Exercise: Graph Design

In this exercise you will be working with Tableau and the AirBnB dataset, which collects information about AirBnB listings in New York City. You can find information about the data here: http://insideairbnb.com/get-the-data.html. Look for the NYC listings file.

Instructions

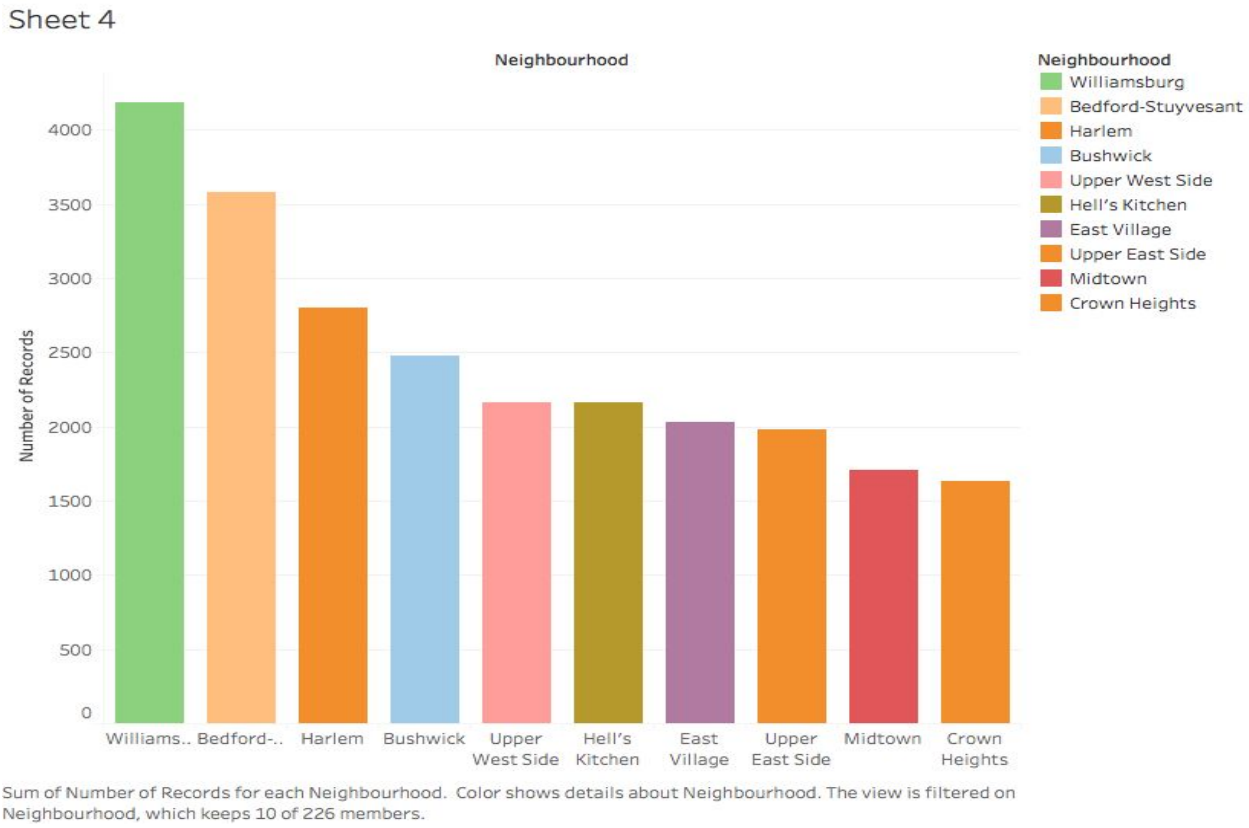
For each of the questions provided below you have to built with Tableau a graph that provides an answer to the question.

