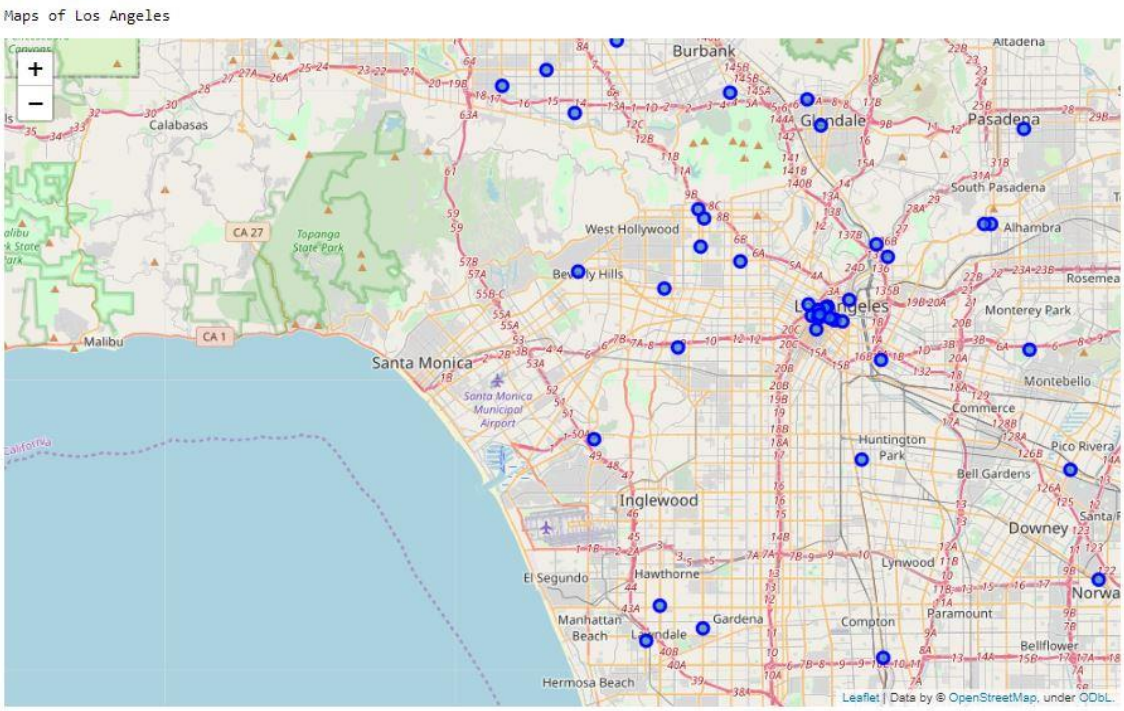
Our main objective is to stay at the hotel that has the best situation in terms of restaurants or Mexican food places depending on the city we are going to visit.

To do this we will analyze each of the maps of each city.

We begin by looking at the map of New York City. As you can see, the blue dots represent a Mexican restaurant. As you can also see, we are almost all very close and well centralized with respect to the center of the city.



