



BRONX OR QUEENS?

Find the ideal place to open a restaurant/food truck



Dan Wyss

Introduction

Background

The background of this analysis is to provide an in-dept analysis for the question where in New York, particular in the district of the Bronx and Queens, is the most appropriate place for opening either a restaurant, a fast food store or just a food truck.

Problem

Besides the used economical instruments for a business creation like a business plan, a SWOT analysis and a investment schedule the focus in this report is in a in-depth environmental analysis of the venues within the above mentioned district to identify the area where the food competition is weak and where a potential growth business can be established either stationary by opening a restaurant or transient by a food truck.

Interest

This may be of interest for any potential owner of a business in the food sector, who is willing to open a location either in the neighborhood of the Bronx or Queens. However, the underlying notbook could potentially also be leveraged for any other place of interest under the condition venues data are available for the location and venues are up-to-date at Foursquare.



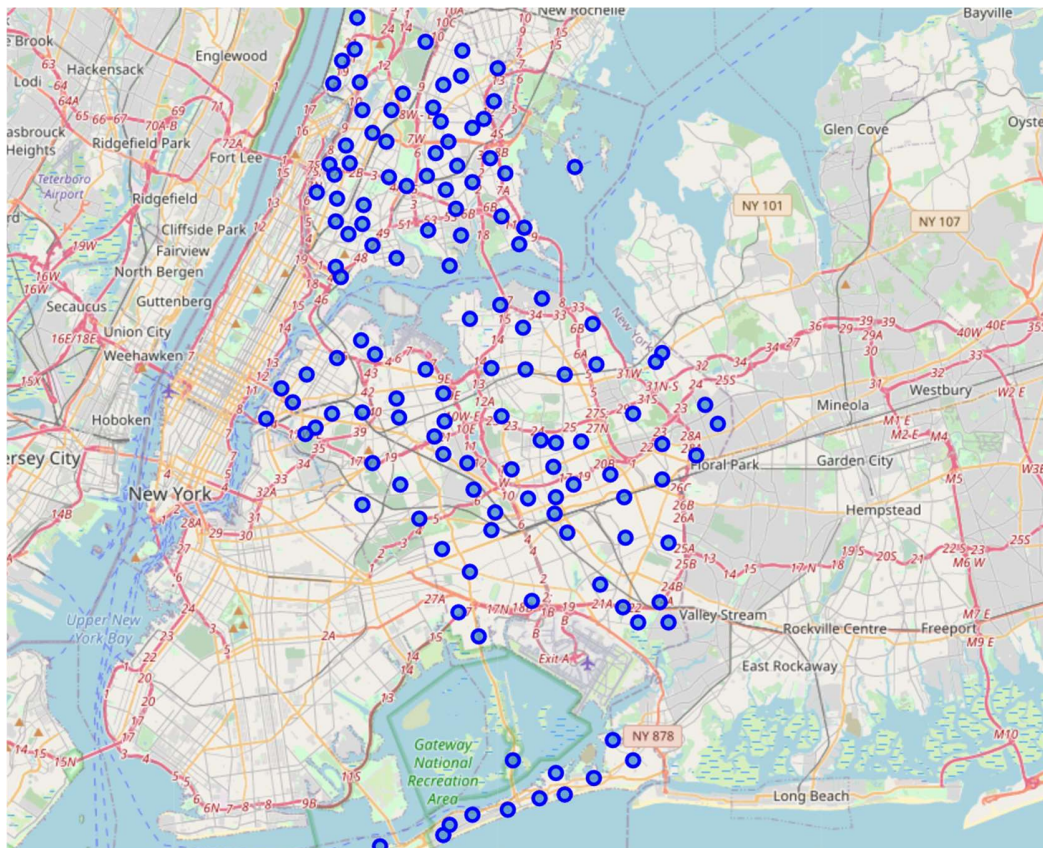
Data acquisition & cleansing

Get the data of New York

As a first step a JSON file containing all the borough and neighborhood name including the corresponding latitude and longitude of the neighborhoods are downloaded.

The structure of a JSON file is quite cumbersome and need therefore to be formatted in a way, in which the data can be further processed. Therefore the needed data will be extracted and inserted in a dedicated generated data frame structure to facilitate further processing.

