Airbnb Case Study

By :

Laltesh Choudhary

Problem Statement:

Suppose that you are working as a data analyst at Airbnb. For the past few months, Airbnb has seen a major decline in revenue. Now that the restrictions have started lifting and people have started to travel more, Airbnb wants to make sure that it is fully prepared for this change.

The different leaders at Airbnb want to understand some important insights based on various attributes in the dataset so as to increase the revenue such as -

The categorization of customers based on their preferences.

- What are the neighborhoods they need to target?
- What is the pricing ranges preferred by customers?
- The various kinds of properties that exist w.r.t. customer preferences.
- Adjustments in the existing properties to make it more customer-oriented.
- What are the most popular localities and properties in New York currently?
- How to get unpopular properties more traction?

Objectives:

- Understand client preferences and trends in user experience for AirBnb in New York City.
- Provide recommendation for fresh acquisitions and customer experience enhancement.

Data Assumptions:

- Assumed that the data prior to the COVID-19 period was achieving the desired goals.
- Airbnb wants to continue to expand in NYC and it has no plans to expand to other territories

Tools Used:

- We have used Python and python libraries for analyzing the dataset. Tableau report is used for reporting purpose.

Steps followed for Airbnb Case Study

1. **Importing the Necessary Libraries:**
   ```
   # Numerical and Data Analysis
   import pandas as pd
   import numpy as np

   # Data Visualization
   import seaborn as sns
   import matplotlib.pyplot as plt
   import plotly.express as px

   #import the warnings and display size.
   import warnings
   warnings.filterwarnings("ignore")
   pd.set_option("display.max_rows", None,"display.max_columns", None)
   ```
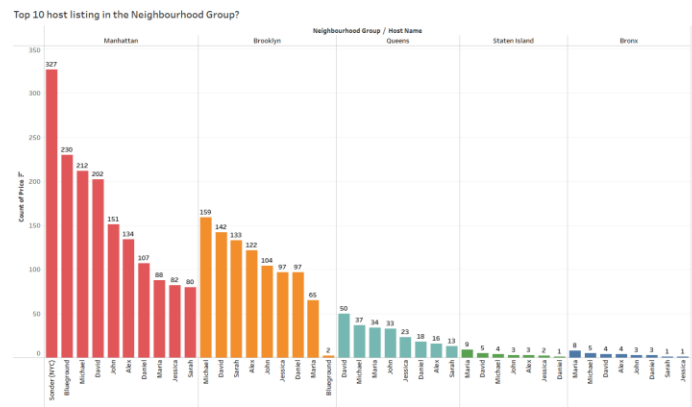
2. **Load the Airbnb Dataset file:**
   ```
   # We need to read the Airbnb CSV file and save it in a dataframe df
   df = pd.read_csv("AB_NYC.csv")
   ```
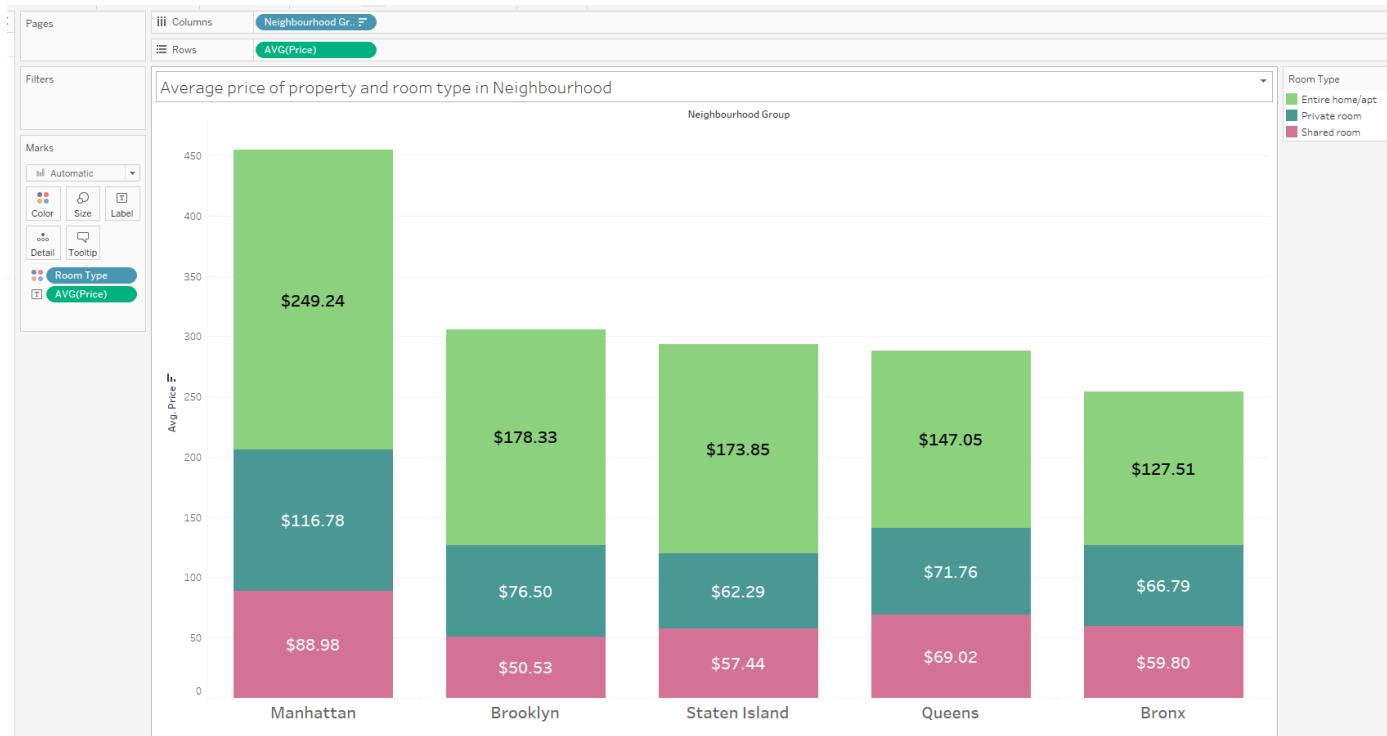
Steps followed for Airbnb Case Study-Continued

