Data Methodology

***Step 1: Storyboarding:***

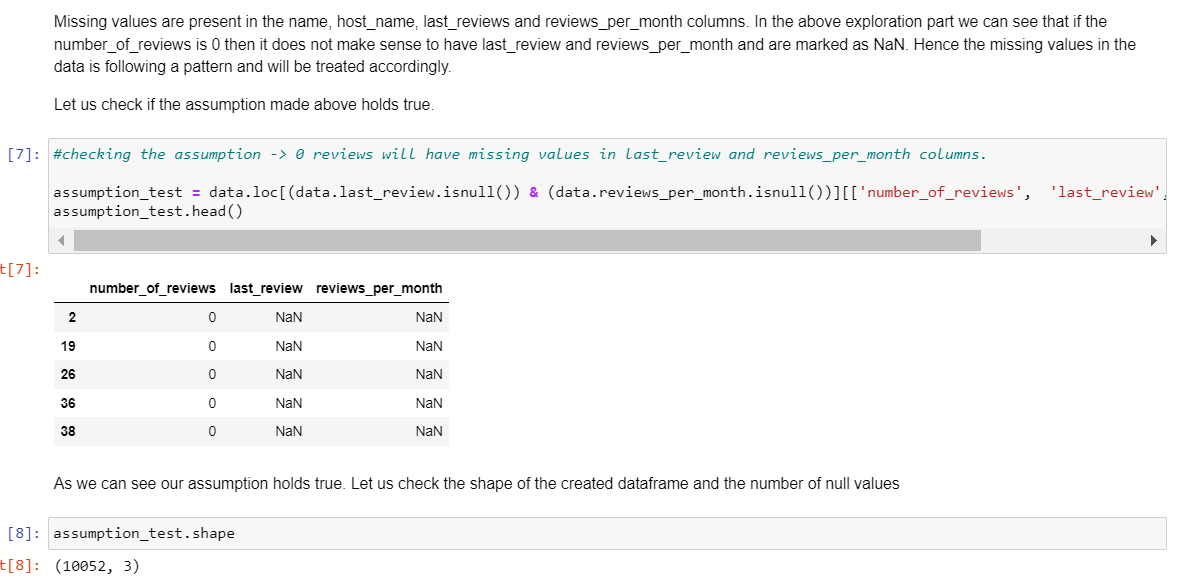
* I familiarized myself with the data by going over it and making a note of key fields.
* Created a mind map for the presentation's slides.
* Made a rudimentary template using this mind map as a guide.

***Step 2: Data Wrangling:***

* Explored all the columns in the dataset by importing it to python notebook
* Checked for the Missing values and found out the columns “name, host\_name, last\_review and reviews\_per\_month” had missing values.



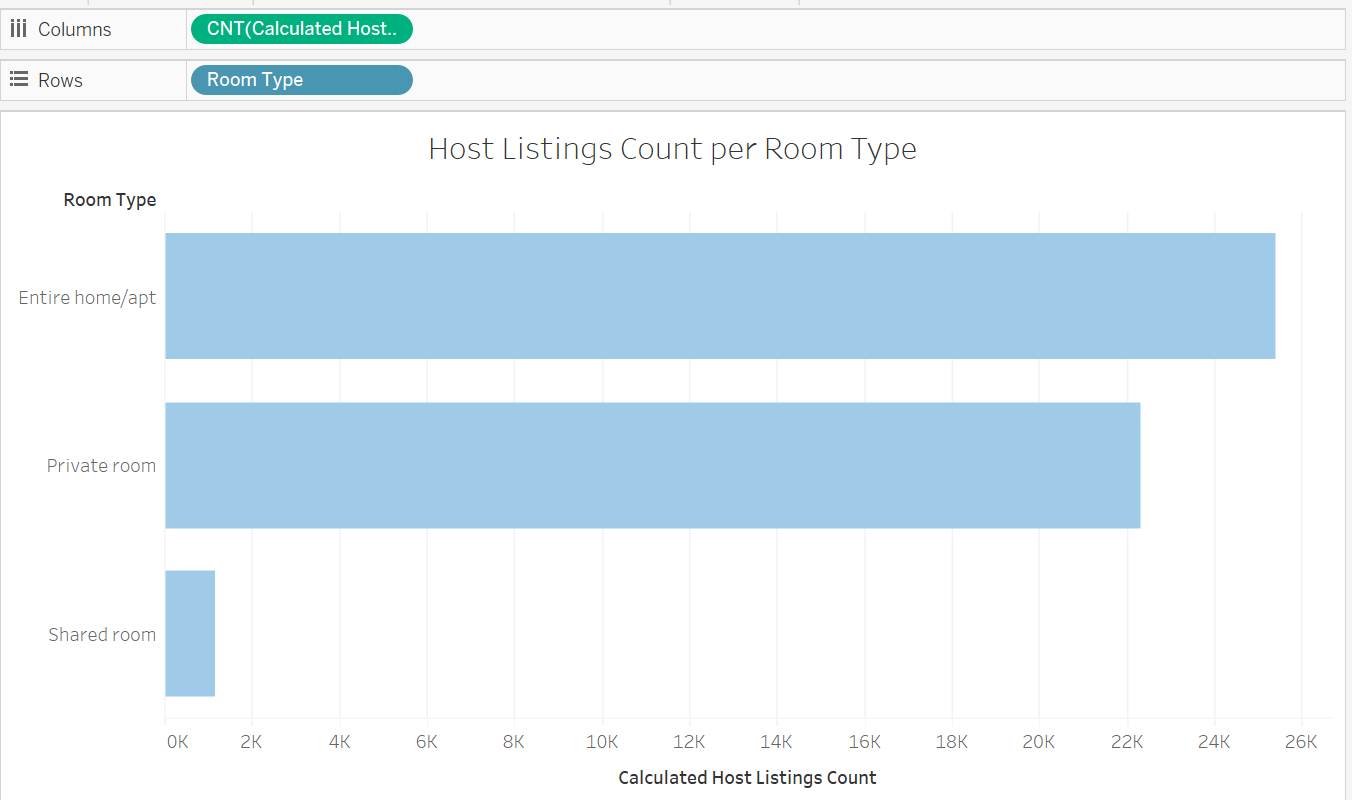
* Missing Values are treated using Python. Below are few snapshots.



A screenshot of a computer

Description automatically generated with medium confidence

* Host Listings count is maximum for entire apartment and private room and is very small for shared room as seen below.



* Checked if any Outliers are present with respect to price.

A screenshot of a computer screen

Description automatically generated with low confidence

* Created a grouped field for Minimum Number of Days assuming null values belonged to the category.

A screenshot of a computer program

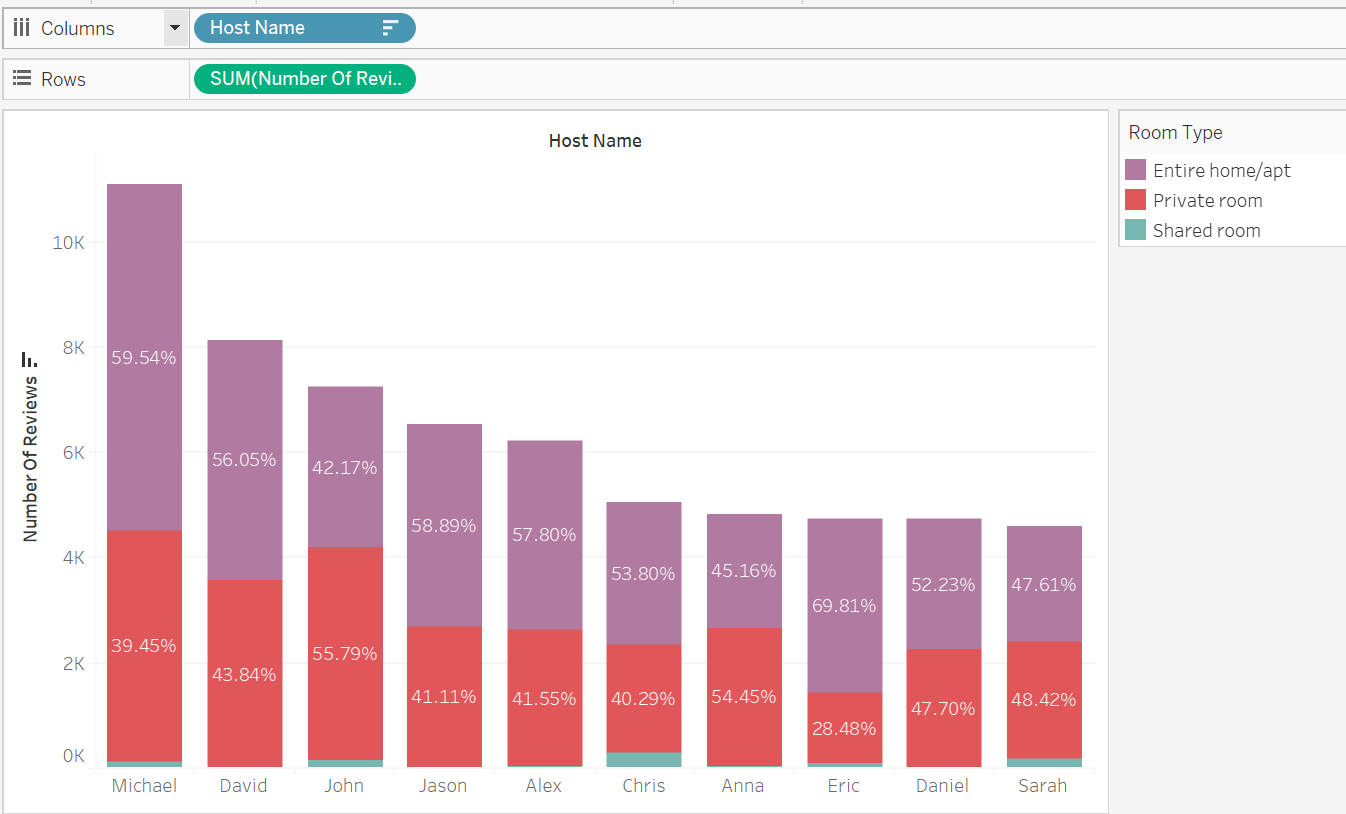
Description automatically generated with low confidence

