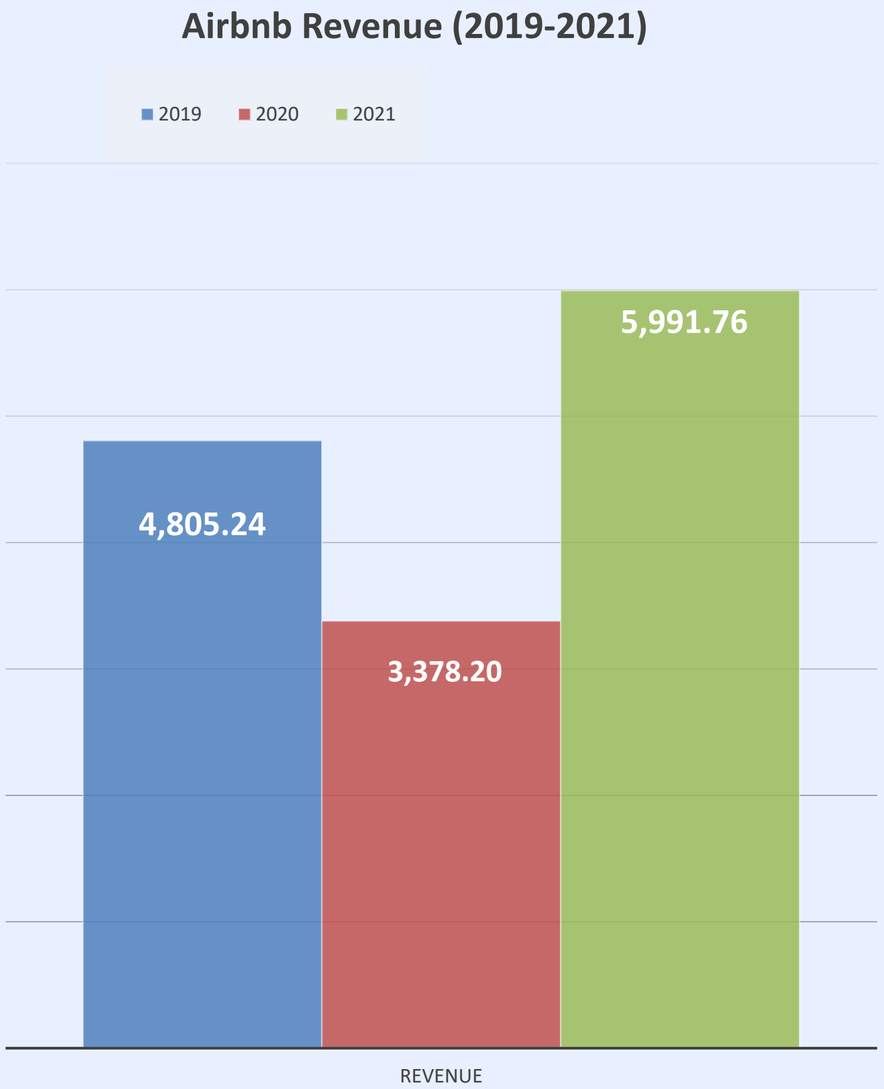
Understanding The Guest Preferences To Stay Ahead of The Game

By Hosain Ahmed

Agenda

- Business Objectives
- Project and Data Assumptions
- Getting to know the dataset
- Modeling Recommendations
- Insights of the findings
- Recommendations

