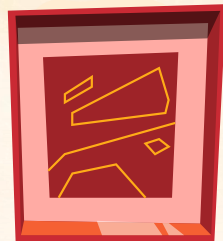


AIRBNB NEW YORK CITY ANALYSIS

Danni Hernando Haryoyudhanto



TABLE OF CONTENTS



01

DATA UNDERSTANDING

Current business problem

02

DATA PREPARATION

Collection and preparation
of required dataset

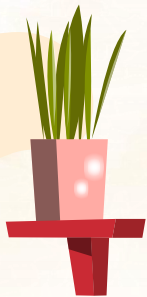
03

DATA VISUALIZATION

Exploring the data by analyzing its statistics and visualizing the values of features and correlations between different features. Explaining the process and the results

01

DATA UNDERSTANDING





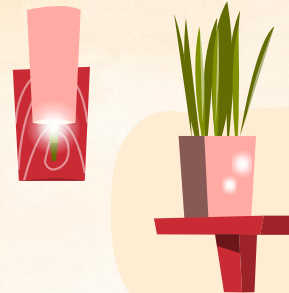
Airbnb: “Air Bed and Breakfast” A service that lets property owners rent out their spaces to travelers looking for a place to stay. Travelers can rent a space for multiple people to share, a shared space with private rooms, or the entire property for themselves. Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present more unique, personalized way of experiencing the world.

GOALS

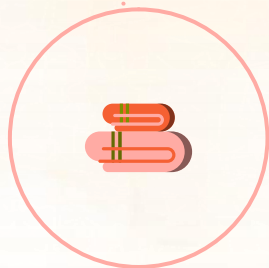


- Converting untidy data into a suitable format for analysis.
- Knowing the correlation between columns
- Knowing how many types of rooms
- What is the average price of Airbnb nights in New York
- How the average monthly Airbnb price compared to the private monthly market
- How Airbnb listing prices compare in the New York area 7.knowing the number of room types and prices in massing" regions

DATASET INFORMATION



48.895 ROWS



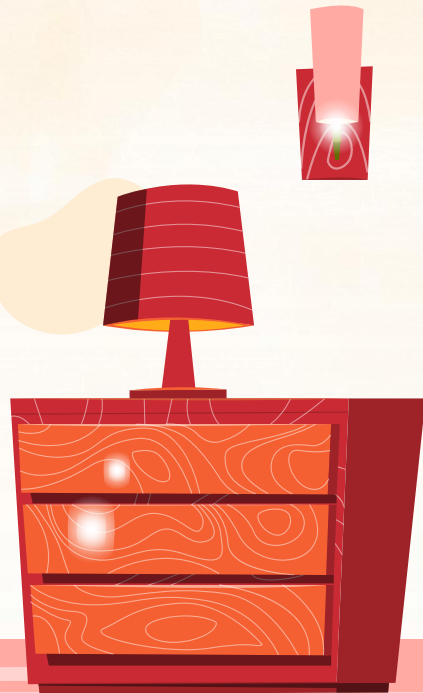
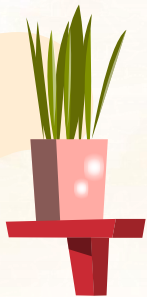
48,895 rows of order data where have
10052 missing values in the last_review
and reviews_per_month columns

