PILLOW PALOOZA DATA INSIGHTS

A Start-up providing short-term rental options in New york City

MY ROLE

As a Data Analyst for Pillow Palooza, my role involves analysing data related to the short-term rental market in New York City to provide valuable insights that can drive the growth and success of the company.

PROJECT SUMMARY

1. Exporting and Cleaning data in Python (separate PDF containing Python code included with the submission)

2. Exploring and Analysing data with SQL (separate PDF containing SQL code included with the submission)

3. Presenting Findings for Business impact using Tableau.

```python
import numpy as np
import pandas as pd
import datetime as dt

#1 Importing the data
prices = pd.read_csv("data/airbnb_price.csv")
xls=pd.ExcelFile("data/airbnb_room_type.xlsx")
room_types = xls.parse(0)
reviews = pd.read_csv("data/airbnb_last_review.tsv", sep='\t')

#2 Cleaning the price column
prices["price"] = prices["price"].str.replace(" dollars", "")
prices["price"] = pd.to_numeric(prices["price"])

#3 Calculating average price
zero_listings = prices["price"] == 0
prices = prices.drop(prices[zero_listings].index)
avg_price = round(prices["price"].mean(), 2)

#4 Comparing costs to the private rental market
prices["prices_per_month"]=prices["price"]*365 / 12
average_price_per_month=round(prices["prices_per_month"].mean(), 2)
difference= round(average_price_per_month-3100,2)

#5 Cleaning the room_type column
room_types["room_type"]=room_types["room_type"].str.lower()
room_types["room_type"]=room_types["room_type"].astype("category")
room_frequencies = room_t

#6 What timeframe are we
reviews["last_review"] = pd.to_datetime(reviews["last_review"])
first_reviewed=reviews["last_review"].dt.date.min()
last_reviewed=reviews["last_review"].dt.date.max()

#7 Joining the DataFrames
rooms_and_prices = prices.merge(room_types, how="outer", on="listing_id")
airbnb_merged=rooms_and_prices.merge(reviews, how="outer", on="listing_id")
airbnb_merged = airbnb_merged.dropna()

#8 Analyzing listing prices by NYC borough
airbnb_merged["borough"] = airbnb_merged["nbhood_full"].str.split(",").str[0]
boroughs =
airbnb_merged.groupby("borough")["price"].agg(["sum","mean","median","count"])
airbnb_merged["price_range"] = pd.cut(airbnb_merged["price"], bins=ranges,
labels=label_names)
prices_by_borough = airbnb_merged.groupby(["borough",
"price_range"])["price_range"].count()

#10 Storing the final result
airbnb_analysis={'avg_price':avg_price,
                 'average_price_per_month':average_price_per_month,
                 'difference':difference,
                 'room_frequencies':room_frequencies,
                'first_reviewed':first_reviewed,
                 'last_reviewed':last_reviewed,
                 'prices_by_borough':prices_by_borough}
print(airbnb_analysis)
```

EXPORTING AND CLEANING DATA IN PYTHON

OUTPUT:

```
shared room          587
Name: room_type, dtype: int64, 'first_reviewed': datetime.date(2019, 1, 1),
'last_reviewed': datetime.date(2019, 7, 9), 'prices_by_borough':
borough        price_range
Bronx          Budget          381
               Average         285
               Expensive        25
               Extravagant       5
               Extravagant     259
attan          Budget         1148
               Average        5285
               Expensive      3072
               Extravagant     810
Queens         Budget         1631
               Average        1505
               Expensive       291
               Extravagant      28
Staten Island  Budget          124
               Average         123
               Expensive        20
               Extravagant       0
```