3. Handling missing Values and Outliers:

- We have checked the missing values and the outliers in the data
- last_review, reviews_per_month, name and host_name has missing values.
- Last_review and reviews_per_month has NaN values and we have replaced it with Zero

4. Data Analysis:

- Exploring the data by plotting different types of charts using Tableau as well as python
- We have created a bar chart for Top 10 host listing in the Neighborhood Group. As per the below graph we have observed that "**sond**er" has the maximum number of listings 327.
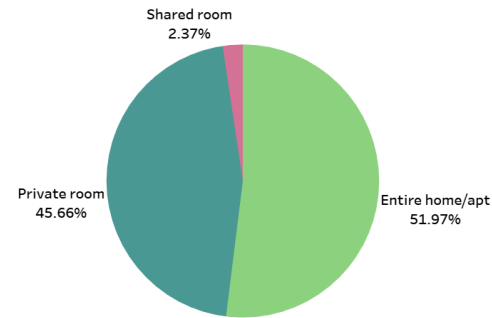
Steps followed for Airbnb Case Study-Continued

- From the below graph we can state that Manhattan has the highest range of prices for the listings with average of $249 followed by Brooklyn. Staten Island and Queens has $173 & $147 respectively. Among all Bronx has the least average of $127
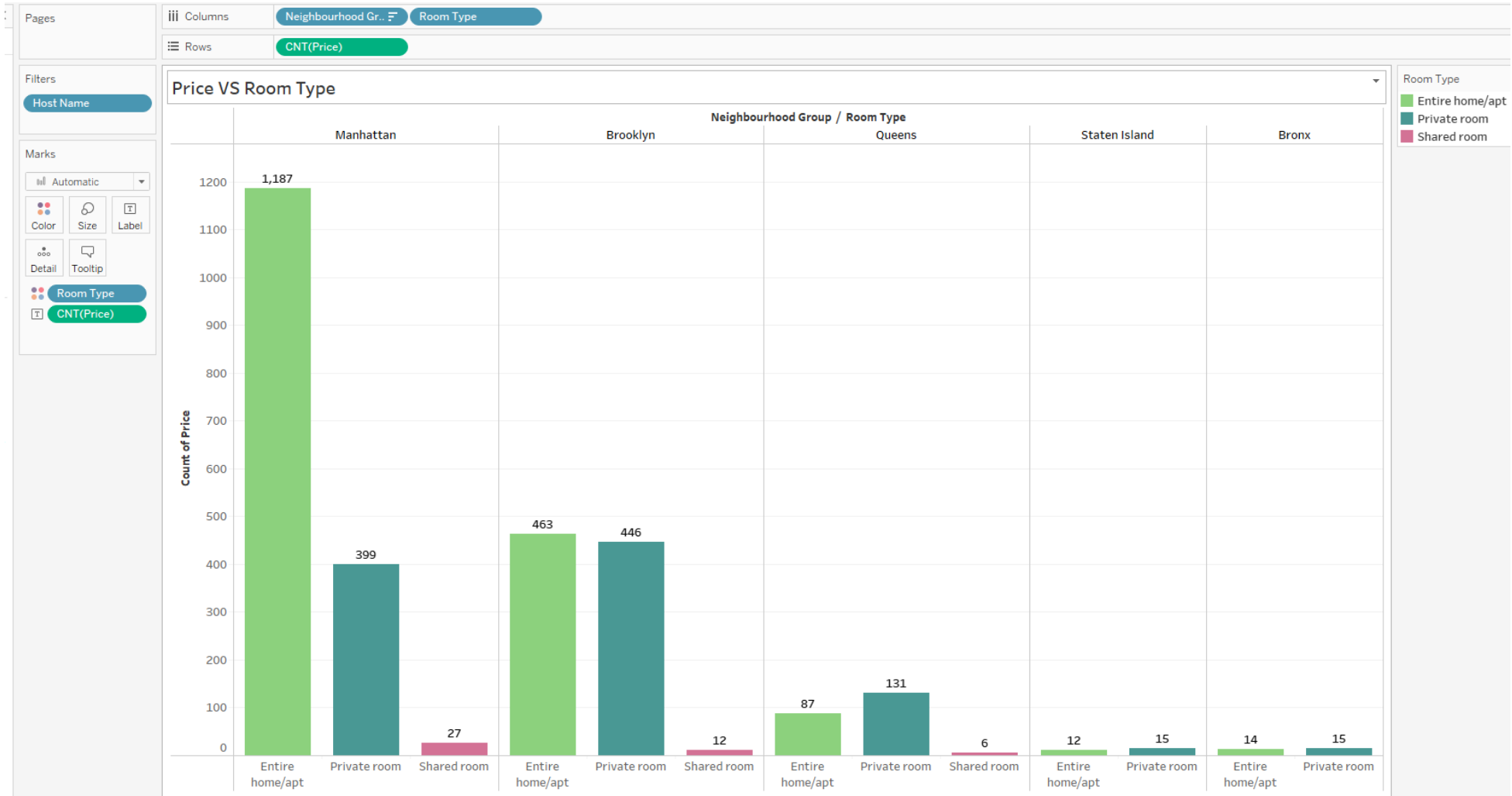
Steps followed for Airbnb Case Study-Continued

