



EMPIRE REALTY

# NYC's Airbnb Landscape

---



# Table of Contents

---

Section I	Meet the Team
Section II	Executive Summary
Section III	New York City Real Estate Market Overview
Section IV	New York Airbnb Open Data 2024
Section V	Data Cleansing
Section VI	Data Visualizations
Section VII	Future Considerations, Recommendations & Conclusion

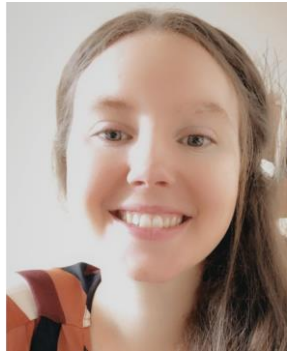


# Meet the Empire Realty Team

---



**Marium Hamdani**  
Data Analyst



**Hannah Redmond**  
Data Analyst



**Gavin Grosswald**  
Data Analyst



**Carolina Sepulveda**  
Data Analyst



# Executive Summary

## About Us

**Empire Realty (“Empire”)** is the premier real estate firm in New York City (“NYC”). With commitment to excellence and innovation, Empire is positioned for growth as it executes on its strategic approach of property acquisition. Empire Realty is dedicated to navigating the complexities of the NYC real estate landscape, always ensuring transparency, and maximizing returns for its stakeholders

## Approach for Success

### Property Diligence

- Evaluate Airbnb’s property landscape across NYC’s five Boroughs
- Understand drivers of successful listings and conduct acquisition diligence

### Property Acquisition

- Real estate portfolio expansion
- Footprint throughout the 5 boroughs
- Increased rental income returns



# New York Real Estate Market Overview

	Manhattan	Brooklyn	Queens	Bronx	Staten Island
Population	19,034,000	2,679,620	2,177,805	1,297,660	492,925
Average Income	\$79,205	\$53,626	\$56,592	\$56,723	\$54,960
Average Home Price	\$1,355,955	\$949,782	\$678,850	\$564,270	\$670,000

Note: Data sourced from Rockethomes.com



# New York Airbnb Open Data 2024

Airbnb, founded in 2008, is an online marketplace connecting hosts who offer short-term lodging accommodations with travelers seeking unique stays. Airbnb augments the travel experience through its global offerings and accommodations. The Airbnb data set was sourced from Kaggle.com and is a comprehensive list of the Airbnb rental market as of January 5, 2024

**22** Variables



**20,758** Observations

**456,676** Data Points

## Key Attributes

- |                      |                           |
|----------------------|---------------------------|
| • ID                 | • Number of Reviews       |
| • Name               | • Last review             |
| • Host ID            | • Reviews per Month       |
| • Host Name          | • Host Listing Count      |
| • Neighborhood Group | • Availability (365)      |
| • Neighborhood       | • Number of Reviews (LTM) |
| • Latitude           | • License                 |
| • Longitude          | • Rating                  |
| • Room Type          | • Bedrooms                |
| • Price              | • Beds                    |
| • Minimum Nights     | • Baths                   |

**Note:** Variable definitions can be found in the written report





# Data Cleansing

---

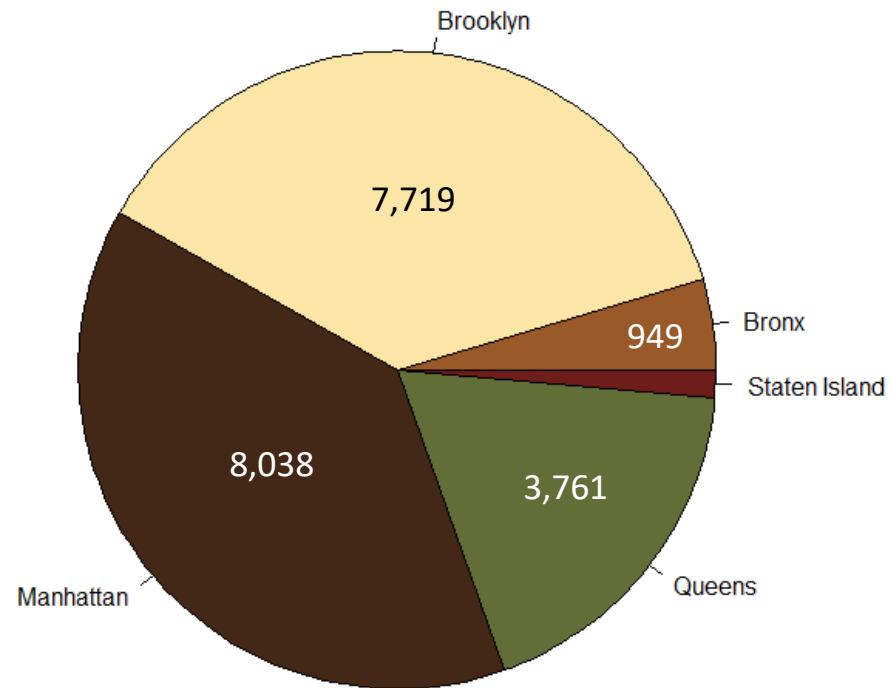
- No missing data or duplicated data
- Variable name, room\_type, bedrooms, rating, neighbourhood, neighbourhood\_group, host\_name and baths were changed from character to factor variable
- Removed unnecessary columns (id, name, host\_id, host\_name, license)



