

Coursera Capstone Project

The Battle of Neighbourhoods

Introduction

New York City is the largest and most populous city in the United States, 2018 it's population was 8,398,748. It consists of five boroughs – Brooklyn, Queens, Manhattan, The Bronx, and Staten Island, each of which is a separate county of the State of New York. Each borough has its own demographic, cultural, linguistic and financial specificities.

But you don't have to be Italian to like pizza. I mean a good pizza, made according to an old Italian recipe.

Then a delicious espresso. Yes, a well crafted espresso tastes great. And ice cream.

In this project, I'm going to **explore the location** of good, high rated **Italian restaurants** in New York City and **find a place where such a restaurant is missing**. We will look for highly populated areas with not many other Italian restaurants in vicinity.

Goal and stakeholders

The goal of this project is to explore the following questions:

- 1. What is best location in New York City for Italian cuisine?**
- 2. Which areas lack Italian restaurants?**

Stakeholders interested in the problem:

Investors planning to open an Italian restaurant.

Data

We will use the following data:

- 1. New York City data to explore various neighborhoods of New York City.**

This dataset covers boroughs along with their neighborhoods, latitude and longitude

Dataset source:

https://geo.nyu.edu/catalog/nyu_2451_34572

- 2. GeoSpace data**

The geo space data allow to get the New York Borough boundaries. It let us visualize choropleth map.

Source:

<https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm>

- 3. Location of the Italian restaurants in New York City**

- a) By using the Fousquare API we will get all the venues in each neighborhood.
 - b) Then we can filter out only Italian restaurants.
 - c) We can also find ratings, tips and like count for each restaurant.
- That allows us to sort and visualize the data.

Source:

Fousquare API

Methodology

In this project we want to detect areas of New York that have a low density of Italian restaurant bit a high population density.

First, We have collected the data:

- location of Boroughs and Neighborhoods
- location of Italian restaurant in each Neighborhood (within 1500 m from their center)
- the population of the Boroughs

In the next step we analysed the '**restaurant density**' across the Boroughs of New York. In the final step we focused on the most populated areas and filtered out the Neighborhoods with the least number of Italian restaurants.

These are the most promising areas which should be attractive to investors.

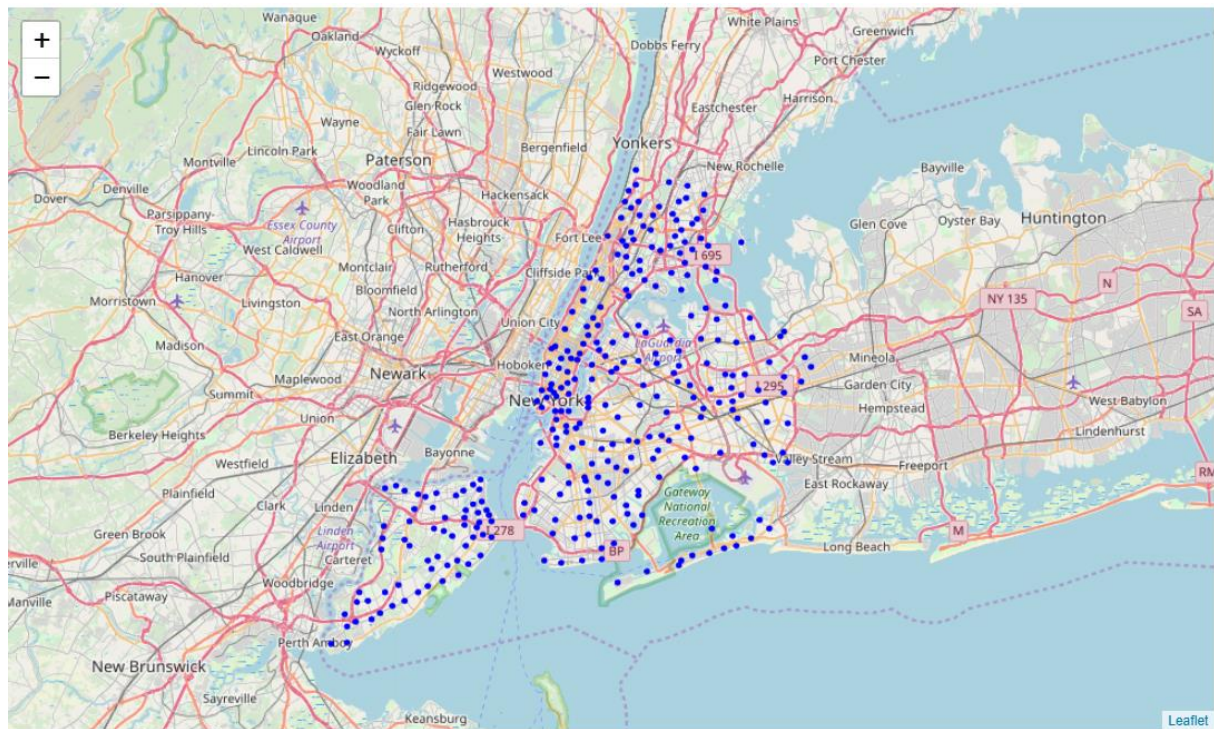
Distribution of Italian Restaurants in New York City

The geographical coordinates of New York City, NY are 40.7127281, -74.0060152.

