exploring-airbnb-nyc-market

November 11, 2024

1 Analyzing the Airbnb NYC 2019 Market

1.0.1 Project Overview

In this project, we will conduct an in-depth exploratory data analysis on the Airbnb NYC 2019 dataset to uncover trends and patterns in rental listings. Additionally, we will create a dynamic dashboard that presents key insights in an interactive and visually engaging format, enabling users to explore the data efficiently.

1.0.2 Background

Airbnb is an online marketplace connecting people who want to rent out their homes with people looking for accommodations in that locale. It currently covers more than 100,000 cities and 220 countries worldwide. For hosts, it's a way to earn money while protecting their property from potential damage. However, for guests, it's a risky venture that they should avoid.

For this project, we are analyzing Airbnb's New York City(NYC) data for 2019. NYC is not only the most famous city in the world but also a top global destination for visitors drawn to its museums, entertainment, restaurants, and commerce. According to the Office of New York State Comptroller, NYC hosted 66.6 million visitors in 2019.

Data analysis on thousands of listings provided through Airbnb is a crucial factor for the company. Our main objective is to find out the key metrics influencing the listing of properties on the platform. For this, we will explore and visualize the dataset from Airbnb in NYC using basic exploratory data analysis (EDA) techniques. We have found out the distribution of every Airbnb listing based on their location, including their price range, room type, listing name, and other related factors. We have analyzed this dataset from different angles and developed interesting insights. This can help in making strategic data-driven decisions by the marketing team, finance team, and technical team of Airbnb.

1.0.3 Dataset

The dataset contains 48895 rows and 16 columns Id: Unique for each Property Listing.

name: Name of each Property Listing.

host id: Unique ID for a host who has listed the property on Airbnb.

host name: Name of host

neighbourhood_group: Name of Each borough of NYC, Manhattan, Brooklyn, Queens, Bronx, State Island.

neighborhood: Area in each borough of NYC

latitude, longitude: Co-ordinates of each listed property

room_type: Different types of rooms available for listing, Private room, Entire home/apt, Shared room.

price: Price of listing.

minimum nigths: Mandatory number of nights to be booked for available for each type of property.

number of review: Number of reviews for each Listed property

last_review: Date on which last time the listing was reviewed

review_per_month: Number of reviews per month

calculated host listings count: Number of listings each host owns

availablity_365: Number of days the given listing is available for booking

1.1 Let's Begin our basic EDA

Import all the necessary libraries

```
[1]: import numpy as np
  import pandas as pd
  import matplotlib.pyplot as plt
  %matplotlib inline
  import seaborn as sns
  import warnings
  warnings.filterwarnings("ignore")
```

Import the dataset

```
[2]: data = pd.read_csv("Airbnb.csv")
  data.head()
```

```
[2]:
          id
                                                                host_id \
                                                          name
     0 2539
                                                                   2787
                            Clean & quiet apt home by the park
     1 2595
                                         Skylit Midtown Castle
                                                                   2845
                           THE VILLAGE OF HARLEM...NEW YORK !
     2 3647
                                                                4632
     3 3831
                               Cozy Entire Floor of Brownstone
                                                                   4869
     4 5022
             Entire Apt: Spacious Studio/Loft by central park
                                                                   7192
         host_name neighbourhood_group neighbourhood latitude
                                                                 longitude \
     0
               John
                               Brooklyn
                                           Kensington
                                                       40.64749
                                                                 -73.97237
           Jennifer
     1
                              Manhattan
                                              Midtown
                                                      40.75362 -73.98377
     2
          Elisabeth
                              Manhattan
                                               Harlem 40.80902
                                                                 -73.94190
     3
       LisaRoxanne
                               Brooklyn Clinton Hill 40.68514 -73.95976
                                          East Harlem 40.79851 -73.94399
              Laura
                              Manhattan
```

room_type price minimum_nights number_of_reviews last_review \

```
225
                                                1
     1
        Entire home/apt
                                                                   45
                                                                       2019-05-21
                                                3
     2
           Private room
                             150
                                                                    0
                                                                               NaN
                                                                  270
                                                                       2019-07-05
        Entire home/apt
                             89
                                                1
        Entire home/apt
                             80
                                               10
                                                                       2018-11-19
                            calculated_host_listings_count
        reviews_per_month
                                                               availability 365
     0
                      0.21
                                                            6
                                                                             365
                      0.38
                                                            2
                                                                             355
     1
     2
                       NaN
                                                            1
                                                                             365
     3
                      4.64
                                                                             194
                                                            1
     4
                      0.10
                                                                               0
[4]: data.tail()
[4]:
                   id
                                                                               host_id \
                                                                       name
     48890
            36484665
                         Charming one bedroom - newly renovated rowhouse
                                                                               8232441
     48891
            36485057
                           Affordable room in Bushwick/East Williamsburg
                                                                               6570630
                                  Sunny Studio at Historical Neighborhood
                                                                              23492952
     48892
            36485431
                                     43rd St. Time Square-cozy single bed
     48893
            36485609
                                                                              30985759
     48894
            36487245
                       Trendy duplex in the very heart of Hell's Kitchen
                                                                              68119814
                 host_name neighbourhood_group
                                                       neighbourhood
                                                                       latitude
                   Sabrina
     48890
                                                  Bedford-Stuyvesant
                                       Brooklyn
                                                                       40.67853
     48891
                   Marisol
                                       Brooklyn
                                                             Bushwick
                                                                       40.70184
     48892
            Ilgar & Avsel
                                      Manhattan
                                                               Harlem
                                                                       40.81475
     48893
                                      Manhattan
                                                      Hell's Kitchen
                                                                       40.75751
                       Taz
     48894
                Christophe
                                      Manhattan
                                                      Hell's Kitchen
                                                                       40.76404
                                                                   number_of_reviews
            longitude
                                                  minimum_nights
                               room_type
                                          price
     48890
            -73.94995
                           Private room
                                              70
                                                                2
                                                                                    0
                                                                4
                                                                                    0
     48891
            -73.93317
                           Private room
                                              40
                                                               10
     48892
            -73.94867
                        Entire home/apt
                                             115
                                                                                    0
     48893
            -73.99112
                             Shared room
                                              55
                                                                1
                                                                                    0
            -73.98933
                           Private room
                                              90
                                                                7
                                                                                    0
     48894
                         reviews_per_month
                                              calculated_host_listings_count
           last_review
     48890
                    NaN
                                        NaN
                                                                             2
     48891
                    NaN
                                        NaN
                                                                             2
     48892
                    NaN
                                        NaN
                                                                             1
     48893
                    NaN
                                        NaN
                                                                             6
     48894
                    NaN
                                        NaN
                                                                             1
            availability_365
     48890
                            9
     48891
                           36
     48892
                           27
```

1

2018-10-19

0

Private room

149

```
48893
                       2
48894
                       23
```

Shape of the dataset

```
[5]: data.shape
```

[5]: (48895, 16)

Information of the dataset

[6]: data.info()

<class 'pandas.core.frame.DataFrame'> RangeIndex: 48895 entries, 0 to 48894 Data columns (total 16 columns):

