**Overview:**

Airbnb is an online company that provides a platform for finding unique accommodations for short-term stays. This platform enables guests to browse through a wide variety of listings, including apartments, villas, houses, and even treehouses and yurts. Airbnb listings can be found in almost every country in the world, making it a popular option for travellers seeking a local and authentic experience.

One of the unique features of Airbnb is the ability for guests to rent out an entire home or just a room within a host's home. This can provide guests with more privacy and space than a typical hotel room, as well as the opportunity to meet and interact with locals. The range of pricing for Airbnb listings is also very diverse, which can make it an affordable option for many travellers.

After staying at an Airbnb listing, guests are encouraged to leave a review and rating of their experience. This feedback is beneficial for both future guests and hosts, as it helps to ensure that high-quality standards are maintained throughout the platform. Hosts are also able to leave a review of their guests, which can help to ensure that guests are respectful and responsible during their stay.

Overall, Airbnb has become a popular option for travellers seeking unique and affordable accommodations, as well as hosts who are interested in renting out their homes or rooms. With its global reach and commitment to quality standards, Airbnb has revolutionized the travel industry and continues to attract a growing number of users.

**Objective and Investigation strategy:**

For this coursework, the objective is to investigate Airbnb's short-term home market in New York City (NYC) through exploratory data analysis (EDA) and create visualizations and dashboards to illustrate findings and answer questions of concern. The investigation will focus on specific questions of interest, including the median price for each room type in a particular neighbourhood, the room types available in a particular zip code and their respective prices and rating scores, and which zip codes have the highest ratings with over 100 reviews and relatively cheaper listing prices.

To begin, we will obtain the relevant data from Inside Airbnb, a publicly available dataset of Airbnb listings in NYC, in CSV format. We will import this data into R using the read.csv() function and clean it as necessary, such as removing missing values and standardizing variable names.

Next, we will conduct EDA to gain insights into the data. We will use techniques such as summary statistics, histograms, scatterplots, and correlation analysis to understand the relationships and trends in the data. Based on the results of the EDA, we will create visualizations and dashboards to illustrate our findings and answer specific questions of interest.

By performing EDA and answering specific questions of interest, we will gain a deeper understanding of Airbnb's short-term home market in NYC and identify areas of opportunity or concern for hosts and guests. This analysis can help inform decisions related to pricing, location, and marketing strategies for Airbnb hosts, as well as travel planning for Airbnb guests in the city.

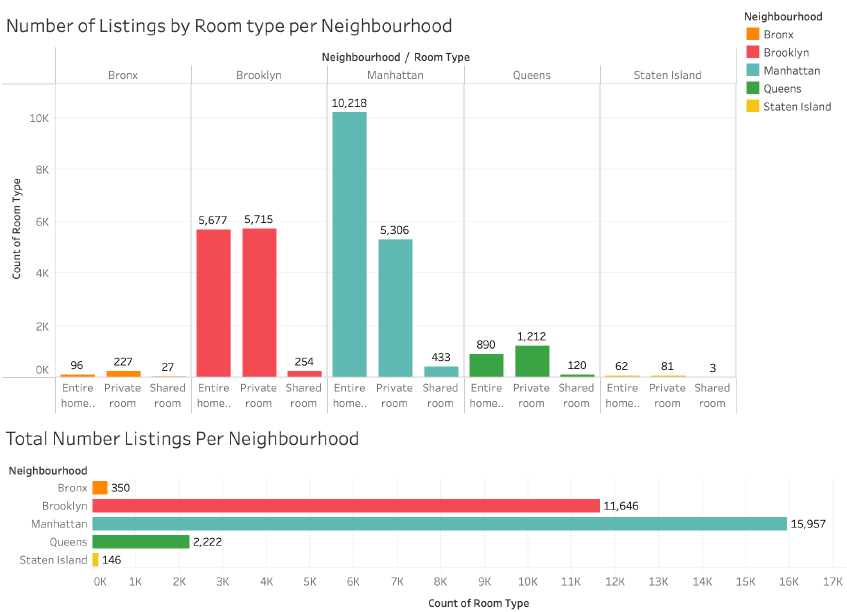
**Data and Resource:**

| **Remote Field Name** | **Datatype** | **Definition** |
| --- | --- | --- |
| Host Id | Number | Airbnb's unique identifier for the host/user |
| Host Since | Date | The date the host/user was created. For hosts that are Airbnb guests this could be the date they registered as a guest. |
| Name | String | Name of the listing |
| Neighbourhood | String | Neighbourhood in NYC |
| Property Type | String | Self selected property type. Hotels and Bed and Breakfasts are described as such by their hosts in this field |
| Review Scores Rating (bin) | Number | The rating bin of the listing as provided by the reviewers |
| Room Type | String | All homes are grouped into the following three room types: Entire place / Private room / Shared room / Entire place |
| Zipcode Geo | Geographic | Geographic long-lat of the listing |
| Beds | Number | The number of bed(s) |
| Number of Records | Number | The number of rows in the table |
| Number Of Reviews | Number | The number of reviews the listing has |
| Price | Number | Daily price in local currency (USD) |
| Review Scores Rating | Number | Score rating |

The Data set has been extracted from Kaggle site and processed with different fields as per the coursework requirement. The dataset can be found in the below link:

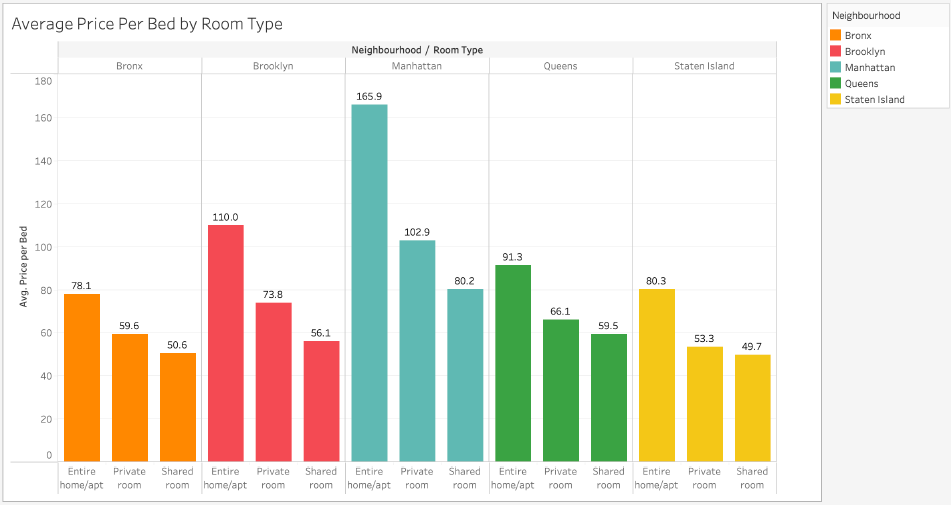
<https://www.kaggle.com/datasets/arianazmoudeh/airbnbopendata>