Business Objectives



Economics of Airbnb

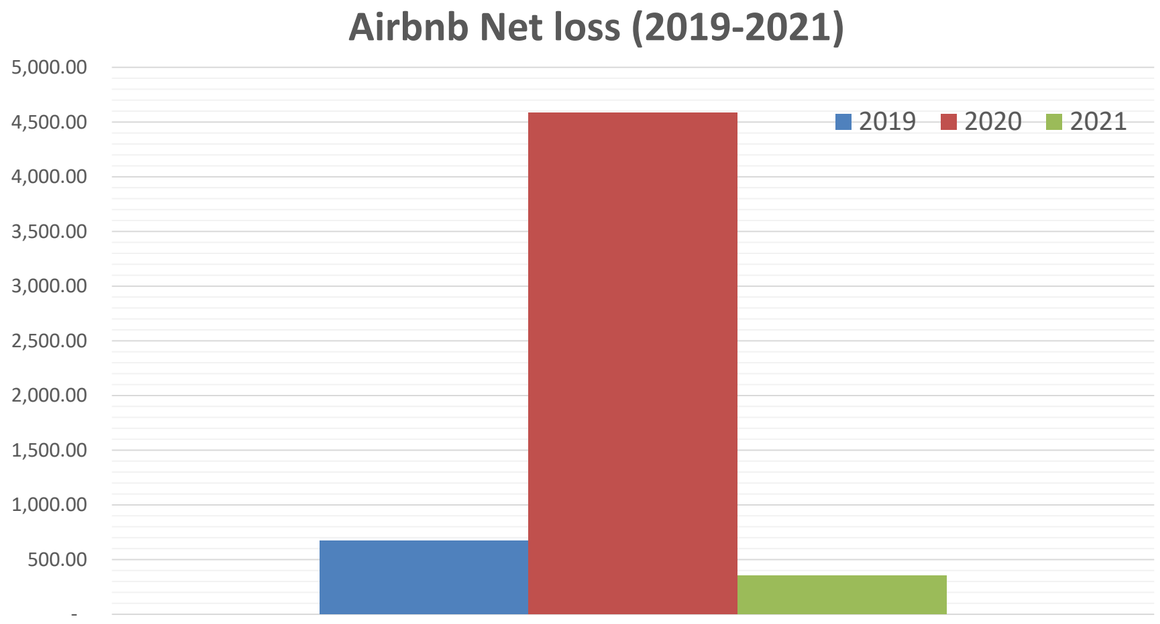
Scenario-1: Devastation of Covid Pandemic

- Airbnb hit badly due to covid pandemic:
  - Revenue decreased by 30% in 2020.
  - Net Loss increased by 3910 million only in 2020.

Scenario-2: Recovery and Glimpse of Hope

- Revenue situation recovered thanks to quarantine and Work-From-Home trends.

Business Objectives Contd.



Scenario-3: Recession looms

- Revenue stream recovered in 2021:
  - New normal trends like work from home still proves to be beneficial for the company.
  - However, recent global political unrest halted economic reform and threatened with a recession.
  - Hence, Airbnb needs to understand and ensure basic the customer preferences to boost revenue during economic hardship.

- **Therefore, Airbnb needs a model to predict the basic guests' preferences while booking an Airbnb.**

Project and Data Assumptions

- **Dataset on Hand**
  - Customer data of New York City in 2019.
- **Assumptions**
  - New York is the perfect example of cultural melting pot.
  - It is one of the most visited cities by the tourists all over the world.
  - New York is also known for few unique tourism i.e. birth tourism and immigration tourism.
  - Data of 2019 will give us the basic booking preferences of the guests.

