# Capstone Project - The Battle of the Neighborhoods

### Applied Data Science Capstone by IBM/Coursera

## Table of contents



## Introduction: Business Problem

In this project we will try to find the best location for to opening an coffee shop business that caters to senior citizens. Specifically, this report will be targeted to stakeholders interested in opening a **Senior Citizen Coffee Shop** in **New York**, New York. Since there are lots of coffee shops in New York City and surrounding five boroughs we will try to detect **locations that are most commonly visited with coffee shops**. We are also particularly interested in **areas with the top demographic of people who drinks coffee in the vicinity**. We would also prefer locations **as close to city center as possible**.
We will use our data science powers to generate a few most promising business locations based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

## Data

*Based on definition of our problem, factors to we will use are:*

- Number of existing coffee shops in the city of any type of coffee shop
- Number of and distance to coffee shops in the city, if any
- Distance of neighborhoods from city center
- New York Senior Population, Geography of Aging, Current and Projected Numbers

*Following data sources will be needed to extract/generate the required information:*

- Number of coffee shops and their type and location in every neighborhood will be obtained using **Foursquare API**
- Demographic of people by age group that consumes the most coffee will be obtained using **https://www.statista.com/**
- Candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using **Google Maps API reverse geocoding**

- Coordinate of New York center will be obtained using **Google Maps API geocoding**
- Blue print of Growing Senior Population in New York using **https://comptroller.nyc.gov/**

# Methodology

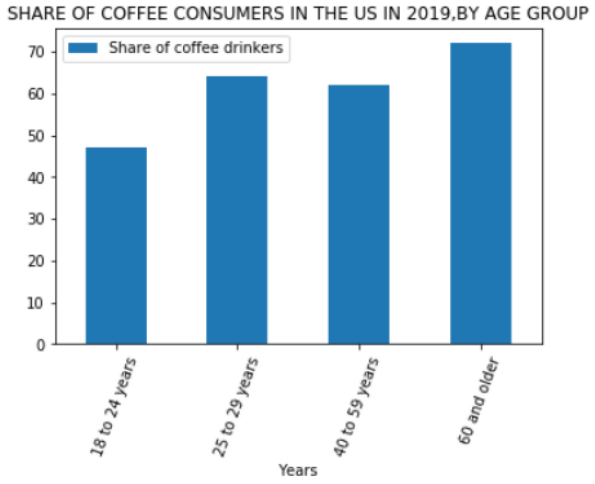
As a database, I used GitHub repository this study. The Data has the main components of current coffee shops, Boroughs, Demographic and Population of Seniors who drink coffee, Latitude and Longitude information of the city.

I utilized the Foursquare API to explore the all coffee shops in New York and boroughs and segment them with the limit as 30 venues. Here is a head of the list Venues name, street, and latitude and longitude information from Forsquare API.

Below is a sample of venues provided by Foursquare.

| | name | location/address | location/lat | location/lng |
|---|---|---|---|---|
| **0** | Stumptown Coffee Roasters | 212 Pacific St | 40.688942 | -73.992141 |
| **1** | East One Coffee Roasters | 384 Court St | 40.681128 | -73.996526 |
| **2** | NYC Love Street Coffee | NaN | 40.706538 | -74.009628 |
| **3** | Gregorys Coffee | 80 Broad St | 40.704341 | -74.011857 |
| **4** | Gregorys Coffee | 100 Wall St | 40.705200 | -74.007090 |
| **5** | Coffee Project NY | 78 Rockwell Pl | 40.687751 | -73.979387 |
| **6** | Gregorys Coffee | 42 Broadway | 40.706110 | -74.012850 |
| **7** | For Five Coffee Roasters | 1 Liberty Plz | 40.709554 | -74.010576 |
| **8** | Periscope Coffee On John Street | 158 John Street | 40.706390 | -74.004480 |
| **9** | For Five Coffee Roasters | NaN | 40.712240 | -74.014915 |
| **10** | Allegro Coffee Company | 238 Bedford Ave | 40.716160 | -73.959530 |
| **11** | Joe Coffee | 102 Hicks St | 40.698673 | -73.994358 |
| **12** | Proof Coffee Roaster | 65 Nassau St | 40.709476 | -74.008546 |
| **13** | Partners Coffee | 125 N 6th St | 40.718012 | -73.959174 |
| **14** | Coffee Cart C0549 | NaN | 40.707532 | -74.008876 |
| **15** | Periscope Coffee On Pearl Street | 123 Pearl Street | 40.704558 | -74.009284 |
| **16** | D'Amico Coffee Roasters | 309 Court St | 40.683593 | -73.995307 |
| **17** | Nick's Coffee Truck | 200 Water St | 40.707458 | -74.004318 |
| **18** | Coffee Project | 239 E 5th St | 40.727153 | -73.989527 |

