

# **Project Report – Opening of a new Bar**

Demetris Englezos

## **1. Introduction/Business Problem:**

### **1.1 Background:**

The owners of a Sports bar chain let's call it "Sports Bars" for the sake of this project want to expand their business by opening two (2) more bars. They already have one sports bar in Queens which is successful and that is why is expanding. In all the bars of the "Sports Bars" chain, they broadcast basketball and baseball. With the current situation with covid-19, the games are played behind closed doors without fans. All these fans are now going to bars to watch their favorite team and the "Sports Bar" chain in order to take advantage of this, they are opening two more bars. The stakeholders of the Bars, want to open the first bar around Manhattan and Bronx to serve the New York Knicks fans (based in Manhattan) and the New York Yankees (based in Bronx). The second bar is going to be around Brooklyn to serve Brooklyn Nets fans (based in Brooklyn). Also after the opening of the bars, they want to find out what are the busy days of the week to have more staff working these days. In order to find the best location to open these two bars in the two areas and to find out which days to have more staff, they hire a Data Scientist or a Data Analyst.

### **1.2 Problem**

As mentioned in the Background section, the problem is to find the best two locations around Manhattan-Bronx and Brooklyn for the two new sports bars. Also which days are expecting more people to overstaff their bars. This means that target audience of the project are the owners of the "Sports Bar" chain alongside with the stakeholders. They will review the project and the presentation and decide if they proceed with the expansion (location and staff schedule) that the Data Scientist propose.

## **2. Data**

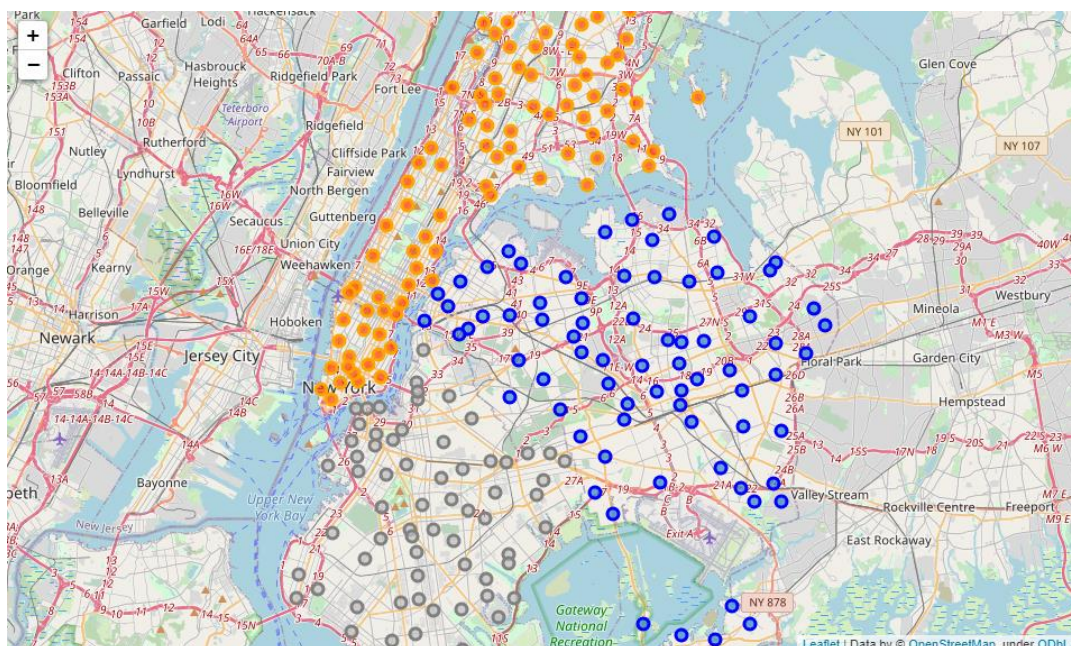
Data that are going to use for the purpose of finding the best locations, are first data from Foursquare with the venues of New York. These are going to be used to see all the location of all bars (sport's bar also attract not sport fans) and check which locations can potentially be suitable. For example a good location might be somewhere with no other sports bars close, or a populated area that might have another sports bar but is always crowded. Also foursquare can be used to check where there are a lot of other venues (like restaurants, theaters etc.) which means that people are hanging in this area and they might look for a bar after their visit in a different venue. In addition, datasets with the game dates of the four teams mentioned above (Knicks/Yankees for Manhattan/Bronx and Nets or Brooklyn). This is because we expect more people visiting the bars to watch their team. The schedule of the above teams was found from [www.basketball-reference.com](http://www.basketball-reference.com).

### 3. Data Analysis

First we have the Data for the Borough and Neighborhoods of New York which we are splitting to areas to study each Borough. We can see from the Dataset that we have 81 Neighborhoods in Queens where the existing bar is in one of them, 82 in Manhattan-Bronx for the first new bar and 70 in Brooklyn for the second bar.

	Neighborhood	Latitude	Longitude
Borough			
Bronx	52	52	52
Brooklyn	70	70	70
Manhattan	40	40	40
Queens	81	81	81
Staten Island	63	63	63

After splitting the Dataset to each of the Boroughs, we visualize the Neighborhoods of each of the Boroughs we are interested in a Map. With Orange color are the Neighborhoods in Manhattan-Bronx, with Grey in Brooklyn and finally Blue in Queens.



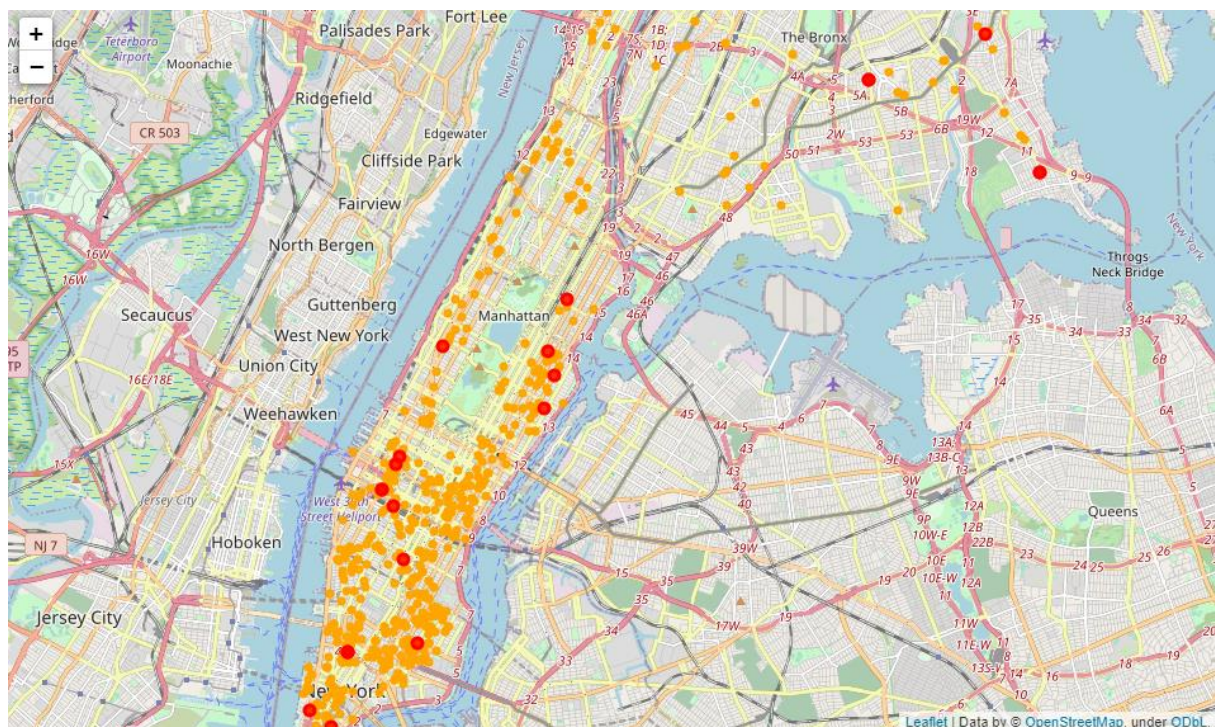
In these neighborhoods now, we have to get the existing venues from Foursquare and after analyze these data, get the Bars of each neighborhood. We started the analysis for Manhattan-Bronx and then continue with Brooklyn.