New York City has 5 boroughs and 306 neighborhoods

	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

Visualize the neighborhoods



Italian restaurants within a radius of 1500 m from the center of every neighborhood

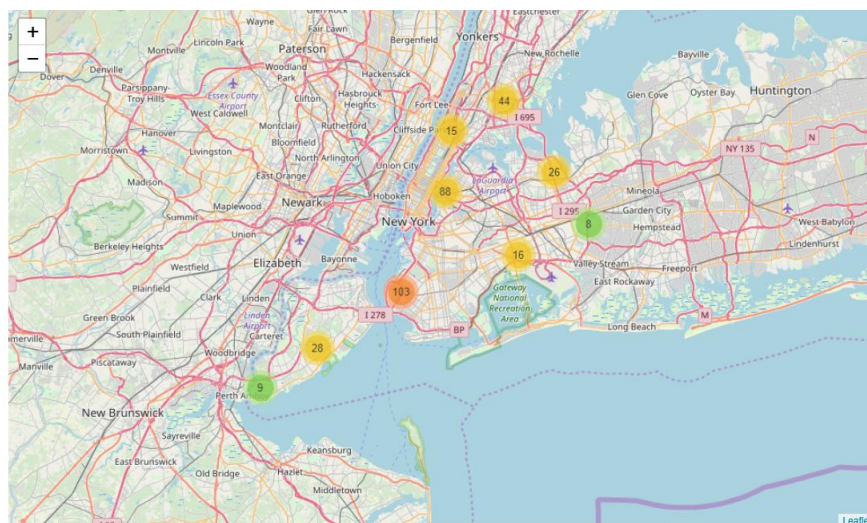
Let's use Foursquare API to get info on Italian restaurants in each neighborhood. We're interested in venues in proper Italian restaurants, we won't care about places like coffee shops or bakeries etc.

There are 337 Italian restaurants in 306 neighborhoods.

Visualize the location of the Italian restaurants in New York



Let's cluster and count the restaurants



Population of New York City

Let's look at the distribution of population density in New York.

	Borough	2010 Population
1	Brooklyn	2504700
3	Queens	2230722
2	Manhattan	1585873
0	Bronx	1385108
4	Staten Island	468730

Analysis

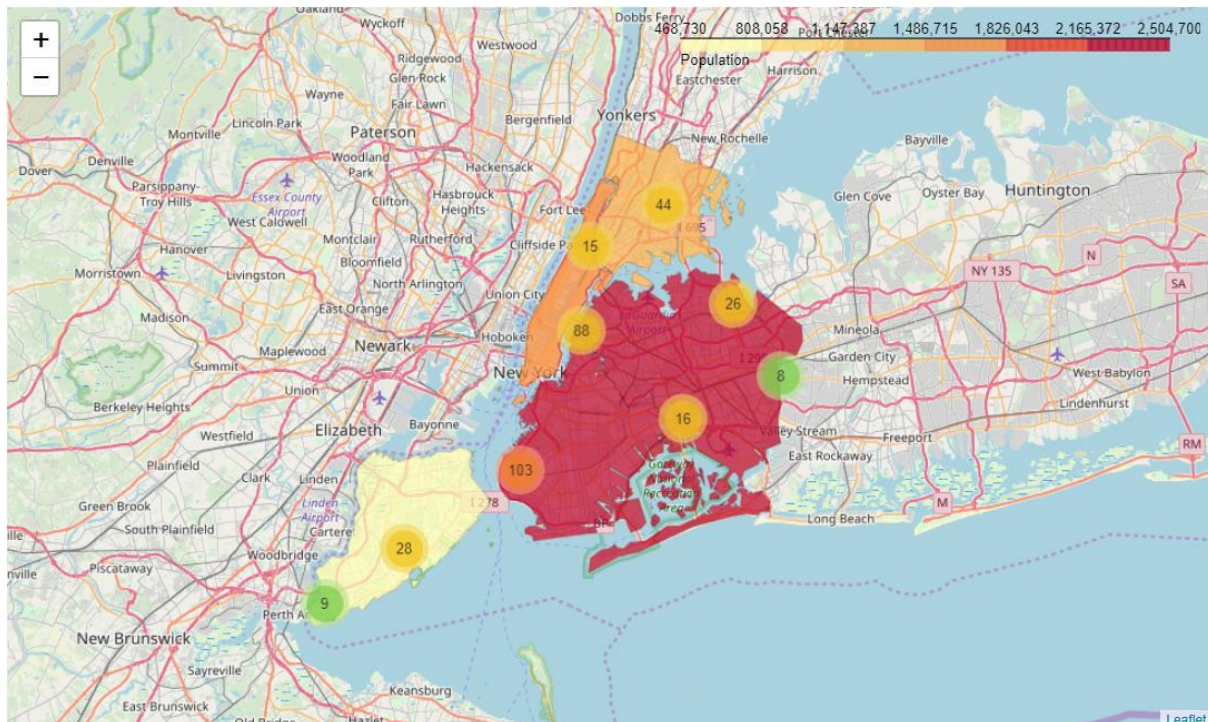
Bring it together: Density of population and Italian restaurants