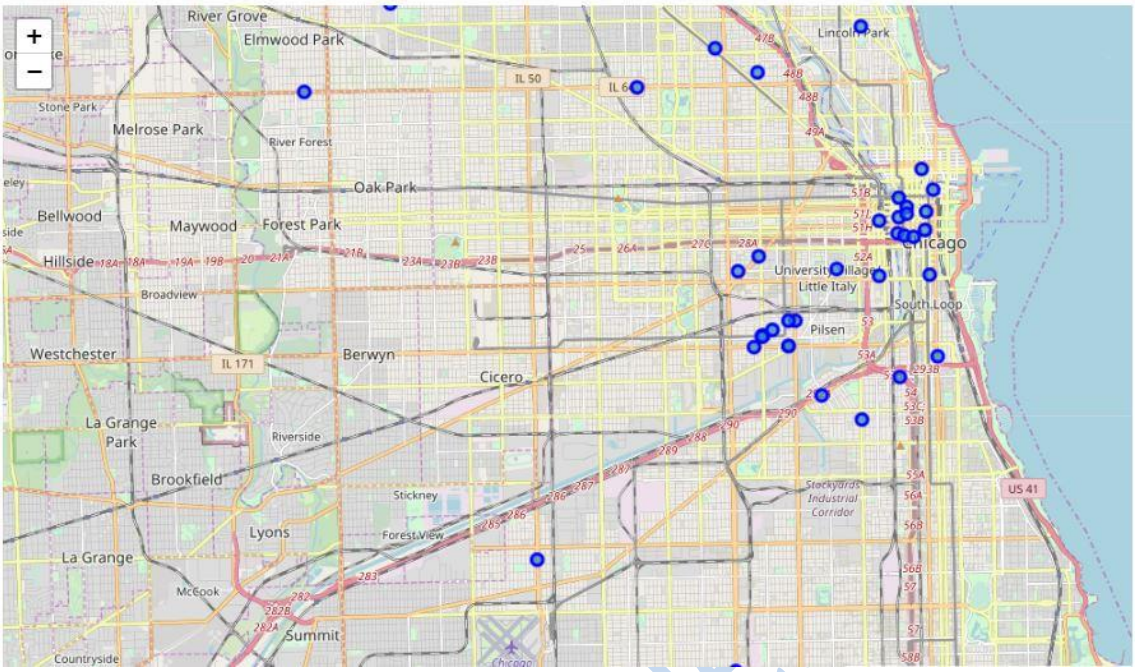
Let's now look at the map of the city of Los Angeles which is on the other coast of North America.



In this case, we observe that the restaurants or places of Mexican food, are much more dispersed that in the city of New York, we are a very small quantity near the center of Los Angeles.

We now look at the map of Chicago.

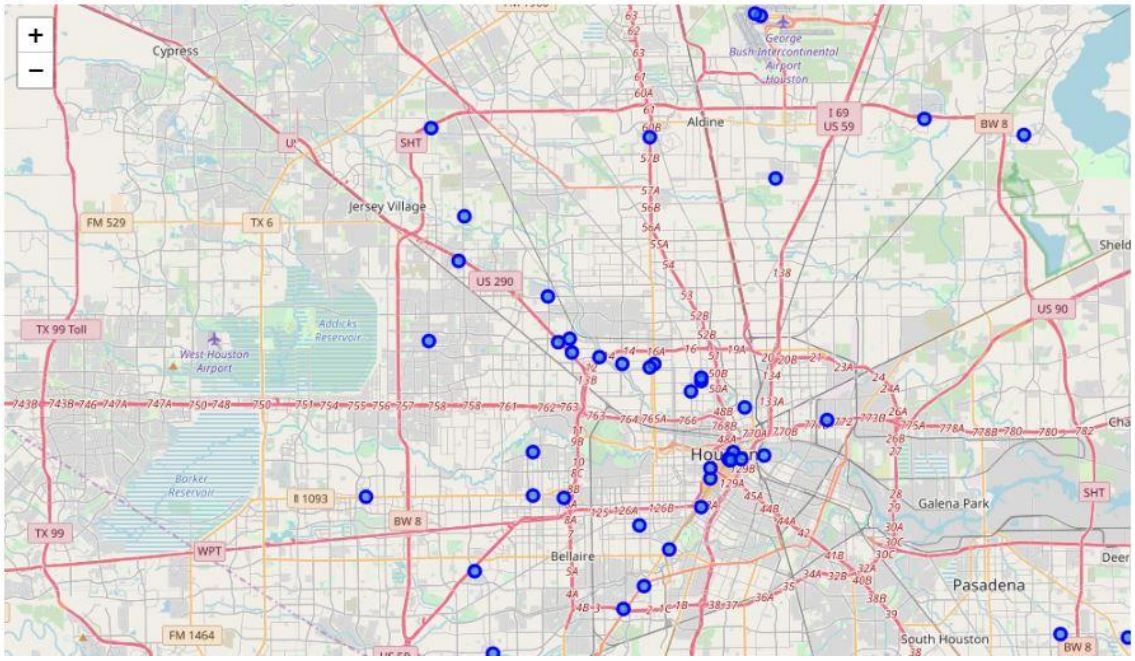
Maps of Chicago



In this case, there are far fewer restaurants and Mexican food places and they are focused mainly in the centre and in a slightly more distant neighbourhood.