***Step 3: Data Analysis:***

Performed Data Analysis using Tableau with the cleaned Data file. Below are few analysis the Data Analytics team can use.

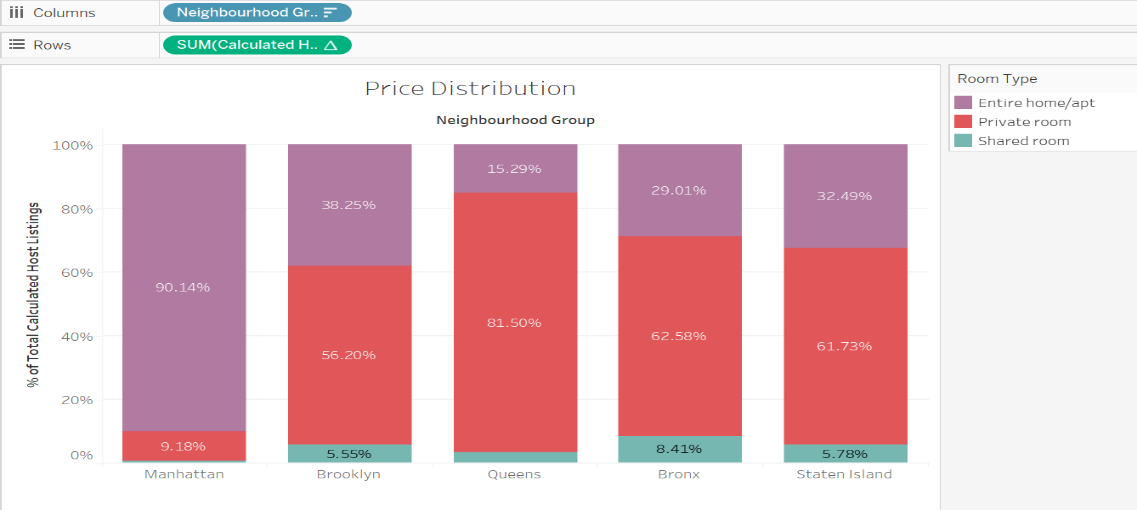
1. Top 10 Hosts by Reviews:

* Shared rooms account for the least number of reviews of the total listed properties.
* More than 50% of the hosts prefer renting out the entire home/apartment.
* Private room & Entire home/apartment seems to be popular (more than 90% reviews).



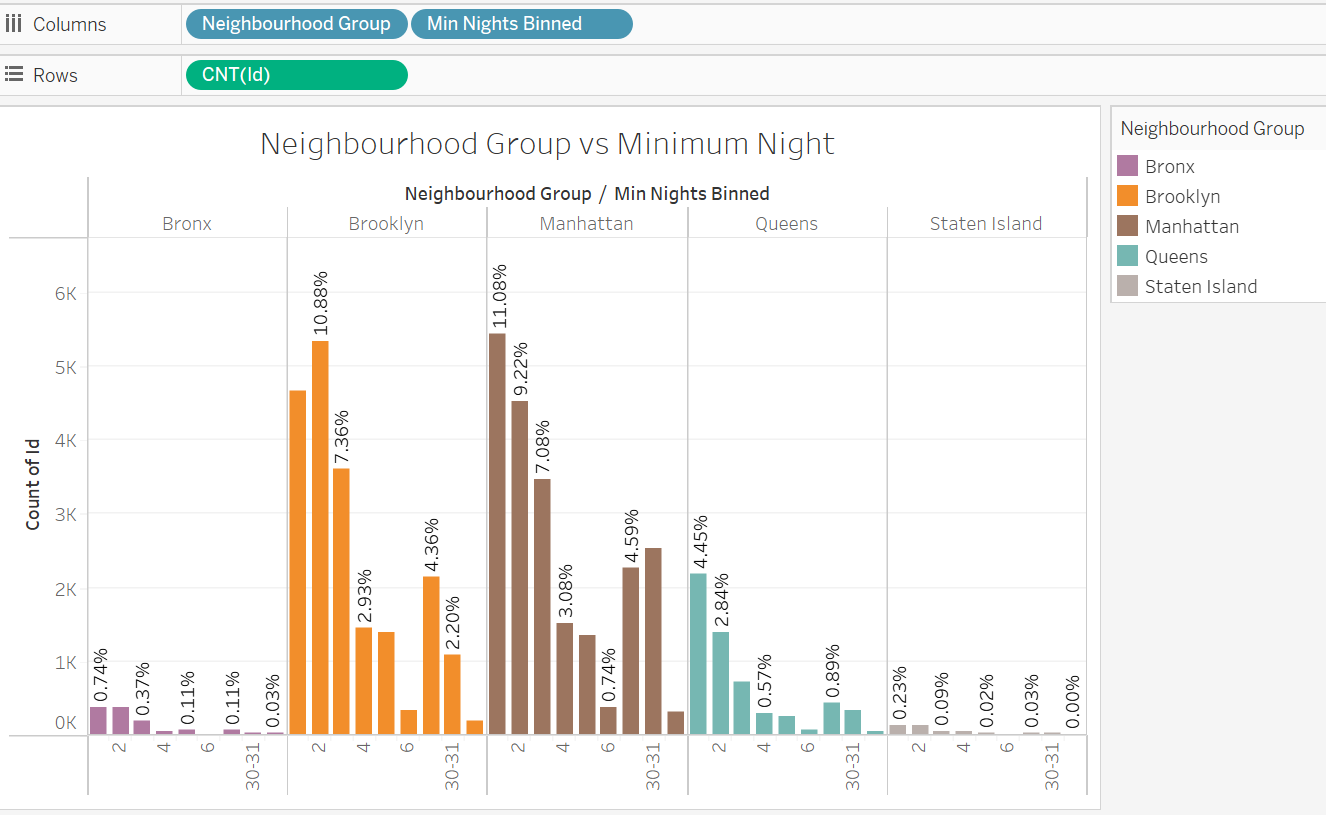
1. Customer Preferences of Properties in NYC w.r.t Price:

* The properties in Manhattan are the most expensive than any other area.
* Manhattan has the highest contribution of ‘Entire home/apt’ compared to the overall contribution of ‘Entire home/apt’.
* Queens has a higher contribution of ‘Private room’ compared to the overall contribution of ‘Private room’.



1. Preferences of Neighborhood Group w.r.t Minimum Nights:

* 20% of the bookings are made either in Manhattan or Brooklyn.
* As Manhattan & Brooklyn are expensive areas, the majority of the people prefer staying less than 3 nights.
* Staten Island is still a developing place while Bronx is the poorest borough, only 10% of the people prefer renting out.



***Step 4: Presentation:***

* Made the presentation adhering to best practices.
* Data Analysis Managers & Lead Data Analyst are our audience for Presentation 1.
* Added recommendations for the respective departments.