Advice Engine for building a Business in a Chosen Area

Mahdi Arfaoui

Wise choosing of B2C business place is important

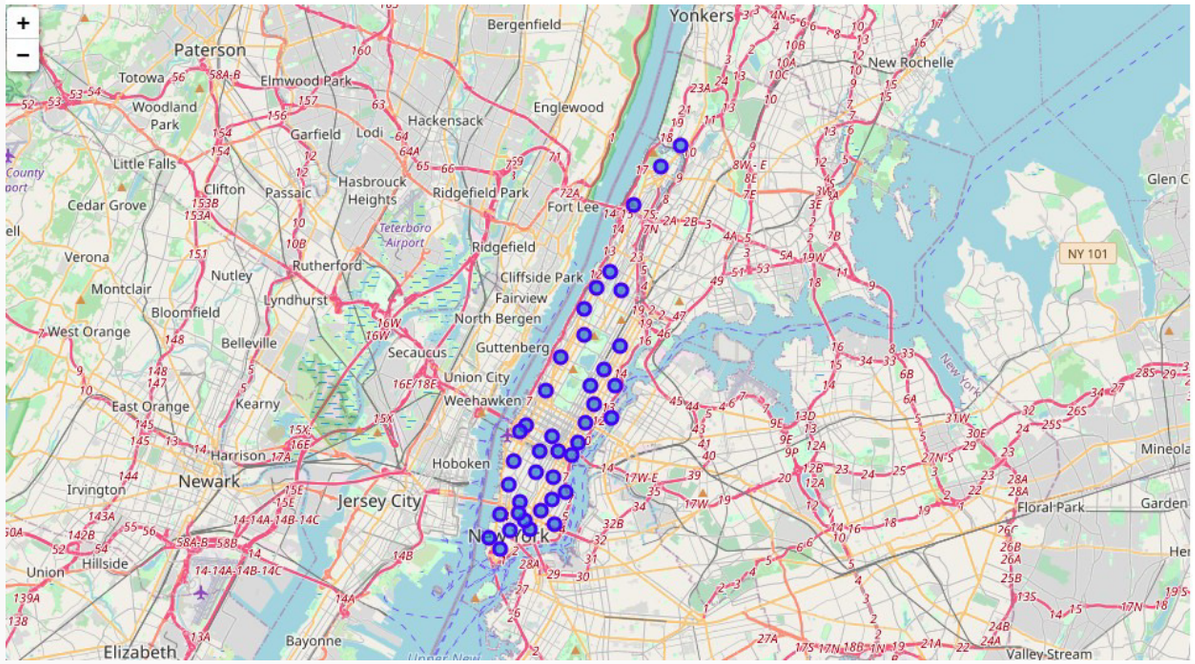
- If you ever thought of building a project, one of the most important pillar you would think of is where the project is going to be built.
- That’s why giving an important amount of time and concentration to choose a place, is a wise thing to do.
- To find such a place you should put a lot of factor in consideration (nature of client , nature of product, supplying …)

Data acquisition

- After receiving data we do some data cleaning and structuring to obtain a clean data frame

	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

	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



- Choosing the city we'll work on and show it's neighborhoods on the map.

Out[34]:

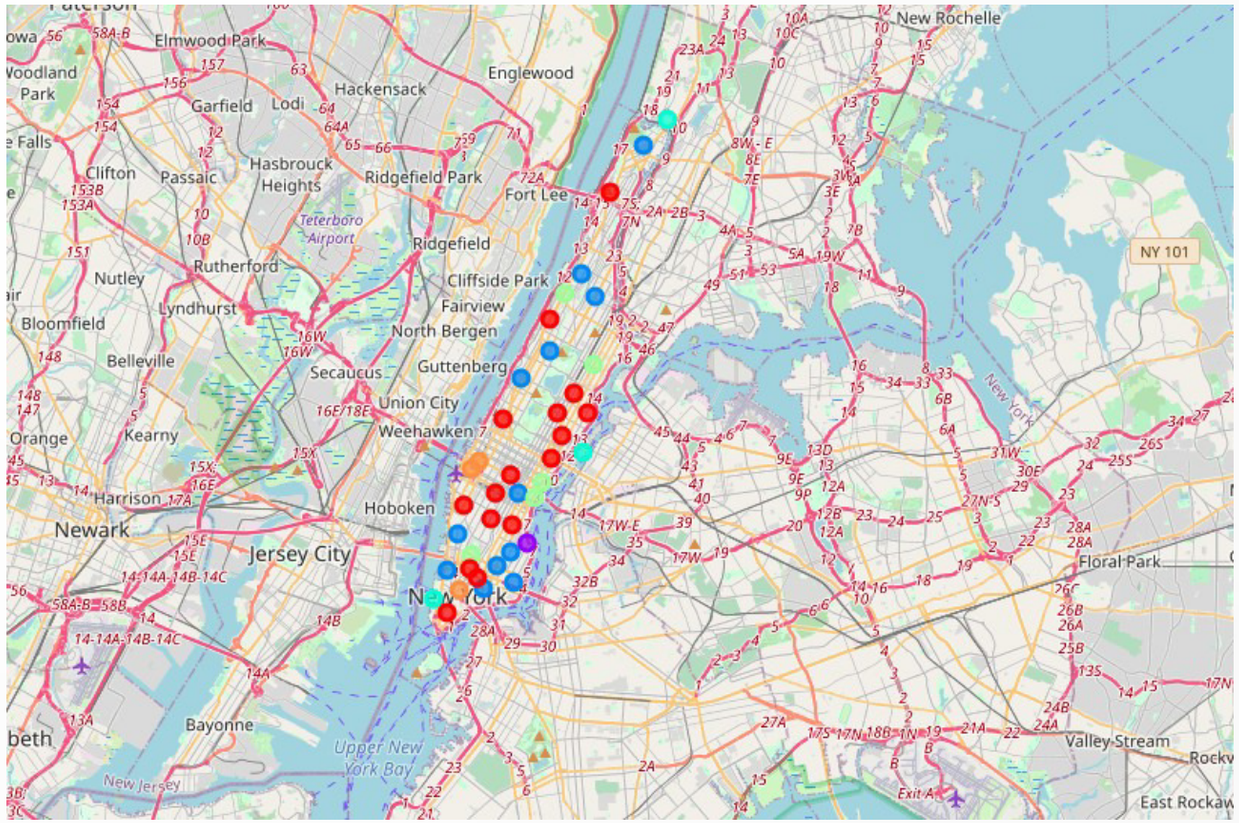
	index	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	9	Marble Hill	40.876551	-73.910660	Land & Sea Restaurant	40.877885	-73.905873	Restaurant
1	24	Chinatown	40.715618	-73.994279	Spicy Village	40.717010	-73.993530	Restaurant
2	27	Chinatown	40.715618	-73.994279	Kiki's	40.714476	-73.992036	Restaurant
3	29	Chinatown	40.715618	-73.994279	Wah Fung Number 1 Fast Food 華豐快餐店	40.717278	-73.994177	Restaurant
4	32	Chinatown	40.715618	-73.994279	Da Yu Hot Pot 大渝火锅	40.716735	-73.995752	Restaurant
5	35	Chinatown	40.715618	-73.994279	Xi'an Famous Foods	40.715232	-73.997263	Restaurant
6	39	Chinatown	40.715618	-73.994279	Forgtmenot	40.714459	-73.991546	Restaurant
7	41	Chinatown	40.715618	-73.994279	Dimes	40.714830	-73.991719	Restaurant
8	44	Chinatown	40.715618	-73.994279	Ling Kee Malaysian Beef Jerky	40.714713	-73.991538	Restaurant
9	47	Chinatown	40.715618	-73.994279	Cervo's	40.714763	-73.991455	Restaurant

```
Entrée [78]: manhattan_reduced['Venue Category'].unique()
Out[78]: array(['Restaurant', 'Bar', 'Hotel', 'Café', 'Gym', 'Shop'], dtype=object)
```

