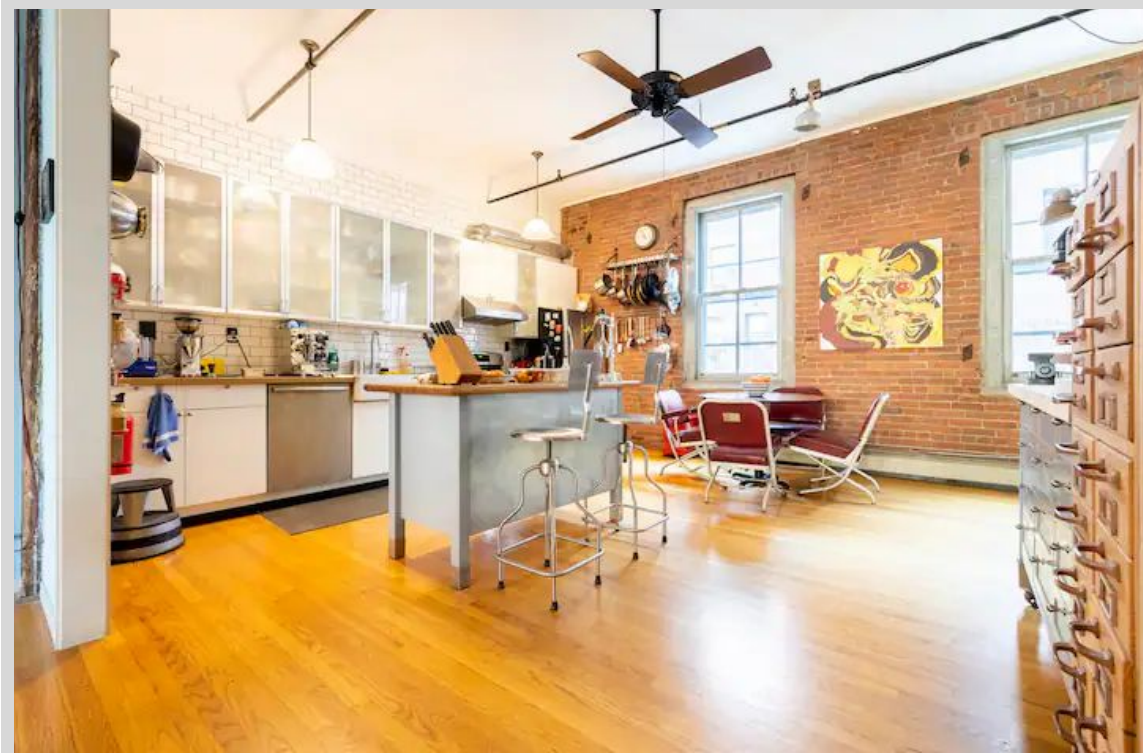# airbnb

Airbnb in New York City

Tableau Data Visualization and Dashboards

By Thanh Nguyen
Data from Inside Airbnb Public Database

Project Flow

1. Context & Data Connection

- Assuming being hired by a competitor company to provide key competitive intelligence in the New York city rental market of Airbnb.
- Historical dataset provided from the public database Inside Airbnb
- Connect the "airbnb-nyc.xlsx file"
- Detect different data types and variables

2. EDA & Data Cleaning

- Treating missing values, duplicates by visualization
- Detect outliers and mismatch by visualization and maps
- Explore variable relationships and patterns by visualization
- EDA and data cleaning by Tableau / Excel
- Export clean data under name "airbnb-nyc-clean".xlsx

3. Insights & Dashboard

To answer these business questions:
- **Q1:** How are Airbnb listings are distributed in NYC?
- **Q2:** How does neighborhood (location) impact Airbnb listings in NYC?
- **Q3:** Which type of property is most preferred by guests in NYC Airbnbs?
- **Q4:** What is the customer preference when choosing an Airbnb in NYC?
- **Q5:** What are some characteristics among NYC's Airbnb hosts?

1. Airbnb Overview

A web platform where users/host can rent out their space to host travelers and vice versa:

Network Economies

- Two-sided online marketplace
- Plenty of supply (marketplace liquidity)
- Niche: Millennials wanted a greater sense of adventure at an affordable cost

Double Commissions & Fee

- Hosts make money when hosting, pay a commission to Airbnb
- Travelers/guests pay service fee

Share Experience & Culture

- Local Connection To The City
- Sell experience, not physical rooms
- Multiple customers touchpoint via their own community

![airbnb logo]

Consultancy Brief

We're hired by a competitor company to provide key competitive insights into Airbnb's market in the New York City area.
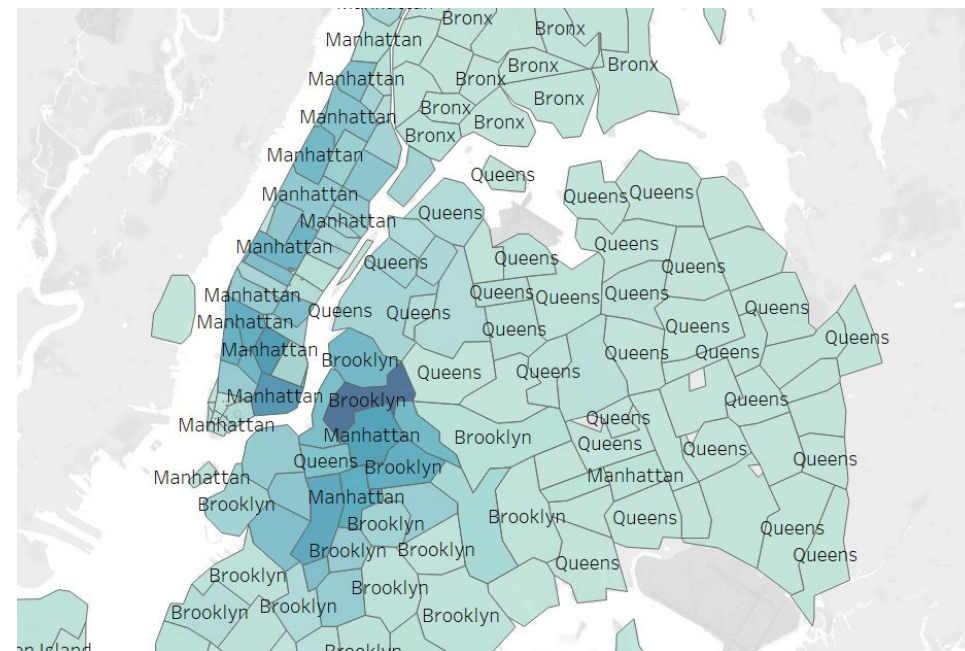
Key questions:

- **Q1: How are Airbnb listings are distributed in NYC?** *- To provide a high-level view*
- **Q2: How does neighborhood (location) impact Airbnb listings in NYC?** *- Location Analysis*
- **Q3: Which type of property is most preferred by guests in NYC Airbnbs?** *- Property Type Analysys*
- **Q4: What is the customer preference when choosing an Airbnb room in NYC?** *- Room Type Analysis*
- **Q5: What are some characteristics among NYC's Airbnb hosts?** *- Host Analysis*

2. EDA & Data Cleaning

- **Treating missing values, duplicates by visualization:** Using Tableau visualization to explore the data, cleaning data using Excel. Cleaned results are saved to a new Excel file.
- Scope of work (1):
  - *Impute missing values in 'Host Since' with existing data (3 rows)*
  - *Drop rows with missing values in zipcode (134 rows). Fix funky zipcode outside of NYC (6 rows). Fix mismatching zipcode vs. neighborhood (67 rows).*

EDA-MissingValues-Chart

|	
Count of Host Id	30,478
Count of Host Since	30,475
Count of Zipcode	30,344
Count of Beds	30,393
Count of Number of Records	30,478
Count of Number Of Reviews	30,478
Count of Price	30,478
Count of Review Scores Rating	22,155
Count of Review Scores Rating (bin)	22,155
Count of Name	30,478
Count of Neighborhood	30,478
Count of Property Type	30,475
Count of Room Type	30,478