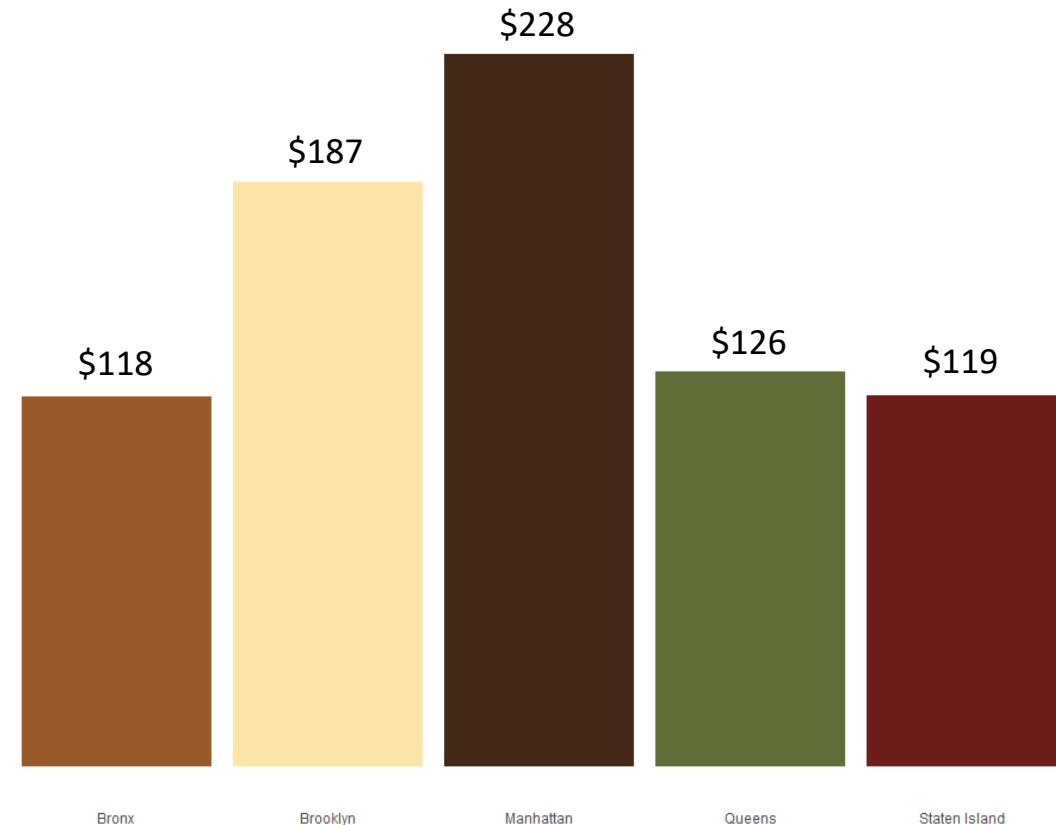


# Statistics by Borough

Count of Listing by Borough



Average Price of Listing Borough

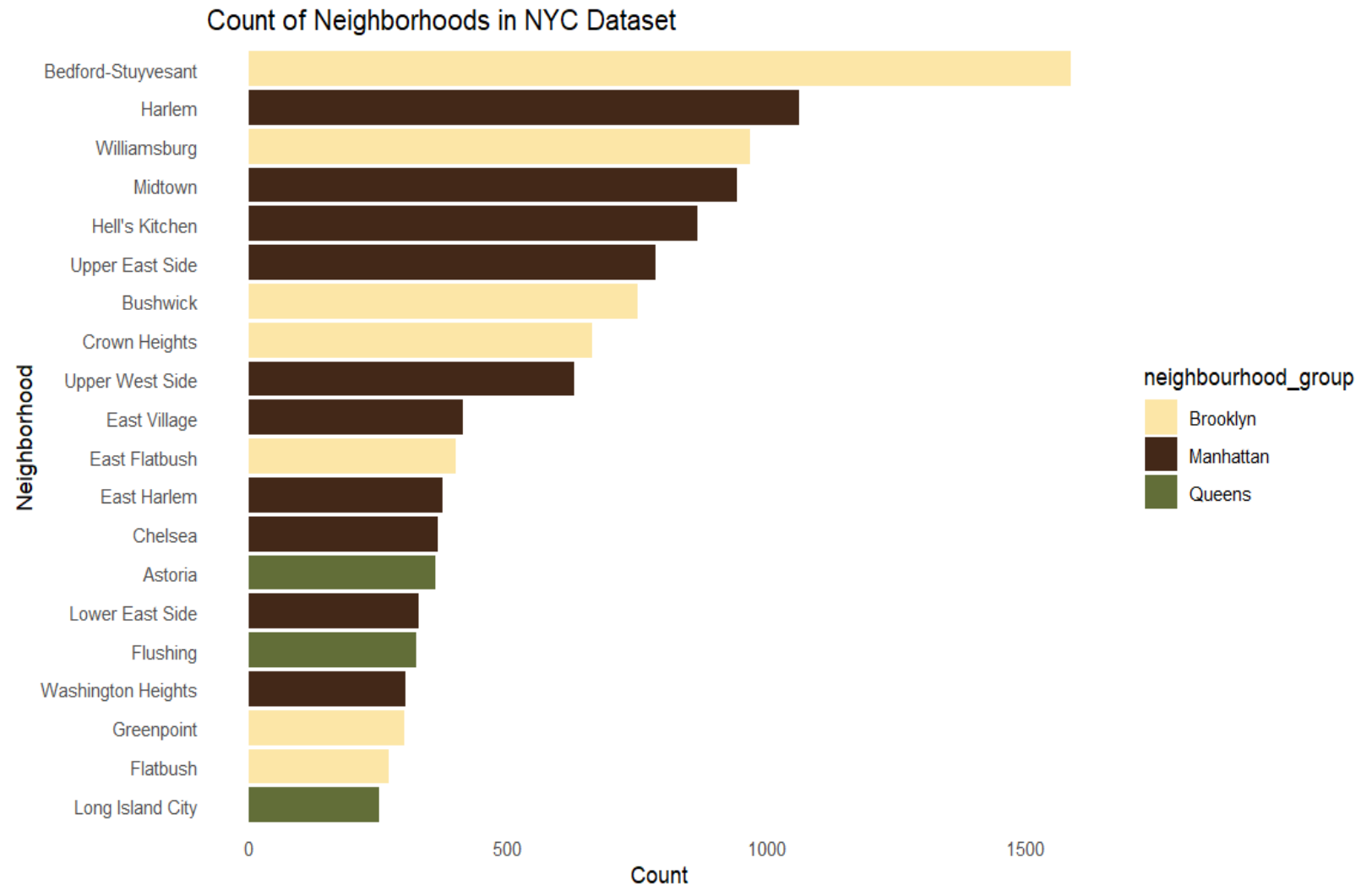