### 3.1 Manhattan-Bronx analysis

The next picture is a small part of a Dataset with all the bars (sport bars or not) for these Boroughs.

	Neighborhood	Venue	Venue Category	Venue Latitude	Venue Longitude	Borough
0	Co-op City	Juice Bar	Juice Bar	40.876156	-73.829506	Manhattan/Bronx
1	Co-op City	The City Lounge	Bar	40.876755	-73.833405	Manhattan/Bronx
2	Eastchester	Gardens Bar & Grill	Bar	40.892959	-73.831187	Manhattan/Bronx
3	Riverdale	Blackstone Bar & Grill	Bar	40.887253	-73.907251	Manhattan/Bronx
4	Riverdale	Fenwick's Bar & Grill	Bar	40.887170	-73.904526	Manhattan/Bronx
5	Kingsbridge	100% Juice Bar Cafe	Juice Bar	40.883705	-73.897896	Manhattan/Bronx
6	Kingsbridge	Madden's Uptown Bar & Grill	Pub	40.883327	-73.901777	Manhattan/Bronx
7	Kingsbridge	Harpurs Bar	Bar	40.878500	-73.903331	Manhattan/Bronx
8	Kingsbridge	Triffiffe Bar	Bar	40.878915	-73.906074	Manhattan/Bronx
9	Kingsbridge	Blackstone Bar & Grill	Bar	40.887253	-73.907251	Manhattan/Bronx
10	Kingsbridge	Rowes Wharf Bar	Bar	40.883200	-73.910000	Manhattan/Bronx
11	Kingsbridge	JC Mac's	Bar	40.884605	-73.899895	Manhattan/Bronx
12	Marble Hill	shooter's bar and grille	Pub	40.878581	-73.914491	Manhattan/Bronx
13	Marble Hill	Triffiffe Bar	Bar	40.878915	-73.906074	Manhattan/Bronx
14	Woodlawn	Lark's Nest Bar	Bar	40.897960	-73.866995	Manhattan/Bronx

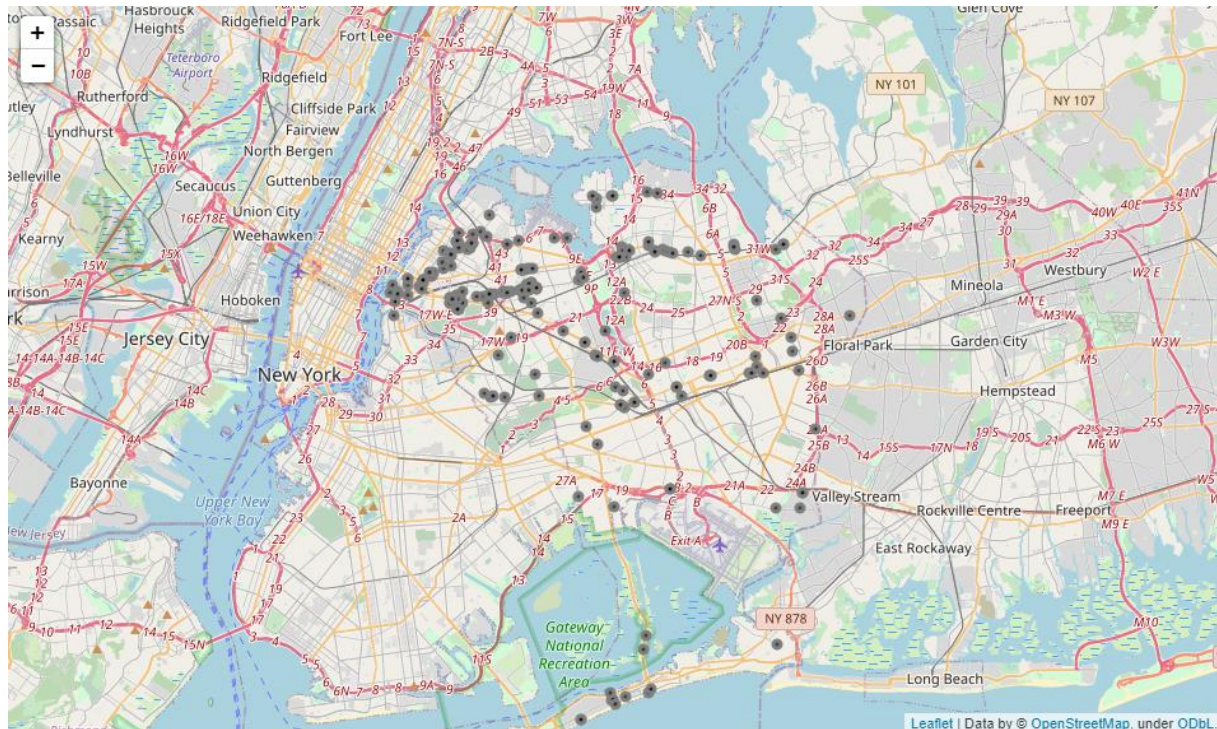
To have a better idea of these bars and their locations, first we splitted the sports bars from the rest bars and then visualize the results. With orange in the map are all the existing bars in Manhattan-Bronx and with red color are the Sports Bars.





We can see from the above map that compare to Bronx, more bars are in the area of Manhattan which is the center of New York. This can potentially means that is a better place for the first bar compare to Bronx since more people are hanging around Manhattan with more venues or bars to visit.

Now that we got the first idea of the situation in New York, in order to find a good location, we have to compare the neighborhood of the existing bar in Queens to find a potentially good location for the bar in Manhattan-Bronx. To do that, we are using Foursquare again to get some data for venues in Queens. In the next map, the venues with the category bar in Queens are presented.