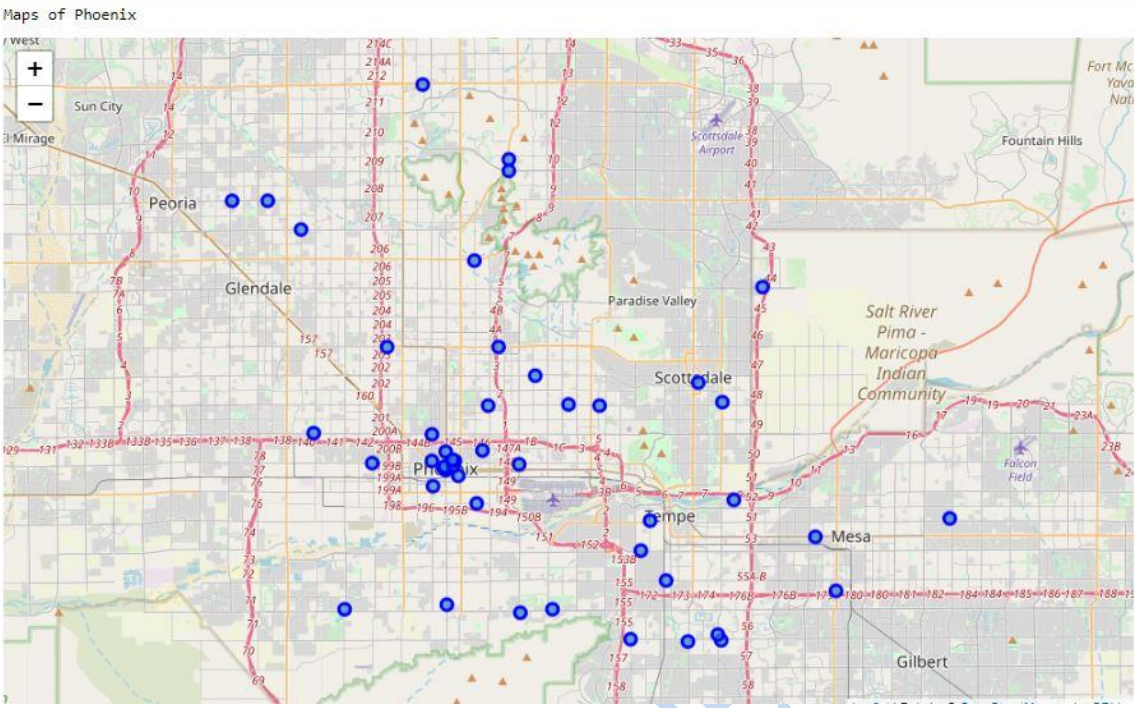
The next map we will see is that of Houston.

Maps of Houston



In this case, as in Los Angeles, there are many scattered and very little in the middle.

Finally we will see the map that we have left, the one of Phoenix



That just like in Los Angeles and Houston, we find a few restaurants or Mexican food places in the downtown area and the rest are more scattered.

The next thing we do now is use some basic statistics. We will take the average distance of the places from the mean coordinates. You will see how close the points found on the map are to the center.

This is a sample of the average distance of the places from the center of the cities.

```
executed in 450ms, finished 14:59:10 2020-03-07
New York, NY
Mean Distance from Mean coordinates
0.021533484982899625
Los Angeles, CA
Mean Distance from Mean coordinates
0.1054176392208922
Chicago,IL
Mean Distance from Mean coordinates
0.06078129479216396
Houston, TX
Mean Distance from Mean coordinates
0.11810189915281487
Phoenix, AZ
Mean Distance from Mean coordinates
0.10243862120978395
```