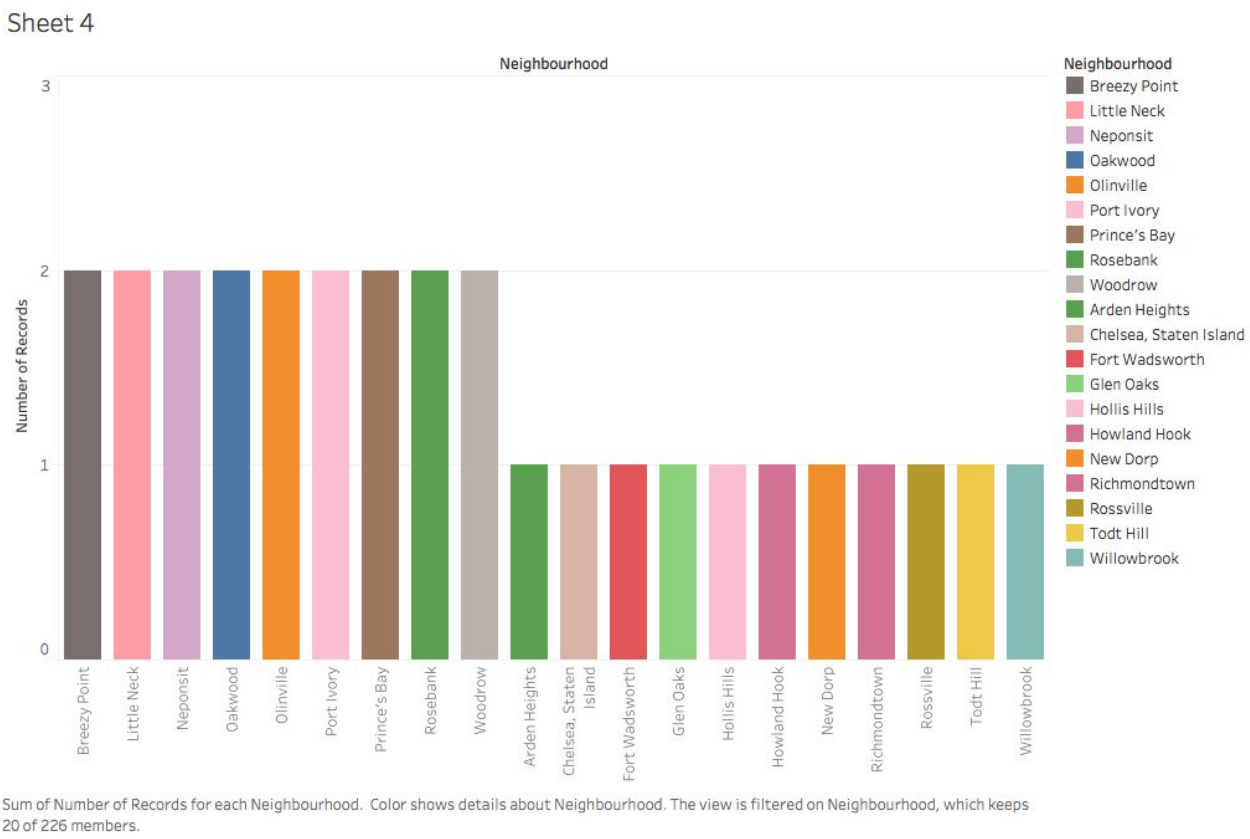
---

1) Which areas in NYC have the highest/lowest number of listings?