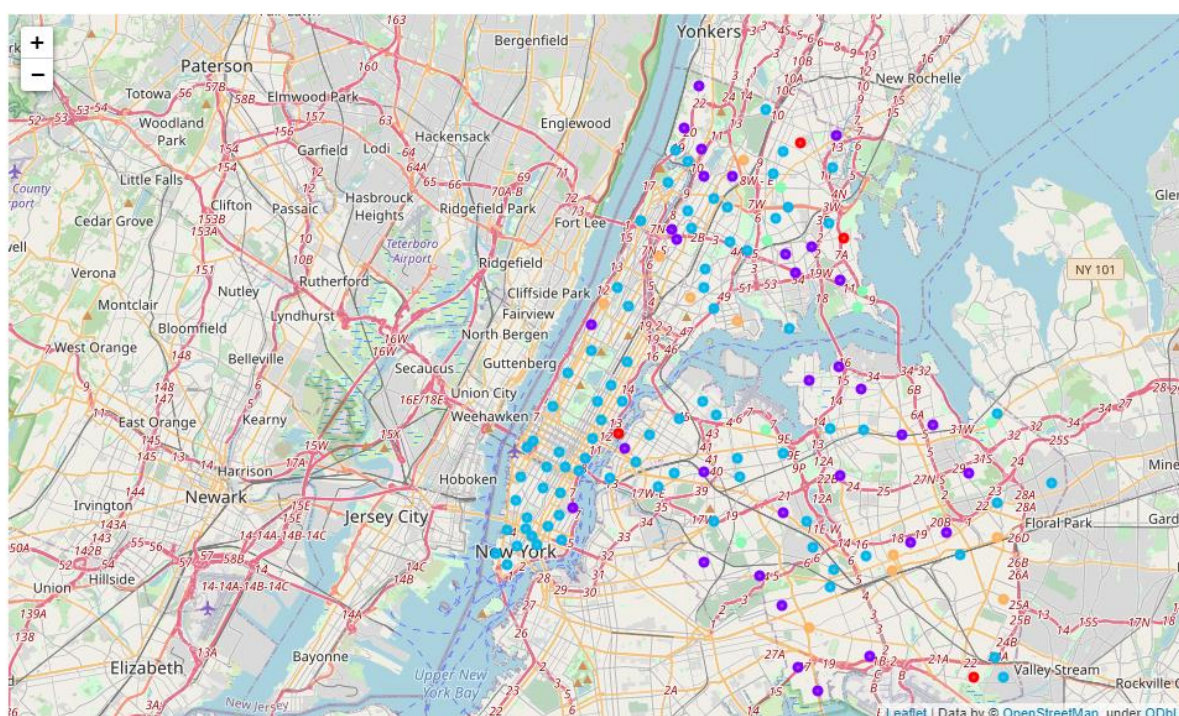
Before continuing with the Machine Learning technique, we also found the most common venues with the category bar in each Neighborhood of Manhattan-Bronx and Queens. These are shown in the next Dataset sample.

	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Allerton	Sports Bar	Cocktail Bar	Wine Bar	Sake Bar	Pub	Lounge	Karaoke Bar	Juice Bar	Hotel Bar	Hookah Bar
1	Astoria	Bar	Sports Bar	Juice Bar	Hookah Bar	Dive Bar	Pub	Lounge	Wine Bar	Sake Bar	Karaoke Bar
2	Astoria Heights	Hotel Bar	Wine Bar	Sports Bar	Sake Bar	Pub	Lounge	Karaoke Bar	Juice Bar	Hookah Bar	Gay Bar
3	Auburndale	Bar	Wine Bar	Sports Bar	Sake Bar	Pub	Lounge	Karaoke Bar	Juice Bar	Hotel Bar	Hookah Bar
4	Battery Park City	Bar	Juice Bar	Hotel Bar	Cocktail Bar	Wine Bar	Sports Bar	Pub	Lounge	Gastropub	Sake Bar

### 3.2 Manhattan-Bronx neighborhoods Clustering

Now that we have data for the bars in Queens too, we have to use a Machine Learning technique to cluster the neighborhoods of Manhattan-Bronx and Queens to find similar neighborhoods in Manhattan-Bronx with the neighborhood of the existing Sports Bar in Queens. To cluster the neighborhoods, we used the KMeans machine learning technique and we split them in 5 Clusters. Next there is a Dataset with the column 'Cluster' and a map with the clustered neighborhoods.

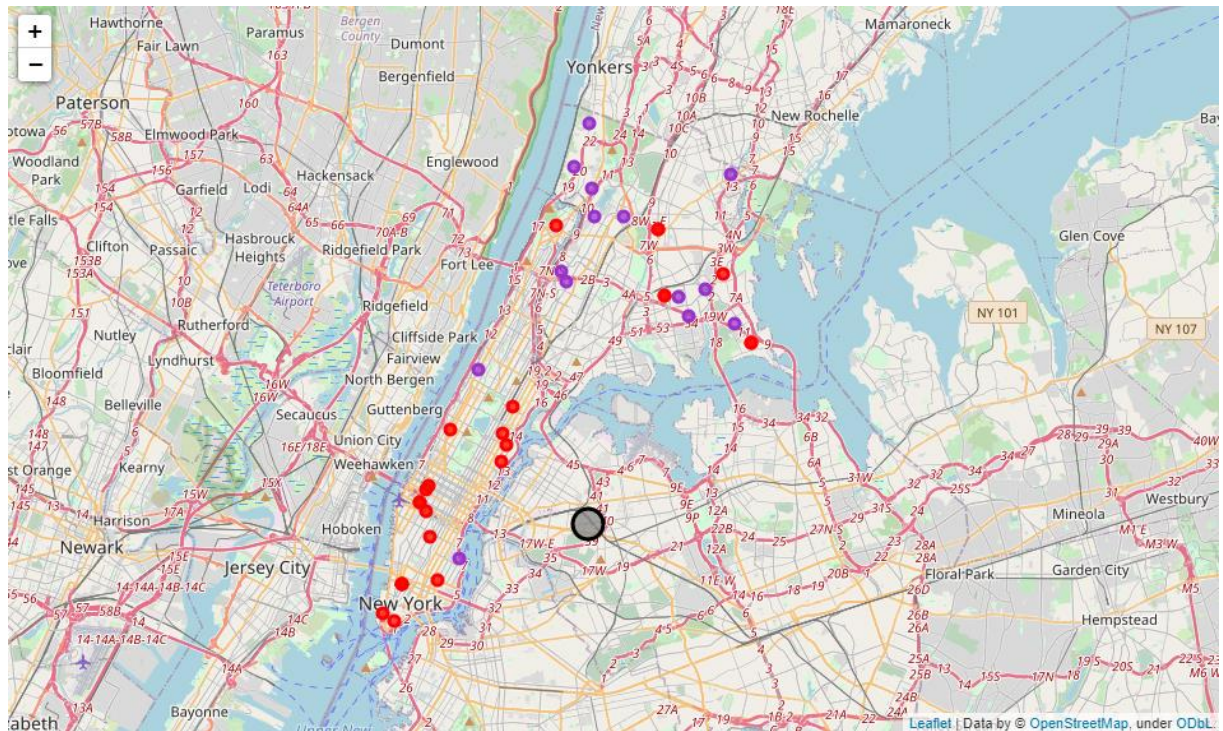
	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Bronx	Co-op City	40.874294	-73.829939	2	Juice Bar	Bar	Wine Bar	Sports Bar	Sake Bar
2	Bronx	Eastchester	40.887556	-73.827806	1	Bar	Wine Bar	Sports Bar	Sake Bar	Pub
4	Bronx	Riverdale	40.890834	-73.912585	1	Bar	Wine Bar	Sports Bar	Sake Bar	Pub
5	Bronx	Kingsbridge	40.881687	-73.902818	1	Bar	Pub	Juice Bar	Wine Bar	Sports Bar
6	Manhattan	Marble Hill	40.876551	-73.910660	2	Pub	Bar	Wine Bar	Sports Bar	Sake Bar
7	Bronx	Woodlawn	40.898273	-73.867315	2	Pub	Bar	Wine Bar	Sports Bar	Sake Bar
8	Bronx	Norwood	40.877224	-73.879391	4	Juice Bar	Bar	Wine Bar	Sports Bar	Sake Bar
9	Bronx	Williamsbridge	40.881039	-73.857446	2	Dive Bar	Bar	Wine Bar	Sports Bar	Sake Bar
11	Bronx	Pelham Parkway	40.857413	-73.854756	2	Cocktail Bar	Wine Bar	Sports Bar	Sake Bar	Pub
13	Bronx	Bedford Park	40.870185	-73.885512	1	Bar	Wine Bar	Sports Bar	Sake Bar	Pub



The five clusters are with the colors turquoise, purple, red, orange and light blue.