16 FEATURES

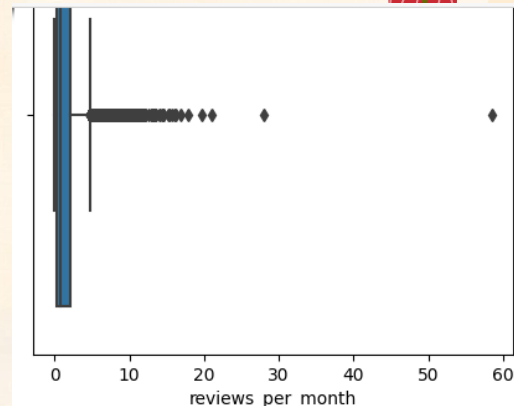
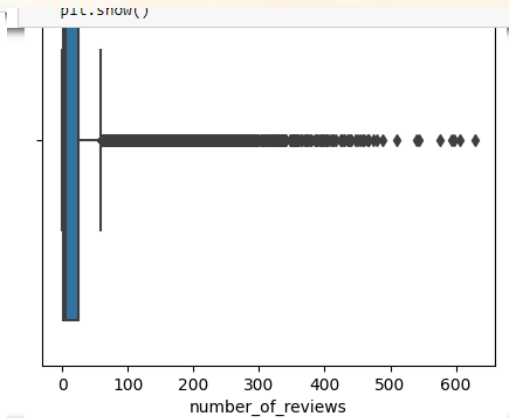
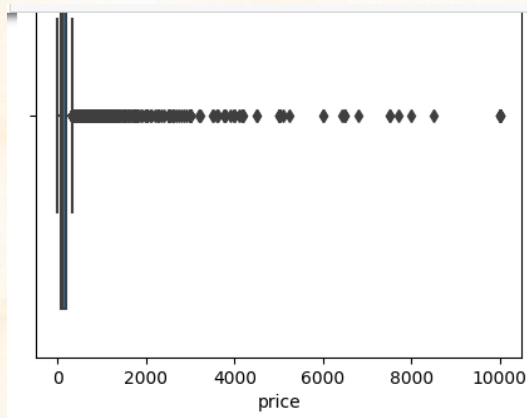
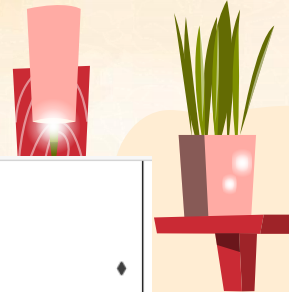
- | | |
|-----------------------|-----------------------------------|
| 0 id | |
| 1 name | |
| 2 host_id | |
| 3 host_name | |
| 4 neighbourhood_group | |
| 5 neighbourhood | |
| 6 latitude | 8 room_type |
| 7 longitude | 9 price |
| | 10 minimum_nights |
| | 11 number_of_reviews |
| | 12 last_review |
| | 13 reviews_per_month |
| | 14 calculated_host_listings_count |
| | 15 availability_365 |

02

DATA PREPARATION



DATA CLEANSING



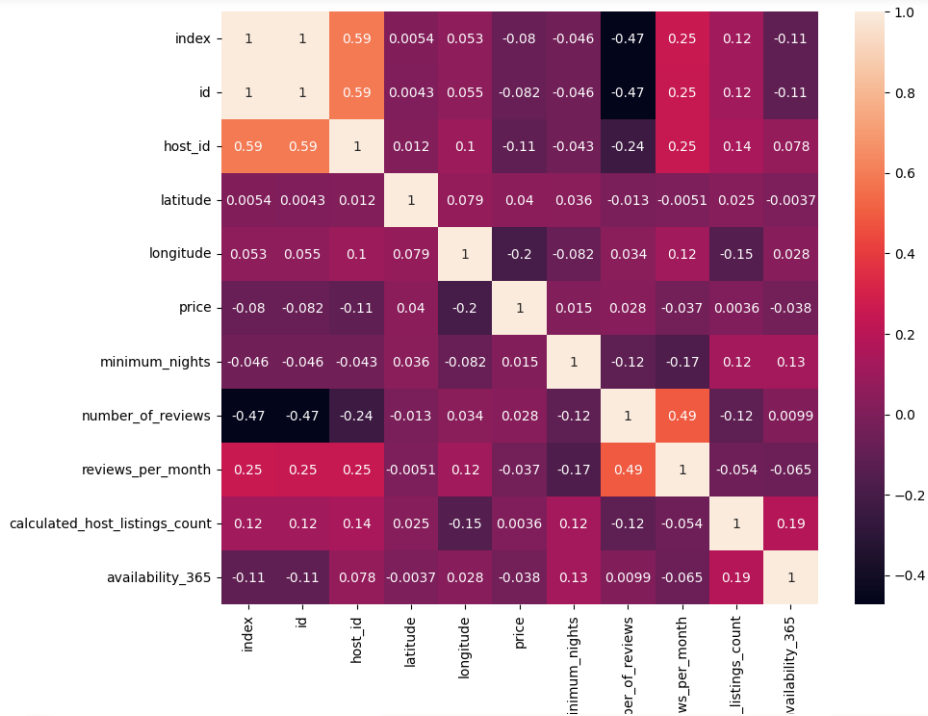
Featuring Outliers

```
In [3]: #Menghapus data yang kosong NaN  
df_airbnb = pd.read_csv('./airbnb.csv')  
df_airbnb.dropna(inplace=True)  
df_airbnb.shape
```

```
Out[3]: (38821, 16)
```