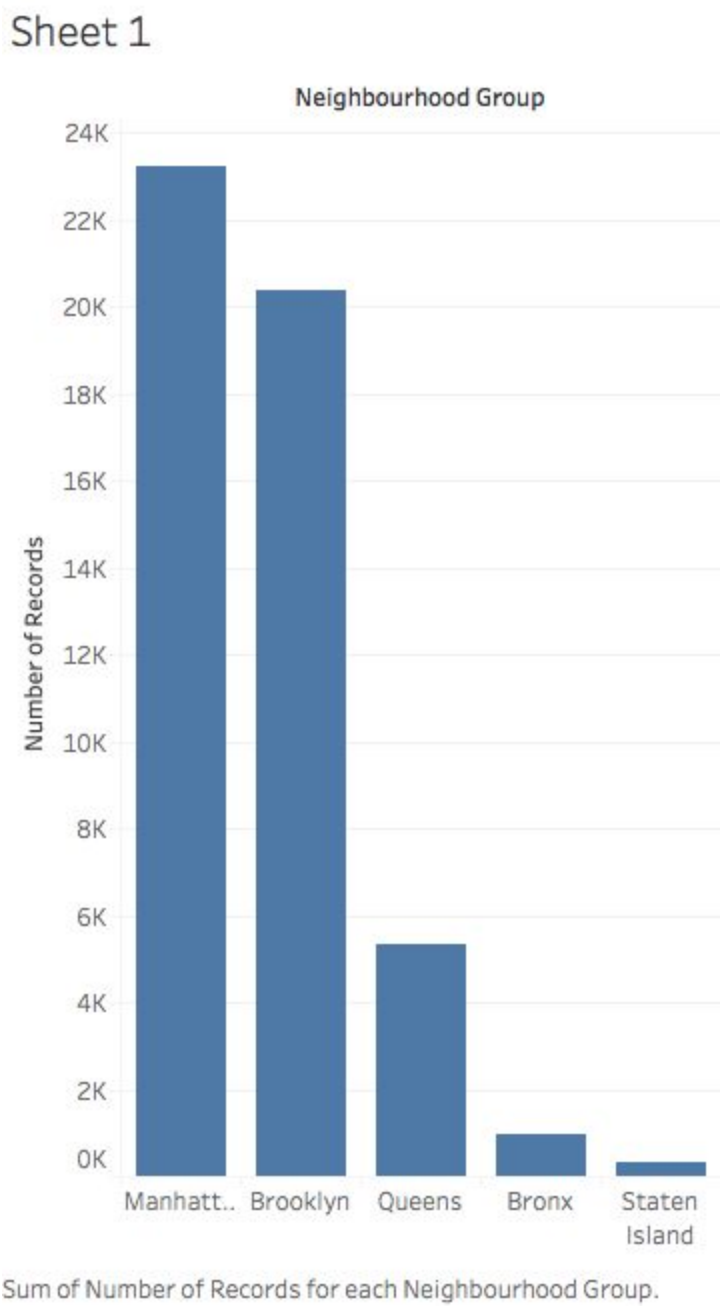
IMAGE + DESCRIPTION

Number of listings are highest in Williamsburg if we consider neighbourhood attribute. I have considered a filter of top 10 for neighbourhood attribute.





Number of listing are lowest in following areas : Arden Heights, chelsea (staten Island) etc. if we consider neighbourhood attribute. I have considered a filter of below 20 for neighbourhood attribute.

