

Introduction

The problem I try to answer is "what is the recommended location for a new cafe in Paris". Suppose that a young entrepreneur wants to start his bussiness and he want to begin with a cafe.

- Should he place his cafe in a place where there are a lot of existing cafes or in a place where there is none or should he try to find some kind of balance between these two extremes?
- What is secret of a successful cafe? Can we get some tips or information based on the data provided by foursquare?

EDIT : I didn't find a Dataset in Paris so I make my report on New York

Data will be used to solve the problem.

To answer at this question, we have to :

- Find the high rating cafe in Paris and identify them visually on a map.
- Create several "clusters" of these cafes to identfiy potential districts to choose from. (We should choose one from these business district and locate his cafe in the center of it.) For the problem two, researcher will collect the tips (customer reviews) for the cafe in that area and filter three types of tips : high ratings, low ratings and long comments because these tips will more likely include heeds for providing better services for the customers.

1. Search for top-rated "cafe" in Paris - Foursquare API

- We should use the "explore" endpoint.
- We should pass following parameters: "section=coffee" and "near=Paris". In the response, we are interested in groups.items.categories, groups.items.venue.name, group.items.venue.location. To simplify the problem, we will only use the geographic location to cluster the cafes.

EDIT : I didn't find a Dataset in Paris so I make my report on New York

2. Find the tips - Foursquare API

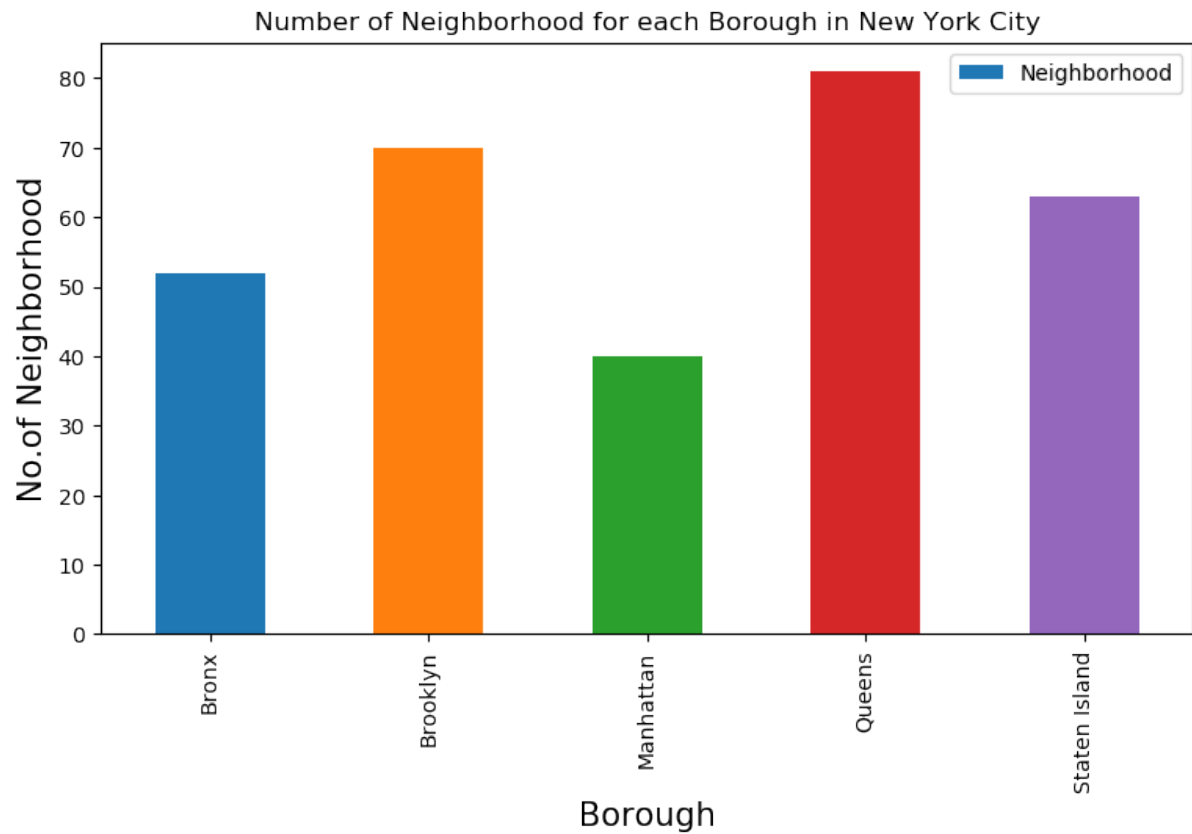
We can use the "geographic mean"(the center of all cafes) as the best location. Based on this location, we use the search endpoint to get all the tips for coffee shops in the surrounding area (1km). For this purpose, the Tips endpoint will be applied.

