Exploring Airbnb of New York City

By Christian, Alexis, Huiting

TABLE OF CONTENTS

01 → **Dataset**

02 → **Data Cleaning**

03 → **Graphs**

04 → **Summary**

Key Variables

neighbourhood_ group
The main region of the city

room_type
Type of room

number_of _reviews
Number of reviews received

name
Property description of the room

Latitude & longitude
Property longitude & latitude

price
The price for one night

last_review
The date of the last review

Combined dataset

```{r}
dim(AB_NYC) # the data from 2019
dim(AB_NYC_2023) # the data from 2023
dim(combined_dataset) # we combined the datasets and remove the two new columns from the 2023 dataset
```

```
[1]  48895     16
[1]  42931     18
[1]  91826     16
```

Missing Value

- There are 40764 missing values in the combined dataset.
- Most of the missing values are located at the number of reviews and last date of review columns.
- We choose to remove the missing rows with host name missing. We remove 26 rows from the dataset.
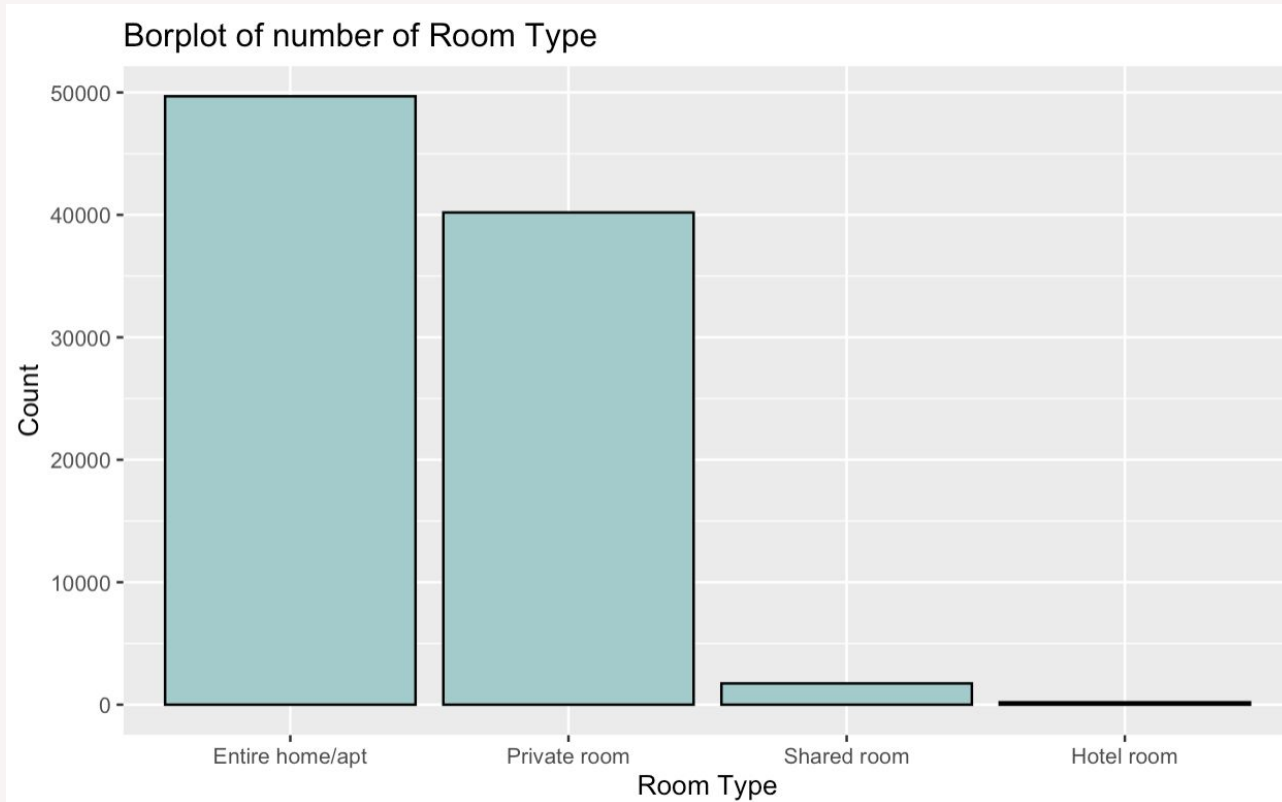
Duplicated