Demographic by age coffee drinker's data. Provided by statista.com



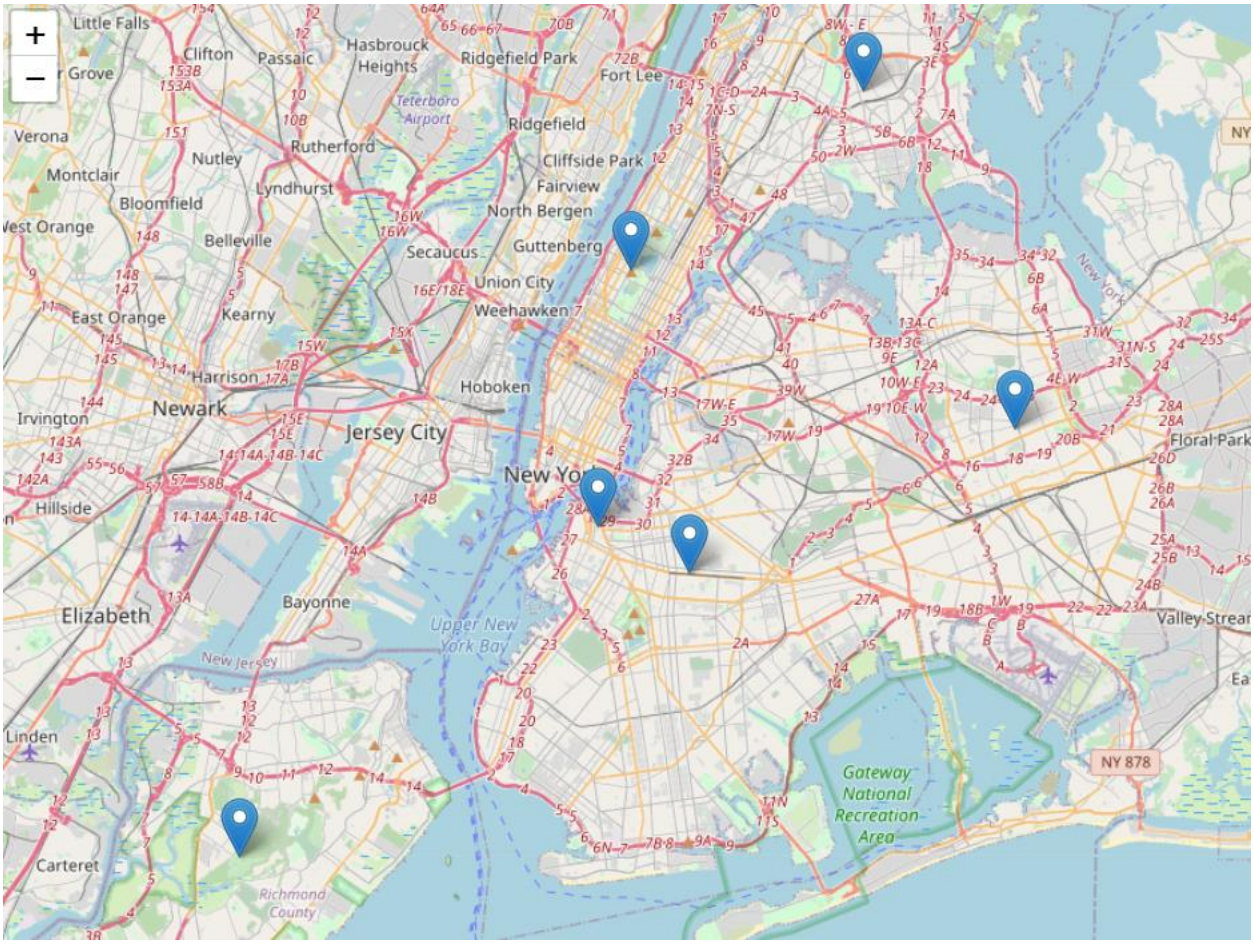

We use matpot to generate a Bar chart showing the demogrhapic of coffee drinkers by age group provided by Statista

The Geography of Ney York City seniors that live in all corners of the five boroughs provided by Comptroller.NYC.Gov

| | BOROUGH | TOTAL 65+ | TOTAL POPULATION (2019) | PERCENT 65+ | LOCATION/LAT | LOCATION/LNG |
|---|---|---|---|---|---|---|
| **0** | BROOKLYN | 326,955 | 2,636,735 | 12.40% | 40.678178 | -73.944158 |
| **1** | QUEENS | 322,803 | 2,339,150 | 13.80% | 40.728224 | -73.794852 |
| **2** | MANHATTAN | 240,100 | 1,644,518 | 14.60% | 40.783060 | -73.971249 |
| **3** | BRONX | 165,921 | 1,455,444 | 11.40% | 40.844782 | -73.864827 |
| **4** | STATEN ISLAND | 71,184 | 474,558 | 15.00% | 40.579532 | -74.150201 |

## ANALYSIS

Let's perform some basic explanatory data analysis and derive some informative info from our raw data collected. We then use python folium library to visualize geographic details of seniors 65+ of the surrounding Boroughs




## RESULTS AND DISCUSSION

Our analysis shows although there is a great number of Coffee shops in the New York and surrounding boroughs area of interest, there are high pockets of Coffee shops Density fairly close to the New York City center. The highest concentration of Coffee shops was detected in New York City and North in Manhattan. So we will focused our analysis to the areas of the South corresponding with the Staten Island Borough. Directing our attention to this more narrow area of interest we created a dense grid of location, the Candidate of location has a dense population of seniors and with 2 or less coffee shops in this zone area. The even though this zone is the less populated of demographic of Seniors upon our Analysis, the results show this is the ideal candidate location for senior coffee shop would be in city of Richmond Town which is center of Staten Island.

## CONCLUSION

The purpose of this was to identify coffee shops close to center with the most populated seniors that drink coffee neighborhoods/boroughs surrounding New York City in order to aid

stakeholders in narrowing down the search for optimal location for a new senior coffee shop. By calculating coffee shop density distribution From **Foursqaure** data and **Folium Library** we have first identified general boroughs that justify further analysis and then generated extensive collection locations to satisfy some basic requirements. Final decision on optimal coffee shops will be made by stakeholders based on specific characteristics of neighborhoods and locations.