Methodologie

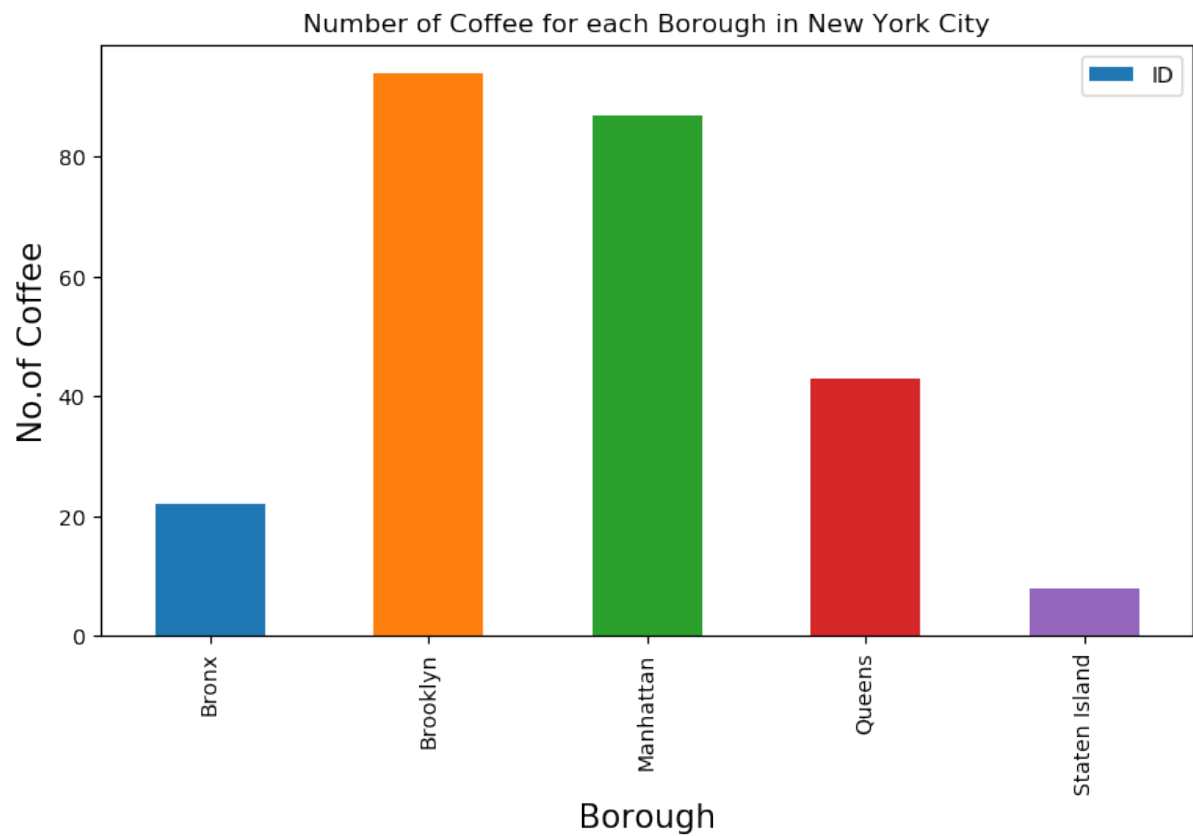
- Collect the new york city data from https://cocl.us/new_york_dataset
- Using FourSquare API we will find all venues for each neighborhood.
- Filter out all venues that are Café.
- Find rating , tips and like count for each Café using FourSquare API.
- Using rating for each coffee, we will sort that data.