The next maps shows the location of the existing sport bar of the 'Sports Bar' chain in Queens in a black circle marker. This bar belongs to Cluster 1 which is the purple in the previous map. As a result, we also visualize the neighborhoods now in Manhattan-Bronx with cluster 1 which means that are similar to the neighborhood of the existing bar. This is presented with purple color and finally with red are the locations of already existing sports bars.



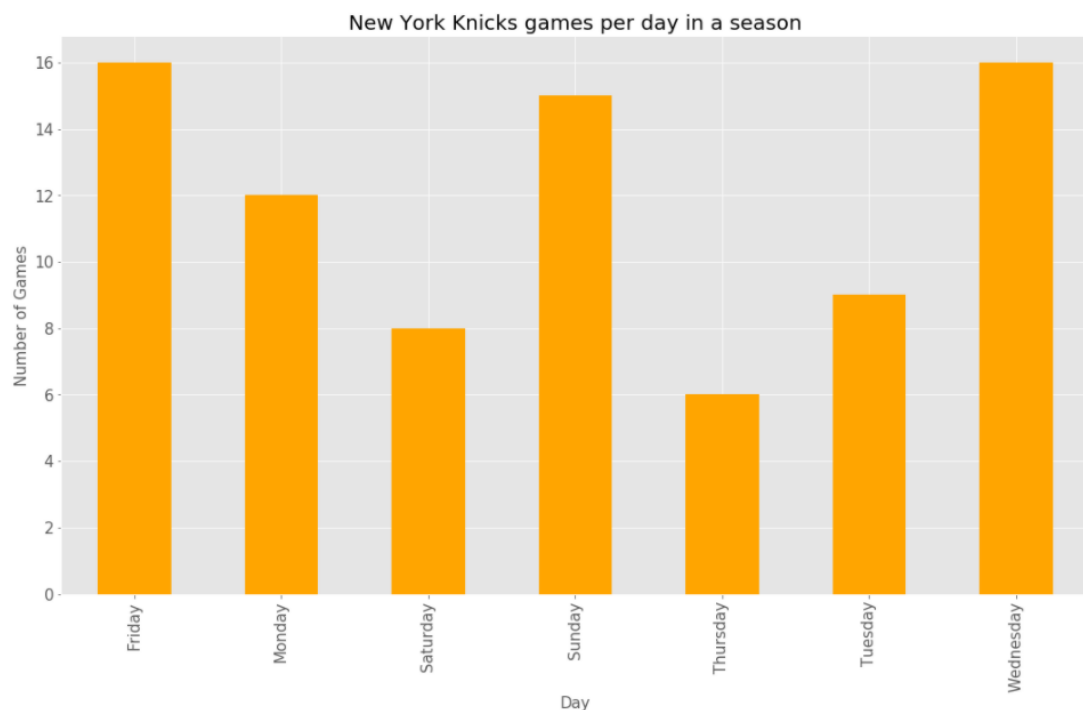
### 3.3 Manhattan-Bronx bar occupancy prediction

We made the first analysis with the potentially chosen locations for the new bar in Manhattan Bronx and now we will use the data for the schedule of New York Knicks and New York Yankees to predict the busy days to overstaff the bar. First we show the Dataset with the schedules of the two teams with the games per day and then the graphs that visualize these.

#### New York Knicks:

	Game	Date	Day
0	1	2018-10-17	Wednesday
1	2	2018-10-19	Friday
2	3	2018-10-20	Saturday
3	4	2018-10-22	Monday
4	5	2018-10-24	Wednesday
5	6	2018-10-26	Friday
6	7	2018-10-29	Monday
7	8	2018-10-31	Wednesday
8	9	2018-11-02	Friday
9	10	2018-11-04	Sunday

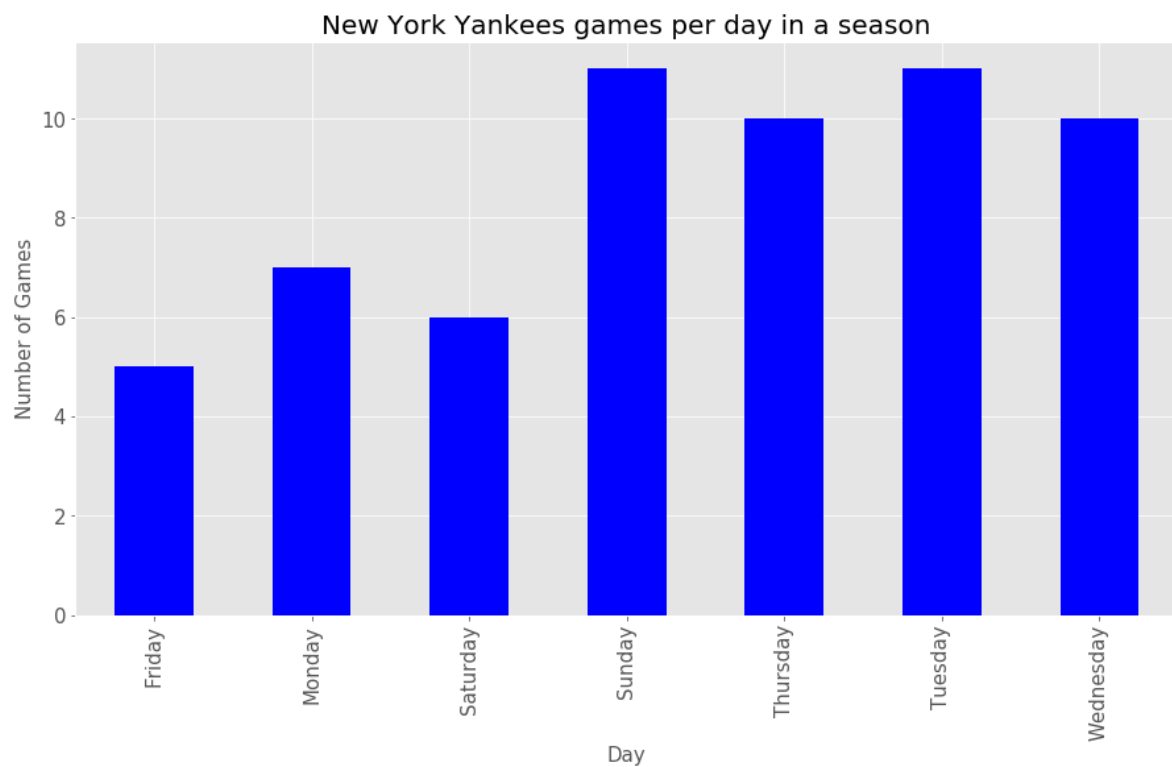
	Day	Game
0	Friday	16
1	Monday	12
2	Saturday	8
3	Sunday	15
4	Thursday	6
5	Tuesday	9
6	Wednesday	16



From the map above we can see that we are expecting Knicks fans in Wednesdays, Fridays and Sundays and it seems that Thursday is a less busy day.

## New York Yankees:

G			Date		G		Day	
0	1		Thursday	Jul 23			Friday	5
1	2		Saturday	Jul 25			Monday	7
2	3		Sunday	Jul 26			Saturday	6
3	4		Wednesday	Jul 29			Sunday	11
4	5		Thursday	Jul 30			Thursday	10
							Tuesday	11
							Wednesday	10



Regarding Yankees, it seems that busy days are Tuesday, Wednesday, Thursday and Sunday. Since Fridays and Saturdays are always busy days, we can say that only Monday is less busy.