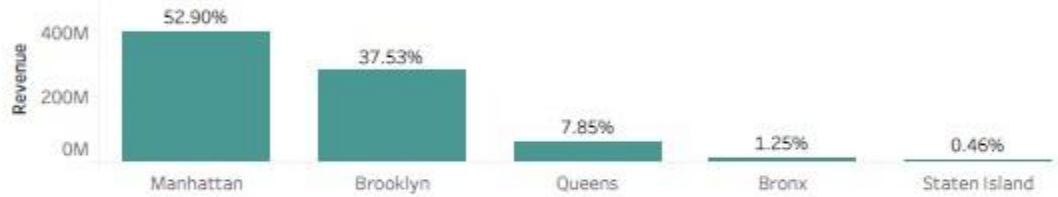
EXPLORING AND ANALYSING DATA WITH SQL

```
        ON reviews.listing_id = room_types.listing_id
GROUP  BY room_type
ORDER  BY total_booked_days DESC;

OUTPUT:
room_type              total_booked_days
entire home/apt            13266
private room               11356
shared room                587

Ans: entire home/apt


**************************************************************************
******

/*2. What is the average price of a listing by room type?*/
SELECT room_type,
       ROUND(AVG(price),2) AS avg_price
FROM   prices
       JOIN room_types
         ON prices.listing_id = room_types.listing_id
GROUP  BY room_type;

OUTPUT:
room_type            avg_price
shared room           53.65
entire home/apt       197.17
private room          81.67

**************************************************************************
****

/*3. Which borough has the highest average price per month?*/

SELECT borough,
       ROUND(AVG(price_per_month)) AS avg_price_month
FROM   prices
GROUP  BY borough
```

```
**************************************************************************
**********

/*4. How many listings of each room type are in each borough?*/

SELECT room_type,
       borough,
       Count(prices.listing_id) AS No_of_listings
FROM   prices
       JOIN room_types
         ON prices.listing_id = room_types.listing_id
GROUP  BY borough,
          room_type;

room_type              borough              no_of_listings
entire home/apt         Staten Island            133
private room            Manhattan                3901
shared room             Brooklyn                 187
shared room             Staten Island            2
shared room             Manhattan                251
entire home/apt         Brooklyn                 5367
private room            Bronx                    403
private room            Brooklyn                 4906
shared room             Queens                   112
entire home/apt         Queens                   1335
private room            Staten Island            132
shared room             Bronx                    33
private room            Queens                   2009
entire home/apt         Bronx                    261
entire home/apt         Manhattan                6170

**************************************************************************
******************

/* 5. How many listings in each room type category have a price of over $500
per night?*/

SELECT room_type,
```

Pillow Palooza NYC Airbnb

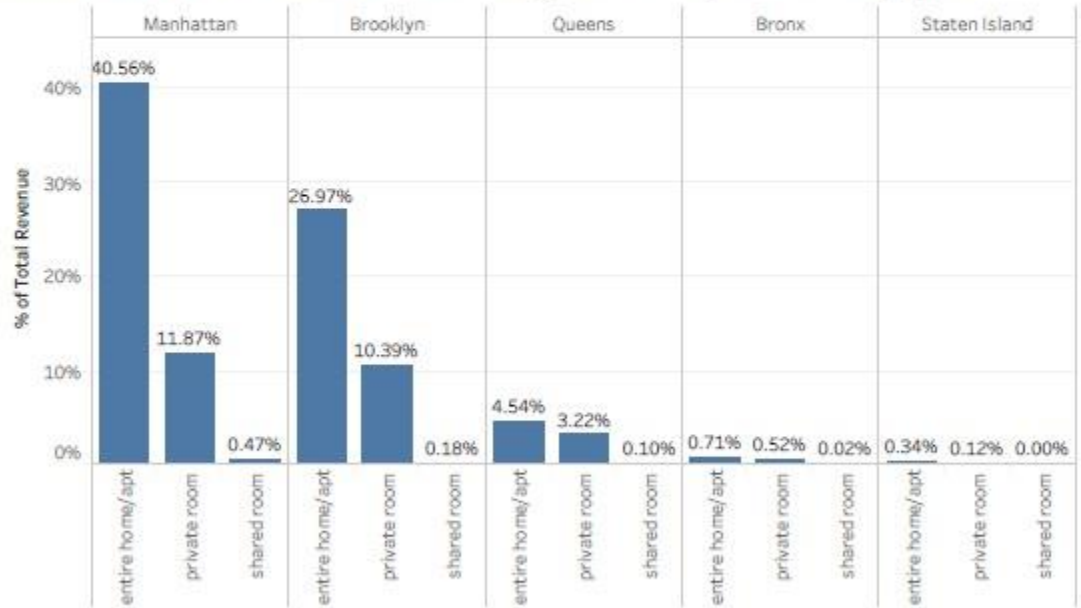
More than 50% of the revenue are generated from Manhattan



Revenue Distribution:
The majority of revenue (52.90%) is generated in Manhattan, followed by Brooklyn (37.53%). With 52.90% of the total revenue being generated in Manhattan indicates that this borough is the most significant contributor to the overall revenue. This suggests that Manhattan has the potential for continued growth.

67% of the revenues come from entire homes/apartments in Manhattan and Brooklyn. Manhattan and Brooklyn are popular tourist destinations, attracting visitors from around the world. The demand for entire homes/apartments indicates that tourists prefer the comfort and privacy offered by such accommodations over Private rooms or shared spaces.

More than 70% of the total revenue is generated by entire home/apt



Pillow Palooza NYC Airbnb

Price vs Borough Box plot

Manhattan is the most expensive of all the boroughs.

Looking at the Manhattan box plot, we conclude the plot is positively skewed. Most of the listings have Avg property prices more than the Median($ 187.50) in that borough. This suggests that there are relatively more listings with higher property prices in Manhattan compared to the lower-priced listings.