How did we get to the model

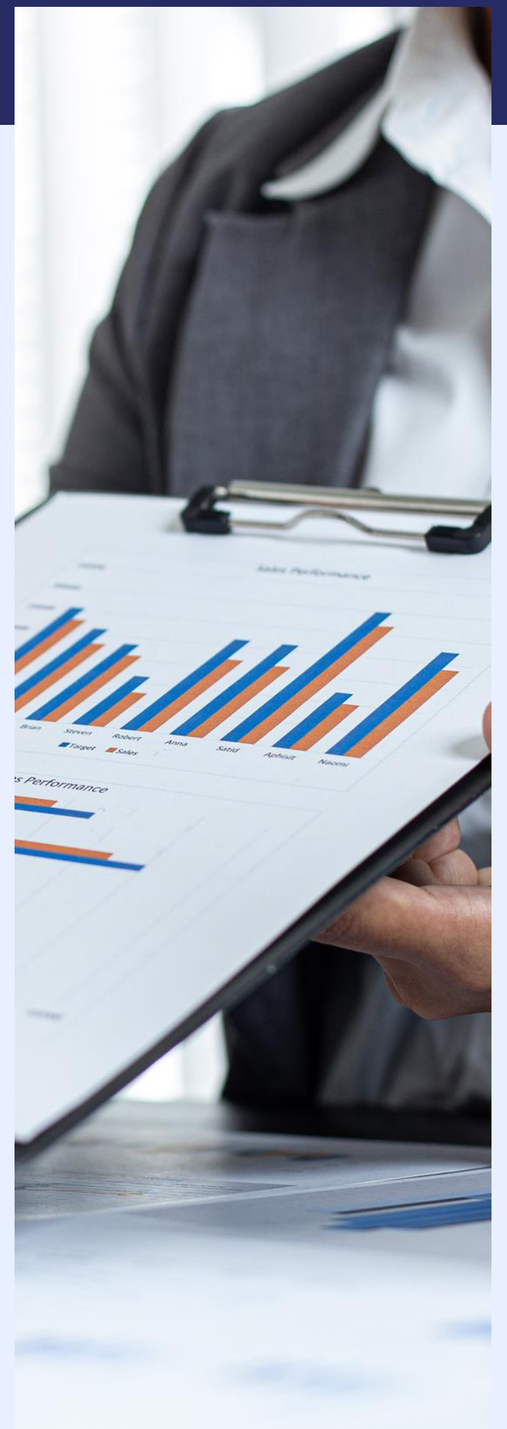
- 4 machine learning models were tested, and the best model was picked based on its ability to predict the guest preferences.
- Root-mean-square error or RMSE score was used to compare the models.
- Summary of all the model used are given in the Appendix.

Features Considered in Building the Model

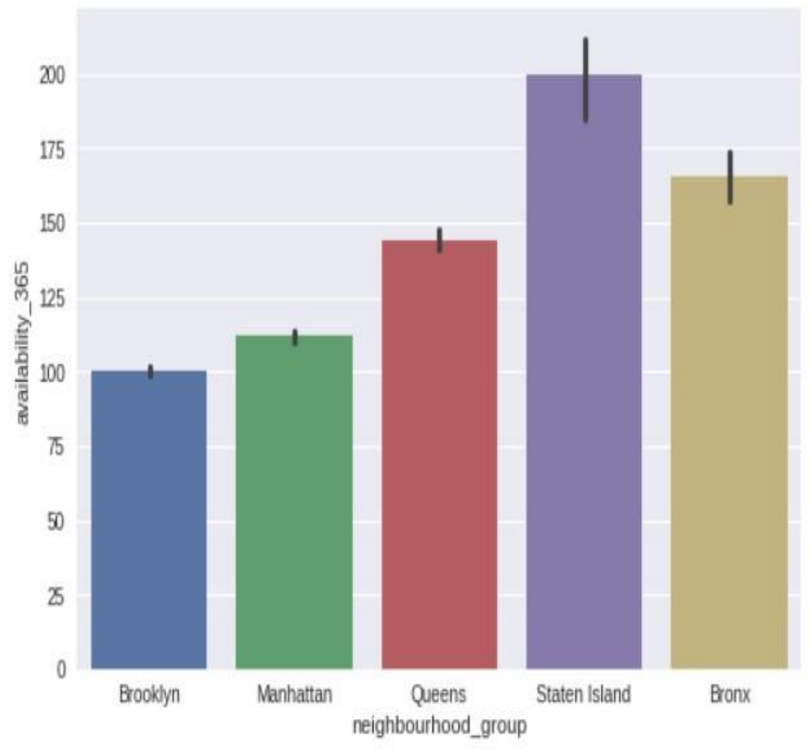
Target: 'availability_365' – Days property was available for booking.

List of variables considered against the target:

- Location Based Features:
  - 'neighbourhood_group'
  - 'neighbourhood'
  - 'latitude'
  - 'longitude'
- Property Related Amenities:
  - 'room_type'
  - 'price'
- Others Features:
  - 'calculated_host_listings_count_group'
  - 'minimum_nights_group'
  - 'number_of_reviews'

