Welcome to New York City, one of the most-visited cities in the world. There are many Airbnb listings in New York City to meet the high demand for temporary lodging for travelers, which can be anywhere between a few nights to many months. In this project, we will take a closer look at the New York Airbnb market by combining data from multiple file types like `.csv`, `.tsv`, and `.xlsx`.

Recall that **CSV**, **TSV**, and **Excel** files are three common formats for storing data. Three files containing data on 2019 Airbnb listings are available to you:

data/airbnb_price.csv This is a CSV file containing data on Airbnb listing prices and locations.

- `listing_id`: unique identifier of listing
- `price`: nightly listing price in USD
- `nbhood_full`: name of borough and neighborhood where listing is located

data/airbnb_room_type.xlsx This is an Excel file containing data on Airbnb listing descriptions and room types.

- `listing_id`: unique identifier of listing
- `description`: listing description
- `room_type`: Airbnb has three types of rooms: shared rooms, private rooms, and entire homes/apartments

data/airbnb_last_review.tsv This is a TSV file containing data on Airbnb host names and review dates.

- `listing_id`: unique identifier of listing
- `host_name`: name of listing host
- `last_review`: date when the listing was last reviewed

In [1]:
```
# We've loaded your first package for you! You can add as many cells as you need.
import numpy as np
import pandas as pd

airbnb_last_review = pd.read_csv("data/airbnb_last_review.tsv", sep = "\t")
airbnb_last_review['last_review'] = pd.to_datetime(airbnb_last_review['last_review'], format = '%B %d %Y')
earliest_date = airbnb_last_review['last_review'].min()
recent_date = airbnb_last_review['last_review'].max()
```

In [2]:
```
airbnb_room_type = pd.ExcelFile('data/airbnb_room_type.xlsx')
print(airbnb_room_type.sheet_names)
df1 = airbnb_room_type.parse('airbnb_room_type')
```

```
['airbnb_room_type']
```

In [3]:
```
df1["proper_room_type"] = df1["room_type"].str.lower()
```

In [4]:
```
room_counts = df1["proper_room_type"].value_counts(dropna = False)
```

In [5]:
```
private_count = room_counts[1]
```

In [6]:
```
airbnb_price = pd.read_csv("data/airbnb_price.csv")
```

In [7]:
```
airbnb_price.head()
airbnb_price["price"] = airbnb_price["price"].str.strip()
```

In [8]:
```
airbnb_price["price_adj"] = airbnb_price["price"].str.replace(" dollars","")
```

In [9]:
```
airbnb_price["price_adj"] = airbnb_price["price_adj"].astype("float")
```

In [10]:
```
df1 = pd.DataFrame(df1)
```

In [ ]:
```

```

In [11]:
```
#merge airbnb_price and df1 together
airbnb_price_room = airbnb_price.merge(df1, on = "listing_id", how = "outer", suffixes= ("_price", "room"))
airbnb_price_room.head()
```

Out[11]:

	listing_id	price	nbhood_full	price_adj	description	room_type	proper_room_type
0	2595	225 dollars	Manhattan, Midtown	225.0	Skylit Midtown Castle	Entire home/apt	entire home/apt
1	3831	89 dollars	Brooklyn, Clinton Hill	89.0	Cozy Entire Floor of Brownstone	Entire home/apt	entire home/apt
2	5099	200 dollars	Manhattan, Murray Hill	200.0	Large Cozy 1 BR Apartment In Midtown East	Entire home/apt	entire home/apt
3	5178	79 dollars	Manhattan, Hell's Kitchen	79.0	Large Furnished Room Near B'way	private room	private room
4	5238	150 dollars	Manhattan, Chinatown	150.0	Cute & Cozy Lower East Side 1 bdrm	Entire home/apt	entire home/apt

In [12]:
```
average_price = airbnb_price_room["price_adj"].mean().round(2)
```

In [13]:
```
my_dict = [{"first_reviewed" : earliest_date,
     "last_reviewed" : recent_date,
     "nb_private_rooms" : private_count,
      "avg_price": average_price}]
```

In [14]:
```
review_dates = pd.DataFrame(my_dict)
```

In [15]:
```
display(review_dates)
```

	first_reviewed	last_reviewed	nb_private_rooms	avg_price
0	2019-01-01	2019-07-09	11356	141.78