Delete empty data

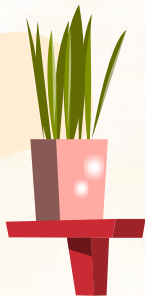
MATRIX CORRELATION



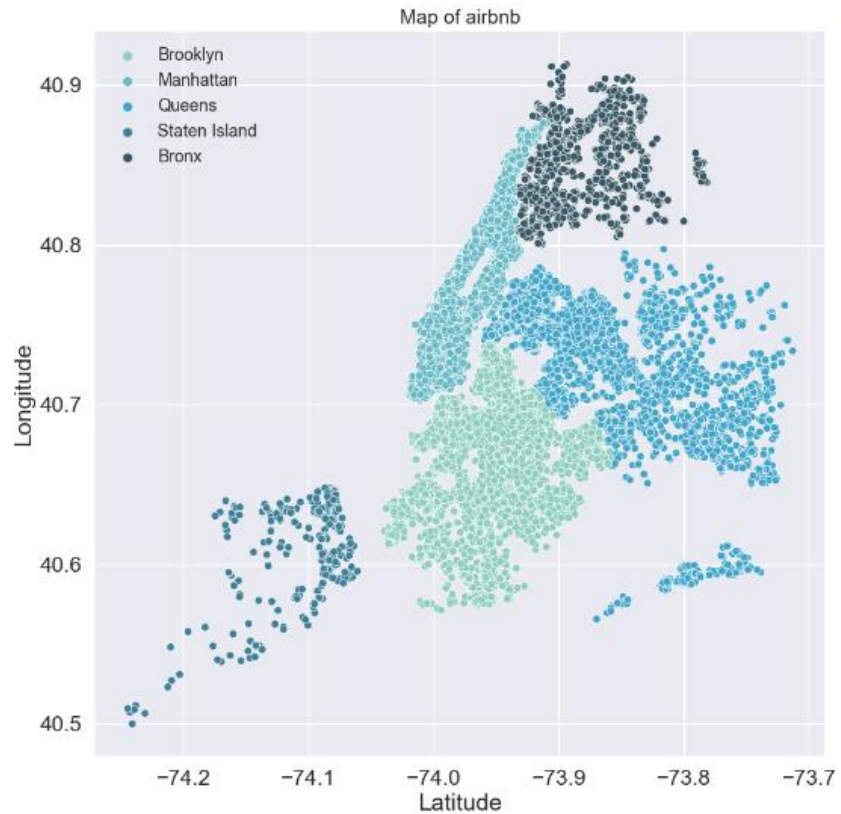
highest correlation is 0.59 which is between "index"- "host_id" and "id"- "host_id" and 0.49 which is between "number_of_reviews"- "review_per_month"