#	Column	Non-Nu	Dtype			
0	id	48895	non-null	int64		
1	name	48879	non-null	object		
2	host_id	48895	non-null	int64		
3	host_name	48874	non-null	object		
4	neighbourhood_group	48895	non-null	object		
5	neighbourhood	48895	non-null	object		
6	latitude	48895	non-null	float64		
7	longitude	48895	non-null	float64		
8	room_type	48895	non-null	object		
9	price	48895	non-null	int64		
10	minimum_nights	48895	non-null	int64		
11	number_of_reviews	48895	non-null	int64		
12	last_review	38843	non-null	object		
13	reviews_per_month	38843	non-null	float64		
14	calculated_host_listings_count	48895	non-null	int64		
15	availability_365 48895 non-null					
dtyp	es: float64(3), int64(7), object	(6)				

memory usage: 6.0+ MB

Unique columns of the dataset

```
[7]: data.columns
```

```
[7]: Index(['id', 'name', 'host_id', 'host_name', 'neighbourhood_group',
            'neighbourhood', 'latitude', 'longitude', 'room_type', 'price',
            'minimum_nights', 'number_of_reviews', 'last_review',
            'reviews_per_month', 'calculated_host_listings_count',
            'availability_365'],
           dtype='object')
```

Statistical Description of the dataset

```
[9]: data.describe().T
 [9]:
                                        count
                                                       mean
                                                                      std \
      id
                                      48895.0
                                              1.901714e+07
                                                             1.098311e+07
     host_id
                                      48895.0
                                              6.762001e+07
                                                             7.861097e+07
     latitude
                                      48895.0 4.072895e+01
                                                             5.453008e-02
                                      48895.0 -7.395217e+01
                                                             4.615674e-02
      longitude
                                      48895.0 1.527207e+02
                                                             2.401542e+02
     price
     minimum_nights
                                      48895.0
                                              7.029962e+00
                                                             2.051055e+01
     number_of_reviews
                                      48895.0
                                               2.327447e+01
                                                             4.455058e+01
      reviews_per_month
                                      38843.0
                                              1.373221e+00
                                                             1.680442e+00
      calculated_host_listings_count
                                      48895.0
                                              7.143982e+00
                                                             3.295252e+01
      availability_365
                                      48895.0 1.127813e+02 1.316223e+02
                                                           25%
                                             min
                                                                         50%
                                      2539.00000
                                                 9.471945e+06
                                                                1.967728e+07
      id
     host id
                                      2438.00000
                                                  7.822033e+06
                                                                3.079382e+07
      latitude
                                        40.49979 4.069010e+01
                                                                4.072307e+01
     longitude
                                       -74.24442 -7.398307e+01 -7.395568e+01
     price
                                         0.00000 6.900000e+01 1.060000e+02
                                                  1.000000e+00
                                                                3.000000e+00
     minimum_nights
                                         1.00000
     number_of_reviews
                                         0.00000
                                                  1.000000e+00
                                                                5.000000e+00
      reviews_per_month
                                         0.01000
                                                  1.900000e-01
                                                                7.200000e-01
      calculated_host_listings_count
                                         1.00000
                                                  1.000000e+00
                                                                1.000000e+00
      availability_365
                                         0.00000 0.000000e+00
                                                                4.500000e+01
                                               75%
                                                             max
      id
                                      2.915218e+07 3.648724e+07
     host_id
                                      1.074344e+08 2.743213e+08
     latitude
                                      4.076311e+01 4.091306e+01
      longitude
                                     -7.393627e+01 -7.371299e+01
     price
                                      1.750000e+02 1.000000e+04
     minimum_nights
                                      5.000000e+00 1.250000e+03
     number of reviews
                                      2.400000e+01 6.290000e+02
      reviews_per_month
                                      2.020000e+00 5.850000e+01
      calculated_host_listings_count
                                     2.000000e+00 3.270000e+02
      availability_365
                                      2.270000e+02 3.650000e+02
     Checking for null values in the dataset
[10] · data isnull() sum()
```

	4454.151411().5411()							
[10]:	id	0						
	name	16						
	host_id	0						

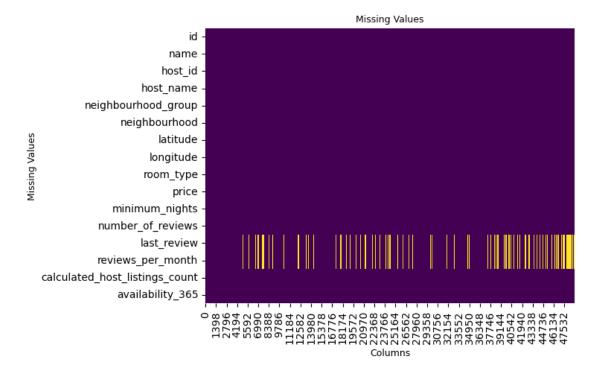
host_name

21

```
neighbourhood_group
                                        0
neighbourhood
                                        0
latitude
                                        0
longitude
                                        0
room_type
                                        0
                                        0
price
minimum_nights
                                        0
number_of_reviews
                                        0
                                    10052
last_review
reviews_per_month
                                    10052
calculated_host_listings_count
                                        0
availability_365
                                        0
dtype: int64
```

Using Heatmaps to Visualize Missing Values

```
[16]: ax = plt.axes()
    sns.heatmap(data.isna().transpose(), cbar=False, ax=ax, cmap="viridis")
    plt.title("Missing Values", fontsize=9)
    plt.xlabel("Columns", fontsize = 9)
    plt.ylabel("Missing Values", fontsize = 9)
    plt.show()
```



Replacing the null values with appropriate values

```
[17]: data['name'].replace(np.nan, 'Other Hotel', inplace =True)
      data['host_name'].replace(np.nan, 'other', inplace = True)
      data['last review'].replace(np.nan, 'Not Reviewed', inplace = True)
      data['reviews_per_month'].replace(np.nan, '0', inplace = True)
[18]: data.isnull().sum()
[18]: id
                                         0
                                         0
     name
     host_id
                                         0
     host name
                                         0
      neighbourhood_group
                                         0
      neighbourhood
      latitude
                                         0
      longitude
                                         0
     room_type
                                         0
                                         0
     price
     minimum_nights
                                         0
     number_of_reviews
                                         0
      last_review
                                         0
      reviews_per_month
      calculated_host_listings_count
                                         0
      availability_365
                                         0
      dtype: int64
```

• Now we can see, the null values in the dataset has been replaced with appropriate values

