

# THE BATTLE OF THE NEIGHBORHOODS

**Mets vs. Yankees**



# TABLE OF CONTENTS

<b>INTRODUCTION .....</b>	<b>1</b>
Background .....	1
Rivalry on the field .....	1
pride off the field .....	1
<b>DATA .....</b>	<b>2</b>
Geographical .....	2
Venue .....	2
<b>METHODOLOGY .....</b>	<b>3</b>
Geography .....	3
New York City .....	4
Flushing, Queens .....	5
Concourse, Bronx .....	5
Venues .....	6
Local Venues .....	6
Summarization, Combination, and Comparison .....	6

<b>RESULTS .....</b>	<b>8</b>
Ballpark Proximity .....	8
Venue Categories.....	9
<b>DISCUSSION .....</b>	<b>10</b>
General observations.....	10
Citi Field .....	11
Yankee Stadium.....	12
<b>CONCLUSION .....</b>	<b>14</b>

# INTRODUCTION

## BACKGROUND

There are two Major League Baseball teams in New York. The New York Mets play in the National League, while the New York Yankees play in the American League. Both teams play in their respective league's East division.

## RIVALRY ON THE FIELD

There is a fierce, intra-city rivalry between the fans of these two teams with each group numbering in the millions. The introduction of interleague play in 1997 allowed this rivalry to be settled on the field, as did the "Subway Series" in 2000.

## PRIDE OFF THE FIELD

The intent of this report is to take the rivalry off the field and to the venues surrounding the Mets' and Yankees' ballparks, Citi Field and Yankee Stadium, respectively. Here, more evidence can be found regarding the differences of the two neighborhoods in which the fanbases can take pride and perhaps gain additional bragging rights.

# DATA

## GEOGRAPHICAL

The first data set utilized in the analysis behind this report was geographical data regarding New York City, which was sourced from the newyork\_data.json file available for download at [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset). This data was used to create a series of maps to familiarize the audience with New York City and its boroughs, ultimately showing the relative proximity of Citi Field and Yankee Stadium.

## VENUE

The second data set was acquired from the FourSquare Endpoints API, available from <https://developer.foursquare.com/docs/api/endpoints>. This was actually a series of smaller data sets used to determine the precise locations of the two teams' ballparks, as well as details of the venues surrounding each of them.

# METHODOLOGY

## GEOGRAPHY

Geography was very important to the story that this report was created to tell, so a large amount of the exploratory analysis focused in that area. The intent was to “zoom in” in order to give audience members who may be less familiar with New York City the context they would need to find the story compelling.

## New York City

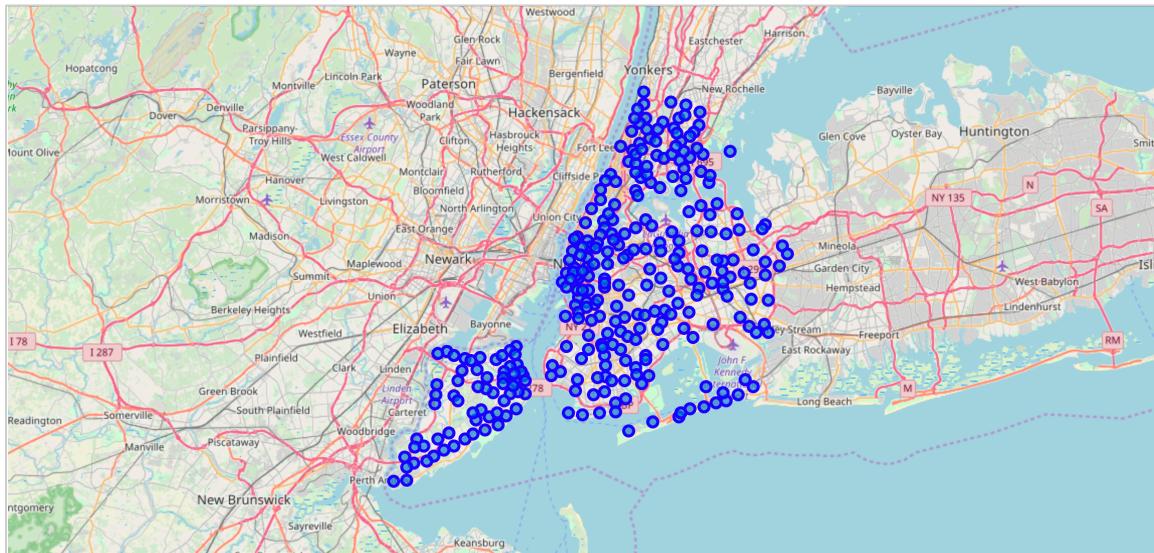


Figure 1

Almost nine-million people call New York City home. These residents are divided among five boroughs:

1. *The Bronx*
2. *Brooklyn*
3. *Manhattan*
4. *Queens*
5. *Staten Island*

Each of these boroughs contains distinct neighborhoods, as shown in Figure 1, totaling dozens throughout the city. However, only two of these neighborhoods are relevant to Mets vs. Yankees.