# Listings by Neighborhood

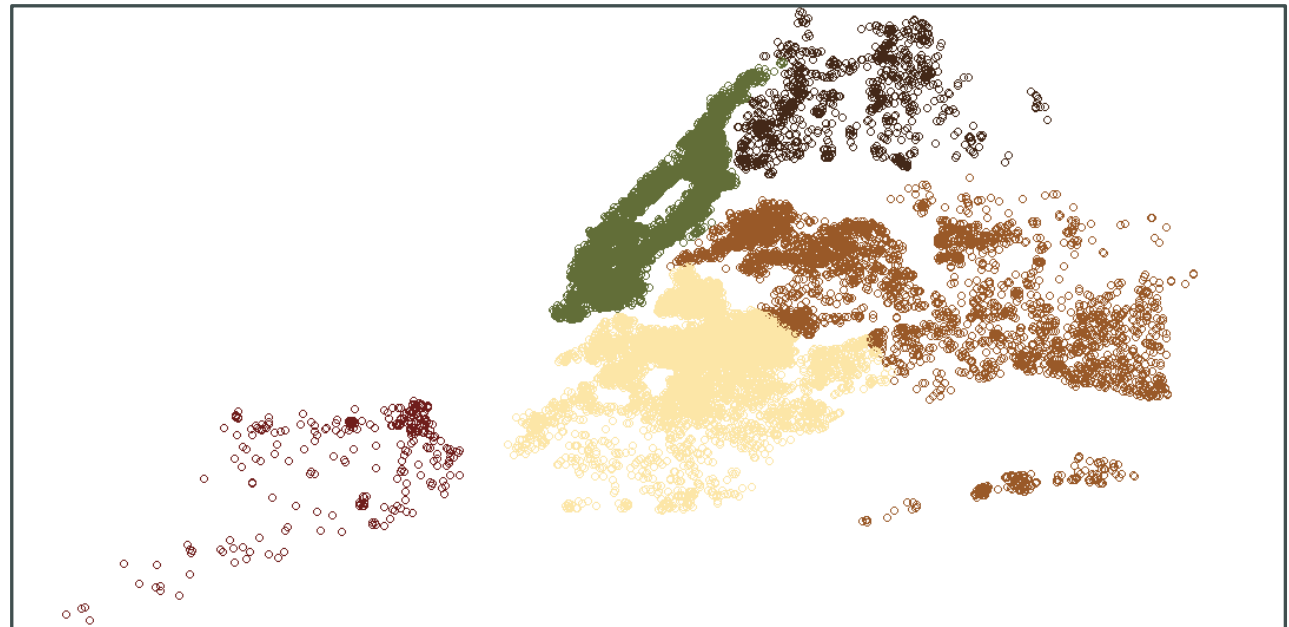
- Included neighborhoods have 250 + listings on Airbnb
- Heavy concentration of listings in Manhattan and Brooklyn