Results

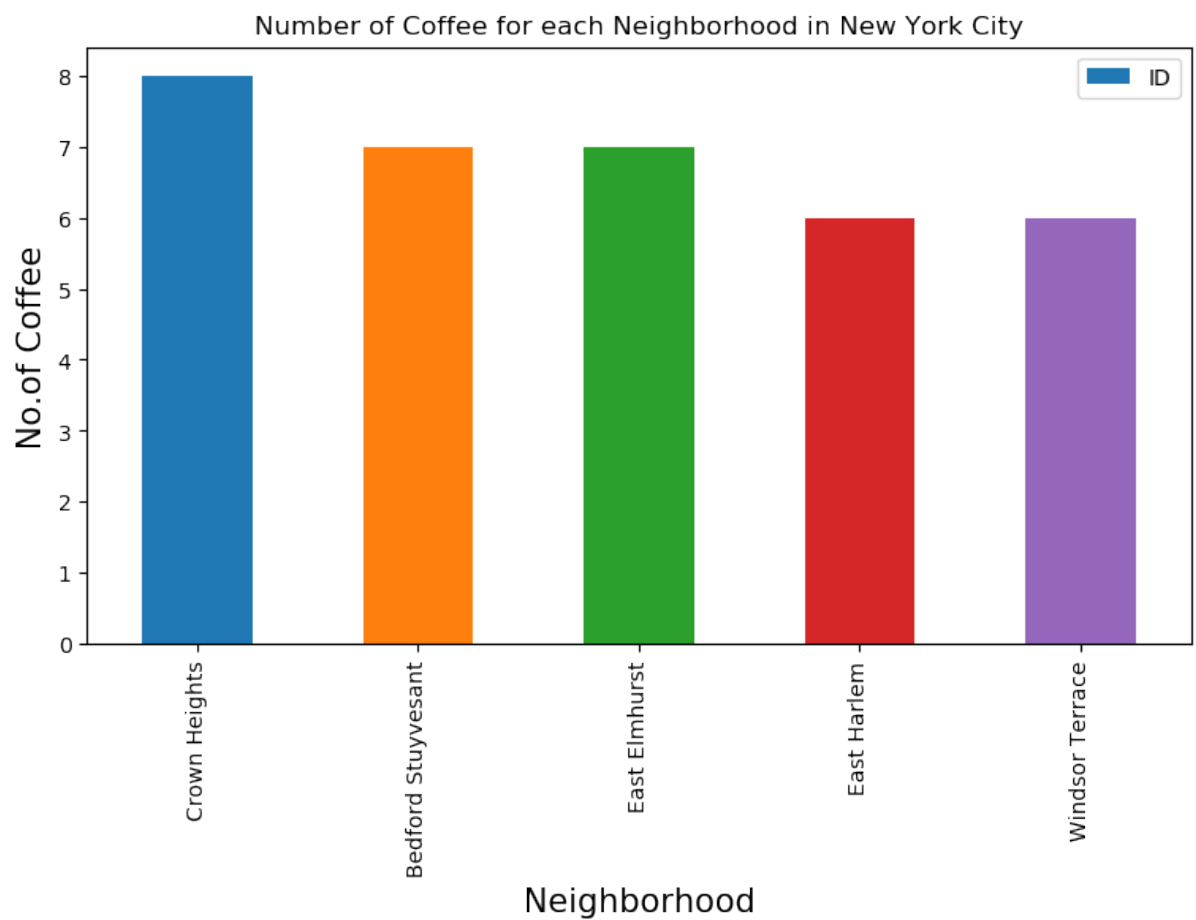
- Queens has the highest number of Neighbourhoods.



