Coursera Capstone Project: The Battle of Neighborhoods

Introduction / Business Problem:

New York City is an economic and cultural hub that has become one of the world's artistic epicenters. As such, a strong community of creatives are drawn to New York for the opportunities it presents for them to build a career. From a business perspective, artist's need spaces to network and exhibit their work and this is typically fostered through gallery representation. For this project, we are going to use our data resources to help a potential gallery owner determine the best location to open a new gallery space. By leveraging our data, we will give the gallery owner valuable insight into areas of popularity, and market saturation to give them the best likelihood for success.

Data:

We will extrapolate the data from the New York Neighborhood data set found at:

https://cocl.us/new_york_dataset

	Borough	Neighborhood	Latitude	Longitude
301	Manhattan	Hudson Yards	40.756658	-74.000111
302	Queens	Hammels	40.587338	-73.805530
303	Queens	Bayswater	40.611322	-73.765968
304	Queens	Queensbridge	40.756091	-73.945631
305	Staten Island	Fox Hills	40.617311	-74.081740

We will partner this data with geo-spatial data information available on Foursquare.com to gain insights into:

- The ideal location to open the gallery
- Areas of art appreciation based on gallery popularity
- Market saturation by neighborhood
- Market opportunity by neighborhood
- Existing gallery density based on geography

A link to the Jupiter Notebook where all the code was written and compiled is available here:

https://gist.github.com/jamisonharper/ca222e9ac1d81ee18ad6c088041aec4b

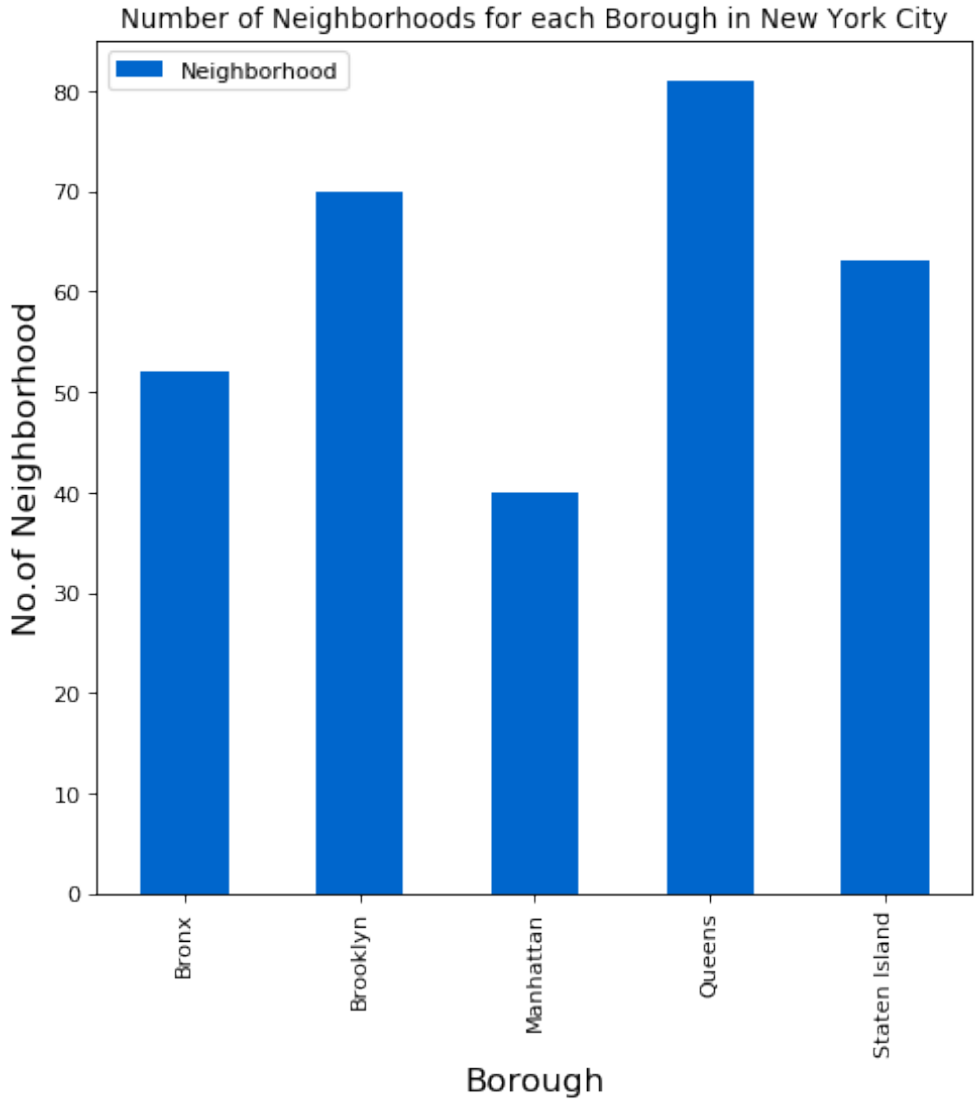
Methodology:

Initial Data

We begin by analyzing the initial data comprising the Boroughs and associated Neighborhoods found in New York City. This data was pulled from the following site:

https://cocl.us/new_york_dataset

Next, we converted the data into panda data frames to make it easier to analyze. Our first exploration into the data was to create a bar chart visualizing our initial data set.



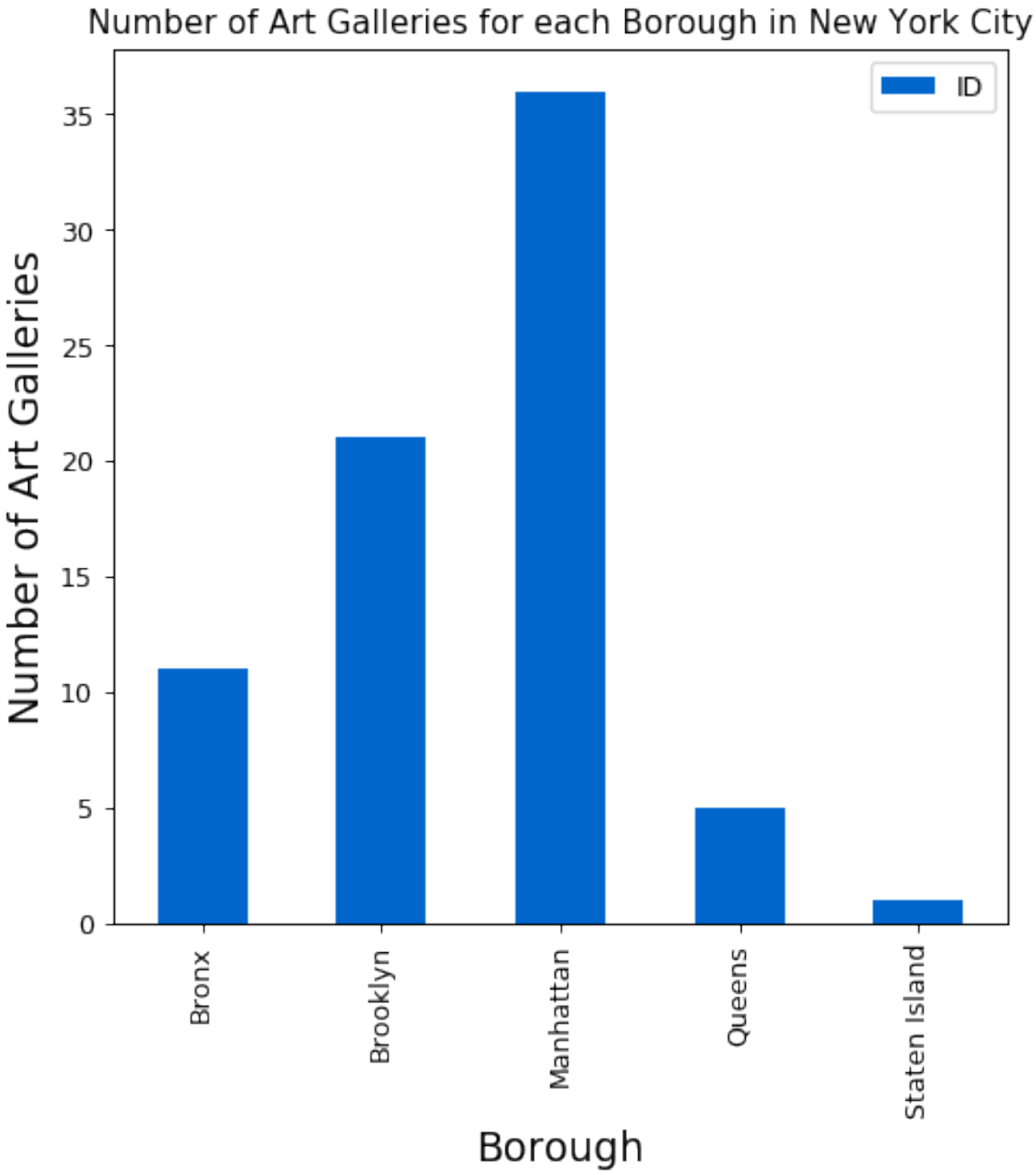
Foursquare

By leveraging the Foursquare platform, we are able to determine valuable information about existing galleries located within the city. Using the venue query for Art Galleries we produce the following list:

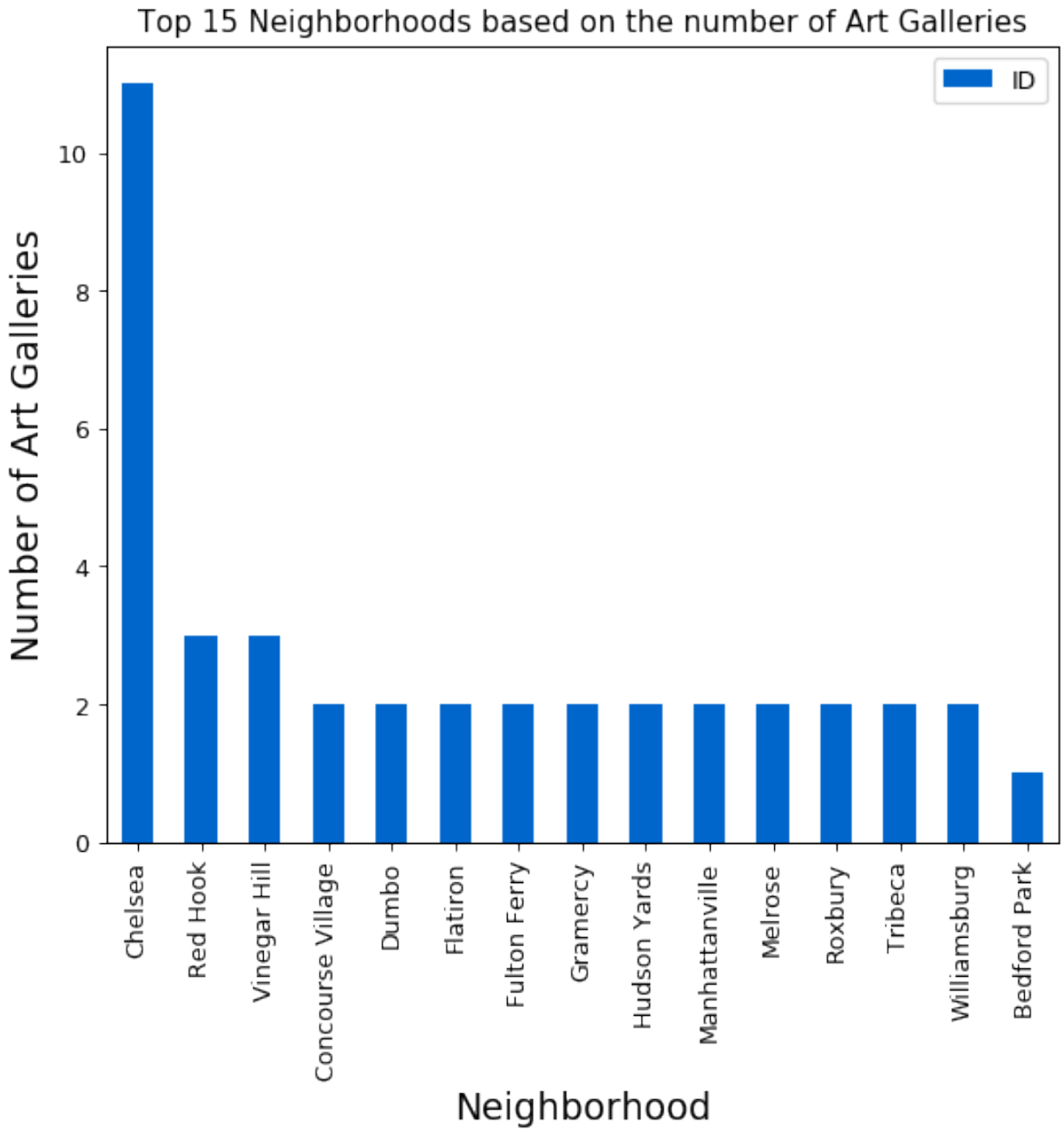
Art Galleries in

Fieldston, Bronx : 1
Riverdale, Bronx : 1
City Island, Bronx : 1
Bedford Park, Bronx : 1
Melrose, Bronx : 2
Mott Haven, Bronx : 1
Longwood, Bronx : 1
Morrisania, Bronx : 1
Flatbush, Brooklyn : 1
Crown Heights, Brooklyn : 1
Williamsburg, Brooklyn : 2
Bedford Stuyvesant, Brooklyn : 1
Cobble Hill, Brooklyn : 1
Red Hook, Brooklyn : 3
Gowanus, Brooklyn : 1
East New York, Brooklyn : 1
Clinton Hill, Brooklyn : 1
Midwood, Brooklyn : 1
Manhattanville, Manhattan : 2
Central Harlem, Manhattan : 1
Upper East Side, Manhattan : 1
Lenox Hill, Manhattan : 1
Roosevelt Island, Manhattan : 1
Clinton, Manhattan : 1
Midtown, Manhattan : 1
Chelsea, Manhattan : 11
Greenwich Village, Manhattan : 1
East Village, Manhattan : 1
Lower East Side, Manhattan : 1
Tribeca, Manhattan : 2
Little Italy, Manhattan : 1
Soho, Manhattan : 1
Manhattan Valley, Manhattan : 1
Gramercy, Manhattan : 2
Long Island City, Queens : 1
St. George, Staten Island : 1
Noho, Manhattan : 1
Civic Center, Manhattan : 1
New Lots, Brooklyn : 1
Concourse Village, Bronx : 2
Sutton Place, Manhattan : 1
Hunters Point, Queens : 1
Flatiron, Manhattan : 2
Fulton Ferry, Brooklyn : 2
Vinegar Hill, Brooklyn : 3
Dumbo, Brooklyn : 2
Roxbury, Queens : 2
Hudson Yards, Manhattan : 2
Queensbridge, Queens : 1
Completed constructing the Art Galleries data set

Continuing our analysis through visualization, we produce the following chart:

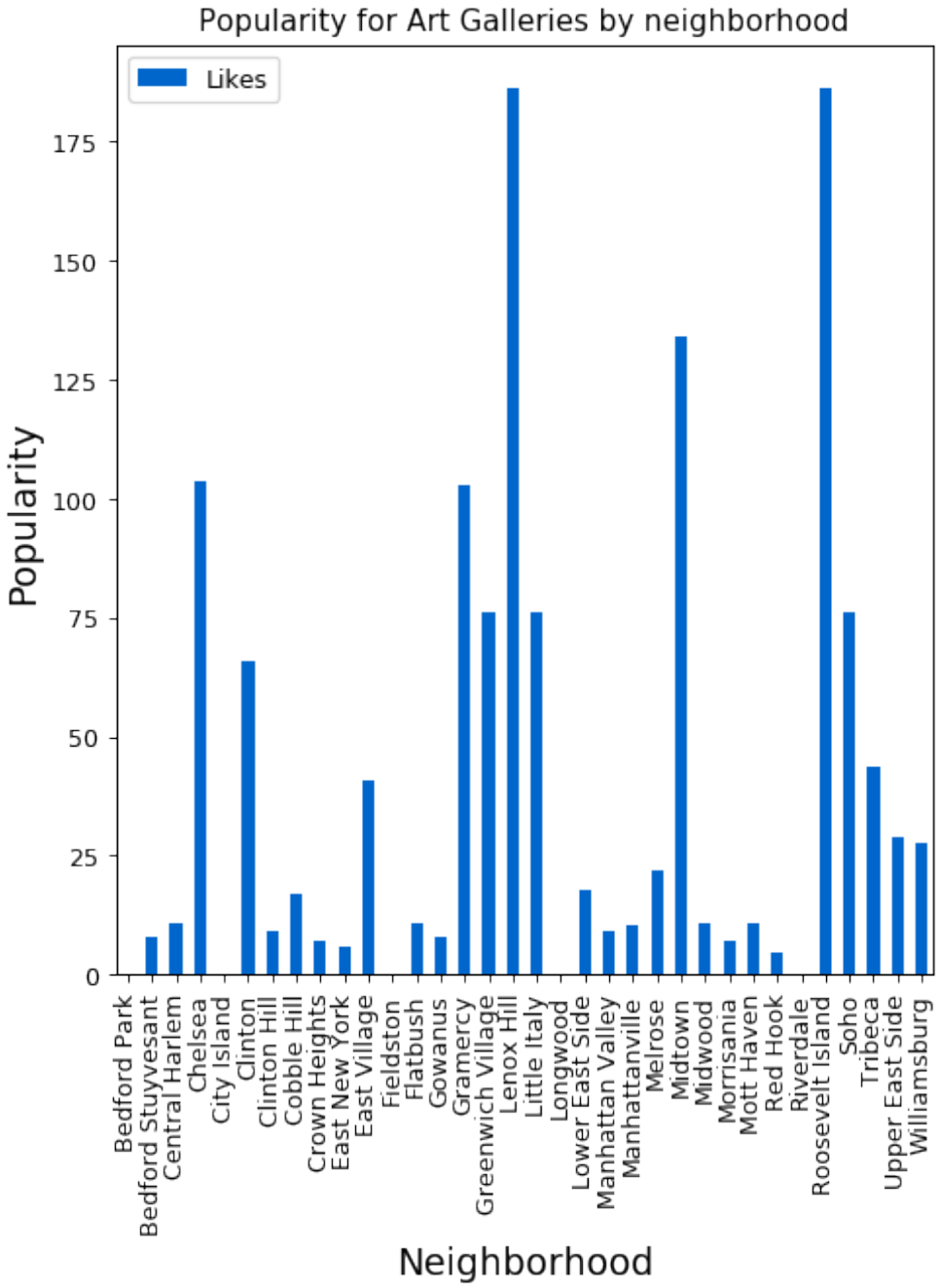


This clearly establishes Manhattan as the borough with the greatest gallery exposure. We can take this one step further by analyzing the neighborhoods within New York that contain art galleries by number. The following chart represents the top 15 neighborhoods:



Chelsea emerges as the neighborhood with the most galleries.

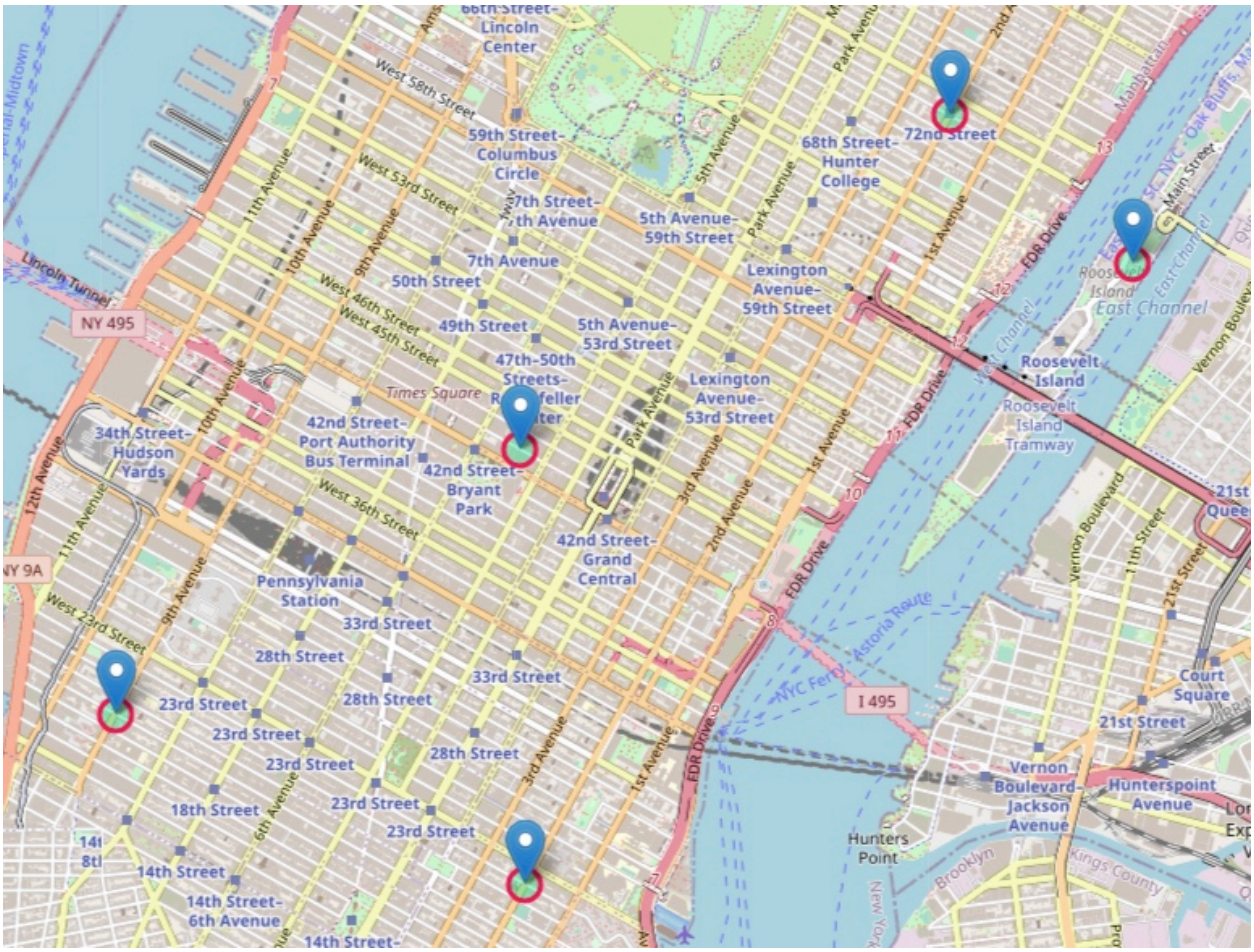
Continuing to examine the data available through Foursquare, we can extrapolate user ratings to determine gallery popularity by analyzing the most liked galleries and their geographies. The following chart gives us the simple visualization:



There is more data here than we really care to use in our final determinations, so we can customize this into the top 5 neighborhoods.

	Neighborhood	Popularity
16	Lenox Hill	186.000000
29	Roosevelt Island	186.000000
23	Midtown	134.000000
3	Chelsea	103.818182
14	Gramercy	103.000000

This provides some context, but we can draw further conclusions by representing these neighborhoods on a map.



Results:

Through our visualizations and data analysis, we were able to accurately determine the Borough with the most existing galleries: Manhattan.

Within Manhattan, Chelsea was the neighborhood with the most galleries.

Through Foursquare queries we could also calculate a popularity metric and apply this to these geographies. As a result we determined the following 5 neighborhoods to have the highest

By cross-referencing our Art Galleries list, we draw the determinations that of our top 5 neighborhoods, two are located in Chelsea - which already has the highest number of galleries. Linux Hill and Roosevelt Island have the highest popularity by user rating, and minimal competition.

Discussion:

During the review of the data through the Methodology, a few of the initial assumptions were corrected or invalidated. For example, it was assumed that visualizing the data of the borough information would help provide insights into the location preferences. However, since Manhattan was over-represented by such a wide margin compared to the other boroughs, the data was skewed. It became more important to dissect Manhattan itself on the neighborhood level to gain insight. Additionally, it was never anticipated that a popularity metric could be applied. After analyzing this information, it helped to determine where support for art was strong and yet there was room for additional galleries.

Boroughs that were under-represented, after reviewing the data, then felt like areas of risk - rather than areas of opportunity. Having an established market that can serve many galleries became more assuring and the data helped clarify this insight. Thoughtful analysis of the data provided the guidance to seek an area that had community support and capacity for growth.

Conclusions:

Based on the Data Analysis and Visualization using Folium and Matplotlib, we have successfully used our data to create graphical visualizations in the forms of charts and maps to gain insights and understanding. This process draws the following conclusions:

The Best Neighborhoods for opening an Art Gallery in New York are:

- Lenox Hill

- Roosevelt Island

- Midtown

- Chelsea (Manhattan)

- Chelsea (Staten Island)

We have determined the areas which lack Art Galleries, and also see that Manhattan is overwhelming the area of key support for existing galleries.

We have determined within Manhattan the neighborhoods that have favorable markets by existing gallery popularity measurements.

We have also been able to identify Chelsea as an area of potential market saturation due to the relationship of existing galleries compared to surrounding neighborhoods.

Lenox Hill presents itself as a favorable market by being included in Manhattan (so it has existing support), and has low saturation by neighborhood with only one existing gallery in competition.
It shares the highest popularity by user ratings with Roosevelt Island. Lenox Hill is more accessible than Roosevelt Island, which if considered, may inhibit traffic and have an insulated audience.

Lenox Hill's proximity to Chelsea means it can share the benefit of established support while having room for growth and through the analysis of the data presented, may be the best opportunity for establishing a new gallery space.