Better Your AirBnB

By Zach Meader, Anne Pizzini, Joshua Yesufu, and Ellen Grove

Agenda

- Methodology & Data Overview
- Visualizations
- Cleaning & Outliers
- Features & Predictions
- Final Considerations
- Demonstration

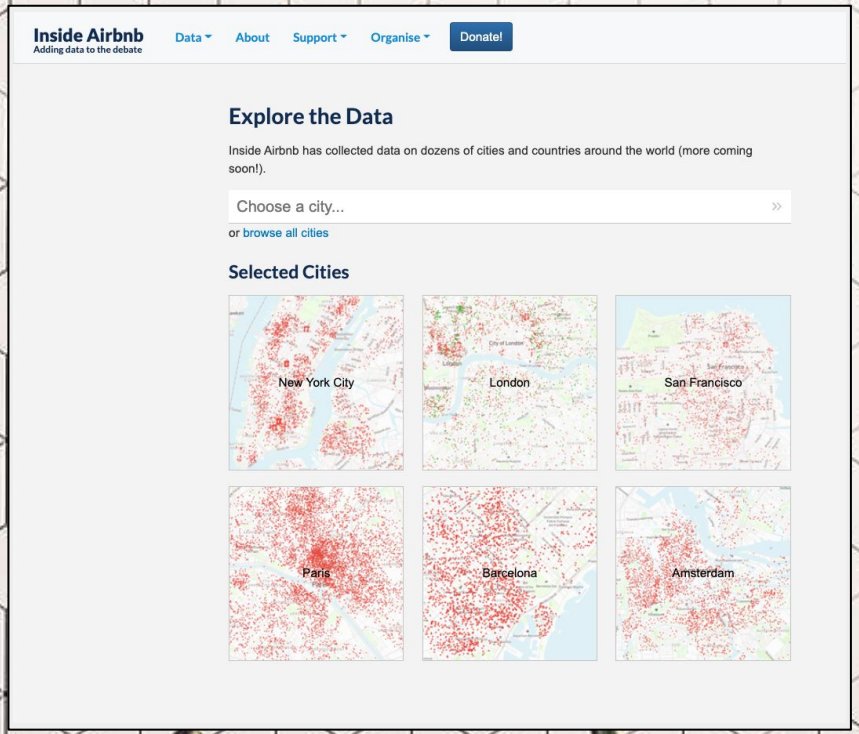
Methodology

Analysis & Visualization

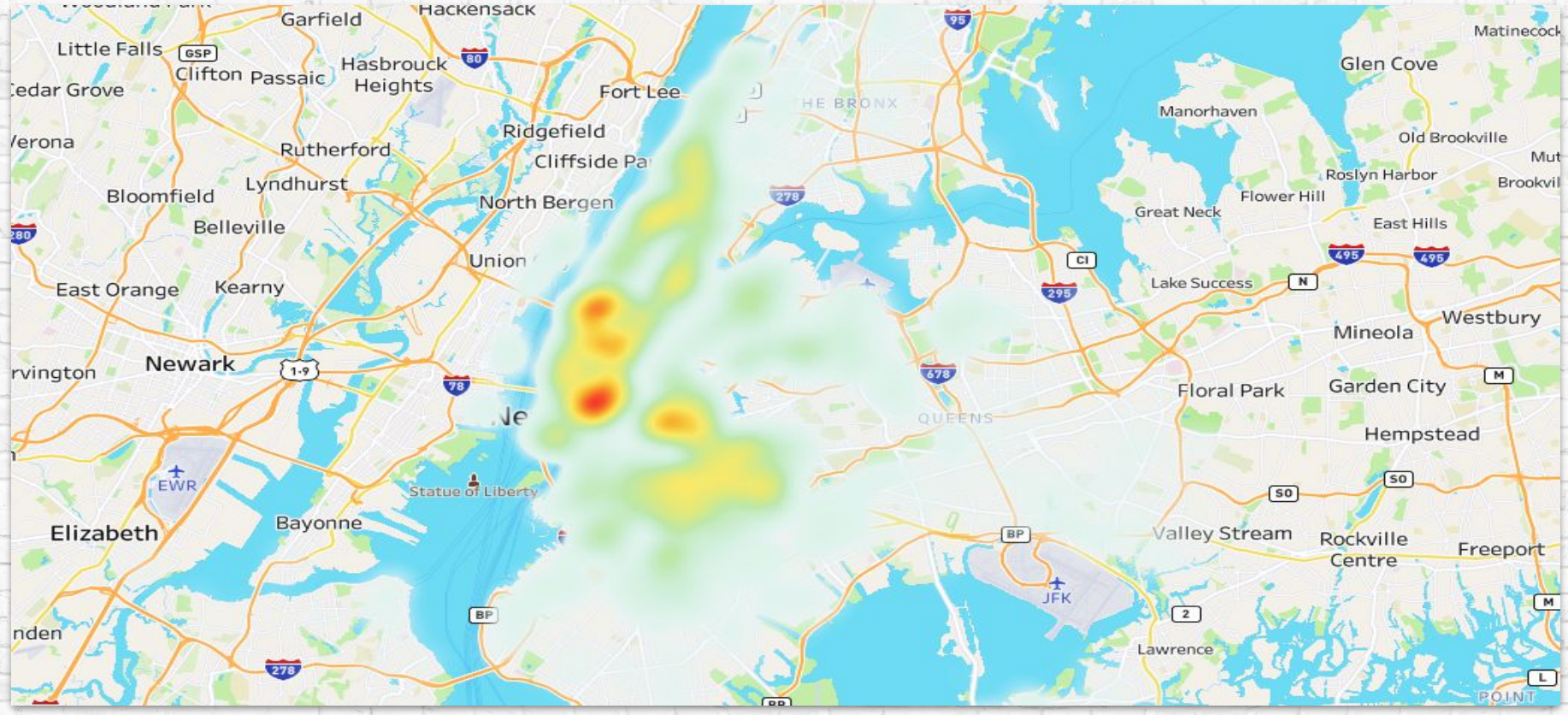
Feature Selection

Model Creation

The source of data, Inside Airbnb, is a mission driven project that provides data and advocacy about Airbnb's impact on residential communities.



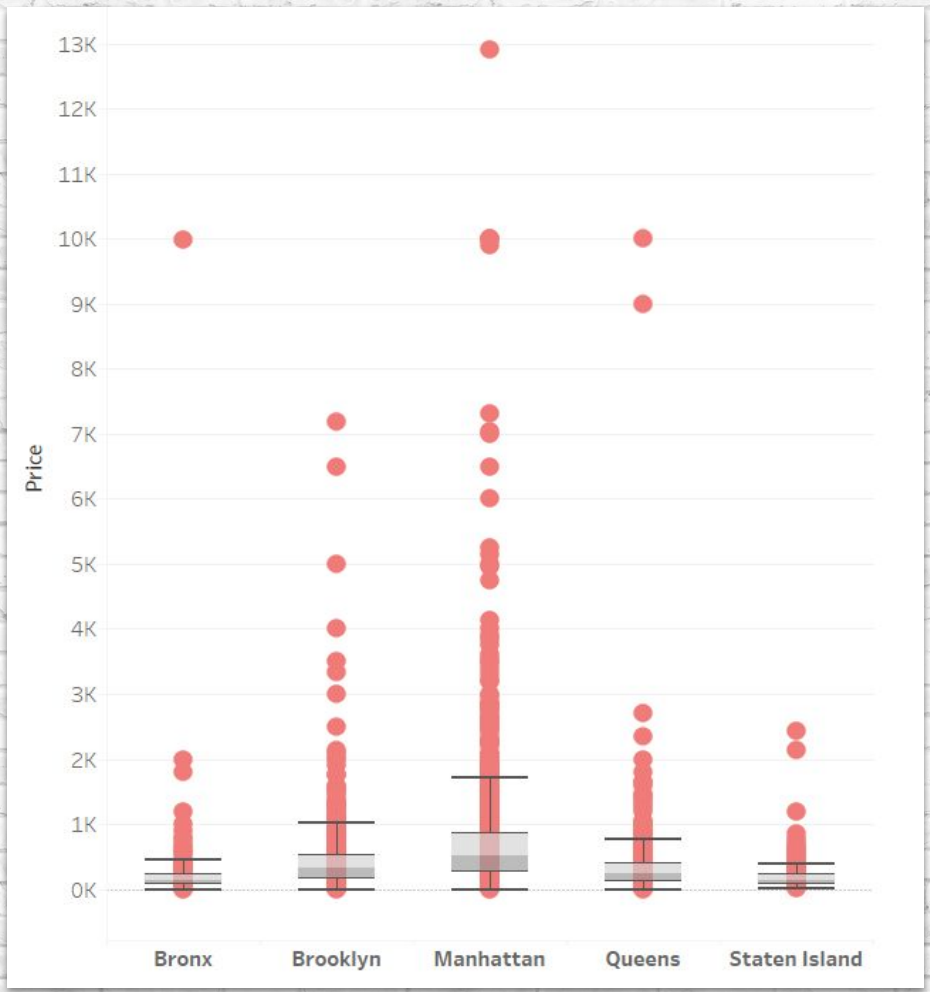
Heat Map of Listing Price



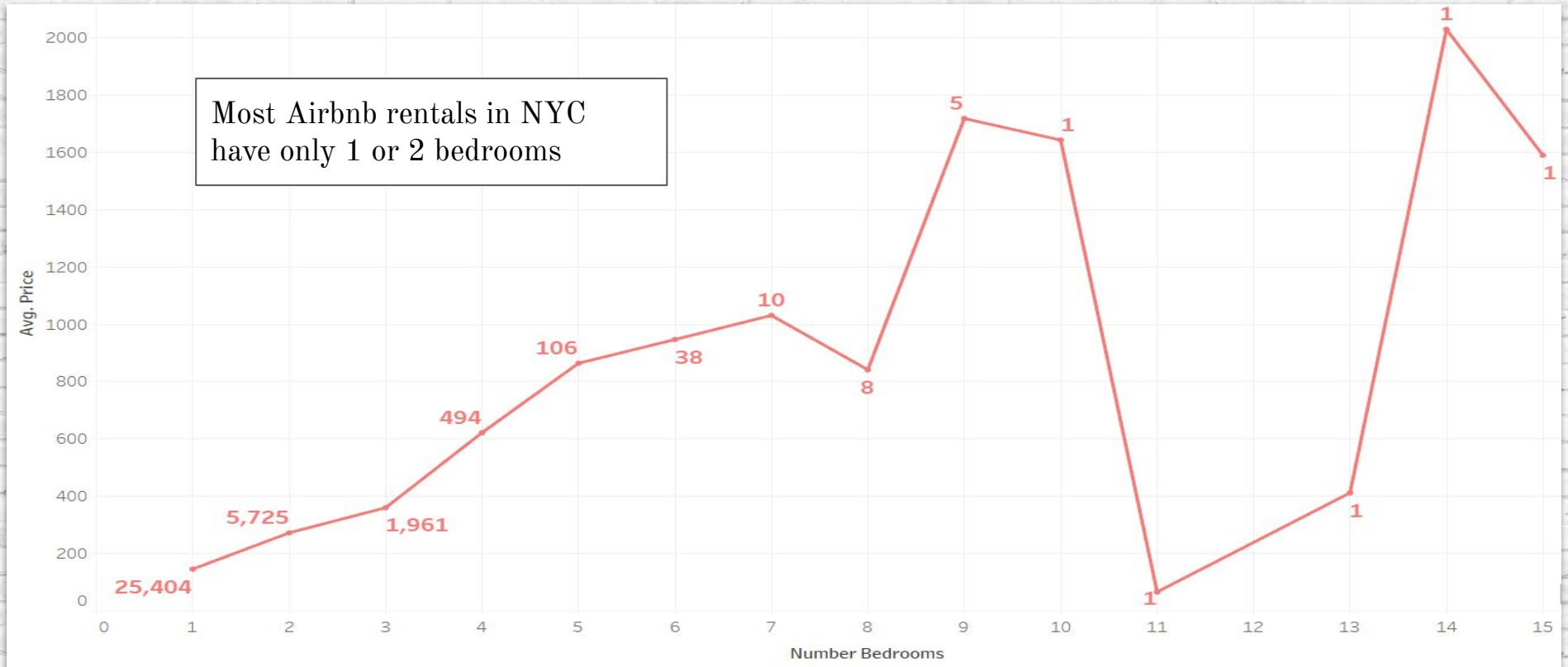
Price by Borough

There are critical differences in the Airbnb rental market by borough.

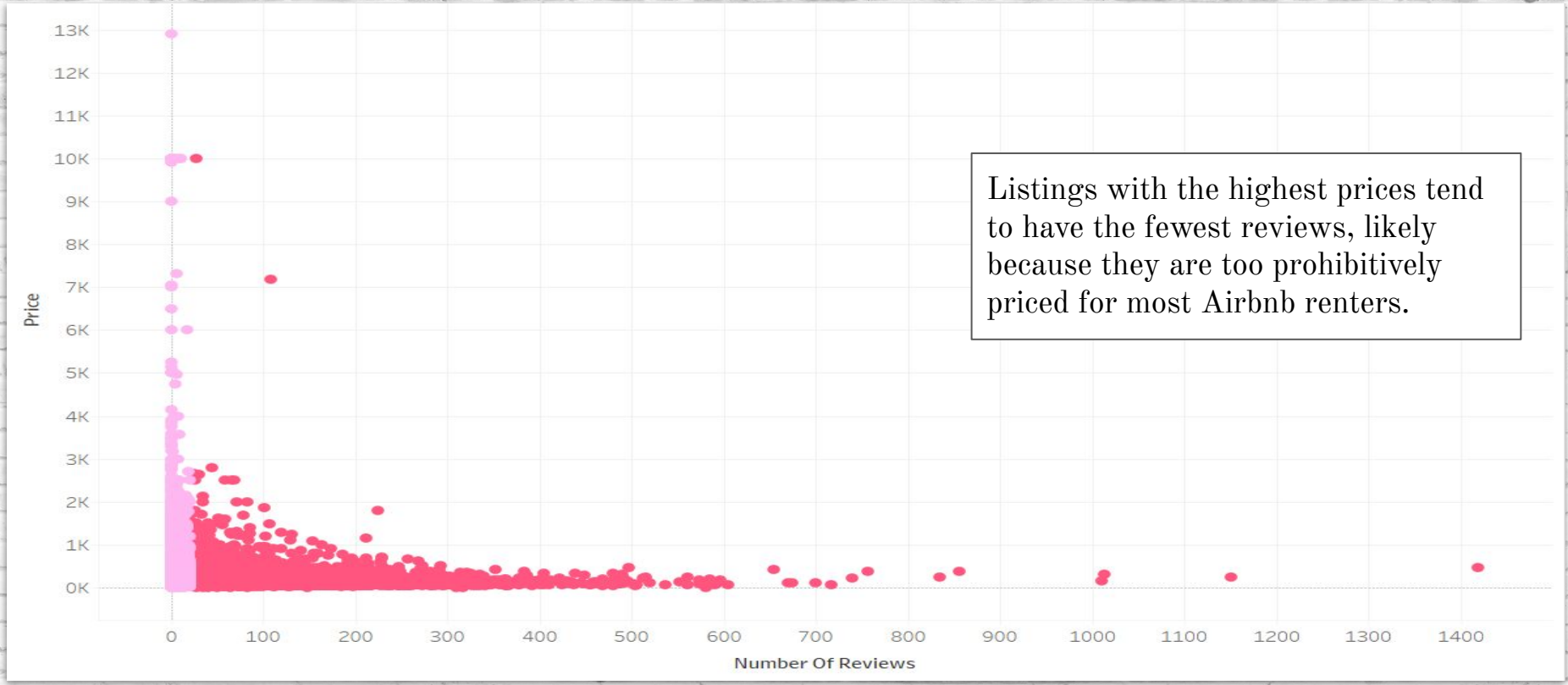
Example: the average price for an Airbnb rental in Manhattan is greater than for a rental in the Bronx.



Average Price Per Number of Bedrooms



Price by Number of Reviews



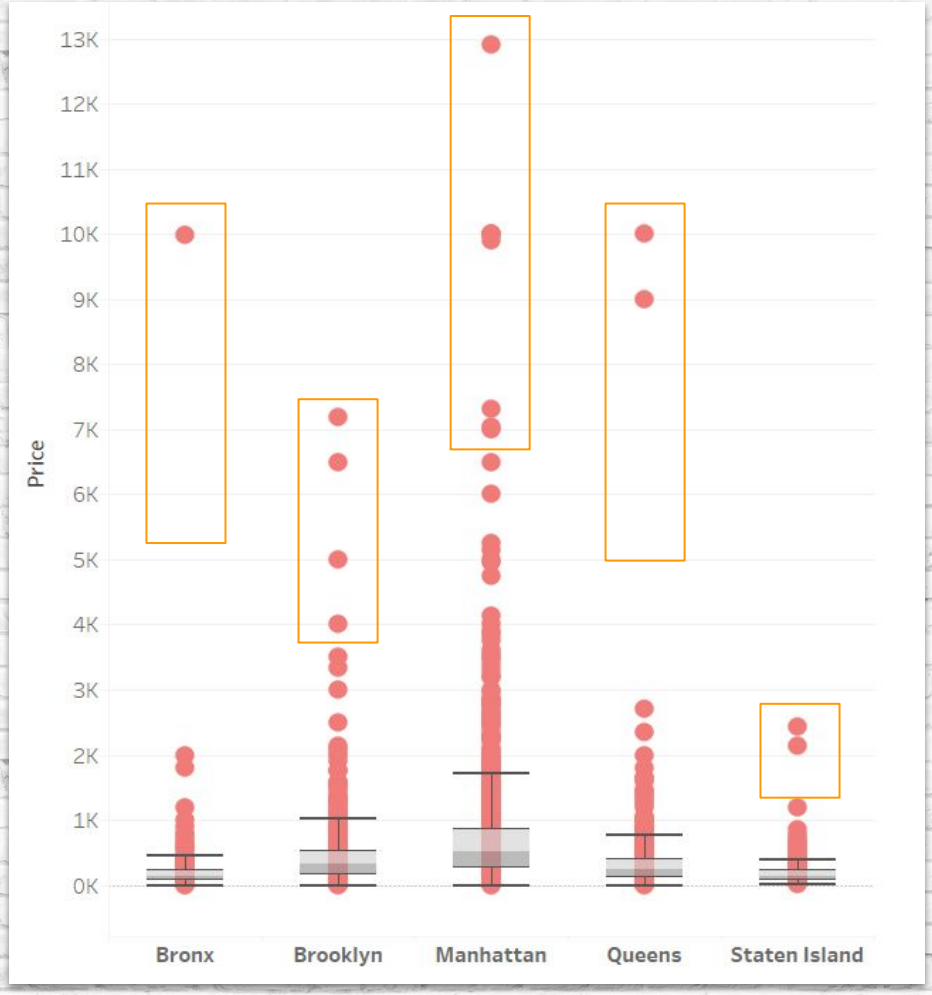
Listings with the highest prices tend to have the fewest reviews, likely because they are too prohibitively priced for most Airbnb renters.

Outliers Included Listings:

- ✗ < 20 Reviews
- ✗ > 10 Bedrooms
- ✗ 1 Year Min Stay
- ✗ 50% Greater Price than Max Price
- ✗ Unusual Property Type

Outliers by Price

Listings within the boxes are greater than 50% of the maximum price by borough and are considered outliers



Features

Borough

Rental Length

Cleanliness Rating

Number of bedrooms

Location Rating

Value Rating

Accommodates

Room Type

Host acceptance rate

Instantly Bookable

Overall Rating

Data Model

Linear Regression

Train R-squared: .4539

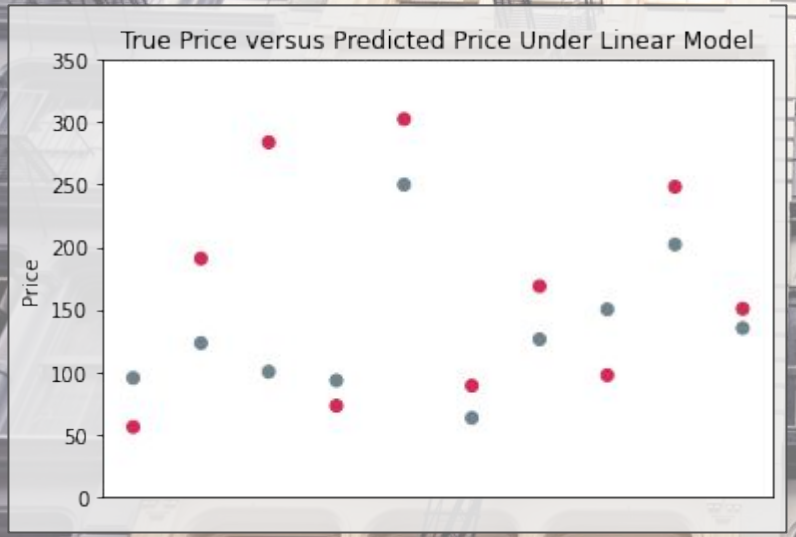
Test R-squared: .4319

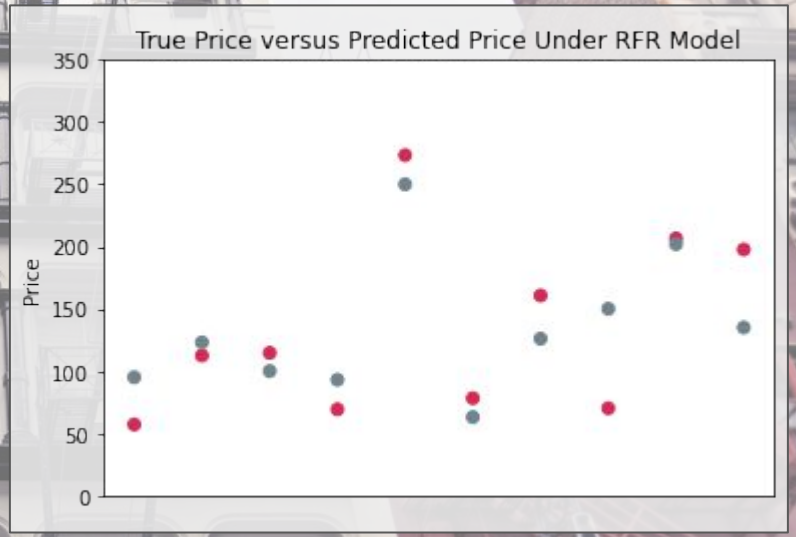
Random Forest Regression

Train R-squared: .9290

Test R-squared: .5144

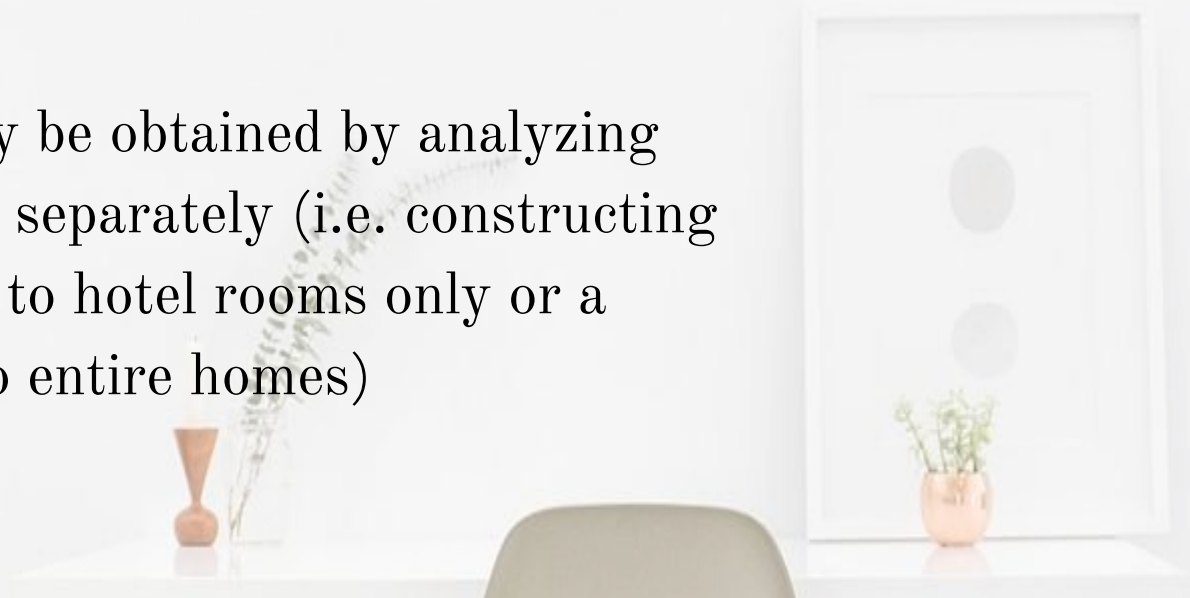
Sample Predictions





Considerations, Limitations, and Closing Thoughts

- Seasonality is an element that is not captured in our analysis
- This analysis can not be extended outside of New York City
- Stronger results may be obtained by analyzing different room types separately (i.e. constructing a model that applies to hotel rooms only or a model that applies to entire homes)




Put it to the test.