



Coursera Capstone IBM

Applied Data Science Capstone

Finding opportunities for opening a delivery-based supermarket in New York City, considering a post-COVID scenario



Author: Tiago de Moraes

June 2021



Introduction

- How will the world be like after the Covid-19 crisis is behind us?
- What new habits acquired during the pandemic will persist?
- Most people agree that will persist is the delivery of groceries and general goods, specially in big cities.
- In 2021 this type of convenience might as well persist as a major delivery sales channel for many businesses and will no longer be a differential, but a mandatory revenue stream for any new business to come.





Objective

- Using Data Science methodology and instruments such as Data Analysis and Visualization, the main goal of this Capstone project is to analyze and select the best locations in the city of New York for the deployment of a delivery-focused market;
- This project aims to provide solutions to answer a core business question: Where in New York City should the investor consider to install a delivery-based supermarket.



Required Datasets

- Geospatial data of the city of New York, containing the neighborhoods and boroughs information;
- Latitude and longitude coordinates of those neighborhoods for plot purposes and to get local venue data.
- Venue data, particularly data related to nearby markets, in order to perform further analysis of the neighborhoods.



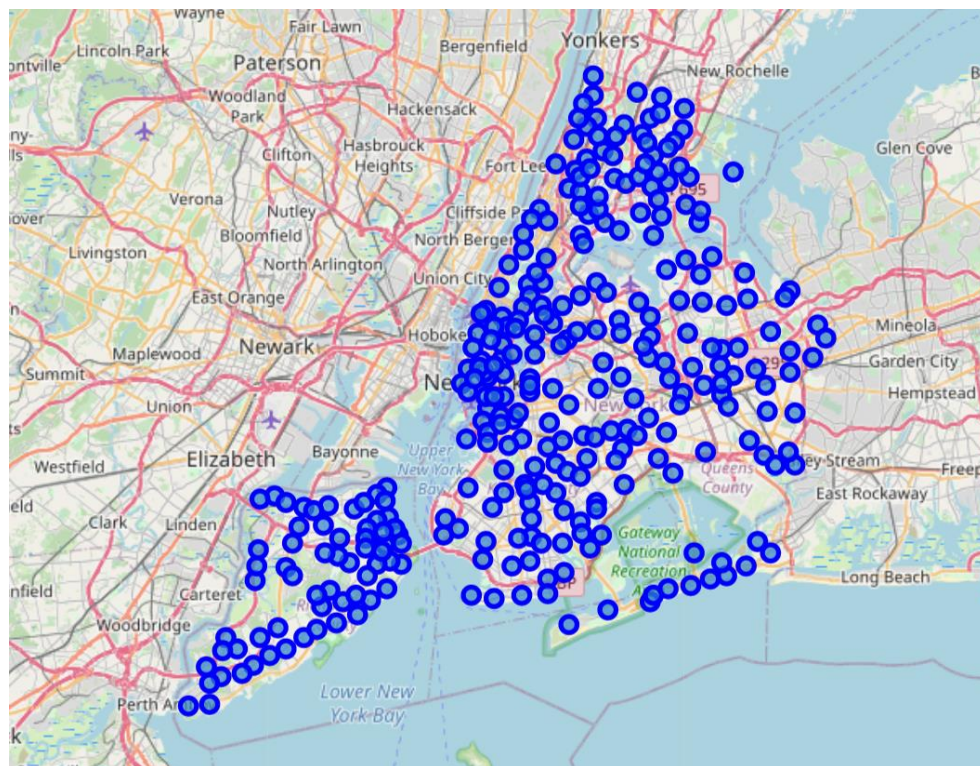


Methodology

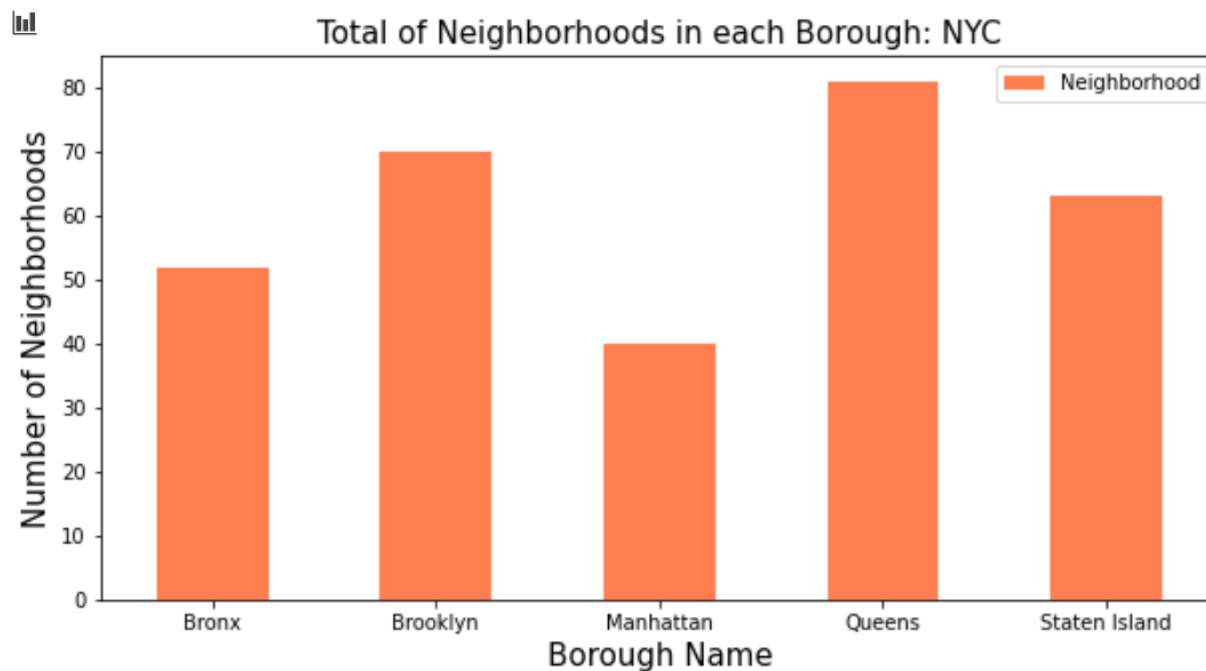
- Data will be collected from https://cocl.us/new_york_dataset, wrangled and then processed into a Pandas dataframe.
- FourSquare API will be used to locate all venues and then filtered by markets in New York. Ratings, tips, and likes by users will be counted and added to the dataframe, and then data will be sorted based on rankings.
- Data be will be visually assessed using Python libraries.
- Maps will be also generated for a more pleasant visual analysis
- k-means algorithms will be used for clueterization of spatial data

Results Discussion

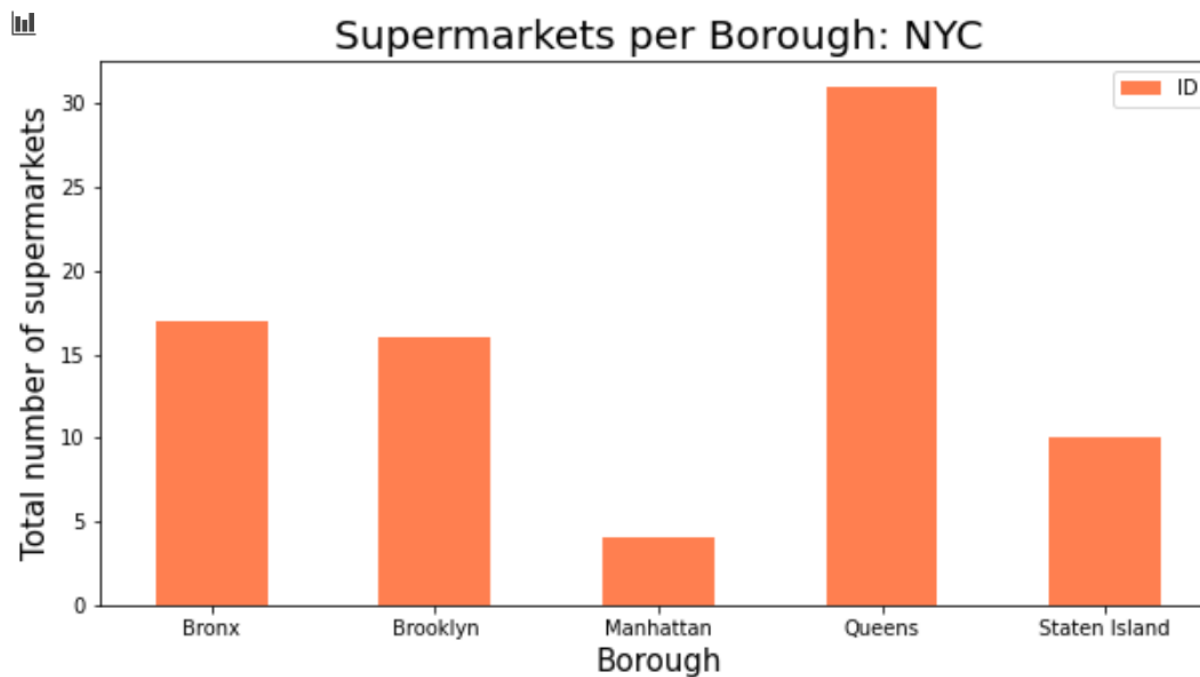
- Queens is the Borough with the most neighborhoods in New York City.
- When plotting the number of supermarkets available in each Borough, it becomes clearer that there is a huge discrepancy between the number of supermarkets in Queens and others Boroughs, such as Manhattan, Brooklyn and Bronx