Checking for Duplicate records

```
[20]: data.duplicated().sum()
```

[20]: 0

• Airbnb dataset do not have duplicated records

1.2 Exploratory Data Analysis – Data Visualization

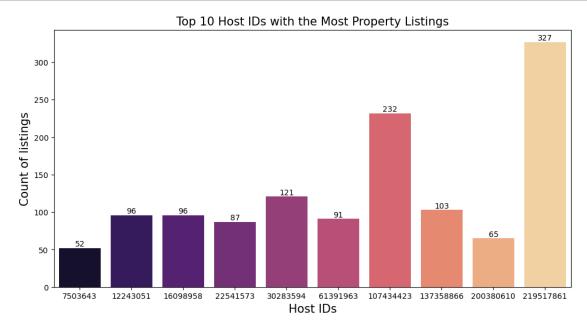
```
[21]: data.head()
[21]:
                                                                 host_id \
           id
                                                           name
      0 2539
                             Clean & quiet apt home by the park
                                                                     2787
      1 2595
                                          Skylit Midtown Castle
                                                                     2845
      2 3647
                            THE VILLAGE OF HARLEM...NEW YORK !
                                                                  4632
      3 3831
                                Cozy Entire Floor of Brownstone
                                                                     4869
      4 5022
             Entire Apt: Spacious Studio/Loft by central park
                                                                     7192
           host_name neighbourhood_group neighbourhood latitude
                                                                  longitude \
                                            Kensington 40.64749
      0
                John
                                Brooklyn
                                                                  -73.97237
```

```
1
             Jennifer
                                Manhattan
                                                 Midtown 40.75362 -73.98377
       2
            Elisabeth
                                Manhattan
                                                  Harlem 40.80902 -73.94190
       3 LisaRoxanne
                                 Brooklyn
                                            Clinton Hill
                                                          40.68514
                                                                    -73.95976
       4
                Laura
                                Manhattan
                                             East Harlem 40.79851 -73.94399
                          price minimum_nights
                                                   number_of_reviews
                                                                       last_review
                room_type
             Private room
                             149
                                                                        2018-10-19
       0
                                                1
       1
         Entire home/apt
                             225
                                                                  45
                                                                         2019-05-21
                                                3
       2
             Private room
                             150
                                                                   0 Not Reviewed
       3 Entire home/apt
                              89
                                                1
                                                                         2019-07-05
                                                                 270
       4 Entire home/apt
                                               10
                              80
                                                                   9
                                                                         2018-11-19
         reviews_per_month
                            calculated_host_listings_count
                                                             availability 365
       0
                      0.21
                                                          6
                                                                           365
                      0.38
                                                          2
                                                                           355
       1
       2
                         0
                                                          1
                                                                           365
       3
                      4.64
                                                          1
                                                                           194
       4
                       0.1
                                                                             0
                                                          1
[22]:
      data.columns
[22]: Index(['id', 'name', 'host_id', 'host_name', 'neighbourhood_group',
              'neighbourhood', 'latitude', 'longitude', 'room_type', 'price',
              'minimum_nights', 'number_of_reviews', 'last_review',
              'reviews_per_month', 'calculated_host_listings_count',
              'availability_365'],
             dtype='object')
      Which are the top 10 host_id with the highest number of property listings?
[27]: data['host_id'].value_counts().reset_index().head(10)
[27]:
            host_id count
       0
        219517861
                       327
       1
        107434423
                       232
       2
           30283594
                       121
       3
        137358866
                       103
           16098958
                        96
       4
       5
           12243051
                        96
       6
           61391963
                        91
       7
                        87
           22541573
       8 200380610
                        65
            7503643
                        52
[105]: # Plotting the bar graph to visualize the top 10 host ids with the highest
        →number of property listings
       plt.figure(figsize=(12, 6))
```

```
ax = sns.barplot(x=data['host_id'].value_counts().iloc[:10].index,__
y=data['host_id'].value_counts().iloc[:10].values, palette="magma")

for bars in ax.containers:
    ax.bar_label(bars)

plt.title("Top 10 Host IDs with the Most Property Listings", fontsize=15)
plt.xlabel("Host IDs", fontsize=15)
plt.ylabel("Count of listings", fontsize=15)
plt.show()
```



• The Host id 219517861 has the highest number of bookings with the total bookings of 327

Which are the unique Neighborhoods in the dataset?

```
[30]: unique_neighborhoods = data['neighbourhood'].nunique()
print("Number of unique neighborhoods:", unique_neighborhoods)
```

Number of unique neighborhoods: 221

```
[33]: data['neighbourhood'].unique()
```

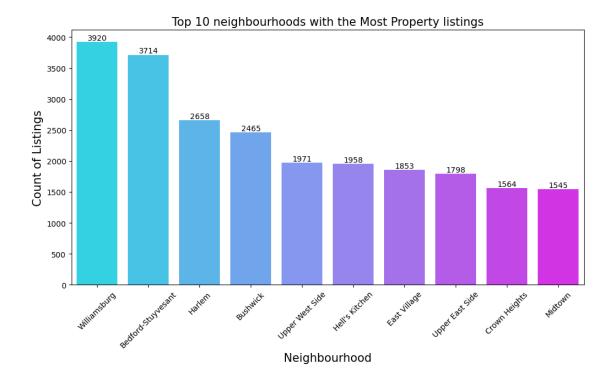
```
'Prospect-Lefferts Gardens', 'Long Island City', 'Kips Bay',
'SoHo', 'Upper East Side', 'Prospect Heights',
'Washington Heights', 'Woodside', 'Brooklyn Heights',
'Carroll Gardens', 'Gowanus', 'Flatlands', 'Cobble Hill',
'Flushing', 'Boerum Hill', 'Sunnyside', 'DUMBO', 'St. George',
'Highbridge', 'Financial District', 'Ridgewood',
'Morningside Heights', 'Jamaica', 'Middle Village', 'NoHo',
'Ditmars Steinway', 'Flatiron District', 'Roosevelt Island',
'Greenwich Village', 'Little Italy', 'East Flatbush',
'Tompkinsville', 'Astoria', 'Clason Point', 'Eastchester',
'Kingsbridge', 'Two Bridges', 'Queens Village', 'Rockaway Beach',
'Forest Hills', 'Nolita', 'Woodlawn', 'University Heights',
'Gravesend', 'Gramercy', 'Allerton', 'East New York',
'Theater District', 'Concourse Village', 'Sheepshead Bay',
'Emerson Hill', 'Fort Hamilton', 'Bensonhurst', 'Tribeca',
'Shore Acres', 'Sunset Park', 'Concourse', 'Elmhurst',
'Brighton Beach', 'Jackson Heights', 'Cypress Hills', 'St. Albans',
'Arrochar', 'Rego Park', 'Wakefield', 'Clifton', 'Bay Ridge',
'Graniteville', 'Spuyten Duyvil', 'Stapleton', 'Briarwood',
'Ozone Park', 'Columbia St', 'Vinegar Hill', 'Mott Haven',
'Longwood', 'Canarsie', 'Battery Park City', 'Civic Center',
'East Elmhurst', 'New Springville', 'Morris Heights', 'Arverne',
'Cambria Heights', 'Tottenville', 'Mariners Harbor', 'Concord',
'Borough Park', 'Bayside', 'Downtown Brooklyn', 'Port Morris',
'Fieldston', 'Kew Gardens', 'Midwood', 'College Point',
'Mount Eden', 'City Island', 'Glendale', 'Port Richmond',
'Red Hook', 'Richmond Hill', 'Bellerose', 'Maspeth',
'Williamsbridge', 'Soundview', 'Woodhaven', 'Woodrow',
'Co-op City', 'Stuyvesant Town', 'Parkchester', 'North Riverdale',
'Dyker Heights', 'Bronxdale', 'Sea Gate', 'Riverdale',
'Kew Gardens Hills', 'Bay Terrace', 'Norwood', 'Claremont Village',
'Whitestone', 'Fordham', 'Bayswater', 'Navy Yard', 'Brownsville',
'Eltingville', 'Fresh Meadows', 'Mount Hope', 'Lighthouse Hill',
'Springfield Gardens', 'Howard Beach', 'Belle Harbor',
'Jamaica Estates', 'Van Nest', 'Morris Park', 'West Brighton',
'Far Rockaway', 'South Ozone Park', 'Tremont', 'Corona',
'Great Kills', 'Manhattan Beach', 'Marble Hill', 'Dongan Hills',
'Castleton Corners', 'East Morrisania', 'Hunts Point', 'Neponsit',
'Pelham Bay', 'Randall Manor', 'Throgs Neck', 'Todt Hill',
'West Farms', 'Silver Lake', 'Morrisania', 'Laurelton',
'Grymes Hill', 'Holliswood', 'Pelham Gardens', 'Belmont',
'Rosedale', 'Edgemere', 'New Brighton', 'Midland Beach',
'Baychester', 'Melrose', 'Bergen Beach', 'Richmondtown',
'Howland Hook', 'Schuylerville', 'Coney Island', 'New Dorp Beach',
"Prince's Bay", 'South Beach', 'Bath Beach', 'Jamaica Hills',
'Oakwood', 'Castle Hill', 'Hollis', 'Douglaston', 'Huguenot',
'Olinville', 'Edenwald', 'Grant City', 'Westerleigh',
```

```
'Bay Terrace, Staten Island', 'Westchester Square', 'Little Neck', 'Fort Wadsworth', 'Rosebank', 'Unionport', 'Mill Basin', 'Arden Heights', "Bull's Head", 'New Dorp', 'Rossville', 'Breezy Point', 'Willowbrook'], dtype=object)
```