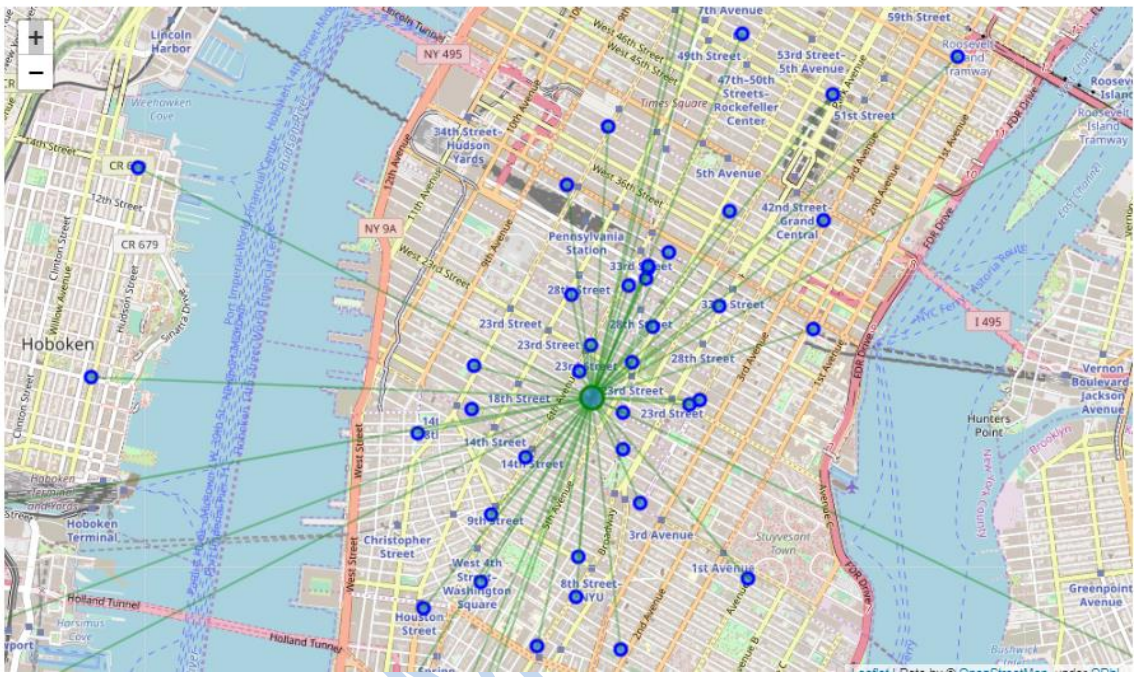
And we can see that the cities where the average distance is less are New York and Chicago. It should be noted that these averages vary depending on whether we find some point further

away from another, which makes the averages worse, but compared with the maps visually, the results coincide.

We see the maps again, seeing the proximity to the center that is represented in each city with the green point and bringing the maps closer.

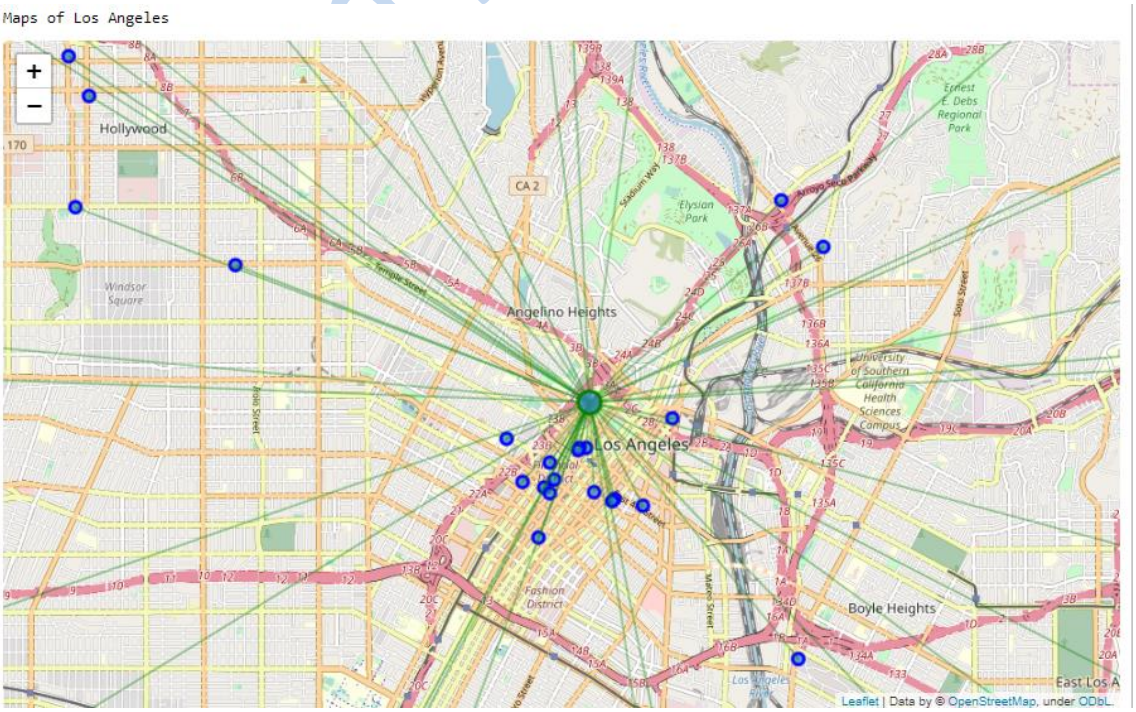
The first one we see is New York.