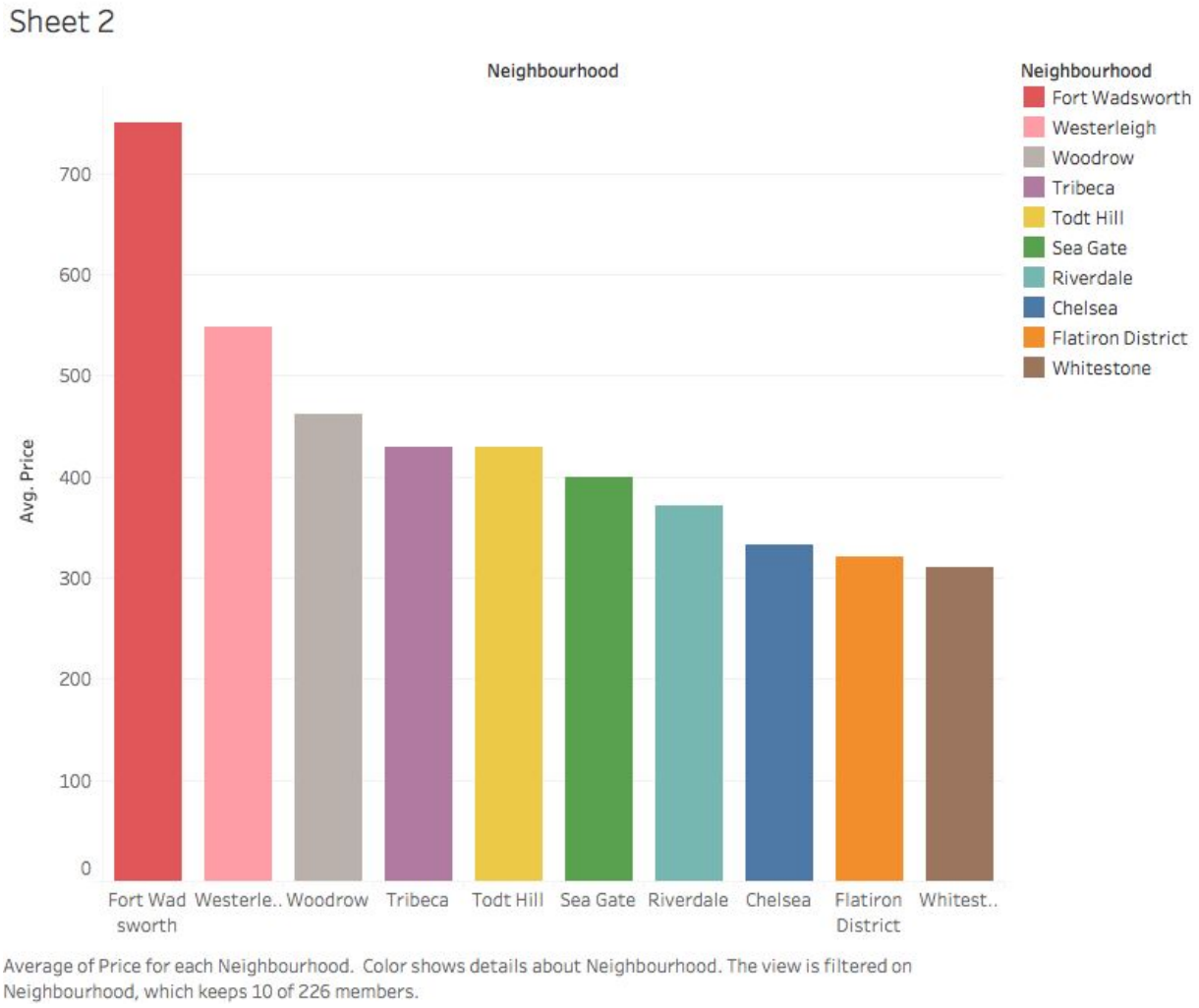


Number of listings are highest in Manhattan and lowest in Staten Island borough if we consider Neighbourhood Group attribute.

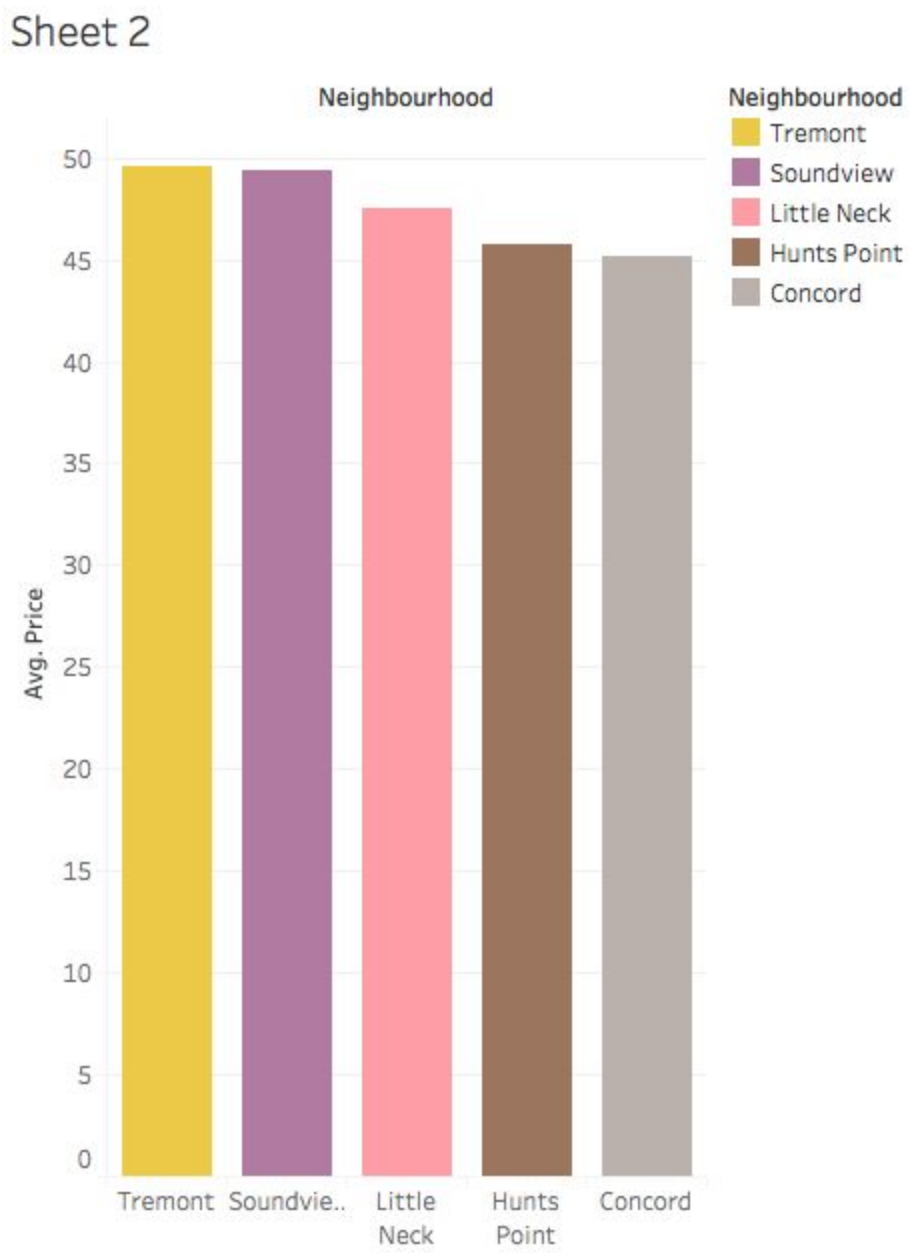
2) How does price distribute across different regions in NYC? Where are the areas with the highest and lowest prices?