- Average cost per listing higher in Manhattan and Brooklyn than the other Boroughs

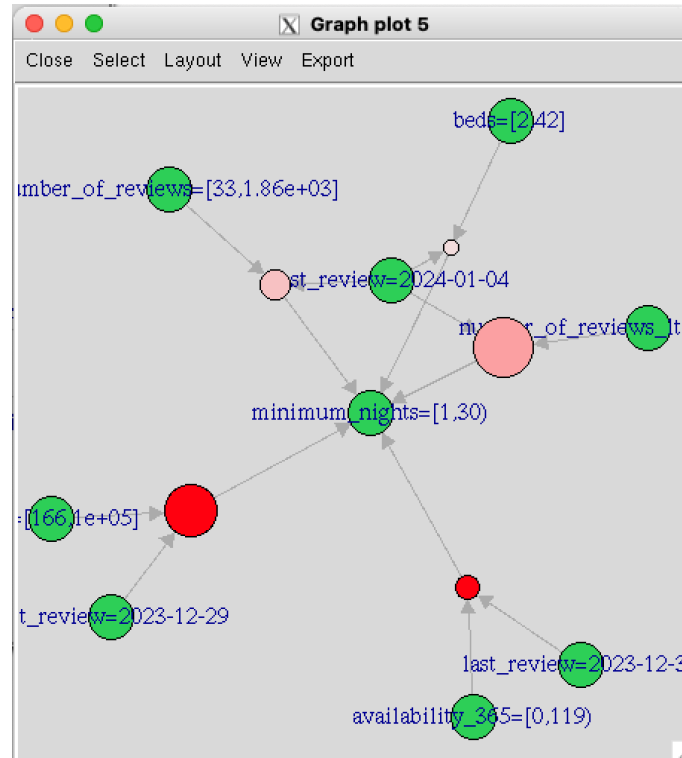
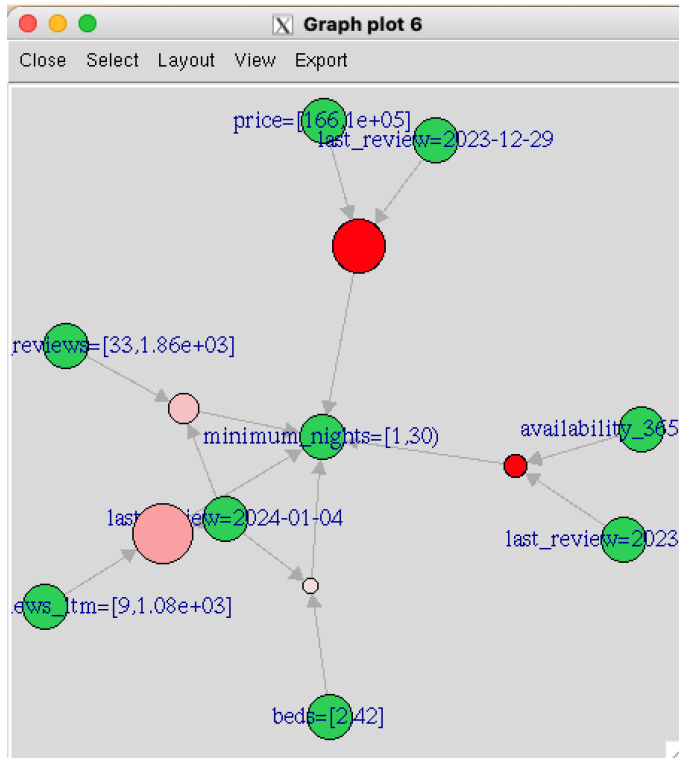




