

Capstone Project-01


AirBnb Booking Analysis

by

Mrutyunjaya Sahoo

What is AirBnb?

- Airbnb is an online community marketplace that connects people looking rent their homes with people who are looking for accomodation.
- Founded in an Francisco in 2008 as a start-up, the company has become a worldwide booking platform.
- Today it contributes the more than 60 million people in 192 countries



Problem

Price is an important concern for customer booking travel online.

Hotels leave you disconnected from the city and its culture.

No easy way exists to book a room with a local or become a host.



Solution

A web platform where users can rent out their space to host travelers to



Save Money
when traveling



Make Money
when hosting



Share Culture
local connection to the city



Market Validation

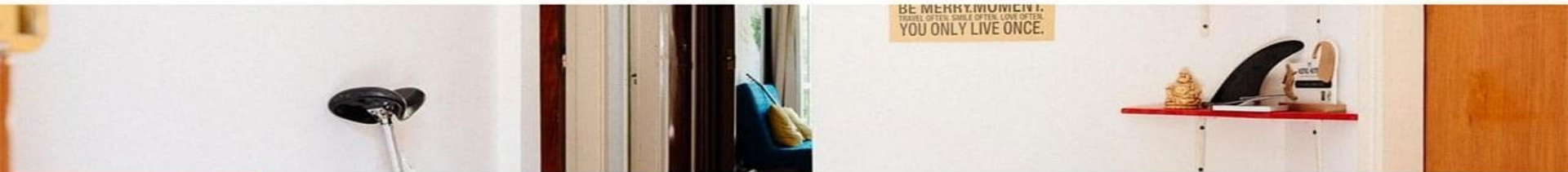
craigslist.com

670,000
TOTAL USERS

couchsurfing.com

17,000
TEMPORARY HOUSING LISTINGS

In SF & NYC from
7/09 to 7/16



BE MERRY, MOMENT.
TRAVEL OFTEN. SMILE OFTEN. LOVE OFTEN.
YOU ONLY LIVE ONCE.



Market Size

2+
Billion

**Trips Booked
(WorldWide)**

Total Available Market

560+
Million

Budget&Online

Serviceable Available
Market

84
Million

Trips w/ AirBnB

Share of Market

15% Available Market

*source: Travel Industry Assn. of America and World Tourism Organization.

**source: comScore

CONTENT:

- **Introduction**
- **Overview of the dataset**
- **Data cleaning**
- **Exploratory Data Analysis**
 - **Observation**
- **Conclusion**

INTRODUCTION



Airbnb is an open online platform where people list their own housing for rent. Since 2008, it has grown in popularity and specially for those community which frequently use to travel. It is becoming a strong competitor to the hotel industry. It has millions of listing, which generate lots of data. I'm analyzing these data for making business decision, for looking best room type etc.

Overview of the Dataset

16 Columns

&

48895 Rows

Column Information

- **id : Unique Id.**
- **name :Name of listing.**
- **host_id : unique id for each listed host.**
- **host_name : Name of the host.**
- **neighbourhood_group : Location**
- **neighborhood : Area**
- **latitude : Latitude coordinates**
- **longitude : Longitude coordinates**

Column Information

- **room_type** : Type of listing
- **price** : Price of listing
- **minimum_nights** : minimum nights to be paid for rooms
- **number_of_reviews** : No. of reviews
- **last_review** : Content of the last review
- **reviews_per_month** : Number of checks per month
- **calculated_host_listings_count** : Total count of hosting list
- **availability_365** : Availability around the year

Data Cleaning

```
[ ] # removing null values
df1['name'].fillna('na',inplace=True)
df1['host_name'].fillna('na',inplace=True)
```

/usr/local/lib/python3.8/dist-packages/pandas/core/generic.py:6392: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

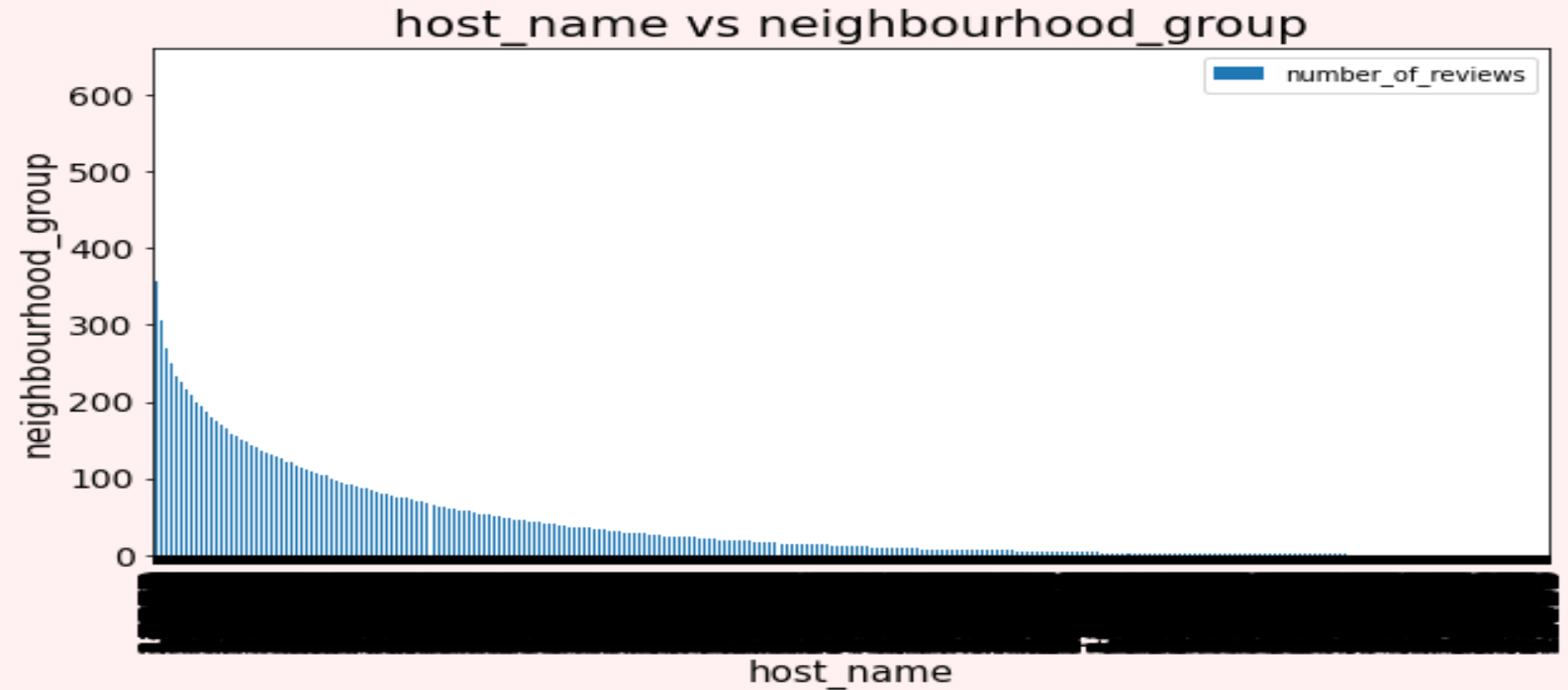
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
return self._update_inplace(result)

```
[ ] df1.isna().sum()
```

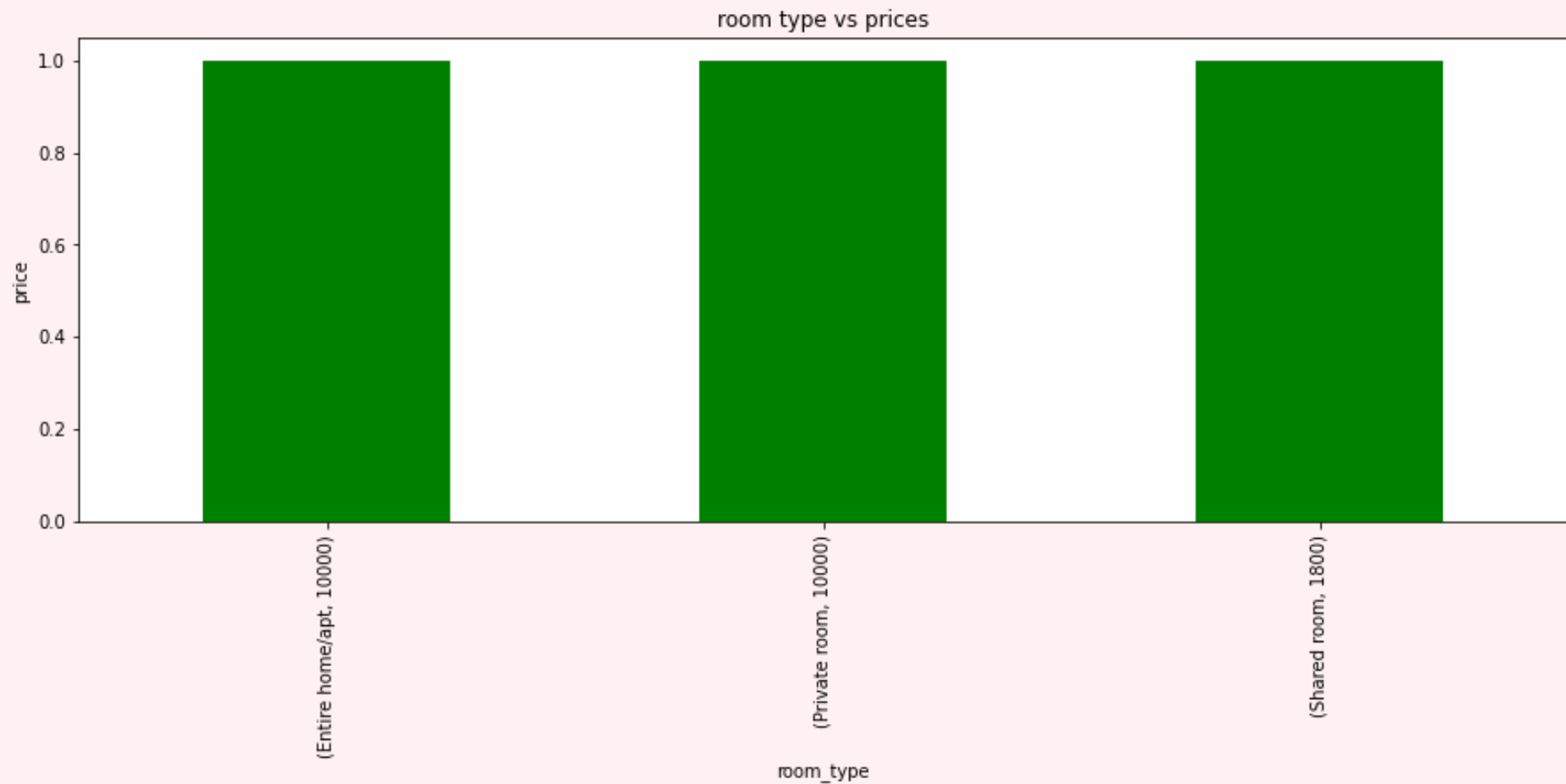
```
id                0
name              0
host_id           0
host_name         0
neighbourhood_group  0
neighbourhood     0
room_type         0
price             0
minimum_nights    0
number_of_reviews  0
calculated_host_listings_count  0
availability_365   0
dtype: int64
```

Exploratory Data Analysis

host vs location analysis:

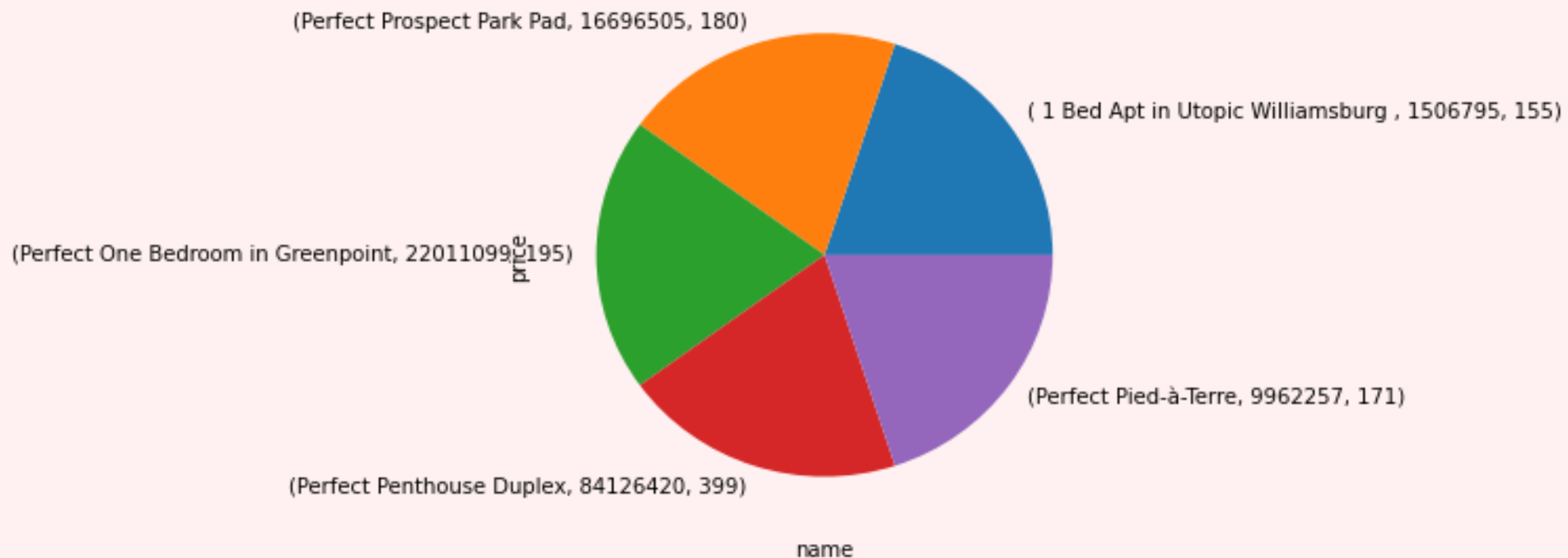


different room type vs prices

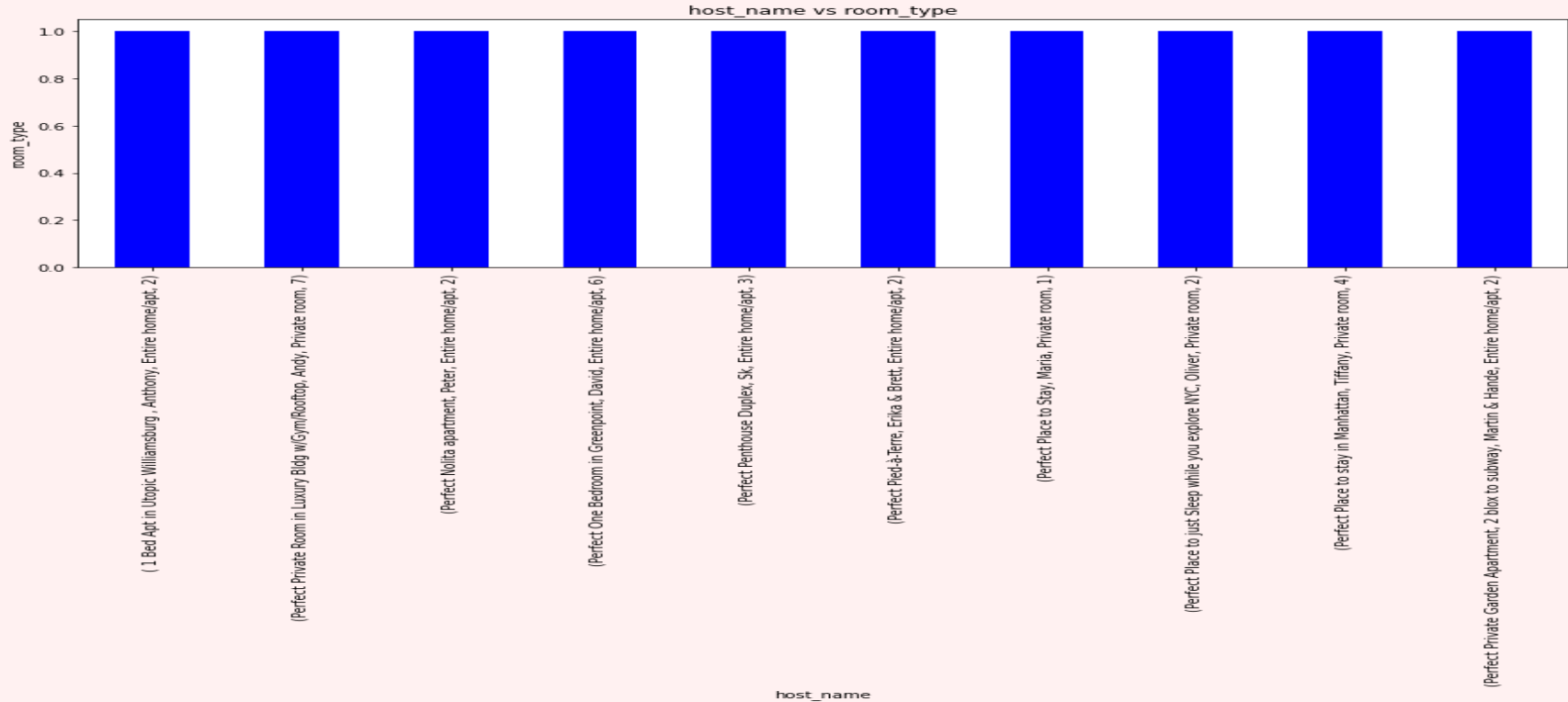