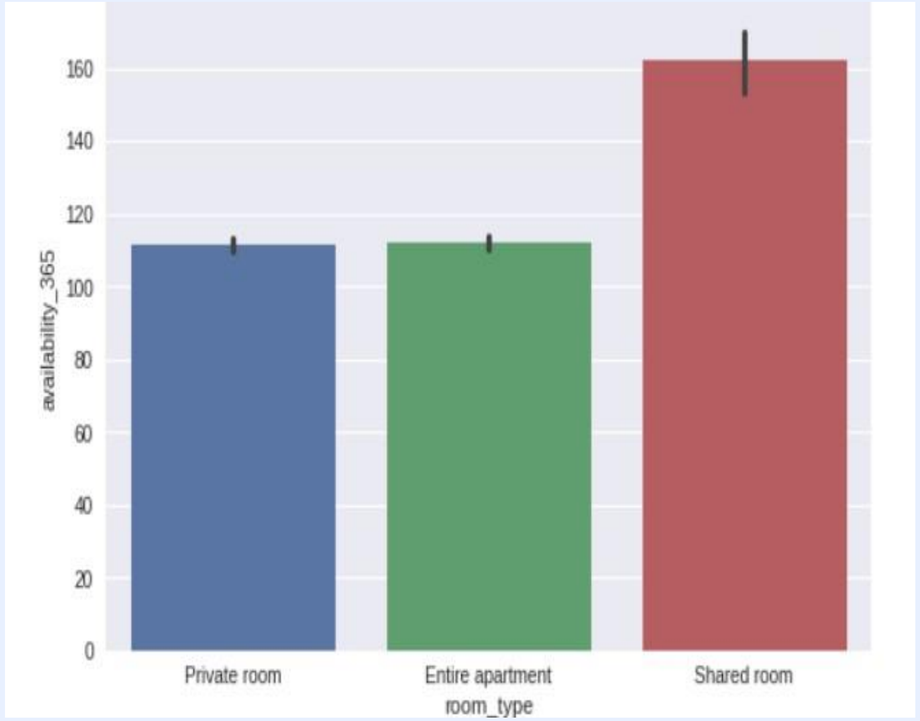


Correlation among features: Location with property demand

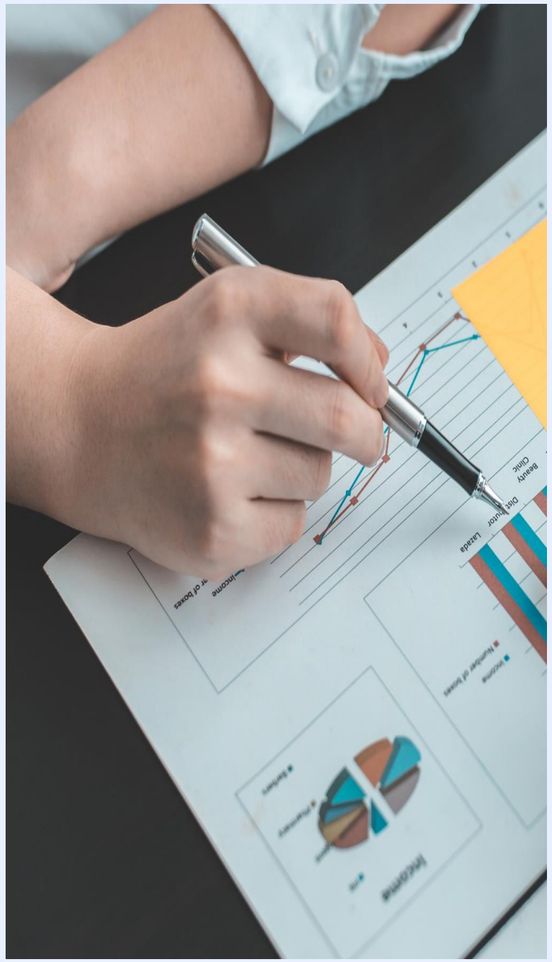


- **Brooklyn and Manhattan are the most preferred location.**