# Concentration of Listings in NYC

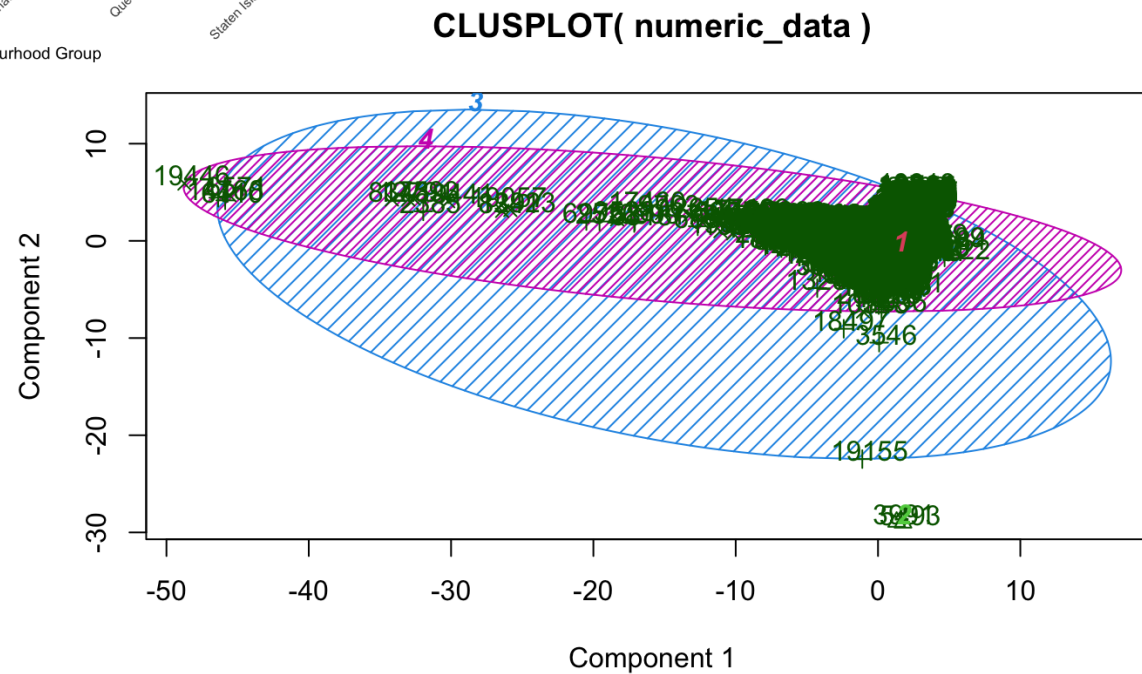
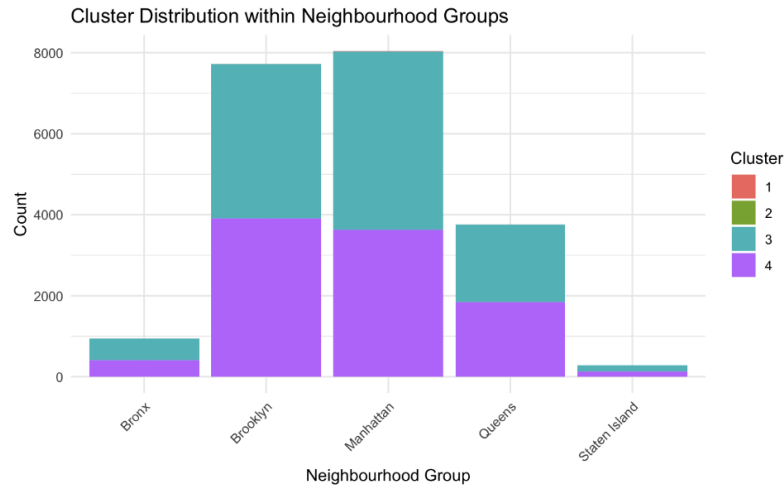
- The top visual shows a concentration of all NYC listing throughout all five boroughs
- The bottom visual breaks down the listing by Borough
- The greatest concentration of listing is in Manhattan and Brooklyn





