

# **Airbnb Feature Importance Analysis & Recommendation System**

**By Nora Adadurova**

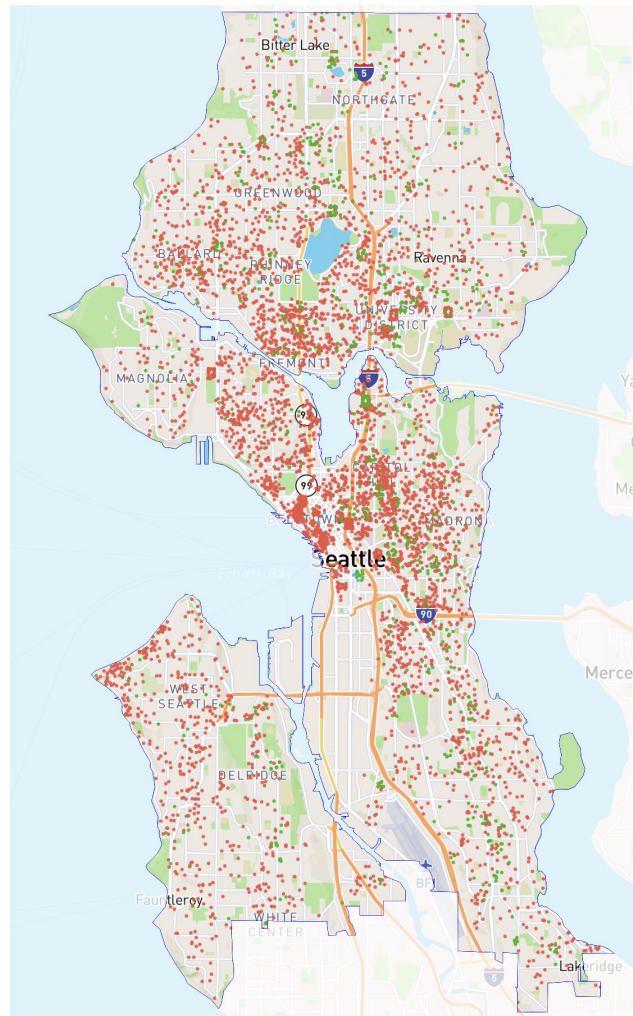
# Itinerary

- Introduction to Data
- Problem Statement
- Data Cleaning
- EDA
- Correlation Analysis
- Text Data EDA
- Feature Engineering
- Feature Importance Analysis
- Recommendation System
- Conclusion & Recommendations



# Dataset Introduction

- Official Airbnb Data
- <http://insideairbnb.com/get-the-data/>
- Seattle, Washington, United States
- 24 March, 2023



# Problem Statement

1. Identify features of importance of Airbnb listings in Seattle, WA.
2. Investigate demand patterns by looking at the relationship between availability and booking-related metrics.
3. Develop a recommendation system to assist future renters in Seattle, WA, in finding listing similar by neighbourhood, room type, or minimum nights and at a comfortable price or availability.

