PROJECT: EXPLORING AIRBNB MARKET TRENDS

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

```
# Import necessary packages
import pandas as pd
import numpy as np

# Begin coding here ...

#1 loading data: (files -> Dfs)
airbnb_price_df = pd.read_csv("data/airbnb_price.csv")

airbnb_room_type_df = pd.read_excel("data/airbnb_room_type.xlsx")

airbnb_last_review_df = pd.read_csv("data/airbnb_last_review.tsv", delimiter = '\t')

print("price cols:",airbnb_price_df.columns)
print("room cols: ",airbnb_room_type_df.columns)
print("last_review cols:",airbnb_last_review_df.columns)
```

```
price cols: Index(['listing_id', 'price', 'nbhood_full'], dtype='object')
room cols:  Index(['listing_id', 'description', 'room_type'], dtype='object')
last_review cols: Index(['listing_id', 'host_name', 'last_review'], dtype='object')
```

```
#2 Merging the three DataFrames

airbnb_df = pd.merge(airbnb_price_df, airbnb_room_type_df, on='listing_id',
how='inner')
airbnb_df = pd.merge(airbnb_df, airbnb_last_review_df, on = "listing_id", how =
"inner")

print(airbnb_df.columns)
print(airbnb_df.head(4))
```

```
Index(['listing_id', 'price', 'nbhood_full', 'description', 'room_type',
       'host_name', 'last_review'],
      dtype='object')
   listing_id        price  ...    host_name   last_review
0        2595  225 dollars  ...     Jennifer   May 21 2019
1        3831   89 dollars  ...  LisaRoxanne  July 05 2019
2        5099  200 dollars  ...        Chris  June 22 2019
3        5178   79 dollars  ...     Shunichi  June 24 2019

[4 rows x 7 columns]
```

```
#3 Determining the earliest and most recent review dates

airbnb_df["last_review"] = pd.to_datetime(airbnb_df["last_review"])

earliest_date = airbnb_df["last_review"].min()
most_recent_date = airbnb_df["last_review"].max()

print("earliest review date: ",earliest_date)
print("most recent review date: ",most_recent_date)
```

```
earliest review date:  2019-01-01 00:00:00
most recent review date:  2019-07-09 00:00:00
```

```
#4 Finding how many listings are private rooms

airbnb_df["room_type"] = airbnb_df["room_type"].str.lower()
value_counts = airbnb_df["room_type"].value_counts()

private_rooms = value_counts["private room"]

print(value_counts)
```

```
entire home/apt    13266
private room       11356
shared room          587
Name: room_type, dtype: int64
```

```
#5 Finding the average price of listings
airbnb_df["price"] = airbnb_df["price"].str.strip("dollars").astype(float)
avg_price = round(airbnb_df["price"].mean(),2)
print("avg price",avg_price)
```

```
avg price 141.78
```

```
#6 Creating a DataFrame with the four solution values
review_dates = pd.DataFrame([[earliest_date, most_recent_date, private_rooms,
avg_price]])
review_dates.columns = ["first_reviewe
d","last_reviewed","nb_private_rooms","avg_price"]
print(review_dates)
```

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