



NYC's next coffee shop

Exploring the city's caffeine-friendly neighborhoods

Pratik Mukherjee

Stakeholders



Entrepreneurs



Venture Capitalists



Caffeine lovers

Problem Statement

- Advocate neighborhoods for opening a coffee shop in the **Manhattan** borough of **New York City** that are most optimal for business profitability and operational effectiveness.

Data Acquisition

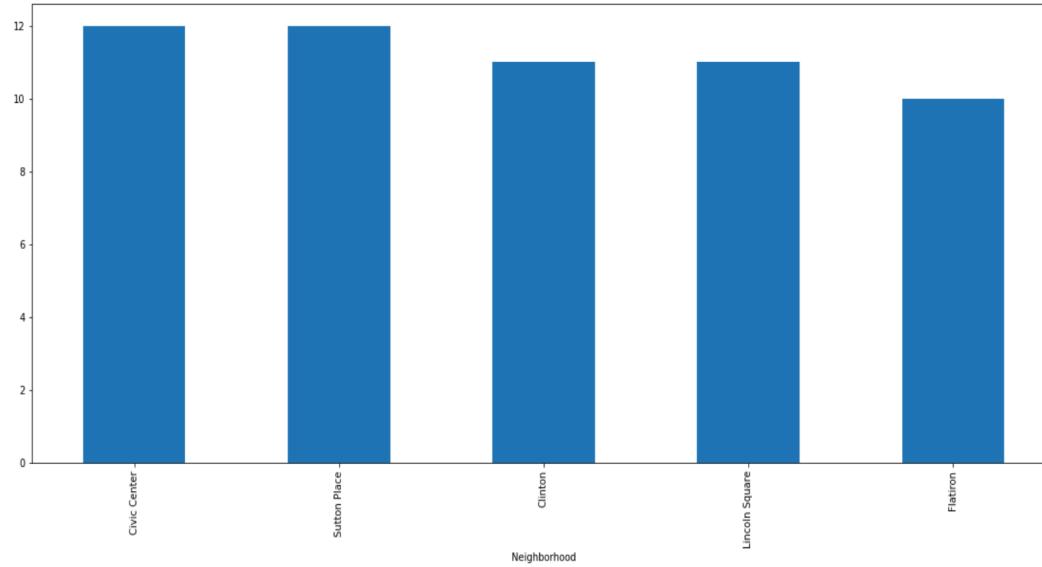
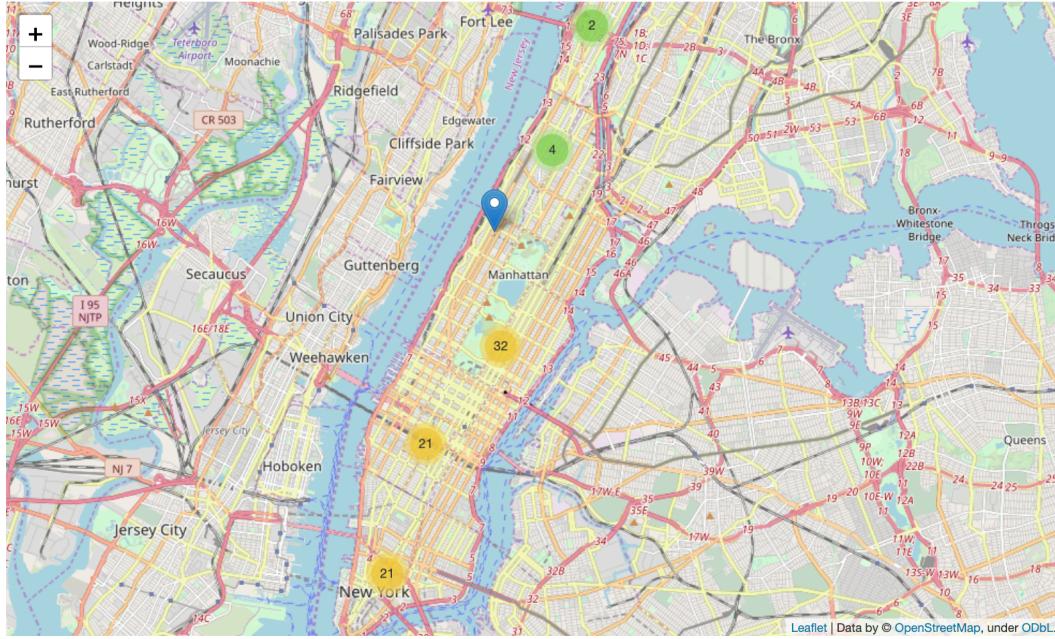
- Data about **neighborhoods, latitude and longitude** was collected from NYC Open Data repository (Link: <https://data.cityofnewyork.us/City-Government/Neighborhood-Names-GIS/99bc-9p23>)
- **Latitude and Longitude** of New York City were collected using the **geopy** library
- **Venue** data for the different neighborhoods were collected from **FourSquare** (Link: <https://api.foursquare.com/v2/venues/explore>)

