Data

In this study, I used the New York City Airbnb Open Data posted on https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data. It is in the form of a CSV file. The dataset describes the listing activity and metrics in NYC, NY for 2019. It includes over 40,000 rows and 16 columns as follows:

Variable	**Notes**
id	listing ID
name	name of the listing
host_id	host ID
host_name	name of the host
neighborhood_group	borough
neighborhood	area
latitude	latitude coordinates
longitude	longitude coordinates
room_type	listing space type
price	price in dollars
minimum_nights	amount of nights minimum
number_of_reviews	number of reviews
last_review	latest review
reviews_per_month	number of reviews per month
calculated_host_listings_count	amount of listing per host
availability_365	number of days when listing is available for booking

In addition, I used the dataset that contains the 5 boroughs in New York City and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood. The dataset is readily available at the website (https://geo.nyu.edu/catalog/nyu_2451_34572). I also used the Foursquare location data which I obtained through API calls.