Profitable Italian Restaurant In New York

22 ŞUBAT

Sarp Özekmekçi

INTRODUCTION

Business Case

My business case is about establishing an Italian restaurant in New York. I want to learn the best position for opening this restaurant. That is the most profitable location for it.

Revenue of a restaurant is directly based on its sales which can be assumed to be depending on two main areas:

- First one is the quality and cost of the product. Food that are not delicious will not sell much and for an expensive food to be profitable it has to bring something unique to the table. Important point with this area is that these are mainly not dependent on the outside factors and can be improved after the restaurant is opened.
- Second one is the location, more explicitly, competition around the restaurant and its compliance with the customer portfolio at that neighbourhood. These cannot be changed after its done or amends would be expensive.

For above reasons, location of a newly established restaurant is very important and needs to be carefully evaluated. On this Project, optimum location for an Italian restaurant will be selected.

DATA

Main data that used for the assessment will be the intensity of Italian restaurants at a certain neighbourhood.

- At a location with high density of Italian restaurants, competetion will be more fierce resulting in potential lower revenues.
    - However there are two cases, if the density is high but the ratings are low then it could be considered as an opportunity. Assumption here is that there is demand for it with no satisfaction.
    - If the density is high with some highly rated Italian Restaurants, that location would not be ideal.
- At a location with lower density of Italian restaurants, competetion will be less resulting in potential high revenues.
    - However again there are two cases, if the density is low both for Italian restaurants and for restaurants in general it can be assumed that there is not demend for a restaurant at the location. In this case, that location would not be recommended.
    - If the density of Italian restaurants is low but in terms of restaurants in general the density is above a certain level, it can be assumed as potentially profitable location. Assumption here is that if there are restaurants, there is a demand for them despite the lack of Italian ones.

For this restaurant density, Foursquare API's database will be used and related Foursquare API subscription will be **Personal** which means there will be limitations to the requests.

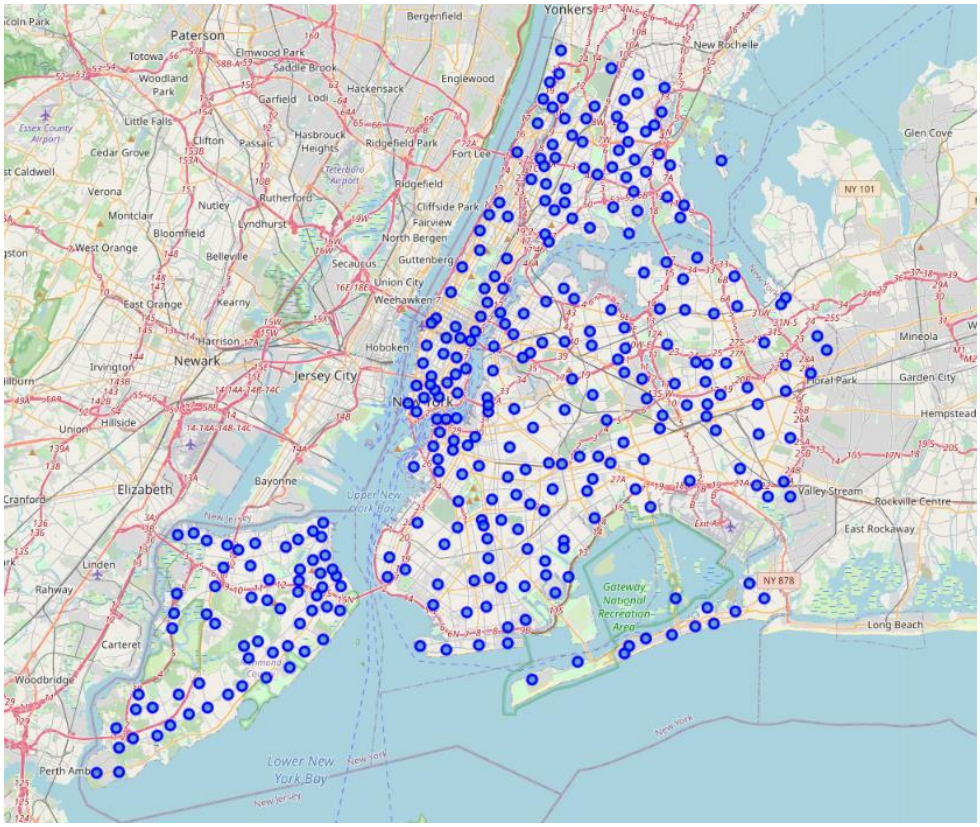

Figure 1: Map with pinpoints of neighbourhoods of New York

METHODOLOGY

When the results returned by the Foursquare API is analysed, it can be seen that there is no apparent cluster amount. Under normal circumstances, this would require DBScan methodology to be used as it doesn't require initial estimate of clusters. However for that, there needs to be large sample so that checks can be done. Due to limitations of the Foursquare API subscription that wasn't possible.

So, as an alternative K-Means method is selected. With that, initial estimate of clusters is required. This number is selected based on try – and – error looking at the distribution of restaurants at the clusters. When the number of restaurants is close to having a normal distribution, the results are accepted to be satisfactory.