IMAGE + DESCRIPTION

If we consider Neighbourhood attribute highest price are in Fort Wadsworth area. I have used filter of top 10 on neighbourhood attribute.



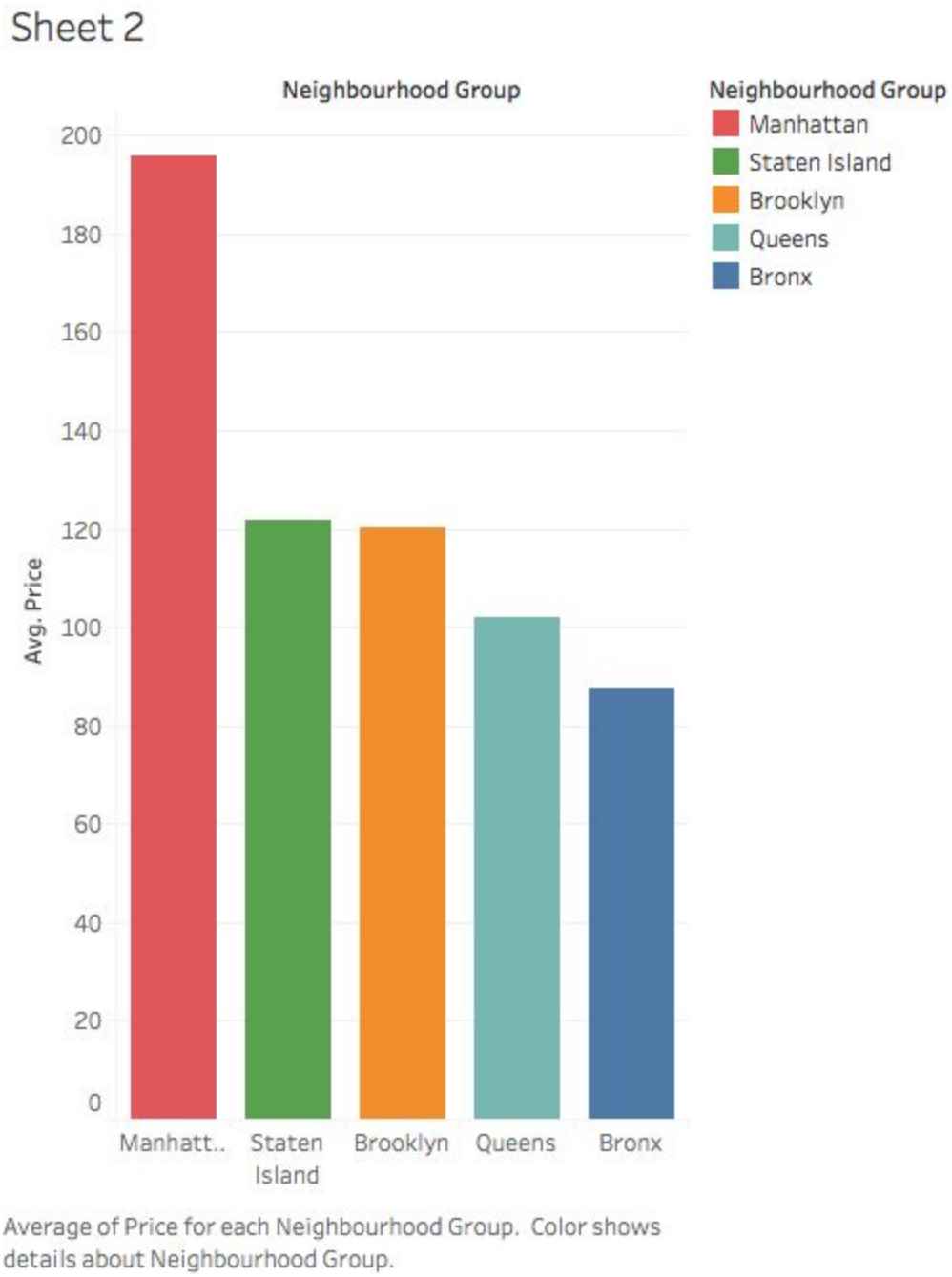
If we consider Neighbourhood attribute lowest price are in Concord area. I have used filter of below 5 on neighbourhood attribute.