The various neighborhoods of the Bronx and Queens are then displayed on the map of New York to get an geographical overview:



Neighbourhood data is not enough – we are looking for the venues and their categories

As already described in the previous section we own now the neighborhood data of the districts of the Bonx and Queens. But is this suitable? No, definitely not. What we need is the data of the various venues within the neighborhood and the categories thereof to better understand and predict the competition for a newly opened food store.

Foursquare

Foursquare provides information in regards of different venues within a city or a place. The venues are categorized in terms of interest, e.g., food, shopping, etc. The use of those additional information is needed to conduct our analysis.

Therefore all venues for the defined neighborhoods are downloaded and moved into dedicated data frame.

For short example of the data for the first 5 records for the neighborhood of Wakefield is presented below:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Ripe Kitchen & Bar	40.898152	-73.838875	Caribbean Restaurant
2	Wakefield	40.894705	-73.847201	Ali's Roti Shop	40.894036	-73.856935	Caribbean Restaurant
3	Wakefield	40.894705	-73.847201	Jackie's West Indian Bakery	40.889283	-73.843310	Caribbean Restaurant
4	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop

There are about 378 categories within this data frame. This is far to much to come to a conclusion where the food competition is most severe.



Further grouping

As a next step, the data is grouped per neighbourhood and the frequency of occurrence per category is calculated.

It is worth to mention that the focus is only on the top three categories identified by the percentage of frequency of occurrence because this may already sum up of between 25% and 40% and we want just to focus on the main and big occurrences. An example of such an outcome for Wakefield is provided below:

```

----Wakefield----
      venue  freq
0      Pharmacy 0.12
1 Caribbean Restaurant 0.08
2      Supermarket 0.08

```

The outcome of this process is then transformed into a data frame structure for defining the clusters by using k-means clustering.

K-means clustering and visualisation

K-means clustering methodology is used to define clusters according to their categories. We set the amount of cluster = 3 and prepared the outcome visually on the map with different colours. The table below shows the outcome of the clustering for a couple of neighbourhoods in the Bronx and the respective cluster based on the most common venues:

Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
Bronx	Wakefield	40.894705	-73.847201	1	Pharmacy	Supermarket	Caribbean Restaurant
Bronx	Co-op City	40.874294	-73.829939	2	Department Store	Mobile Phone Shop	Shopping Mall
Bronx	Eastchester	40.887556	-73.827806	1	Caribbean Restaurant	Deli / Bodega	Diner
Bronx	Fieldston	40.895437	-73.905643	2	Bar	Pizza Place	Bus Station
Bronx	Riverdale	40.890834	-73.912585	2	Park	Pizza Place	Bar



The visualization of the clustering exercise provides a better picture on the map:



We can see on the map above the the three cluster are spread over the two different districts of the Bronx and Queens. But the main question is how to the differentiate?



Examining the clusters (to be found in the appendix section)

Examining the red cluster

The categorization of the red cluster on the map is presented in the table above. It can be stated that there is not really a high occurrence of restaurants or other food stores, but the main venue is related to beach and the other leisure categories.

Examining the purple cluster

The purple cluster is overflowed with restaurants, fast food places, donut shops and pizza stores. There is a huge competition in this area for a new food business.

Examining the green cluster

The green cluster has also a quite high penetration of resuarants, pizza places and donut shops but it seems that the penetration is not as high as compared to the purple cluster. However, despite the lower relative penetration, the competion is still huge and therefore unfavorable for a new food business.

Conclusion and recommendation

Based on the outcome of the analysis above the only cluster to open a food business without too much competition is the red cluster and can therefore be seen as the only recommendation.



Appendix

Red Cluster

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
Broad Channel	Other Nightlife	Park	Playground
Breezy Point	Surf Spot	Pizza Place	Trail
Arverne	Beach	Surf Spot	Deli / Bodega
Rockaway Beach	Beach	Ice Cream Shop	Bar
Neponsit	Beach	Deli / Bodega	Boutique
Belle Harbor	Beach	Boutique	Spa
Rockaway Park	Beach	Pizza Place	Bar
Roxbury	Art Gallery	Theater	Beach
Hammels	Beach	Surf Spot	Donut Shop



Purple cluster

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
Wakefield	Pharmacy	Supermarket	Caribbean Restaurant
Eastchester	Caribbean Restaurant	Deli / Bodega	Diner
Norwood	Pizza Place	Sandwich Place	Donut Shop
Williamsbridge	Caribbean Restaurant	Pizza Place	Fried Chicken Joint
University Heights	Grocery Store	Pizza Place	Fried Chicken Joint
Morris Heights	Food Truck	Pharmacy	Deli / Bodega
Melrose	Mexican Restaurant	Donut Shop	Grocery Store
Longwood	Fast Food Restaurant	Grocery Store	Mexican Restaurant
Morrisania	Fast Food Restaurant	Grocery Store	Discount Store
Soundview	Fried Chicken Joint	Department Store	Fast Food Restaurant
Olinville	Pizza Place	Fried Chicken Joint	Sandwich Place
Pelham Gardens	Donut Shop	Pharmacy	Pizza Place
Concourse	Pharmacy	Pizza Place	Deli / Bodega
Edenwald	Pizza Place	Fast Food Restaurant	Caribbean Restaurant
Hollis	Pizza Place	Fried Chicken Joint	Deli / Bodega
South Jamaica	Fried Chicken Joint	Park	Pizza Place
St. Albans	Caribbean Restaurant	Pizza Place	Convenience Store
Rochdale	Pharmacy	Chinese Restaurant	Sandwich Place
Springfield Gardens	Gym	Donut Shop	Chinese Restaurant
Cambria Heights	Caribbean Restaurant	Pharmacy	Restaurant
Rosedale	Cosmetics Shop	Caribbean Restaurant	Sandwich Place
Far Rockaway	Chinese Restaurant	Pizza Place	Metro Station
Laurelton	Caribbean Restaurant	Cosmetics Shop	Deli / Bodega
Brookville	Intersection	Deli / Bodega	Park
Claremont Village	Discount Store	Donut Shop	Pizza Place
Mount Eden	Grocery Store	Pizza Place	Deli / Bodega
Mount Hope	Grocery Store	Fast Food Restaurant	Pizza Place
Allerton	Pizza Place	Deli / Bodega	Donut Shop
Bayswater	Grocery Store	Playground	Park



Examining the green cluster

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
Co-op City	Department Store	Mobile Phone Shop	Shopping Mall
Fieldston	Bar	Pizza Place	Bus Station
Riverdale	Park	Pizza Place	Bar
Kingsbridge	Pizza Place	Mexican Restaurant	Sandwich Place
Woodlawn	Deli / Bodega	Pizza Place	Pub
Baychester	Clothing Store	Pizza Place	Department Store
Pelham Parkway	Pizza Place	Bus Station	Deli / Bodega
City Island	Harbor / Marina	Seafood Restaurant	Italian Restaurant
Bedford Park	Pizza Place	Park	Diner
Fordham	Italian Restaurant	Pizza Place	Spanish Restaurant
East Tremont	Pizza Place	Donut Shop	Park
West Farms	Pizza Place	Zoo	Supermarket
High Bridge	Baseball Stadium	Lounge	Park
Mott Haven	Pizza Place	Donut Shop	Mexican Restaurant
Port Morris	Baseball Field	Donut Shop	Spanish Restaurant
Hunts Point	Park	Bank	Paintball Field
Clason Point	Park	Bus Stop	Pool
Throgs Neck	Italian Restaurant	Deli / Bodega	Pizza Place
Country Club	Deli / Bodega	Italian Restaurant	Bank
Parkchester	Pizza Place	Donut Shop	Supermarket
Westchester Square	Pizza Place	Fast Food Restaurant	Donut Shop
Van Nest	Pizza Place	Supermarket	Donut Shop
Morris Park	Pizza Place	Sandwich Place	Pharmacy
Belmont	Italian Restaurant	Pizza Place	Bakery
Spuyten Duyvil	Park	Pharmacy	Pizza Place
North Riverdale	Pizza Place	Deli / Bodega	Burger Joint