# Data Cleaning



```
graph LR; A[Missing Values] --> B[Duplicates]; B --> C[Evaluating Data]; C --> D[Outliers];
```

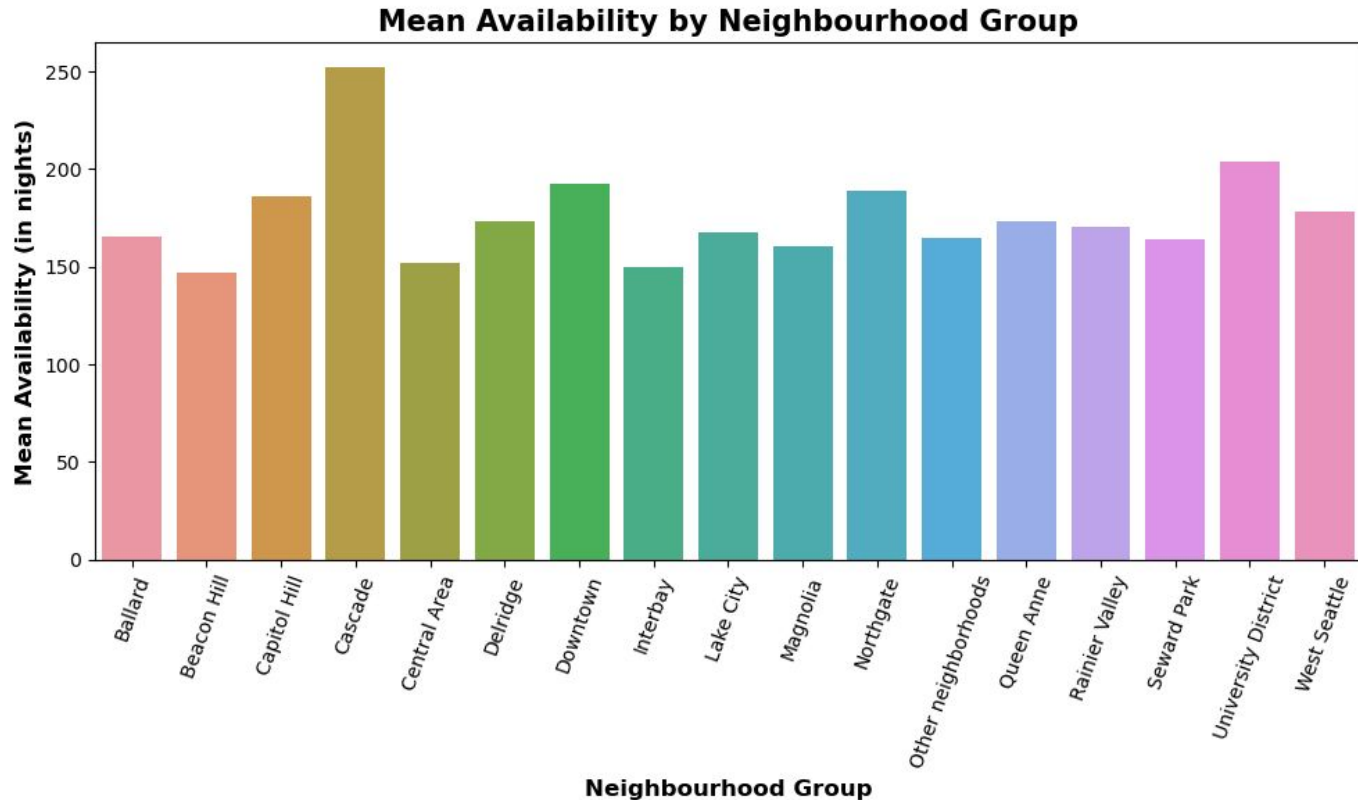
Missing  
Values

Duplicates

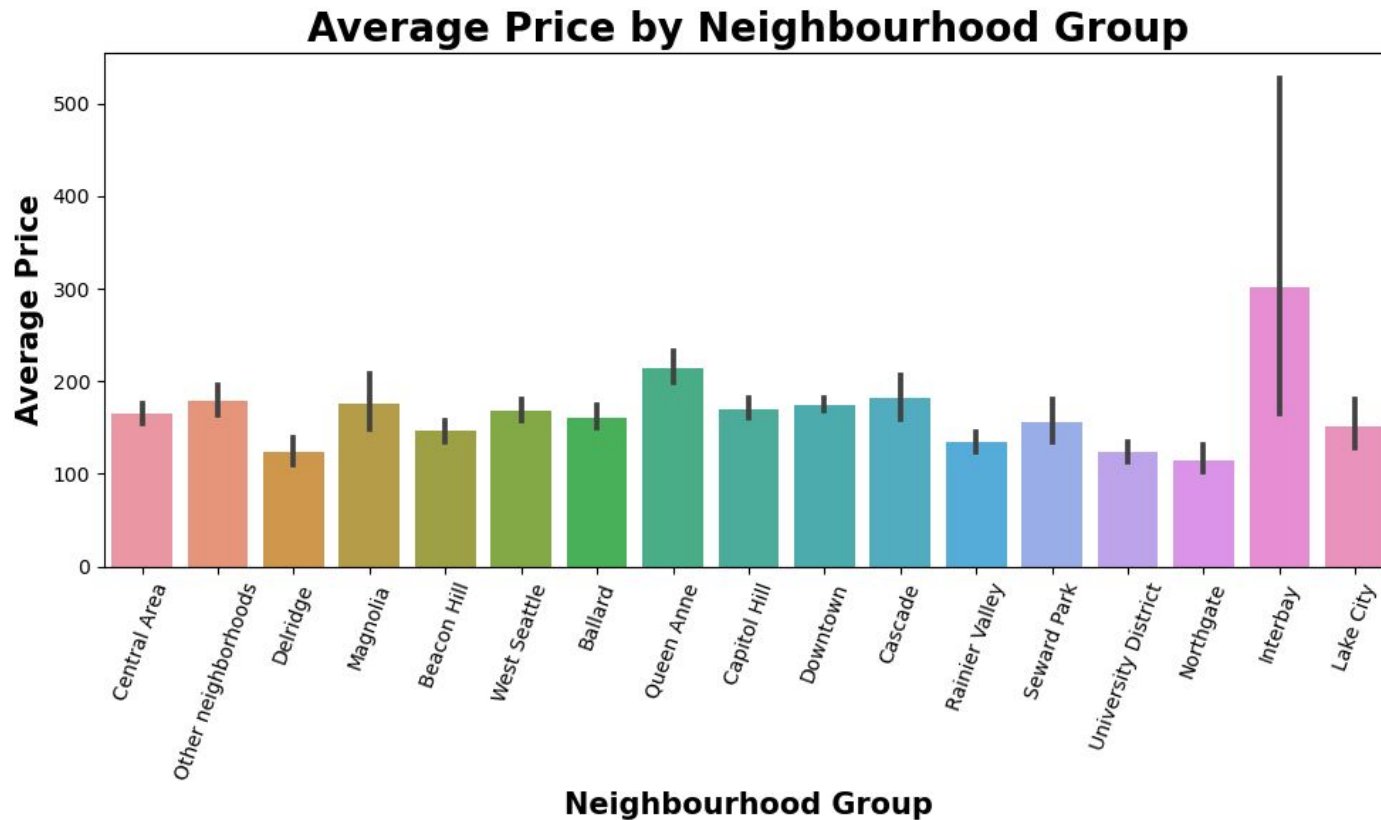
Evaluating  
Data

Outliers