Average of Price for each Neighbourhood. Color shows details about Neighbourhood. The view is filtered on Neighbourhood, which keeps Concord, Hunts Point, Little Neck, Soundview and Tremont.

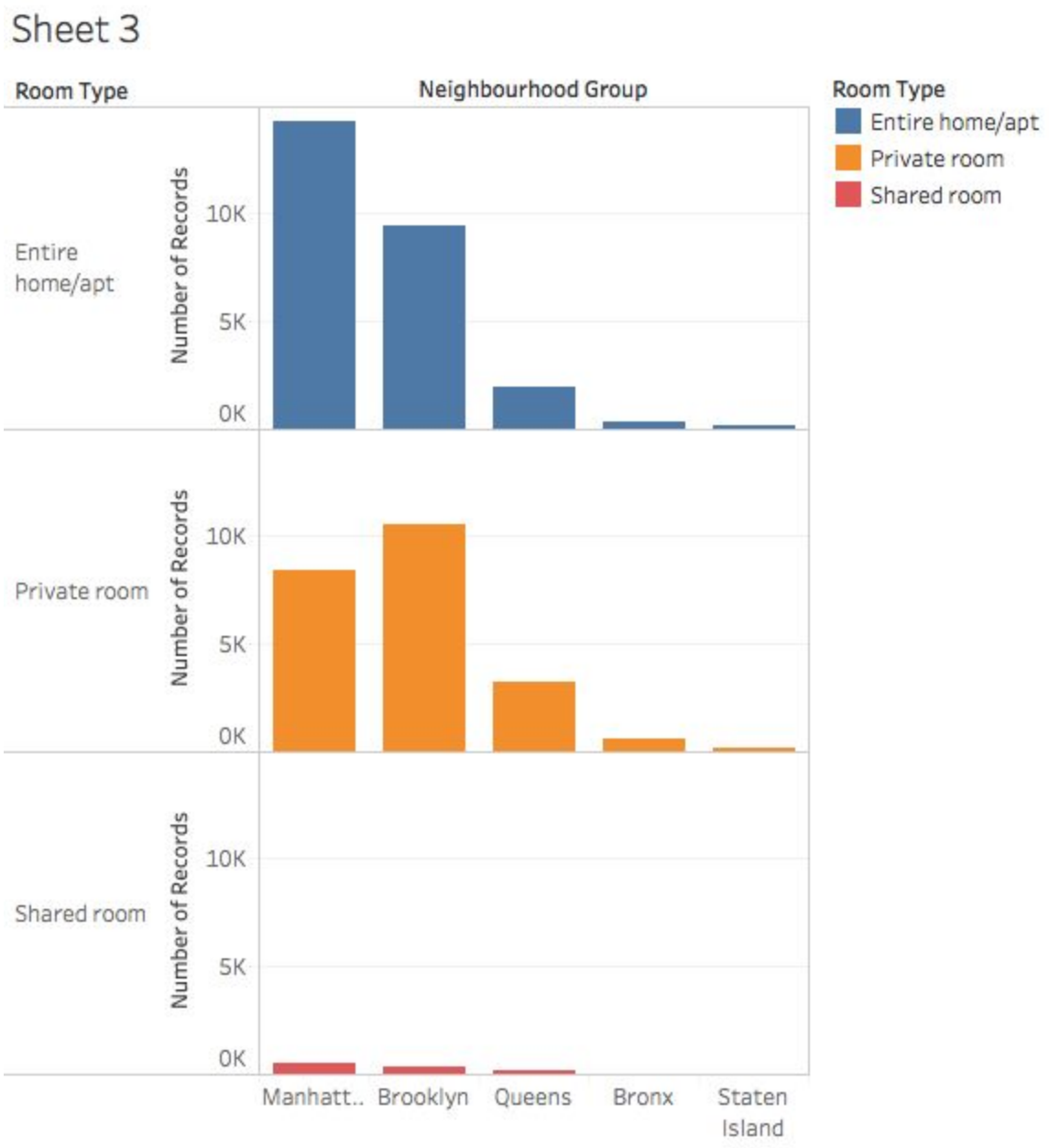
If we consider Neighbourhood Group attribute :