- There are 11 duplicated data.
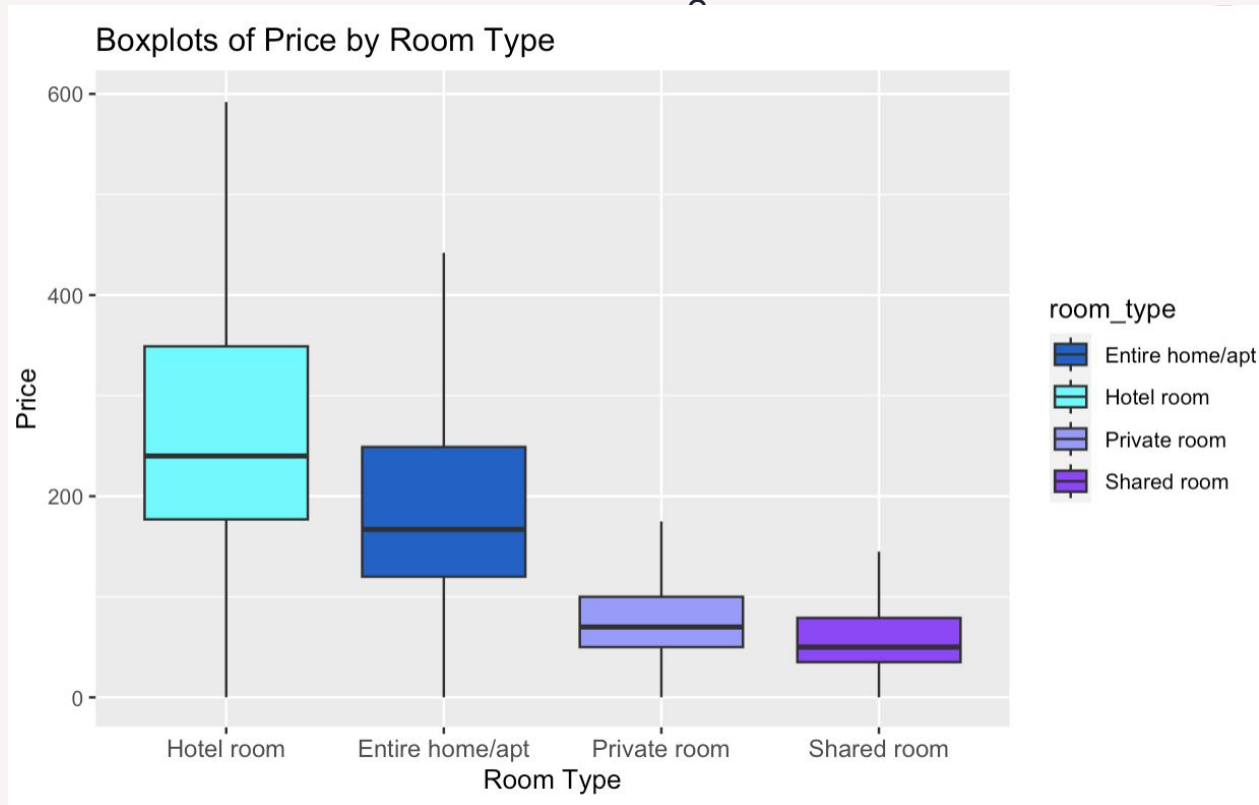- After remove them, there are 91789 rows and 16 columns in the combined dataset