• There are 221 unique neighborhoods present in the dataset

What are the top 10 neighborhoods with the most property listings?

```
[45]: data['neighbourhood'].value_counts().reset_index().head(10)
[45]:
             neighbourhood
                           count
              Williamsburg
                            3920
      0
        Bedford-Stuyvesant
                            3714
      2
                    Harlem
                            2658
      3
                  Bushwick
                            2465
           Upper West Side
                            1971
      4
            Hell's Kitchen
      5
                            1958
      6
              East Village
                            1853
      7
           Upper East Side
                            1798
             Crown Heights
      8
                            1564
                   Midtown
      9
                            1545
[106]: # Plotting the bar graph to visualize the top 10 neighborhoods with the highest
       →number of listings
      plt.figure(figsize=(12, 6))
      ax = sns.barplot(x=data['neighbourhood'].value_counts().iloc[:10].index,__
       for bars in ax.containers:
          ax.bar_label(bars)
      plt.title("Top 10 neighbourhoods with the Most Property listings", fontsize=15)
      plt.xlabel("Neighbourhood", fontsize=15)
      plt.ylabel("Count of Listings", fontsize=15)
      plt.xticks(rotation=45)
      plt.show()
```



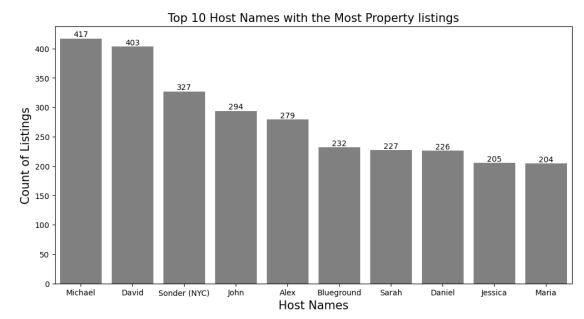
• Williamsburg is the neighborhood with the most listings in NYC with total 3920 listings

```
What are the top 10 host_name with the highest number of listings?
```

```
data['host_name'].value_counts().reset_index().iloc[:10]
[43]:
[43]:
             host_name
                         count
      0
               Michael
                           417
      1
                 David
                           403
         Sonder (NYC)
      2
                           327
                  John
      3
                           294
      4
                  Alex
                           279
      5
            Blueground
                           232
      6
                 Sarah
                           227
      7
                Daniel
                           226
      8
               Jessica
                           205
      9
                 Maria
                           204
```

```
ax.bar_label(bars)

plt.title("Top 10 Host Names with the Most Property listings", fontsize=15)
plt.xlabel("Host Names", fontsize=15)
plt.ylabel("Count of Listings", fontsize=15)
plt.show()
```



• We can observe that, Michael is the host who is the owner of highest number of property listings with the total of 417 listings

Which are the unique neighbourhood_groups in the dataset?

```
[47]: unique_neighborhood_grps = data['neighbourhood_group'].nunique()
print("Number of unique neighborhood grpups:", unique_neighborhood_grps)
```

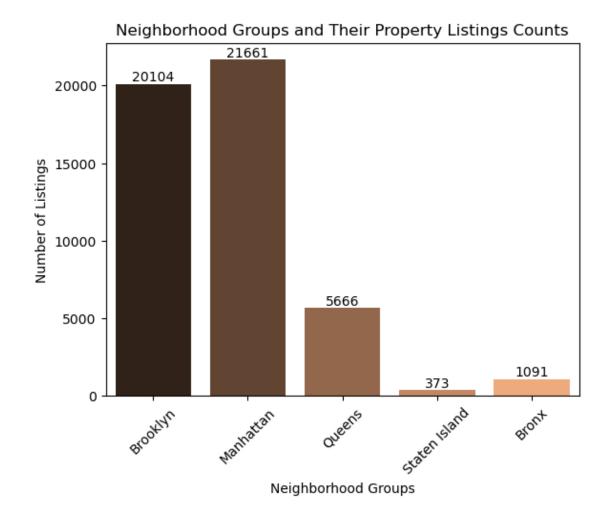
Number of unique neighborhood grpups: 5

```
[48]: data['neighbourhood_group'].unique()
```

```
[48]: array(['Brooklyn', 'Manhattan', 'Queens', 'Staten Island', 'Bronx'], dtype=object)
```

What are the property listing counts for each neighborhood group?

```
Number of listings in each neighborhood group:
       neighbourhood_group
      Manhattan
                       21661
      Brooklyn
                       20104
      Queens
                        5666
      Bronx
                        1091
      Staten Island
                         373
      Name: count, dtype: int64
[108]: listings_per_neighborhood_group = sns.countplot(x='neighbourhood_group',__
        ⇔data=data, palette="copper")
       for bars in listings_per_neighborhood_group.containers:
          listings_per_neighborhood_group.bar_label(bars)
       plt.title("Neighborhood Groups and Their Property Listings Counts")
       plt.xlabel("Neighborhood Groups")
       plt.ylabel("Number of Listings")
       plt.xticks(rotation=45)
       plt.show()
```



• Data shows that, Manhattan is the neighborhood group with the highest number of property listings and Staten Island contributes to the lowest number of listings.

Which neighborhoods belong to each neighborhood group?

```
[50]: neighborhoods_by_group = data.groupby('neighbourhood_group')['neighbourhood'].

ounique()
print(neighborhoods_by_group)
```

neighbourhood_group

Bronx [Highbridge, Clason Point, Eastchester, Kingsb...
Brooklyn [Kensington, Clinton Hill, Bedford-Stuyvesant,...
Manhattan [Midtown, Harlem, East Harlem, Murray Hill, He...
Queens [Long Island City, Woodside, Flushing, Sunnysi...
Staten Island [St. George, Tompkinsville, Emerson Hill, Shor...
Name: neighbourhood, dtype: object

What are the different Types of rooms available in the dataset?

```
[60]: data['room_type'].unique()
```

[60]: array(['Private room', 'Entire home/apt', 'Shared room'], dtype=object)

• As we can see, there are mainly 3 categories of rooms available in the Airbnb NYC listings

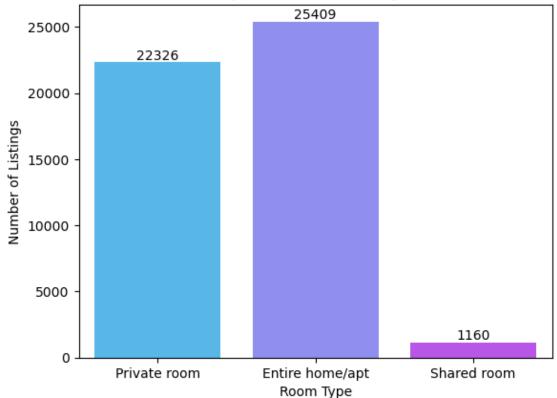
Which room type contributes to the highest number of listings in Airbnb NYC?

```
[109]: room_type_listings = sns.countplot(x='room_type', data=data, palette="cool")

for bars in room_type_listings.containers:
    room_type_listings.bar_label(bars)

plt.title("Room types and their listing counts")
plt.xlabel("Room Type")
plt.ylabel("Number of Listings")
plt.show()
```

Room types and their listing counts



- We can observe clearly that "Entire home/apt" room type has the highest number of listings in NYC
- The most common room type is "Entire home/apt" room type

• And Shared room contributes to the lowest number of listings

What is the Average price for all listings?

```
[64]: average_price = data['price'].mean()
print("Average price for all listings:", average_price)
```

Average price for all listings: 152.7206871868289

• The average price for all listings in the dataset is approximately \$152.72.