03

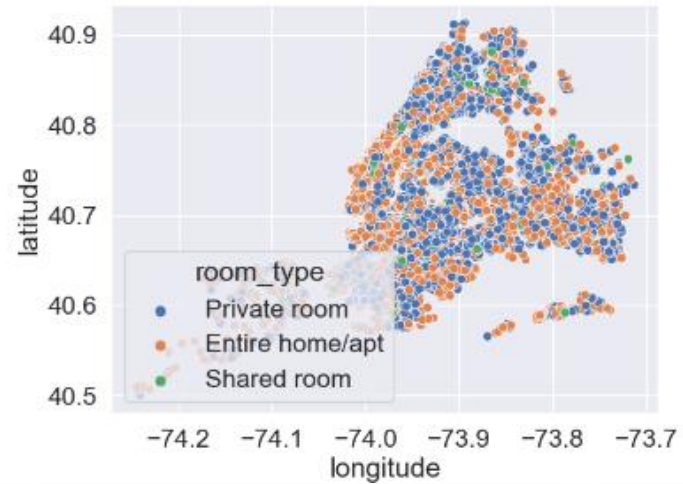
DATA VISUALIZATION



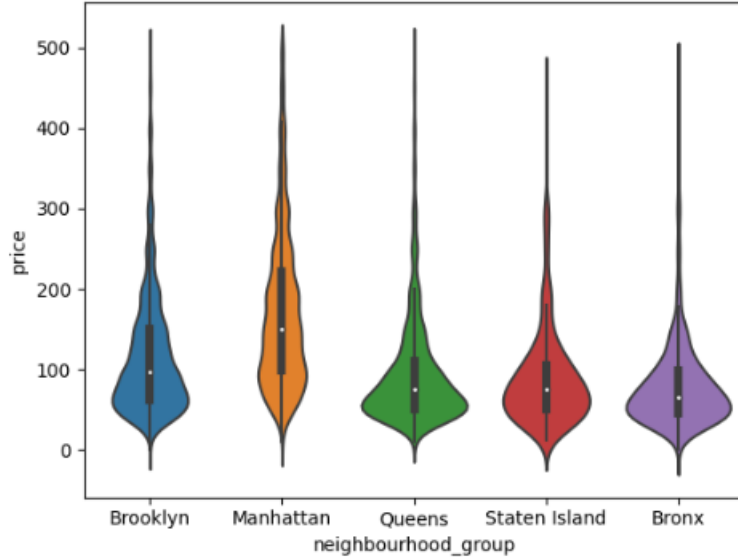
neighborhood_group deployment



Type_rooms deployment

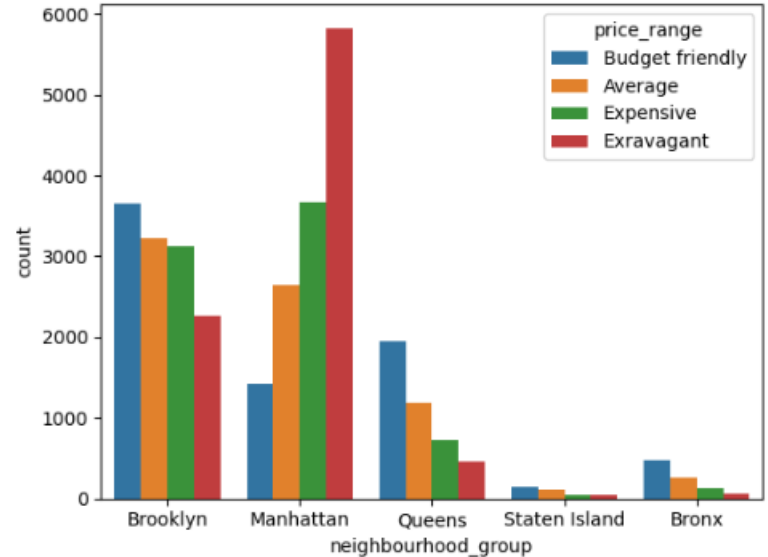


Visualization of the average" price of each area



Visualization of a comparison of several prices ranging from cheapest to most expensive

Out[36]: <Axes: xlabel='neighbourhood_group', ylabel='count'>



THANKS!

Does anyone have any questions?

daninando098@gmail.com

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