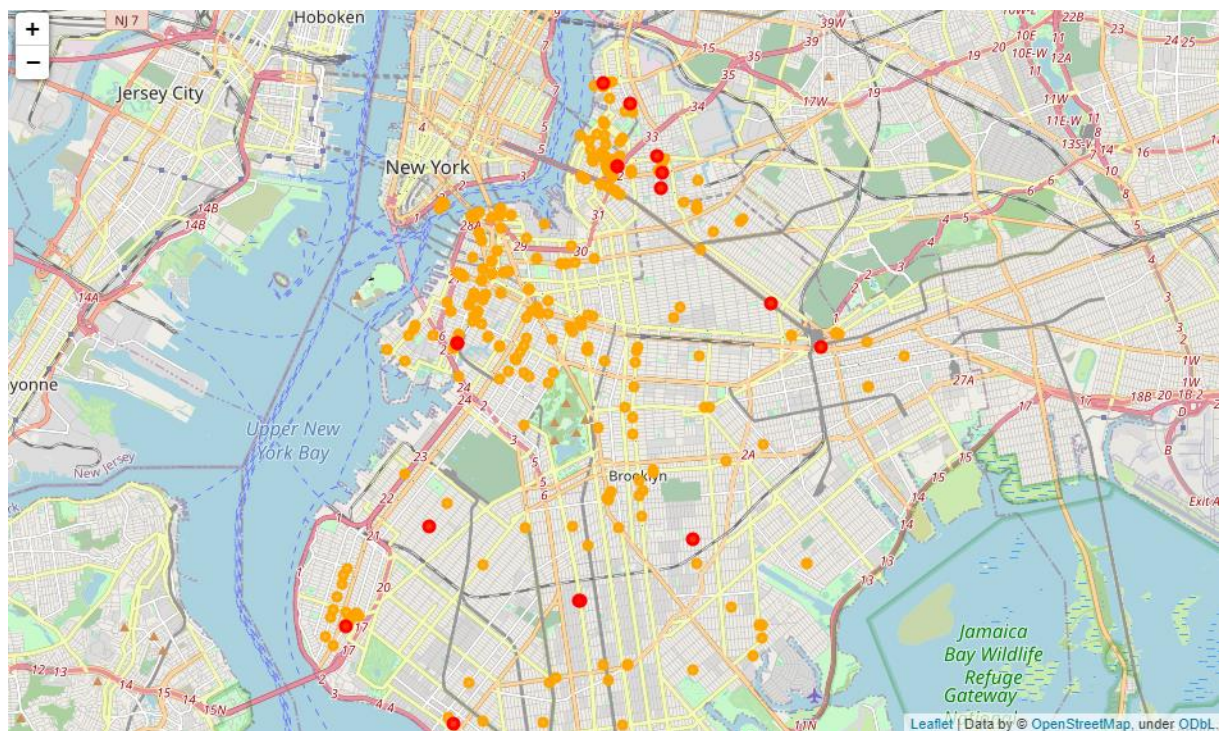
Next we are presenting the same analysis for the second bar in Brooklyn. The steps we followed are the same as Manhattan Bronx that is why there will be less description.



### 3.4 Brooklyn analysis

First is a Dataset with the sports bar in Brooklyn and next a map with all the venues in Brooklyn with the category bar with orange color and with red color are the sports bars in Brooklyn.

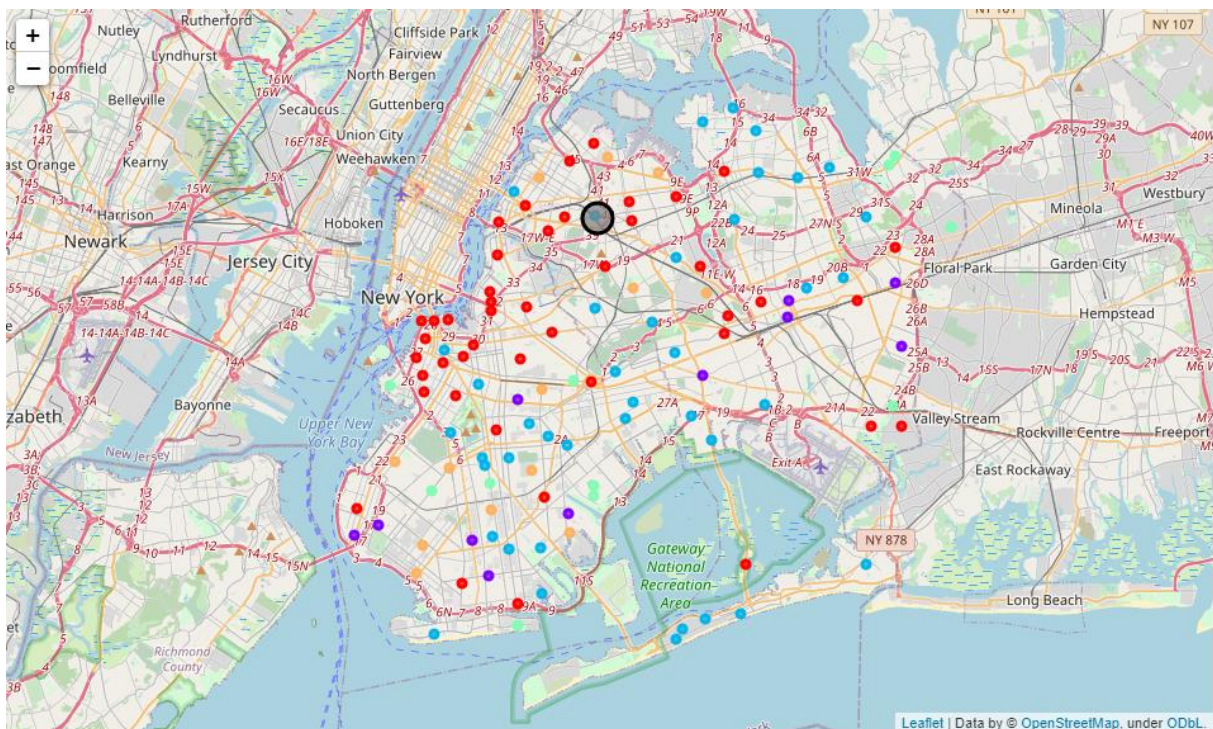
	Neighborhood	Venue	Venue Category	Venue Latitude	Venue Longitude
0	Sunset Park	Winners Circle Cafe Sports Bar	Sports Bar	40.639652	-74.005592
1	Greenpoint	Lake Street Bar	Sports Bar	40.725646	-73.951820
2	Greenpoint	Red Star NY	Sports Bar	40.729734	-73.958923
3	East Flatbush	The SandHills Sports Bar & Lounge	Sports Bar	40.637203	-73.934959
4	Williamsburg	Eddie JR's Sports Bar	Sports Bar	40.712948	-73.955081
5	Carroll Gardens	New York Sports Bar, Terminal B, LaGuardia Air...	Sports Bar	40.676990	-73.998180
6	Gowanus	New York Sports Bar, Terminal B, LaGuardia Air...	Sports Bar	40.676990	-73.998180
7	Bath Beach	TequilaVille Mexican Bar & Grill	Sports Bar	40.599518	-73.999237
8	Ocean Hill	Zona Viva Sports Bar	Sports Bar	40.684960	-73.914015
9	Midwood	Jack's Sports Bar	Sports Bar	40.624596	-73.965240



## 3.2 Brooklyn neighborhoods Clustering

Again, we used KMeans for clustering the neighborhoods with 5 Clusters.

