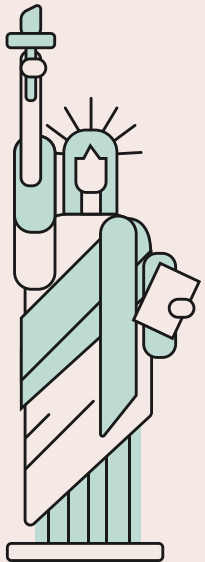
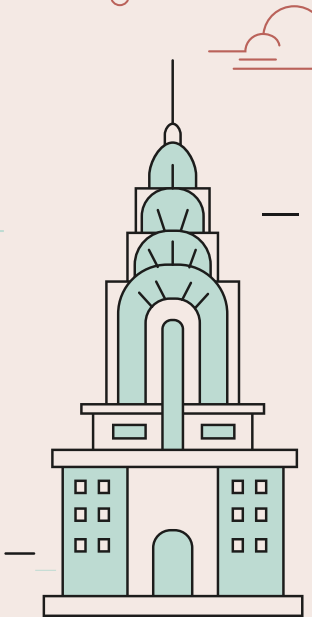
Price Prediction Model

Airbnb





Developed and Presented by Varshini Chellapilla for Springboard

Overview

My intention was to build a machine learning model that could predict a listing's price based on its features, utilizing exploratory data analysis, supervised learning, statistical inference, Thus, the aim of this model is to provide a highly-qualified aid to hosts as they decide their prices.

The questions posed at the beginning of the project are:

- What does the distribution of listing prices across New York look like?
- What are the factors that influence a listing's price?
- How does the price vary across different locations and types of dwellings?

.

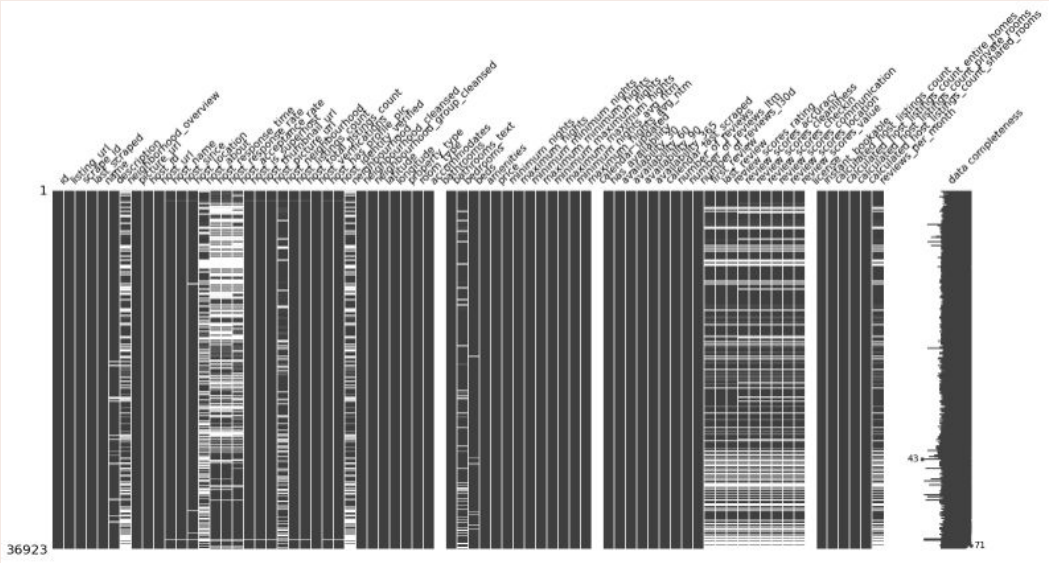
Data Acquisition and Wrangling

36,923

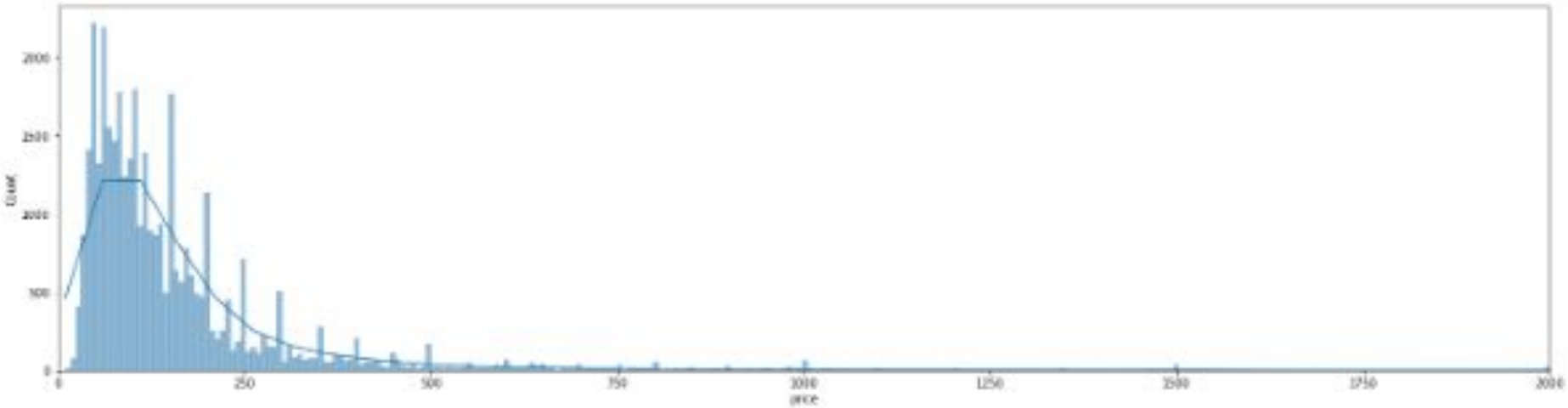
Number of listings in the original dataset