Data Cleansing

- Venue Categories were classified into **Customers** and **Competitors**
- **Customers** = ['College Theater', 'Gym / Fitness Center', 'Gym / Fitness Center', 'Cooking School', 'Tourist Information Center', 'Tourist Information Center', 'Music School', 'Community Center', 'Community Center', 'Weight Loss Center', 'Weight Loss Center', 'School', 'College Academic Building', 'College Academic Building', 'Garden Center', 'Garden Center', 'High School', 'College Bookstore', 'Recreation Center', 'Recreation Center', 'Paper / Office Supplies Store', "Doctor's Office", 'College Arts Building', 'Office', 'Medical Center', 'Medical Center', 'Spiritual Center', 'Spiritual Center']
- **Competitors** = ['Café', 'Tea Room', 'Cafeteria', 'Bubble Tea Shop', 'Coffee Shop', 'College Cafeteria', 'College Cafeteria']

Data Cleansing (contd.)

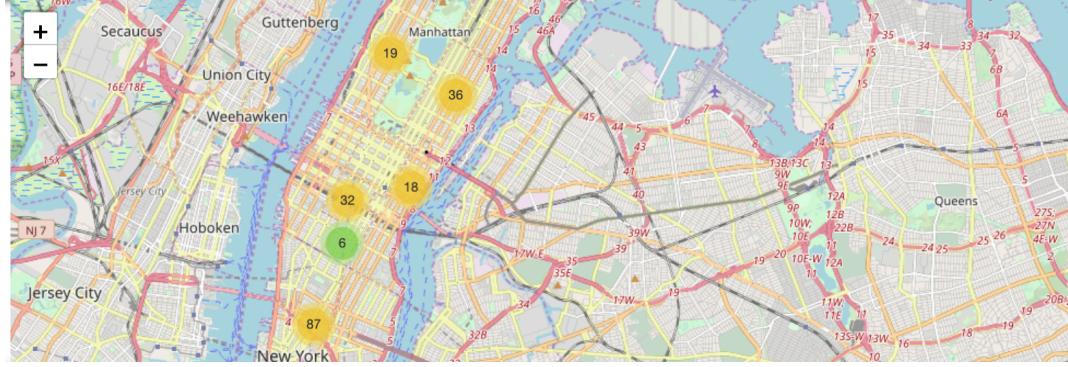
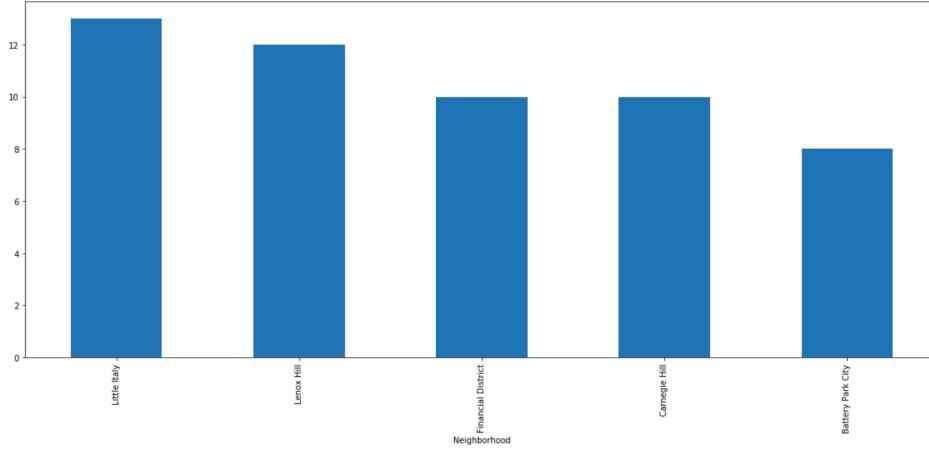
- Records with borough not in **Manhattan** were filtered out.
- Data-frame created with **Neighborhood**, **Latitude**, **Longitude**, **Customers** and **Competitors**
- **Customers:** Venue count where venue category in **Customers**
- **Competitors:** Venue count where venue category in **Competitors**