house vs price

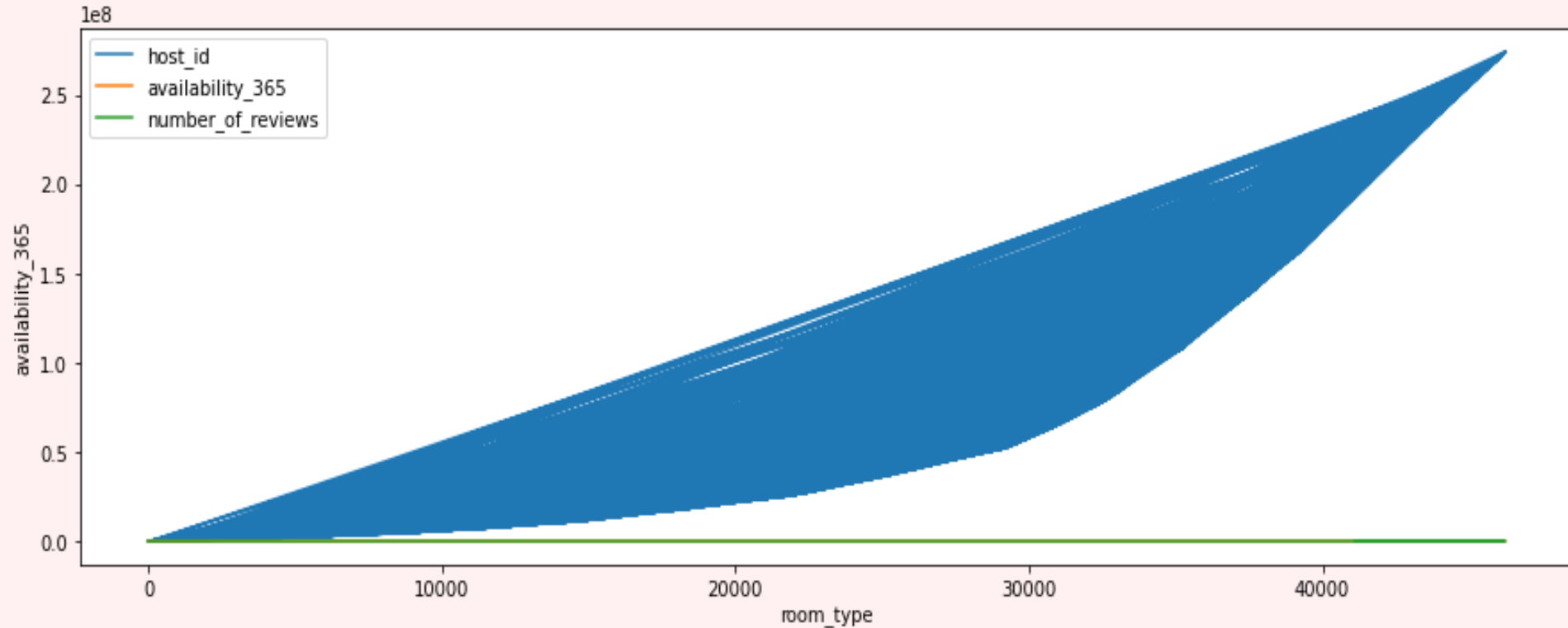
housevsprice



