The DataViz Challenge - Transforming EDA Projects to Dashboards

Create Stunning Tableau Desktop/Tableau Public/PowerBI Dashboard from the same dataset on which EDA was performed for the capstone project.

Suppose, **if someone performed EDA on Airbnb operations** then the following steps need to be followed for building an adequate dashboard through which one can analyse the performance of the company to take necessary actions.

Problem Statement:

In the context of Airbnb operations, how can the utilisation of Tableau Desktop/Tableau Public/PowerBI facilitate a comprehensive comparative examination between Chicago and New Orleans, two diverse urban environments?

This inquiry seeks to leverage Tableau Desktop/Tableau Public/PowerBI's visual analytics capabilities to uncover and illustrate the shared attributes, disparities, and distinctive patterns inherent to Airbnb's presence in these cities, thus elevating the depth and insightfulness of the study.

Dataset Selection:

For this EDA project, we have chosen the "Airbnb Listings Data" dataset from 2 major cities: Chicago and New Orleans. This dataset provides a comprehensive snapshot of various attributes related to Airbnb listings, such as property type, neighbourhood, pricing, availability, and more. The dataset is ideal for conducting an in-depth exploration of the local Airbnb market and deriving actionable insights.

Why Airbnb:

Airbnb, a prominent online platform, enables individuals to reserve accommodations spanning a spectrum from beds and rooms to apartments and entire homes across global locales. This user-centric platform serves as a conduit for seamless property rentals, negating the need for intricate intermediaries or substantial capital outlays. Notably, users can secure lodgings at significantly competitive rates relative to traditional hotels. Distinctively, Airbnb extends its reach to regions where conventional hotel presence might be limited, offering an avenue for lodging acquisition in underserved locales. Moreover, the inclination towards immersive local experiences often steers individuals towards selecting accommodations embedded within native communities, fostering a distinctive preference for authenticity and cultural engagement.

Airbnb Statistics • Over 4 million listings worldwide • 150 million users in 191 countries • Worldwide value is $32 billion • Global growth rate since 2009 - 153%

Dataset Details:

- **Dataset Name:** Airbnb Listings Data
- **Source:** Link to dataset
- **Cities:** Chicago & New Orleans
- **Description:** The Airbnb Listings Data contains information about different properties available for rent on Airbnb in a specific city. Each record represents a unique listing and includes attributes such as property type, neighbourhood, number of bedrooms, pricing, availability, host information, and more.

Key Attributes:

1. **Id:** Unique identifier for each listing.
2. **name:** The title or name of the listing.
3. **host_id:** Unique identifier for the host of the property.
4. **host_name:** Name of the host.
5. **neighbourhood_group:** The broader area or group that the neighbourhood belongs to.
6. **neighbourhood:** Specific neighbourhood where the property is located.
7. **latitude:** Latitude coordinate of the property.
8. **longitude:** Longitude coordinate of the property.
9. **room_type:** Type of room (e.g., Private room, Entire home/apt, Shared room).
10. **price:** Price of the listing per night.
11. **minimum_nights:** Minimum number of nights required for booking.
12. **number_of_reviews:** Total number of reviews received for the listing.
13. **last_review:** Date of the last review.
14. **reviews_per_month:** Average number of reviews per month.
15. **availability_365:** Number of days the listing is available for booking in a year.

Problem Areas to Explore:

1. Which are the popular **neighbourhoods**, their average prices and no. of listings?
2. What is the percent share of different **property types** and **room types**?

3. How the **pricing** is varying with location, property type, and reviews?

4. What are the different correlations between type of **hosts** and factors like- reviews & price?

Divide the visualisation findings into 4 categories:

- Overview of Airbnb
- Property analysis
- Pricing analysis
- Host analysis

How to proceed with the dashboard:

1. Data Cleaning

Begin by addressing the disorder and inconsistency within the dataset. Utilise Jupyter Notebook and Tableau Desktop/Tableau Public/PowerBI Prep to systematically cleanse the data, rectifying discrepancies, eliminating duplicates, and standardising formats.

2. Data Transformation

Generate supplementary columns by utilising pre-existing categorical data. These columns will be derived from extensive descriptive text, which, in its original form, proved arduous to comprehend and unsuitable for visualisation purposes. The extra columns that we created gave a much clear sense of how to approach and make an effective visualisation.

3. Tableau Desktop/Tableau Public/PowerBI

Employ Tableau Desktop/Tableau Public/PowerBI Prep to leverage its distinctive "Group and Replace" feature. Under the column denoted as Neighbourhood there are instances where identical entities are variably represented due to disparities in letter casing, spelling variations, or phonetic similarity. The "Group and Replace" algorithm inherent to Tableau Desktop/Tableau Public/PowerBI Prep proved instrumental in mitigating this issue.

1. **Airbnb Analysis - Chicago and New Orleans - Dashboard Link :-**

https://drive.google.com/file/d/1e1HksRqbry2SeDxLDc2cplnO6lJvMmC7/view?usp=drive_link

2. **Airbnb Dataset (Copy of new_orleans_listings.csv) Link :-**

https://drive.google.com/file/d/1OTR5fZ_-Dhj9pMiCPmCkKTlatgKiYq3s/view?usp=sharing

3. **Airbnb Dataset (Copy of chicago_listings.csv) Link :-**

https://drive.google.com/file/d/1OTR5fZ_-Dhj9pMiCPmCkKTlatgKiYq3s/view?usp=sharing

4. **GitHub Link -**

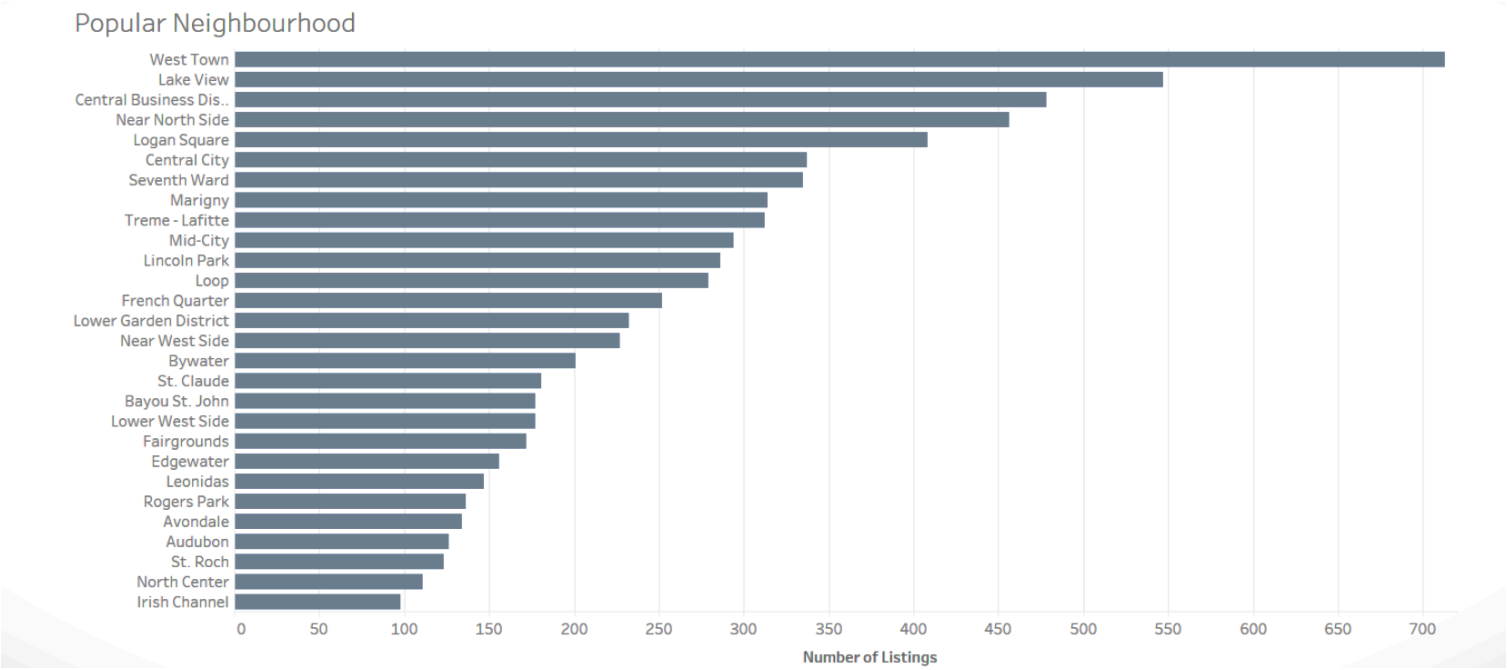
https://github.com/Manish-Kumar950407/Chicago-and-New-Orleans-Listing---Data-Visualization-

A) Overview of Airbnb

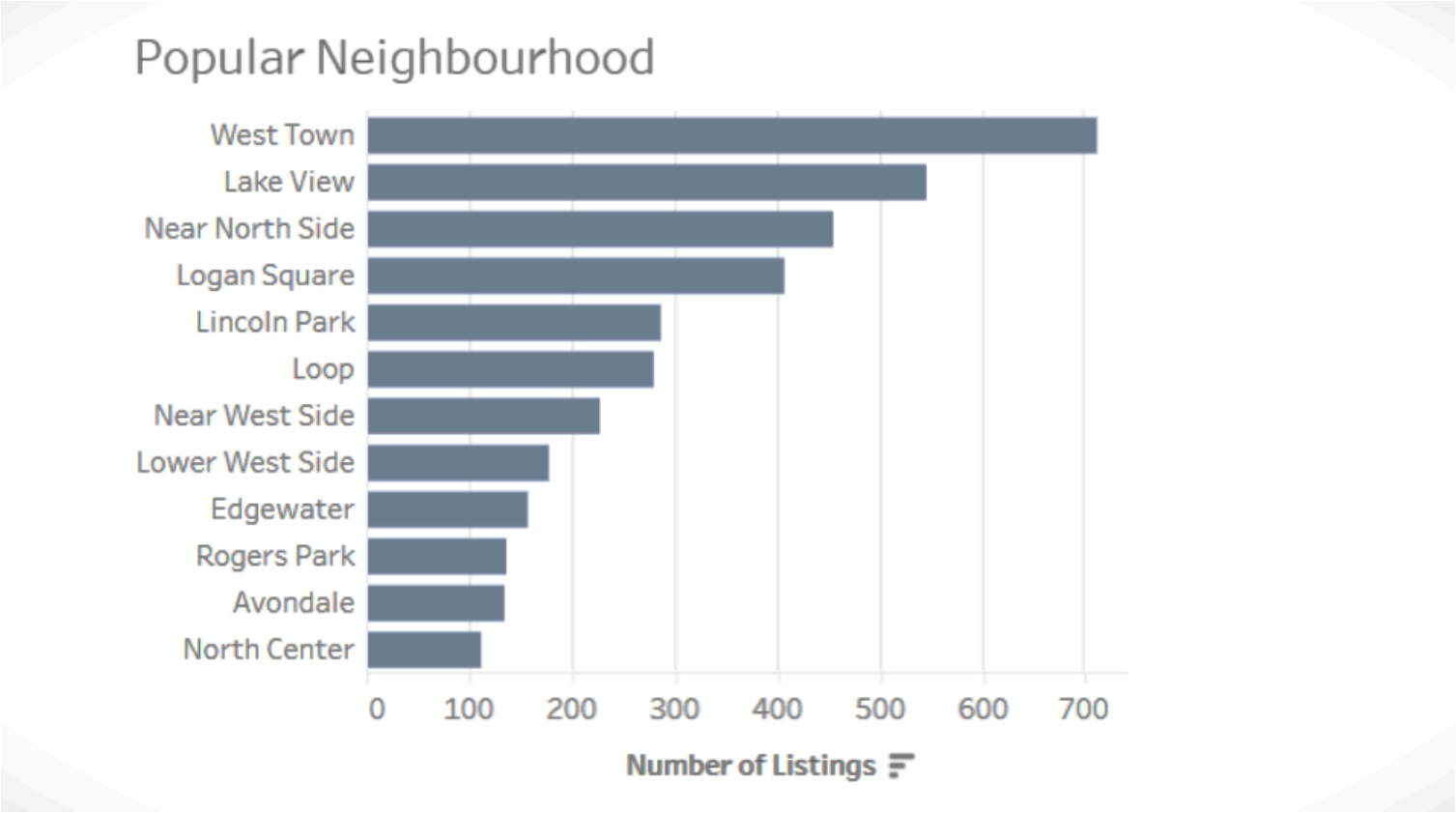
1. Which are the popular **neighbourhoods**, their average prices and no. of listings.

a. Popular neighbourhood areas in the ranking from highest number of listings to the least :-

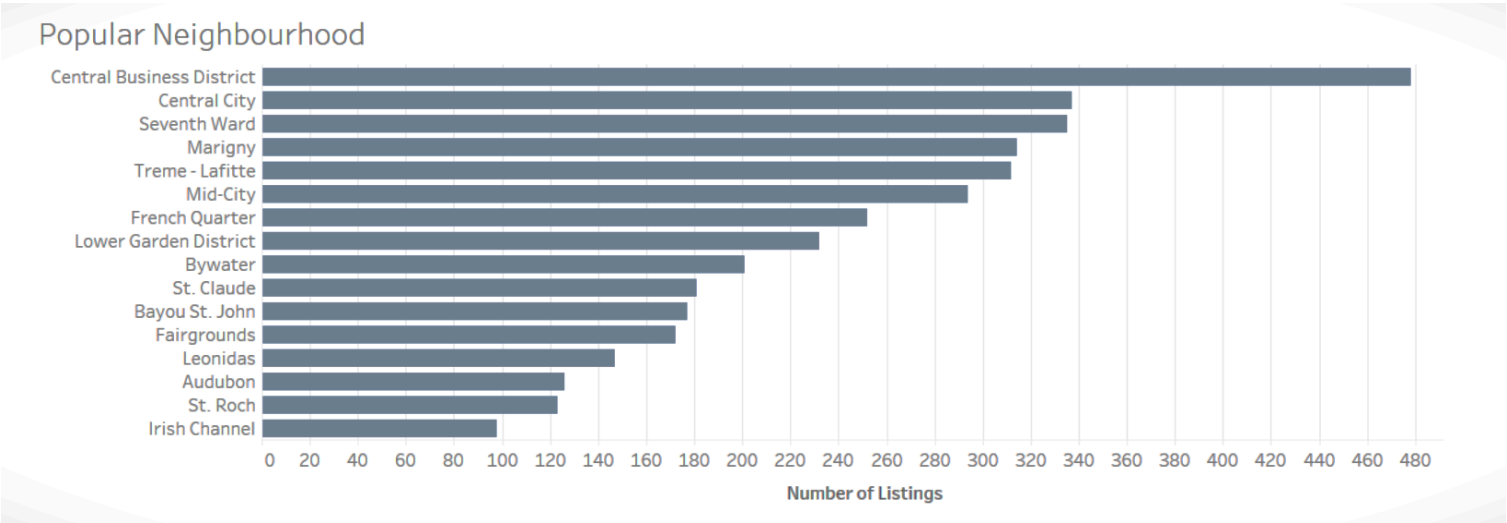
1. **Overall**



2. **Chicago**



3. New Orleans



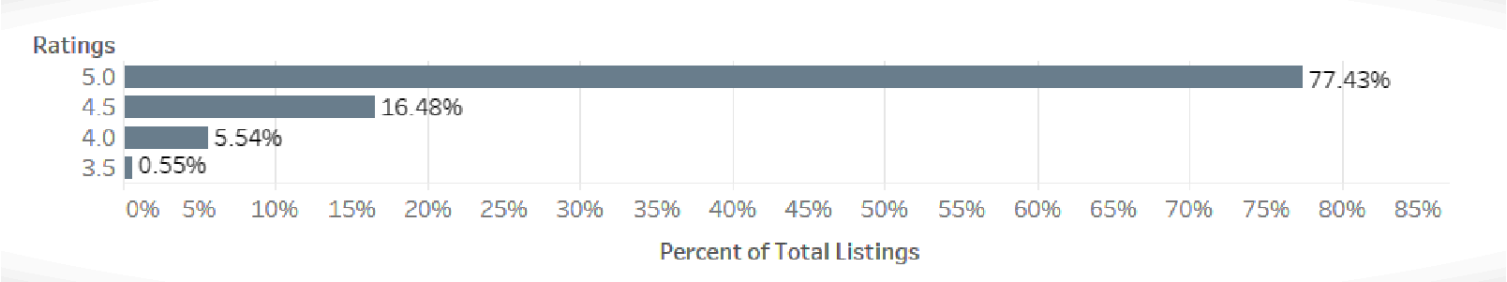
Overall, the West Town area, in Chicago, has 713 number of listings (highest). West town is one of the best places to live in Illinois as it is very posh, located west of downtown Chicago and is a centre of attraction for people due to its liveliness.

On the other hand, the Irish Channel (in New Orleans) has the least number of listings i.e. 98. One of the possible reasons could be that the Irish Channel is predominantly known for its crime activities. It is not considered safe to live in the area, especially during night time. So, there are high chances that people don't prefer to take accommodation in this area which has affected the number of listings as well.

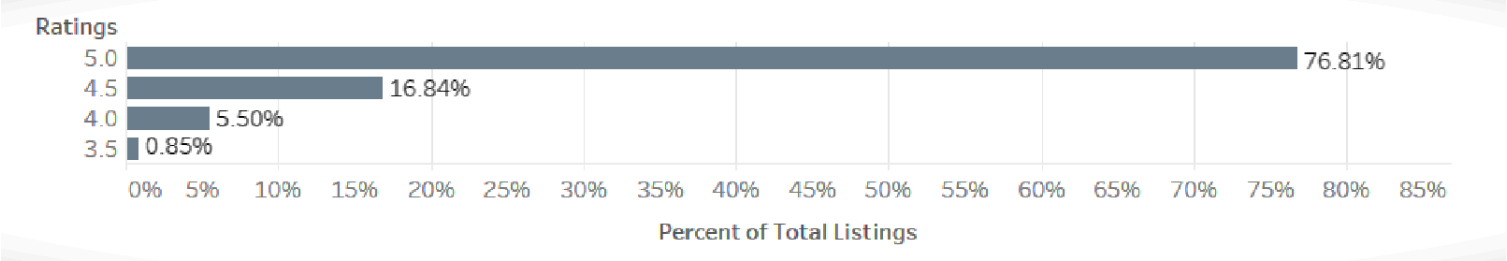
---

b. Ratings :-

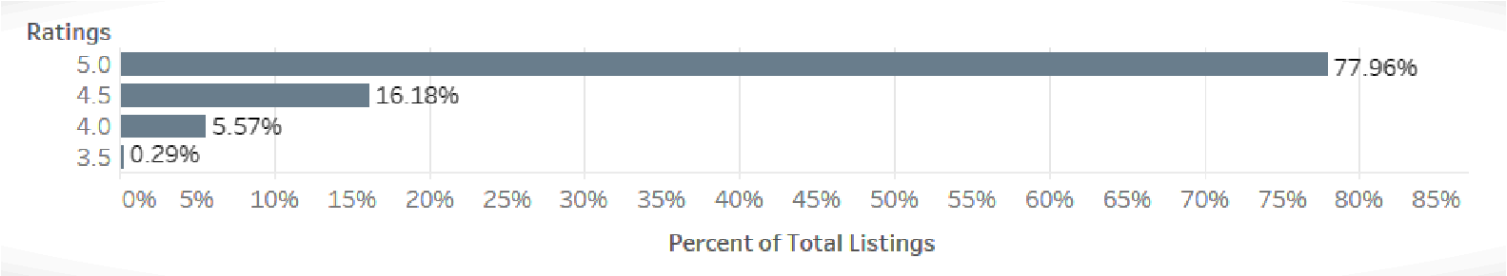
1. Overall



2. Chicago



3. New Orleans



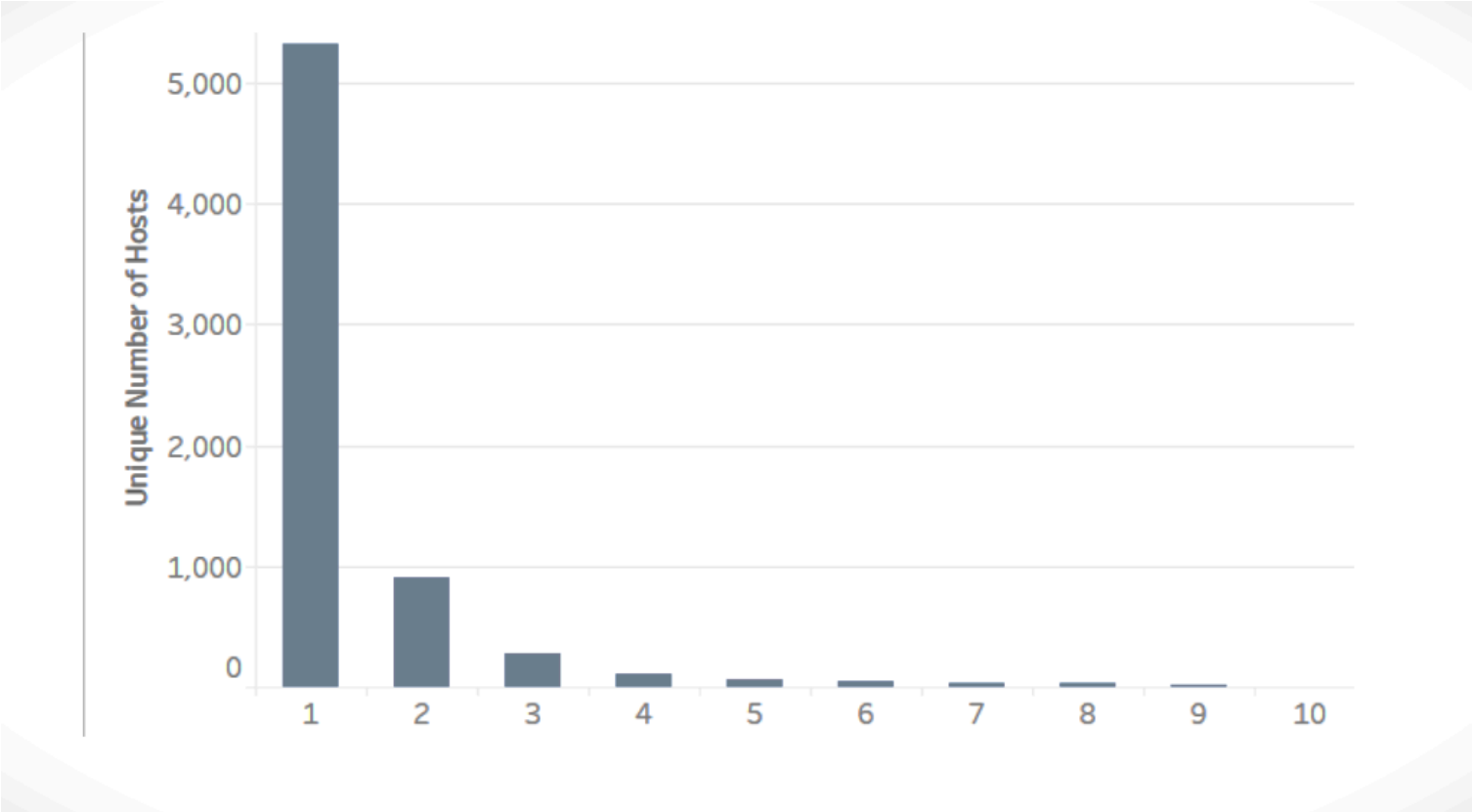
The rating pattern is quite similar in both Chicago and New Orleans. We can see that more than 75% of the listings have attained a rating of 5.0; and less than 1% have a rating of 3.5 or lower, which suggests that people are highly satisfied with the kind of accommodation, price and service they are getting through Airbnb.

B) Property Analysis

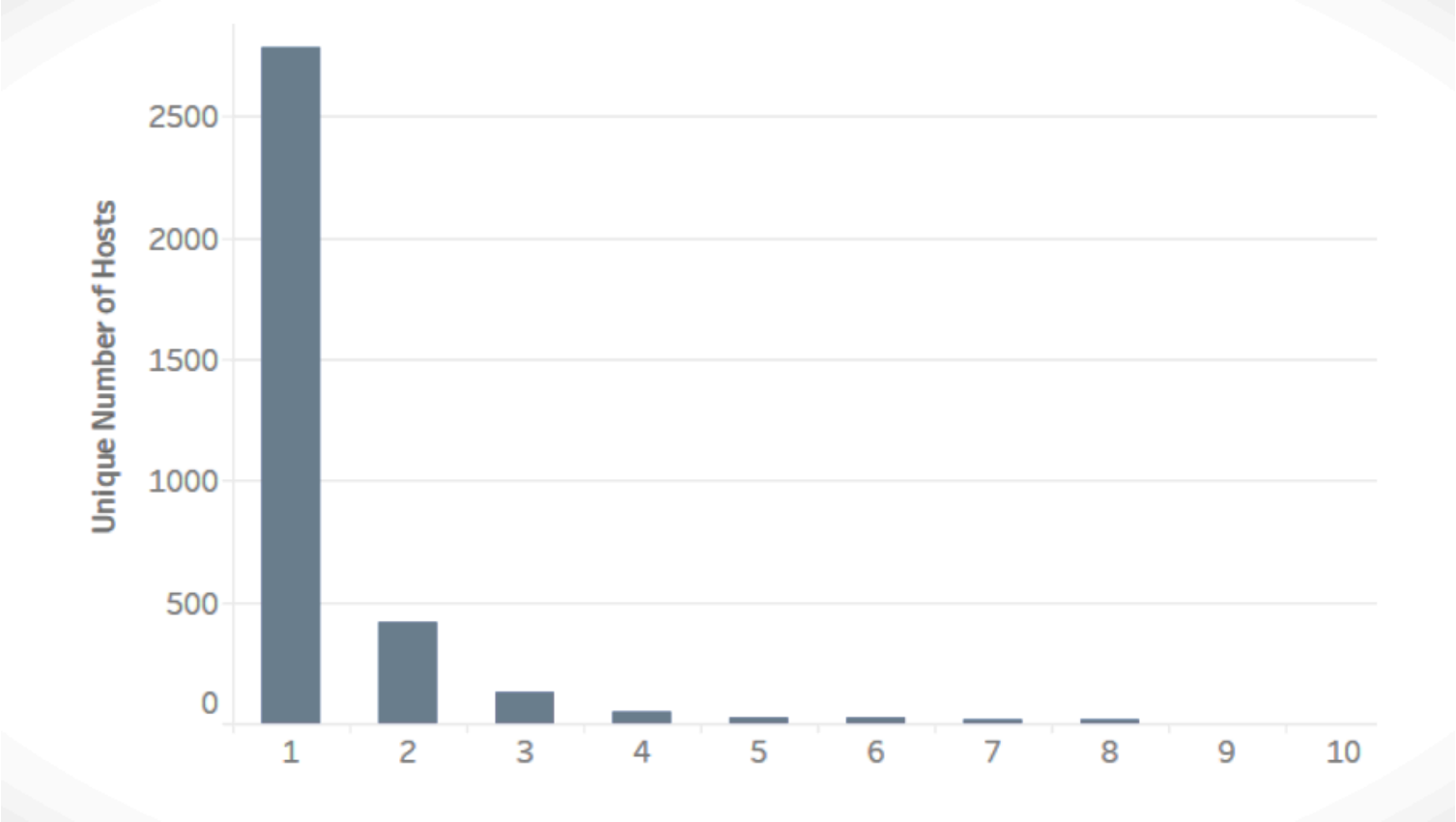
2. What is the percent share of different **property types** and **room types?**

a. Number of listings per host :-

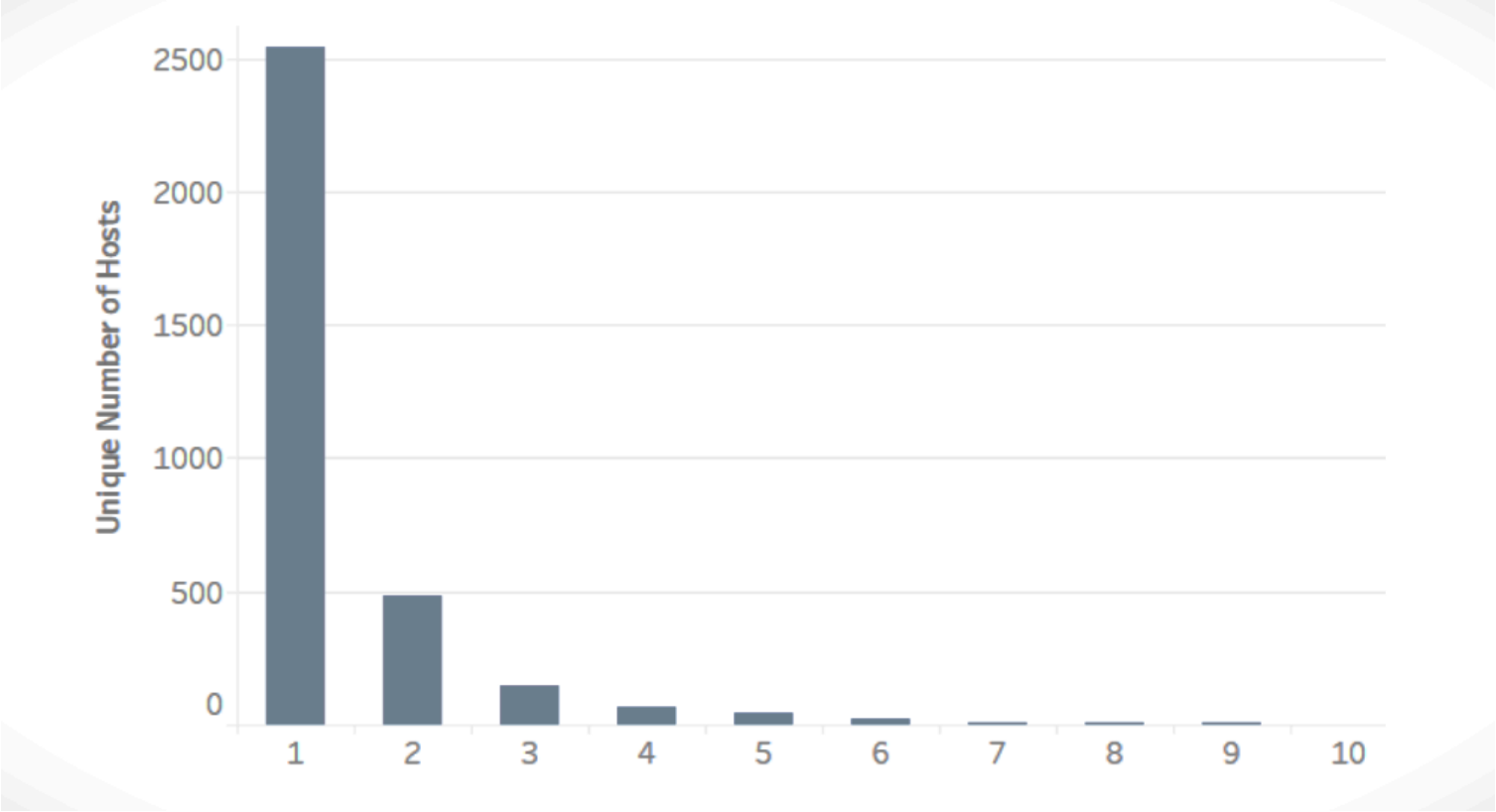
1. **Overall**


Unique Number of Hosts
5,000
4,000
3,000
2,000
1,000
0
1
2
3
4
5
6
7
8
9
10

2. **Chicago**



3. **New Orleans**

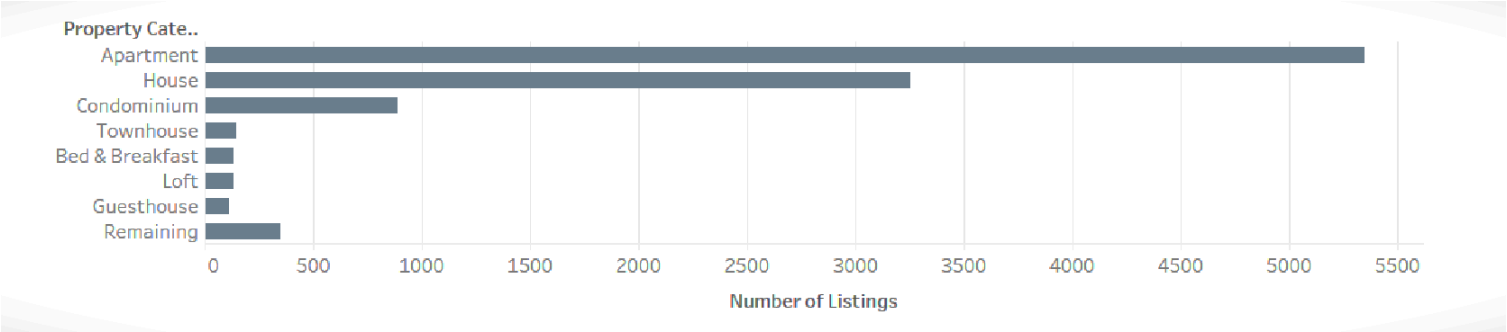


Both cities have the same pattern. Maximum number of hosts are having only one property. Also, there are only 4 hosts having 10 listings. Overall, more than 78% hosts have only 1 listing. This can be attributed to the fact that most of the hosts lend their place as a side business without getting into the mainstream business, by lending their place or a room of their apartment at their convenience.

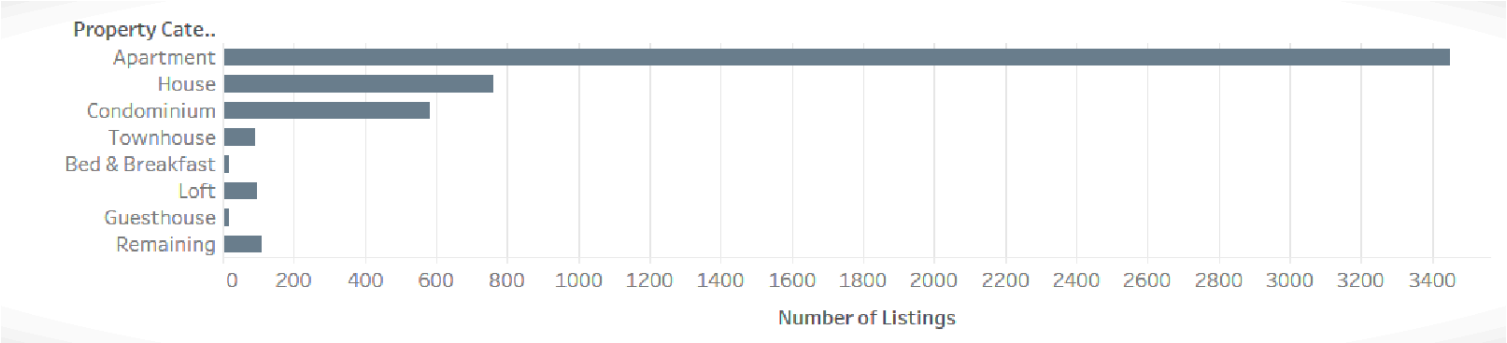
---

b. Property types :-

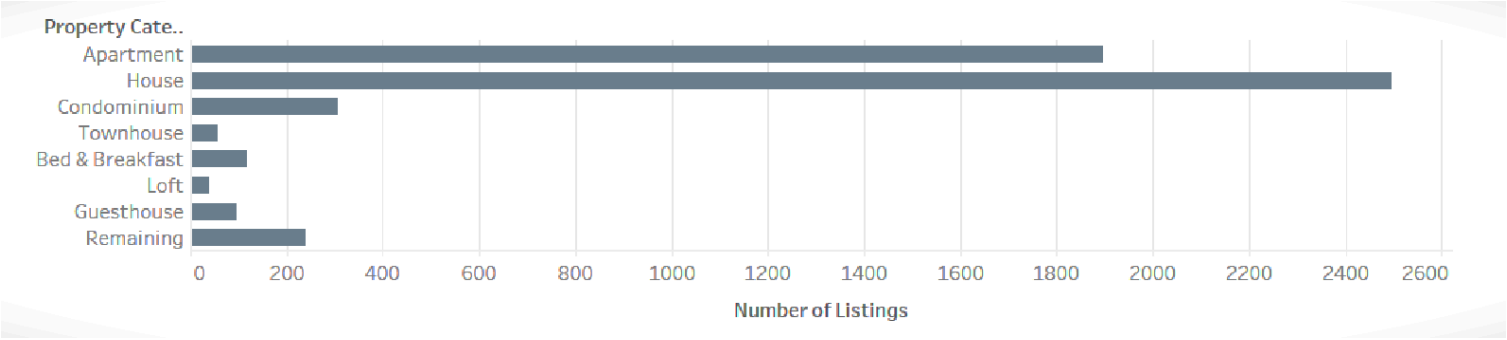
1. **Overall**



2. Chicago



3. New Orleans

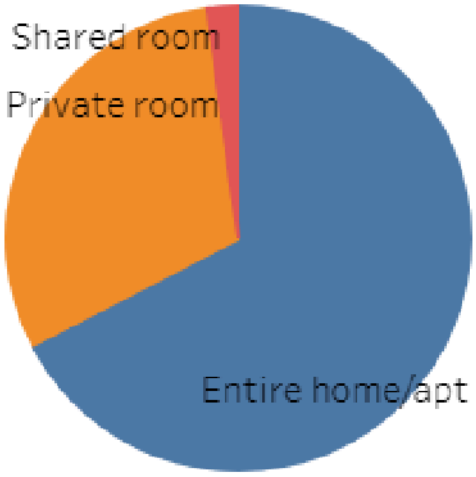


Predominantly, 'Apartments' and 'houses' have the highest number of listings among all the property types. Overall, there are around 50% apartments and 30% Houses. In Chicago, 'Apartments' outnumber all the other property types whereas in New Orleans, majority of the property types are 'House'.

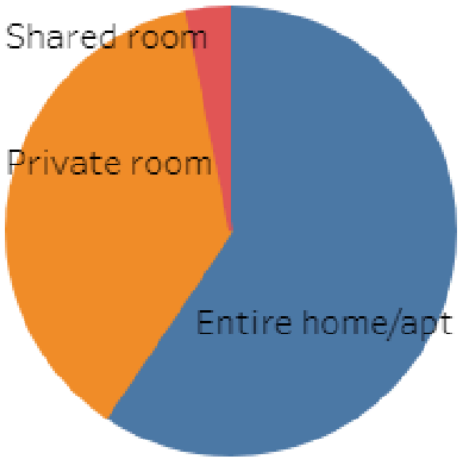
_____

c. Room Type :-

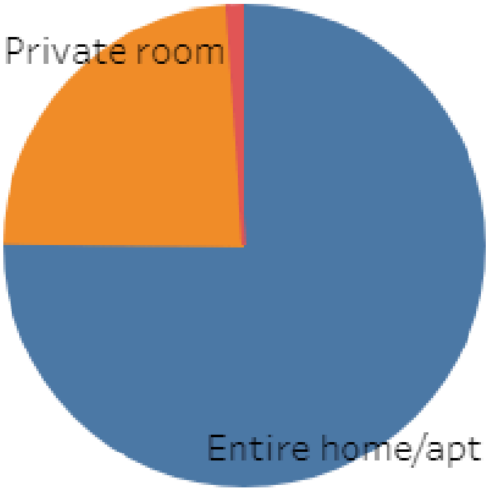
1. **Overall**



2. **Chicago**



3. **New orleans**

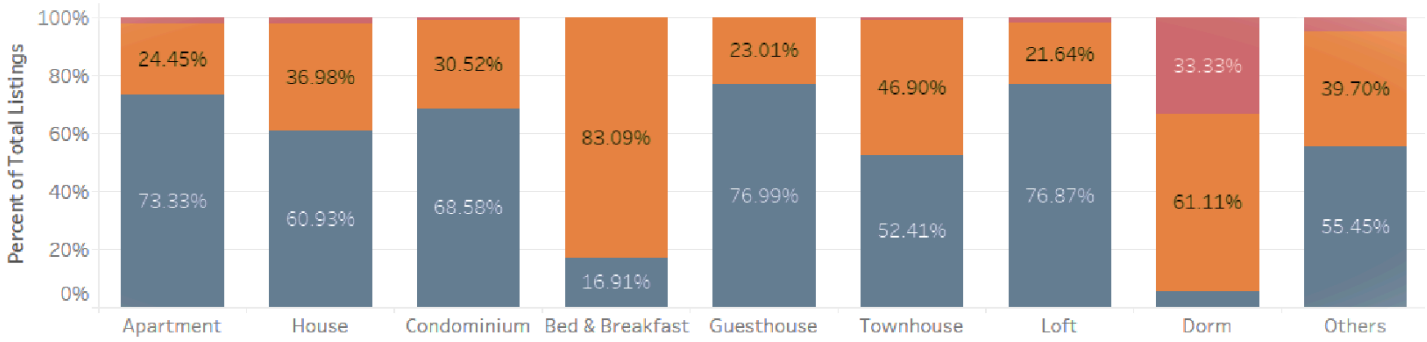


Entire home/apt have the highest share among 'Room Types' followed by Private Room and Shared Room. Both the cities depict the same pattern.

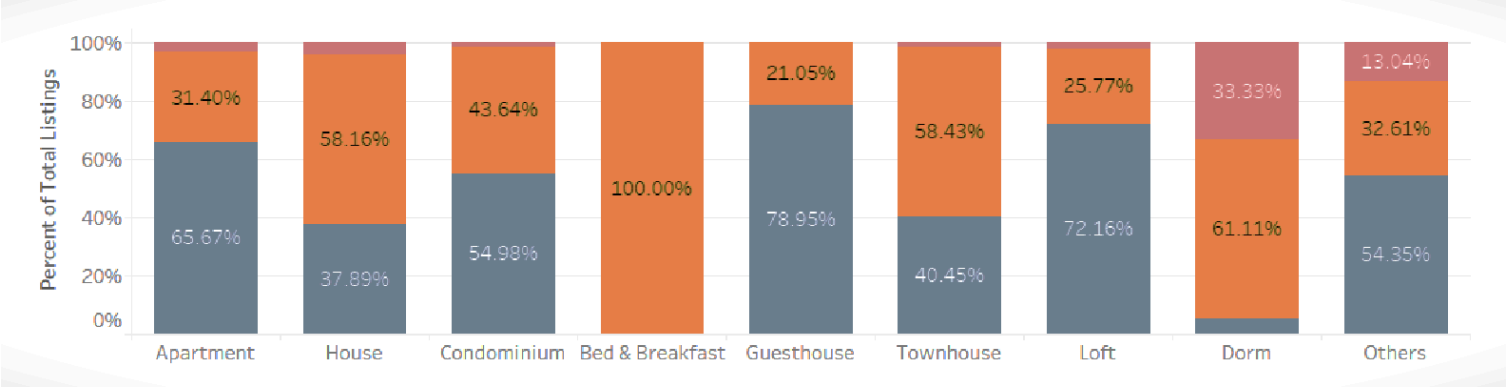
---

d. Room Type by Property Type :-

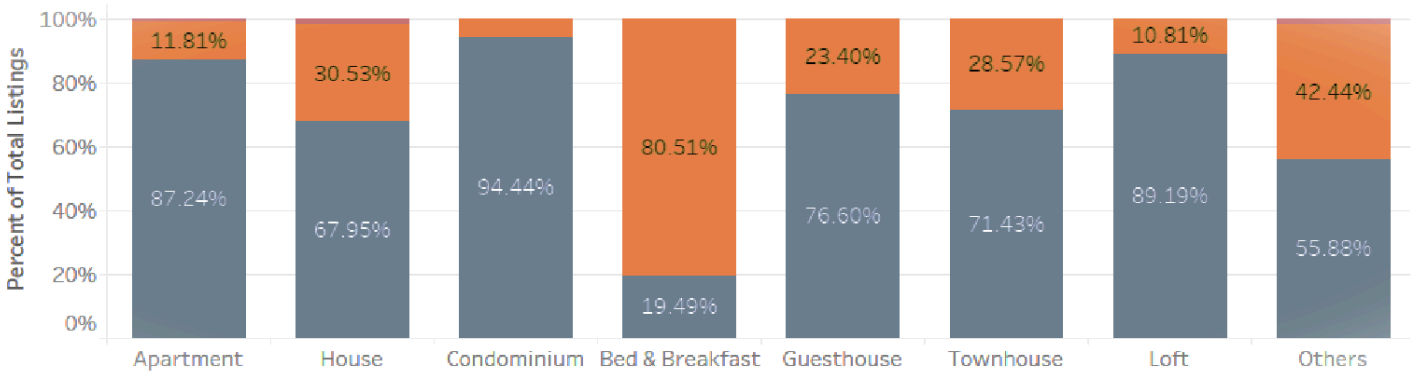
1. **Overall**



2. Chicago



3. New Orleans



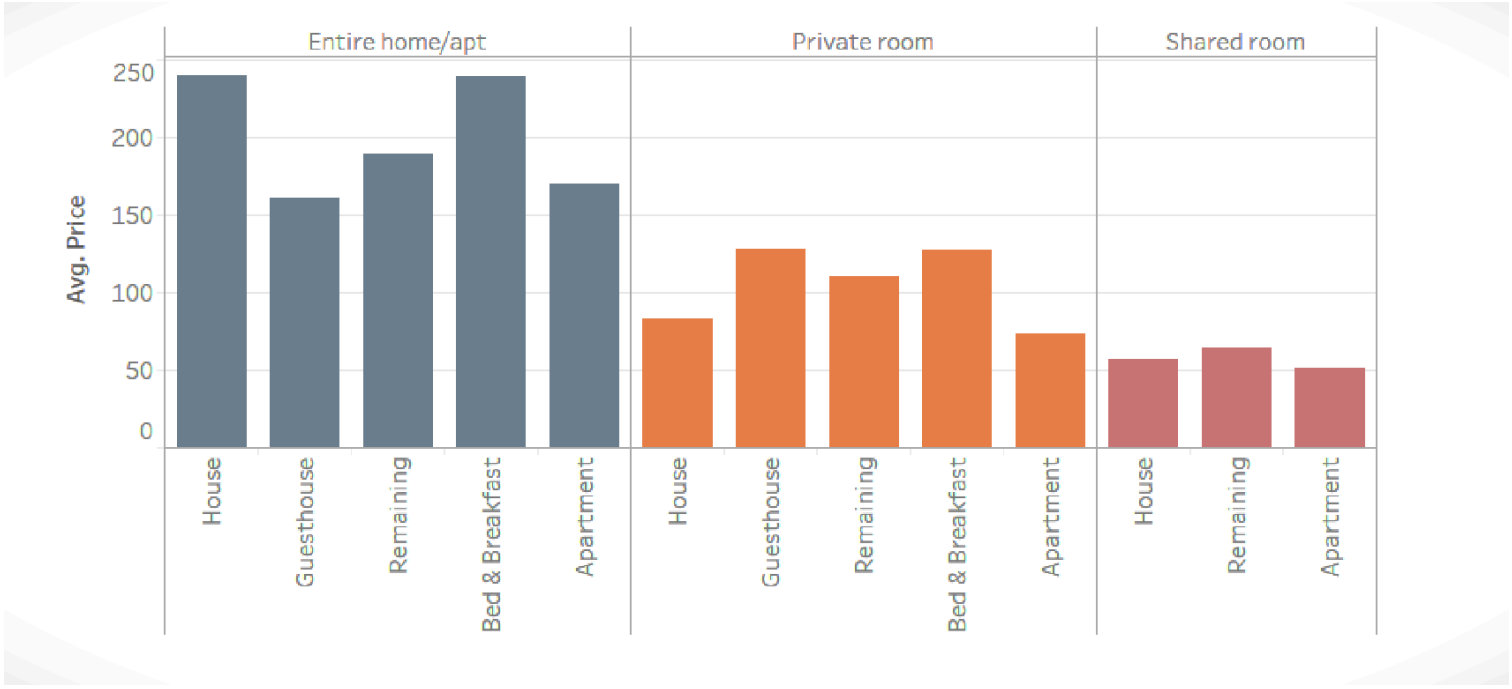
All the property types, either in Chicago or in New Orleans, have the similar pattern that they all have more than 50% of the Room type as Entire Home/Apartment except Bed & Breakfast and Dorm type. These exceptional property type have more than 50% room type as Private Room, and this is very intuitive as people who are or may be solo travellers, would not like to book an the entire home instead they want to live in a private room plus it would be great for them if they are being provided by Breakfast, the next day.

C) Pricing analysis

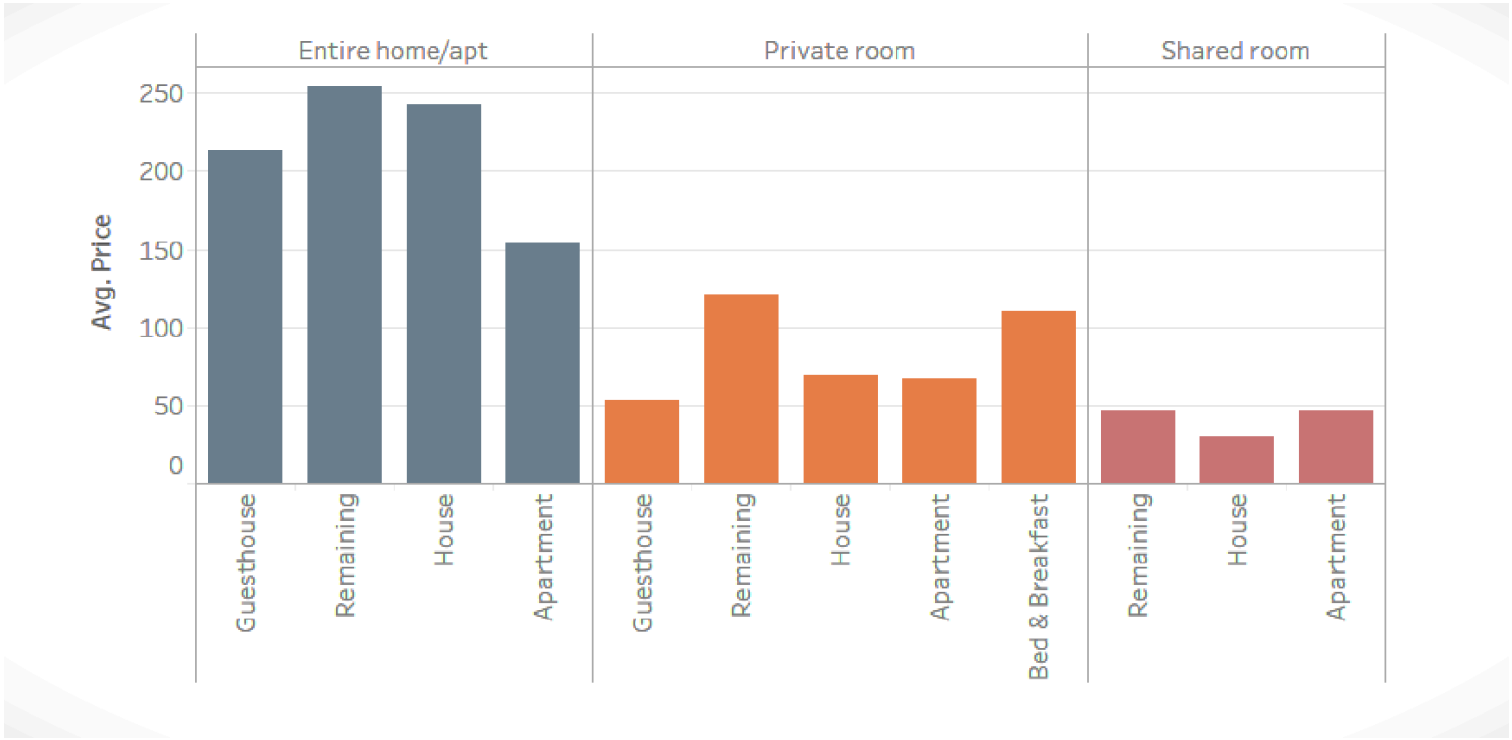
3. How the **pricing** is varying with location, property type, reviews.

a. Room Type/ Property Type :-

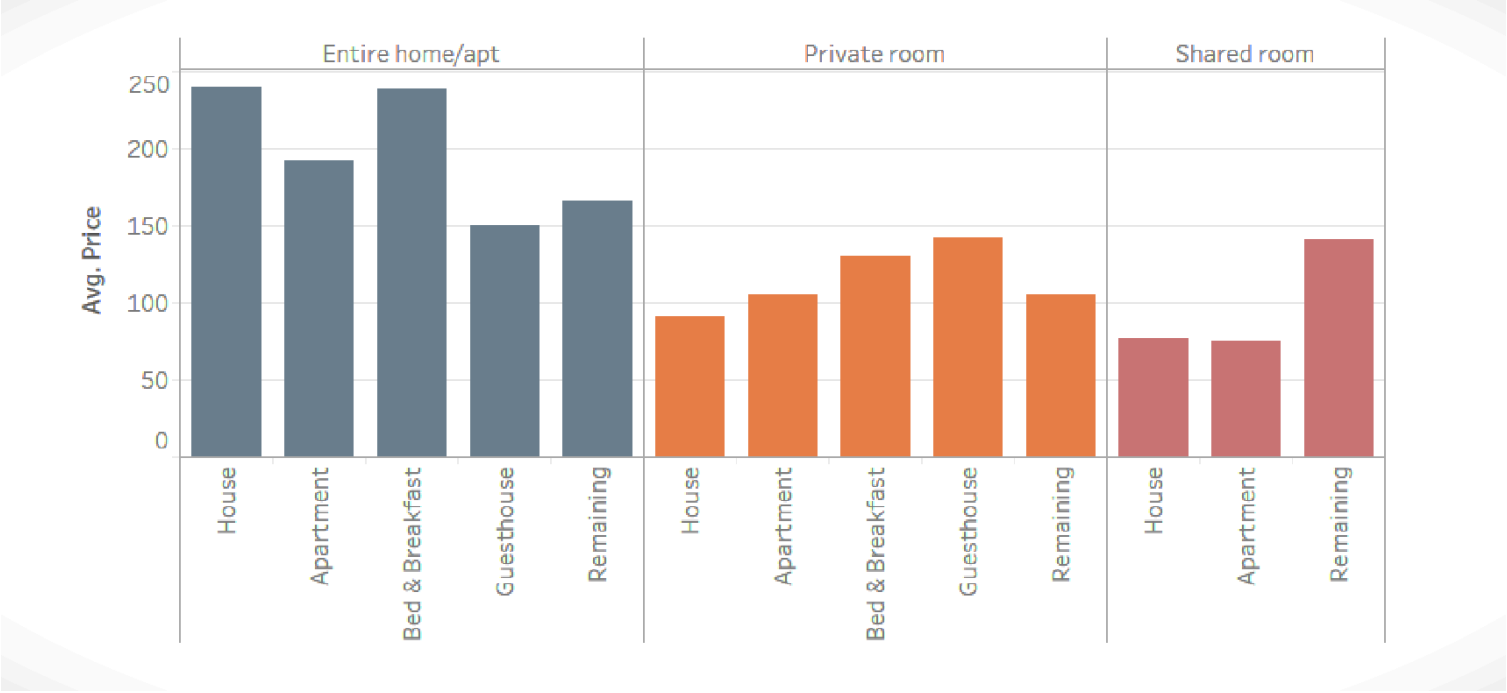
1. **Overall**



2. Chicago



3. New Orleans

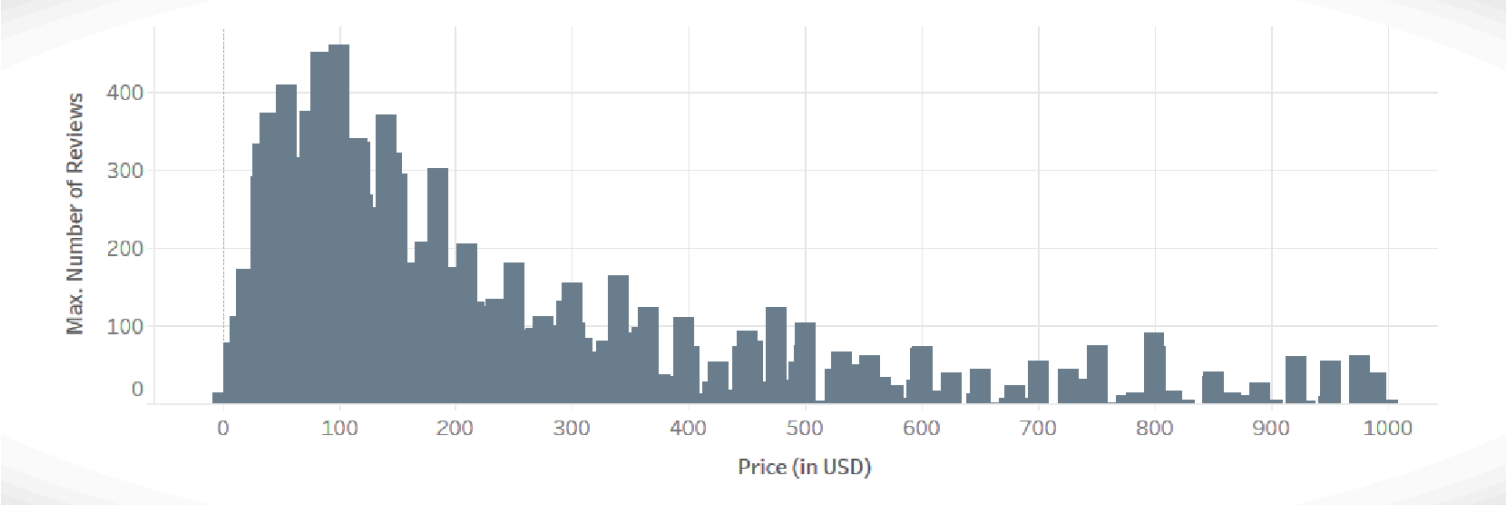


Average Price of Entire home/apt is the highest for both cities, followed by Private Room and Shared Room. An interesting finding to note is that though the number of listings for 'Apartment' type is the highest (~ 50%) as discussed in *Property Analysis- Dashboard 2,* Average price of 'Apartments' is never maximum in any of the 'Room Type'. This can be attributed by the fact that there are enough 'Apartments' available and hence there is no tight market for the 'Apartments' leading to no severe price competition.

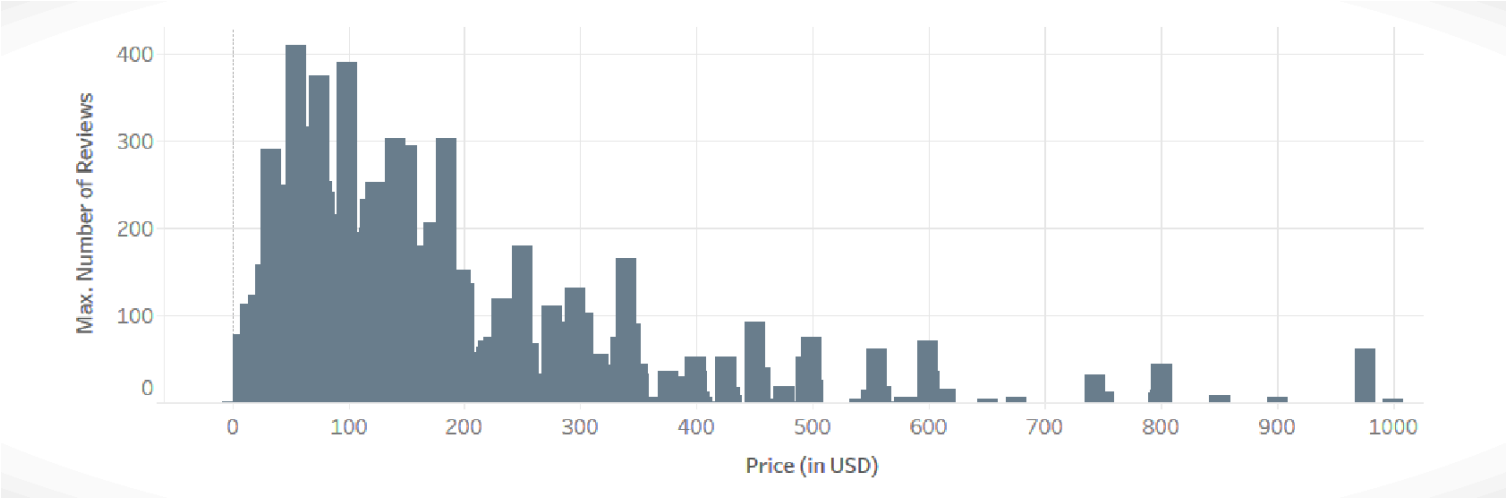
---

b. Maximum number of Reviews :-

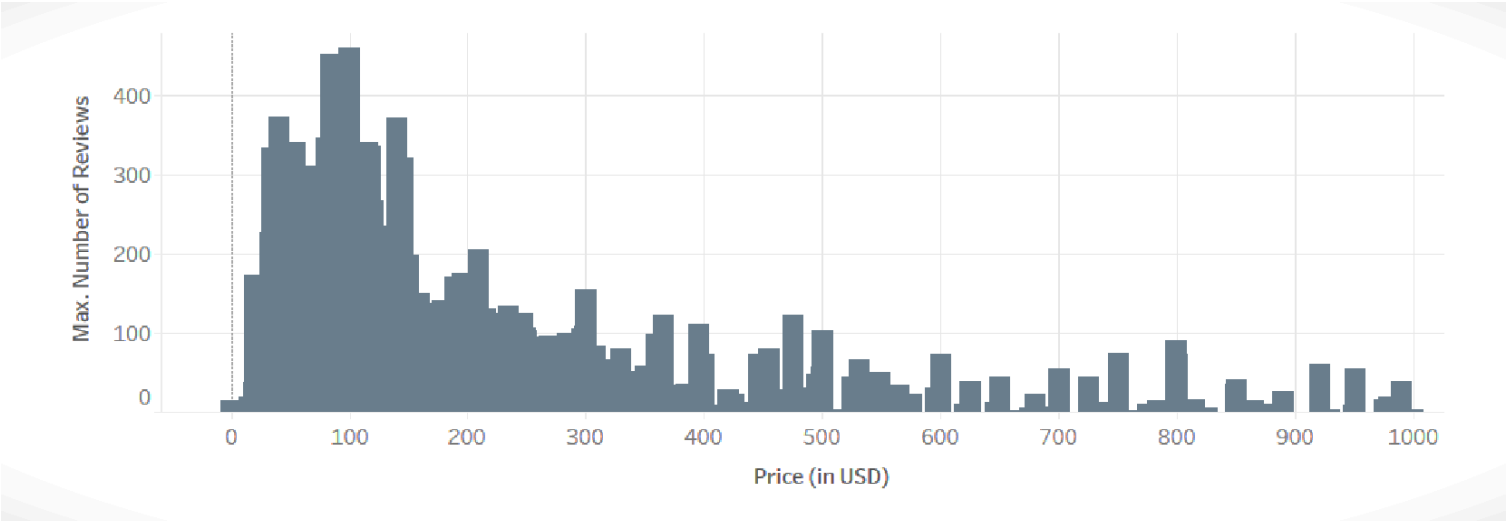
1. **Overall**



2. Chicago



3. New Orleans



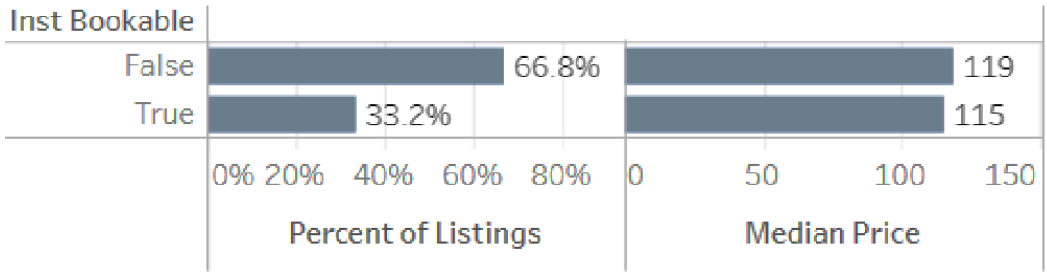
Both the cities show similar patterns where maximum number of reviews are for the listings which lie in the price range of $25-$200. There are two possible reasons for it.

- Most of the people prefer affordable and cheaper options to stay. Hence, there is a high probability that very few people take up high priced accommodations and consequently have less number of reviews. On the other hand, a major portion of people interested in affordable accommodations will correspond to more reviews. We do not have any data regarding the number of times a listing has been booked. So, we cannot validate the fact.
- Mostly, the elite/rich people opt for pricey options to stay and generally they are least bothered about posting their reviews to public forums. That could be a major reason for less reviews for high price accommodations. On the other hand, middle class people, opting for affordable options, care much more about writing reviews about the service they get.

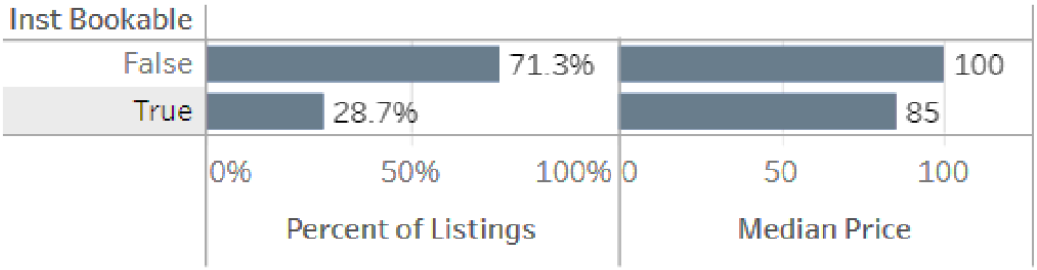
---

c. Instant Bookable v/s Non-Instant Bookable listings :-

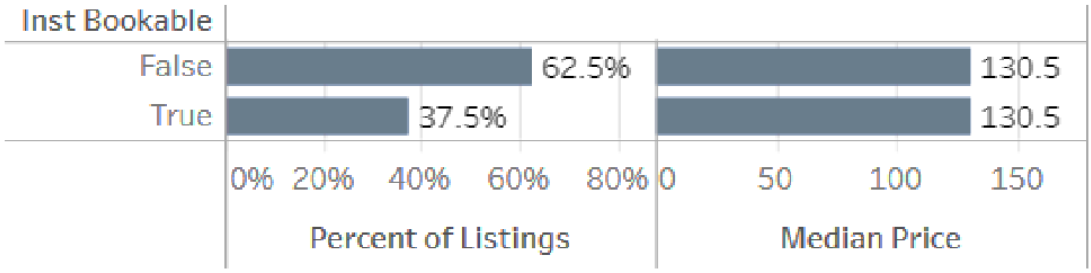
1. Overall



2. Chicago



3. New Orleans

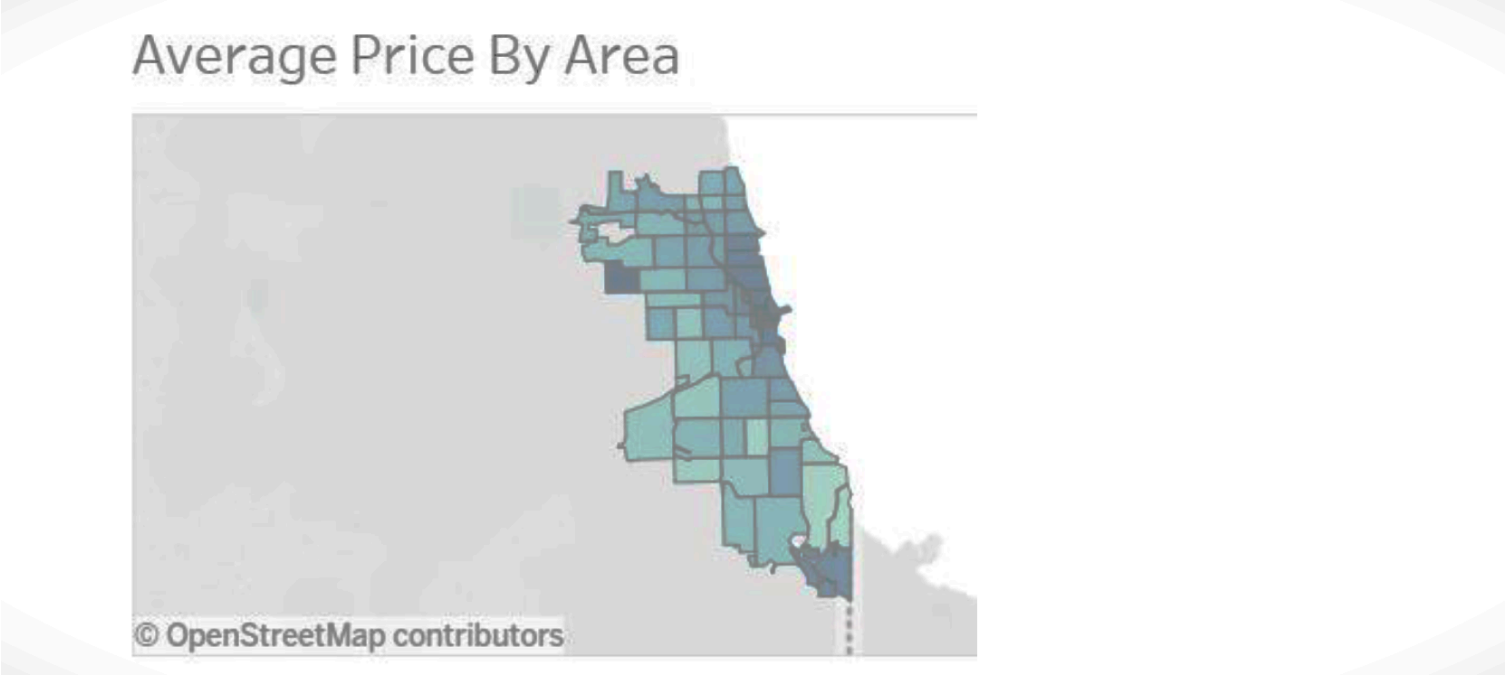


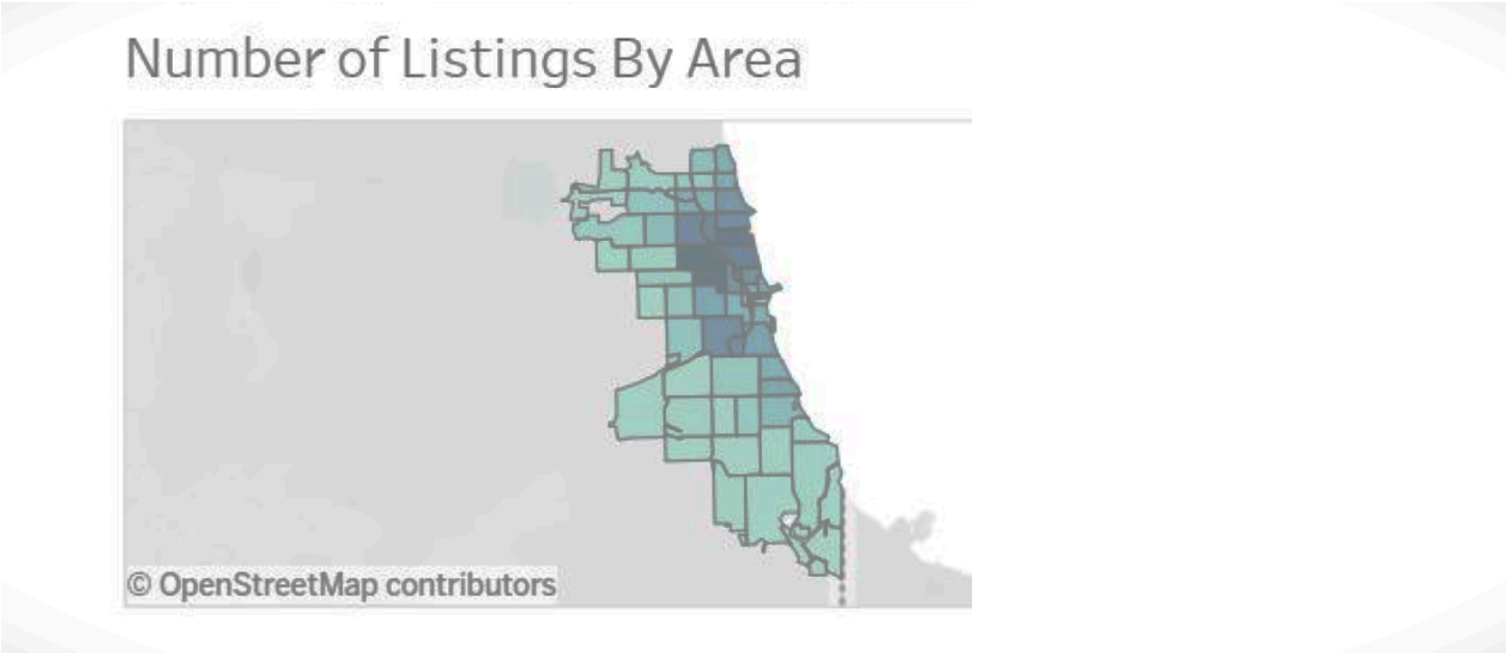
Accommodations available for 'Instant Booking' are usually considered to be lesser and the statistics also verifies the fact that overall there are only 33% listings which are available for instant booking. It is also considered a common notion that instant bookable accommodations are usually pricey as compared to non-instant bookable ones. But, in contrast to our assumptions, it can be seen that the prices of both instant bookable and non-instant bookable listings are comparable.