Neighborhoods by Customers

- Top 5 Neighborhoods:-
- **Civic Center**
- **Sutton Place**
- **Clinton**
- **Union Square**
- **Flatiron**

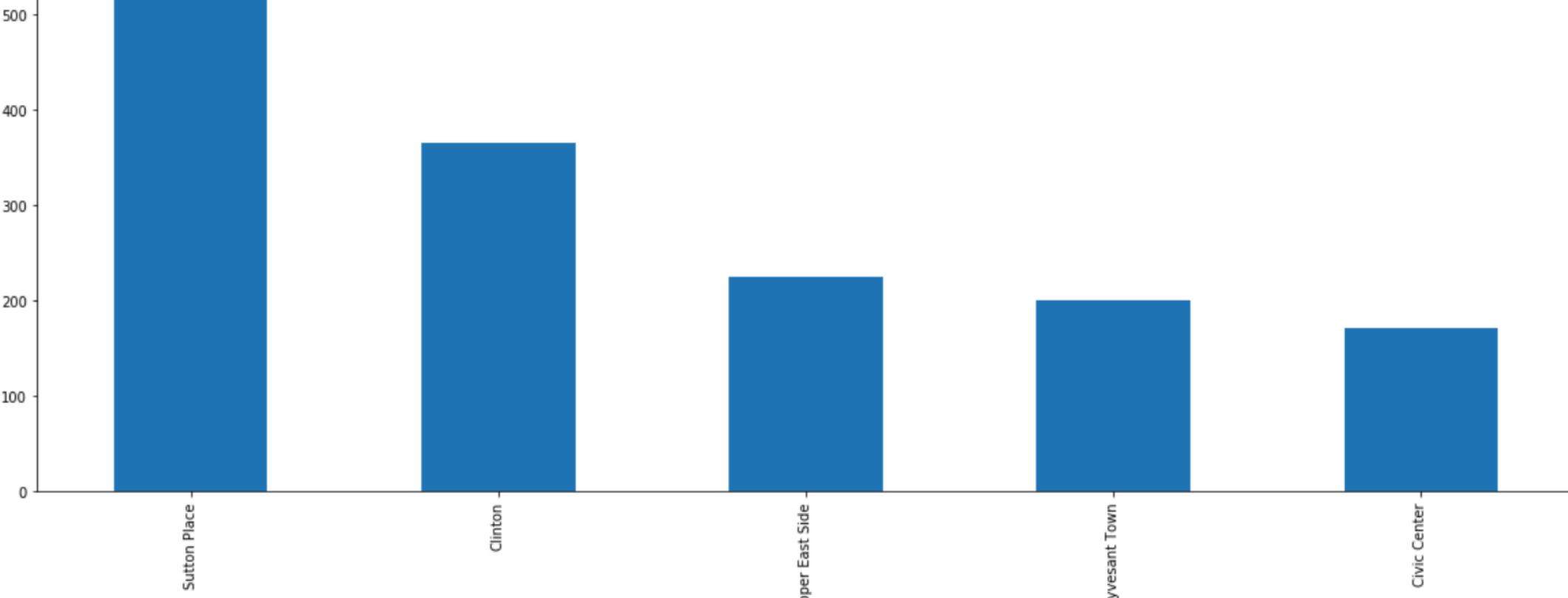
Neighborhoods by Competitors



- Top 5 neighborhoods:-
 - **Little Italy**
 - **Lenox Hill**
 - **Financial District**
 - **Carnegie Hill**
 - **Battery City Park**