How many hosts have more than one listing?

```
[65]: hosts_listings = data['host_id'].value_counts()
hosts_multiple_listing = (hosts_listings > 1).sum()
print("Number of hosts with more than one listing:", hosts_multiple_listing)
```

Number of hosts with more than one listing: 5154

Which neighborhood has the highest average price?

```
highest_avg_price_neighborhood = data.groupby('neighbourhood')['price'].mean().
idxmax()
highest_avg_price = data.groupby('neighbourhood')['price'].mean().max()
highest_avg_price_neighborhood, highest_avg_price
```

[76]: ('Fort Wadsworth', 800.0)

How many listings are available throughout the entire year?

```
[23]: listings_available_all_year = data[data['availability_365'] == 365] listings_available_all_year
```

[23]:		id			na	me host_i	id \
	0	2539		Clean & quie	t apt home by the pa	rk 278	37
	2	3647		THE VILLAGE O	F HARLEMNEW YORK !	4632	
	36	11452		Clean	and Quiet in Brookl	.yn 735	55
	38	11943		Cou	ntry space in the ci	ty 4544	1 5
	97	21644		Upp	er Manhattan, New Yo	ork 8268	35
		•••			***	•••	
	48744	36415840		A BEAUTIFUL SPACE IN	HEART OF WILLIAMSBU	JRG 22371546	30
	48844	36453952	West	Village Studio on qu	iet cobblestone stre	et 11549189	}6
	48868	36473253		Heaven	for you(only for gu	y) 26133817	77
	48880	36481315	The Rac	coon Artist Studio i	n Williamsburg New	. 208514239	
	48887	36483152	Garde	n Jewel Apartment in	Williamsburg New Yo	rk 20851423	39
		ho	st_name	neighbourhood_group	neighbourhood	latitude \	\
	0		John	Brooklyn	Kensington	40.64749	
	2	E1	isabeth	Manhattan	Harlem	40.80902	
	36		Vt	Brooklyn	Bedford-Stuyvesant	40.68876	

```
38
                 Harriet
                                      Brooklyn
                                                            Flatbush
                                                                       40.63702
97
                                                                       40.82803
                 Elliott
                                     Manhattan
                                                               Harlem
                                       •••
       Simon And Julian
                                                        Williamsburg
48744
                                      Brooklyn
                                                                       40.71091
48844
                    Will
                                     Manhattan
                                                        West Village
                                                                       40.73620
48868
                   Diana
                                      Brooklyn
                                                           Gravesend
                                                                       40.59118
48880
                   Melki
                                      Brooklyn
                                                        Williamsburg
                                                                       40.71232
                   Melki
48887
                                      Brooklyn
                                                        Williamsburg
                                                                       40.71232
       longitude
                                      price
                                              minimum_nights
                                                               number_of_reviews
                          room_type
                                                            1
0
       -73.97237
                       Private room
                                         149
                                                                                 9
                                                            3
2
       -73.94190
                       Private room
                                         150
                                                                                 0
                                                           60
36
       -73.94312
                       Private room
                                         35
                                                                                 0
38
       -73.96327
                       Private room
                                         150
                                                            1
                                                                                 0
97
       -73.94731
                       Private room
                                         89
                                                            1
                                                                                  1
       -73.96560
                   Entire home/apt
                                         499
                                                                                 0
48744
                                                           30
48844
       -74.00827
                   Entire home/apt
                                         205
                                                            1
                                                                                 0
                                                            7
       -73.97119
                        Shared room
                                         25
                                                                                 0
48868
                   Entire home/apt
48880
       -73.94220
                                         120
                                                            1
                                                                                 0
48887
       -73.94220
                   Entire home/apt
                                         170
                                                            1
      last_review
                    reviews_per_month
                                         calculated_host_listings_count
       2018-10-19
0
                                   0.21
2
               NaN
                                    NaN
                                                                          1
                                    NaN
36
               NaN
                                                                          1
38
                                    NaN
               NaN
                                                                          1
97
       2018-10-09
                                   0.11
                                                                          1
48744
                                    NaN
               NaN
                                                                          1
48844
               NaN
                                    NaN
                                                                          1
                                                                          6
               NaN
                                    NaN
48868
                                                                          3
48880
               NaN
                                    NaN
                                                                          3
48887
               NaN
                                    NaN
       availability_365
0
                      365
2
                      365
36
                      365
38
                      365
97
                      365
48744
                      365
48844
                      365
48868
                      365
                      365
48880
                      365
48887
```

[1295 rows x 16 columns]

36

38

NaN

NaN

```
[26]: listings_365 = listings_available_all_year.shape[0] print("Number of listings available throughout the entire year:", listings_365)
```

Number of listings available throughout the entire year: 1295

What is the average minimum nights required for a stay?

```
[5]: avg_minimum_nights = data['minimum_nights'].mean()
print("Average minimum nights required for a stay is:", avg_minimum_nights)
```

Average minimum nights required for a stay is: 7.029962163820431

How many listings have never been reviewed?

```
[28]: listings_never_reviewed = data[data['number_of_reviews'] == 0] listings_never_reviewed.head()
```

[28]:		id					name	host_id	\
[20]	2	3647		THE VIL	LAGE O	F HARLEMNEW YOR		4632	`
	19	7750				pper East Centa		17985	
	26	8700	Magnifique Su	•	-	nhattan - vue Cl			
	36	11452			Clean	and Quiet in Br	ooklyn	7355	
	38	11943		Cou	ntry space in th	ne city	45445		
		h	hbourhoo	d grou	p neighbou	ırhood	latitude	\	
	2	Elisabeth			nhatta	-	Harlem		
	19		Ma	nhatta	n East H	Harlem	40.79685		
	26	Claude	Ma	nhatta	n]	nwood	40.86754		
	36		В	Brooklyn Bedford-Stuyvesant Brooklyn Flatbush		40.68876			
	38		Harriet			40.63702			
		longitu	longitude room		price	minimum_nights	numbe	r_of_revie	ws \
	2	-73.941			150	3			0
	19	-73.948	372 Entire ho	me/apt	190	7			0
	26	-73.926	39 Privat	e room	80	4			0
	36	-73.943	312 Privat	e room	35	60			0
	38	-73.963	327 Privat	e room	150	1			0
		last_rev	riew reviews_	per_mont	h cal	culated_host_lis	stings_	count \	
	2	_	NaN	Na			0 -	1	
	19		NaN	Na	.N			2	
	26		NaN	Na	.N			1	