Outliers

- There are a lot of outliers when we create the graphes.
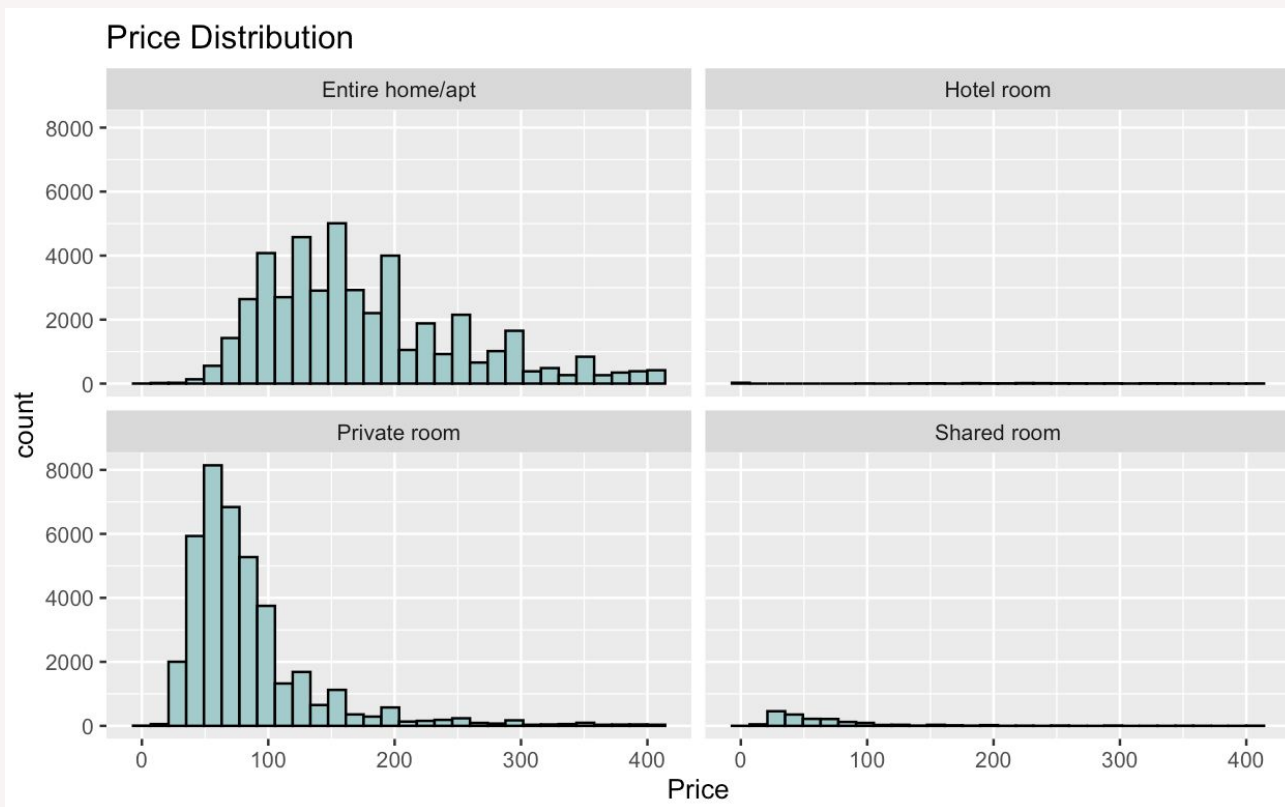- We choose to not showing the outliers for the graphs for better visualization.

Barplot: What kind of room is more common in New York City Airbnb?

