

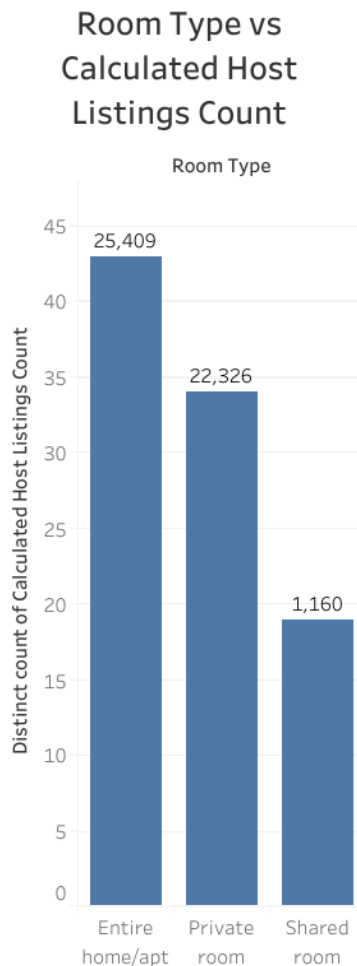
# Methodology Document

## Step 1: Storyboarding

- Went through the data to get familiarized with it and noted down important fields
- Made a mind map of the various slides of the presentation
- Made a rough template based on this mind map

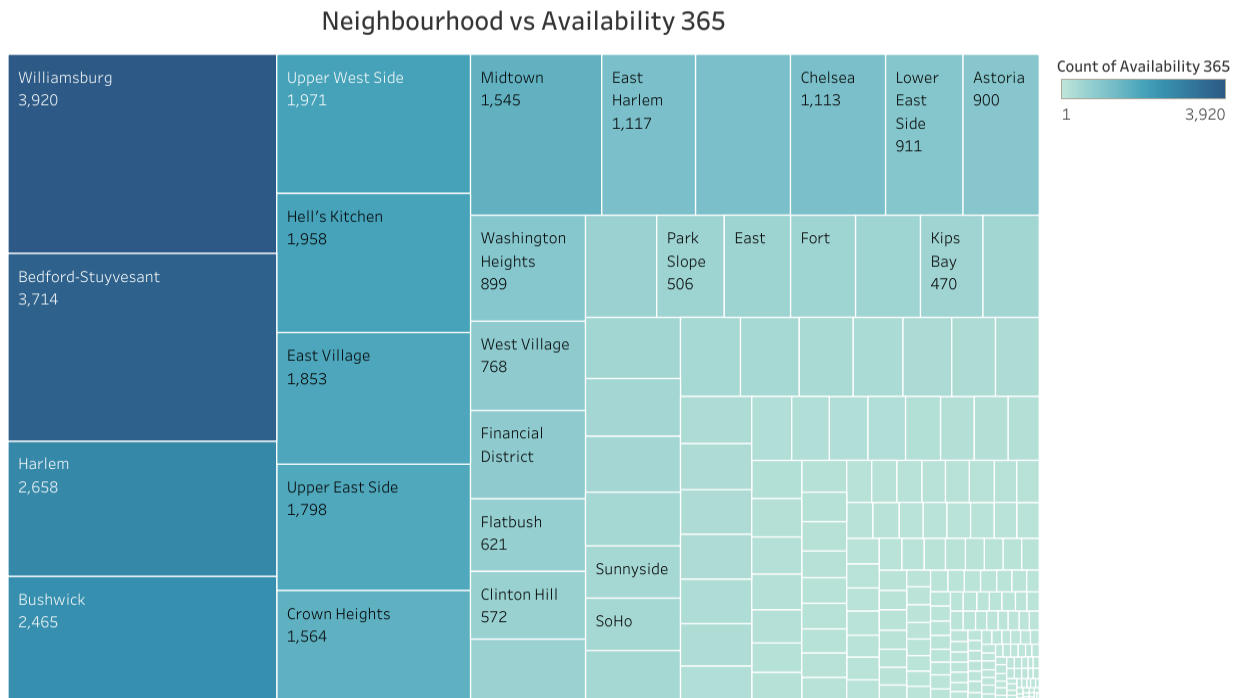
## Step 2: Data Wrangling

- Did univariate analysis using Tableau on the fields to see their distributions, the unique values in a field, the missing values and to check for outliers if any
- Since price has outliers, used median instead of mean as the measure for price
- Host Listings count is maximum for Entire Home/Apartment and is very small for Private and Shared Rooms.

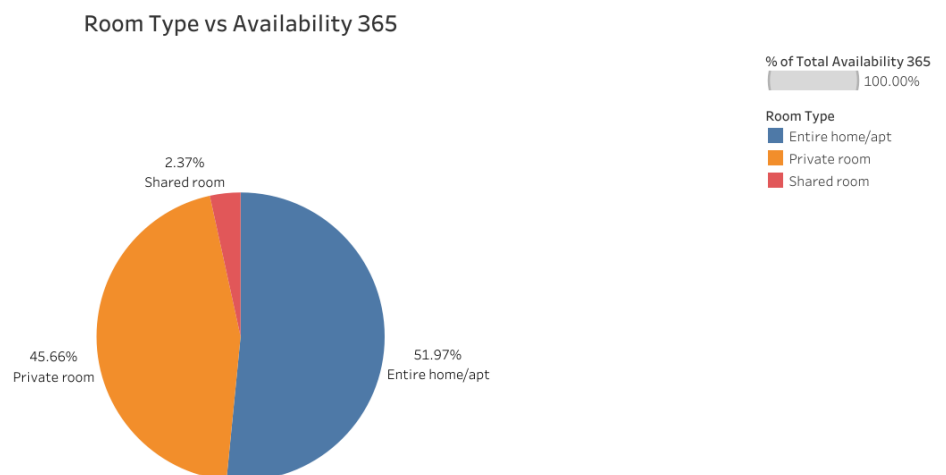


### Step 3: Data Analysis

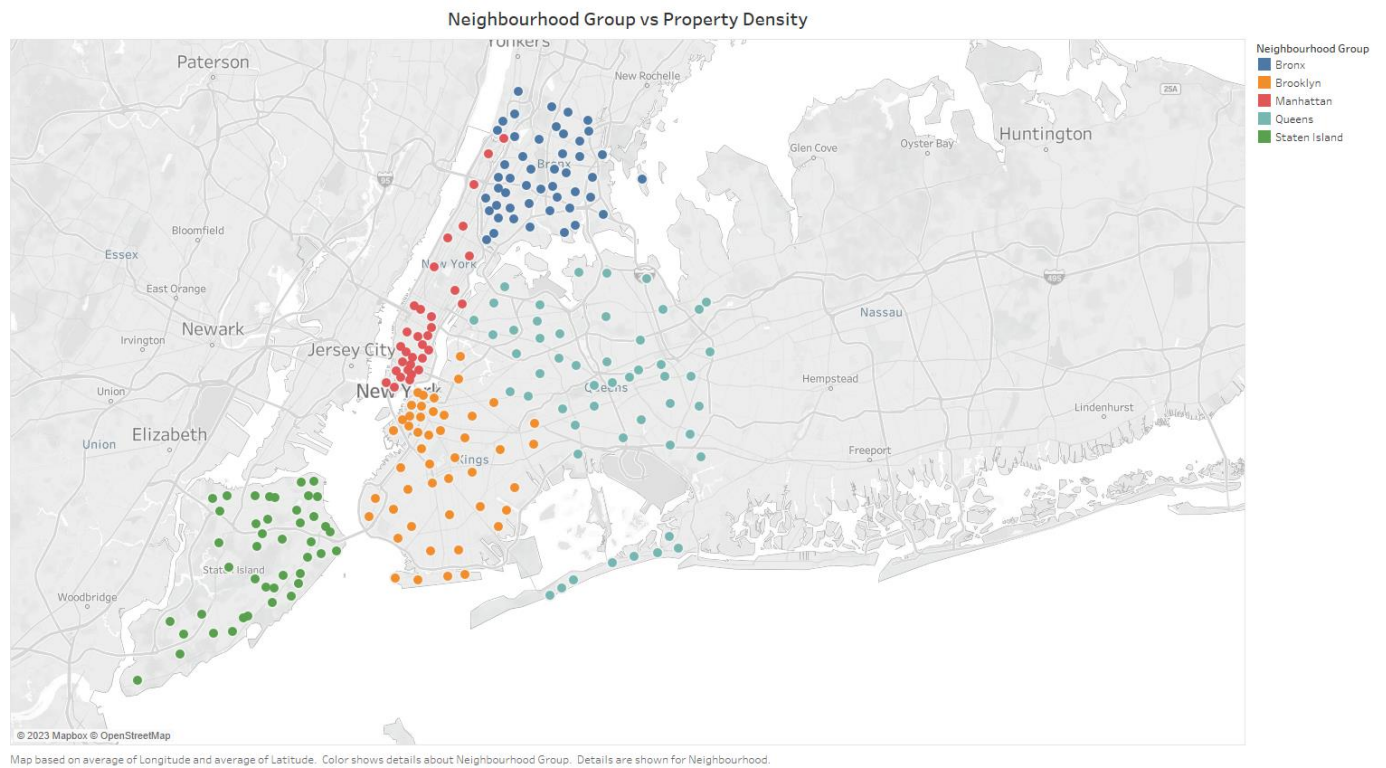
- Checked Neighborhood distribution based on availability



- Checked Room Type distribution based on availability



- Neighborhood Group based on property density



- Through these clearly:
  - Most of the hosts prefer to list their Entire home/Apartment.
  - Williamsburg and Bedford-Stuyvesant neighborhood have the highest availability of the rooms throughout the year.
  - Most of the guest prefer Entire home apartment (51.97%).
  - The density of listed properties in Brooklyn and Queens is quite low so we can focus more on acquiring more properties in these two Neighborhood Groups.

#### Step 4: Presentation

- Made the presentation adhering to best practices and pyramid principle
- Added recommendations for the respective departments