2. EDA & Data Cleaning

- **Treating missing values, duplicates by visualization:** Using Tableau visualization to explore the data. Cleaning results are saved to a clean Excel file.
- Scope of work (2):
  - *Fill in null values in 'Property Type' (3 rows) with most common type: 'Apartment'*
  - *Fill in null values in 'Beds' (84 rows) with info from other columns: Apartment: Usually 1 bed - House: Usually 1 bed - Dorm: Usually 1 bed*
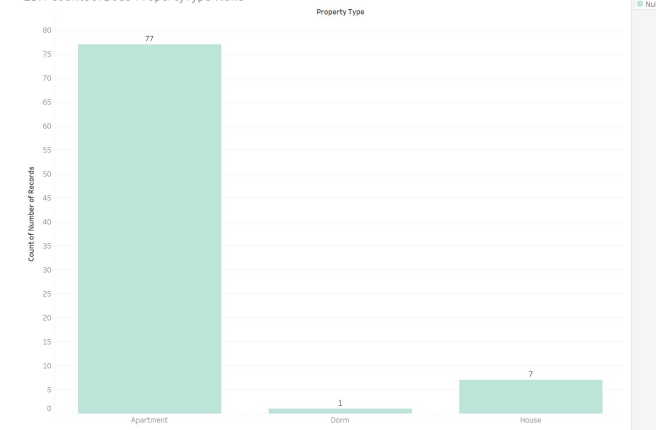

EDA-CountsOfBeds-PropertyType-Nulls

EDA-PropertyType-Nulls

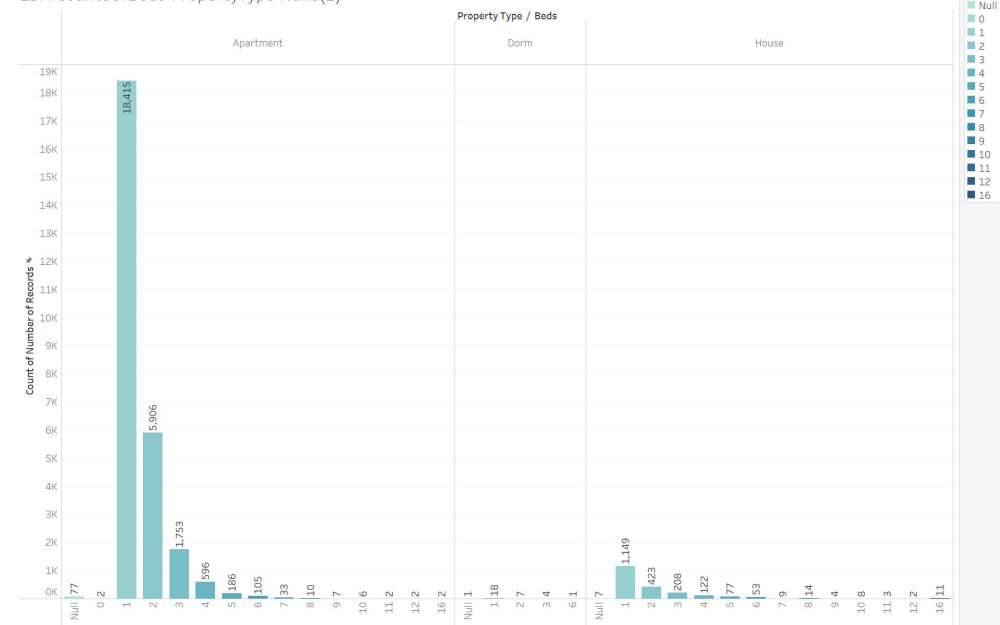| Property Type	Room Type	Name	Beds	Count of Number of Records
Null	Entire home/apt	Entire home/apt in New York	2.000	1.000
		Hip & Spacious 1BD in the LES	1.000	1.000
	Private room	Bright and Cheery Chelsea Studio	1.000	1.000

EDA-PropertyType-Chart
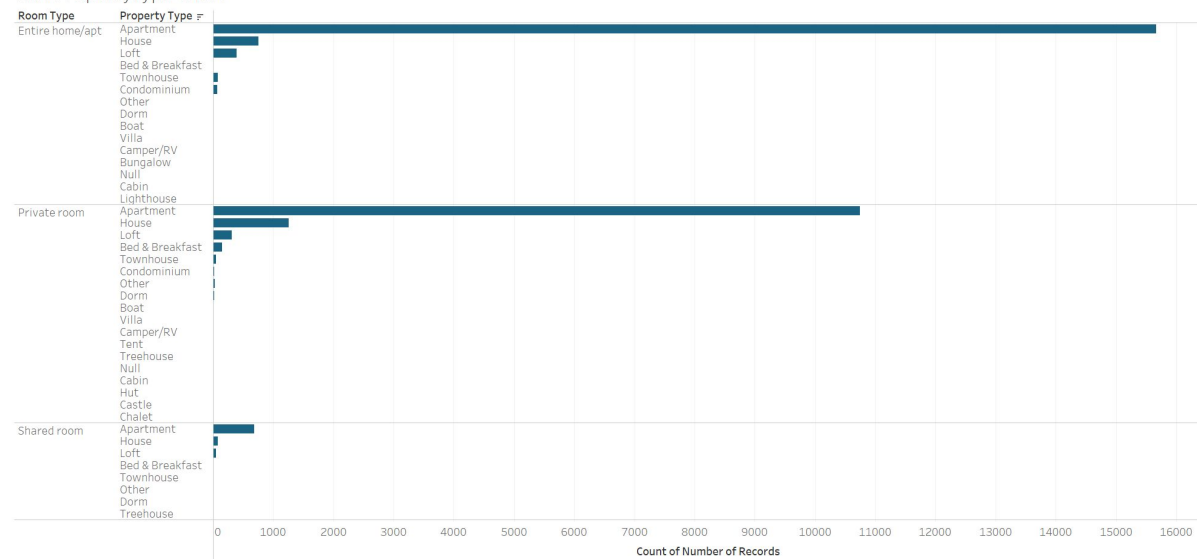
EDA-CountsOfBeds-PropertyType-Nulls(2)

2. EDA & Data Cleaning

- **Treating missing values, duplicates by visualization:** Using Tableau visualization to explore the data. Cleaning results are saved to a clean Excel file.
- Scope of work (3):
  - *Drop the Review Scores Rating (bin) column because it doesn't have much meanings in the dataset*
  - *We can't drop the rows with null values in 'Review Scores Rating', because there are more than 8.3K rows like that.*
  - *The majority of missing values are due to having zero 'Number of Reviews'. We'll just drop 509 over 30,478 rows for those having >= 1 number of reviews but having no reviews. The rest of missing values here, we choose to keep them them as 'Null' and add a 'Zero Review' clarification column with value 1 and 0 - 1 is when there is a missing value in 'Review Scores Rating' when Number of Reviews' = 0 and vice versa.*
 - *Drop all duplicates*

EDA-ReviewScoresRating-Nulls
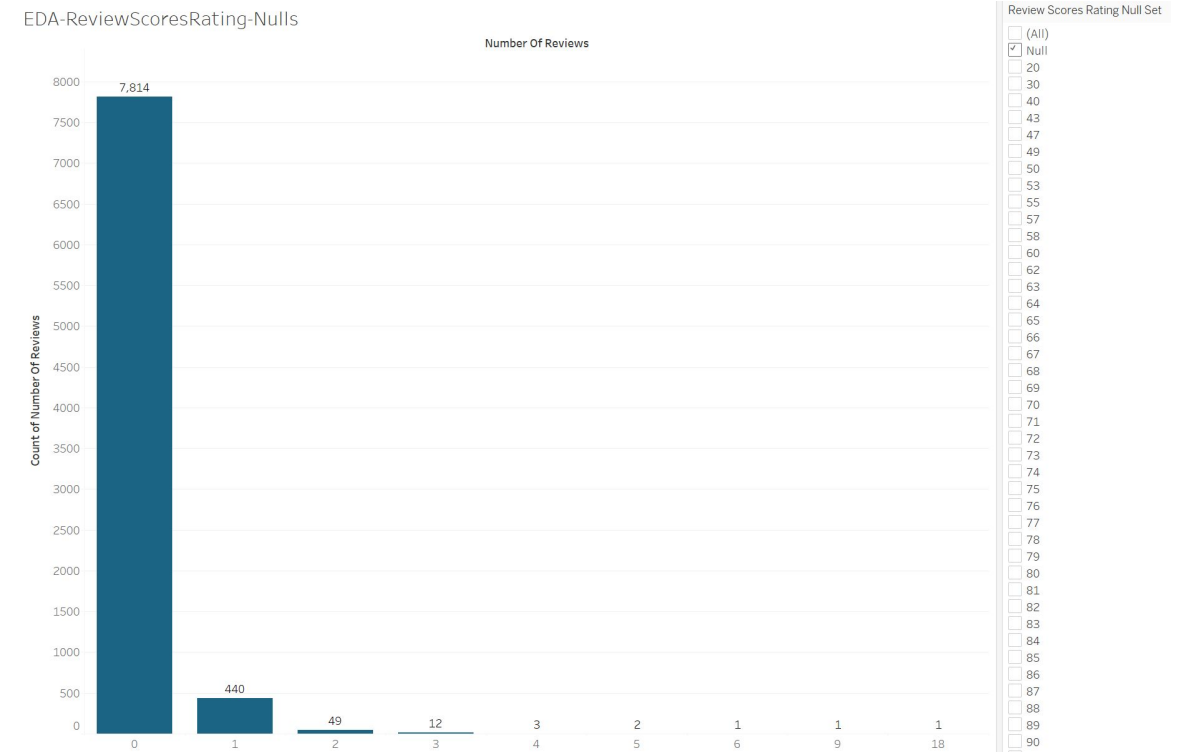
Number of Records	Number Of Reviews	Price	Review Scores Rating	Zero Review
1	39	549	96	0
1	0	199		1
1	4	149	100	0
1	9	250	100	0
1	80	90	94	0
1	95	270	90	0