- Target rental units primarily in Manhattan
- Rating greater than 4.60
- Featuring at least 1 bed and 1 bath

# Associative Rule Mining Analysis



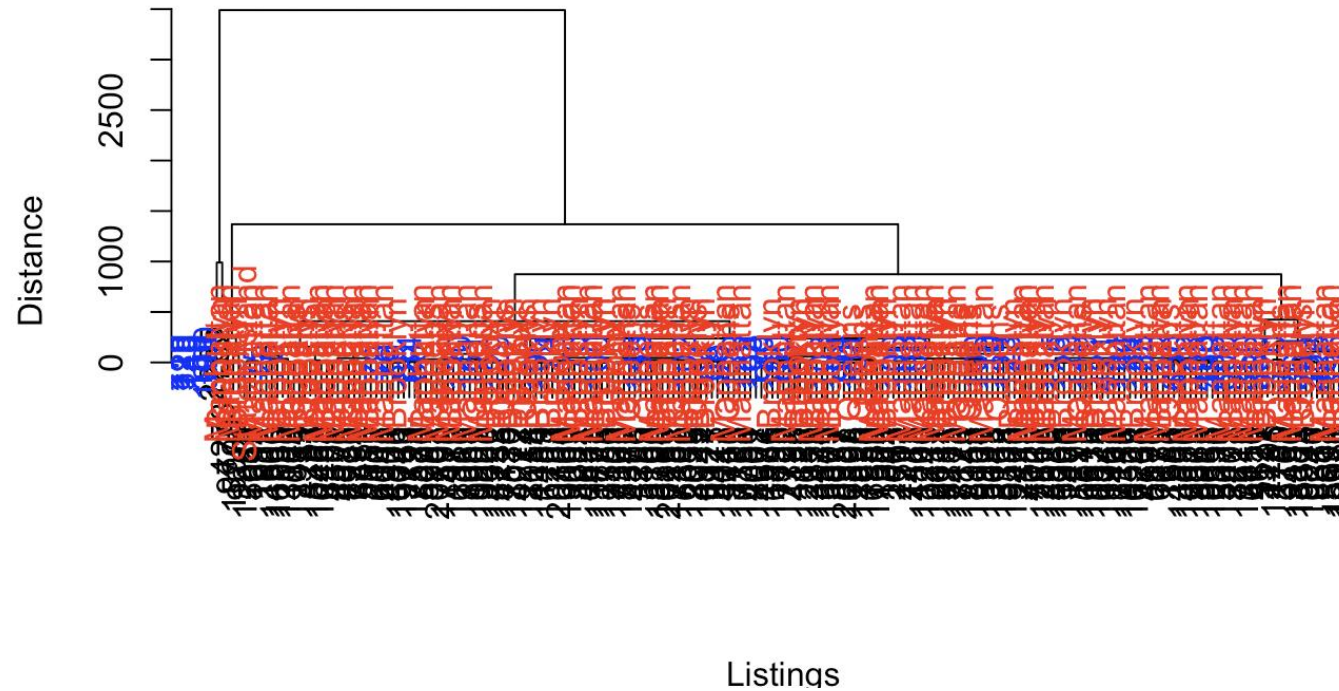
These two components explain 36.09 % of the point variability.

