Lab-4-H-AirBnB

Answer following question with AirBnB dataset

In [ ]:
```
##Import the libraries
import pandas as pd
import numpy as np
import seaborn as sns
```

In [ ]:
```
#Load the dataset
df =
```

In [ ]:
```
#check the top 5 rows
```

Task 1

Basic Descriptive Statistics

- 1.1 How many AirBnB listings are there in the city of Amsterdam?

In [ ]:
```
##Write your code here
```

Answer:

- 1.2 What is the average price of all the AirBnB listings in Amsterdam?

In [ ]:
```
##Write your code here
```

Answer:

- 1.3 - What is the average rating received by all the AirBnB listings in Amsterdam? *Hint* - Use the `review_scores_rating` column for answering this question

In [ ]:
```
##Write your code here
```

Answer:

Task 2

Plot a histogram for the following variables and observe their distribution. Choose the parameters like bin width, number of bins, etc. as per your choice.

- `price`
- `number_of_reviews`
- `review_scores_rating`

In [ ]:
```
##Write your code here for plotting the distribution of price
```

Observation:

In [ ]:
```
##Write your code here for plotting the distribution of reviews
```

Observation:

In [ ]:
```
##Write your code here for plotting the distribution of ratings
```

Observation:

Task 3

Plot a visualization to show the number of listings for each `room_type`

Which `room_type` has the highest number of listings?

In [ ]:

Observation -

Task 4

You want to observe the relationship between the reviews(given by `number_of_reviews`) and the ratings received (`review_scores_rating`) by different AirBnB listings.

For this, plot both a scatterplot and a jointplot. What can you say about the relationship between ratings and reviews?

In [ ]:
```
##Write your code here for scatterplot
```

In [ ]:
```
##Write your code here for jointplot
```

Observation -

Task 5

AirBnB has been adding quite a few listings in the city of Amseterdam since they started operating in 2008. Plot a lineplot to observe how the average `review_scores_rating` has changed across the different years (use `host_since_Year` column) AirBnB has been operating in the city.

In [ ]:
```
### Write your code for lineplot here
```

Observation -

Task 6

You wish to identify the relationship between the various ratings each of the listing has received from the customers. These ratings have been summarized below:

In [ ]:
```
review_columns = ['review_scores_rating',
 'review_scores_accuracy',
 'review_scores_cleanliness',
 'review_scores_checkin',
 'review_scores_communication',
 'review_scores_location',
 'review_scores_value']
```

Plot a heatmap of the correlation matrix of the above ratings variables and document your observations. You can use this link to understand these variables further - https://www.airbnb.co.in/help/article/1257/star-ratings (https://www.airbnb.co.in/help/article/1257/star-ratings)

In [ ]:

Observation -

Task 7

Analyze the listings' prices across the following neighborhoods using a categorical boxplot

- 'Westerpark', 'Oud-Noord', 'Noord-West','Zuid'

In [ ]:
```
nc = ['Westerpark', 'Oud-Noord', 'Noord-West','Zuid']
```

In [ ]:
```
### Write your code for subsetting the data for only the above neighborhoods here
```

In [ ]:
```
### Write your code for plotting the categorical boxplot here
```

Observation -

In [ ]:
```

```