NaN

NaN

1

1

```
availability_365
      2
                        365
      19
                        249
      26
                          0
      36
                        365
      38
                        365
[19]: zero_reviews_count = listings_never_reviewed.shape[0]
      zero_reviews_count
[19]: 10052
        • As per data, 10052 listings have never been reviewed
[44]: data.head()
[44]:
                                                                   host_id \
           id
                                                             name
         2539
                              Clean & quiet apt home by the park
                                                                       2787
         2595
                                            Skylit Midtown Castle
      1
                                                                       2845
                             THE VILLAGE OF HARLEM...NEW YORK !
      2
         3647
                                                                    4632
      3 3831
                                 Cozy Entire Floor of Brownstone
                                                                       4869
        5022
               Entire Apt: Spacious Studio/Loft by central park
                                                                       7192
           host_name neighbourhood_group neighbourhood latitude
                                                                     longitude \
      0
                John
                                 Brooklyn
                                             Kensington
                                                          40.64749
                                                                     -73.97237
      1
            Jennifer
                                Manhattan
                                                 Midtown
                                                          40.75362
                                                                    -73.98377
           Elisabeth
                                Manhattan
                                                  Harlem 40.80902
                                                                    -73.94190
      3
        LisaRoxanne
                                 Brooklyn Clinton Hill 40.68514 -73.95976
                                            East Harlem 40.79851
               Laura
                                Manhattan
                                                                    -73.94399
                                  minimum_nights
                                                   number_of_reviews last_review
               room_type
                          price
      0
            Private room
                             149
                                                                       2018-10-19
      1
         Entire home/apt
                             225
                                                1
                                                                   45
                                                                       2019-05-21
      2
            Private room
                                                3
                             150
                                                                    0
                                                                              NaN
         Entire home/apt
                              89
                                                1
                                                                  270
                                                                       2019-07-05
      3
```

```
reviews_per_month
                        calculated_host_listings_count
                                                           availability_365
0
                 0.21
                                                                          365
                                                        6
1
                 0.38
                                                        2
                                                                          355
2
                                                                          365
                  NaN
                                                        1
3
                 4.64
                                                        1
                                                                          194
                 0.10
                                                                            0
```

Dropping latitude, longitude columns

80

Entire home/apt

```
[48]: data = data.drop(columns=['latitude', 'longitude'])
```

10

2018-11-19

```
[50]: data.columns
[50]: Index(['id', 'name', 'host_id', 'host_name', 'neighbourhood_group',
              'neighbourhood', 'room_type', 'price', 'minimum_nights',
             'number_of_reviews', 'last_review', 'reviews_per_month',
             'calculated_host_listings_count', 'availability_365'],
            dtype='object')
     Outlier Detection in Price attribute
[51]: data.head()
[51]:
           id
                                                                    host id \
                                                             name
      0
         2539
                              Clean & quiet apt home by the park
                                                                       2787
      1 2595
                                            Skylit Midtown Castle
                                                                       2845
      2 3647
                             THE VILLAGE OF HARLEM...NEW YORK !
                                                                    4632
      3 3831
                                 Cozy Entire Floor of Brownstone
                                                                       4869
      4 5022
              Entire Apt: Spacious Studio/Loft by central park
                                                                       7192
           host_name neighbourhood_group neighbourhood
                                                                 room_type
                                                                            price
      0
                John
                                 Brooklyn
                                              Kensington
                                                                              149
                                                             Private room
      1
            Jennifer
                                Manhattan
                                                 Midtown
                                                         Entire home/apt
                                                                              225
      2
           Elisabeth
                                Manhattan
                                                  Harlem
                                                             Private room
                                                                              150
                                                         Entire home/apt
      3
         LisaRoxanne
                                 Brooklyn
                                           Clinton Hill
                                                                               89
               Laura
                                Manhattan
                                             East Harlem
                                                         Entire home/apt
                                                                               80
         minimum_nights
                          number_of_reviews last_review
                                                         reviews_per_month
      0
                                                                        0.21
                       1
                                           9
                                              2018-10-19
                       1
                                                                        0.38
      1
                                         45
                                              2019-05-21
                       3
      2
                                           0
                                                     NaN
                                                                         NaN
      3
                                        270
                                              2019-07-05
                                                                        4.64
                       1
                                              2018-11-19
                                                                        0.10
      4
                      10
         calculated_host_listings_count
                                          availability_365
      0
                                                        365
                                        2
      1
                                                        355
      2
                                        1
                                                        365
      3
                                        1
                                                        194
      4
                                                          0
                                        1
[53]: data['price'].describe()
               48895.000000
[53]: count
                 152.720687
      mean
                 240.154170
      std
      min
                   0.000000
      25%
                  69.000000
```

```
75%
                 175.000000
      max
               10000.000000
      Name: price, dtype: float64
[58]: data['price'].median()
[58]: 106.0
     5 Number Summary: 1. Minimum = 0 2. First Quartile range (Q1) = 69 3. Median = 106 4.
     Third Quartile range (Q3) = 175 5. Maximum = 10000
[55]: # Calculating Interquartile Range (IQR)
      Q1 = 69
      Q3 = 175
      IQR = Q3-Q1
      print("Interquartile range, IQR: ",IQR)
     Interquartile range, IQR:
[59]: Lower_Fence = Q1-(1.5*(IQR))
      Higher Fence = Q3 + (1.5*(IQR))
      print("Lower Fence is: ",Lower_Fence)
      print("Higher Fence is: ", Higher_Fence)
     Lower Fence is: -90.0
     Higher Fence is:
                       334.0
     This means that the price below -90 and above 334 is considered as an outlier
[61]: # Finding the data within the range of Lower fence and upper fence
      data_within_fences = data[(data['price'] >= Lower_Fence) & (data['price'] <=__
       →Higher Fence)]
      data_within_fences.head()
[61]:
           id
                                                                  host_id \
                                                            name
      0 2539
                             Clean & quiet apt home by the park
                                                                      2787
      1 2595
                                           Skylit Midtown Castle
                                                                      2845
      2 3647
                            THE VILLAGE OF HARLEM...NEW YORK !
                                                                   4632
      3 3831
                                 Cozy Entire Floor of Brownstone
                                                                      4869
      4 5022 Entire Apt: Spacious Studio/Loft by central park
                                                                      7192
           host_name neighbourhood_group neighbourhood
                                                                room_type
                                                                          price \
                                Brooklyn
      0
                John
                                             Kensington
                                                            Private room
                                                                             149
      1
            Jennifer
                               Manhattan
                                                Midtown Entire home/apt
                                                                             225
           Elisabeth
                                                            Private room
                               Manhattan
                                                 Harlem
                                                                             150
      3 LisaRoxanne
                                Brooklyn Clinton Hill Entire home/apt
                                                                              89
               Laura
                                            East Harlem Entire home/apt
                                                                              80
                               Manhattan
```

50%

106.000000

```
minimum_nights
                         number_of_reviews last_review reviews_per_month \
                                                                       0.21
      0
                      1
                                             2018-10-19
                                                                       0.38
                      1
                                         45
                                             2019-05-21
      1
      2
                      3
                                          0
                                                    NaN
                                                                        NaN
      3
                      1
                                        270
                                             2019-07-05
                                                                       4.64
                                                                       0.10
      4
                     10
                                             2018-11-19
         calculated_host_listings_count availability_365
      0
                                                        355
      1
      2
                                       1
                                                       365
      3
                                       1
                                                        194
      4
                                                         0
                                       1
[62]: data_within_fences.shape
[62]: (45923, 14)
[65]: # Calculate the percentage of data within fences
      percentageofdata_within_fences = (len(data_within_fences) / len(data)) * 100
      percentageofdata_within_fences
[65]: 93.9216688822988
```

Approximately 94% of data lies within Lower fence and upper fence.