# Clustering Analysis



# HAC Analysis

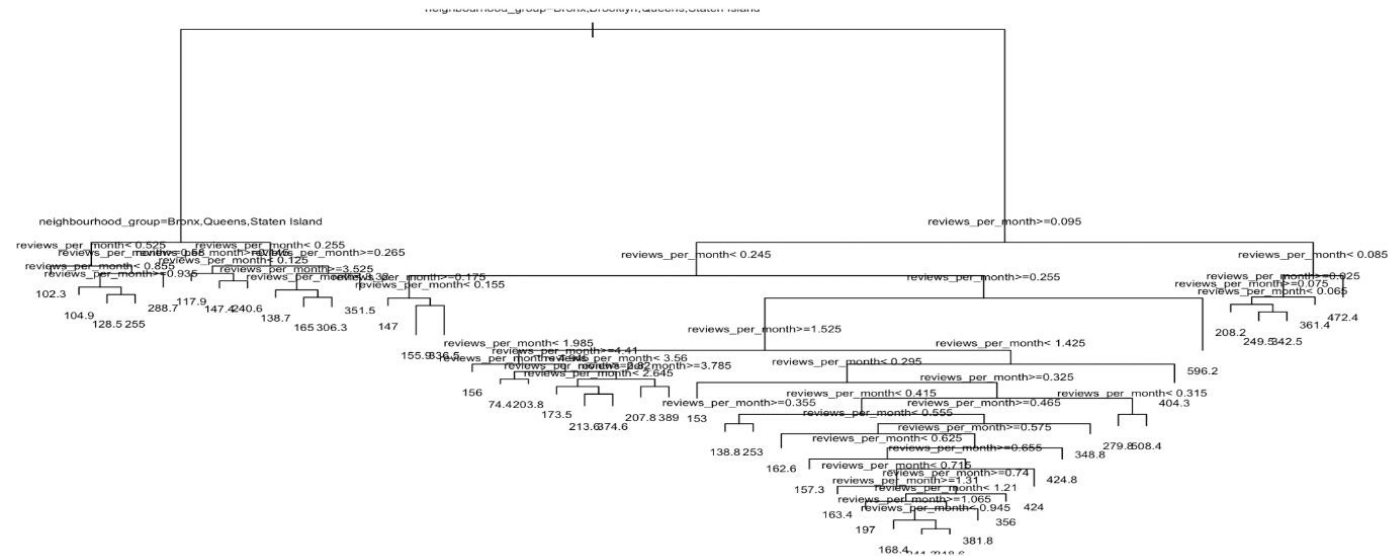
## HAC Analysis of Price and Neighbourhood Groups





# Decision Tree

- Trained using 10% of the total dataset
- Predict rental prices based on the number of reviews and property borough
- Determined that location plays a significant role in determining price
- Properties with a greater number of reviews tend to have lower price suggests that more affordable properties are more popular among renters