Correlation among features: Room Type with property demand

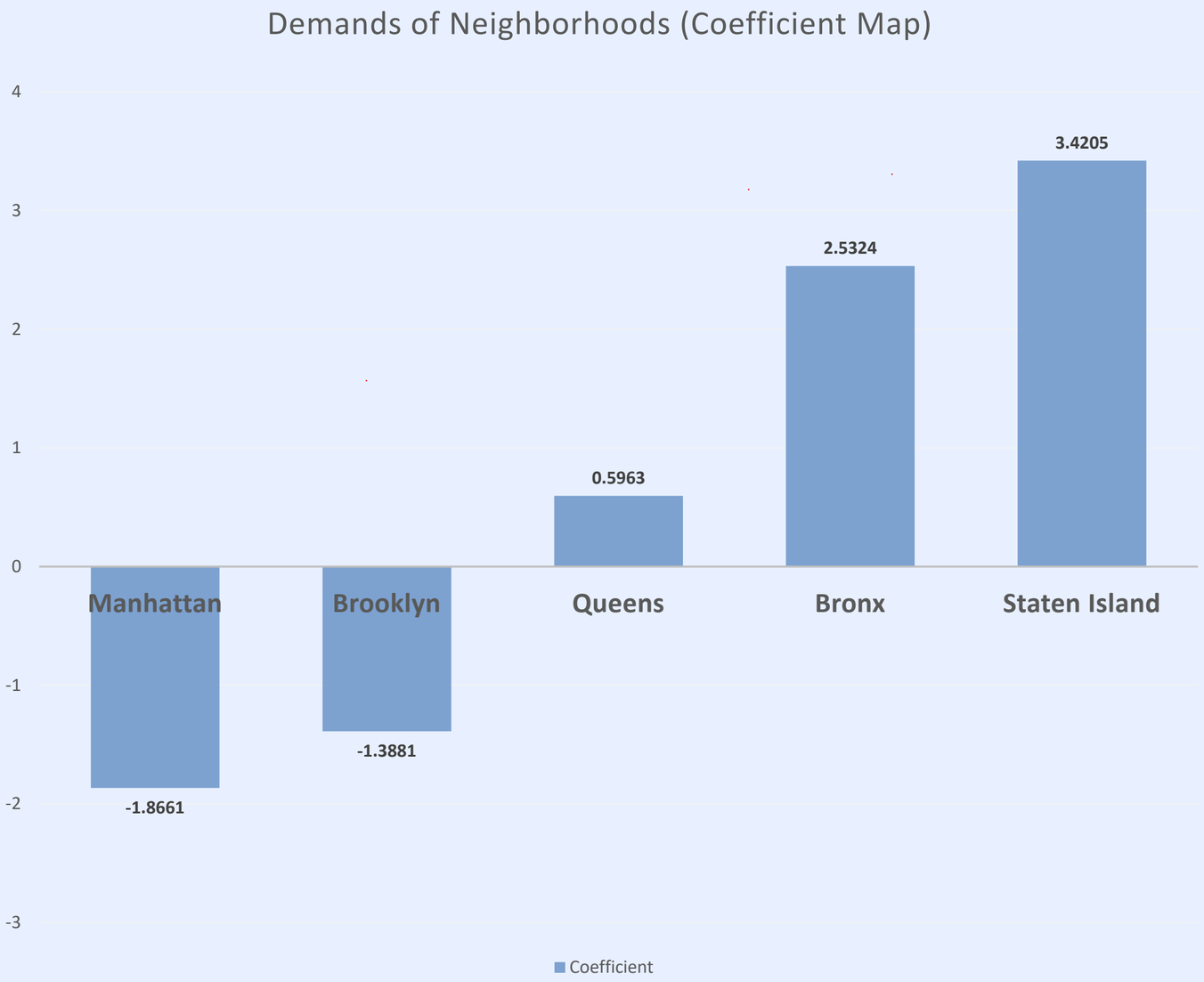


- **Entire Apartments are on demand as people value privacy.**



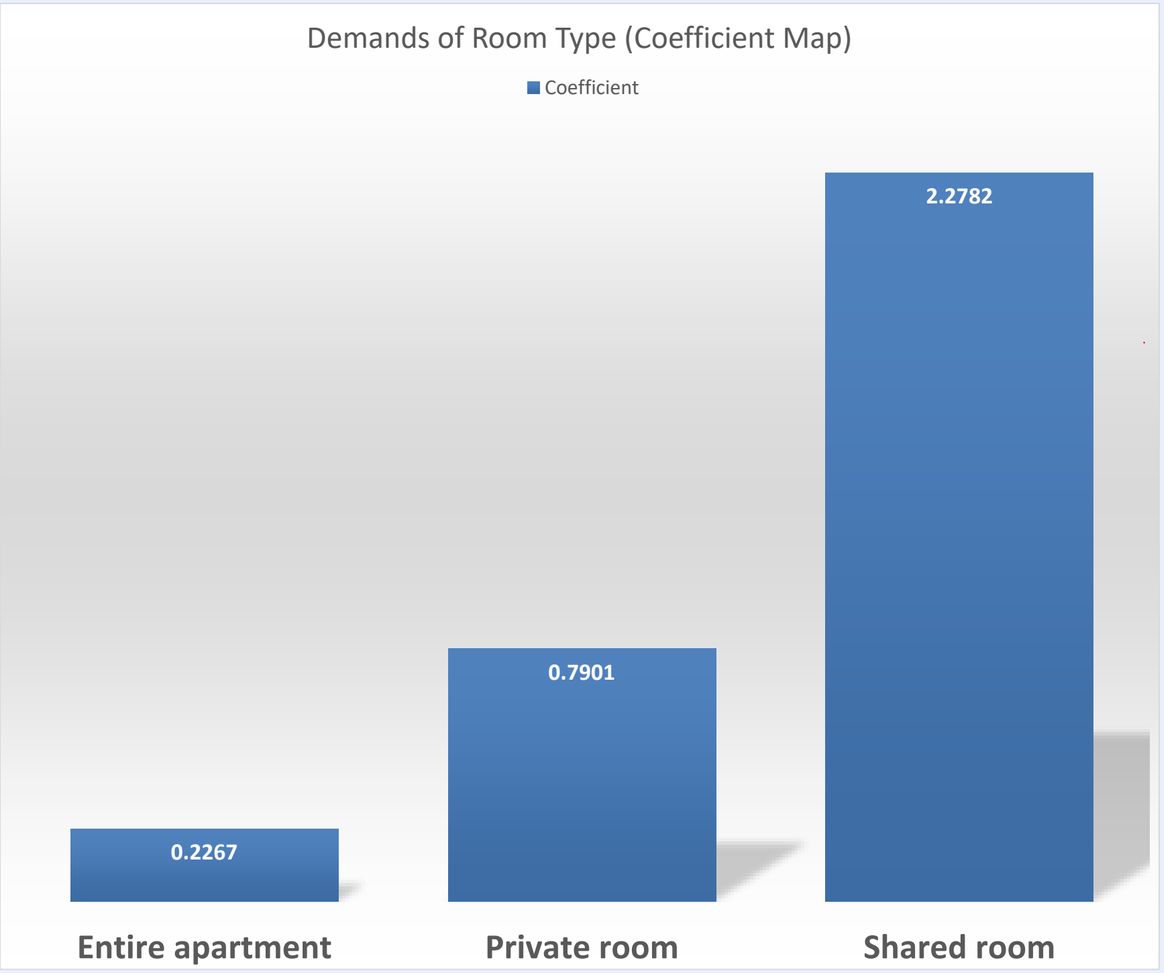
The Best Model: Multiple Linear Regression

- The Model has the best Root-mean-square error score of only 6.22
- The score means the error of the model in predicting customer preference is only 6.22%