## Flushing, Queens

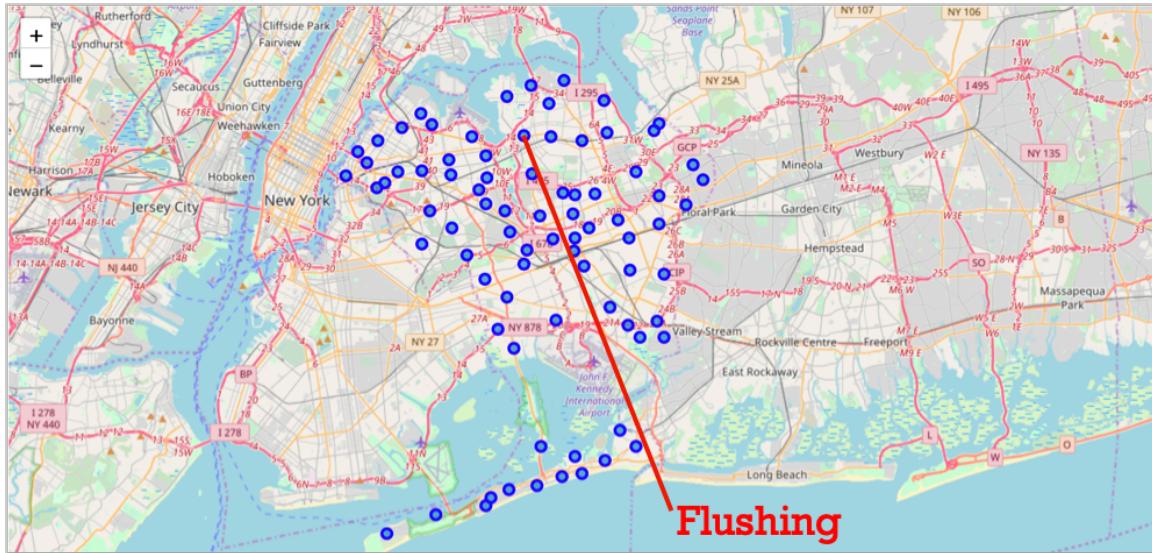


Figure 2

The New York Mets call the neighborhood of Flushing, labeled in Figure 2, in the borough of Queens home.

## Concourse, Bronx

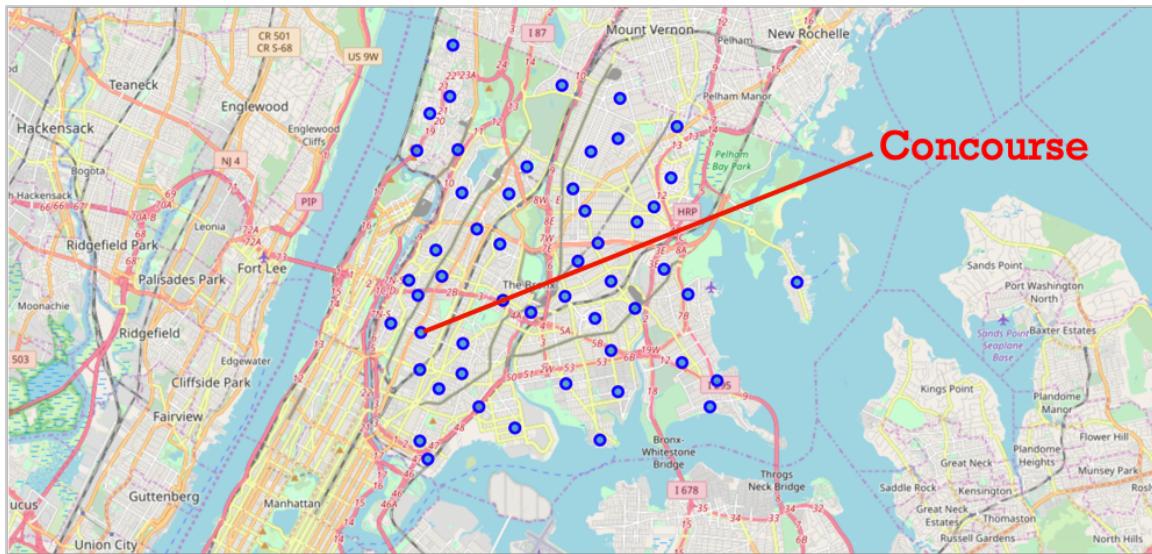


Figure 3

The New York Yankees play in the Concourse neighborhood, called out in Figure 3, of The Bronx. Notice that Queens is just to the north.