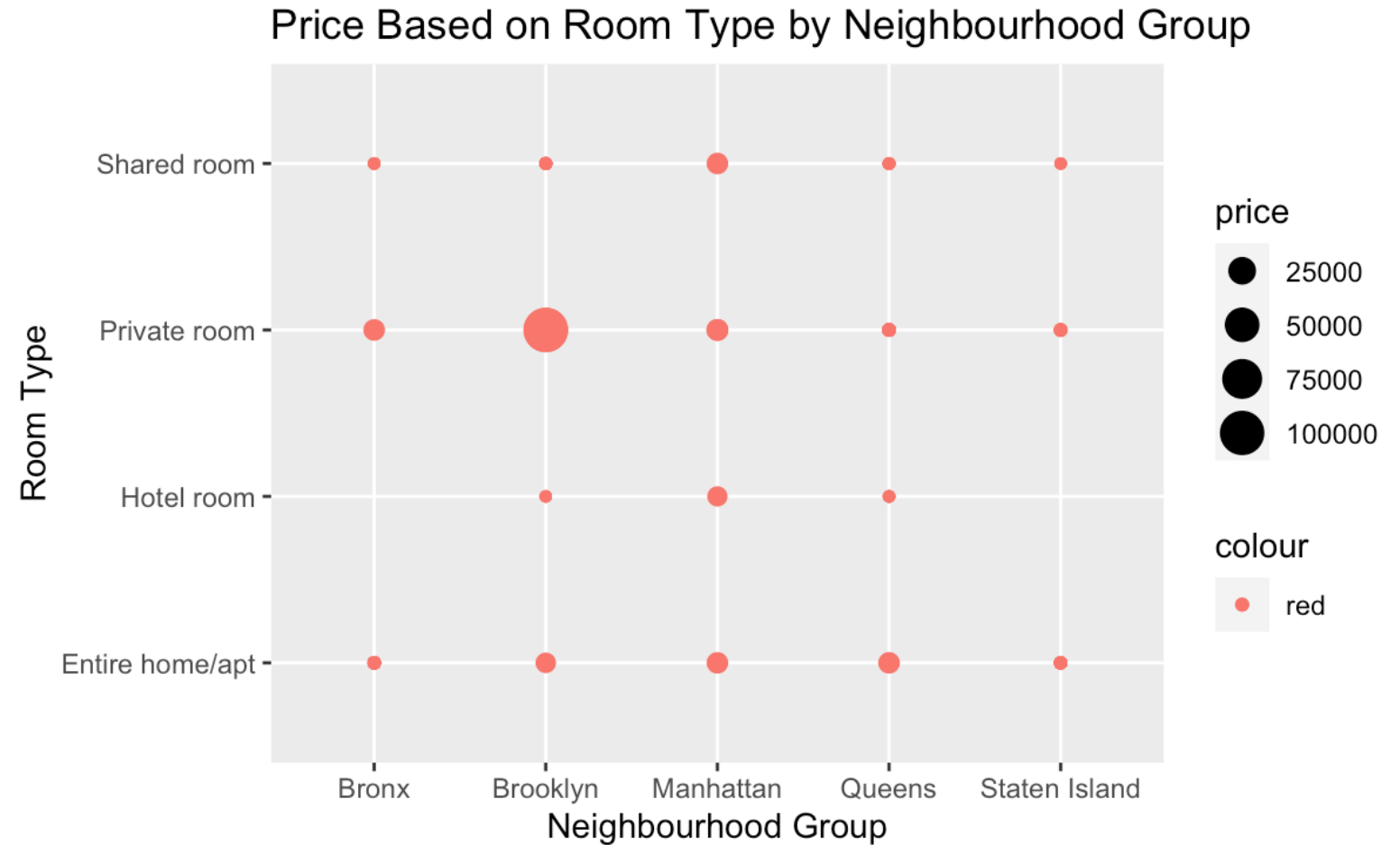






# NYC Listing Price / Room Type

- Most expensive private rooms are in Brooklyn
- Private room in Brooklyn is more expensive than the entire house/apt.
- Manhattan room types are similarly priced





# Future Considerations, Recommendations & Conclusion

## Dataset Size

- The dataset was quite large, consisting of ~500k datapoints
- Executing machine learning models often crashed RStudio or took an extensive amount of time
- Taking a random subset of rows resulted in faulty data and inaccurate analysis

## Airbnb vs. Leases

- Airbnb provides short term rentals while leases are typically longer term and executed via formal contract
- Using Airbnb data to predict longer term leases may not be the most relevant indicator
- Evaluating acquiring properties without mortgage related data

## Recommendations

- NYC is the tourist hub of North America – Volume is crucial
- Manhattan and Brooklyn have highest volume and highest prices
- Staten Island hindered by size, location, and access to transportation

## Conclusion

Empire Realty continues to position itself at the forefront of the NYC real estate industry due to its ability to successfully execute on its growth strategy and acquire properties with potential for long term sustainability. Through leveraging the Airbnb dataset, Empire is able to make informed decisions which result in a diverse and strong portfolio of properties