2. EDA & Data Cleaning

- **Using pairplots to study the relationship between numeric variables:**
  - *There doesn't seem to have linear correlation between Beds , Price, Number of Reviews, Review Scores Rating. The relationships between the numerical variables somewhat resemble the log-linear trend line.*
  - *We removed an extreme outlier from the pairplot, and save the clean data results to a new Excel file named "airbnb-nyc-clean.xlxs".*
  - *The clean dataset contains 29,816 rows after the EDA & data cleaning process.*
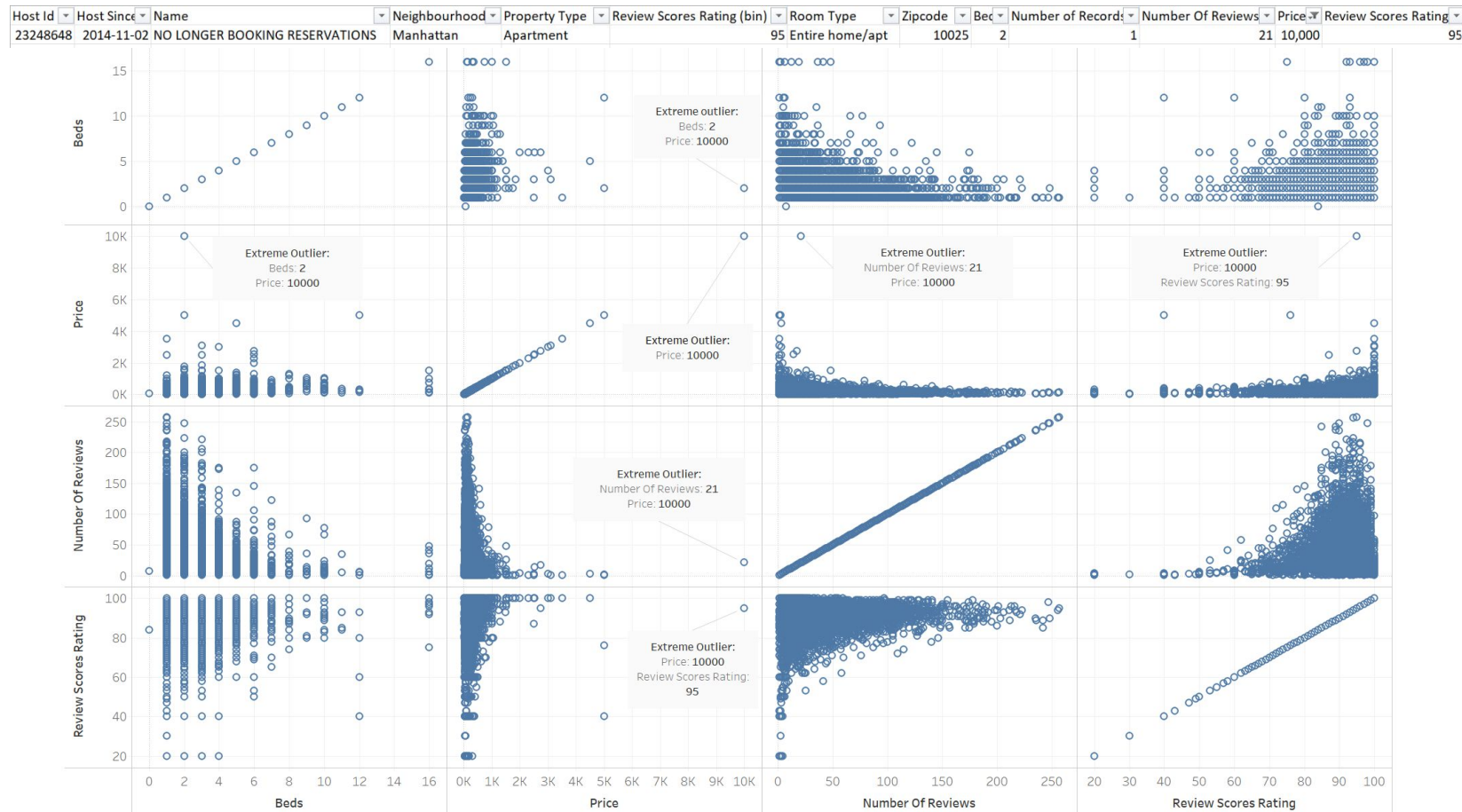
3. Insights & Dashboard

Question 1: How are Airbnb listings are distributed in NYC?

- There are 29,816 Airbnb listings from our dataset. Entire home/apartment accounts for 55,86%, seconded by Private Room (41.40%) and then Shared Room (2.75%).
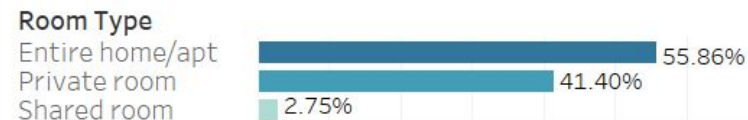- The four neighborhoods that Airbnb listings are located: Manhattan, Brooklyn, Queens, Staten Island

Total Listings: 29,816

Total Beds: 45,570

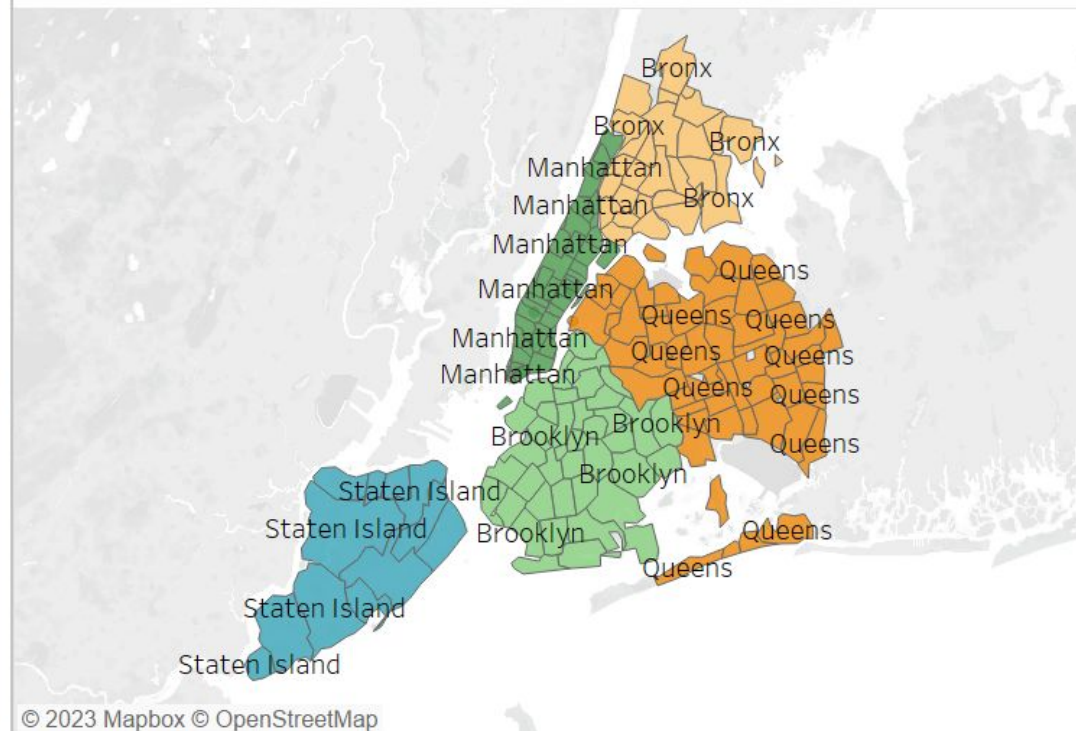**Total Hosts:** 23,957

Avg. Price Per Listing: $164

Listings by Room Type

Room Type
- Entire home/apt — 55.86%
- Private room — 41.40%
- Shared room — 2.75%

Room Type
- ☑ (All)
- ☑ Entire home/apt
- ☑ Private room
- ☑ Shared room

Neighborhood Overview

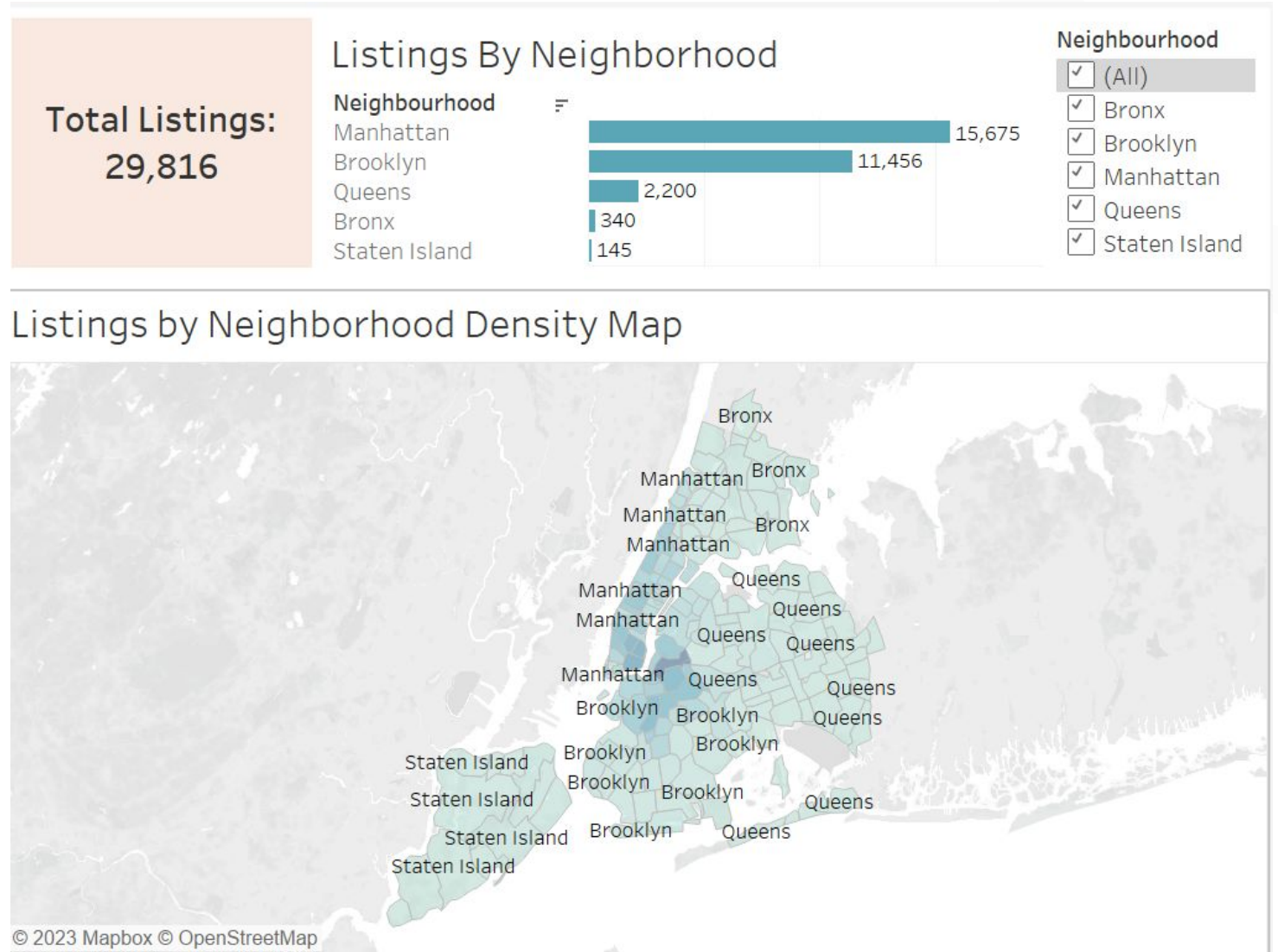© 2023 Mapbox © OpenStreetMap

3. Insights & Dashboard

Question 2: How does neighborhood (location) impact Airbnb listings in NYC?

- Manhattan and Brooklyn are the highest density locations for Airbnb listings in New York city. Especially in the hip Williamsburg neighborhood (Brooklyn), and Lower East Side of Manhattan (connected by a bridge).
- It is easy to relate because they are where the majority of attractions, activities, restaurants & bars and all the "actions" are located at.

Neighborhoods to Know in NYC

Listings By Neighborhood

Total Listings: 29,816

Neighbourhood	
Manhattan	15,675
Brooklyn	11,456
Queens	2,200
Bronx	340
Staten Island	145

Neighbourhood
- ✓ (All)
- ✓ Bronx
- ✓ Brooklyn
- ✓ Manhattan
- ✓ Queens
- ✓ Staten Island

Listings by Neighborhood Density Map

© 2023 Mapbox © OpenStreetMap

3. Insights & Dashboard

Question 2: How does neighborhood (location) impact Airbnb listings in NYC?

- Not only Manhattan's Airbnb locations are generally more dense, they are more expensive than of Brooklyn on average.
- Neighborhood are partially related to the clusters of Airbnb locations, on top of price.
- These findings may be indirectly justified by the fact that Manhattan is among the most expensive places in New York city and the city to rent/own a place with very limited space.
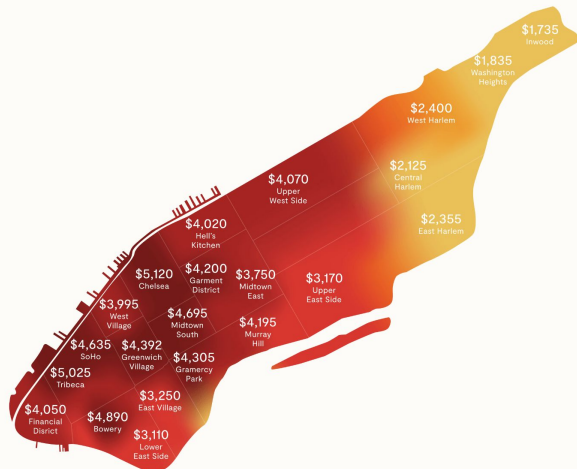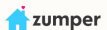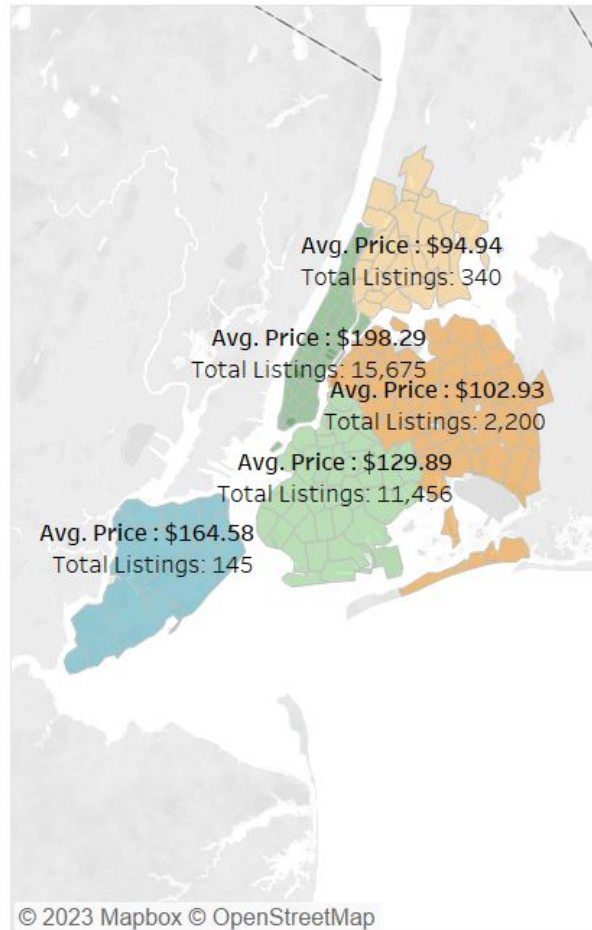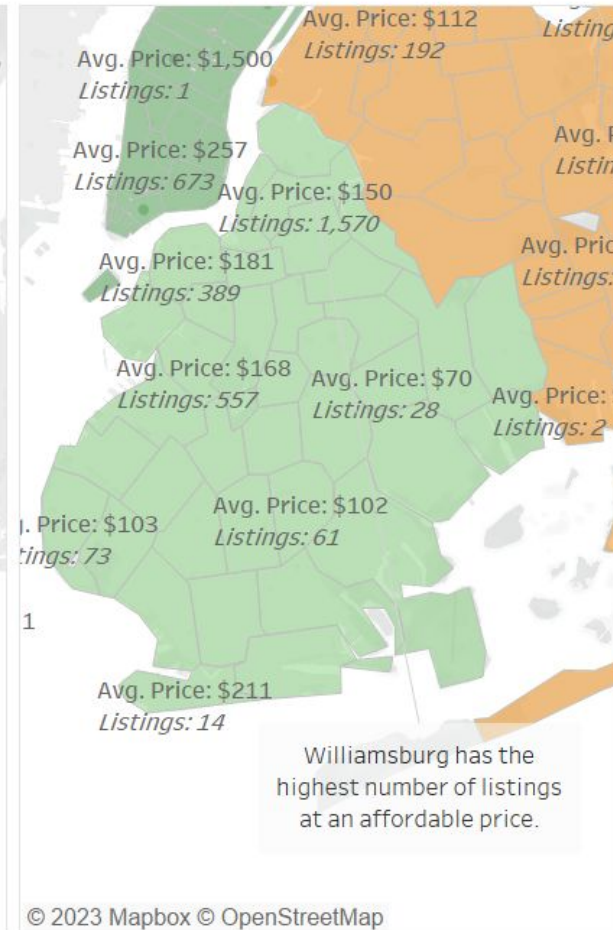

Listings by Neighborhood - Price per Neighborhood

Listings by Neighborhood - Price Per Zipcode

Neighbourhood
- Manhattan
- Brooklyn
- Queens
- Bronx
- Staten Island

Neighbourhood
(All) ▼

3. Insights & Dashboard

Question 3: Which type of property is most preferred by guests in NYC Airbnbs?

- The top 5 most common property types Airbnb guests like to stay in New York City are:
  - *1. Apartment*
  - *2. House*
  - *3. Loft*
  - *4. Bed & Breakfast*
  - *5. Condominium*
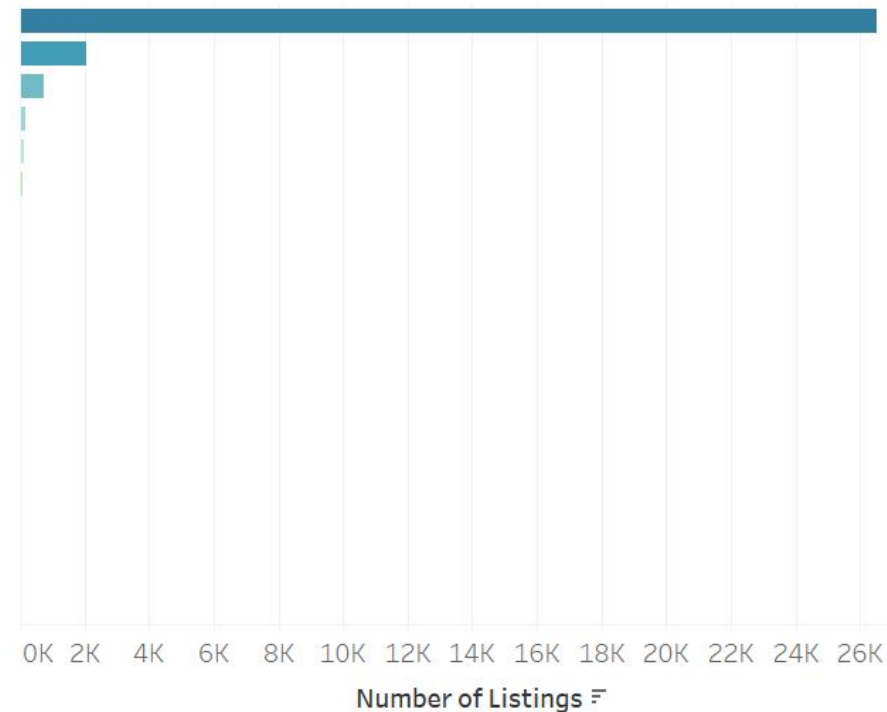
Total Listings: 29,816

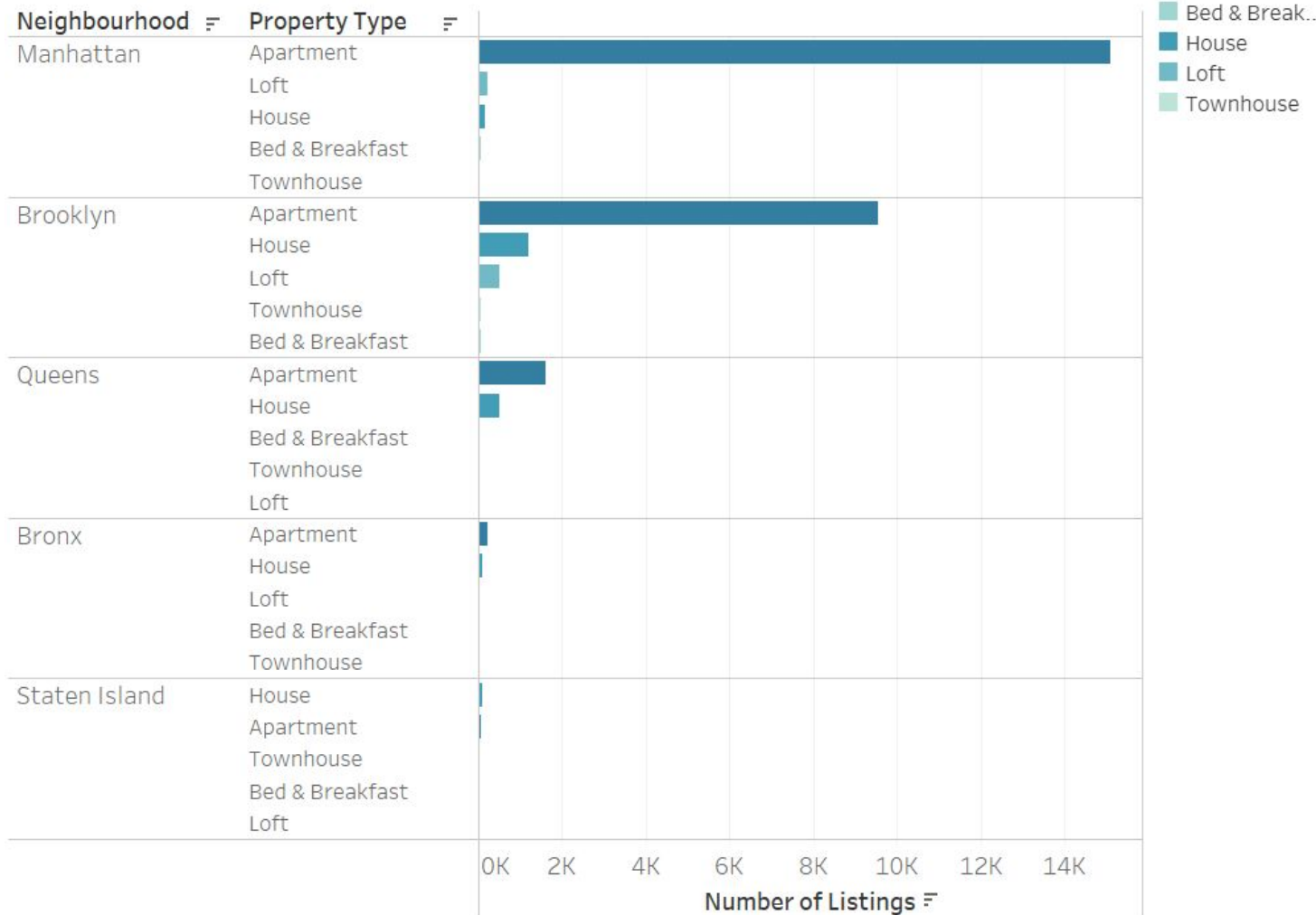Room Type
(All) ▾

Neighbourhood
(All) ▾

3. Insights & Dashboard

Question 3: Which type of property is most preferred by guests in NYC Airbnbs?

- Apartment is the clear winner across all the neighborhood.
- There's not much houses for Airbnb rentals in Manhattan which seems to make sense, since Manhattan usually tops the board for most expensive places to live. Plus, the proximity to famous attractions like Time Square, Empire State Building, Statue of Liberty, Randal Park and other museums.

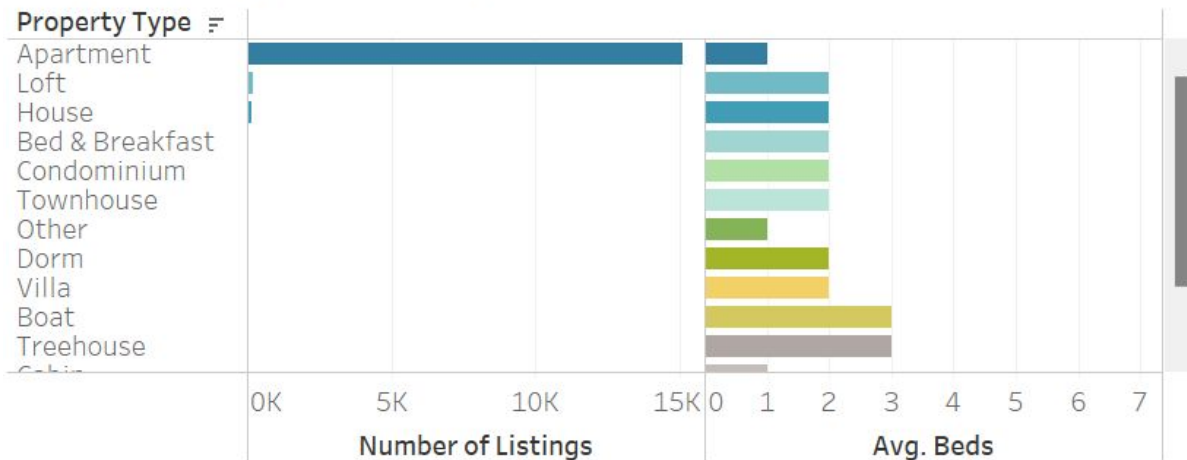## Top 5 Property Type by Neighborhood

Property Type
- Apartment
- Bed & Break..
- House
- Loft
- Townhouse

Neighbourhood	Property Type
Manhattan	Apartment
	Loft
	House
	Bed & Breakfast
	Townhouse
Brooklyn	Apartment
	House
	Loft
	Townhouse
	Bed & Breakfast
Queens	Apartment
	House
	Bed & Breakfast
	Townhouse
	Loft
Bronx	Apartment
	House
	Loft
	Bed & Breakfast
	Townhouse
Staten Island	House
	Apartment
	Townhouse
	Bed & Breakfast
	Loft

Number of Listings

3. Insights & Dashboard

Question 4: What is the customer preference when choosing an Airbnb room in NYC?

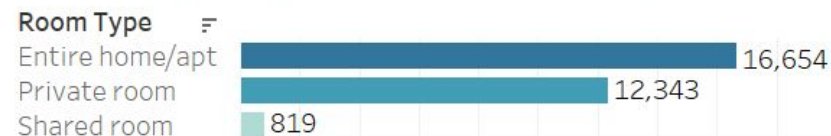- Customers show a preference for renting the entire place or a private room; all within the Apartment property type. It may depends on the number of guests, and number of beds accordingly.

Total Listings: 29,816

Room Type By Number of Listings

Room Type	
Entire home/apt	16,654
Private room	12,343
Shared room	819

Neighbourhood
- ☑ (All)
- ☑ Bronx
- ☑ Brooklyn
- ☑ Manhattan
- ☑ Queens
- ☑ Staten Island

Room Type vs. Property Type
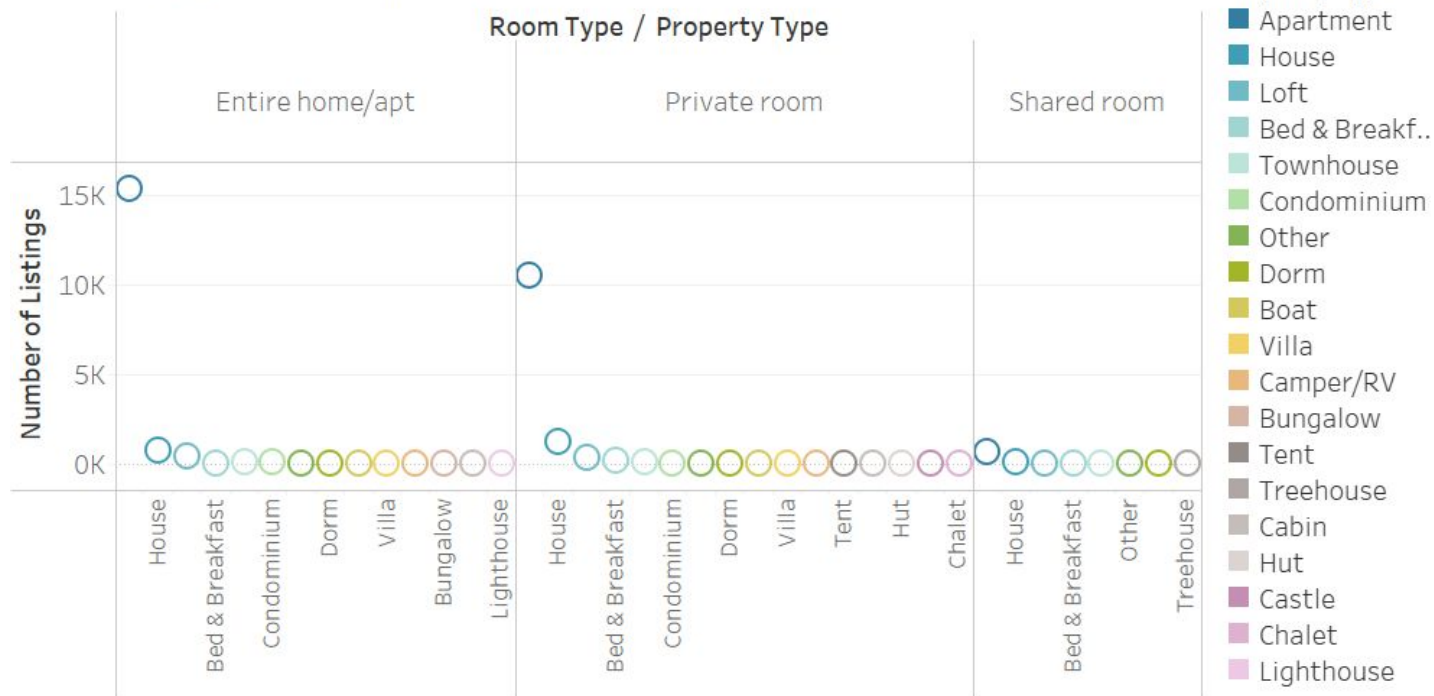
Room Type / Property Type

Entire home/apt   Private room   Shared room

Property Type
- ■ Apartment
- ■ House
- ■ Loft
- ■ Bed & Breakf..
- ■ Townhouse
- ■ Condominium
- ■ Other
- ■ Dorm
- ■ Boat
- ■ Villa
- ■ Camper/RV
- ■ Bungalow
- ■ Tent
- ■ Treehouse
- ■ Cabin
- ■ Hut
- ■ Castle
- ■ Chalet
- ■ Lighthouse

3. Insights & Dashboard

Question 4: What is the customer preference when choosing an Airbnb room in NYC?

- Another indicator that Manhattan seems to attract customers that are looking for more comfort (bigger place, more beds) with much higher price.
- Airbnb customers staying in Brooklyn are 50/50 between renting an entire place or just a room. That might come from families that are willing to pay a bit more for more bed; or solo-travellers looking for a private room with a budget.
- Staten Island is an interesting case with very little number of listings, yet stands out with the average price (and average bed) per Entire place. It is due to the hike-up of some places costing $500 - $5,000 for 3-16 beds.

Room Type by Neighborhood, Avg. Beds and Avg. Price
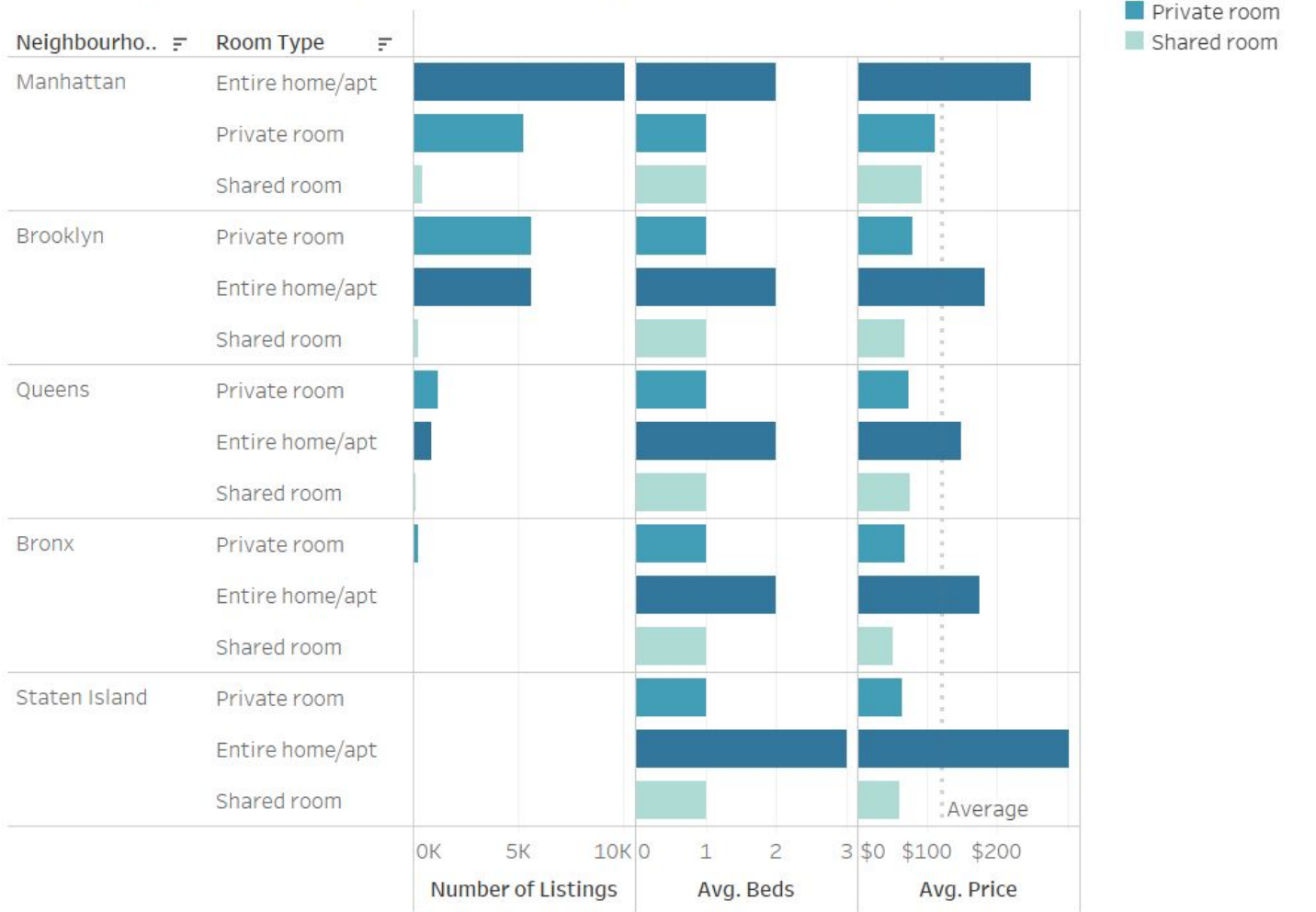
3. Insights & Dashboard

Question 5: What are some characteristics among NYC's Airbnb hosts?

- The average listing per host is only 1.24 within the whole New York city, although some hosts can have up to 32 listings per head!
- The biggest concentration of hosts in NYC is weighted in the renting the Entire place room type in Manhattan.

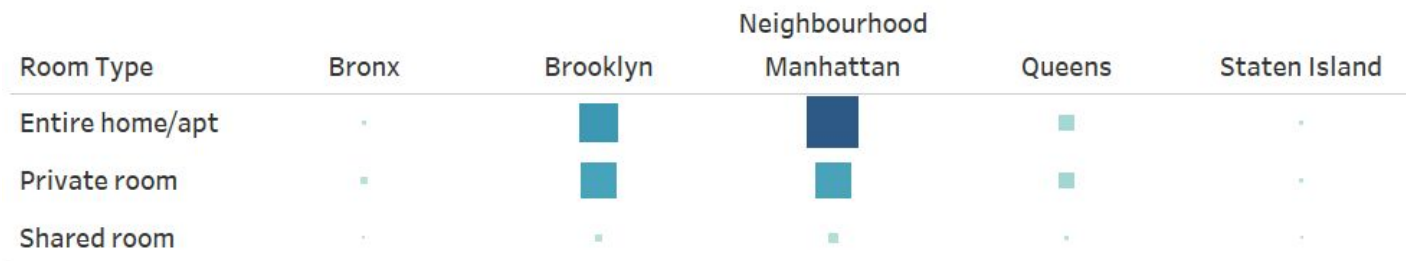Total Hosts: 23,957

Avg. Listing Per Host: 1.24

Neighbourhood: (All)

Room Type: (All)

Number of Hosts Per Room Type and Neighborhood

Room Type	Bronx	Brooklyn	Manhattan	Queens	Staten Island
Entire home/apt					
Private room					
Shared room					

Top 10 Hosts by Number of Listings

Host Id

| Host Id	Number of Listings
12122046	32
26377263	30
417504	28
35927005	25
20780082	21
3250450	17
1217923	15
4887492	14
1411399	

Number of Listings

3. Insights & Dashboard

Question 5: What are some characteristics among NYC's Airbnb hosts?

- Overall, there seems to be a big dip in the number of hosts that joins Airbnb since 2015. The breakdown per neighborhood shows the biggest decrease in Manhattan and Brooklyn.
- The changing pattern of listing prices regarding 2015 "new hosts" shows tightening of some forms, for prices at the upper limits.
- According to this source: *"Governor Cuomo mulls over signing into law another ban on full-apartment Airbnb rentals for less than 30 days"*. Regardless of the actual numbers, a report says more than 50% of listings in NYC in 2015 offered to book an entire apartment for less than 30 days.

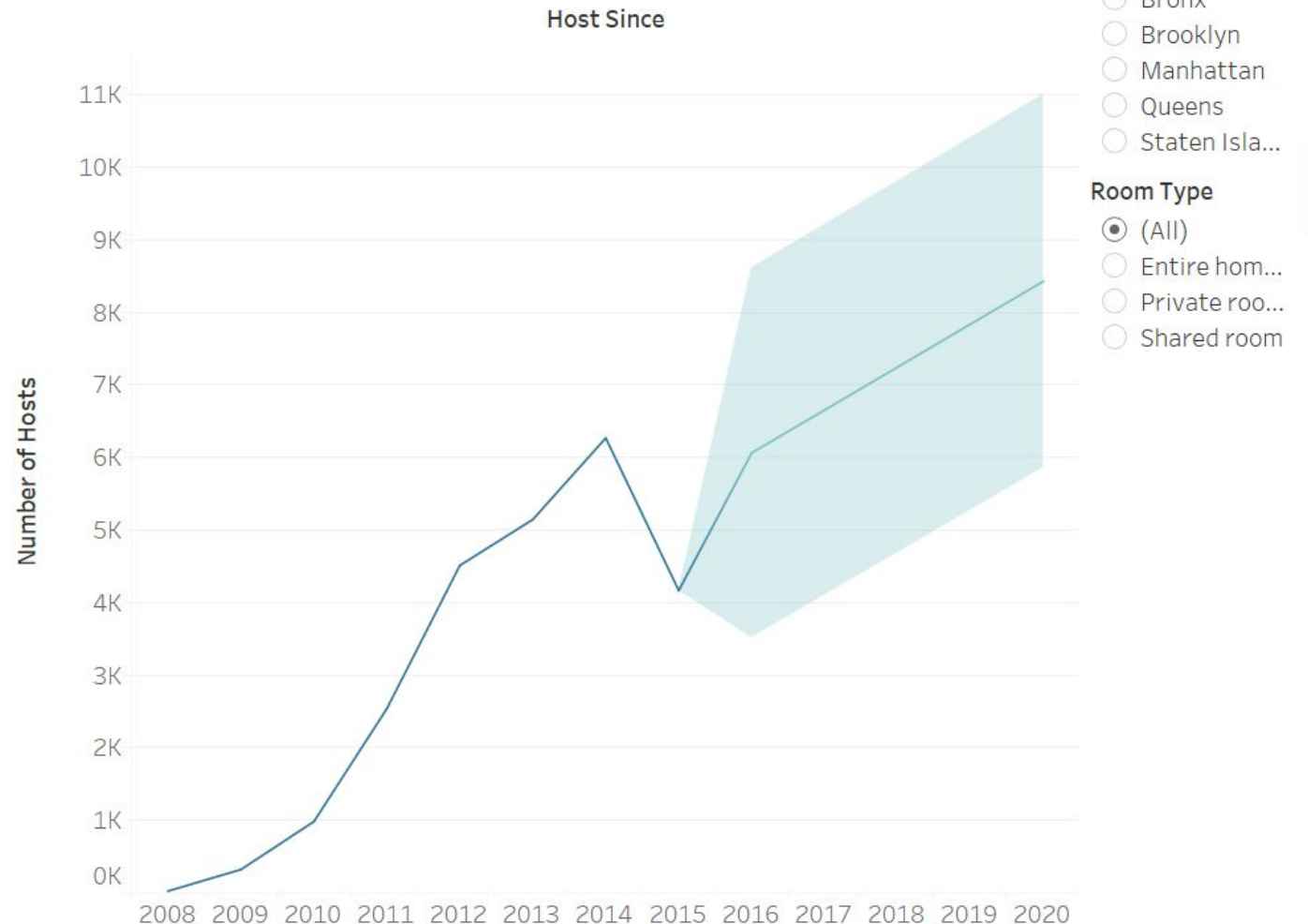Number of Hosts by Neighborhood and Avg. Price Over The Years on Airbnb
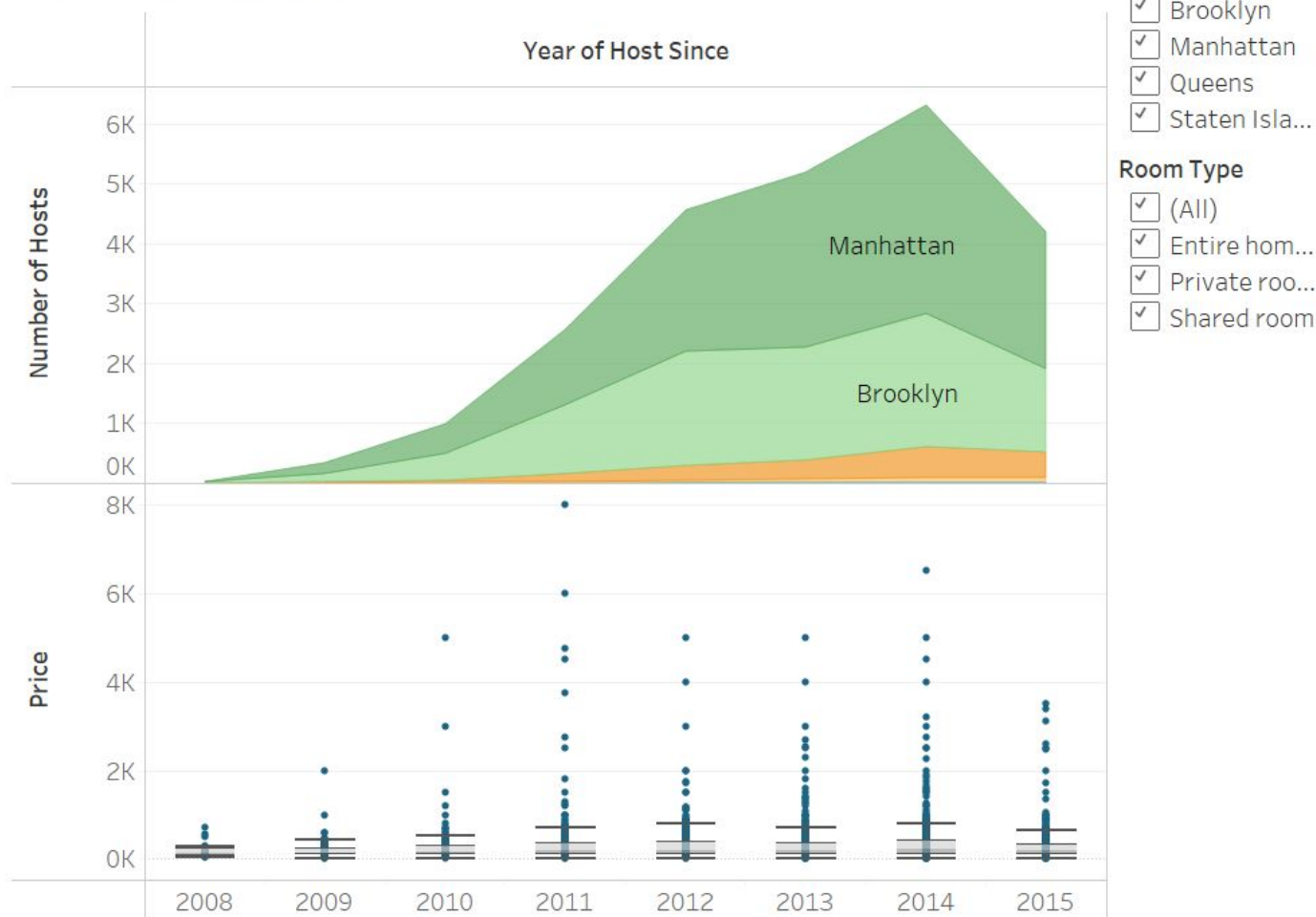
3. Insights & Dashboard

Question 5: What are some characteristics among NYC's Airbnb hosts?

- **Will the regulation mentioned sanction impact new hosts joining from 2015 onwards?**
  - The answer is it really depends, as we're dealing with an abrupt external business environment! Legal complications are always bearing a high degree of risks.
  - However, our forecast based on historical data shows that new hosts (aka the suppliers of the Airbnb platform) may recover and resumed growth. This is based on the assumption that homes & rooms rentals for vacations are always in high demand in New York City.
  - If we have more data on the Short-term and Long-term attributes of the current dataset, it'll help tremendously in improving the forecasts.

Forecasted Recovery in the Number of Hosts 2016 - 2020

Neighbourhood
- (All)
- Bronx
- Brooklyn
- Manhattan
- Queens
- Staten Isla...

Room Type
- (All)
- Entire hom...
- Private roo...
- Shared room

Main Project Challenges

- Time constraint to perform EDA, create visualization, clean data and outliers
- Hard to find the timeline of the dataset, except for the 'Host Since' information. However, it's not fully accurate to make assumptions that the dataset is from the 2008 - 2015 period.
- The need for extra time to dig into other interesting information outside the scope of this project, such as Review Score and its relationship with other variables.
- Might be interesting to have more information on the Short-term and Long-term rental aspects of this dataset, if available.