74

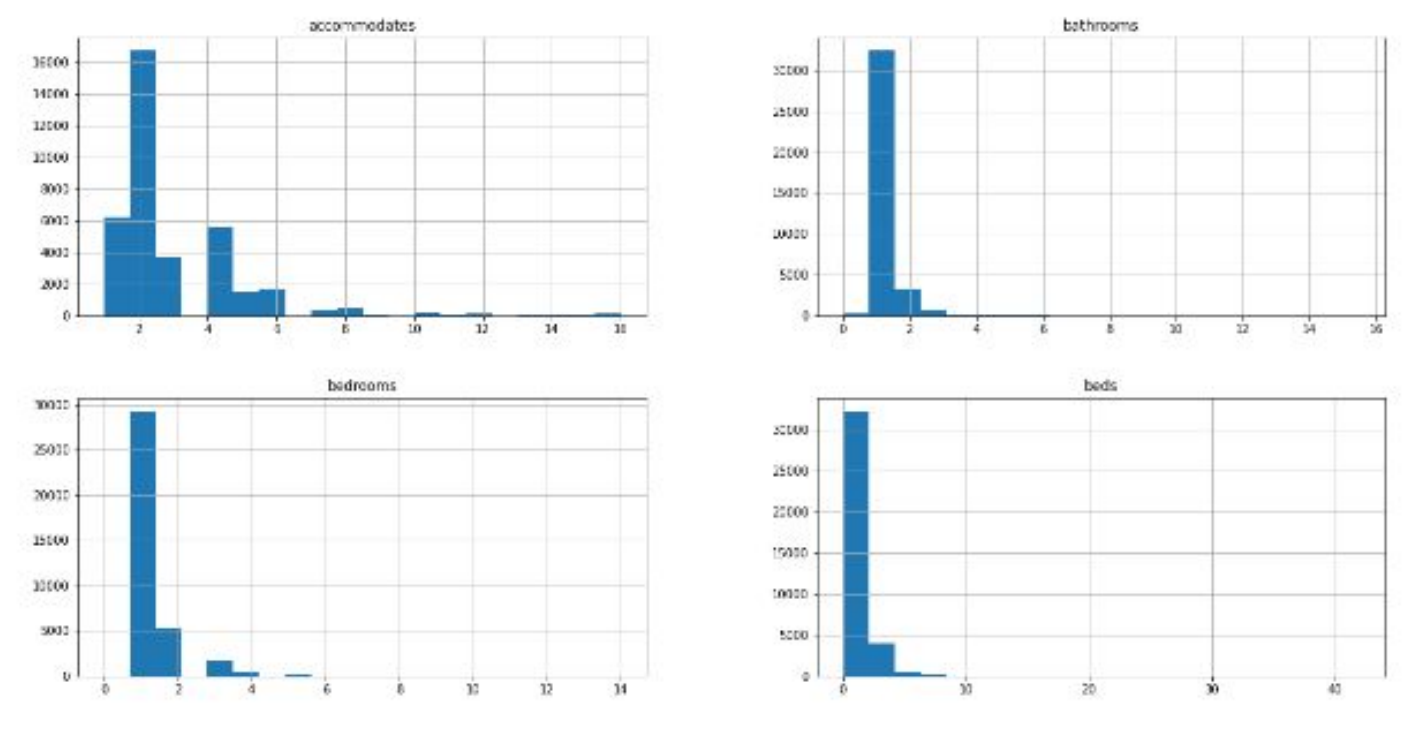
Number of features in the original dataset



Initial Distribution of Price



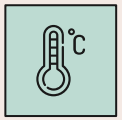
EDA: Amenities, Bathrooms, Bedrooms, and Beds