Customer_Comp_{etitor} Ratio

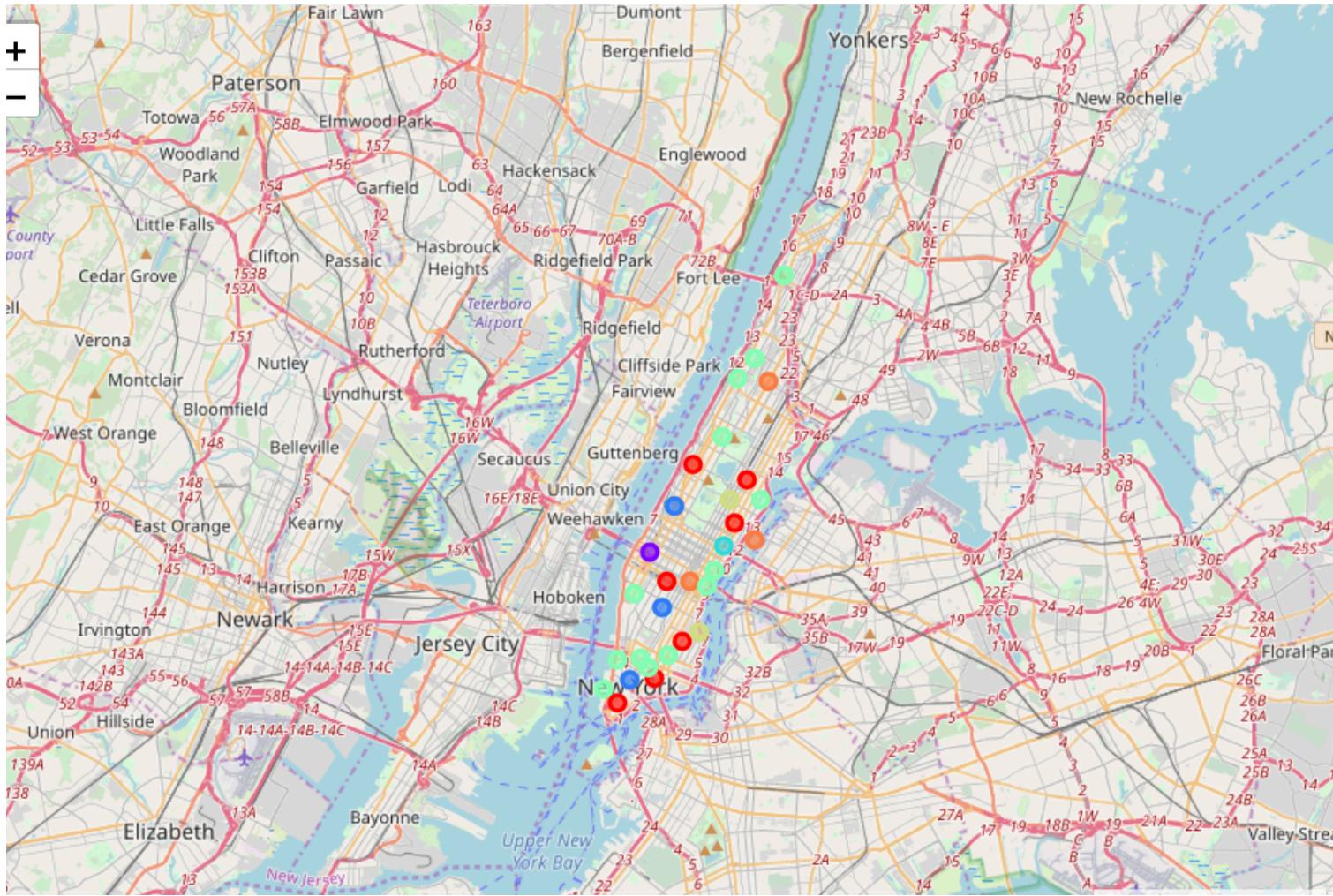
- New column called **Customer_Comp_{etitor}** was added to the data-frame.
- $$\text{Customer_Competitor} = \frac{\text{Number of Customers} * 100}{\text{Number of Competitors}}$$



Neighborhoods by Customer_Compiler

- Top 5 neighborhoods:-
 - **Sutton Place**
 - **Clinton**
 - **Upper East Side**
 - **Stuyvesant Town**
 - **Civic Center**

Clusters Map



CLUSTER LABEL	NEIGHBORHOODS	CUSTOMER_COMPETITOR RATIO
0	Carnegie Hill Chinatown East Village Financial District Lenox Hill Midtown South Upper West Side	60 - 75
1	Clinton	366
2	Civic Center Flatiron Lincoln Square	142 - 171
3	Sutton Place	600
4	Chelsea Hamilton Heights Little Italy Manhattan Valley Manhattanville Noho Soho Tribeca Tudor City Turtle Bay Washington Heights YorkVille	12 - 42
5	Stuyvesant Town Upper East Side	200 - 225
6	Central Harlem Murray Hill Roosevelt Island	85 - 100

Clusters Results

Discussion

- Neighborhoods that have a proportionate distribution of commercial enterprises, tourists' spots, and residential areas are best suited to set up a coffee shop.
- Neighborhoods that are at the geographical extremes of the city are not as well suited as they lack accessibility and public footfall.
- Tourist and business neighborhoods are poorly suited as the number of existing coffee shops exceeds the number of customers.

Future Improvements

- Incorporate factors like land prices, neighborhood rent, subway stops, number of residential building in the decision-making process.
- Use techniques like Heat Maps for better visualization process.