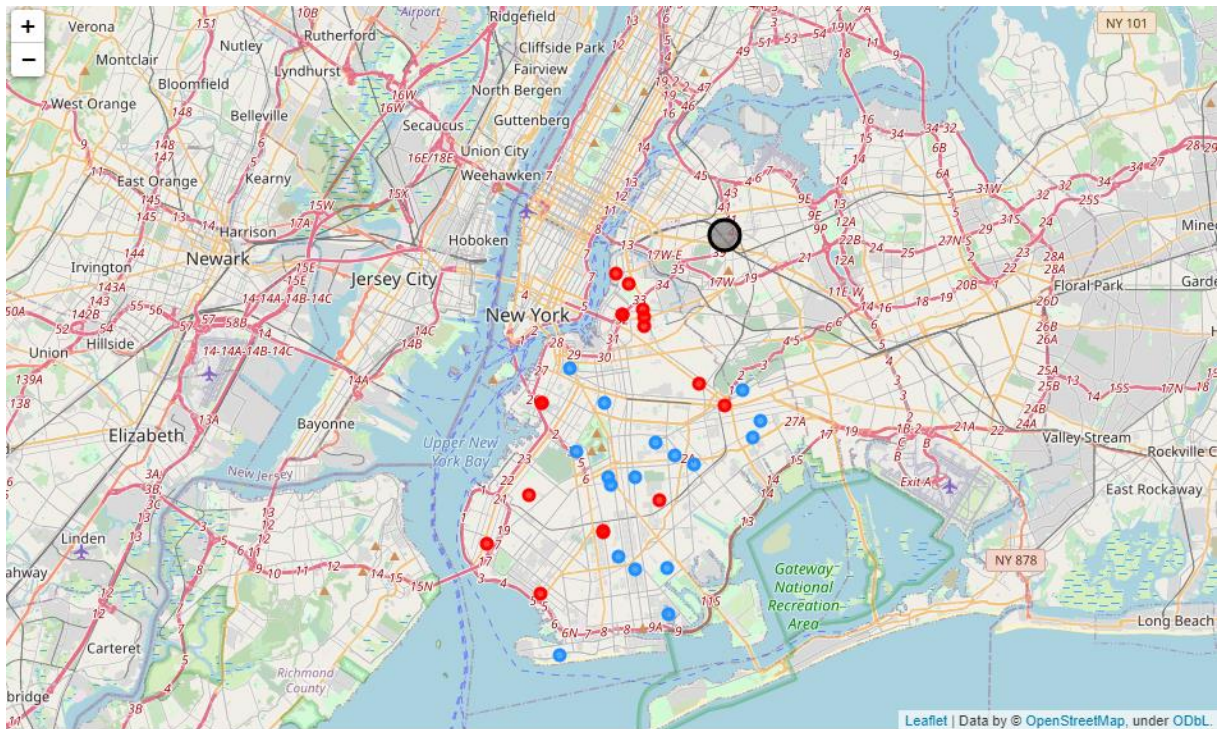
Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Brooklyn	Bay Ridge	40.625801	-74.030621	0.0	Bar	Juice Bar	Wine Bar	Hookah Bar	Cocktail Bar
Brooklyn	Bensonhurst	40.611009	-73.995180	4.0	Wine Bar	Sports Bar	Pub	Lounge	Karaoke Bar
Brooklyn	Sunset Park	40.645103	-74.010316	4.0	Sports Bar	Dive Bar	Cocktail Bar	Wine Bar	Pub
Brooklyn	Greenpoint	40.730201	-73.954241	0.0	Bar	Sports Bar	Wine Bar	Cocktail Bar	Pub
Brooklyn	Gravesend	40.595260	-73.973471	0.0	Bar	Lounge	Hookah Bar	Wine Bar	Sports Bar



In the above map with black circle is the existing bar and the clustering of the Neighborhoods in Brooklyn and Queens is shown with the other colors. Same colors seems similar neighborhoods.



The next maps shows the location of the existing sport bar of the 'Sports Bar' chain in Queens in a black circle marker. This bar belongs to Cluster 1 which is the turquoise/blue in the previous map. As a result, we also visualize the neighborhoods now in Brooklyn with cluster 1 which means that are similar to the neighborhood of the existing bar. This is presented with blue color and finally with red are the locations of already existing sports bars.

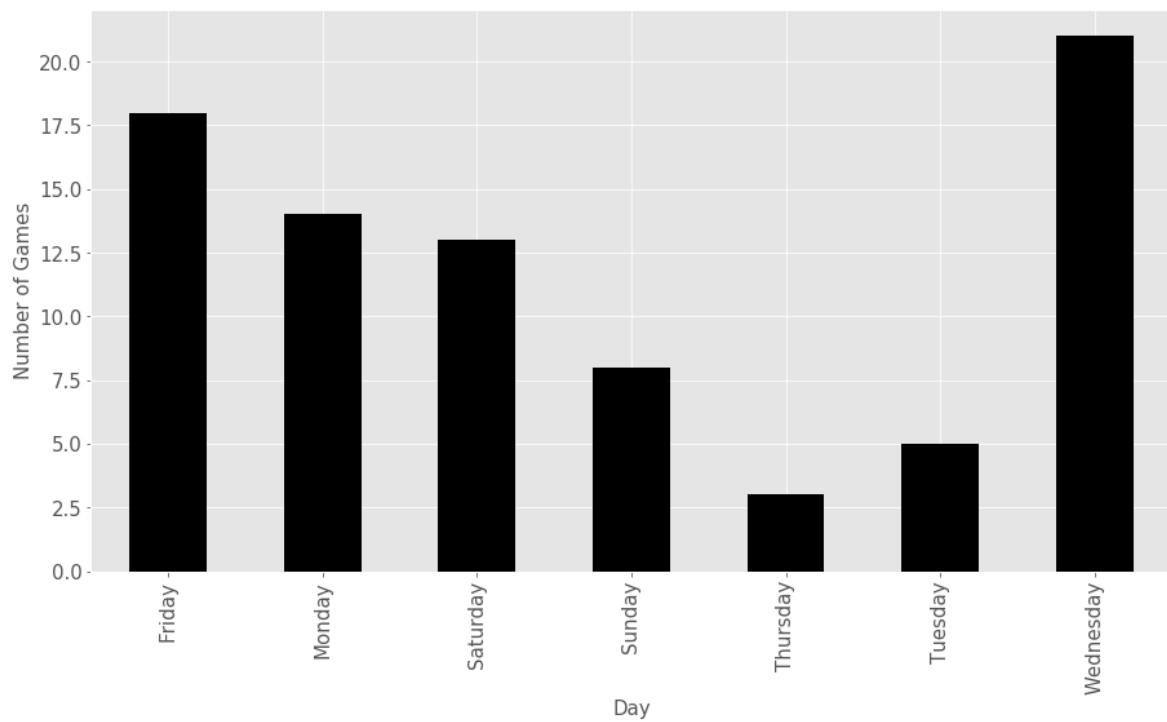




### 3.6 Brooklyn bar occupancy prediction

#### Brooklyn Nets:

	G	Date	Day
0	1	2018-10-17	Wednesday
1	2	2018-10-19	Friday
2	3	2018-10-20	Saturday
3	4	2018-10-24	Wednesday
4	5	2018-10-26	Friday