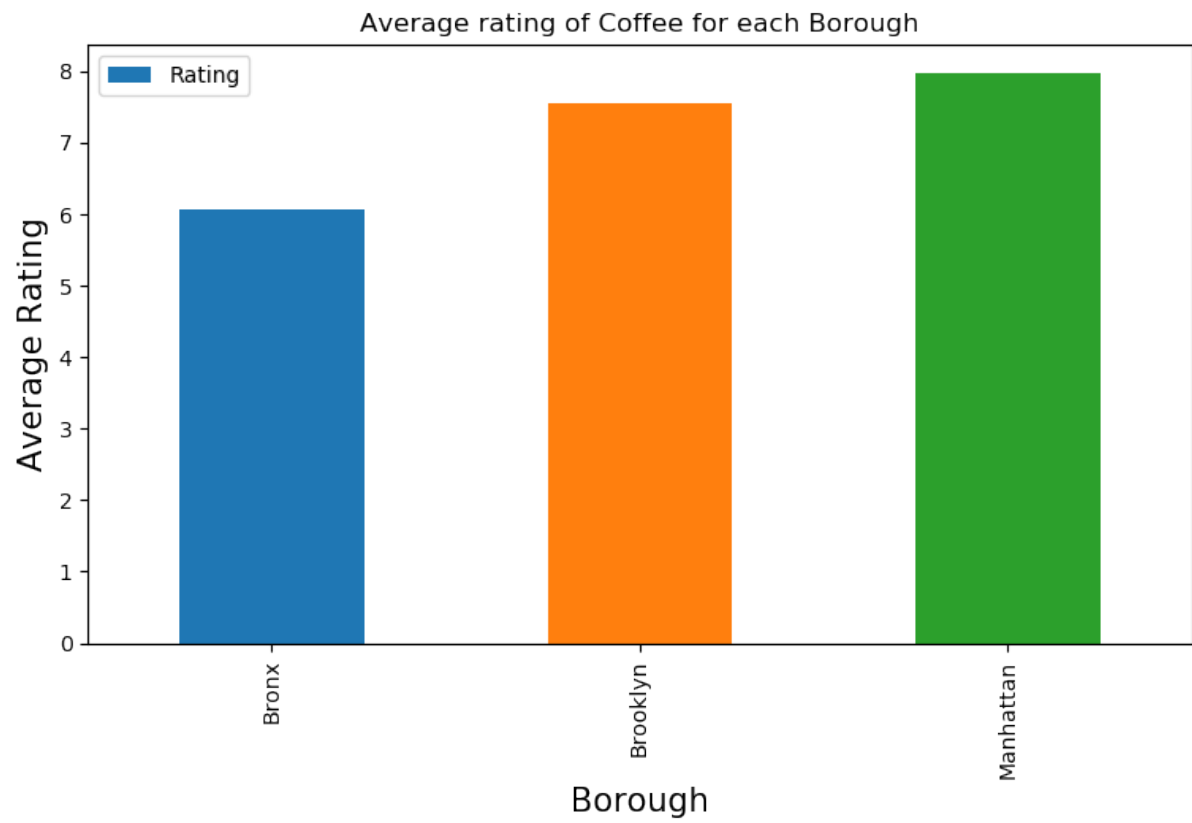
- Brooklyn has the highest number of Coffee.



- Crown Heights in Brooklyn has the highest number of Coffee.



- Manhattan has the highest average rating for Coffee.



Conclusion

- Eastchester (Bronx) is the best neighborhoods for coffee.
- Manhattan is the best place to open a new coffee.

In []: