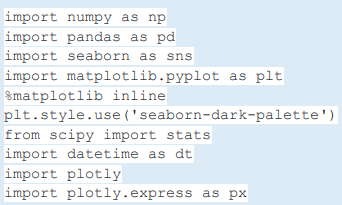
# **METHODOLOGY DOCUMENT**

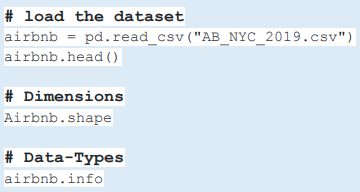
* **Research Problem**
* For the past few months, Airbnb business has seen a significant decline in revenue due to travel restrictions because of the Covid-19 pandemic.
* The revenue took the largest hit in NYC in the Q2 of 2020.
* Now that the restrictions have started lifting and people have started traveling, Airbnb wants to make sure it is fully prepared for the change.
* **Objectives**
* Improve our strategies to revive the impact of Covid-19 on the economic and market conditions of Airbnb, NYC.
* Understand the customer preference and user experience trends for Airbnb, NYC.
* Provide recommendations for new acquisitions and improve customer experience.
* **Data Assumptions**
* Assumed that the data prior to the Covid-19 period was achieving the desired goals.
* Airbnb wants to continue its business in NYC and has no plans of expanding to other territories.
* The strategies decided were considered keeping in mind that there will be no further travel restrictions.
* **Data Methodology**
* **Tools used -** Python for Data Understanding, Pre-processing, and general Univariate Analysis. Tableau and Excel were used for in-depth Bi-Multivariate Analysis.

1. Data Understanding and Preparation:

* The following relevant libraries were imported.



* The dataset was loaded, datatypes of variables were checked and along with that the dimensions and size of the data frame was checked.

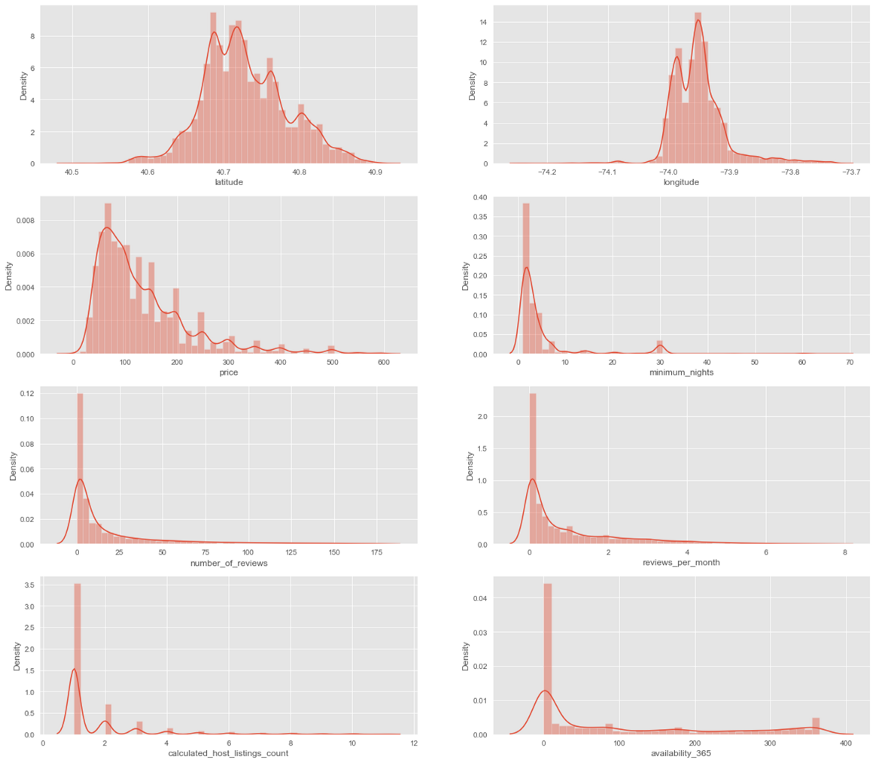


1. Handling Missing Values and Outliers:

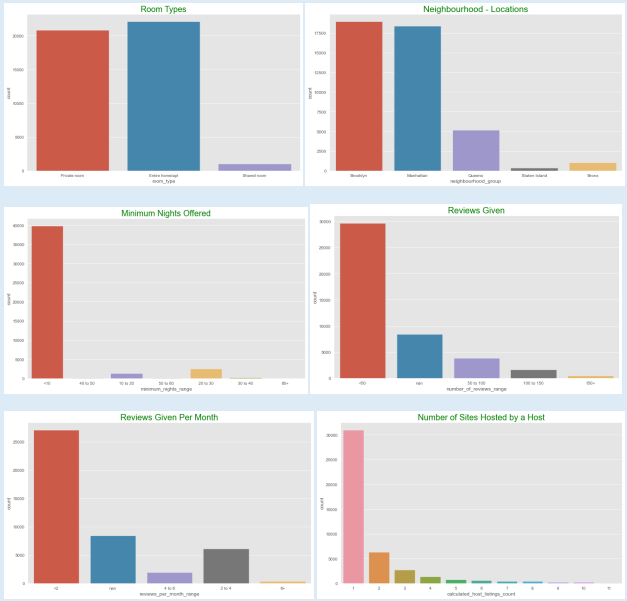
* The missing values and outliers were checked in the data frame.
* The following columns had missing values - last\_review, reviews\_per\_month, host\_name, and name.
* These columns had NaN values - last\_review and reviews\_per\_month indicating some listed properties didn’t receive reviews.
* The following columns had outliers - price, minimum\_nights, number\_ of\_reviews, reviews\_per\_month, and calculated\_host\_listings\_count.
* The missing values and outliers were not treated for.

1. Data Analysis:

* Univariate Analysis done on Numerical Variables

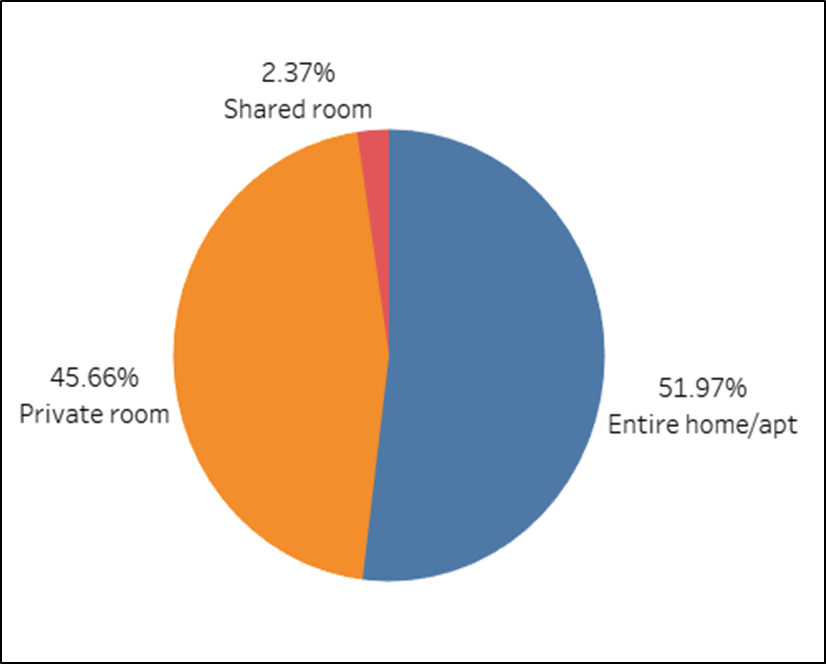


* Univariate Analysis done on Categorical Variables



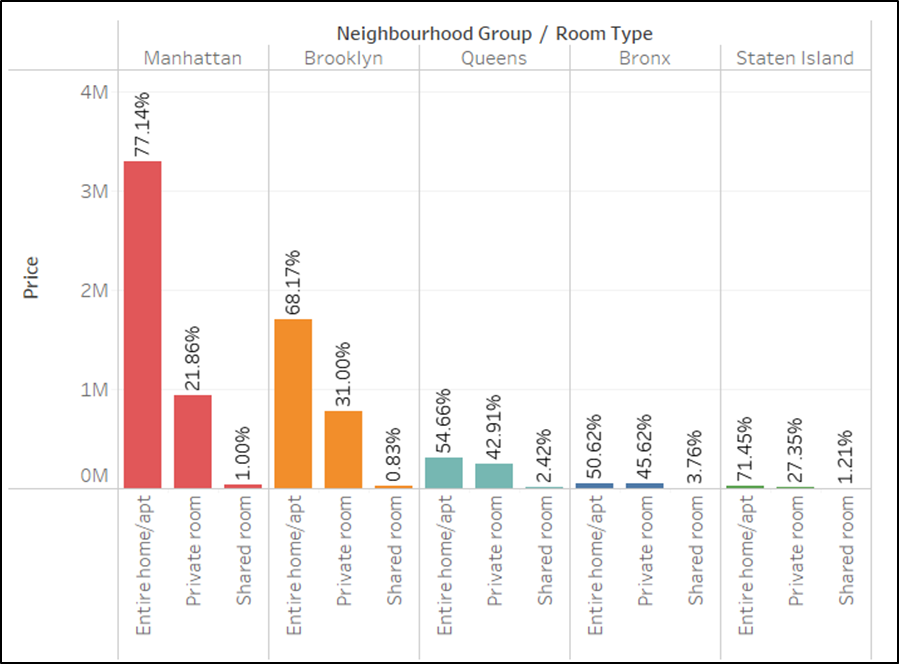
* Bi-Multivariate Analysis done in Tableau

1. Customer Preferences of the Three Property Types



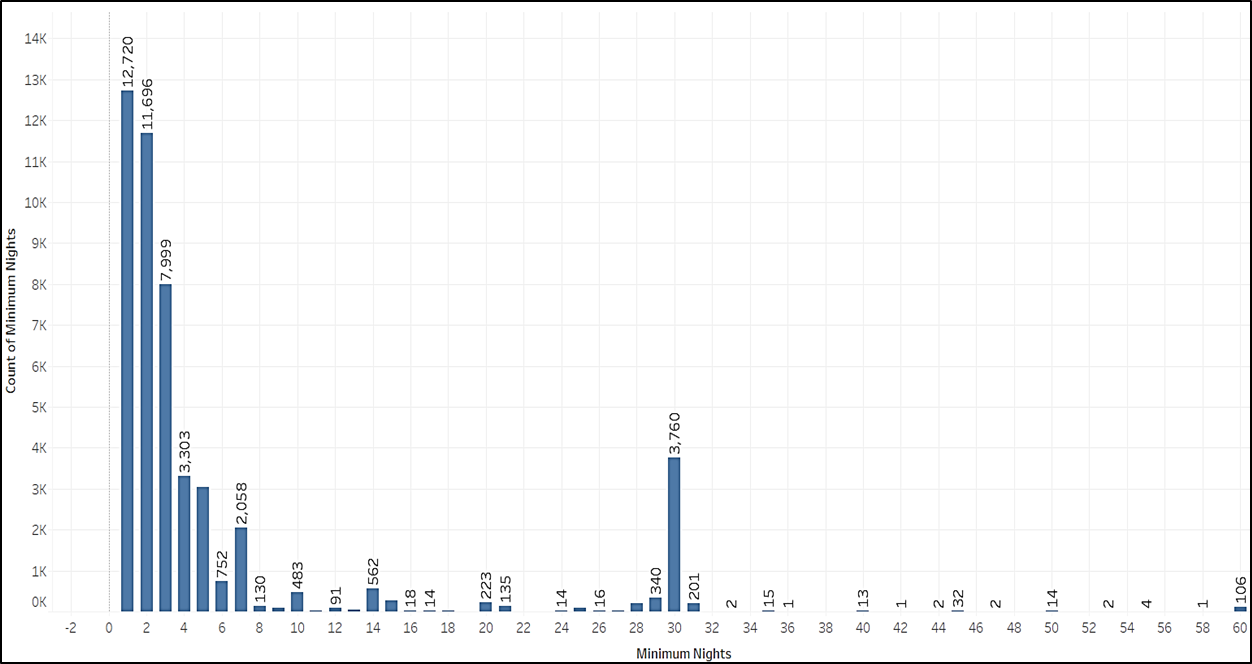
* The properties of Entire home/apt and Private room are preferred over shared rooms by Airbnb hosts offering rentals in NYC.
* The room types ‘Entire home/apt’ and ‘Private room’ accounts for around 97.6% of the listed properties in NYC.
* Shared rooms account for only 2.4% of the total listed properties.

1. Prices vs Room Type in Neighbourhood Group



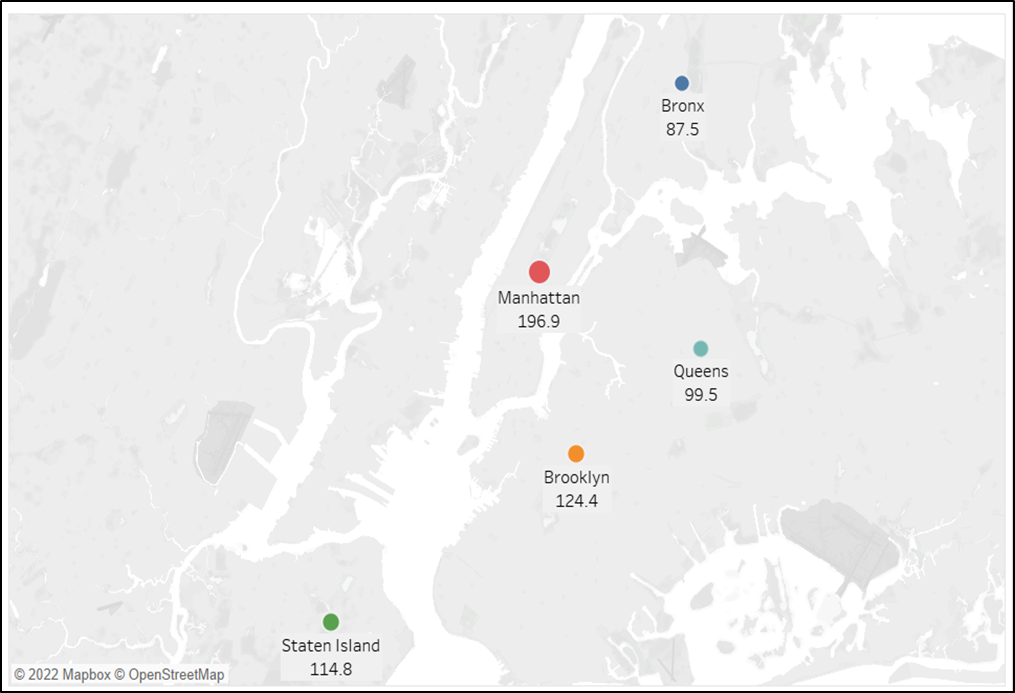
* Manhattan has the most expensive properties followed by Brooklyn while the Bronx and Staten Island are the cheapest.
* Entire homes/apt in both Manhattan, as well as Brooklyn, are the most expensive room type followed by Private rooms.
* In Queens and Bronx, the prices of room types of entire homes/apt and private rooms have a close difference of 5-10% as compared to the other neighbourhood groups.

1. Customer Preferences of Minimum Nights Per Stay



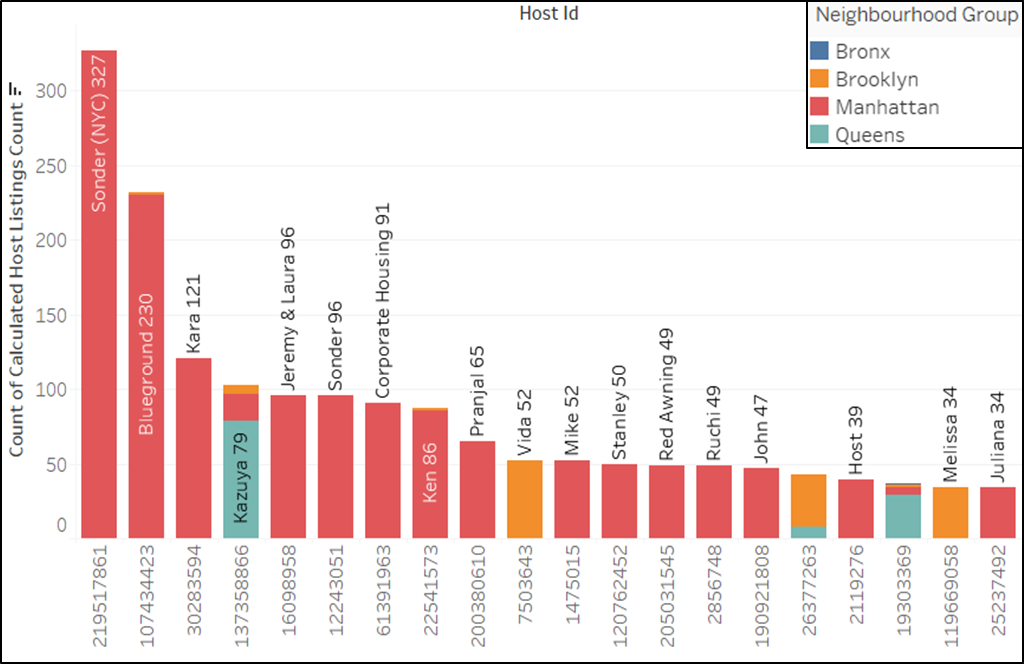
* 12.7k customers prefer to stay for a day on their visit closely followed by 11.6k customers who stay for 2 nights.
* Customers also prefer to stay for 3-7 nights.
* Customer preferences for night stay rise on 14 nights, 30 nights - 1 month, and 60 nights - 2 months.

1. Average Price Range of the Neighbourhood Groups



* Manhattan is the only neighbourhood group that has the most expensive average price of approx. $200 on the properties listed.
* Brooklyn and Staten Island have an average price of $124 and $114 resp. which is less expensive compared to Manhattan.
* Queens and Bronx have an average price of $100 and $88 resp. which is the most affordable neighbourhood in NYC.

1. Count of Top 20 Host Listing in the Neighbourhood Group



* Sonder(NYC) has the highest number of rooms listed in Manhattan – 327 followed by Blueground at 230.
* Manhattan is the most listed neighbourhood group followed by Brooklyn and Queens.
* Bronx has only 1 listing whereas Staten Island does not fall in the Top 20 category.

1. Neighbourhood vs Avg Price – Neighbour Group

Chart, bar chart

Description automatically generated

* + - Top two neighbourhood with highest average price of properties - Fort Wadsworth & Woodrow are in Staten Island.
    - Manhattan has 12 neighbourhood in the top 20 leading with Tribeca & Battery Park City.
    - Brooklyn, Queens & Bronx have only 1 neighbourhood each – Sea Gate, Riverdale & Neponsit respectively in the Top 20 category.

1. Host/Neighbour Group vs No. of Reviews

Chart, waterfall chart

Description automatically generated

* + Majority of reviews are for properties based out of Manhattan. The host Sonder (NYC) has maximum reviews.
  + Queens and Brooklyn hardly make it to the review list while Staten Island and Bronx are not present at all.
  + Need to inform the hosts specifically the non-Manhattan neighbourhood group to ensure receiving reviews since most bookings are based on the reviews.