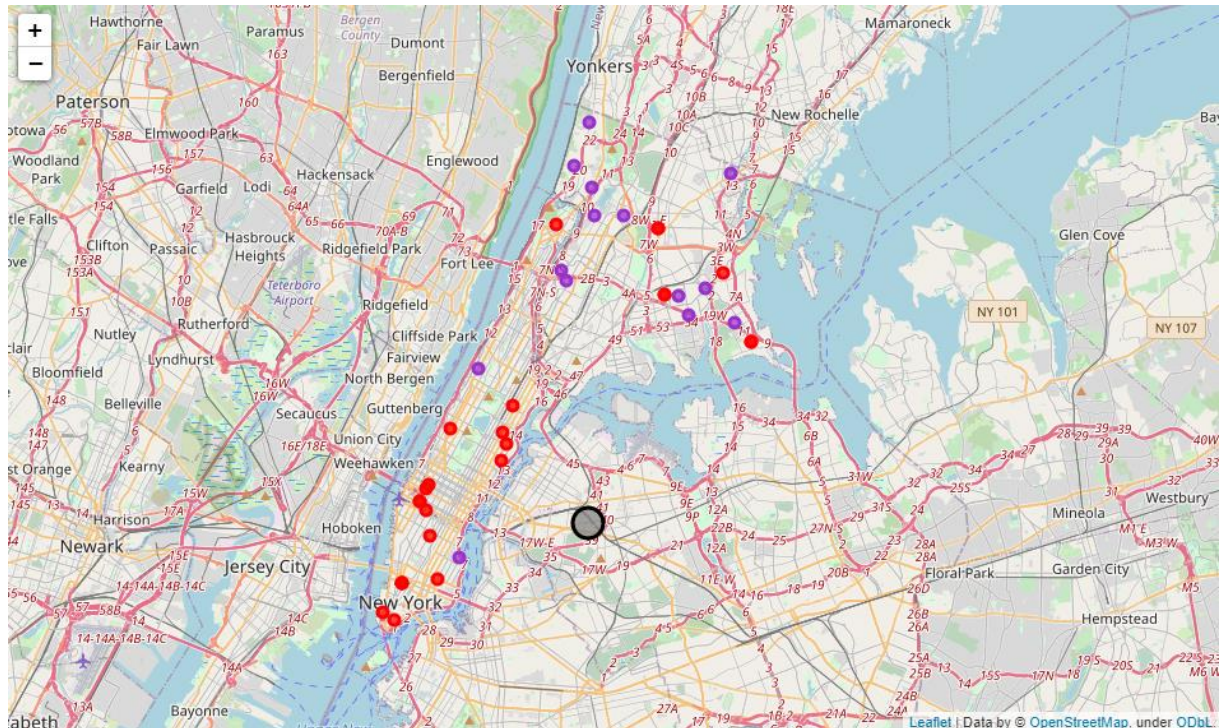


From the above graph, we can say that we expect more Brooklyn fans to come to the bar on Wednesday so is good to have more staff. In addition, it seems that in Tuesday and Thursday we do not expect Brooklyn Nets fans.

## 4. Results

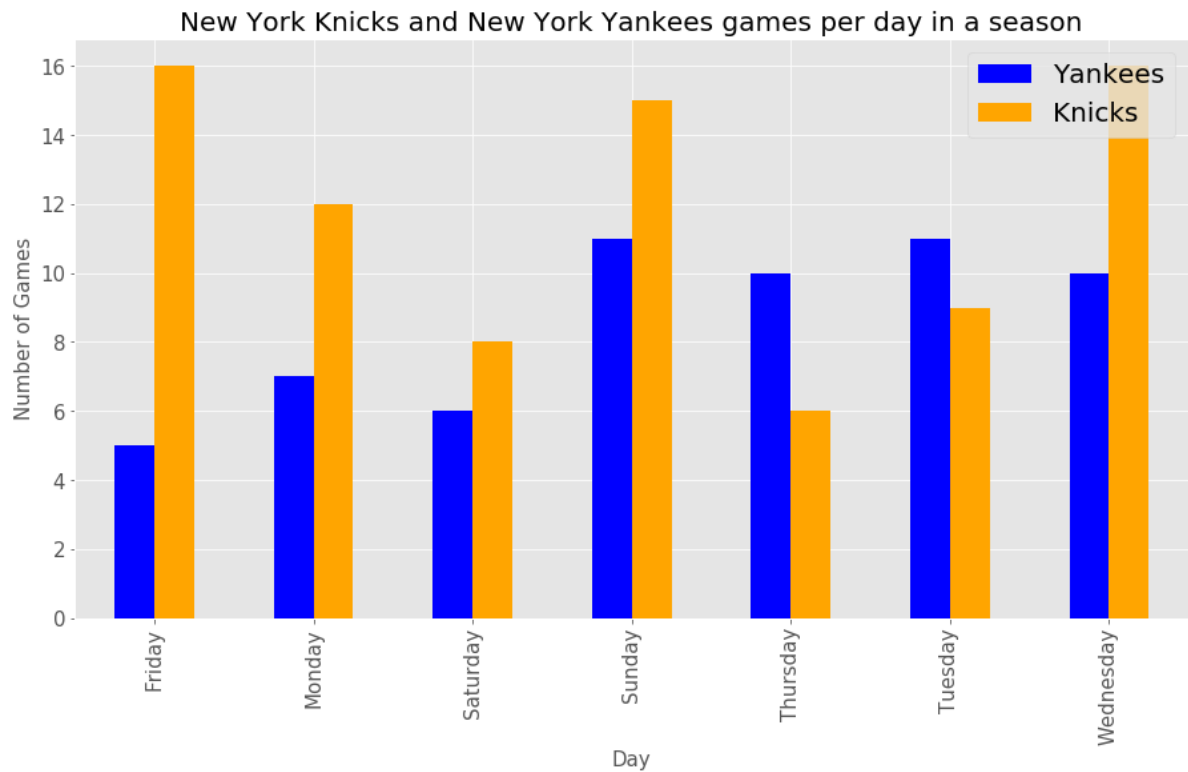
### 4.1. Bar in Manhattan-Bronx Location

For the location of the first bar, we present the map that shown in a chapter before.

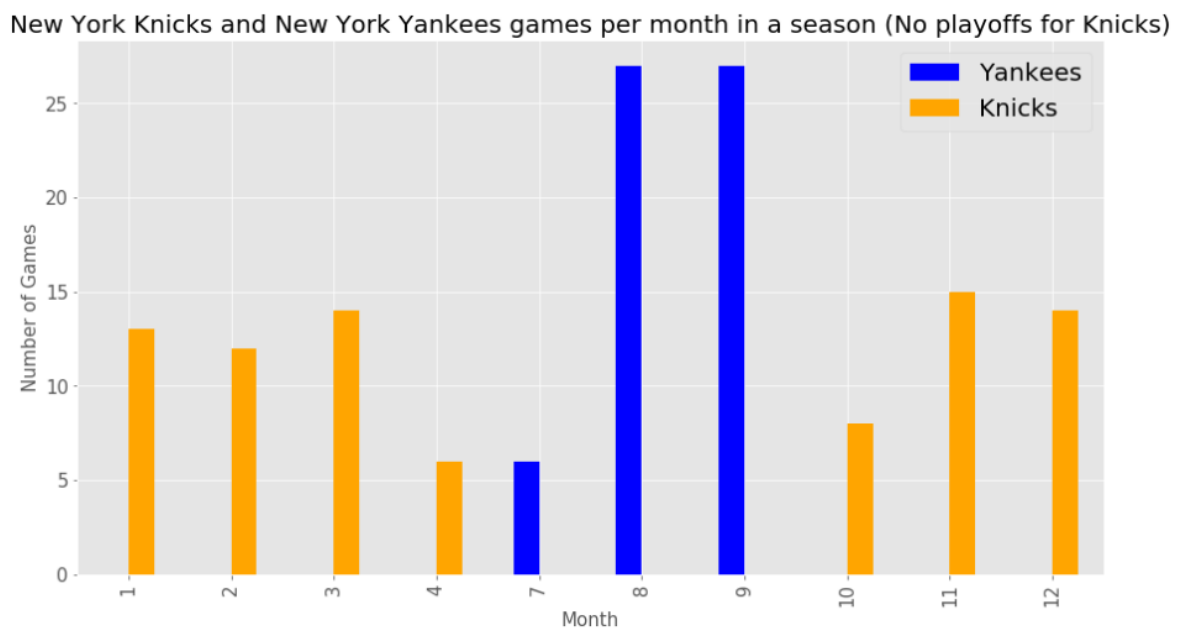


We can see that in purple are locations suitable for opening a new bar since the neighborhood is similar with the one in Queens. Also, we want to avoid sports bars close (red markers) so from a first view the purple neighborhood next to the river is a good location. Manhattan area is more suitable for the new bar since is more crowded with more venues. In addition, Manhattan is more suitable because New York Knicks are playing more games in a season than Yankees. They play for 7 months compare to 3 for Yankees. If further studies show that Yankees fans are fulling the bars more than Knicks then a location closer to Queens can be chosen like purple markers norther.