- **Bivariate Analysis in Tableau:**
  1. **Customer preference of three property types:** As per the below chart Entire home/apt and private room are preferred over shared rooms. However, the shared rooms accounts only for 2.4% of the total listed properties.
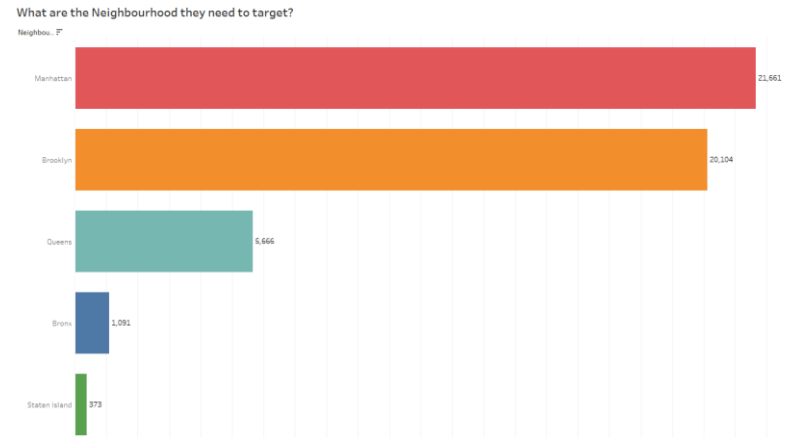

  2. **Prices vs Room Type in Neighborhood Group:** Manhattan has the most expansive properties followed by Brooklyn while the Bronx and Staten Island are the cheapest.

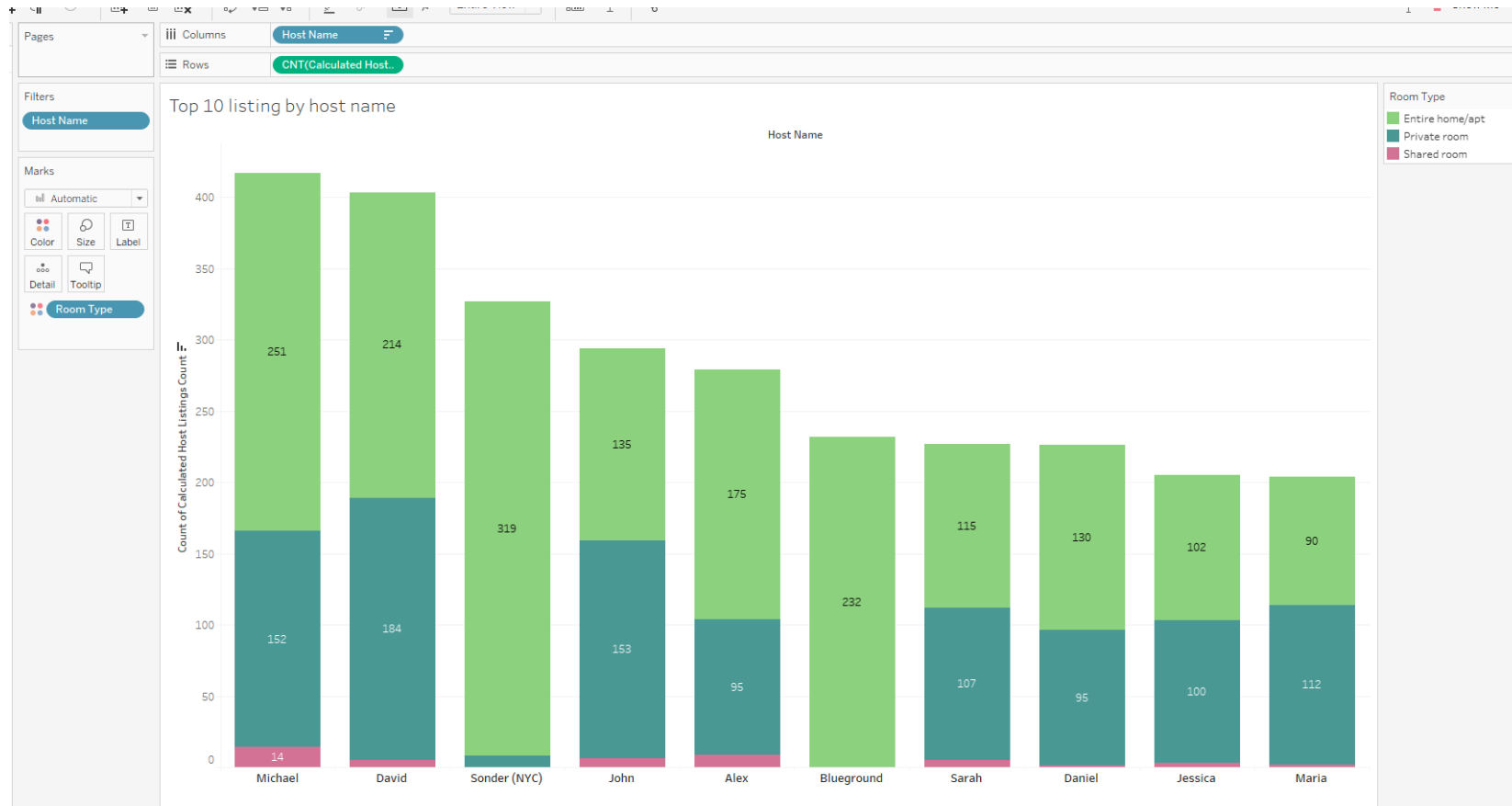# Steps followed for Airbnb Case Study-Continued