Tribeca, Flatiron District, NoHo, SoHo, Midtown, West village, and Murray Hill are some of the outliers for Manhattan.

Pillow Palooza NYC Airbnb

Top 10 Expensive Neighbourhoods

Borough	Neighbourhood	Avg. Price	Count of Listing Id (Room Types.Cs..	Number Of Reviews
Brooklyn	Sea Gate	805	2	7
Manhattan	Tribeca	397	61	1,667
	Flatiron District	342	36	1,245
	NoHo	335	41	749
	SoHo	299	191	6,460
	Midtown	272	608	16,655
	West Village	259	358	12,185
	Murray Hill	240	189	3,778

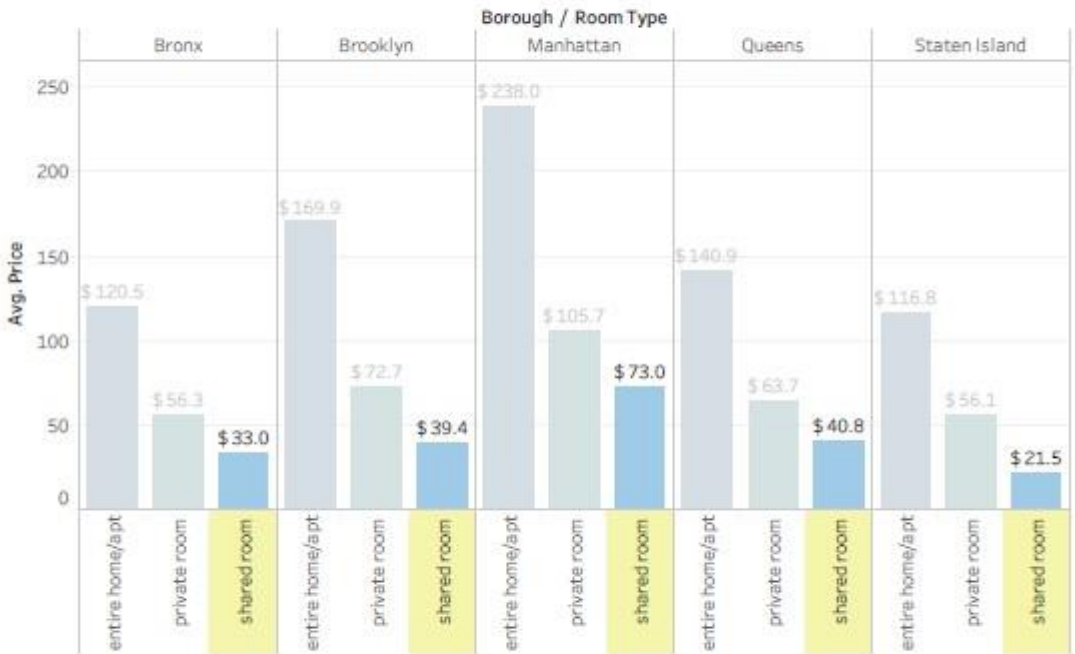
Neighbourhood
- ✔ Allerton
- ✔ Arden Heights
- ✔ Arrochar
- ✔ Arverne
- ✔ Astoria
- ✔ Bath Beach
- ✔ Battery Park City
- ✔ Bay Ridge
- ✔ Bay Terrace
- ✔ Baychester
- ✔ Bayside
- ✔ Bayswater

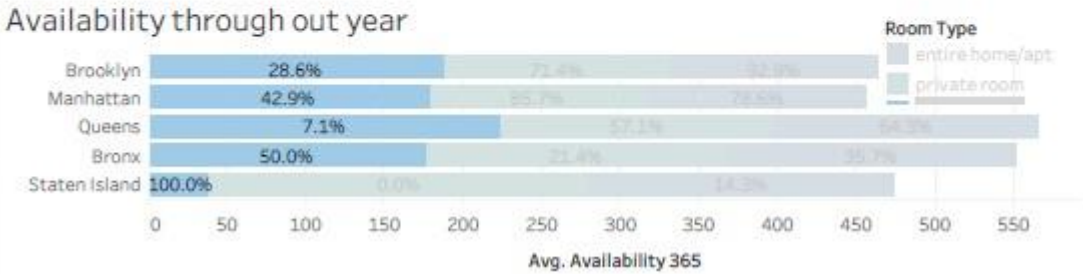
Limit
Top 10 by AVG([Price])