Borplot of number of Room Type

Box Plots: What is the price difference between different types of

Boxplots of Price by Room Type

Histograms

Price Distribution

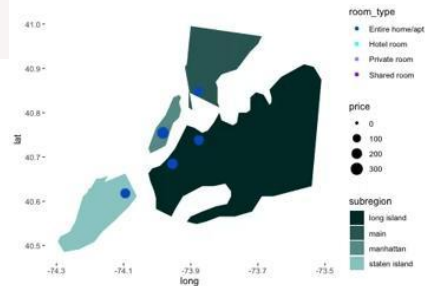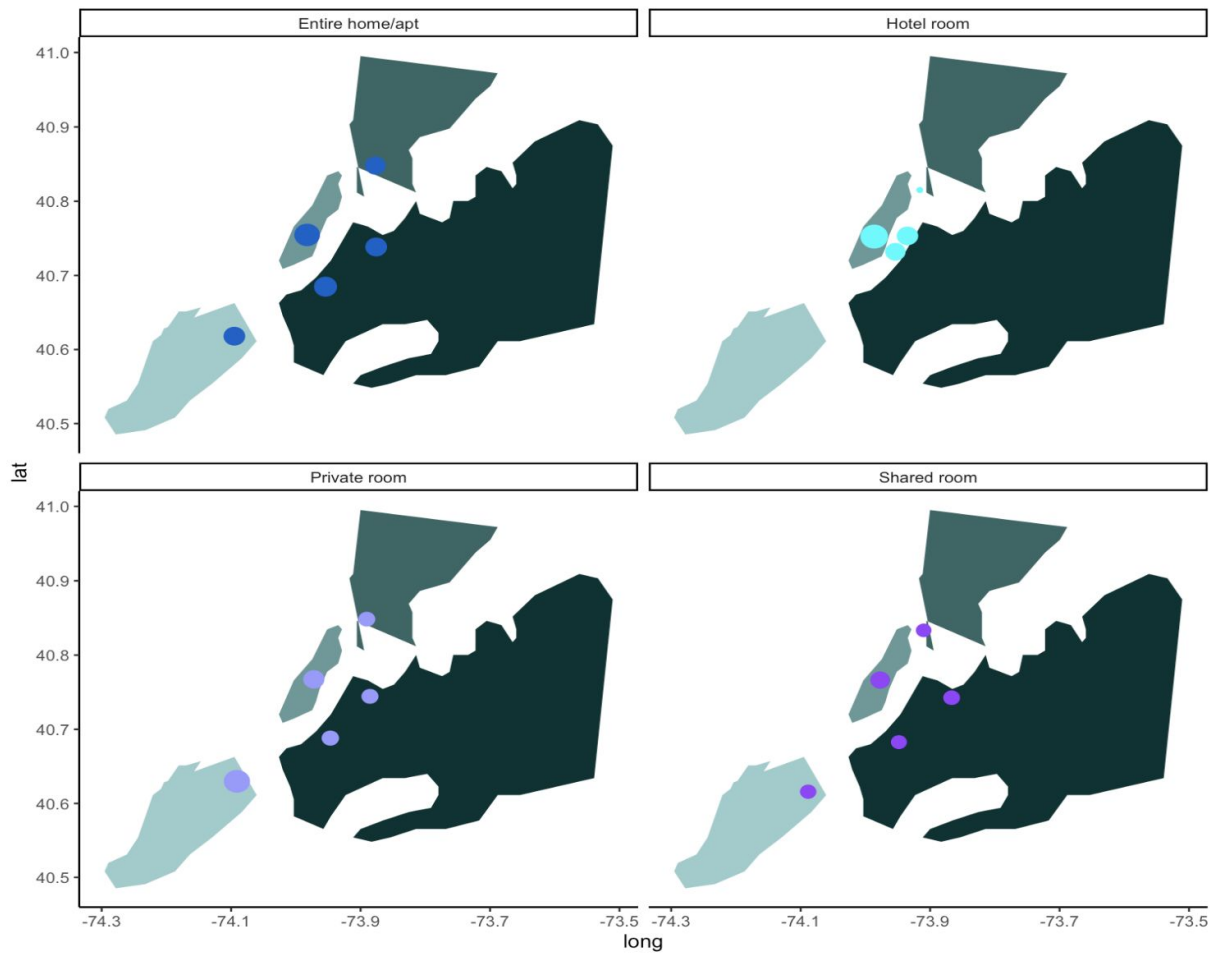# Box Plots: What are the most expensive regions to stay in New York City?