- Using Foursquare Api, we extract the venues of each neighborhood along with their locations, we then customize their category to be convenient with our list of categories we'll work with.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	3	Shop	Gym	Restaurant	Hotel	Café	Bar
1	Manhattan	Chinatown	40.715618	-73.994279	2	Restaurant	Shop	Bar	Hotel	Gym	Café
2	Manhattan	Washington Heights	40.851903	-73.936900	0	Restaurant	Shop	Café	Gym	Bar	Hotel
3	Manhattan	Inwood	40.867684	-73.921210	2	Restaurant	Shop	Bar	Café	Hotel	Gym
4	Manhattan	Hamilton Heights	40.823604	-73.949688	2	Restaurant	Shop	Bar	Café	Hotel	Gym

- We organize the data with top visited venues, and here's a simple classification with KNN where we could easily compare it to our data to extract some insight, for example we could relate red dots (K = 2) to the places there is more restaurants then the other categories.



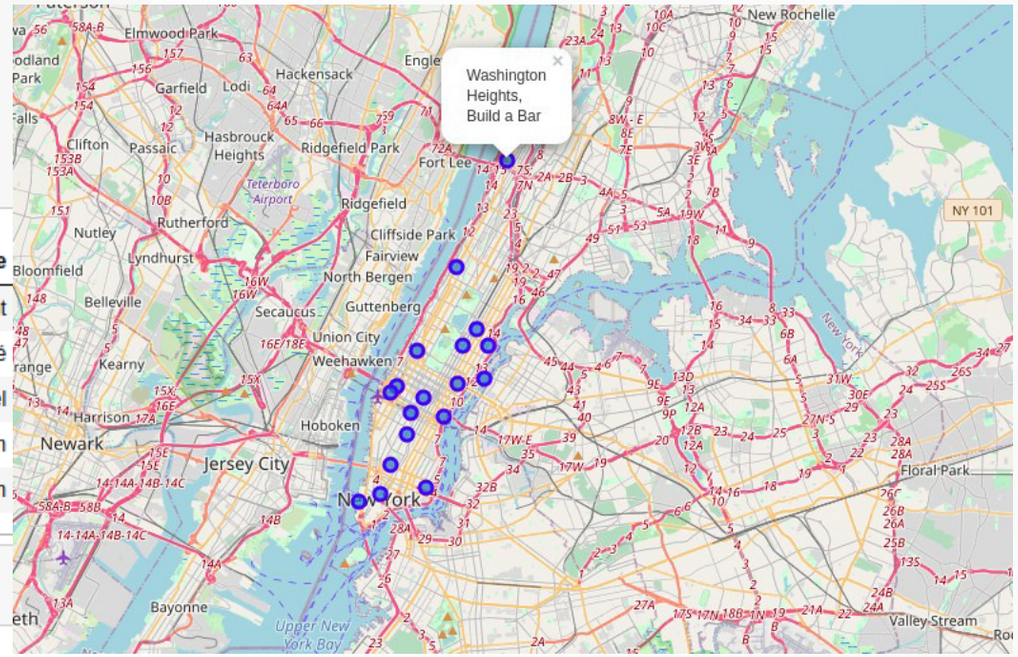
```
#Let's polt the places where the licent want to build a Bar
input = "Bar"

target_df = adviceMe(input)
target_df.head()
```

	Neighborhood	Latitude	Longitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Marble Hill	40.876551	-73.910660	Shop	Gym	Restaurant
2	Washington Heights	40.851903	-73.936900	Restaurant	Shop	Café
8	Upper East Side	40.775639	-73.960508	Restaurant	Shop	Hotel
9	Yorkville	40.775930	-73.947118	Restaurant	Shop	Gym
11	Roosevelt Island	40.762160	-73.949168	Shop	Restaurant	Gym

```
target_df.shape
```

```
(19, 6)
```



- Here we can see that we choose best places where a client could build a bar, by eliminating all the places that have a bar in their top 3 visited venues.