Maps of New York



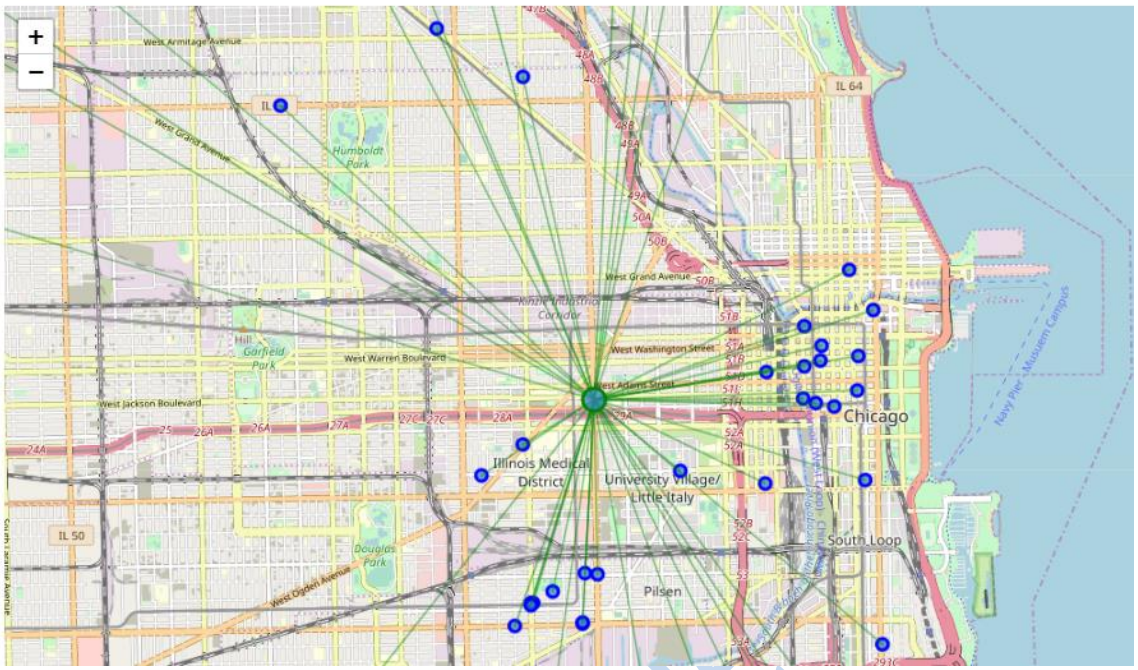
Now let's go look at the map of Los Angeles.