## VENUES

The first challenge requiring the Foursquare Endpoints API was to more specifically locate the stadiums where the Mets and Yankees play, so that they could be mapped as well. Requests were submitted for both Flushing, Queens and Concourse, Bronx to gather the top 100 venues, which was assumed and confirmed to contain the baseball stadiums.

### Local Venues

Additional requests were sent to the Foursquare Endpoints API, moving the centers of interest from Flushing and Concourse to the ballparks themselves. One of the challenges here was the size of the stadiums. The original 500m radius was increased in order to return closer to the 100-venue mark and reduce the contribution of venues within the ballparks.

### Summarization, Combination, and Comparison

After moving the venue data pulls into two distinct Pandas dataframes, Table 1 and Table 2, a summary of the venues surrounding each ballpark was created by counting the venues by category.

*Table 1*

	<b>name</b>	<b>categories</b>	<b>lat</b>	<b>lng</b>
<b>0</b>	Citi Field	Baseball Stadium	40.756652	-73.845849
<b>1</b>	Mikkeller Brewing NYC	Brewery	40.757007	-73.844354
<b>2</b>	Mets Clubhouse	Sporting Goods Shop	40.756892	-73.845792
<b>3</b>	Shea Bridge	Bridge	40.757096	-73.845779
<b>4</b>	Jackie Robinson Rotunda at Citi Field	Plaza	40.755501	-73.846272

Table 2

	<b>name</b>	<b>categories</b>	<b>lat</b>	<b>lng</b>
<b>0</b>	Delta SKY360° Suite	Lounge	40.829464	-73.926936
<b>1</b>	Bleacher Creatures	Athletics & Sports	40.829309	-73.925509
<b>2</b>	Yankee Stadium	Baseball Stadium	40.829581	-73.926525
<b>3</b>	NYCFC @ Yankee Stadium	Soccer Field	40.829597	-73.926213
<b>4</b>	New York Yankees Museum	History Museum	40.829622	-73.926192

These dataframes were then combined through a merge, using an outer join, as seen in Table 3. This provided a view of all venue categories, even if a given category were only present at one stadium.

Table 3

	<b>Mets</b>	<b>Yankees</b>
	<b>Category</b>	
<b>American Restaurant</b>	2.0	2.0
<b>Art Museum</b>	NaN	1.0
<b>Arts &amp; Crafts Store</b>	NaN	2.0
<b>Athletics &amp; Sports</b>	NaN	1.0
<b>BBQ Joint</b>	1.0	1.0

This view allowed a direct comparison of the number of each type of venue surrounding Citi Field and Yankee Stadium.

# RESULTS

## BALLPARK PROXIMITY

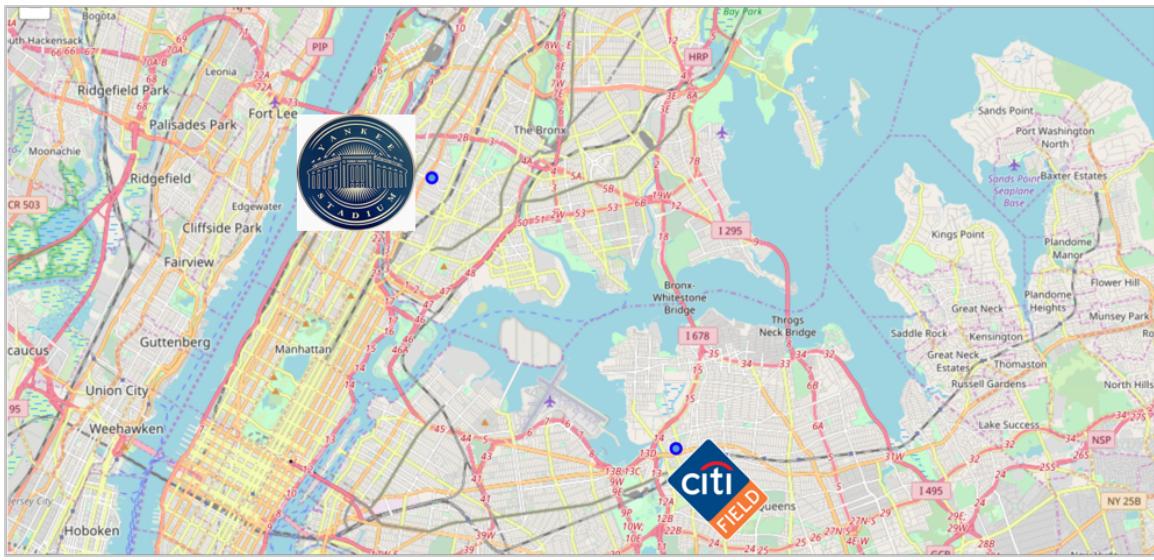


Figure 4

The respective ballparks of the Mets and Yankees, Citi Field and Yankee Stadium, are in very close proximity, as shown in Figure 4. In fact, they are just 9.5 miles apart. This is a large driver of the rivalry between the two fan bases.