After the overlapping of population density and the location of Italian restaurants we will see in which areas of the city there are large concentrations of people who do not have near such a restaurant.

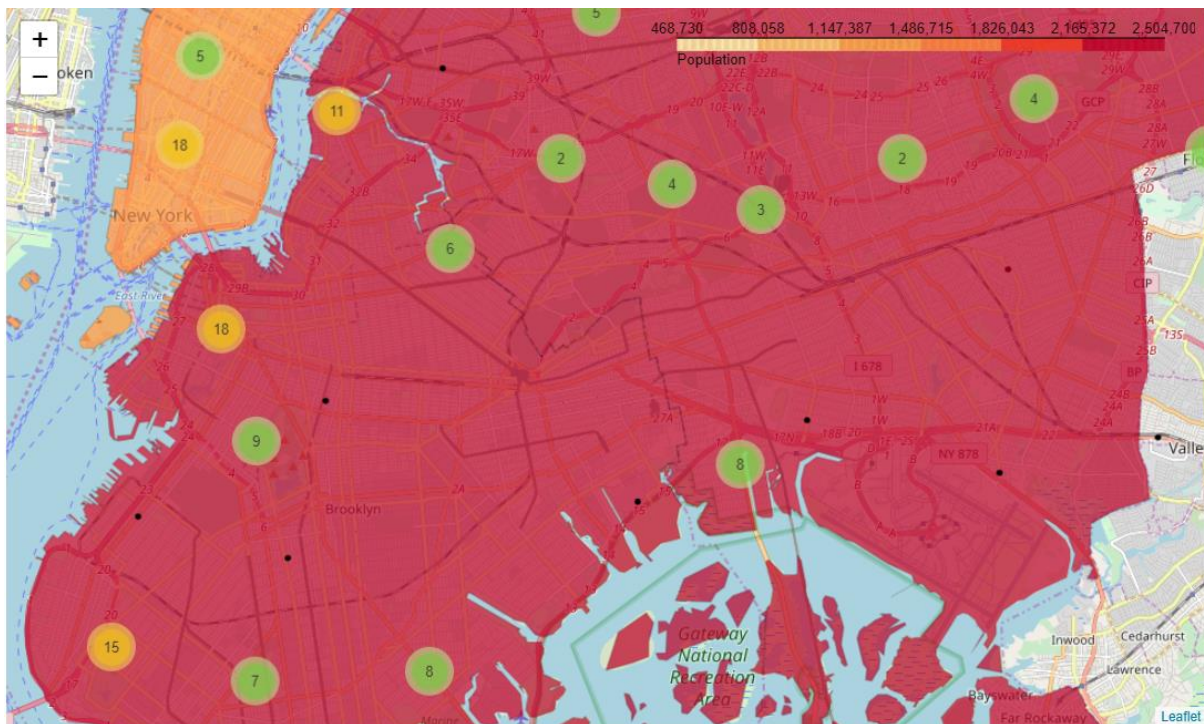
By changing the zoom we will be able to obtain data with the selected accuracy.

The next step in the analysis will focus on densely populated areas.

We will try to find the areas with the lowest number of Italian restaurants there.



Let's take a closer look at the most populated areas:



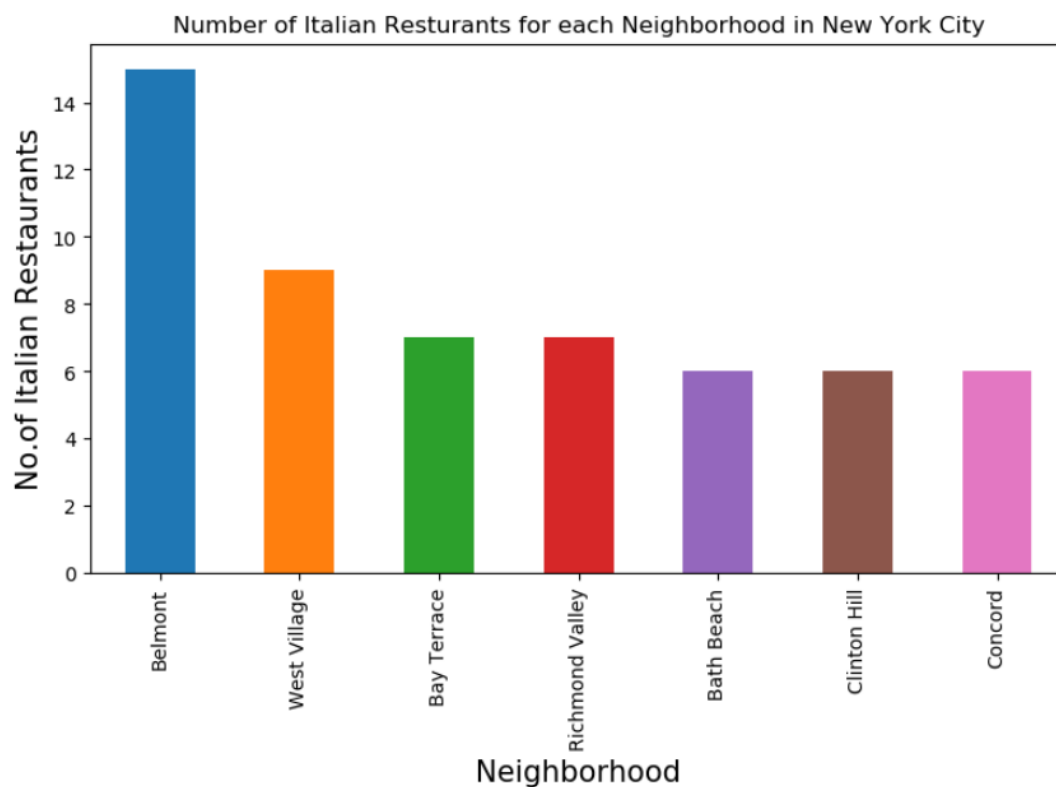
The most densely populated boroughs (red) are Brooklyn and Queens. We can see which areas of these Boroughs have only few Italian restaurants. **They are the areas we should care about.**



Let's look for neighborhoods.

Neighborhood

Belmont	15
West Village	9
Bay Terrace	7
Richmond Valley	7
Clinton Hill	6



As we can see, most Italian restaurants are in the Belmont area.

This is not a coincidence.

As a result of its cultural history and wide array of Italian businesses, Belmont is known as the "**Little Italy**".

Results

Now, let's check out the neighborhoods of **Brooklyn** and **Queens**.

We select the neighborhood in both boroughs with no or one italian restaurant:

Candidates in Brooklyn:

Boerum Hill

Brighton Beach

Coney Island

Ditmas Park

Downtown

Erasmus

Fulton Ferry

Madison

Manhattan Beach

Prospect Heights

Prospect Lefferts Gardens

Starrett City

Sunset Park

Candidates in Queens:

Astoria

Bellaire

Blissville

Briarwood

Brookville

East Elmhurst

Holliswood

Jackson Heights

Maspeth

North Corona

Pomonok

Rockaway Park

Rosedale

South Ozone Park

St. Albans

Utopia

Whitestone

Discussion

In this project we tried to find the best places to open an Italian restaurant. The approach presented here is, of course, simplified. We have only considered two factors. In order to get the appropriate results it would be necessary to consider many more factors, including the average income of the inhabitants, their ethnic structure, type of the area (whether it is for example a typical industrial, office or living district).

However, the study presented above allows us to narrow down the area of New York to places where a newly opened Italian restaurant would be competitive and would have a chance of success.

Conclusion

Purpose of this project was to find areas of New York with low number of Italian restaurants and high density of population. The goal was to help the stakeholders narrowing down the search for optimal location for a new Italian restaurant.

By calculating the restaurant distribution from Foursquare data and population density we have identified the boroughs and the neighborhoods within them, which could be used as starting points for final exploration by stakeholders.

The final decision requires consideration of a number of additional factors, specific characteristics of neighborhoods and locations in every recommended area as well the preferences of the stakeholders.