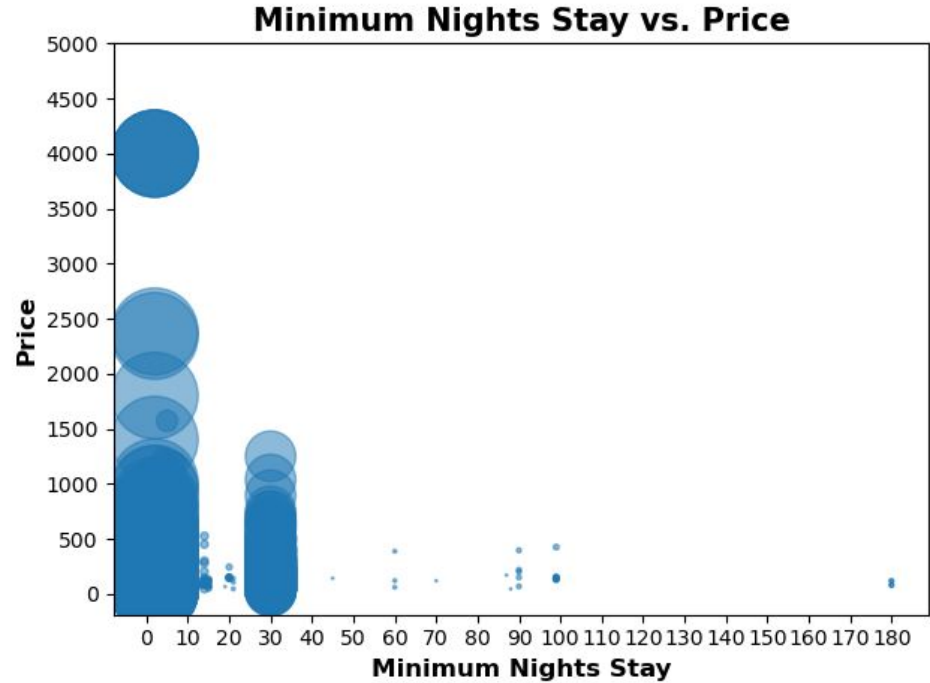
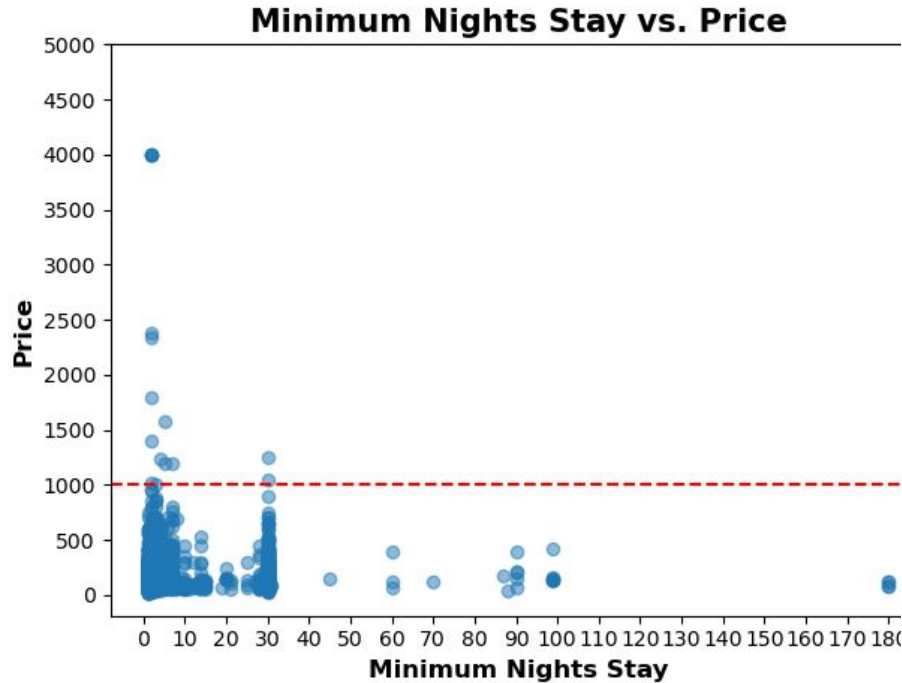
# EDA: Availability vs. Neighbourhood Groups