## 4.2. Bar in Manhattan-Bronx Location Staff



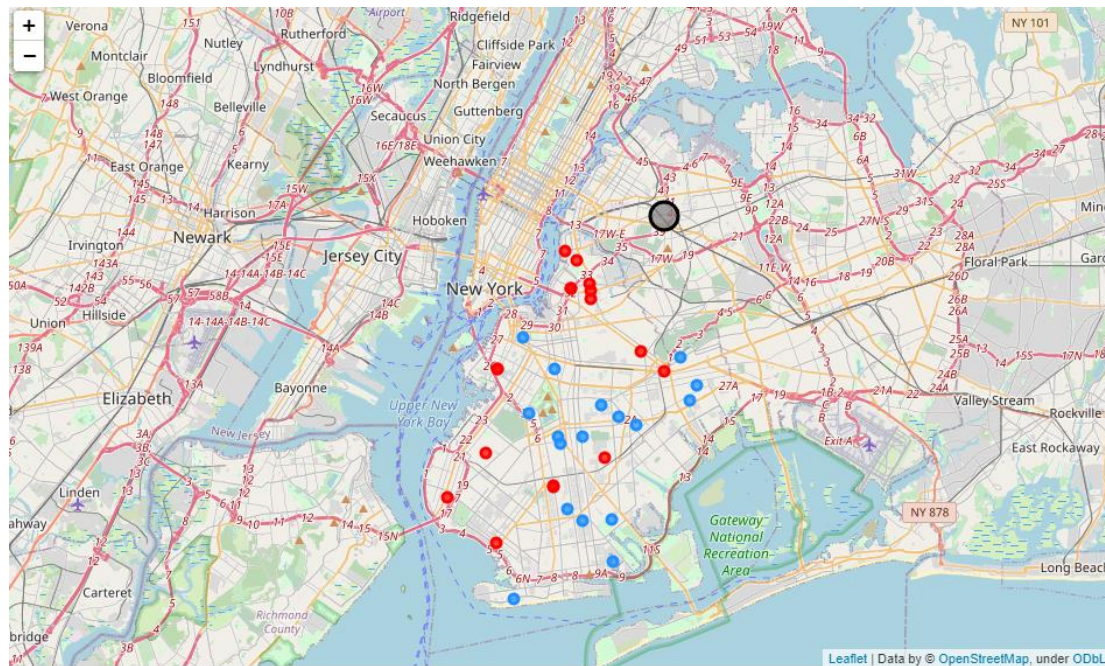
From the above plot, we can say that except the expected days Saturday and Sunday, we expect New York Knicks fans on Wednesday and Friday so is good to have some extra staff these days (less day offs these days). Also maybe Thursdays are not busy



As mentioned above, since Yankes are playing only for 3 months in a year, and Knicks for 7-8 (depend- ing if they enter playoffs) is better if the new bar is in Manhattan instead of Queens

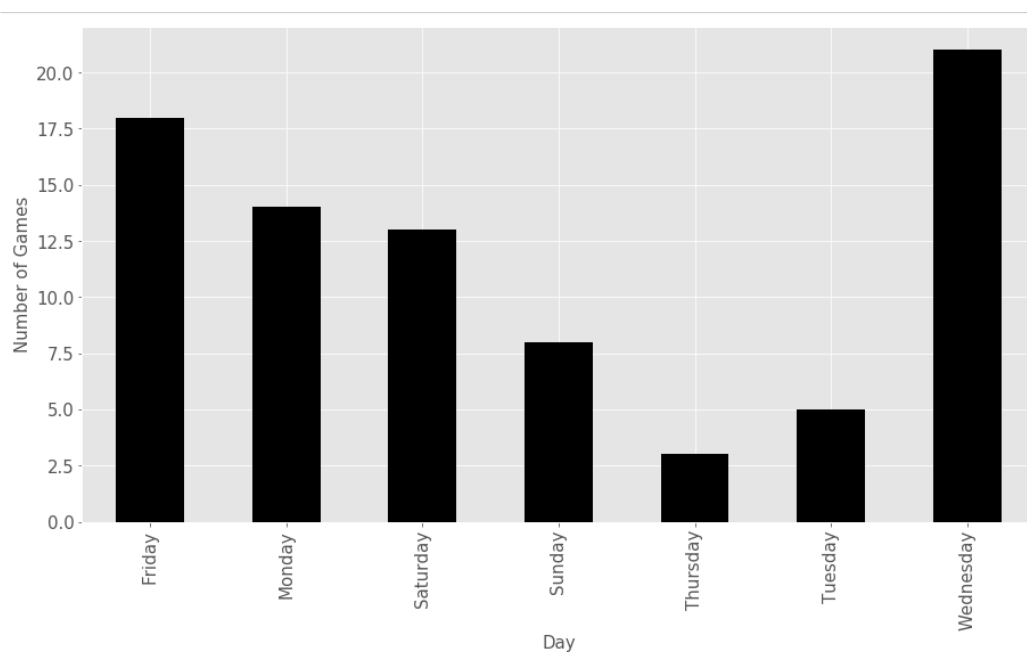


### 4.3. Bar in Brooklyn Location



A potentially good location for the Bar in Brooklyn are the Neighborhoods with Blue in the center of Brooklyn because there is no sports bar close and are similar with the one in Queens.

### 4.4. Bar in Brooklyn Location Staff



We can say that definitely we expect fans in Wednesdays and Fridays so is good to overstaff the bar these days alongside the weekend.

## **5. Discussion**

The final recommendations will listed in this section of the report:

### **Manhattan Bar:**

For the first bar in Manhattan, we can say that the Neighborhoods 'Morningside Heights' and 'North Riverdale' have increased chances to be successful for opening the new bar. However further analysis with more data regarding other factors than venues is needed.

Regarding the busy days we know for sure that Knicks and Yankees fans are expected in the bar Wednesdays, Fridays and Sundays. However once opening the bar maybe will attract more people than basketball and baseball fans.

### **Brooklyn Bar:**

On the other hand, in Brooklyn location with less risk are ' Windsor Terrace' and ' Prospect Heights'. As in Manhattan however more data are needed for the final decision.

Busy days with Brooklyn fans as mentioned are Wednesdays and Fridays and the weekend of course.

## **6. Conclusion**

In this study we analyze the best locations for an opening a new venue (Sport Bar in this case) after knowing a successful first location of the same venue. The results are quite interesting and the given notebook can also be used for similar cases except of Sports Bars. In advance someone can collect more data for these locations and proceed with further analysis of the neighborhoods to cluster them with higher groups/clusters.

In addition, in this analysis we predict the busy days of the venue by knowing what kind of people we expect to crowd the venue.