## VENUE CATEGORIES

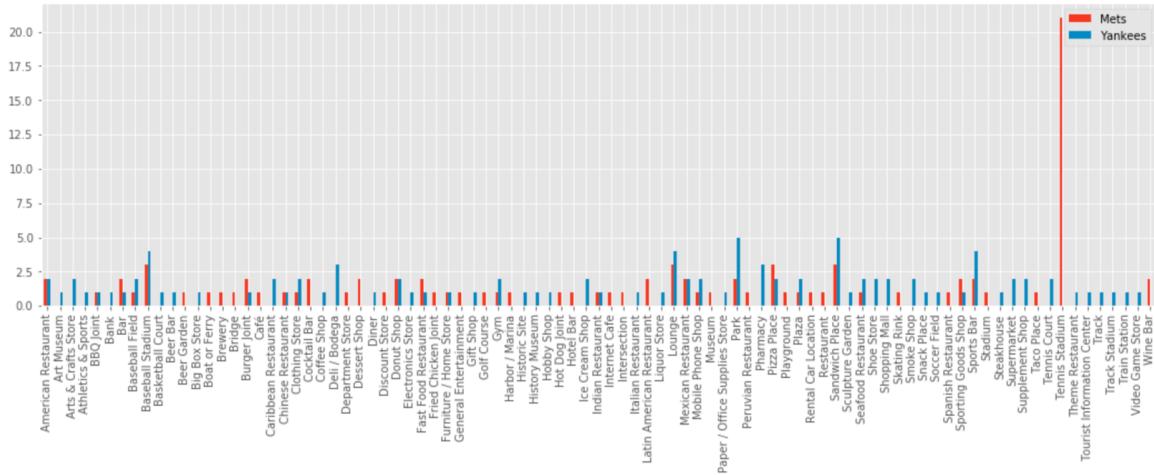


Figure 5

The full dataframe, the head() of which was displayed in Table 3, was fairly lengthy and hard to digest. Thus, a paired bar chart, Figure 5, was used to visualize the number of venues of each category surrounding Citi Field (red) and Yankee Stadium (blue). This made the conclusions much easier to draw.

# DISCUSSION

## GENERAL OBSERVATIONS

One of the surprising things was how many seemingly common categories of venues are in the proximity of only one ballpark. This could be due to factors such as the type of area the stadium is in, i.e. residential vs. industrial. However, one would think that a major entertainment and tourist destination would draw at least one cocktail bar and ice cream shop for example. That did not turn out to be the case.

One disappointment with the project was the apparent lack of easy (and free) access to the reviews of the venues. The analysis would have admittedly been more compelling if there were a quality component as well as a volume one.

## CITI FIELD



There are four categories of venues exclusive to the area surrounding Citi Field with exactly two examples each:

- *cocktail bars*
- *dessert shops*
- *Latin American restaurants*
- *wine bars*

There are also several categories that had a lone exclusive example, e.g. brewery.

However, the most interesting exclusive category, both from a data and general interest standpoint, has to be tennis stadiums with a

whopping 21 venues according to Foursquare. This appears to be a bit misleading, as all 21 are likely part of the United States Tennis Association (USTA) Billie Jean King National Tennis Center, home of the annual US Open tournament. However, the international importance of this set of venues is certainly unique to Flushing and the area surrounding Citi Field.

## YANKEE STADIUM



The Yankee Stadium neighborhood has a distinct advantage in the number of exclusive categories. Delis / bodegas and pharmacies are unique and have three venues each nearby. In addition, seven more exclusive categories have two venues each:

- *arts & crafts stores*

- *ice cream shops*
- *shoe stores*
- *shopping malls*
- *supermarkets*
- *supplement shops*
- *tennis courts*

All of that said, the last such category is somewhat ironic, because it highlights what Yankee Stadium is not near: Anything else of great general interest. A couple local tennis courts don't get to host the US Open.

# CONCLUSION

Clearly there are many sources of pride among the fan bases of both the Mets and Yankees. While the on-field bragging rights are left to others' capable hands to document, this report hopefully provided some lesser known, off-field bragging rights.

Mets fans can point to a small handful of exclusive venue categories but can also take great pride that the USTA National Tennis Center is in their backyard, bringing international exposure on annual basis.

Though Yankees fans do not have such a single destination, they do have significantly more exclusive categories of venues, including some pretty fundamental ones, like pharmacies, supermarkets, and shopping malls.