# EDA: Price vs. Neighbourhood Groups

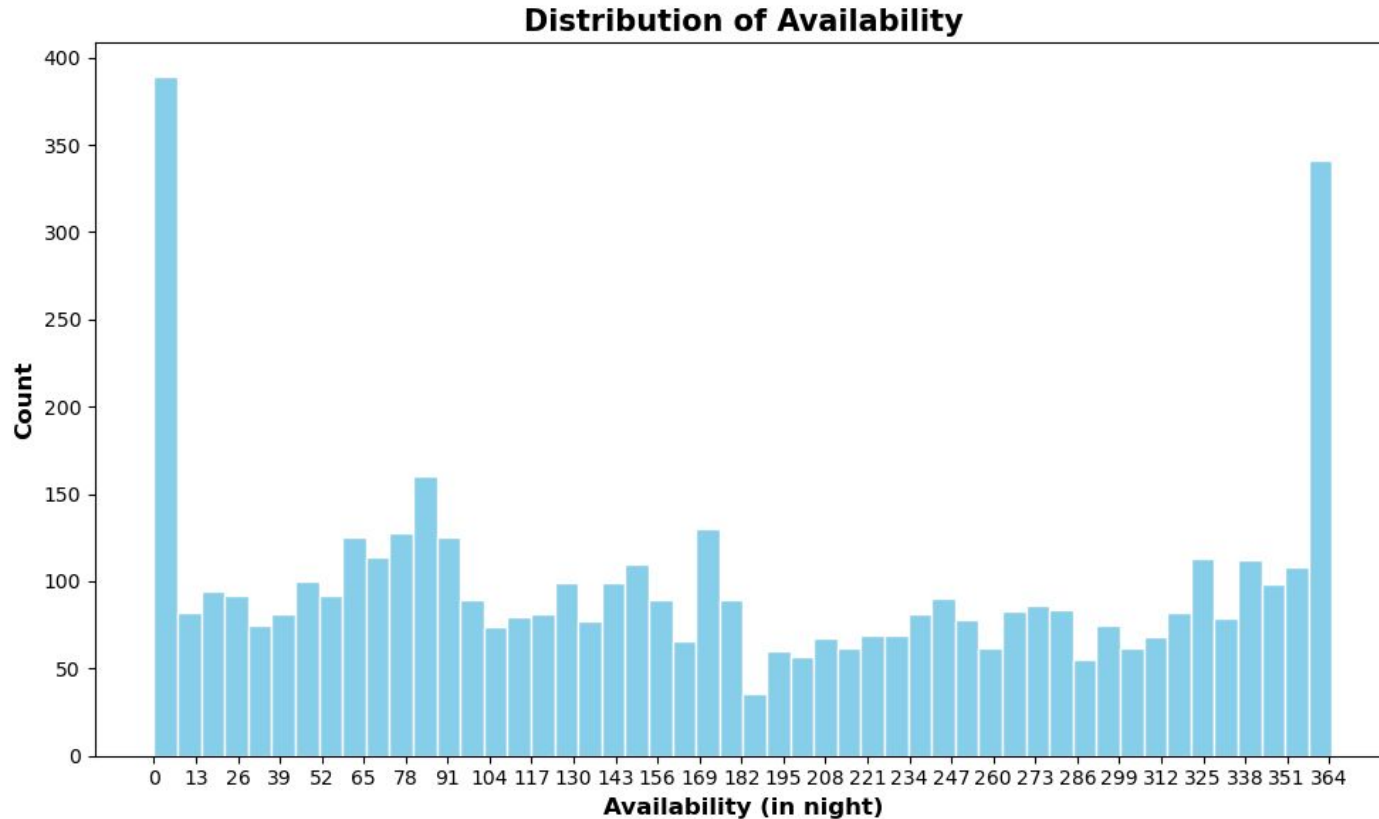


# EDA: Number of Minimum Nights vs. Price

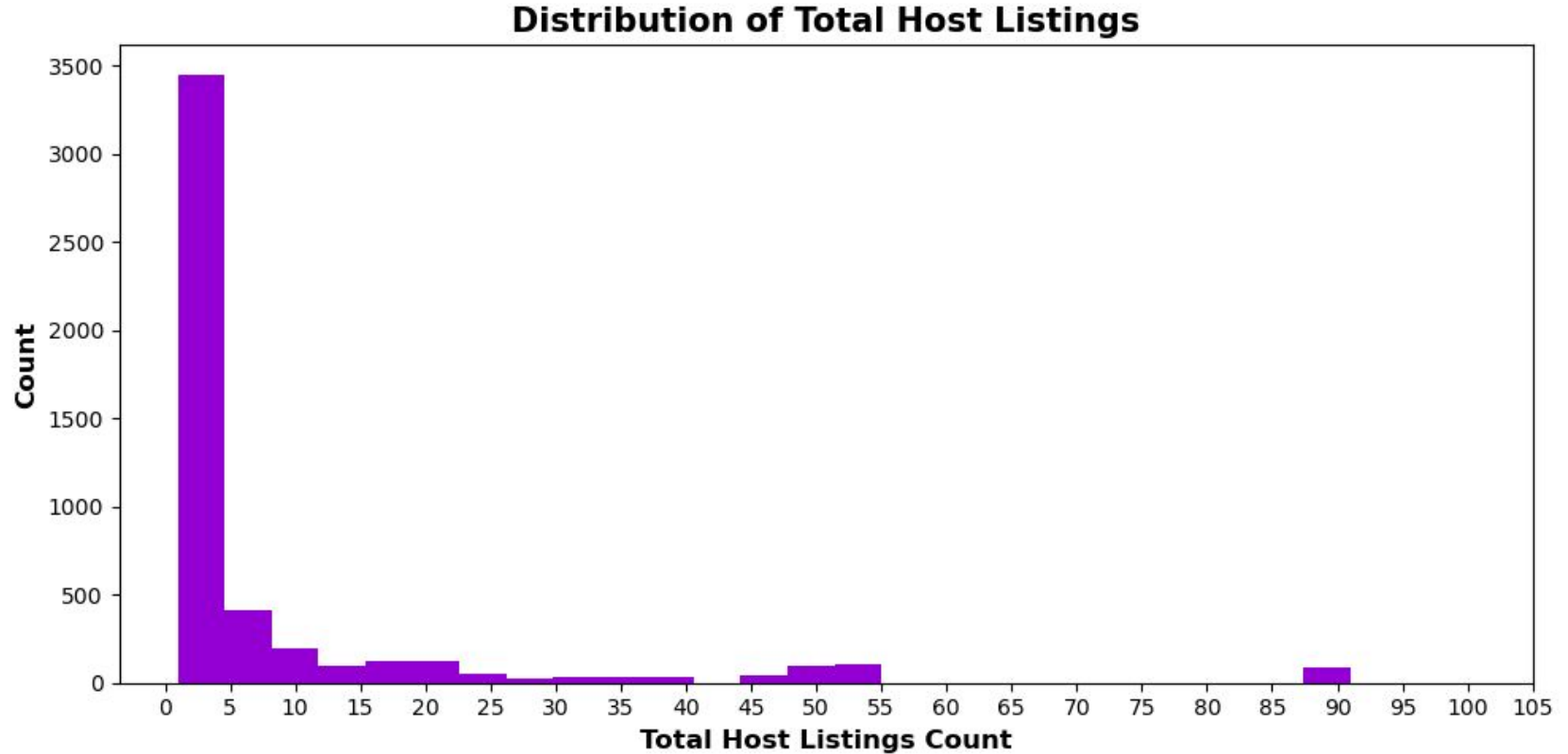




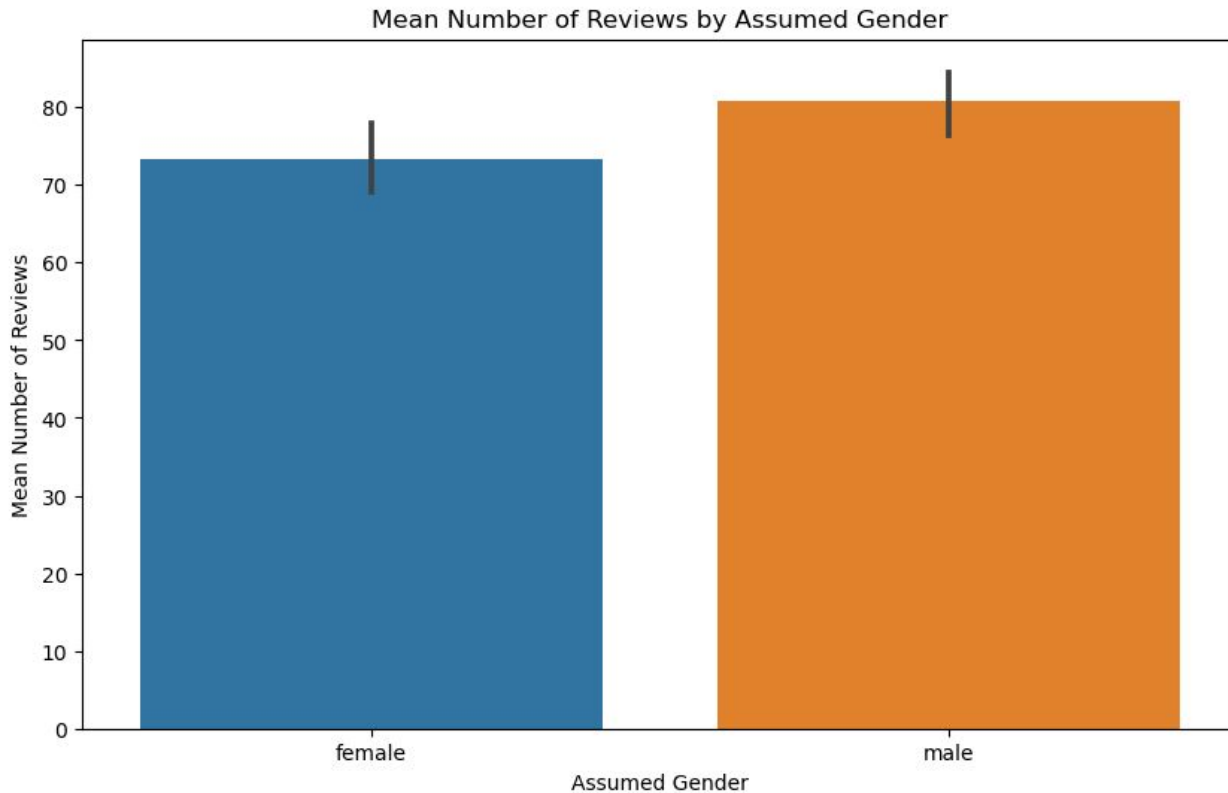
# EDA: Availability



# EDA: Total Listings per One Host



# EDA: Gender Analysis



# Text Data EDA



# Text Data EDA: Cheapest vs. Expensive

### Word Cloud for Cheapest Listings



### Word Cloud for Most Expensive Listings



# Feature Engineering



```
graph LR; A[Adding Columns] --> B[Transforming Columns]; B --> C[Scaling]; C --> D[Changing Categories]; D --> E[Choosing Features]
```