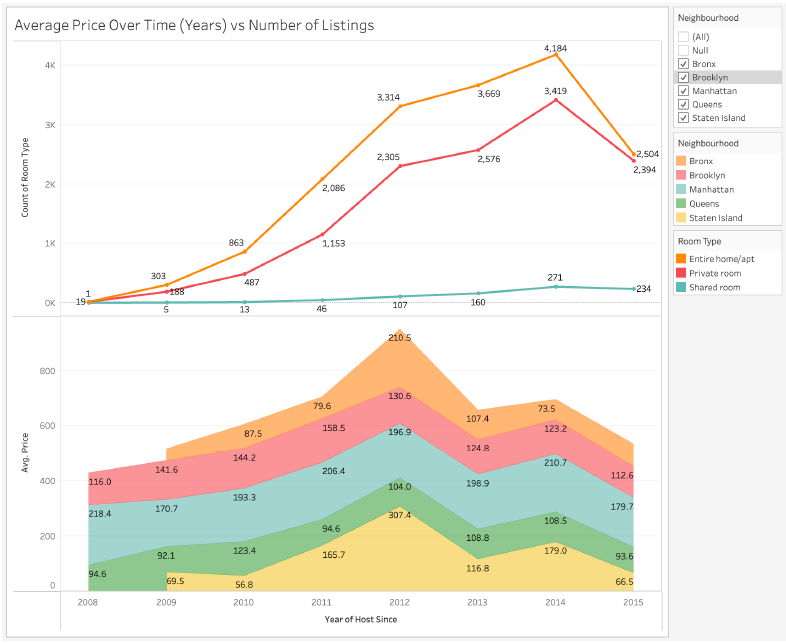
**Exploration Dive Deep:**

Upon analysing the data, it was observed that Manhattan and Brooklyn neighbourhood had the highest number of Airbnb listings in comparison to another neighbourhood. The reason for this could be attributed to the high popularity of tourism in these areas. Tourists often prefer to stay in the heart of the city, close to popular attractions, restaurants, and other amenities, which are mostly found in Manhattan and Brooklyn.

Furthermore, when comparing the different types of listings, it was noticed that the Bronx, Queens, and Staten Island had more listings for private rooms than entire homes or shared rooms. This could be because these neighbourhood are more residential, and property owners are more likely to rent out individual rooms rather than their entire homes. This trend could also be a result of the lower rental rates in these areas, making it more affordable for individuals to rent a private room rather than an entire home. Overall, these findings provide valuable insights into the Airbnb market in New York City, allowing hosts and visitors to make informed decisions about their listings and accommodations.

**Average Price Calculation per Bed:**

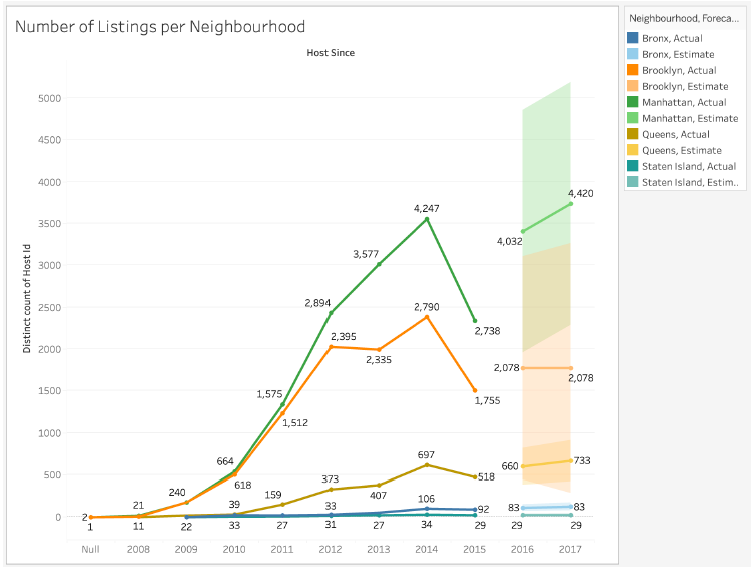
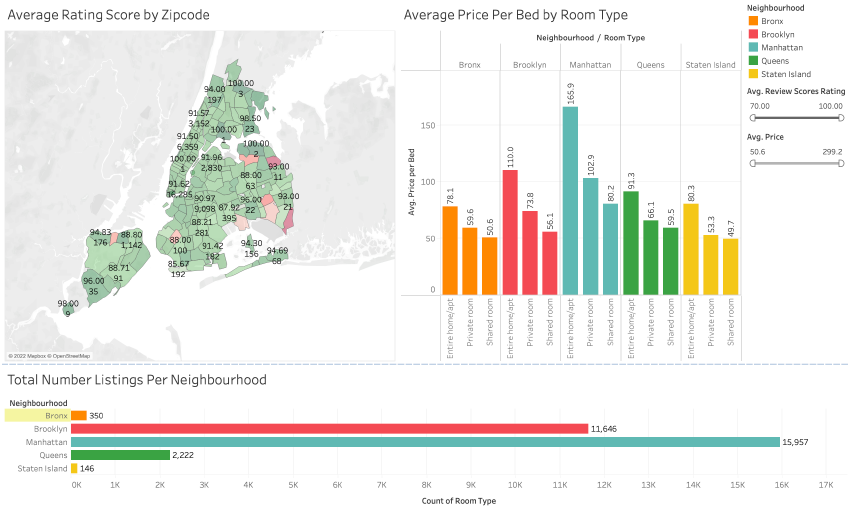
 From the below dashboarding, it was observed that the average price of a bed increases with the space of the room provided, specifically for entire homes and apartments. In other words, the larger the space of the listing, the higher the daily rental price tends to be. Moreover, it was noticed that the average price of a bed for an entire home/apartment is observed to be the most expensive in all neighbourhood.

