**Project Information**

The data source is from inside Airbnb; they are activists who gather information from the Airbnb websites. The main objective for them “*Provide data that quantifies the impact of short-term rentals on housing and residential communities; and also provides a platform to support advocacy for policies to protect our cities from the impacts of short-term rentals.”*

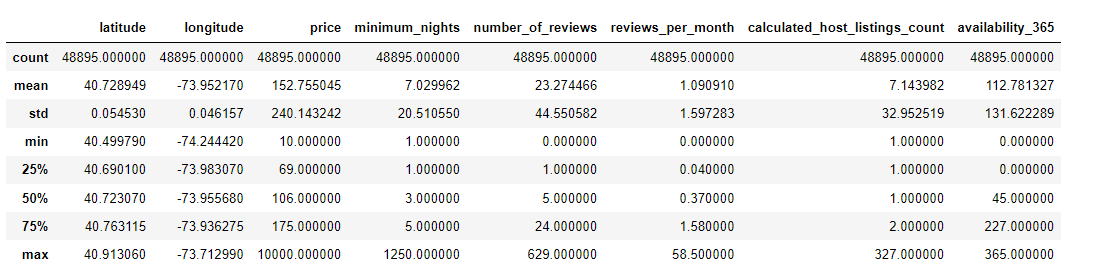
I have chosen this data set because it would be interesting to know more regarding what causes rent prices to hike and whether there are exciting factors that I can see.

**Data profile**

All the changes that have been made on the data set to provide cleaner data.

|  |  |  |
| --- | --- | --- |
| Column | Reason for change | resolution |
| Name | Mix data type | From object to string |
| Host\_name | Mix data type | From object to string |
| Last\_review | Mix data type | Object to datetime64 |
| Reviews\_per\_month | Null values | Change all null values to 0 |
| Last\_review | Null Values | Change all null values with the mean of this column |
| Price | 0 as the price | Change all the 0 price with the mean of price column |

**Descriptive Analysis**



The descriptive analysis provides information of all the variables where it would provide us a certain information like popular location listing, price, and minimum nights. The descriptive analysis also helps us provide information on where there may be any dirty columns.

**Limitations and Ethics**

1. Since the data that was provided by the Airbnb website, that could be a sample bias where Airbnb does not release all New York listings and only a certain listing.
2. The data was gathered during that particular time, which means many listing can be deleted and added, resulting in inaccurate data.
3. The address of the actual listings have been anonymized by Airbnb, which is good in terms of privacy, but this could cause an issue in the analysis since the data provided is not accurate.

There is no ethics issue that I can see; Airbnb has worked to the best of its ability to protect its customer’s data.

**Questions:**

1. Which location would be the most popular? Is there a reason for it?
2. How many houses and apartments are being rented out frequently to tourists and not to long-term residents?
3. How many houses and apartments are being rented out for people with long-term stays?
4. How much are hosts making from renting to tourists compared to long-term rentals?
5. What is the most popular price that provides the best revenue?
6. Does the number of reviews have an impact on the revenue?
7. Which room type is the most popular, and which room type generates revenue?
8. Is there any specific host that stands out in generating revenue, and why?