

The Battle of Neighborhoods

WEEK 1

Presentation of the Problem

Initial Presentation

The Battle of Neighborhoods
Week 1

We can suppose that we are thinking about opening a exotic restaurant in Manhattan, more specifically a Japanese or Asiatic one. This could be a good business opportunity but we need to carry out a market research in order to establish a long-term success.

To start with, we will analyse the existing restaurants of this category in Manhattan. And we will sort them by neighborhood, in order to identify the best possible location.

At the end, we will identify, based on a clustering, the best possible location in Manhattan.

We take data from all Boroughs

Presentation of the Problem

Analysis of the New York Perspective and concentrate the focus on Manhattan

```
We create a Dataframe with the downloaded data
```

```
[44] > MI
neighborhoods_data = newyork_data['features']
# We define the columns of the dataframe
column_names = ['Borough', 'Neighborhood', 'Latitude', 'Longitude']

# We start and initialize the dataframe
neighborhoods = pd.DataFrame(columns=column_names)

for data in neighborhoods_data:
    borough = neighborhood_name = data['properties']['borough']
    neighborhood_name = data['properties']['name']

    neighborhood_latlon = data['geometry']['coordinates']
    neighborhood_lat = neighborhood_latlon[1]
    neighborhood_lon = neighborhood_latlon[0]

    neighborhoods = neighborhoods.append({'Borough': borough, 'Neighborhood': neighborhood_name,
                                         'Latitude': neighborhood_lat, 'Longitude': neighborhood_lon}, ignore_index=True)

# We show the Dataframe
neighborhoods.head()
```

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

And then concentrate on Manhattan

```
We analyze of the neighborhoods in Manhattan
```

```
[46] > MI
manhattan_data = neighborhoods[neighborhoods['Borough'] == 'Manhattan'].reset_index(drop=True)
manhattan_data.head()
```

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

The Battle of Neighborhoods
Week 1

We take data from Foursquare about all the Venues in Manhattan, filtering by Category

```
We take the venues in Manhattan of Japanese and Sushi Restaurants

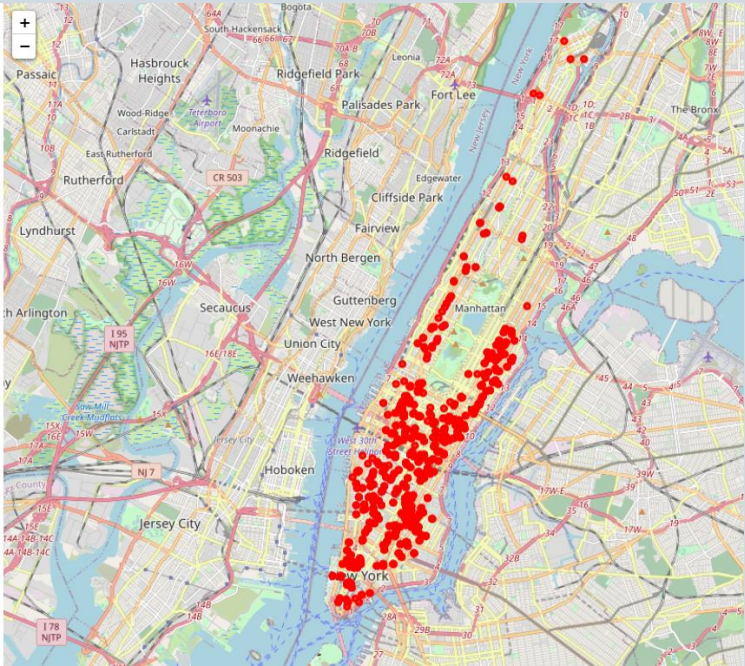
[50] > MI
neighborhoods = neighborhoods[neighborhoods['Borough'] == 'Manhattan'].reset_index(drop=True)
newyork_venues_sushi = getNearbyVenues(names=neighborhoods['Neighborhood'], latitudes=neighborhoods['Latitude'], longitudes=neighborhoods['Longitude'], radius=1000, categoryIds='4bf58dd8d48988d1d2941735')
newyork_venues_sushi.head()

Neighborhood Neighborhood Latitude Neighborhood Longitude Venue Venue Latitude Venue Longitude Venue Category
0 Harlem Hill 40.876551 -73.910660 Planet Tokyo 40.886233 -73.909479 Sushi Restaurant
1 Chinatown 40.715618 -73.994279 Nakeji 40.715912 -73.996597 Sushi Restaurant
2 Chinatown 40.715618 -73.994279 Shinsen 40.715608 -73.996611 Japanese Restaurant
3 Chinatown 40.715618 -73.994279 Sushumi Asian Fusion 40.721155 -73.987337 Sushi Restaurant
4 Chinatown 40.715618 -73.994279 Bondi Bar 40.721247 -73.996264 Sushi Restaurant

[51] > MI
print("To answer the question How many Sushi Restaurants are there in Manhattan?")
newyork_venues_sushi.shape

To answer the question How many Sushi Restaurants are there in Manhattan?
(1100, 7)
```

And show them in a map



Presentation of the Problem

Analysis of the Venues in Manhattan

We analyze the different neighborhoods in Manhattan by frequencies, according to the venues