EDA: Amenities

Long-Term Stay

Kitchen

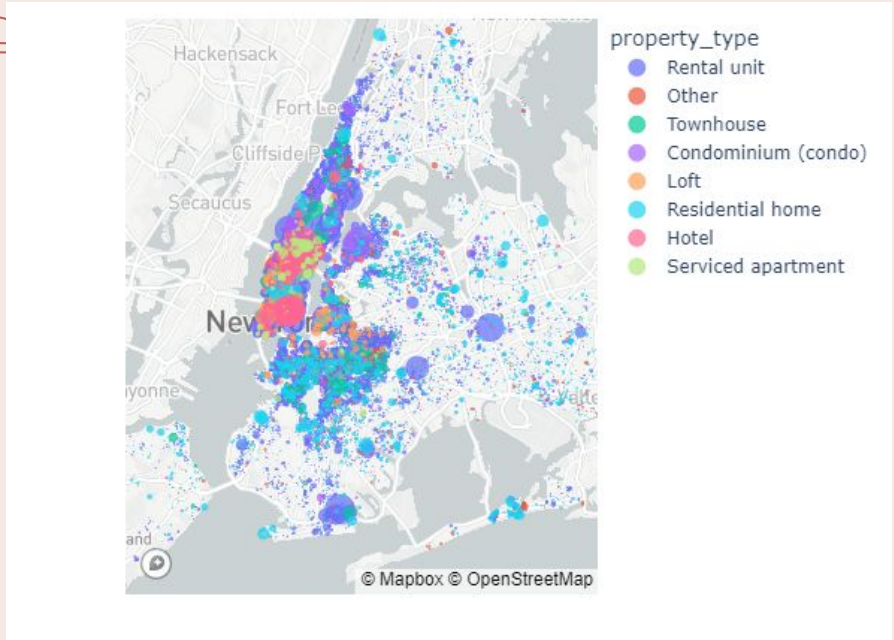


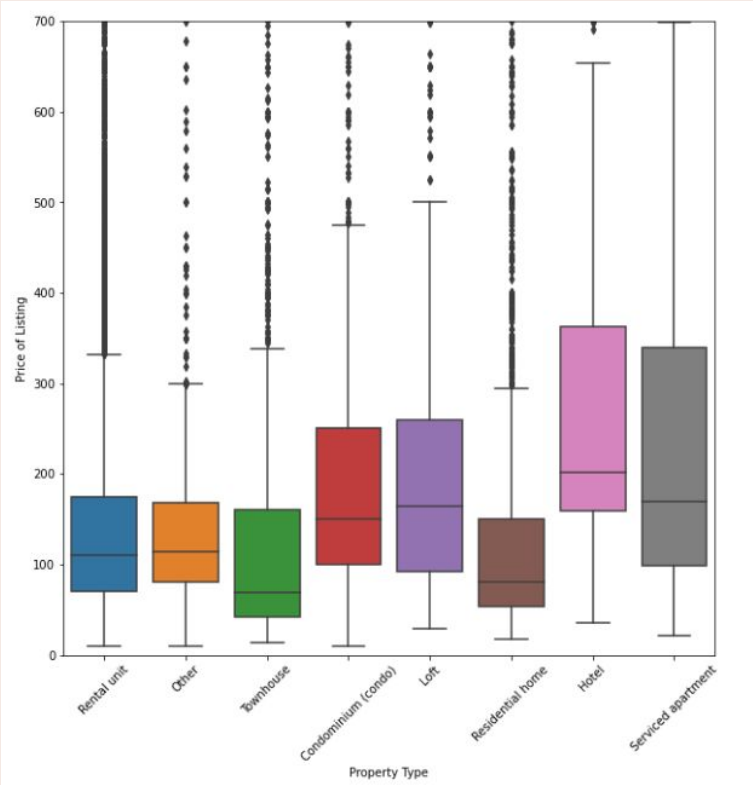
Air Conditioning

Hangers

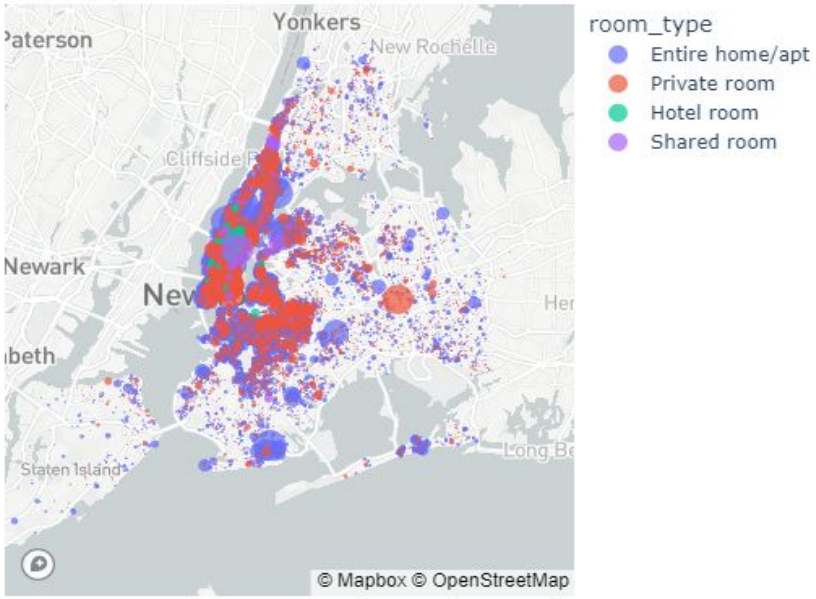
Free Street Parking

