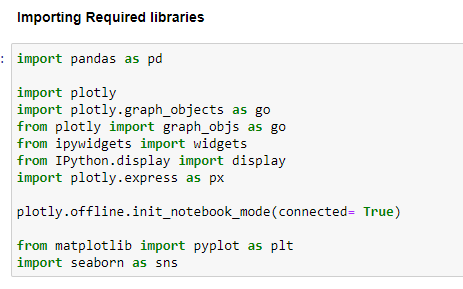
**Methodology Document**

**Libraries were imported using the below code**

****

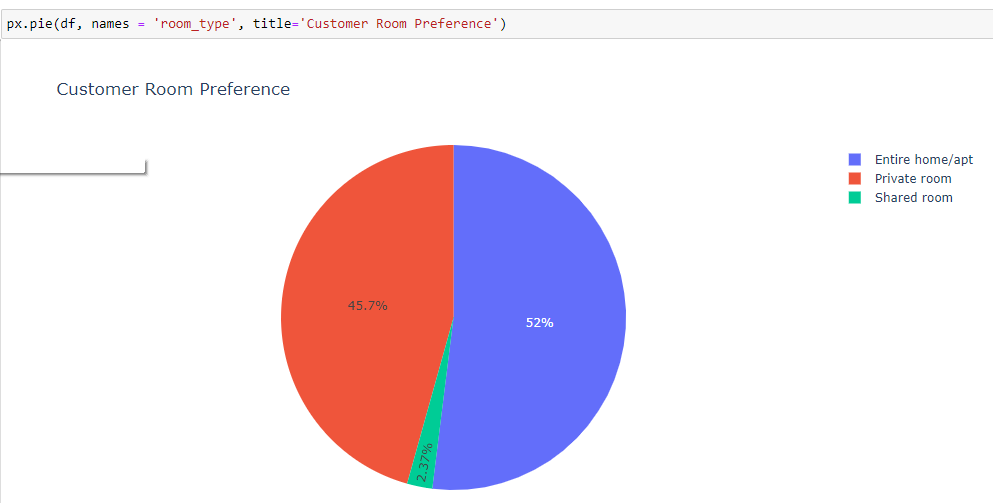
**Data has been imported using the below code**

****

**Assumption:** The data analysis was done based on the following data assumption that a host creates a listing based on customer preference- higher is the number of listings in a particular place higher is the demand among the customers for that place.

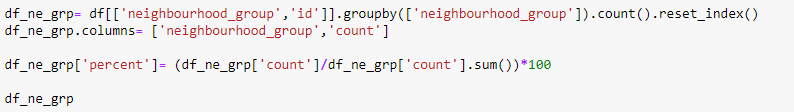
**Exploratory Data Analysis was performed on the data to understand the data :**

* **CUSTOMER PREFERENCES FOR ROOM TYPE:** Since the cardinality is less for room type, a pie chart has been used to classify the percentage constitution for a given room type**.**

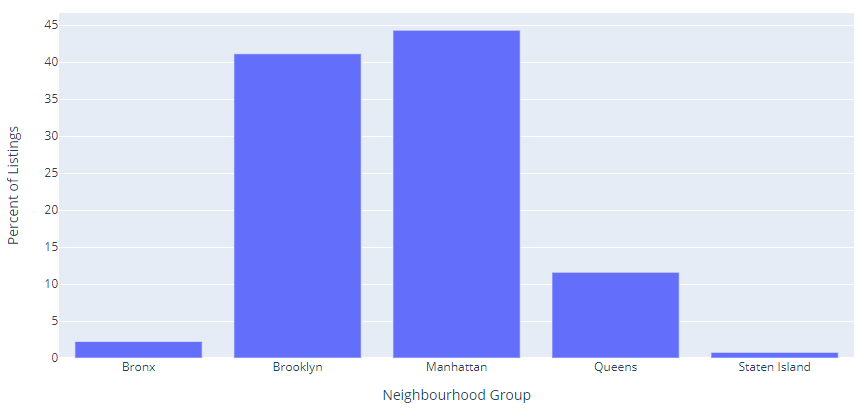
****

Customers prefer Entire home/apt or private rooms more than sharing the room with others.

* **CUSTOMER PREFERENCES FOR NY CITY BOROUGHS:** A bar chart has been used to classify the most popular neighborhoods in NYC.

****

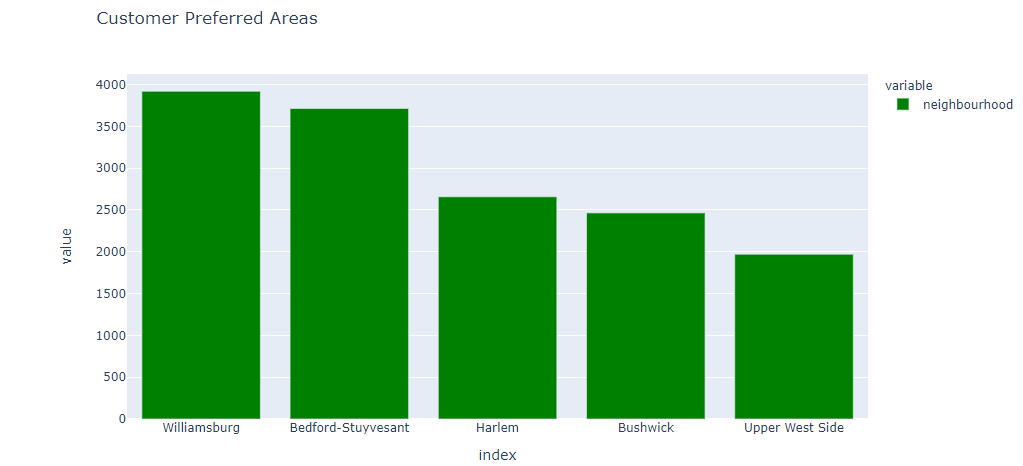
****

****

Customers prefer to live in the most developed parts of New York city like Manhattan and Brooklyn than the under developed parts like Bronx and Staten Island.

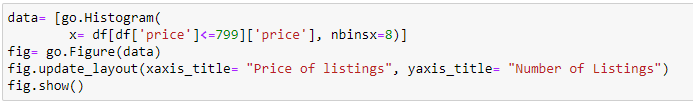
* **CUSTOMER PREFERRED AREAS IN NEW YORK CITY:** A bar chart has been used to understand the customer preferred areas



****

Customers prefer to stay in the more developed areas like Manhattan and Brooklyn.

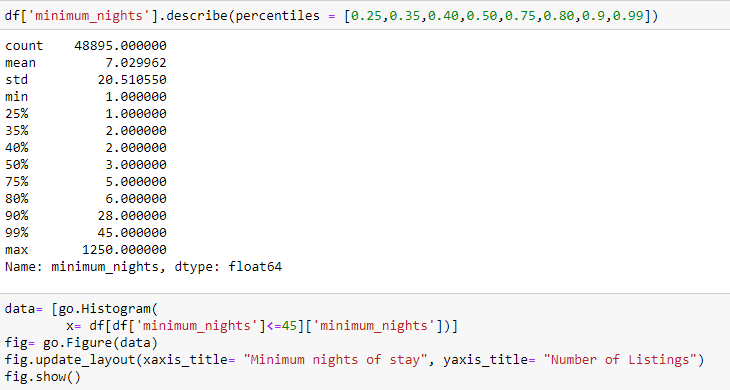
* **CUSTOMER PREFERENCES BASED ON PRICE:** A histogram has been used to display the price range that is most popular among the customers. The data is classified into 8 bins.

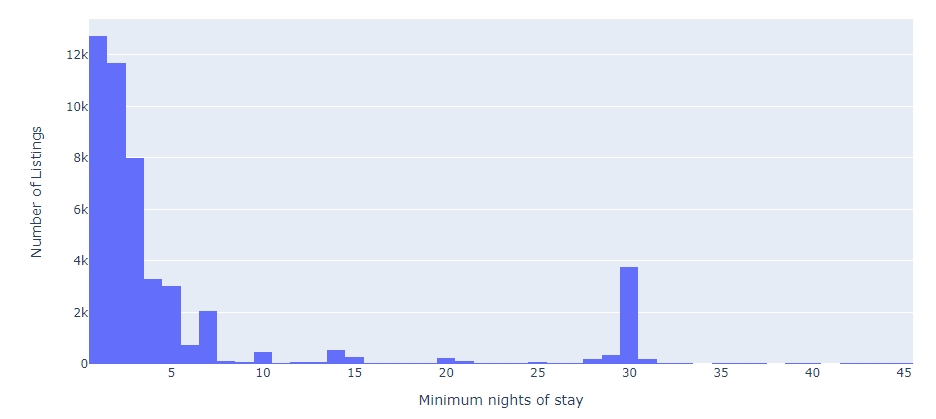




Majority of the listings fall under the 0 – 200 dollar bucket. Customers prefer the listings which fall in the lower bucket range.

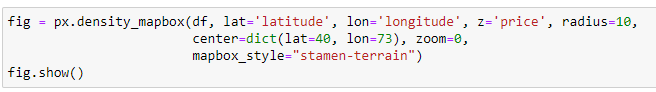
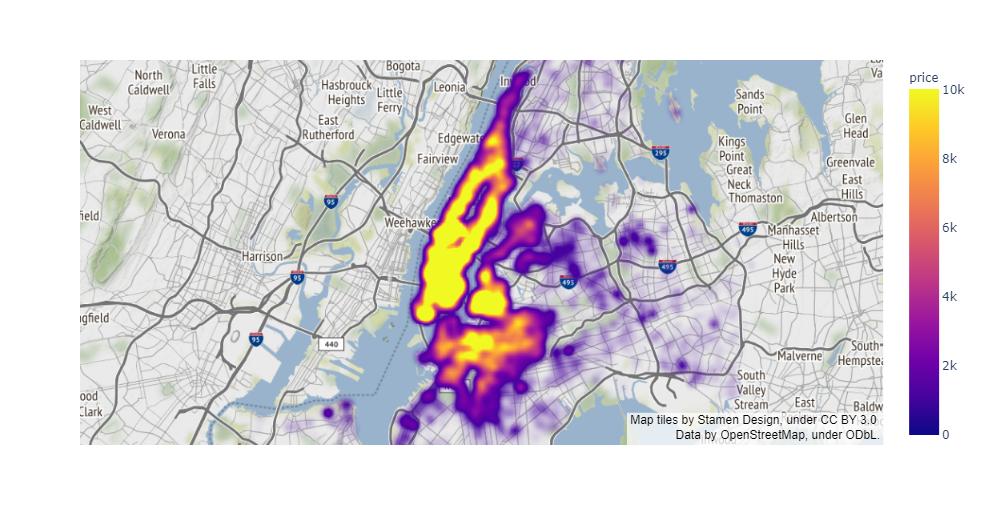
* **CUSTOMER PREFERENCES ON MINIMUM NIGHTS CRITERIA:** Owing to the high cardinality in the minimum night’s variable, a bar plot has been used to display the in-demand requirement among the customers.

****

****

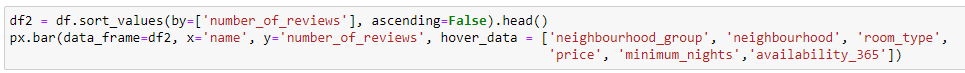
Customers prefer the listings where the minimum nights criteria is less. Majority of the listings are observed wherein the minimum nights required is between 1 to 7 days.

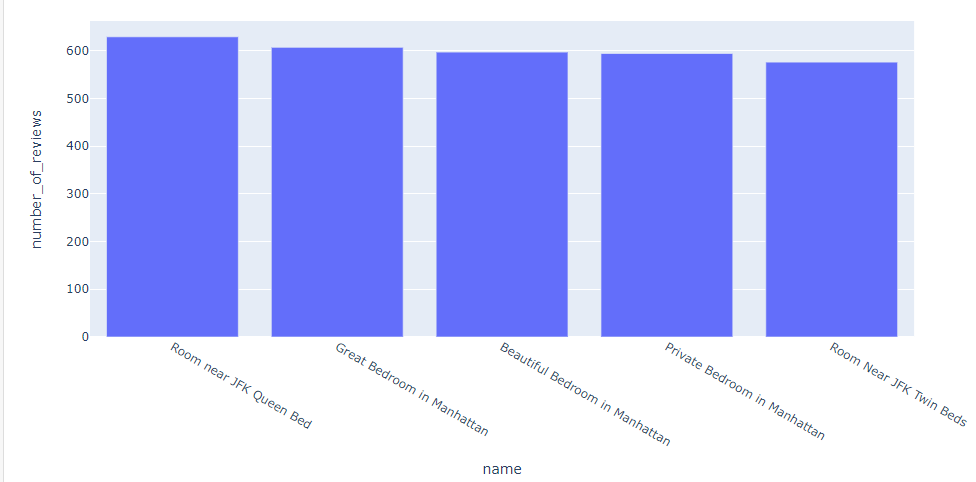
* **GEOGRAPHICAL INSIGHTS OF THE LISTINGS:** With the help of a map, we were able to identify the correlation of price and the listing location**.**

**** 

Listings near to the rivers are priced higher when compared to others. Since Manhattan is mostly enclosed by river in its sides, the listings are priced higher.

* **CUSTOMER PREFERRED LISTINGS:** A bar plot has been used to understand the most responsive customers are from which location.

****

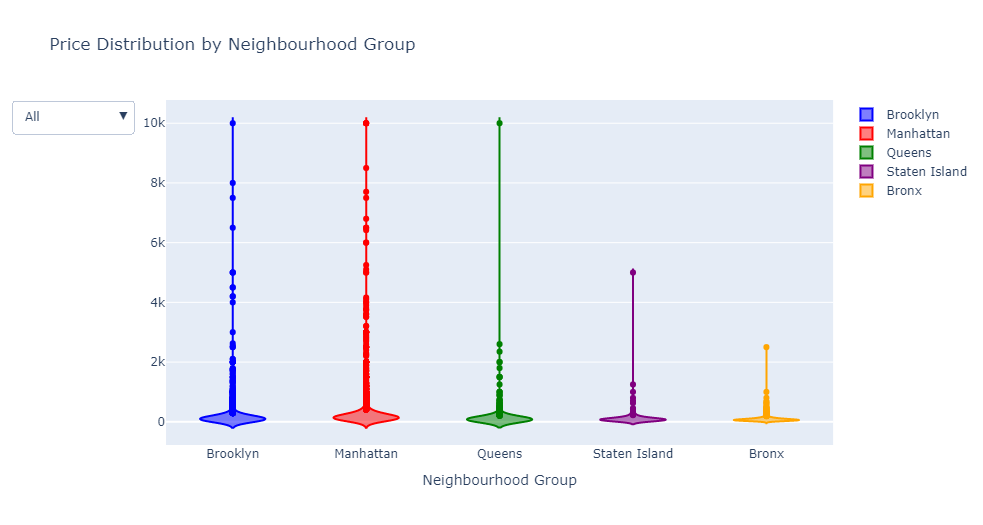
****

The "Room near JFK Queen Bed" is the most popular with 629 reviews and is situated in Queens

* **PRICE DISTRIBUTION BY NEIGHBOURHOOD GROUP:** Violin plot has been used to get the price distribution based on location.

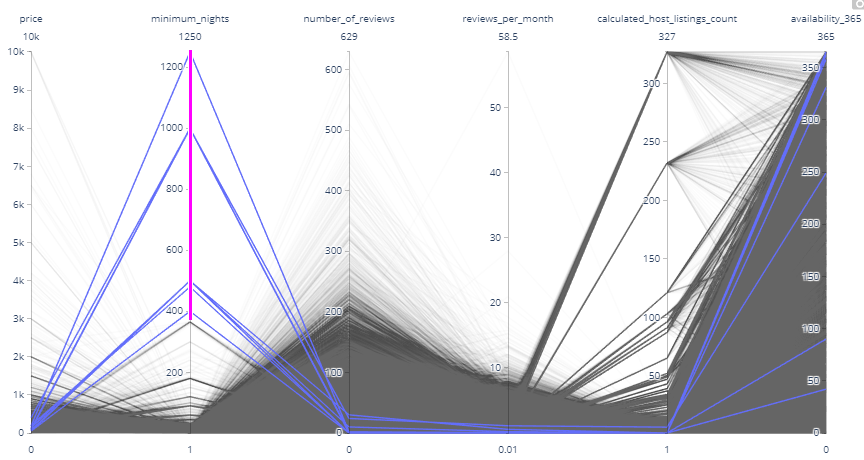




The highest priced Airbnb's are in Brooklyn, Manhattan, Queens and Bronx being the lowest. Manhattan has a median price of 150 dollars whereas Brooklyn has a median price of 90 dollars.

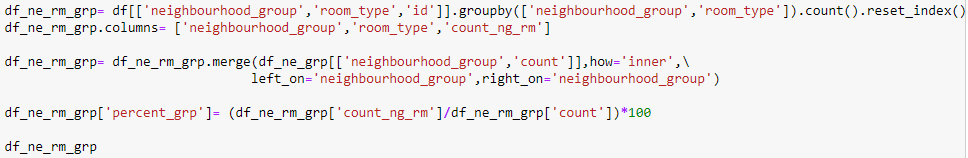
* **LISTINGS WITH MINIMUM NIGHTS MORE THAN 400 DAYS:** Parallel coordinates plot is used to compare here the several individual observations on the set of the numerical variables present here. It has helped to visualize what is the price ,number of reviews, availability, etc for the listings which have minimum number of nights more than a year.

****

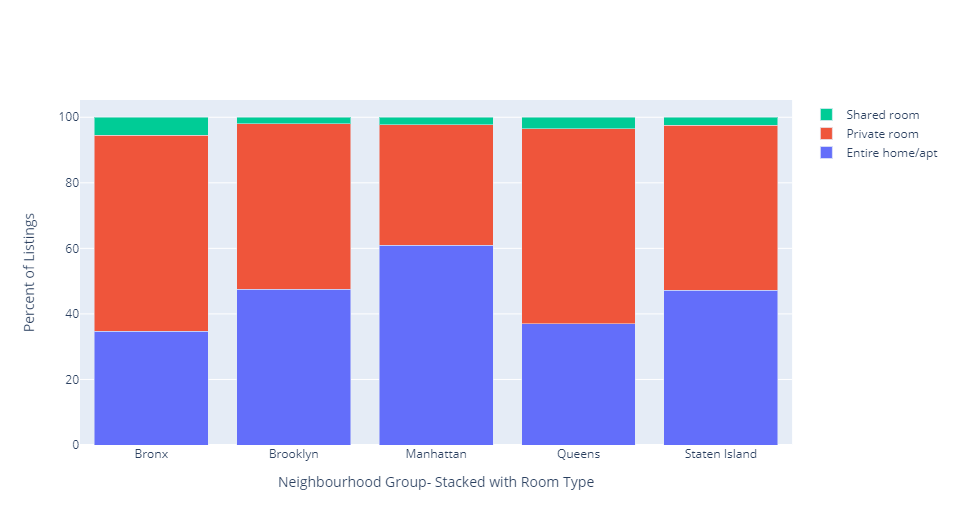
****

There are some listings where customers can stay for over a year or more during their work stays or studies. These listings are of very lesser price and are available 365 days which allow customers to stay for more than a year**.**

* **CUSTOMER PREFERENCES OF APARTMENTS NY CITY BOROUGHS:** Stacked Bar chart has been used to analyse what kind of apartments are preferred by customers in each of the NY City Boroughs

****

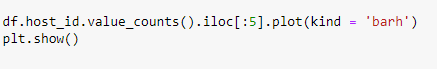
****

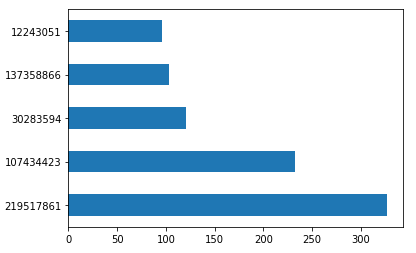
****

It can be seen that Customers prefer to live in either Entire home/apt or Private room in Manhattan, Brooklyn and Queens, whereas shared apartments are preferred in Bronx.

* **MOST POPULAR HOST:**

The Bar plot has been used to see who is the most popular host in NY City who has transacted the most over the year of 2019.

****

****

The host Sonder with host id 219517861 is the most preferred host.