What is the average price for each neighborhood group, and which neighborhood group has the highest average price?

[75]: data_within_fences.groupby(['neighbourhood_group'])['price'].describe().T

5]:	neighbourhood_group	Bronx	Brooklyn	Manhattan	Queens
	count	1070.000000	19415.000000	19506.000000	5567.000000
	mean	77.365421	105.699614	145.952835	88.904437
	std	47.110940	60.937808	70.473076	53.536041
	min	0.000000	0.00000	0.000000	10.000000
	25%	45.000000	60.000000	90.000000	50.000000
	50%	65.000000	90.000000	135.000000	74.000000
	75%	95.000000	140.000000	199.000000	108.000000
	max	325.000000	333.000000	334.000000	325.000000
	neighbourhood_group	Staten Island	i		
	count	365.000000)		
	mean	89.235616	3		
	std	57.700147	7		
	min	13.000000)		
	25%	50.000000)		

50%	75.000000
75%	105.000000
max	300.000000

- The data shows that Manhattan has the highest average property prices, making it the most expensive area to live in NYC. In contrast, the Bronx has the lowest average property prices, making it the most affordable area in the city.
- Queens and Staten Island appear to have similar property prices.

What is the average number of listings per host?

```
data.head()
[76]:
[76]:
           id
                                                                     host id \
                                                              name
         2539
                              Clean & quiet apt home by the park
      0
                                                                        2787
         2595
                                            Skylit Midtown Castle
      1
                                                                        2845
      2
         3647
                             THE VILLAGE OF HARLEM...NEW YORK !
                                                                     4632
      3
         3831
                                  Cozy Entire Floor of Brownstone
                                                                        4869
         5022
               Entire Apt: Spacious Studio/Loft by central park
                                                                        7192
           host_name neighbourhood_group neighbourhood
                                                                  room_type
                                                                             price
      0
                 John
                                  Brooklyn
                                              Kensington
                                                              Private room
                                                                                149
            Jennifer
      1
                                 Manhattan
                                                  Midtown
                                                           Entire home/apt
                                                                                225
      2
           Elisabeth
                                 Manhattan
                                                   Harlem
                                                              Private room
                                                                                150
                                                           Entire home/apt
      3
         LisaRoxanne
                                  Brooklyn
                                            Clinton Hill
                                                                                 89
      4
               Laura
                                              East Harlem
                                                           Entire home/apt
                                                                                 80
                                 Manhattan
         minimum_nights
                          number_of_reviews last_review
                                                           reviews_per_month
      0
                       1
                                           9
                                              2018-10-19
                                                                         0.21
                                              2019-05-21
                                                                         0.38
      1
                       1
                                          45
      2
                       3
                                           0
                                                                          NaN
                                                      NaN
      3
                       1
                                         270
                                              2019-07-05
                                                                         4.64
      4
                                              2018-11-19
                                                                         0.10
                      10
         calculated_host_listings_count
                                           availability_365
      0
                                                         365
                                        2
      1
                                                         355
      2
                                        1
                                                         365
      3
                                        1
                                                         194
      4
                                                           0
                                        1
[80]: total_listings = data.shape[0]
      unique_hosts = data['host_id'].nunique()
      average_listings_per_host = total_listings / unique_hosts
      print("Average Listings per host: ", average_listings_per_host)
```

Average Listings per host: 1.3053634834610353

• Average number of listings per host is 1.30

How many listings are there per neighborhood group by room type?

[84]:	neighbourhood_group	room_type	
	Bronx	Private room	652
		Entire home/apt	379
		Shared room	60
	Brooklyn	Private room	10132
		Entire home/apt	9559
		Shared room	413
	Manhattan	Entire home/apt	13199
		Private room	7982
		Shared room	480
	Queens	Private room	3372
		Entire home/apt	2096
		Shared room	198
	Staten Island	Private room	188
		Entire home/apt	176
		Shared room	9

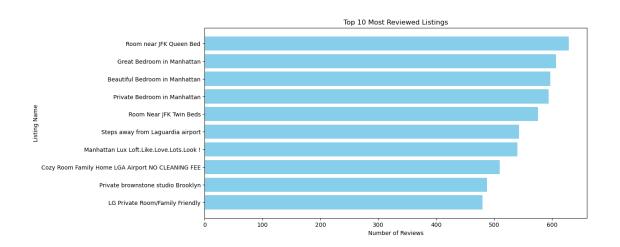
Name: count, dtype: int64

What are the top 10 most reviewed listings?

[88]: data head()

00]:	aa	ta.nea	α()					
38]:		id				name host_id	\	
	0	2539		Clean & quie	t apt home by	-		
	1	2595		•	Skylit Midtow	n Castle 2845		
	2	3647		THE VILLAGE O	F HARLEMNEW Y	ORK ! 4632		
	3	3831		Cozy Enti:	re Floor of Br	ownstone 4869		
	4	5022	Entire Apt:	Spacious Studio	o/Loft by cent	ral park 7192		
		hos.	t_name neigh	bourhood_group	neighbourhood	room_type	price	\
	0		John	Brooklyn	Kensington	Private room	149	
	1	Je	nnifer	Manhattan	Midtown	Entire home/apt	225	
	2	Elia	sabeth	Manhattan	Harlem	Private room	150	
	3	LisaR	oxanne	Brooklyn	Clinton Hill	Entire home/apt	89	
	4		Laura	Manhattan	East Harlem	Entire home/apt	80	
		minim	um_nights n	number_of_reviews	s last_review	reviews_per_mont	h \	
	0		1		9 2018-10-19	0.2	1	
	1		1	4.	5 2019-05-21	0.3	8	

```
2
                      3
                                          0
                                                     NaN
                                                                        NaN
      3
                                             2019-07-05
                                                                       4.64
                      1
                                        270
      4
                                                                       0.10
                     10
                                             2018-11-19
         calculated_host_listings_count
                                          availability_365
      0
                                       6
                                                        365
                                       2
                                                        355
      1
      2
                                       1
                                                        365
      3
                                       1
                                                        194
      4
                                                          0
                                       1
[95]: top_10_most_reviewed_listings = data.sort_values(by='number_of_reviews',_
       ⇒ascending=False).head(10).reset index()
      print(top_10_most_reviewed_listings[['id', 'name', 'number_of_reviews']])
               id
                                                                 name
                                                                       \
     0
         9145202
                                              Room near JFK Queen Bed
          903972
                                           Great Bedroom in Manhattan
     1
     2
                                      Beautiful Bedroom in Manhattan
          903947
                                        Private Bedroom in Manhattan
     3
          891117
     4
        10101135
                                              Room Near JFK Twin Beds
     5
         8168619
                                   Steps away from Laguardia airport
                            Manhattan Lux Loft.Like.Love.Lots.Look!
     6
          834190
     7
       16276632
                  Cozy Room Family Home LGA Airport NO CLEANING FEE
         3474320
                                  Private brownstone studio Brooklyn
     8
     9
                                     LG Private Room/Family Friendly
          166172
        number_of_reviews
     0
                       629
     1
                       607
     2
                       597
     3
                       594
     4
                       576
     5
                       543
     6
                       540
     7
                       510
     8
                       488
     9
                       480
[99]: plt.figure(figsize=(12, 6))
      ax = plt.barh(top_10_most_reviewed_listings['name'],__
       stop_10_most_reviewed_listings['number_of_reviews'], color='skyblue')
      plt.xlabel('Number of Reviews')
      plt.ylabel('Listing Name')
      plt.title('Top 10 Most Reviewed Listings')
      plt.gca().invert_yaxis()
      plt.show()
```



What is the average price for listings with availability less than 100 days?