EDA: Property Type and Room Type

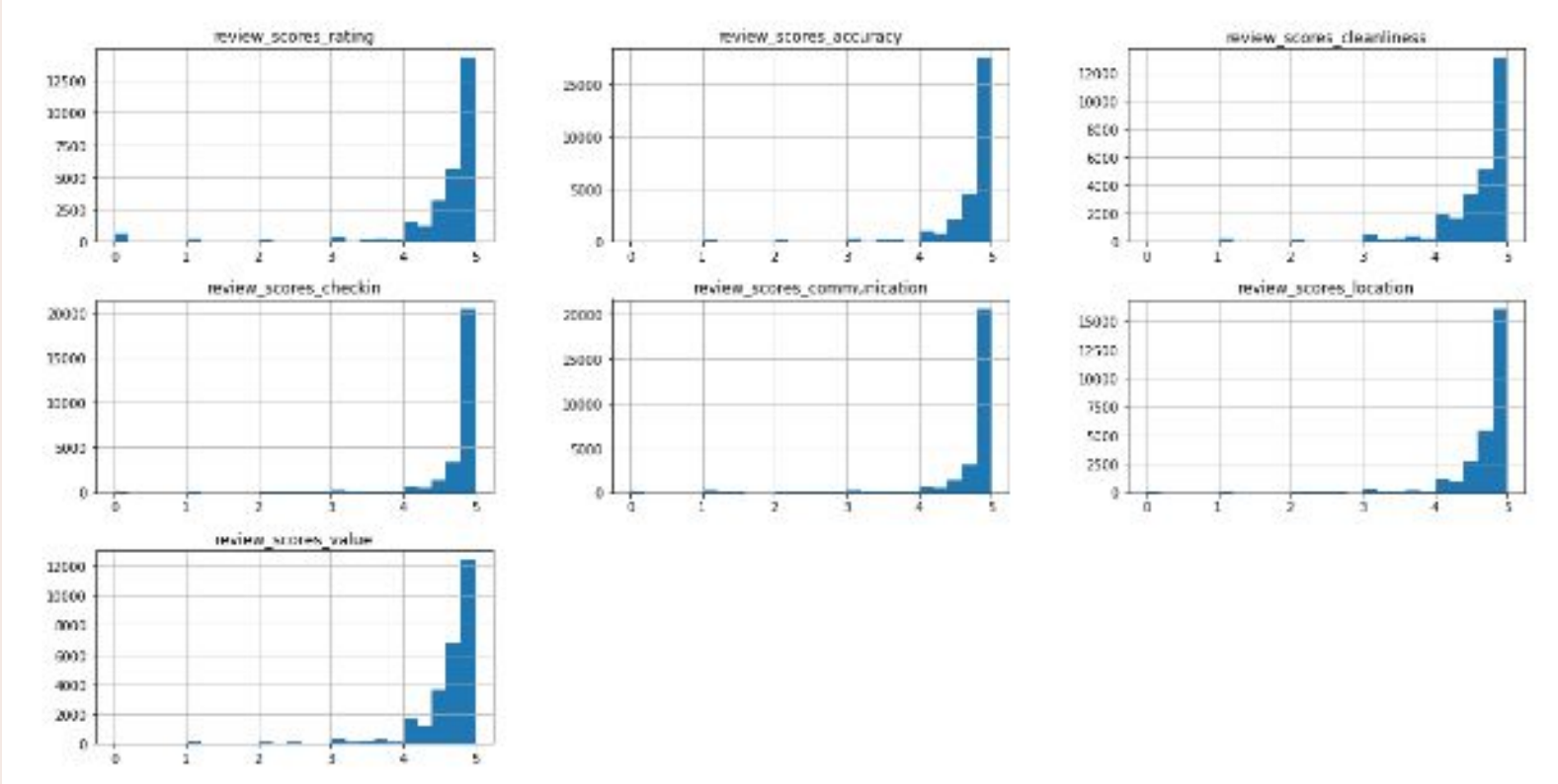




EDA: Room Type

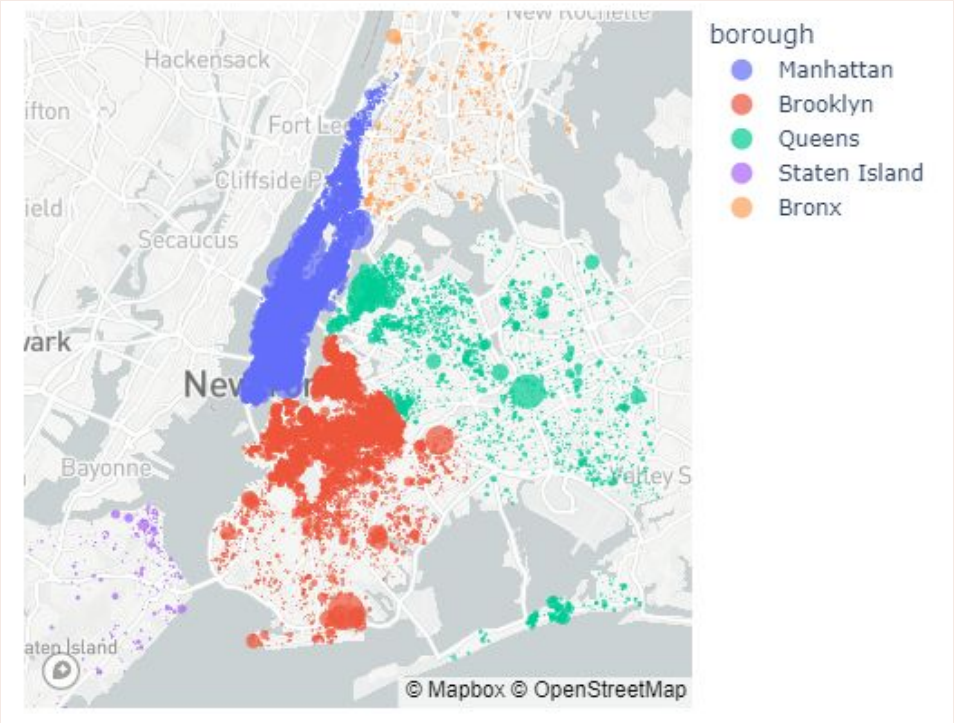


EDA: Reviews

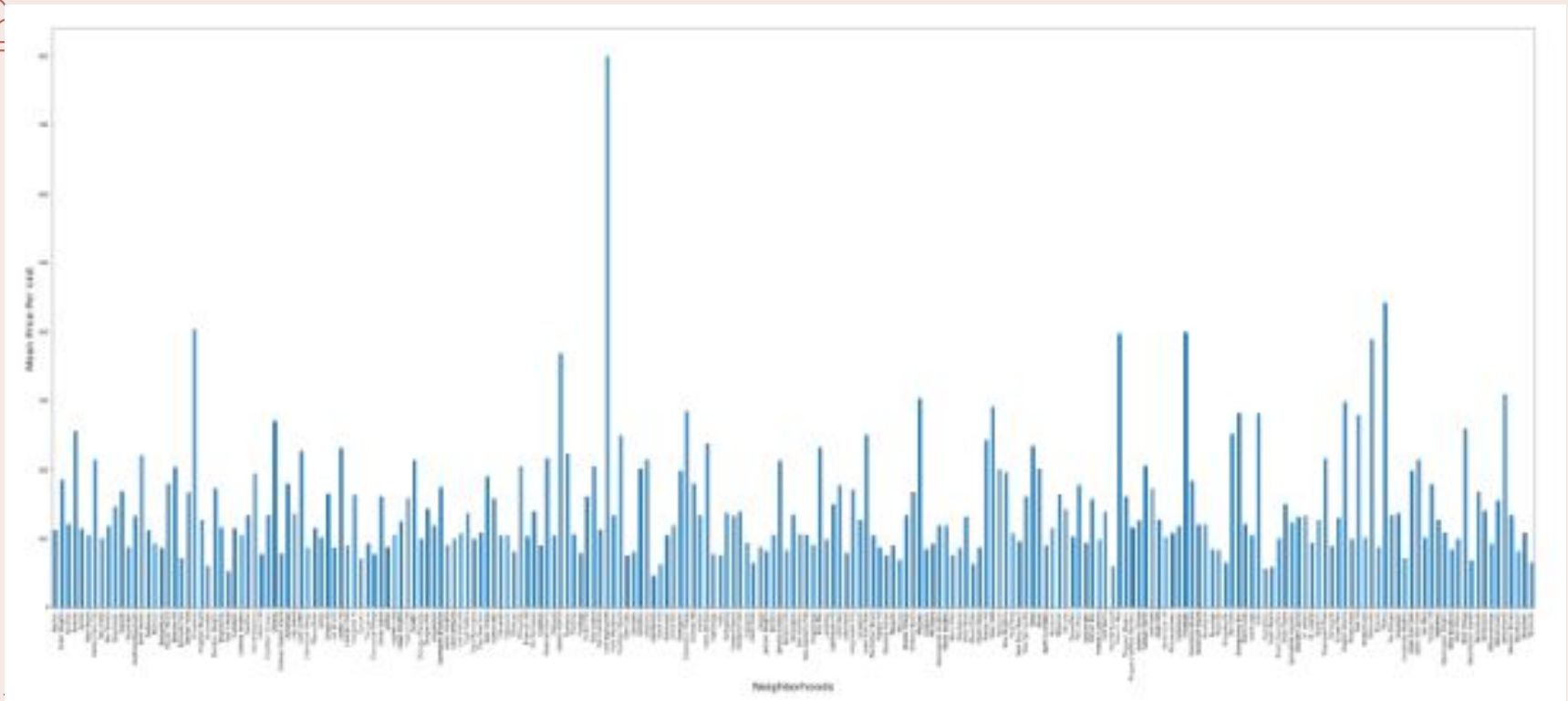

review_scores_rating
review_scores_accuracy
review_scores_cleanliness
review_scores_checkin
review_scores_communication
review_scores_location
review_scores_value

EDA: Location, Location, Location (Borough)

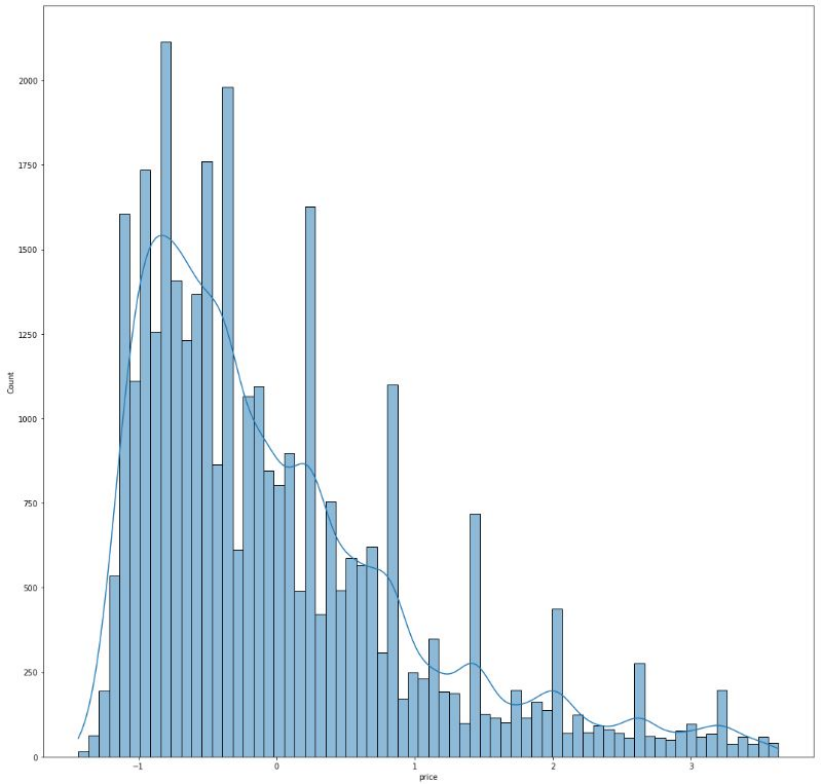
- Manhattan had the highest number of listings at 16,182, while Staten Island had the lowest with 313
- Manhattan once again claimed the highest value at $214, Bronx had the lowest at $104.
- Manhattan has the largest range of variability of prices while Bronx has the smallest. However, a large majority of the listings in each borough are listed at a price of less than $1,000.

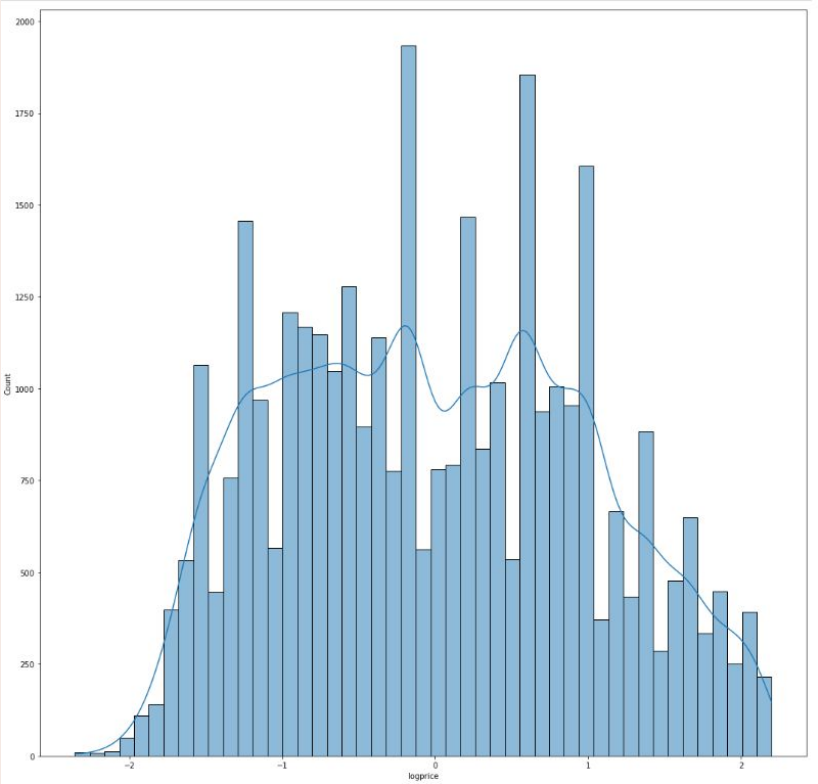


EDA: Location, Location, Location (Neighborhood)

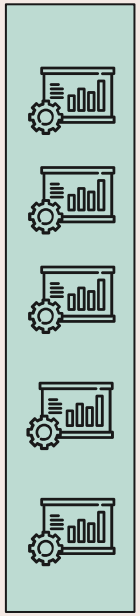


Pre-Processing: Target Variable (Price)