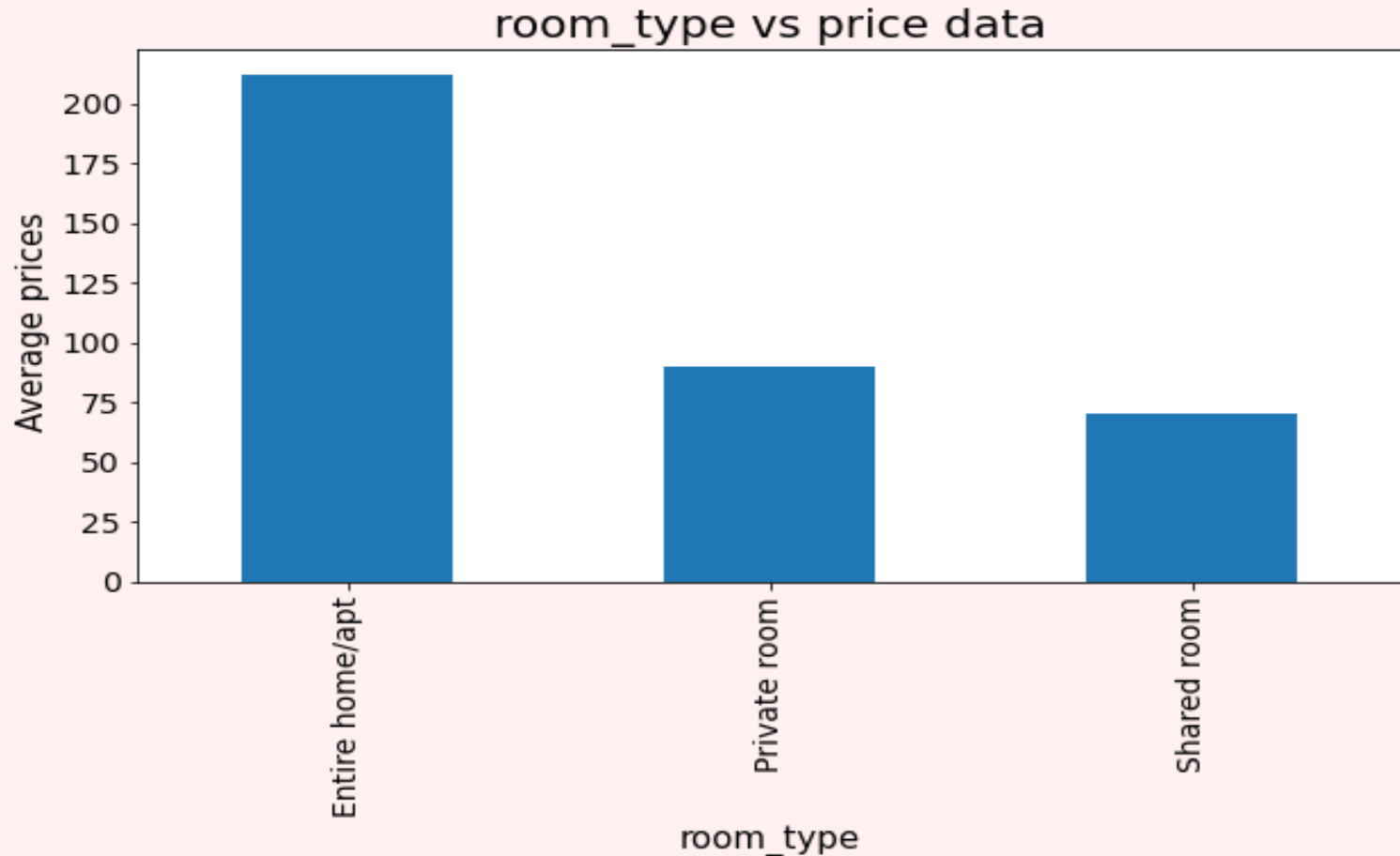
who has hosted different room type with respect of minimum nights