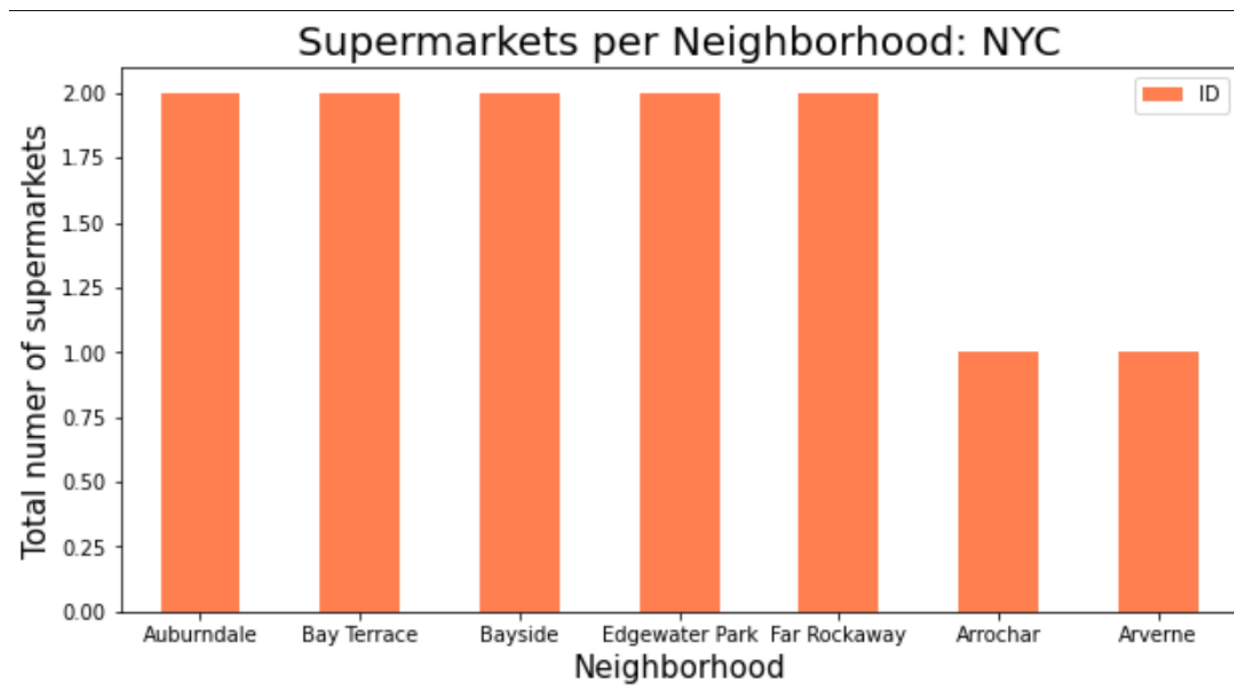
Results Discussion



Results Discussion



Results Discussion



Results Discussion

- Auburndale, Bay Terrace, Bayside, Eggewaterpark and Far Rockaway are the neighborhoods with the highest number of supermarkets.
- The most interesting piece from this data shows that supermarkets are far away from each other, showing that even in the Borough with the most number of Supermarkets (Queens) and its neighborhood most populated with markets (Auburndale), you will only find 2 supermarkets nearby.
- Brooklyn and Bronx are the best candidates for the installation of the proposed delivery-based market.





Conclusions

- Considering the post-Covid scenario, this final project explored the best locations for a delivery-focused market throughout New York City;
- Data showed what was already suspected, that central areas, such as Manhattan, which have a less residential profile, don't have as many supermarkets available as other areas, such as Queens and Brooklyn;
- The main takeaway from this project is that there is still much more room for many residential areas, in special Bronx and Brooklyn, to expand to a more decentralized market model;
- The post-covid world may points towards to a more decentralized market model, which this project shows that there are a tremendous potential to expand such model in suburban areas, specially Bronx and Brooklyn.