[101]: listings_available_lessthan100_days = data[data['availability_365']<100]

[101]:	listings_available_lessthan100_days.head()										
[101]:		id			name host_id \						
	4	5022 Entire A	Apt: Spacious Studio	o/Loft by central	-						
	6	5121	•	BlissArtsSp	_						
	8	5203	Cozy Clean Guest Room - Family Apt 7490								
	10	5295	Beautiful :	lbr on Upper West	Side 7702						
	11	5441	Central Ma	anhattan/near Broa	adway 7989						
		host_name neig	ghbourhood_group	neighbourhood	room_type	price	\				
	4	Laura	Manhattan	East Harlem	Entire home/apt	80					
	6	Garon	Brooklyn Be	edford-Stuyvesant	Private room	60					
	8	MaryEllen	Manhattan	Upper West Side	Private room	79					
	10	Lena	Manhattan	Upper West Side	Entire home/apt	135					
	11	Kate	Manhattan	Hell's Kitchen	Private room	85					
		minimum_nights	number_of_reviews	s last_review rev	views_per_month \						
	4	10)	9 2018-11-19	0.10						
	6	45	49	9 2017-10-05	0.40						
	8	2		3 2017-07-21	0.99						
	10	5	5 53	3 2019-06-22	0.43						
	11	2	2 188	3 2019-06-23	1.50						
	calculated_host_listings_count availability_365										
	4		1	0							
	6		1	0							
	8		1	0							
	10		1	6							
	11		1	39							

```
[102]: listings_available_lessthan100_days['price'].mean()
[102]: 138.74156660949114
[103]:
       data.head()
[103]:
            id
                                                                      host_id \
                                                                name
       0
          2539
                                Clean & quiet apt home by the park
                                                                          2787
          2595
                                              Skylit Midtown Castle
                                                                          2845
       1
          3647
                               THE VILLAGE OF HARLEM...NEW YORK !
       2
                                                                      4632
       3
          3831
                                   Cozy Entire Floor of Brownstone
                                                                          4869
          5022
                Entire Apt: Spacious Studio/Loft by central park
                                                                          7192
            host_name neighbourhood_group neighbourhood
                                                                   room_type
                                                                               price
       0
                  John
                                   Brooklyn
                                                Kensington
                                                                Private room
                                                                                 149
       1
             Jennifer
                                  Manhattan
                                                   Midtown
                                                            Entire home/apt
                                                                                 225
       2
            Elisabeth
                                  Manhattan
                                                    Harlem
                                                                Private room
                                                                                 150
       3
          LisaRoxanne
                                   Brooklyn
                                             Clinton Hill
                                                             Entire home/apt
                                                                                  89
       4
                 Laura
                                  Manhattan
                                               East Harlem
                                                             Entire home/apt
                                                                                  80
          minimum nights
                           number_of_reviews last_review
                                                             reviews_per_month
       0
                        1
                                             9
                                                2018-10-19
                                                                           0.21
       1
                        1
                                            45
                                                2019-05-21
                                                                           0.38
       2
                        3
                                                                            NaN
                                             0
                                                       NaN
                                                                           4.64
       3
                        1
                                           270
                                                2019-07-05
       4
                       10
                                                2018-11-19
                                                                           0.10
          calculated_host_listings_count
                                             availability_365
       0
                                          6
                                                           365
       1
                                          2
                                                           355
       2
                                          1
                                                           365
       3
                                          1
                                                           194
       4
                                          1
                                                             0
```

1.3 Addressed the below questions

- 1. Which are the top 10 host ids with the most listings?
- 2. Which are the unique neighborhoods in the dataset?
- 3. What are the top 10 neighborhoods with the most property listings?
- 4. What are the top 10 host names with the most listings?
- 5. Which are the unique neighborhood groups in the dataset?
- 6. What are the property listing counts for each neighborhood group?
- 7. Which neighborhoods belong to each neighborhood group?
- 8. What types of rooms are available in the dataset?

- 9. Which room type contributes to the highest number of listings in NYC?
- 10. What is the average price of all listings?
- 11. How many hosts have more than one listing?
- 12. Which neighborhood has the highest average price?
- 13. How many listings are available throughout the year?
- 14. What is the average minimum number of nights required for a stay?
- 15. How many listings have never been reviewed?
- 16. What is the average price for each neighborhood group, and which neighborhood group has the highest average price?
- 17. What is the average number of listings per host?
- 18. How many listings are there per neighborhood group by room type?
- 19. What are the top 10 most reviewed listings?
- 20. What is the average price for listings with availability of less than 100 days?

1.4 Key Insights

The Airbnb NYC 2019 dataset is considered to be a very rich dataset to explore and uncover trends and patterns in rental listings. The dataset contains 16 attributes and 48895 rows in total. We analyzed the dataset using Python and data visualization libraries and uncovered useful insights.

Firstly we identified the top 10 hosts with the highest number of rental listings on Airbnb. The host with the host ID 219517861 is recorded with the highest number of property listings on Airbnb with a total of 327 listings.

It was seen that throughout the entire year, only 1295 listings were available. Data also reveals that the average number of listings per host is just 1 listing, suggesting that most hosts manage only a single property.

Our dataset includes 221 unique neighborhoods, with Williamsburg standing out as the neighborhood with the most property listings, totaling 3,920 listings. It is followed by Bedford-Stuyvesant and Harlem, with 3,714 and 2,658 listings, respectively. Fort Wadsworth has the highest average price, at \$800.

We analyzed the top 10 hosts to identify the one with the most property listings on Airbnb. Our findings show that Michael holds the top spot with 417 listings, making him the host with the highest number of properties. He is followed by David, who has the second-highest number of listings at 403. Additionally, our data reveals that 5,154 hosts have more than one listing on Airbnb.

The dataset includes five distinct neighborhood groups: Brooklyn, Manhattan, Queens, Staten Island, and the Bronx. It was found that Manhattan has the highest number of property listings on Airbnb, with a total of 21,661 listings, while Staten Island has the fewest property listings. The data shows that Manhattan has the highest average property prices, making it the most expensive area to live in NYC. In contrast, the Bronx has the lowest average property prices, making it the

most affordable area in the city. Whereas Queens and Staten Island appear to have similar property prices.

The Airbnb NYC listings primarily consist of three room categories: Private room, Entire home/apt, and Shared room. Among these, the Entire home/apt has the highest number of listings, making it the most common room type, while the Shared room has the fewest listings. The average price for all listings is \$152.72.

On average, Airbnb listings in this dataset require a minimum stay of 7 nights. This indicates that, generally, hosts prefer longer bookings, possibly to reduce the frequency of turnover and related maintenance tasks.

According to the dataset, there are a total of 10,052 property listings on Airbnb that have not received a single review. This means that these listings have either not been booked by guests or, if they have been, no feedback has been provided by those who stayed. This lack of reviews could impact the visibility of these listings to potential guests, as reviews often play a significant role in helping users decide on accommodations.

[]: