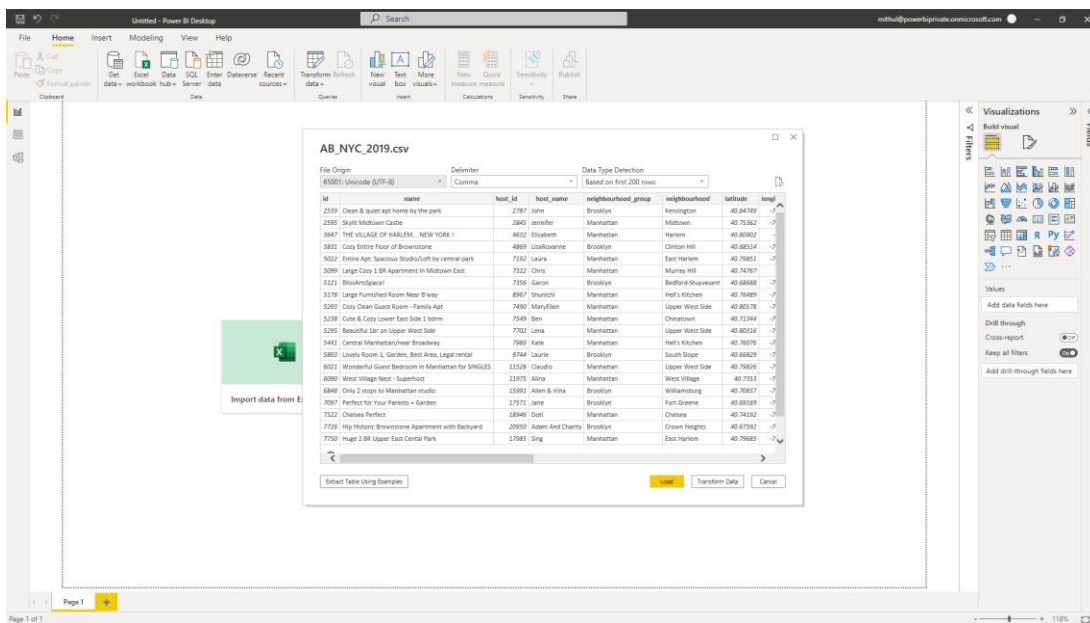


Methodology for Airbnb analysis

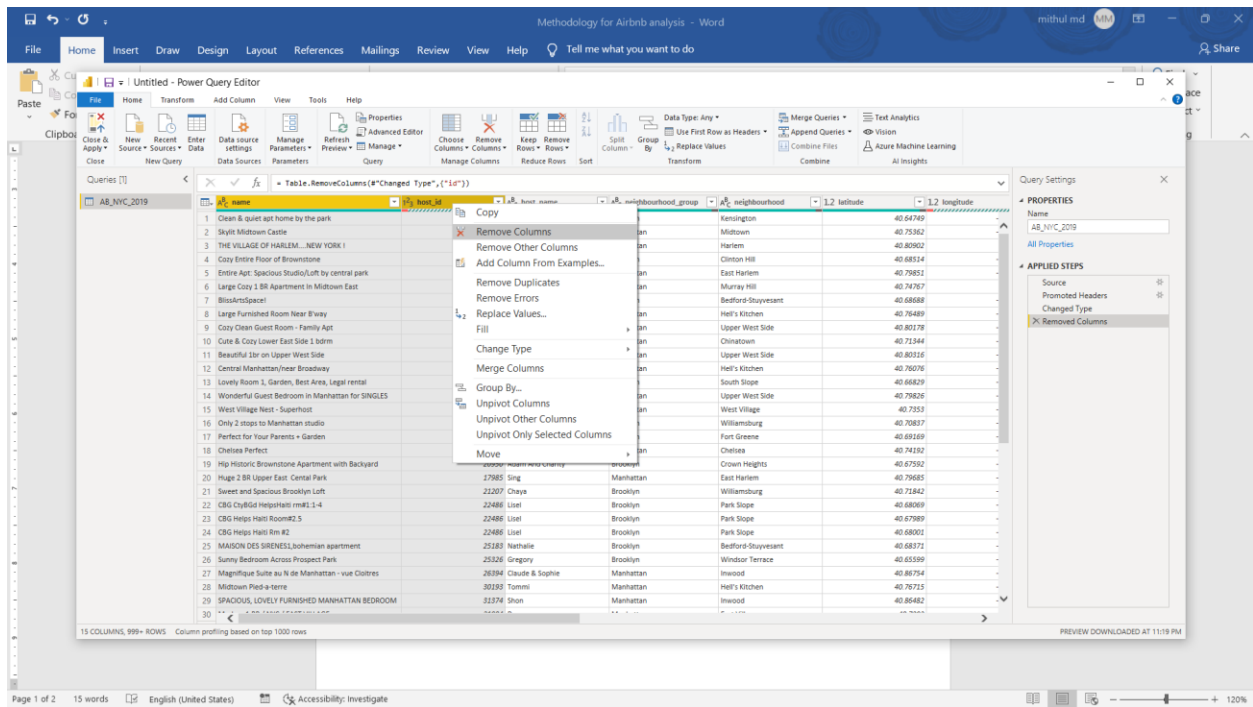
❖ Story boarding

- Got familiarized with the fields available in the dataset
- Made a rough note on to which of the fields to be targeted
- Made a list of questions to be answered through analyzing the dataset

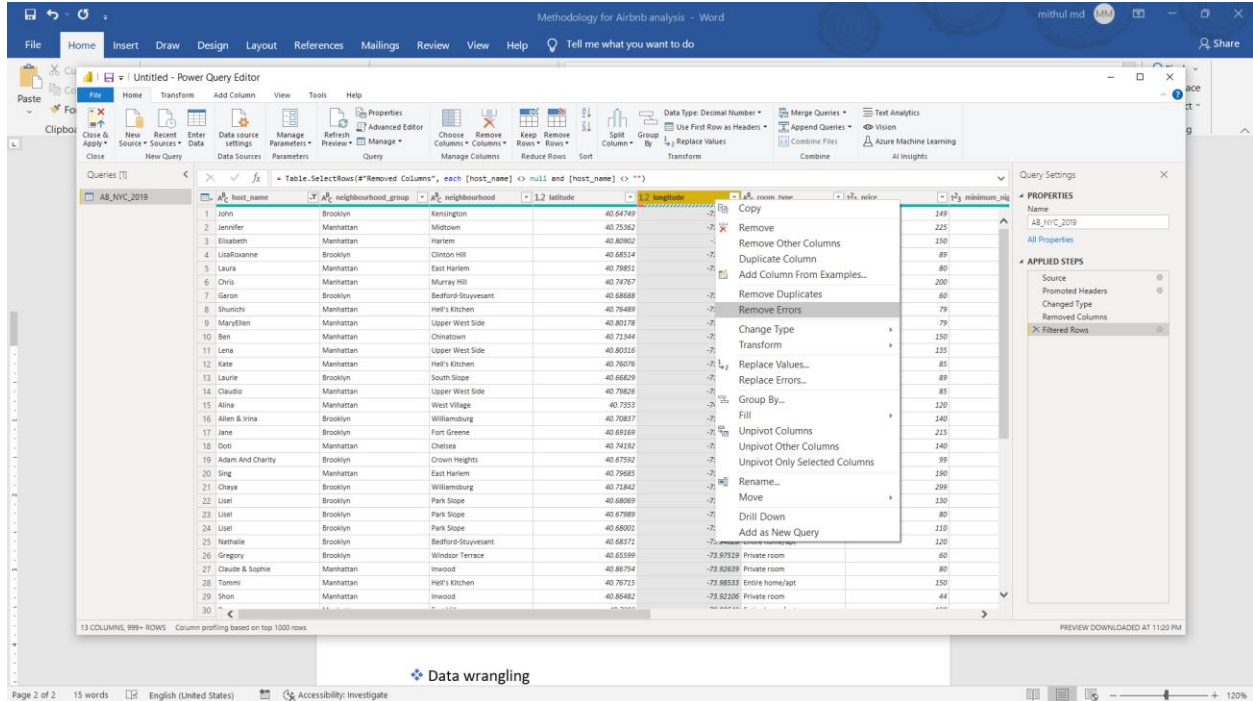
❖ Data preparation



- The downloaded data has been transformed and loaded in PowerBI.



- Removed columns that aren't contributing towards analysis for gaining insights. Some of the columns were Host id, Name.



- Removed Errors from the columns of the dataset.

Methodology for Airbnb analysis - Word

File Home Insert Draw Design Layout References Mailings Review View Help Tell me what you want to do

Clipboard

Power Query Editor

Table.SelectRows(*Removed Errors*, each [last_review] < null and [last_review] < "")

13 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

Data analysis

Page 3 of 3 15 words English (United States) Accessibility: Investigate

- Properties with availability as '0' has been filtered as they are assumed as not available for stay.
- The null values from other columns have been filtered.

❖ Data wrangling

Final - Power BI Desktop

File Home Table tools Column tools

Clipboard

Power Query Editor

Table.SelectRows(*Removed Errors*, each [availability_365] < null and [availability_365] < "")

13 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

Table AB_NYC_2019 (06/07 rows) Column price (564 distinct values)

Filter Rows

Apply one or more filter conditions to the rows in this table.

Basic Advanced

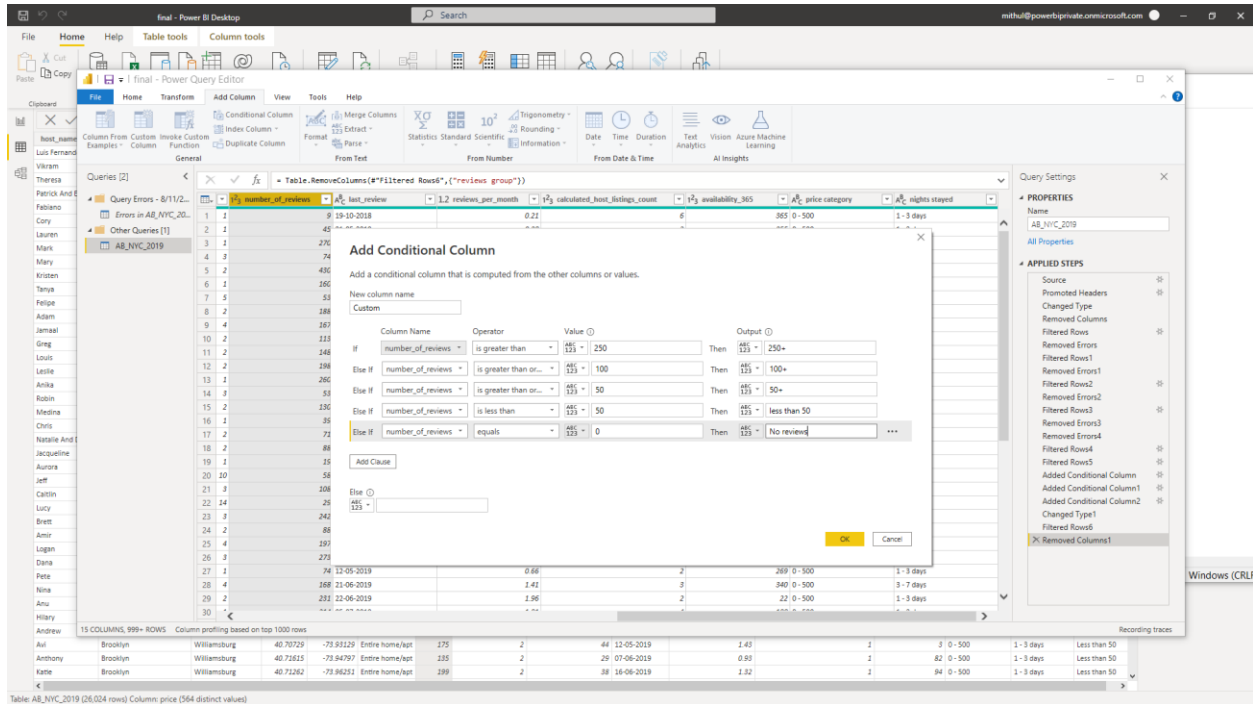
Keep rows where 'minimum_nights'

is less than or equal to 500

And Or Enter or select a value

Page 3 of 3 15 words English (United States) Accessibility: Investigate

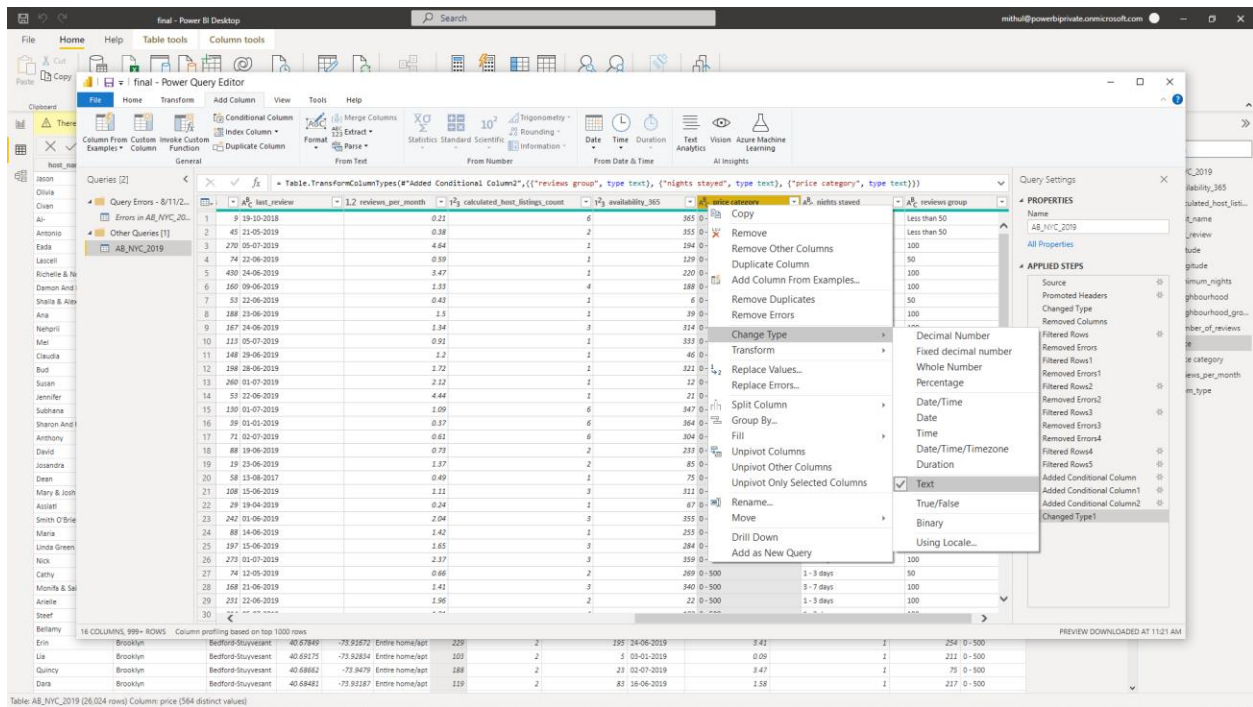
- Outliers have been filtered from the min_nights column.