_____

d. Average Price by Number of Listings :-

1. **Chicago**

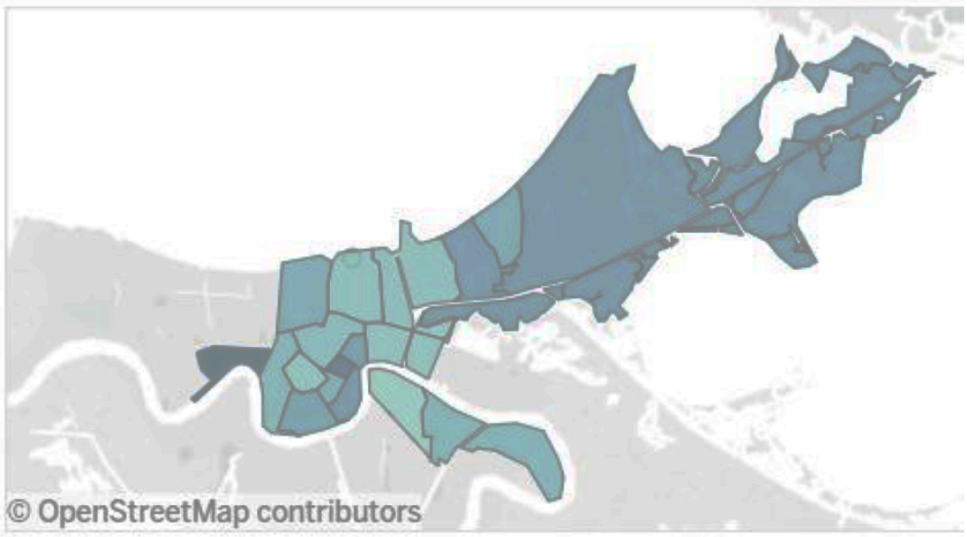




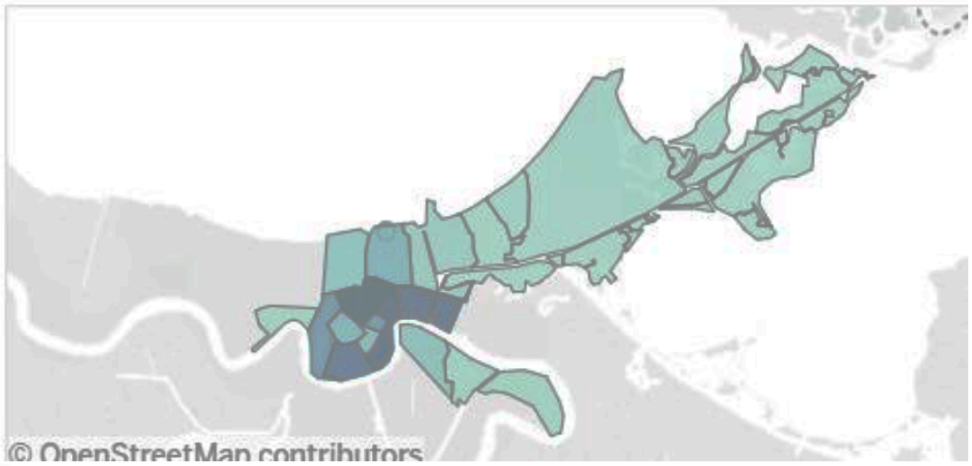
Here, we can see that the 'number of listings' are maximum along the coast line and gradually decreases as we move out of the coastal area. The average price is uniformly distributed around the city with slightly higher range around the coast line.

2. New Orleans

Average Price By Area



Number of Listings By Area



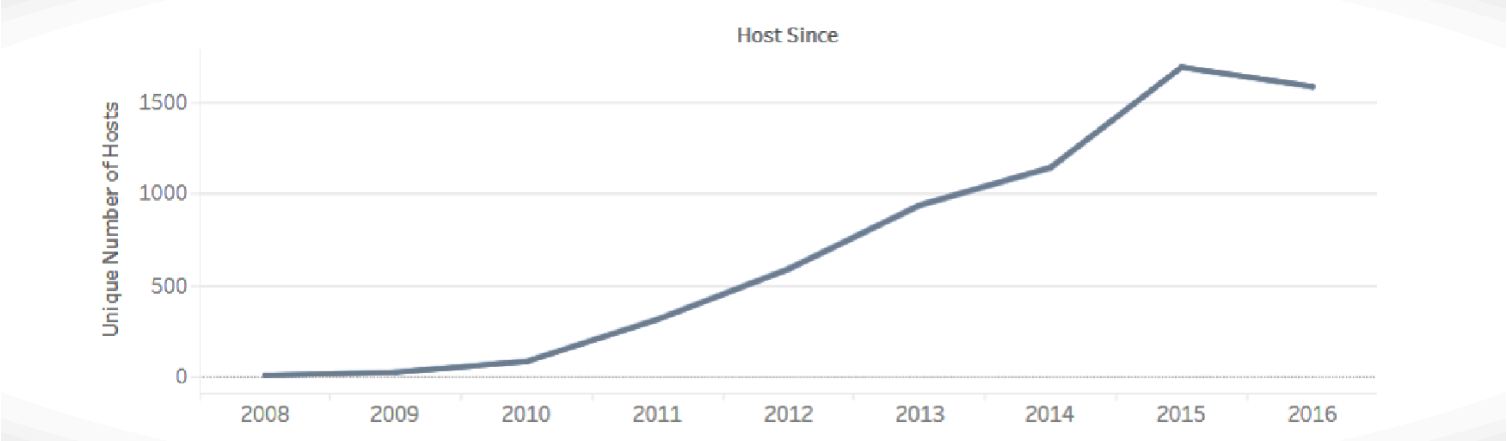
Here, we can see that 'average price by area' and 'number of listing' are contradicting each other. The areas where the number of listings is comparatively more, the average price is less. This can be attributed by the fact that the competition is high and therefore the prices kept by the hosts are reasonable. In the areas where the listings are very less, prices are high. The hosts are probably getting benefitted by high demand due to less staying options in those areas.

D) Host analysis

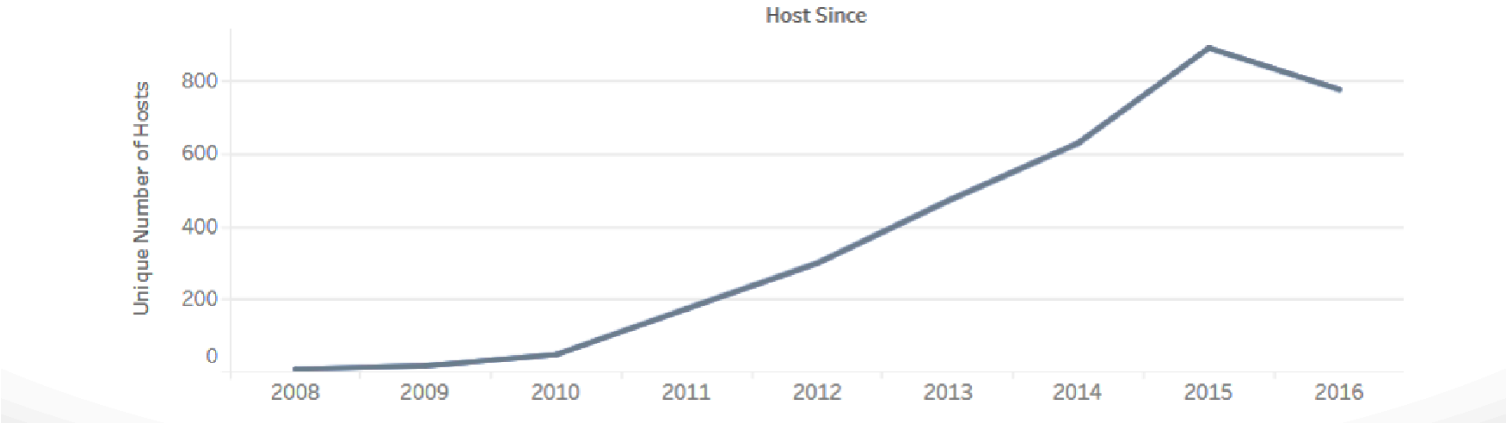
4. What are the different correlations between type of **hosts** and factors like- reviews & price.

a. Number of hosts added each year :-

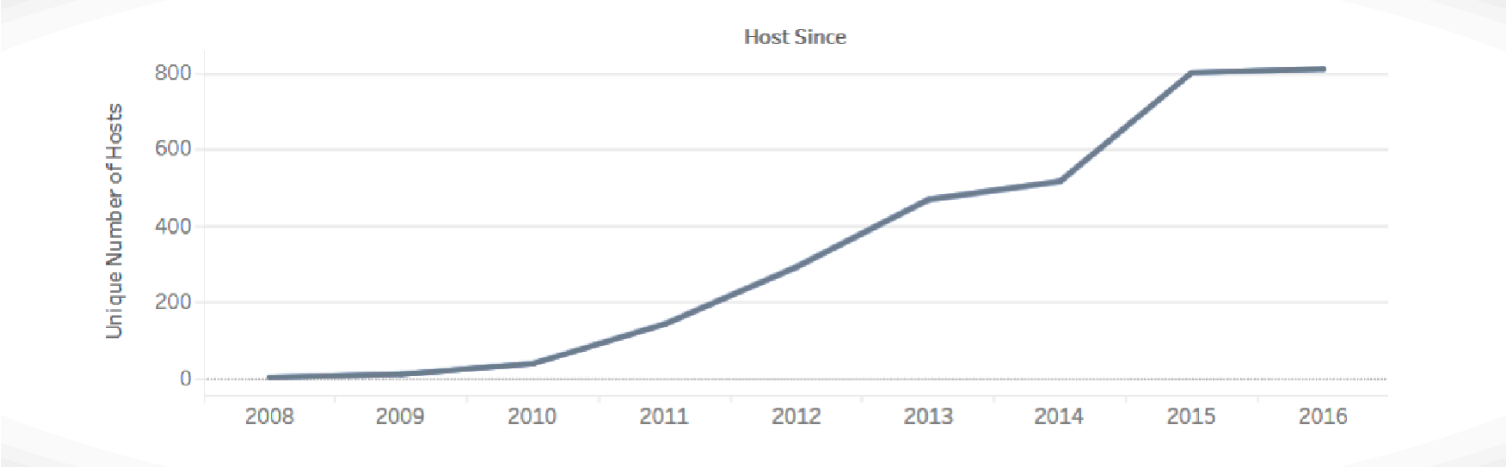
1. **Overall**



2. **Chicago**



3. New Orleans

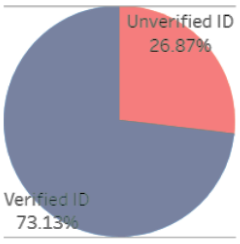


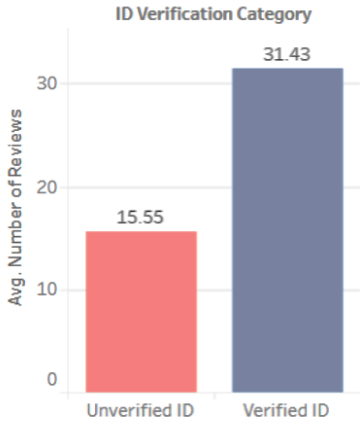
Here, we can see that the number of hosts joining Airbnb every year is continuously increasing at a higher rate, from only 8 hosts joining in 2008 to around 1500 hosts joining in 2016.