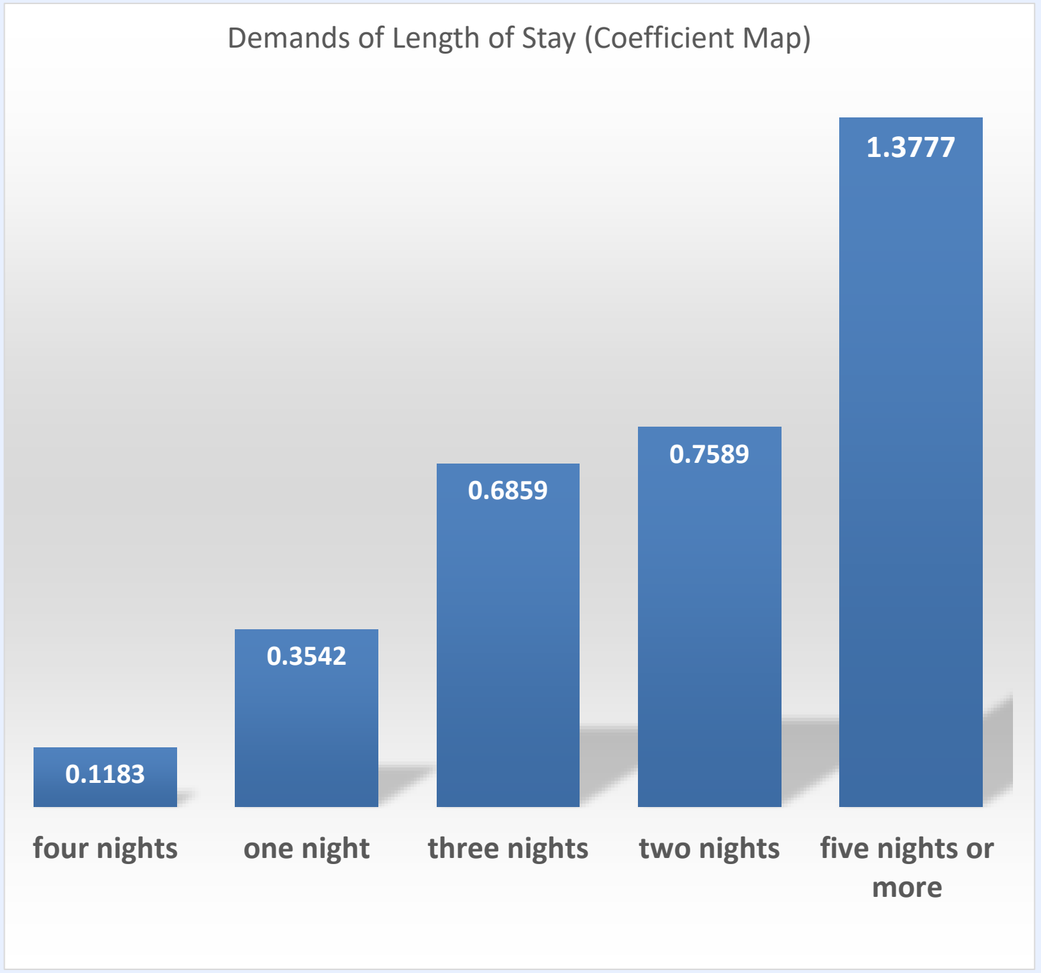
Best Features from the Model-1

- Manhattan and Brooklyn is the most favorite place.
- However, Queens holds the third position.



Best Features from the Model-2

- **People Likes Entire Apartment but avoids shared room.**



Best Features from the Model-3

- People likes to stay for longer period at a cheaper rate
- Coefficient of price with demand is 0.2845, which means demand decreases with price.

Overall Insights of the Findings

- Insights that we can the model tells us:
- People likes to stay I around downtown area where transport and other civil facilities are convenient.
- People prefers to stay for long time at cheaper rates.
- People value their privacy.

Recommendations

- Queens area has demand but its lower than expectations.
  - Probable Reasons:
    - Under the table bookings for longer stays for more than 30days. Usually practiced in situations like birth tourism.
    - Solution: Company must offer extra benefits to the hosts for longer bookings for more than 30days.
- Considering the popularity of Manhattan and Brooklyn Airbnb should acquire more offerings close to subways stations and superstores.
- Should shrink property offering in area like Bronx and Staten Island.
- Offers discounts to encourage longer stays for more than 7days.

Recommendations Contd.

- **Further Analysis:** Airbnb can use text mining of the guests reviews to find out more on privacy and feature preference issues which was beyond the scope of this dataset.
- **Word of Caution**: Since the guest preferences are ever changing therefore Airbnb should monitor the results of this model in every 3 months and consider rebuilding the model if the RMSE score drifts for more than 25% from the original results.

Appendix: Summary of all the models

Model Name	Training Accuracy	Validation Accuracy	RMSE Score (Validation)
Multiple Linear Regression	**27%**	**26%**	**6**
Random Forest Regressor	85%	25%	116
K neighbors Classifier	50%	31%	144
Neural Network	36%	36%	174

THANK YOU