Borough
- ✔ Bronx
- ✔ Brooklyn
- ✔ Manhattan
- ✔ Queens
- ✔ Staten Island

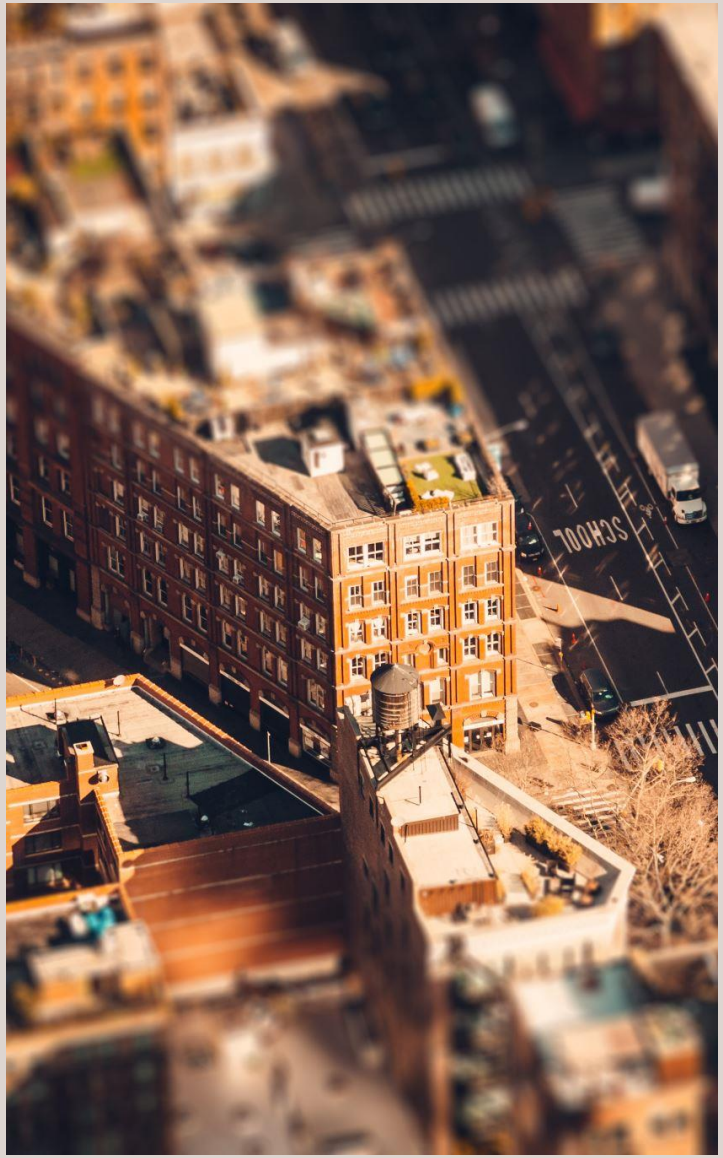
Top 10 Neighbourhoods. Here we can see that the most expensive neighborhoods are all located in Manhattan. Please note Sea Gate(Brooklyn) is an outlier. Despite the higher rental prices, Manhattan and Brooklyn remain popular due to their proximity to NYC tourist attractions.

Pillow Palooza NYC Airbnb





Though Manhattan has higher short-term rental prices, still people prefer to live in this borough. One of the reasons might be its proximity to NYC tourist attractions.
Looking at the availability, Staten Island and Bronx have listings that are most available throughout the year compared to other boroughs.
Also note, Manhattan and Brooklyn are the most desirable boroughs. Manhattan and Brooklyn exhibit less availability i.e. strong occupancy rates, indicating high demand for accommodations. Investing in properties within these boroughs can lead to positive returns and minimize risk.



CONCLUSIONS & RECOMMENDATIONS:

- Focusing on Manhattan and Brooklyn presents a greater opportunity for Revenue generation.
- 70% of the revenues come from entire homes/apartments. These segments outperform shared rooms and private rooms in terms of occupancy rates. The demand for entire homes/apartments indicates that tourists prefer the comfort and privacy offered by such accommodations over private homes or shared spaces. Emphasizing entire homes/apartments and private rooms can enhance revenue performance.
- Manhattan emerges as the most expensive borough in New York City, with a concentration of higher-priced listings and a positively skewed distribution. The top expensive neighbourhoods, such as Tribeca, Flatiron, and Noho, are all located in Manhattan. Despite the higher rental prices, Manhattan and Brooklyn remain popular due to their proximity to NYC tourist attractions.
- Manhattan and Brooklyn are considered the most desirable boroughs overall. Manhattan and Brooklyn consistently exhibit strong occupancy rates, indicating high demand for accommodations. Investing in properties within these boroughs can lead to positive returns and minimize risk.
- However, it is important to acknowledge that Airbnb dynamics are influenced by multiple factors and market conditions that can vary over time. Conducting further analysis with more comprehensive and up-to-date data is necessary for a more precise understanding of the Airbnb market in New York City.