Maps of Los Angeles



Next we'll see the map of Chicago

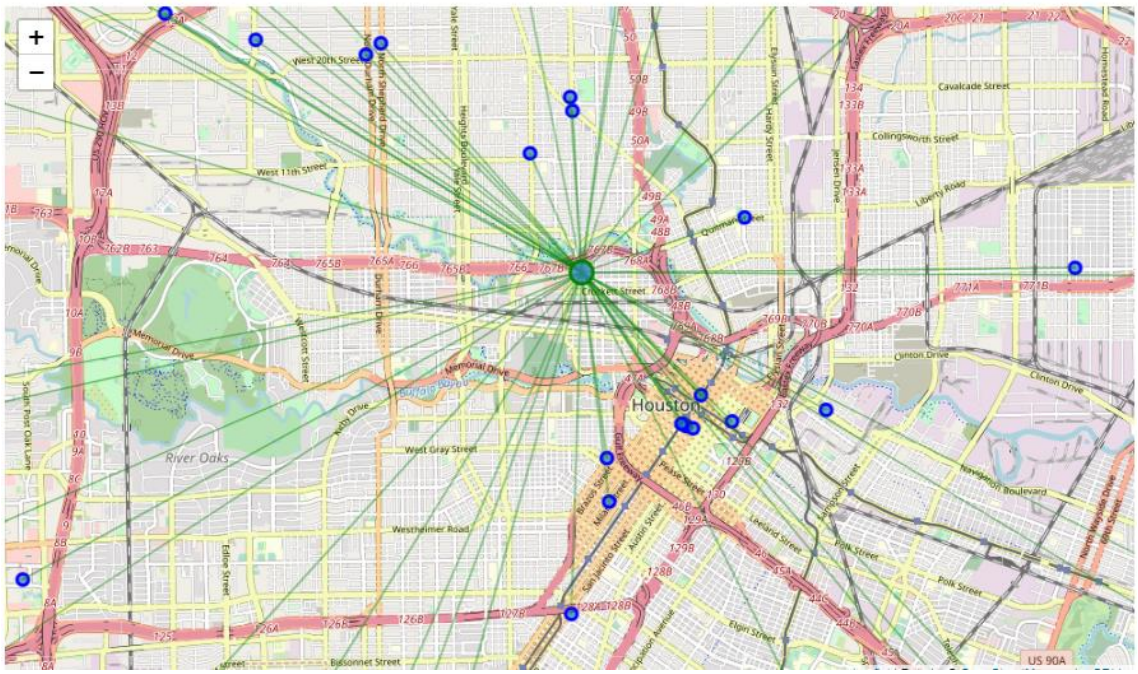
Maps of Chicago



In this specific case, it strikes me that the restaurants or Mexican food places are closer to the coast, and are also very close to each other, which indicates to me that there are possibly more inhabitants of Mexican nationality living in that area.

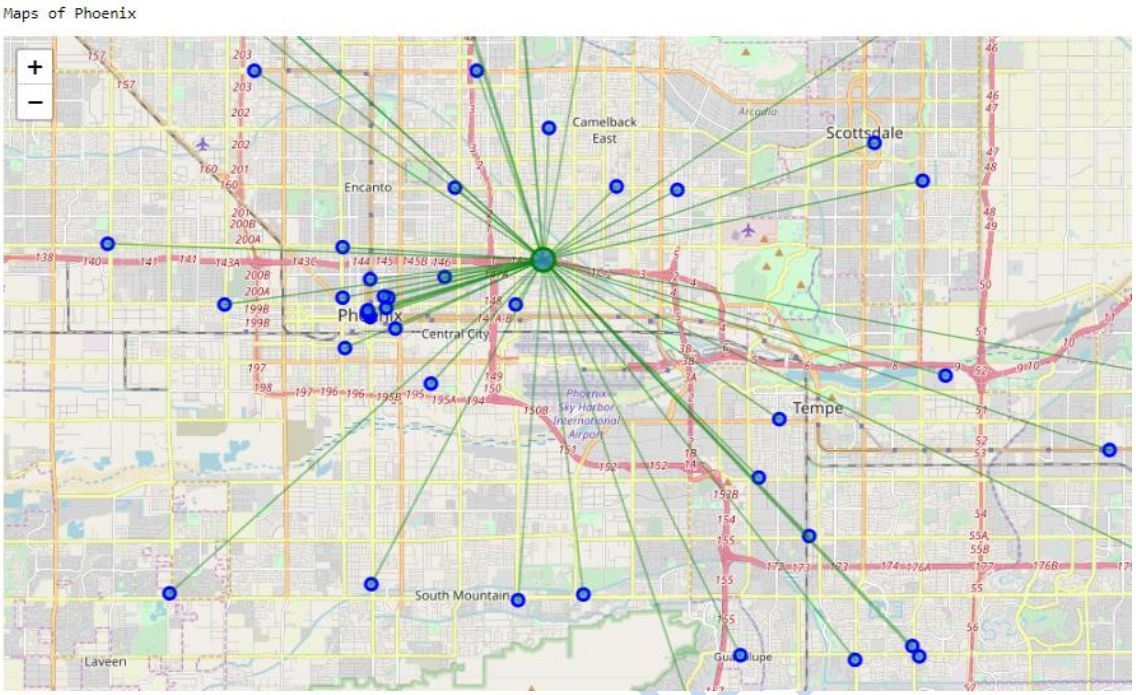
Let's now look at the map of Houston.