Furthermore, when comparing the room types across different neighbourhood, it was discovered that Manhattan had the most expensive room types. This could be due to the high demand for accommodation in this neighbourhood, as it is a popular tourist destination and offers numerous attractions and amenities. Therefore, it is not surprising that the prices of Airbnb listings in this area tend to be higher than in other neighbourhood. These insights provide valuable information for individuals who are interested in listing their property on Airbnb or looking for accommodation in New York City, helping them make informed decisions about their pricing and location choices.

The trend of average listing prices for New York City neighbourhood over the years. Until 2012, the average price of listings increased in every neighbourhood. However, after 2012, there was a sharp decline in the average price of listings for each neighbourhood per year. The decline in prices can be attributed to the increase in listings in 2012, based on the number of room types available for rent, the average price of listings continued to decline gradually between 2013 and 2014, which may be due to a continual and gradual increase of listings. Between 2014 and 2015, there was another decline in the average price of listings, which can be attributed to a combination of increased competition (i.e., an increase in listings) and stricter regulations for short-term housing rentals in NYC. For example, Airbnb removed 1,500 listings from the public data released from November 2014 to November 2015, which likely impacted the supply and demand dynamics for short-term rentals.

The highlights is that the importance of understanding the factors that influence the average price of listings in NYC. While some factors, such as the number of listings, may contribute to a decline in prices, other factors, such as increased regulations, can also have an impact. As such, it is important to consider multiple factors when analysing trends in the NYC real estate market.

**Neighbourhood forecast and facts:**

The analysis presents information about the trend of Airbnb hosts listing their homes for short-term rentals in New York City. It notes that there has been a significant increase in the number of listings from 2008 to 2015, indicating the growing popularity of the short-term rental market in the city. The analysis also compares the increases in listings among

different boroughs, with Manhattan and Brooklyn experiencing the highest growth, while Queens, Staten Island, and the Bronx saw lower increases.

The analysis further provides a forecast analysis that predicts a continuing increase in the number of listings in Manhattan and Queens from 2016 to 2017.

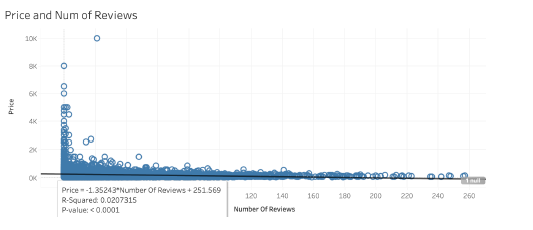
This suggests that the short-term rental market is likely to keep expanding in these two boroughs, potentially impacting the housing market, tourism industry, and local economies.

In addition, the analysis highlights an interesting finding about Airbnb's removal of 2,000 listings in 2015, following reports from the New York attorney general that many listings violated state or city regulations. This finding reveals the impact of regulation on the short-term rental market and the measures taken by Airbnb to comply with legal requirements. The analysis also raises questions about the validity of Airbnb's public data, as it removed 1,500 listings from its data for November 2014 to November 2015, indicating the need for caution when interpreting such data.