Steps followed for Airbnb Case Study-Continued

3. **What are the Neighborhood that they need to Target?** :The majority of listings are primarily located in the neighborhood groups of Manhattan and Brooklyn, and the price range in these two areas is noticeably different from the other three neighborhood groups.

What are the Neighbourhood they need to target?
Neighbou.. F

Manhattan	21,661
Brooklyn	20,104
Queens	5,666
Bronx	1,091
Staten Island	373

4. Top 10 Listing By Host Name: The top 10 hosts primarily offer Entire Home/apt and Private Room rentals, with their prices ranging from 140 to 300.
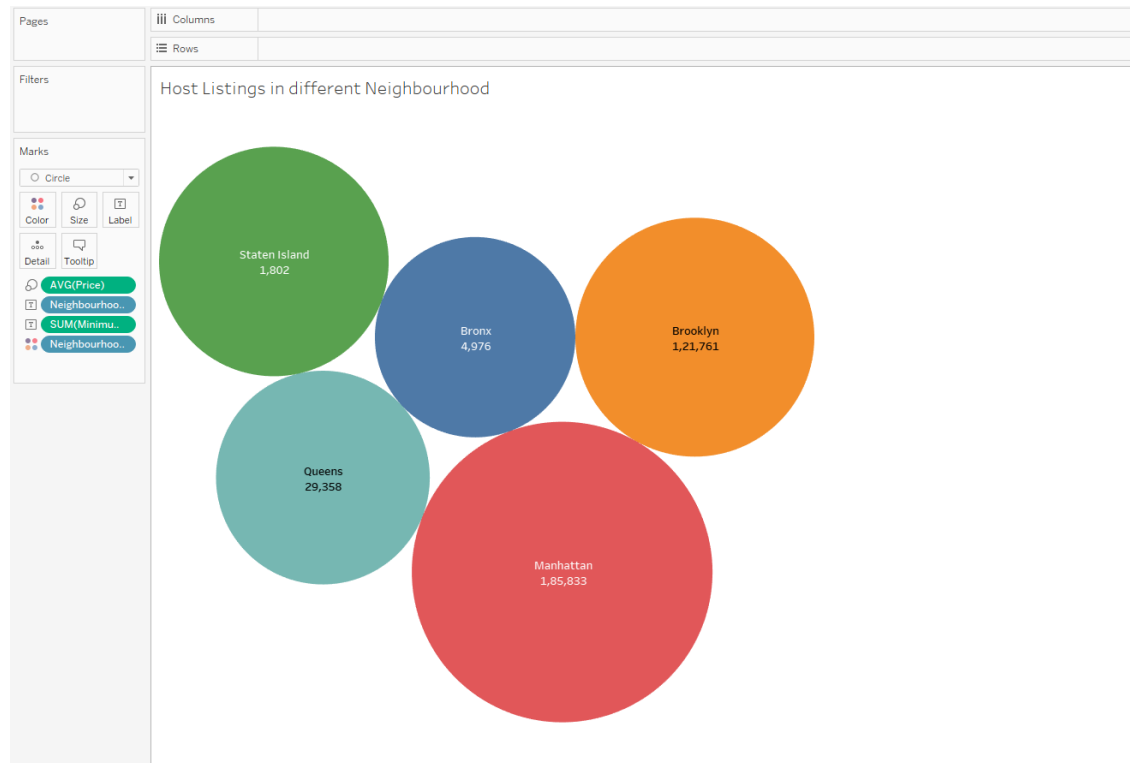
Steps followed for Airbnb Case Study-Continued