Maps of Houston



Here, as we mentioned before, they are further away from the centre.

Let's finally analyze the map of Phoenix.



Here we find them more scattered and an area farther away from the center in which there are a few together.

4. Cluster

Let's focus now on New York City in which we see the different categories we encounter, in our case 6 as seen in the attached image.

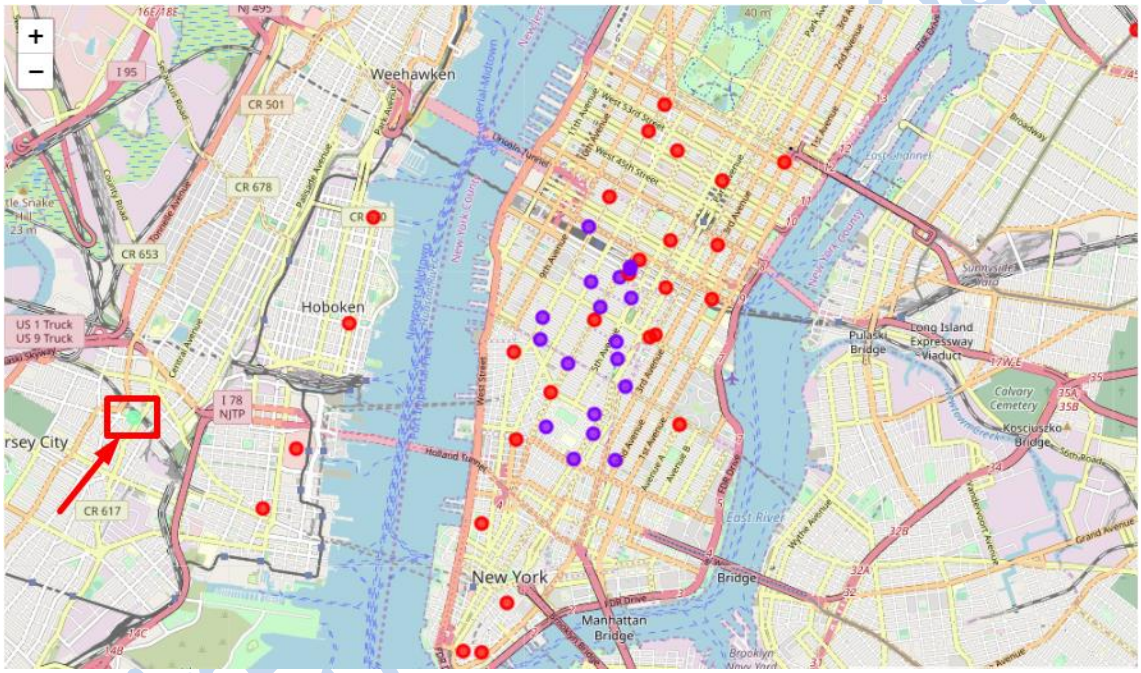
```
In [169]: newyork_venues['categories'].unique().tolist()
          executed in 26ms, finished 21:29:52 2020-03-08

Out[169]: ['Mexican Restaurant',
           'Burrito Place',
           'Food Truck',
           'Gift Shop',
           'Latin American Restaurant',
           'Food']
```

We look at how many guys we have in each zip code

Out[134]:

categories	Name	Address	Latitude	Longitude	PostalCode
Burrito Place	2	2	2	2	2
Food	2	2	2	2	2
Food Truck	5	4	5	5	5
Gift Shop	1	1	1	1	1
Latin American Restaurant	2	2	2	2	2
Mexican Restaurant	38	37	38	38	37

We made 3 clusters that can be seen on the map with the colors, red, purple and light green.