MAP

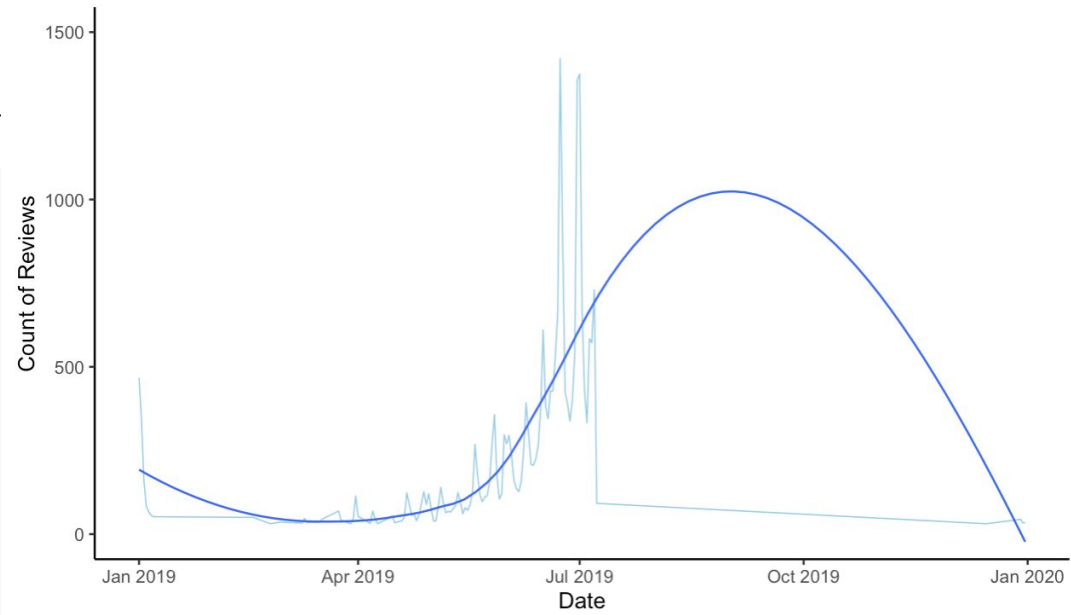# Time Series Plot
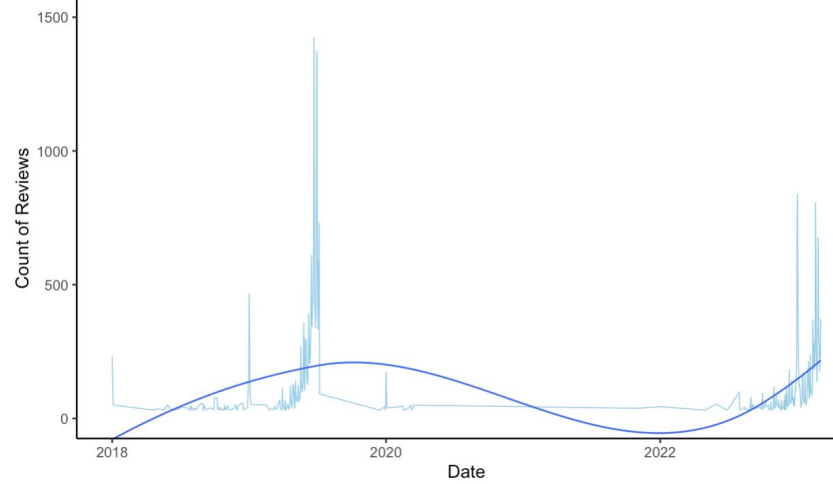
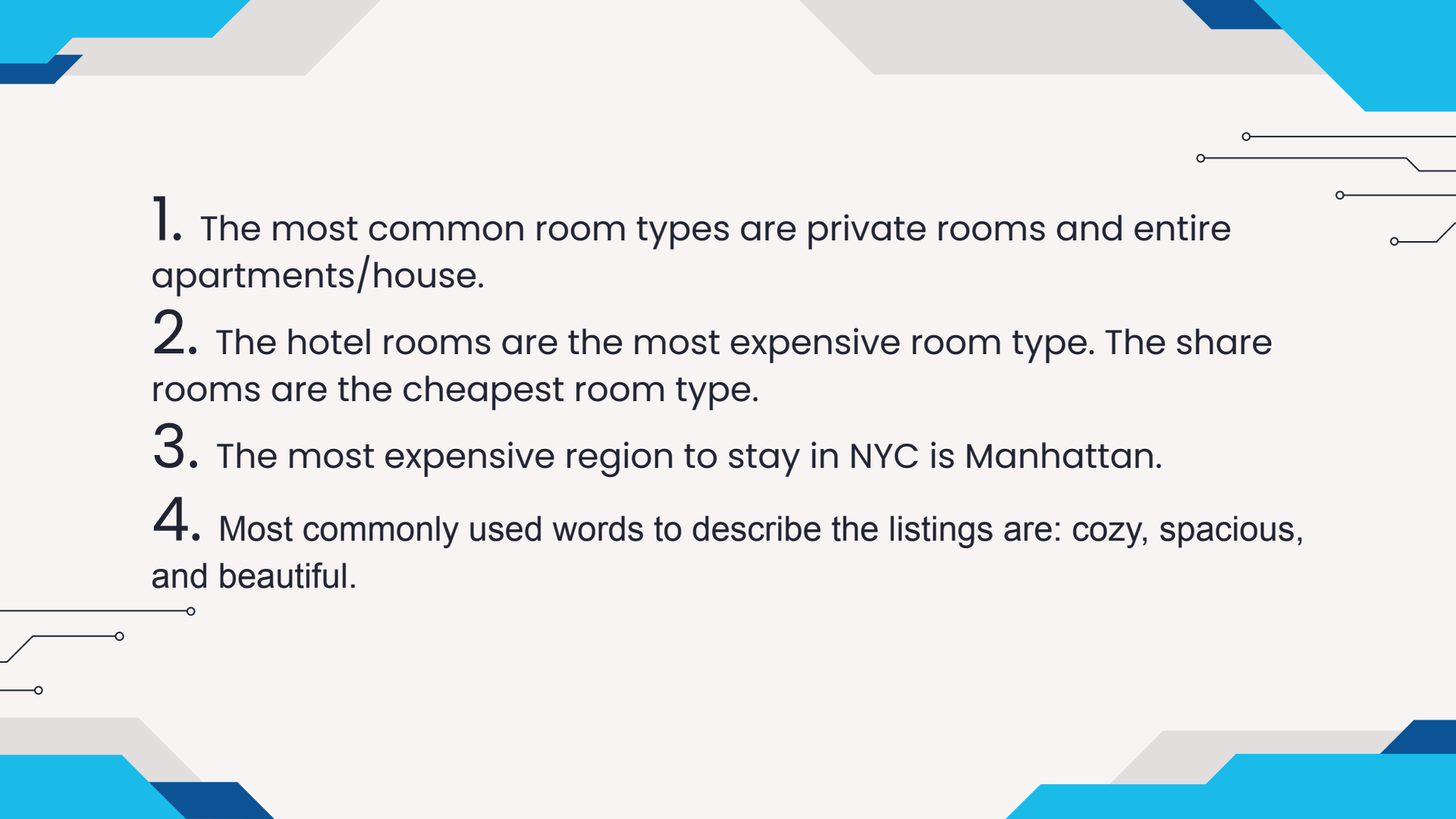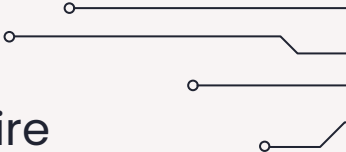| Year <dbl>	Count of Last Reviews <int>
2019	27439
2023	13291
2022	8570
2018	7478
2017	4409
2016	3982
2015	2154
2020	2109
2021	1592
2014	301
2013	79
2012	37
2011	10

What kind of descriptions about the rooms are frequently used by the host?
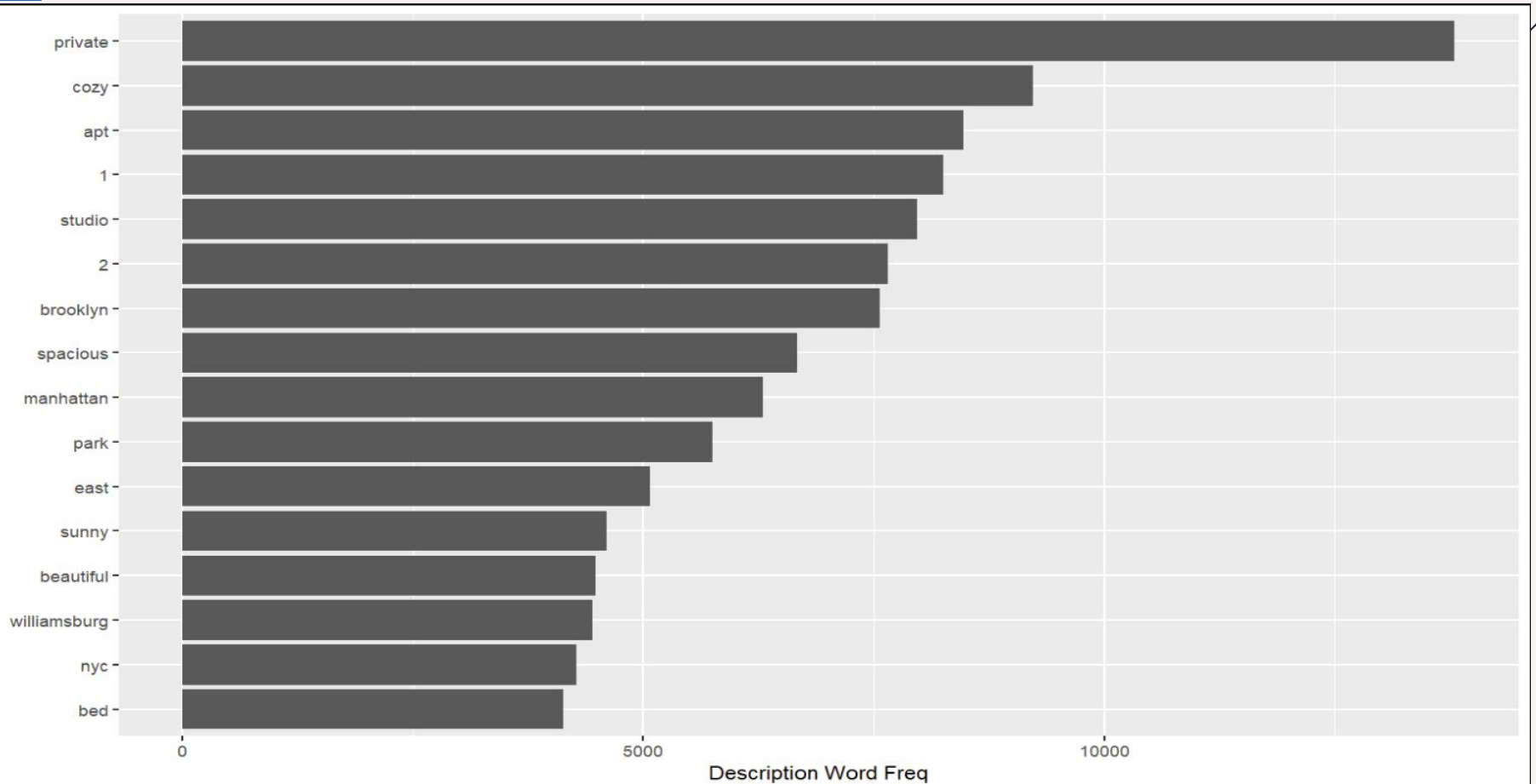
Summary

1. The most common room types are private rooms and entire apartments/house.

2. The hotel rooms are the most expensive room type. The share rooms are the cheapest room type.

3. The most expensive region to stay in NYC is Manhattan.

4. Most commonly used words to describe the listings are: cozy, spacious, and beautiful.

Questions?

Resources

- Dataset:
  https://www.kaggle.com/datasets/dgomonov/new-york-city-airbnb-open-data

  https://www.kaggle.com/datasets/godofoutcasts/new-york-city-airbnb-2023-public-data?select=NYC-Airbnb-2023.csv

Thank You!