Steps followed for Airbnb Case Study-Continued

5. **Average price of Neighborhood**: Top two Neighborhoods Fort Wardsworth and Woodrow falls under Staten Island Neighborhood Group. Manhattan has Tribeca, Battery Park City, Flatiron District and NoHo in the top 10.

Steps followed for Airbnb Case Study-Continued

6. **Average price of Neighborhood**: We have drawn bubble chart in order to conclude the hosts listings in different neighborhoods groups. Manhattan has the highest count of host listings around 21661 and Staten island has the least count of host listings.

Steps followed for Airbnb Case Study-Continued

7. **Scatter Plot:** Scatter plot depicts the latitude and longitude of availability of rooms
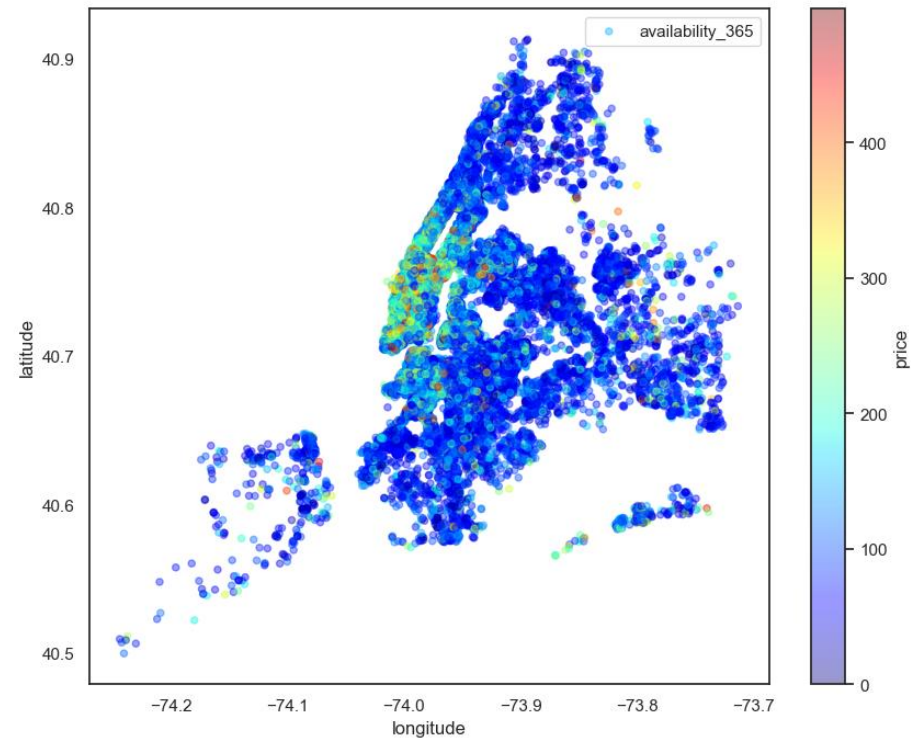
Key Insights & Recommendation:

- It appears to be a very rich dataset with a variety of columns that allowed us to do deep data exploration on each significant column presented. First, we have found hosts that take good advantage of the platform and provide the most listings; we found that our top host has 327 listings.
- After that, we proceeded with analyzing neighborhood listing and what areas were more popular than another.
- Lastly, we found the most reviewed listings and analyzed some additional attributes. Overall, we discovered a very good number of interesting relationships between features and explained each step of the process.