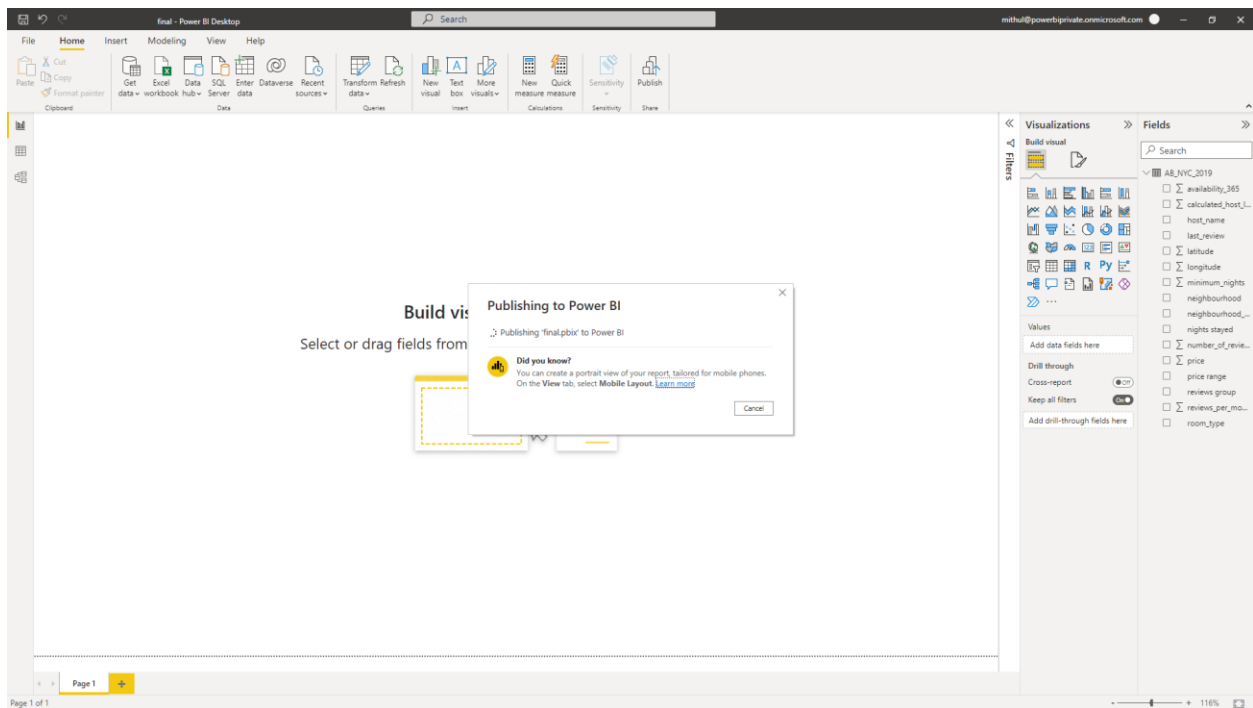
- Bins have been created for nights stayed, price category, reviews.
- Conditional columns were added for binning.

Bins created are as following:

- Nights stayed
 - 1-3
 - 3-7
 - 7-15
 - 15-30
 - 30+
- Price category
 - 0-500
 - 500-1000
 - 1000-2000
 - 2000-5000
 - 5k plus
- No. of reviews
 - 50+
 - 100+
 - 250+
 - less than 50
 - 0

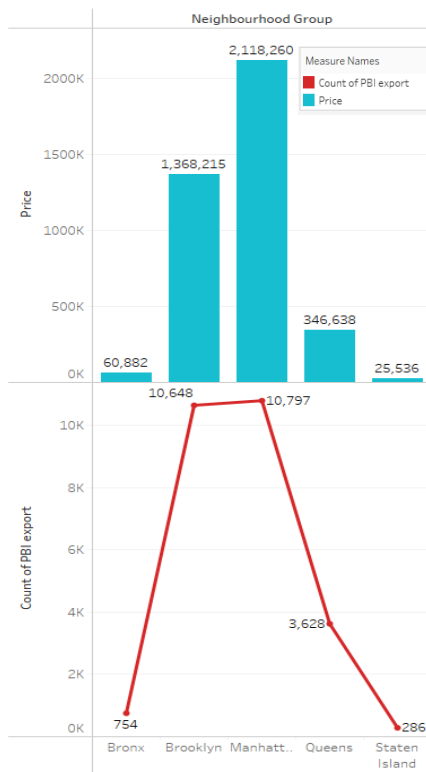


- Changed data types of the created Binned columns.

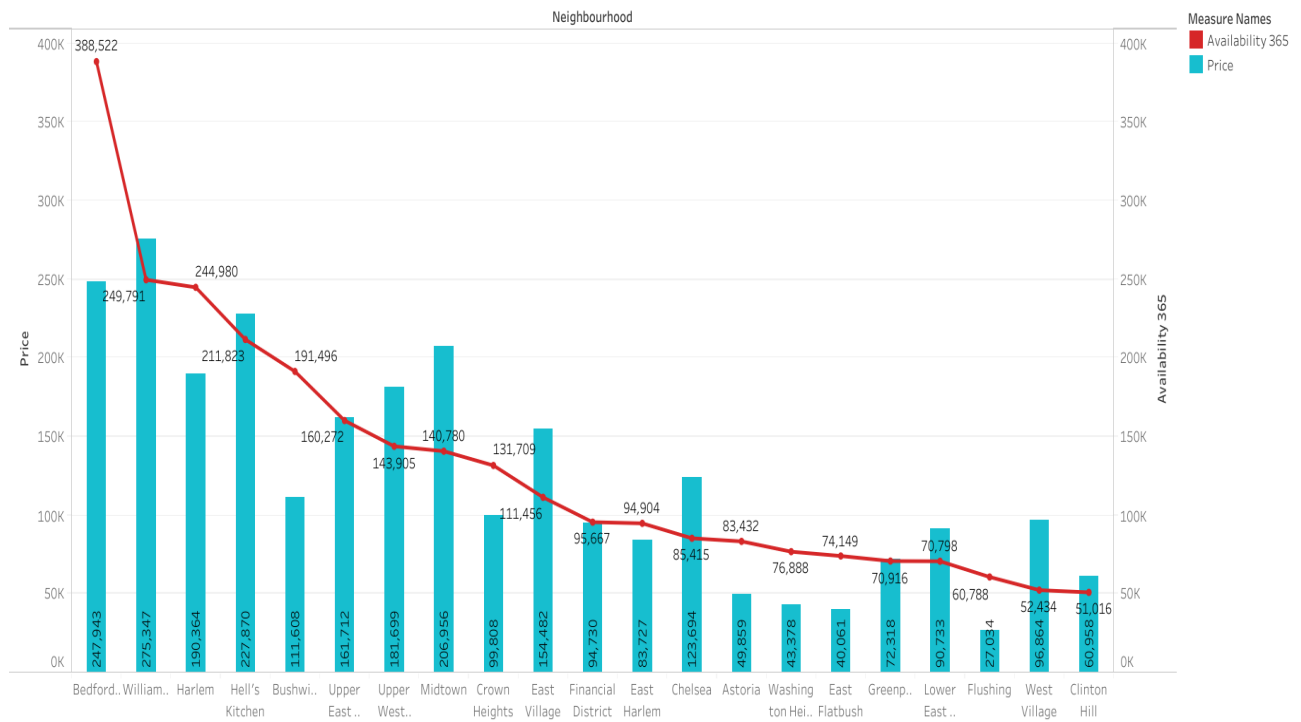


- Published the cleaned dataset to PowerBI and that has been used for the forth coming analysis.
- Exported the data to tableau through Excel.

❖ Data analysis

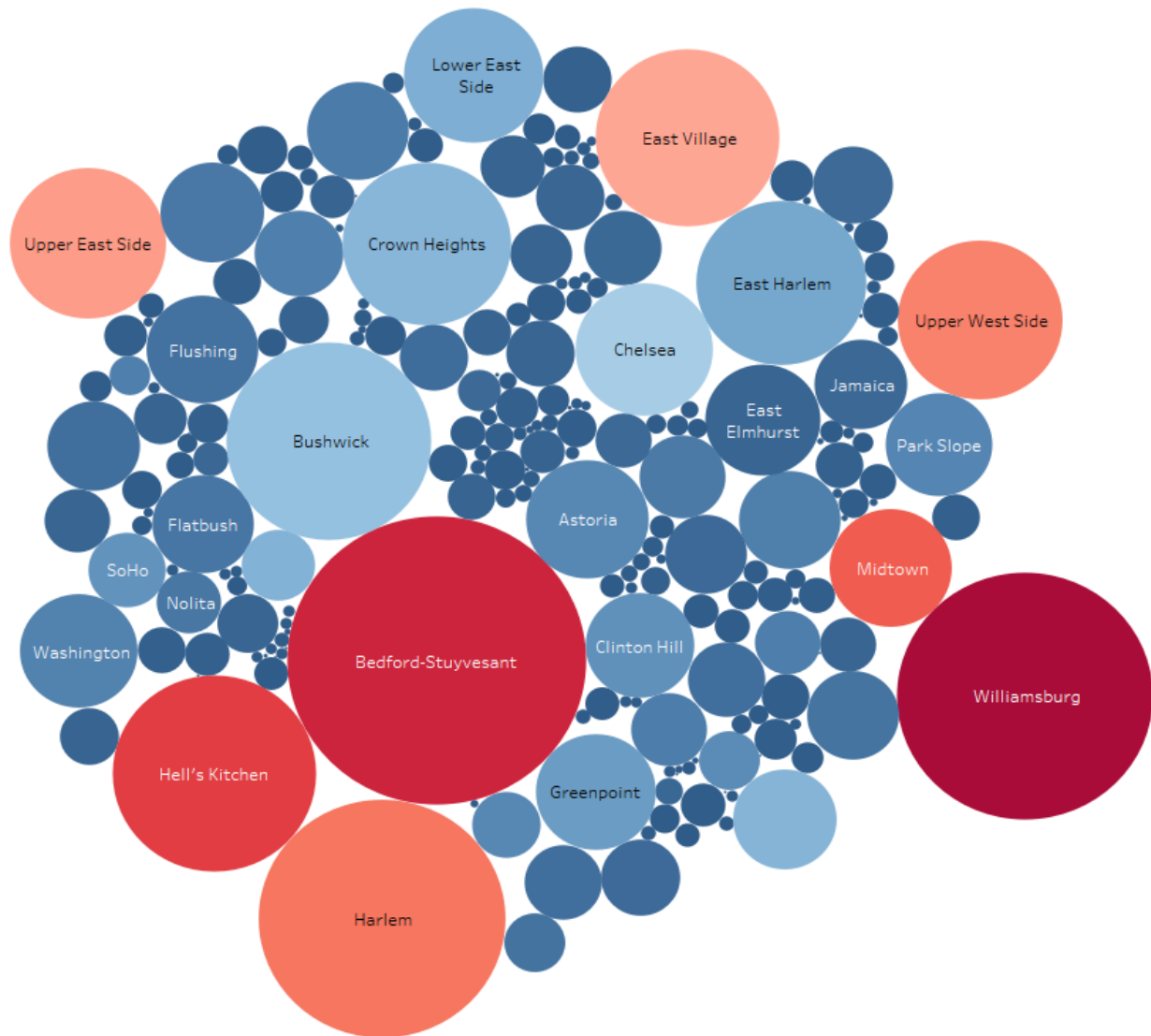


- Manhattan is the expensive neighborhood groups for homestay
- Bronx is the least expensive but also has more visitors than Staten Island
- Staten Island known for its natural spaces has availability of rooms mostly in all seasons and is less populated but has comparatively a less no. of visitors.



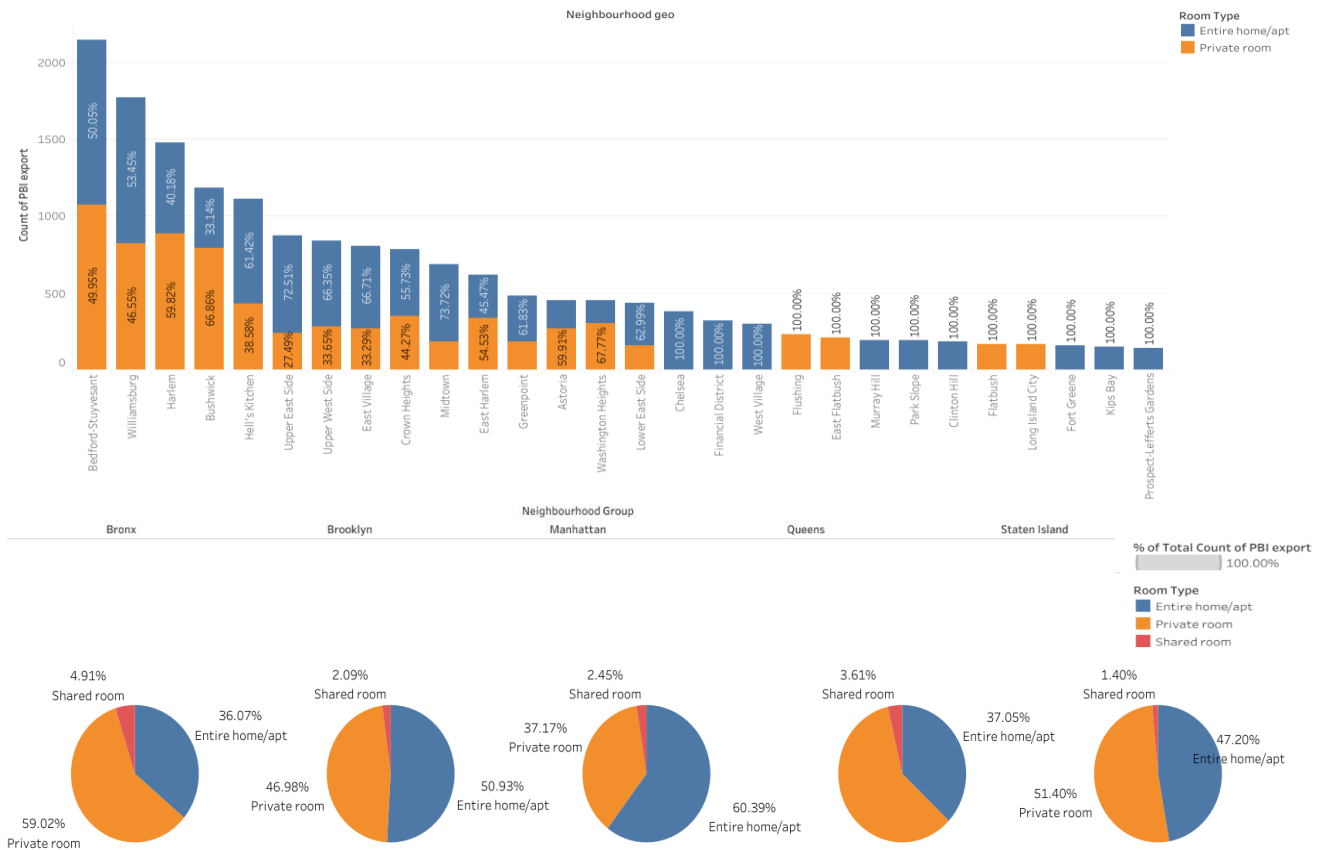
This infers the following points:

- There is a high availability of properties in Bedford , Williamsburg, Harlem, hell's kitchen.
- These places are usually busy as there is a high revenue inflow.
- Visitors prefer to be near the city center because of its vast exposure to various topics.



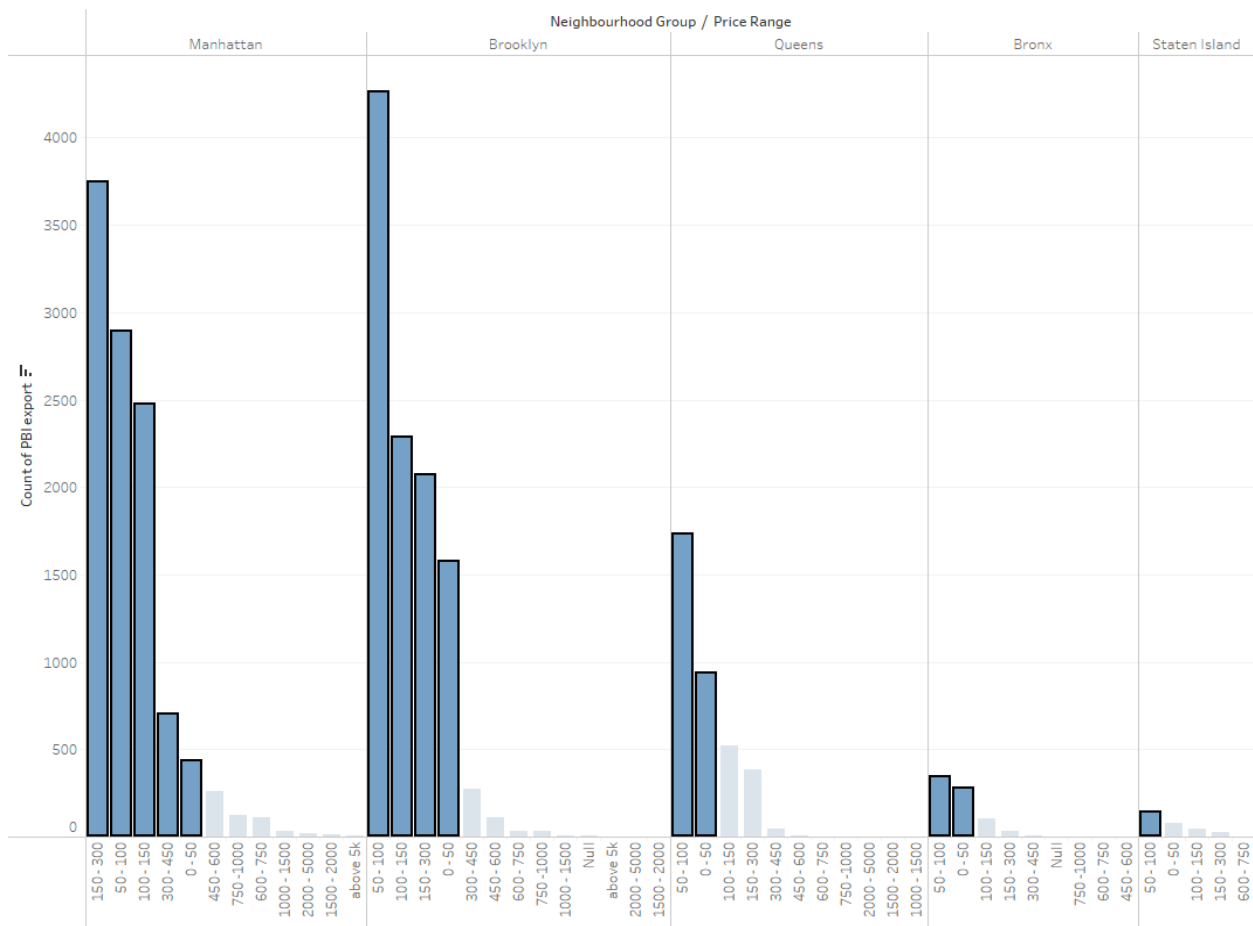
This visual infers the following:

- Popular neighborhoods have more reviews than the rest.
- These neighborhoods are recommended for an affordable stay and are indicating good revenue inflow.
- Bedford, Williamsburg, Harlem hold the highest in the list of no. of visitors.



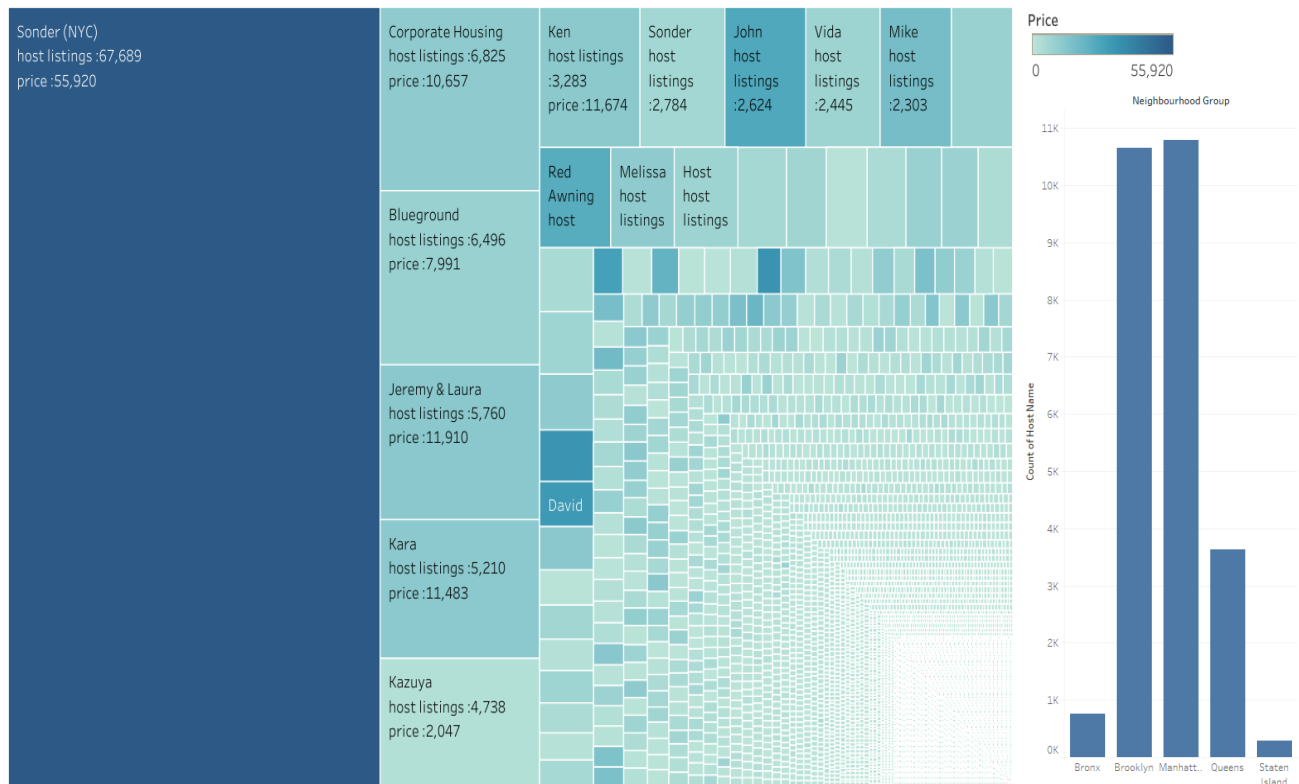
This visual derives the following:

- Entire home/Apt is most preferred in all the neighborhoods. Bedford, Williamsburg, Harlem having more hosting have more visitors as well.
- Manhattan being the costliest, has a high preference on Entire Home/Apt.
- Shared rooms are the least preferred in common.
- Brooklyn having a lesser no. of shared rooms also is the least expensive among all the boroughs.



This graph shows the following:

- Brooklyn and Manhattan have broad range of visitors.
- Mid-price range is preferred in Manhattan and is also the case that the place has more no. of costly properties for stay compared to others in the group.
- On common \$50-\$300 is the preferred price range for the lavish boroughs. While least priced-\$150 is preferred for stays in Queens, Bronx, Staten Island.



This states the following:

- Sonder (NYC) undertakes broad range of properties only Manhattan.
- Most of the properties are hosted by housing corporations/Realtors and are hosted in Manhattan and as the price is comparatively high, the inflow of revenue is high as well from here.
- This is also because Manhattan has high no. of properties hosted.

❖ Conclusion

- The principles of data storytelling have been followed throughout the presentation.
- Visuals were shared along with the inferences
- Recommendations were listed in the end of the presentation