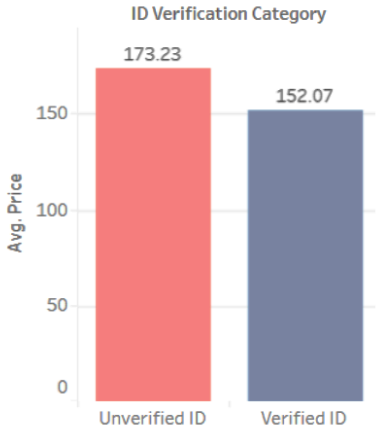
_____

b. Verified/Unverified Hosts :-

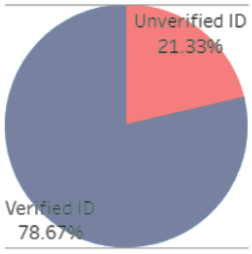
1. Overall

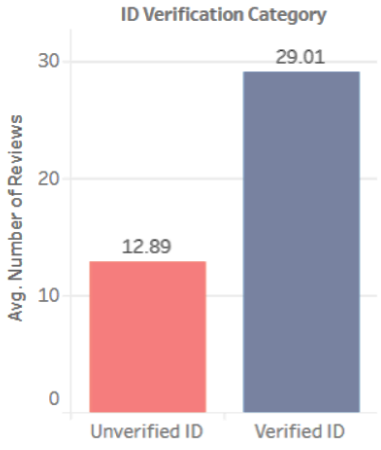

Unverified ID
26.87%
Verified ID
73.13%

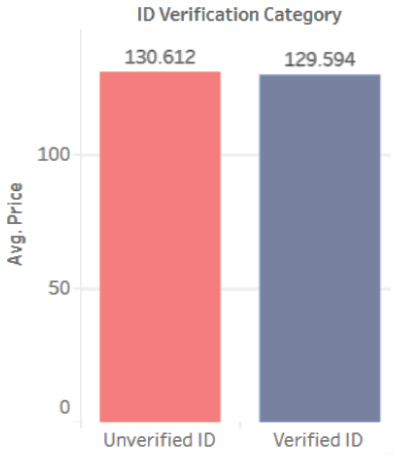

ID Verification Category
Avg. Number of Reviews
31.43
15.55
30
20
10
0
Unverified ID
Verified ID


ID Verification Category
Avg. Price
173.23
152.07
150
100
50
0
Unverified ID
Verified ID

2. Chicago


Unverified ID
21.33%
Verified ID
78.67%


ID Verification Category
Avg. Number of Reviews
29.01
12.89
30
20
10
0
Unverified ID
Verified ID


ID Verification Category
Avg. Price
130.612
129.594
100
50
0
Unverified ID
Verified ID

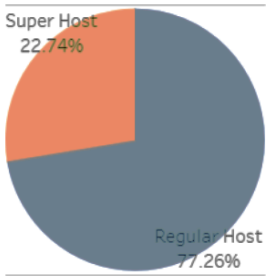
3. New Orleans

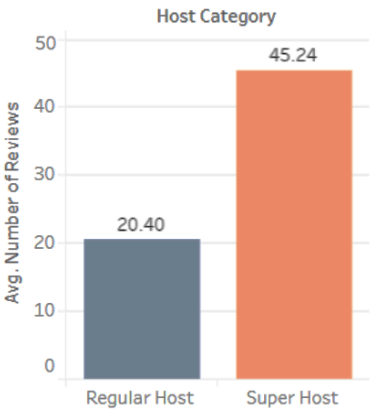
One of the findings which is very unexpected is that around one-third of the hosts are unverified on Airbnb. Although unverified hosts possess higher prices as compared to the verified hosts, they get lesser reviews. There can be two possible reasons for that. First, maybe people do not prefer taking up places by the unverified hosts and hence lesser number of people booked the listings by unverified hosts which leads to comparatively lesser reviews. Since, we do not have any data for the number of times a listing has been booked, we cannot validate this fact. Second, even if people take up the places, they might be reluctant to write reviews for unverified hosts on public forums.

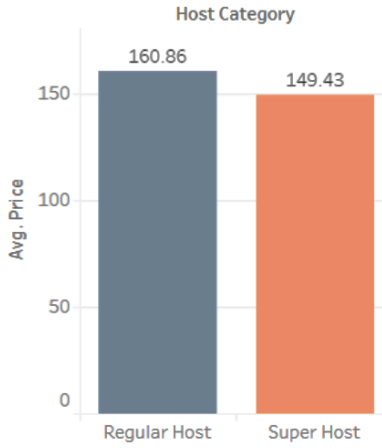
---

c. Super Hosts/Regular Hosts :-

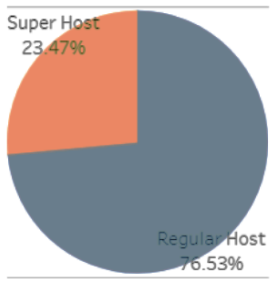
1. Overall

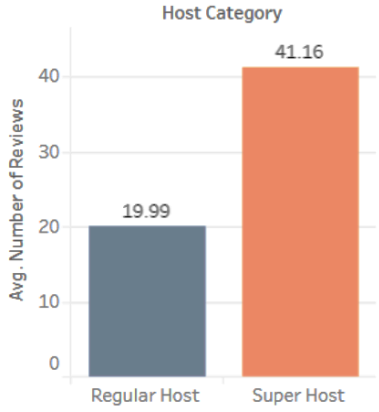


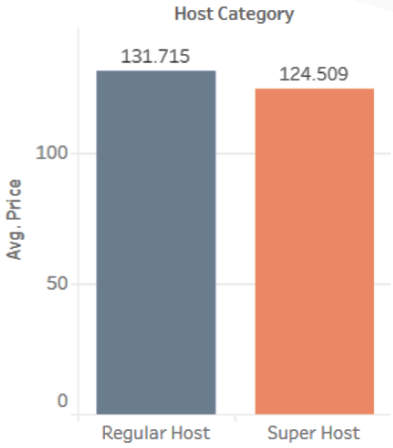




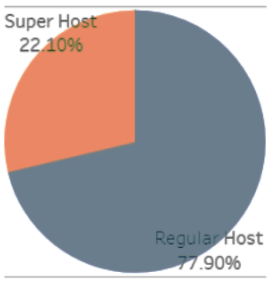
2. Chicago

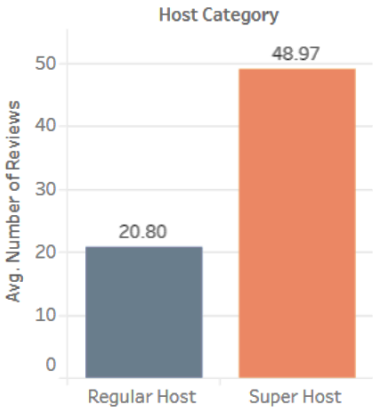


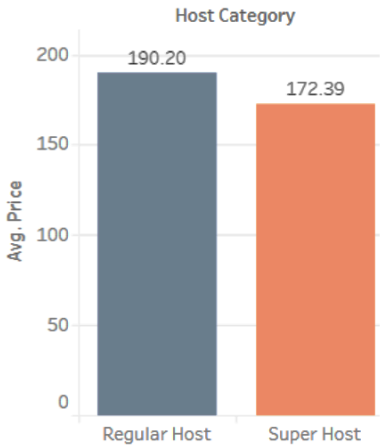




3. New Orleans







Airbnb lists its hosts into two categories- Super Hosts and Regular Hosts.

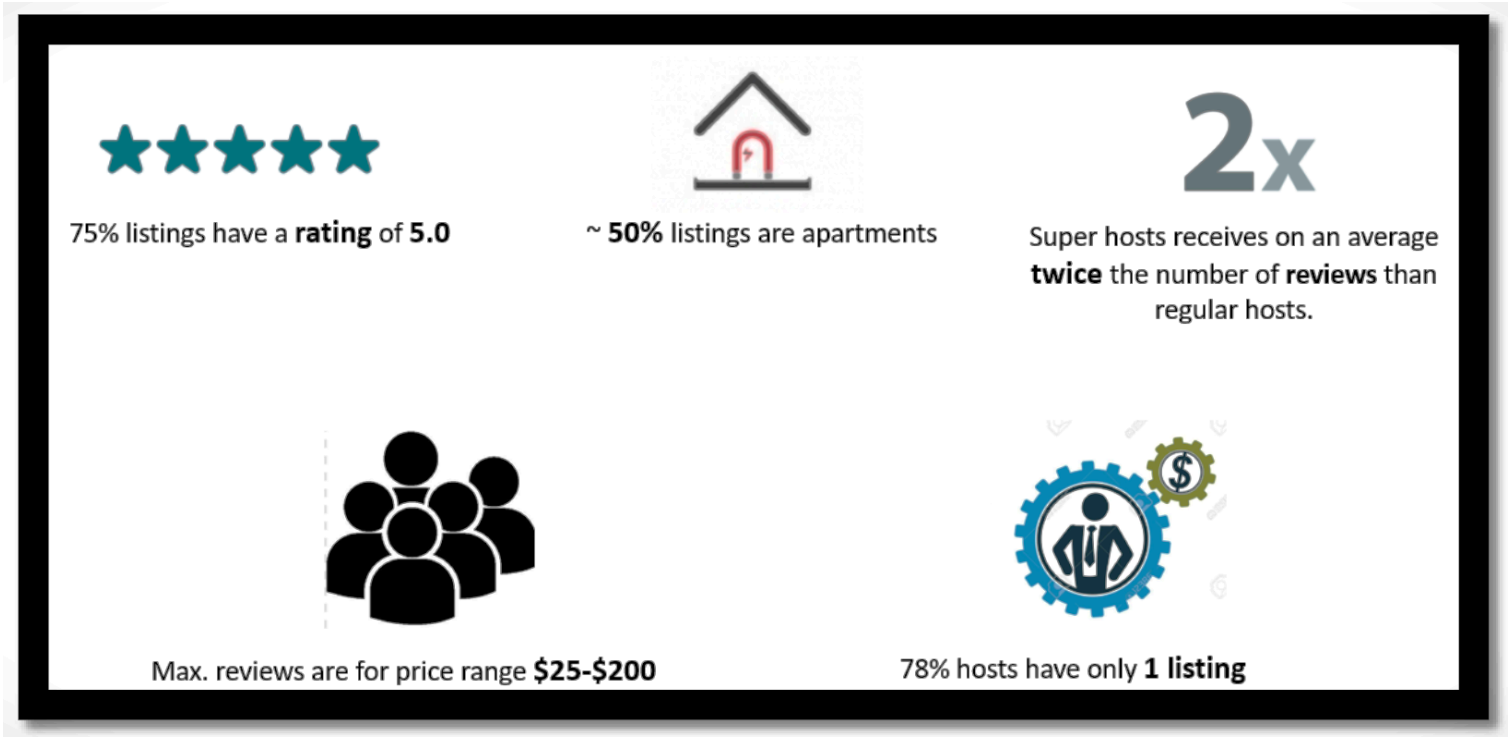
To become a Super Hosts, four criteria should be met by the hosts-

- Provide a minimum of 10 stays in a year.
- Maintain 90% response rate
- More than 80% 5-star reviews.
- Zero cancellations (with exceptions)

From the statistics, it can be seen that around one-third hosts have been designated as Super Hosts which suggests that hosts are doing pretty well in these areas. Also, Super hosts possess comparatively lesser prices and much higher reviews (around twice) as compared to regular hosts.

Major Findings

Some of the major findings through the analysis have been listed below


75% listings have a **rating** of **5.0**
~ **50%** listings are apartments
2x
Super hosts receives on an average **twice** the number of **reviews** than regular hosts.
Max. reviews are for price range **$25-$200**
78% hosts have only **1 listing**