Area with highest price : Manhattan
Area with lowest price : Bronx



3) How does the prevalence of room types (use “Room Type”) change across the main boroughs? Are there any major differences on whether there are more entire apartments, private rooms or shared rooms?

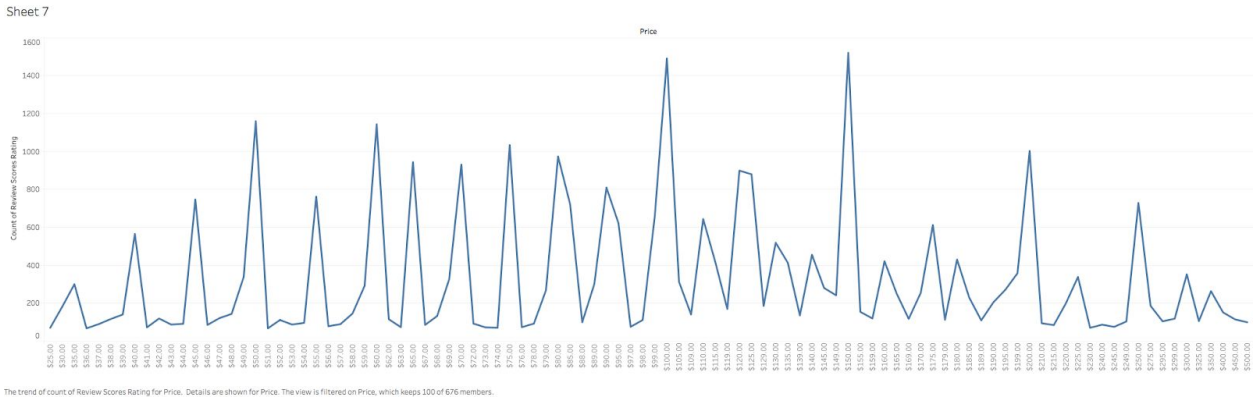
IMAGE + DESCRIPTION



From above visualization we can draw following conclusions:
Manhattan borough : Entire home/Apt is most prevalent
Brooklyn : Private Room is most prevalent
Queens : Private Room is most prevalent
Bronx : Private Room is most prevalent
Staten Island : Entire home/Apt and Private Room almost have same prevalence.

4) Is there a relationship between the price of a listing and the review score (use “Review Score Rating”)? Do listings with higher priced tend to be rated with higher scores?

IMAGE + DESCRIPTION



There is no unique relationship between price and review score as we can see from above visualization. And also listings with higher price does not necessarily have high review score.

--

***Important Notes**

In all of the questions above always consider how you can filter the data to make data patterns more intelligible. There are cases where you need to remove a few outliers (extreme values) before you start seeing something interesting. Make use of the “filter” function in Tableau.

In questions 1) and 2) there are many ways you can interpret the term “area” and “region”. The data set contains a few different attributes you can use to investigate areas (Zipcode, Lat/Lon, Neighborhood, Neighborhood Group). Find the combinations you deem more interesting and appropriate. If necessary produce more than one graph to emphasize different aspects of the same phenomenon.

In question 4) you will see that while it’s not hard to come up with a first graph to investigate this question it’s not easy to create an intelligible graph that shows some kind of pattern. You will probably have to explore a few different options before getting to something that works.