And in each of the clusters we have the following:

Cluster 1:

As you can see in the image below, it is ordered by what it considers to be the most popular and you can see that in second place the Latin American restaurants are the most abundant. They are represented on the map with the red dots.

Out[170]:

	Longitude	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
0	-73.986661	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
1	-73.982866	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
3	-73.999926	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
4	-73.985130	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
5	-74.005063	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
6	-73.975099	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
9	-73.993370	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
13	-73.980739	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
14	-73.974354	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
15	-73.982899	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
18	-73.985294	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
20	-74.006443	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
21	-73.982125	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
22	-74.012958	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
23	-74.010172	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
24	-74.042872	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
27	-74.005448	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
31	-73.981135	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
32	-73.984358	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
33	-73.988366	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
35	-73.965054	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
36	-73.975872	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
37	-74.026414	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place
40	-74.037869	Latin American Restaurant	Gift Shop	Food Truck	Food	Burrito Place

Cluster 2:

In this cluster, as you can see in the image, they come much less, and are more mixed and are represented on the map with the purple points.

1.7.1.2 Cluster 2

```
newyork_merged.loc[newyork_merged['Cluster Labels'] == 1, newyork_merged.columns[[1] + list(range(5, newyork_merged.shape[1]))]]
```

executed in 52ms, finished 21:05:58 2020-03-08

	Longitude	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
2	-73.988824	Gift Shop	Food Truck	Burrito Place	Latin American Restaurant	Food
7	-73.993437	Gift Shop	Food Truck	Burrito Place	Latin American Restaurant	Food
8	-73.990094	Gift Shop	Food Truck	Burrito Place	Latin American Restaurant	Food
10	-73.996533	Burrito Place	Latin American Restaurant	Gift Shop	Food Truck	Food
11	-73.987871	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
12	-73.993970	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
16	-73.992498	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
17	-73.990074	Gift Shop	Food Truck	Burrito Place	Latin American Restaurant	Food
19	-73.993603	Gift Shop	Food Truck	Burrito Place	Latin American Restaurant	Food
25	-74.001218	Mexican Restaurant	Latin American Restaurant	Gift Shop	Food Truck	Burrito Place
26	-73.990269	Gift Shop	Food Truck	Burrito Place	Latin American Restaurant	Food
28	-73.989661	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
29	-74.000698	Burrito Place	Latin American Restaurant	Gift Shop	Food Truck	Food
34	-73.988165	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
38	-73.988171	Mexican Restaurant	Latin American Restaurant	Gift Shop	Food	Burrito Place
39	-73.994328	Food Truck	Latin American Restaurant	Gift Shop	Food	Burrito Place
45	-73.997381	Mexican Restaurant	Latin American Restaurant	Gift Shop	Food Truck	Burrito Place
49	-74.001430	Mexican Restaurant	Latin American Restaurant	Gift Shop	Food Truck	Burrito Place

Cluster 3:

As you can see in the image, we only have one represented on the map with the green dot.

1.7.1.3 Cluster 3

In [166]:
```
newyork_merged.loc[newyork_merged['Cluster Labels'] == 2, newyork_merged.columns[[1] + list(range(5, newyork_merged.shape[1]))]]
```

executed in 40ms, finished 21:08:47 2020-03-08

Out[166]:

	Longitude	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
44	-74.06213	Mexican Restaurant	Gift Shop	Food Truck	Food	Burrito Place

5. Conclusions

In this work, we are working with little data, since from each city we have obtained an approximate sample of about 50 data. Apart from the fact that the topic chosen is Mexican restaurants or places to eat, which is less typical than usual, thus reducing the sample. Anyway, looking at the maps we can extract enough information to advise tourists where to stay since we can clearly see the neighborhoods where we can find more variety of restaurants or places of that particular type of food.