Model Selection



Linear Regression

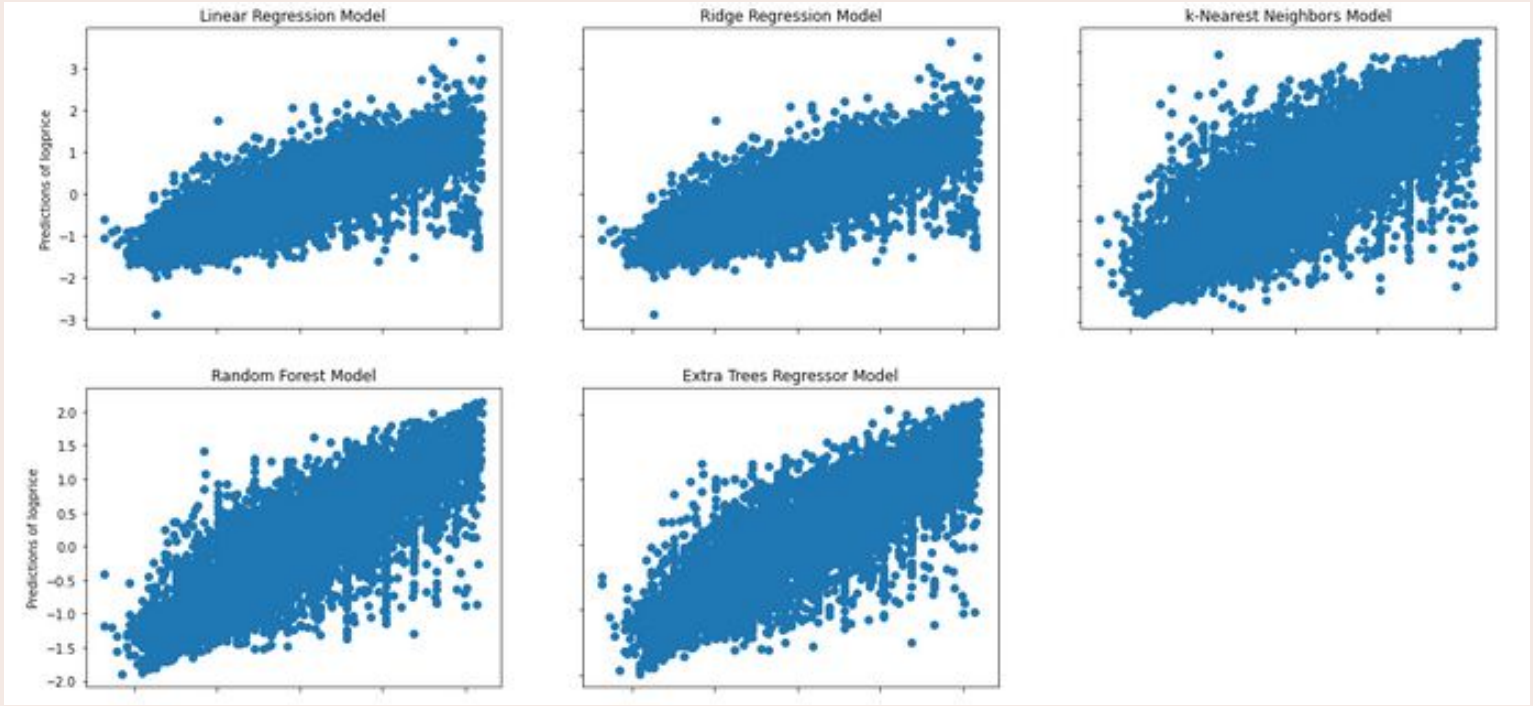
Ridge Regression

k-Nearest Neighbors

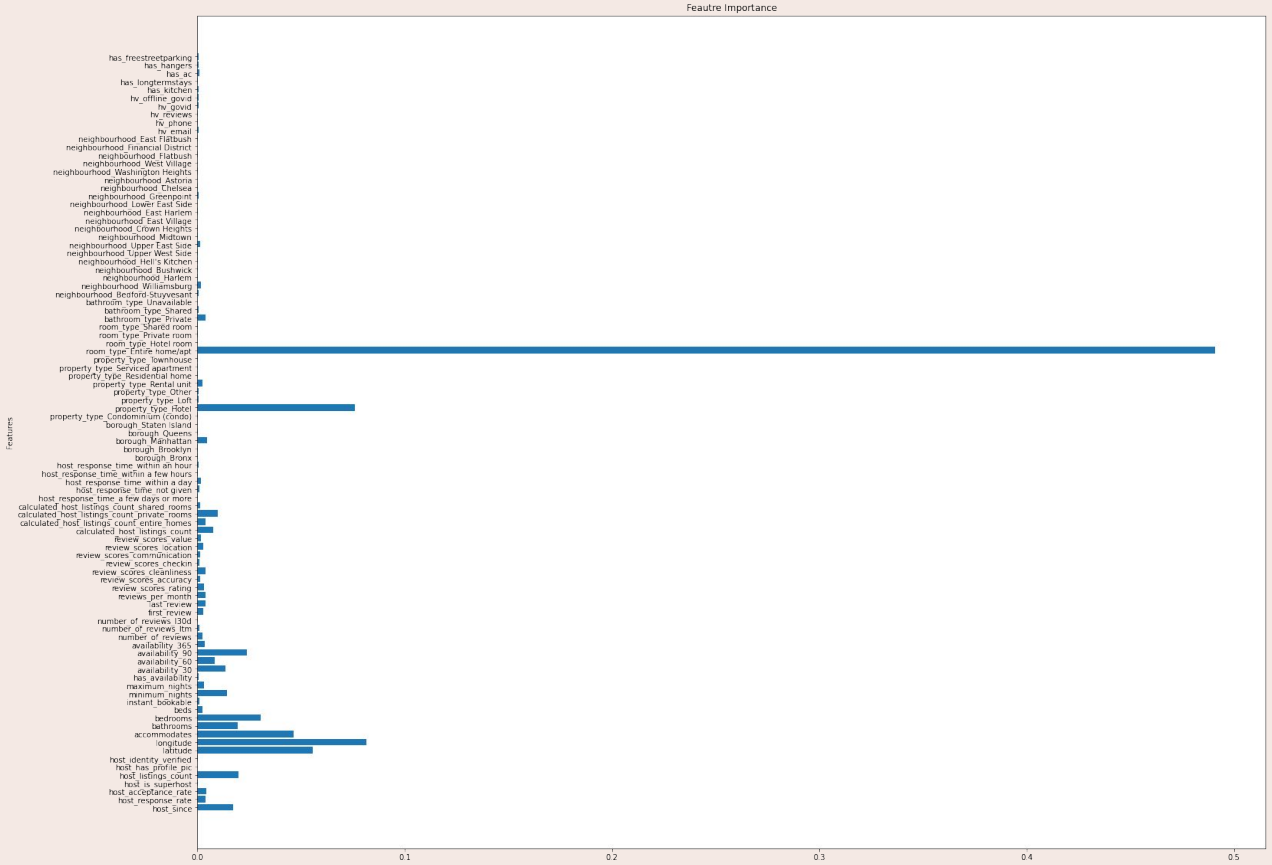
Random Forest

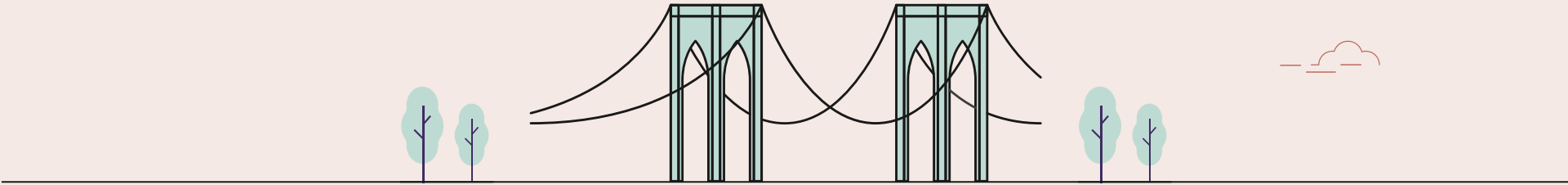
Extra Trees Regressor

Model Selection



Conclusion


Feautre Importance
Features
has_freestreetparking
has_hangers
has_ac
has_longtermstays
has_kitchen
hv_offline_govid
hv_govid
hv_reviews
hv_phone
hv_email
neighbourhood_East Flatbush
neighbourhood_Financial District
neighbourhood_Flatbush
neighbourhood_West Village
neighbourhood_Washington Heights
neighbourhood_Astoria
neighbourhood_Chelsea
neighbourhood_Greenpoint
neighbourhood_Lower East Side
neighbourhood_East Harlem
neighbourhood_East Village
neighbourhood_Crown Heights
neighbourhood_Midtown
neighbourhood_Upper East Side
neighbourhood_Upper West Side
neighbourhood_Hell's Kitchen
neighbourhood_Bushwick
neighbourhood_Harlem
neighbourhood_Williamsburg
neighbourhood_Bedford-Stuyvesant
bathroom_type_Unavailable
bathroom_type_Shared
bathroom_type_Private
room_type_Shared room
room_type_Private room
room_type_Hotel room
room_type_Entire home/apt
property_type_Townhouse
property_type_Serviced apartment
property_type_Residential home
property_type_Rental unit
property_type_Other
property_type_Loft
property_type_Hotel
property_type_Condominium (condo)
borough_Staten Island
borough_Queens
borough_Manhattan
borough_Brooklyn
borough_Bronx
host_response_time_within an hour
host_response_time_within a few hours
host_response_time_within a day
host_response_time_not given
host_response_time_a few days or more
calculated_host_listings_count_shared_rooms
calculated_host_listings_count_private_rooms
calculated_host_listings_count_entire_homes
calculated_host_listings_count
review_scores_value
review_scores_location
review_scores_communication
review_scores_checkin
review_scores_cleanliness
review_scores_accuracy
review_scores_rating
reviews_per_month
last_review
first_review
number_of_reviews_l30d
number_of_reviews_ltm
number_of_reviews
availability_365
availability_90
availability_60
availability_30
has_availability
maximum_nights
minimum_nights
instant_bookable
beds
bedrooms
bathrooms
accommodates
longitude
latitude
host_identity_verified
host_has_profile_pic
host_listings_count
host_is_superhost
host_acceptance_rate
host_response_rate
host_since
0.0
0.1
0.2
0.3
0.4
0.5



THANKS!

CREDITS: This presentation template was created by **Slidesgo,** including icons by **Flaticon,** and infographics & images by **Freepik**