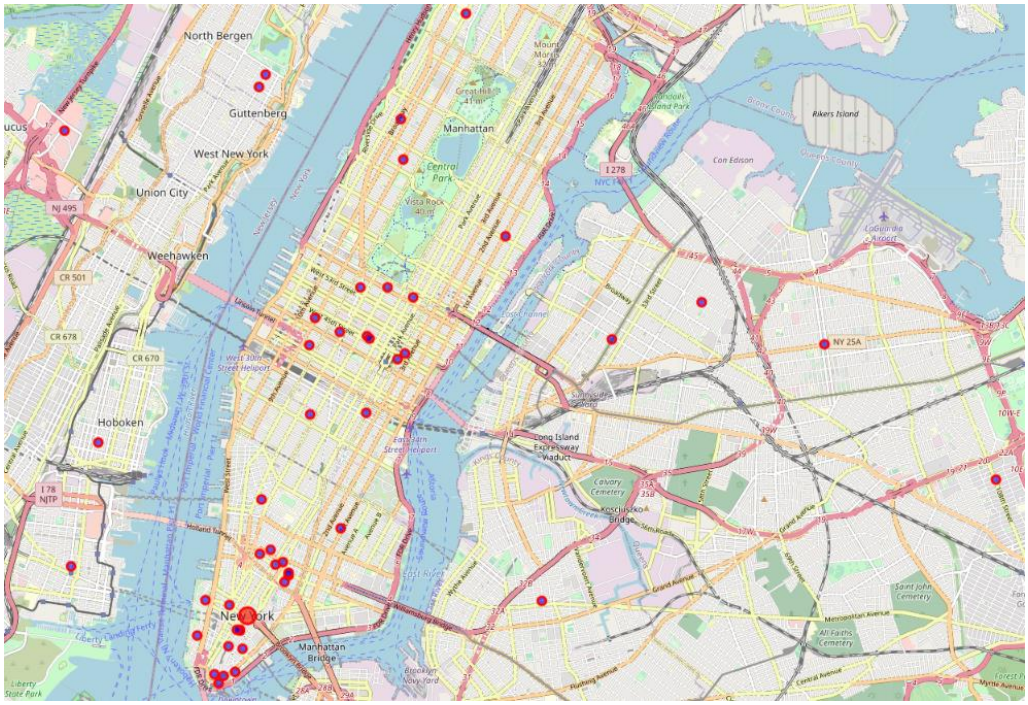

Figure 2: Map of Italian restaurants with pinpoints.

RESULTS

Distribution of the restaurants based on the clusters can be seen below:

	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6
Number of restaurants	6	23	14	4	2	1

According to this, some of the clusters are more crowded and some are less crowded. It is assumed that all the factors except the competition related ones are same at all these neighbourhoods for easier assessment. As a result, neighbourhood with least amount of Italian restaurants is deemed to be the most suitable one for establishing a new Italian restaurant, that is Area 6 which can be seen on the below map.

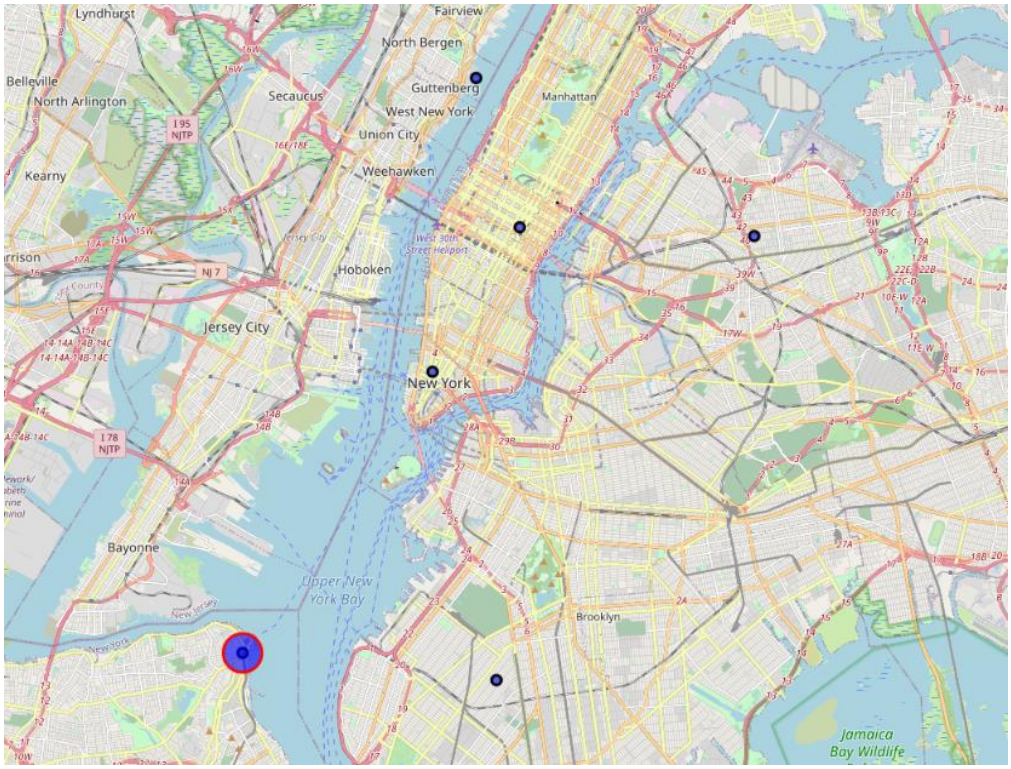

Figure 3: Map with pinpoints on the centroids of the 6 clusters also showing the Area 6 which is the most suitable neighbourhood.

DISCUSSION

There were some limitations to the project due to both the Foursquare API and time constraints. As further improvements, Italian restaurant list with larger sample size can be used. This will lead to better assessment.

As a second imrpovement, with a larger sample DBScan algorithm can be implemented to have a judgement unbiased by the initial cluster that is needed for the K-Means algorithm.

For the third improvement, analysis of restaurants in general can be carried out besides the one only for the Italian restaurants.

CONCLUSION

According to the analysis carried out, optimum location will be Area 6, bottom left of the map which is Thompkinsville because that is the least crowded neighbourhood in terms of Italian restaurants.