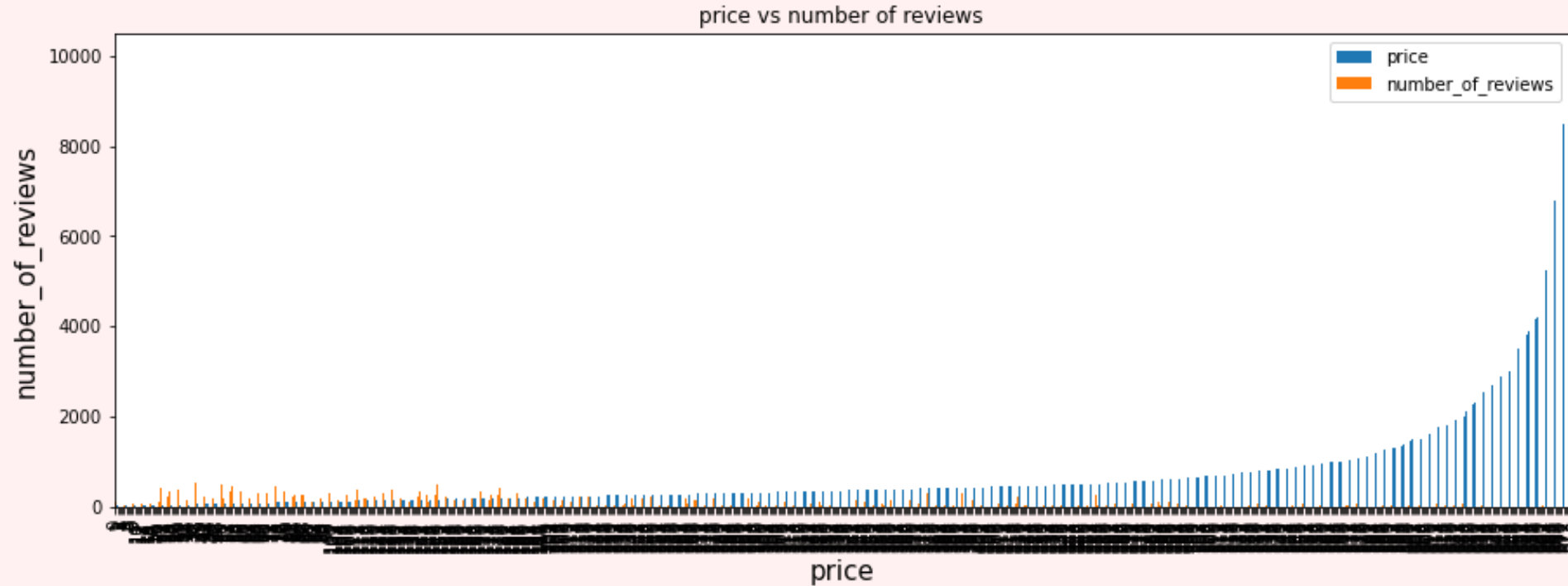
room_type vs availaility based on number of reviews



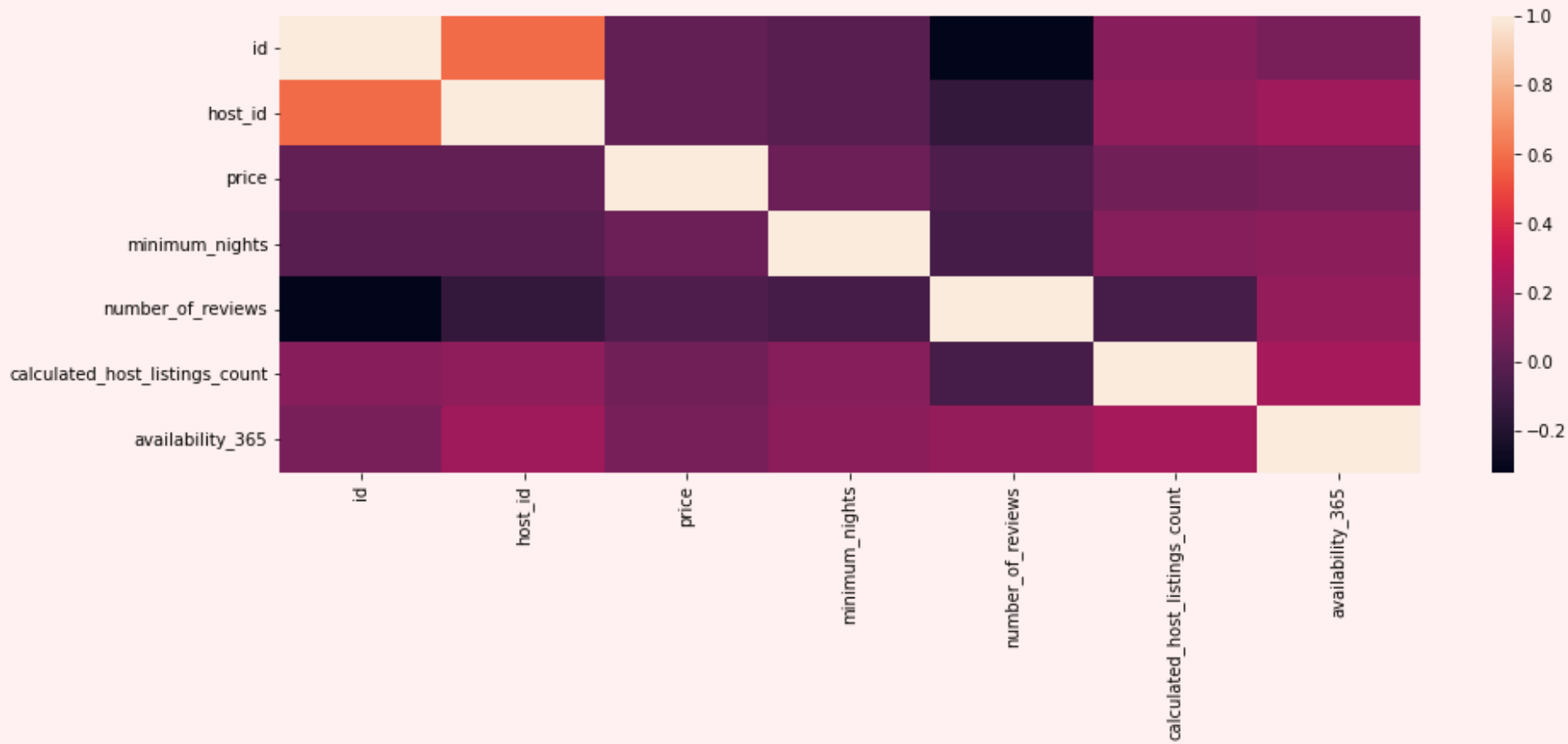
Comparing Average house prices for different room types



price vs number of reviews



Heatmap Discussion



CONCLUSION

THANK YOU!