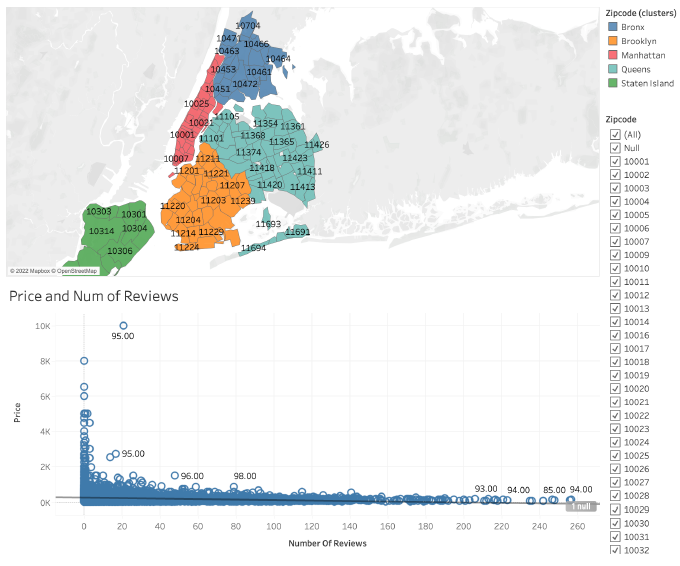
The study provides valuable insights into the growth and regulation of the short-term rental market in New York City, highlighting the impact on different boroughs, forecast analysis, and potential challenges of relying on data provided by short-term rental platforms.

**Higher Rating Zip code:**

Below studies states that there is no linear relationship between the price of Airbnb listings and the number of reviews they receive. This means that the number of reviews a



listing gets, does not increase or decrease in a predictable way based on its price.

* The R-squared value of 0.0207315 indicates that only a small portion of the variation in the number of reviews can be explained by the variation in listing prices. In other words, the price of a listing is not a good predictor of the number of reviews it will receive.
* Despite the lack of a linear relationship, the p-value of the linear model is significant, with a value of less than 0.0001. This means that the linear model is a statistically significant predictor of the relationship between price and the number of reviews.
* It's important to note that statistical significance does not necessarily imply practical significance. In other words, while the linear model is statistically significant, it may not be useful or meaningful in predicting the actual number of reviews a listing will receive based on its price.
* Overall, the analysis provides insight into the complex relationship between Airbnb listing prices and the number of reviews they receive, highlighting the need for caution in using price as a predictor of success on the platform.

**Closure and constrains:**

While working on the coursework the learned lessons are that collecting more data could lead to a more comprehensive and accurate investigation into the relationship between Airbnb listing prices and other factors, such as crime rates. By investigating the effects of crime rates on listing prices, researchers may be able to identify additional factors that influence pricing and occupancy trends, providing valuable insights for both Airbnb hosts and travellers.

Additionally, investigating the time of year when check-ins are most frequent could provide insights into the popularity of occupancy for Airbnb listings. This information could be used to forecast occupancy rates for future time periods, allowing hosts to plan and adjust their pricing strategies accordingly, either increasing prices during high occupancy periods or offering promotions or lower prices during low occupancy periods.

As per the analysis of Airbnb data for New York City, it can be concluded that Manhattan and Brooklyn have the highest number of listings and have seen the most increase in listings compared to other boroughs. The analysis also shows that there is no linear relationship between the price of listings and the number of reviews, but the linear model is still significant. Additionally, by getting more data, researchers could investigate the effects of crime rates on pricing and occupancy, and by studying the time of year when check-ins are most frequent, they could forecast occupancy rates for future periods. It is also important to note that in 2015, Airbnb removed thousands of listings after reports that many of them violated state and city regulations, which raises concerns about the validity of the data. Overall, the insights gained from this analysis could be useful for both Airbnb hosts and travellers to better understand pricing and occupancy trends in New York City.