Presentation of the Problem

Analysis of the Venues in Manhattan

	Neighborhood	Asian Restaurant	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Fish Market	Grocery Store	Hawaiian Restaurant	Indian Chinese Restaurant	Japanese Restaurant	Noodle House	R
0	Battery Park City	0.000000	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.090909	0.045455	0.00
1	Carnegie Hill	0.041667	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.041667	0.125000	0.000000	0.00
2	Central Harlem	0.000000	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00
3	Chelsea	0.045455	0.000000	0.000000	0.000000	0.000	0.022727	0.000000	0.000000	0.000000	0.113636	0.000000	0.00
4	Chinatown	0.000000	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.125000	0.000000	0.00
5	Civic Center	0.000000	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.062500	0.031250	0.00

And the most common venues

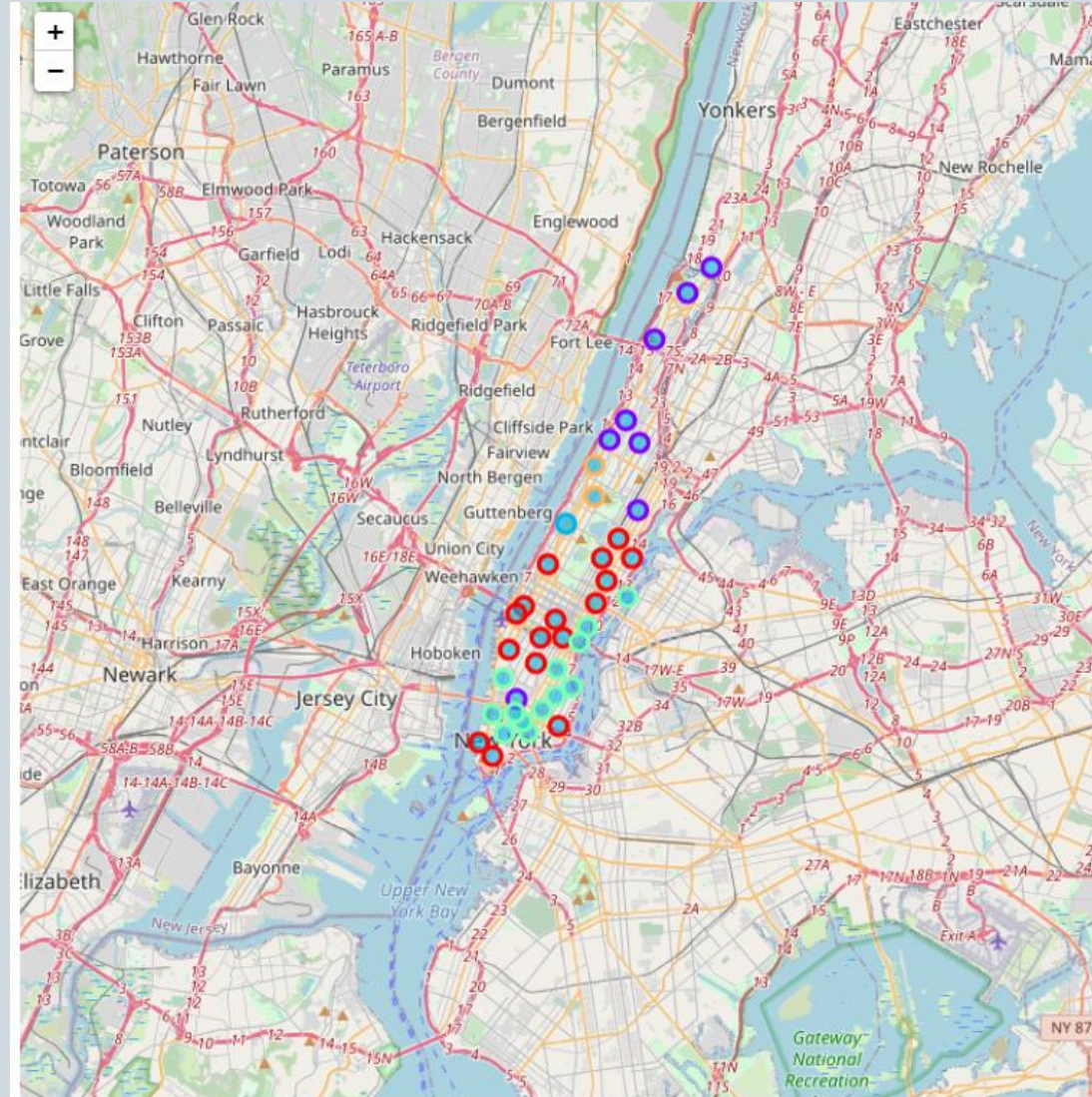
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
0	Battery Park City	Sushi Restaurant	Japanese Restaurant	Noodle House	Theme Restaurant	Indian Chinese Restaurant	Bakery	Chinese Restaurant
1	Carnegie Hill	Sushi Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Indian Chinese Restaurant	Asian Restaurant	Seafood Restaurant	Sandwich Place
2	Central Harlem	Sushi Restaurant	Vegetarian / Vegan Restaurant	Japanese Restaurant	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega
3	Chelsea	Sushi Restaurant	Japanese Restaurant	Asian Restaurant	Fish Market	Vegetarian / Vegan Restaurant	Seafood Restaurant	Sandwich Place
4	Chinatown	Sushi Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega

We cluster them into 5 groups and present it in a map with different colours

Presentation of the Problem

Clustering the Venues

The Battle of Neighborhoods
Week 1



We find the best zones to open our restaurant and highlight them in red

Presentation of the Problem

Finding the best location

The Battle of Neighborhoods
Week 1