Adding  
Columns

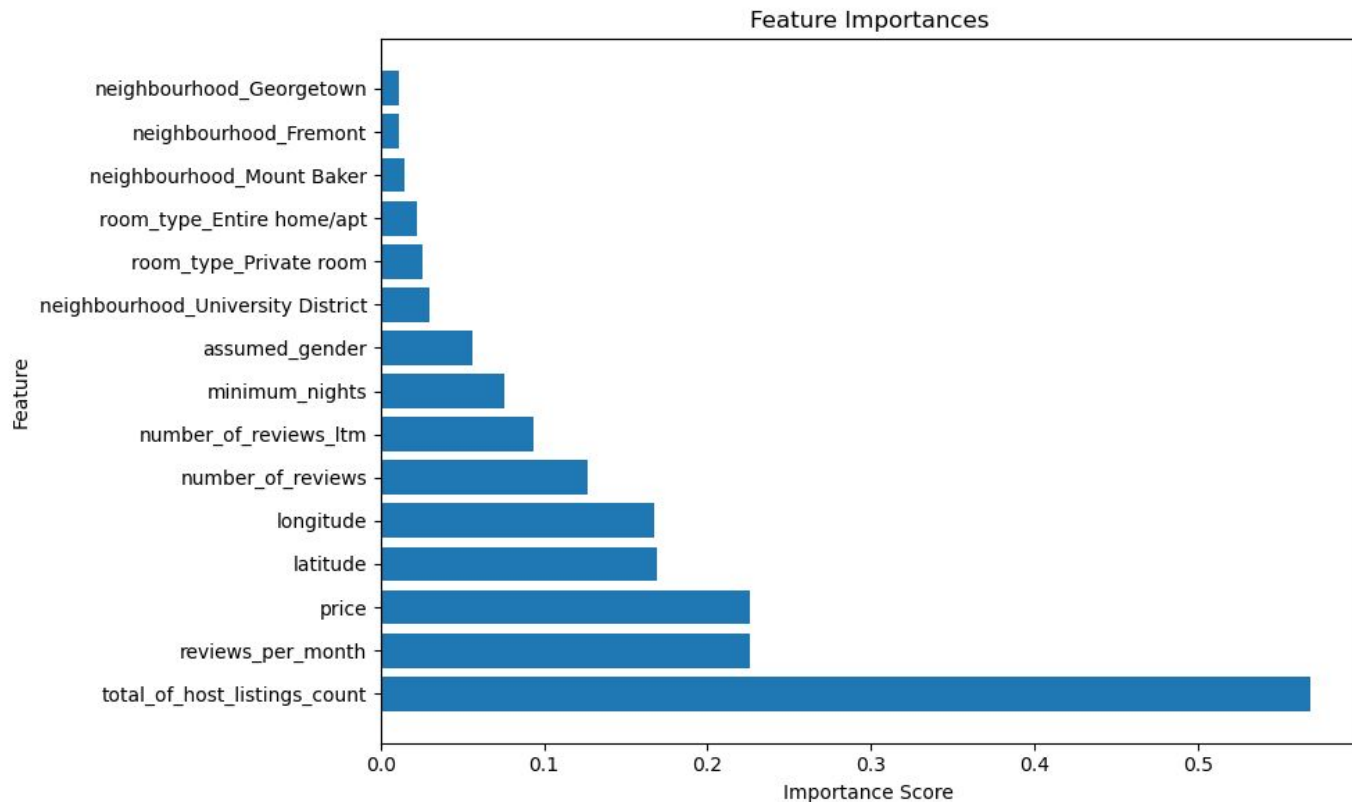
Transforming  
Columns

Scaling

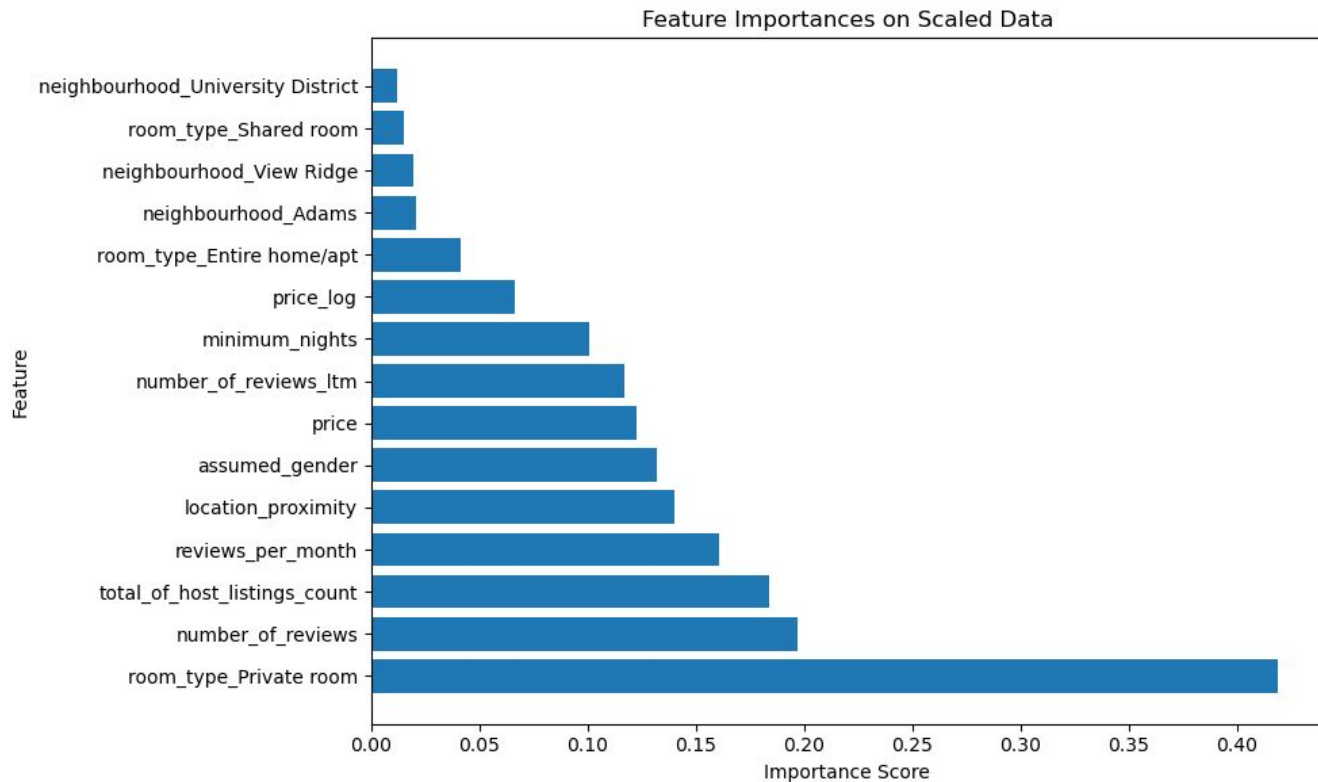
Changing  
Categories

Choosing  
Features

# Feature Importance not Scaled Data

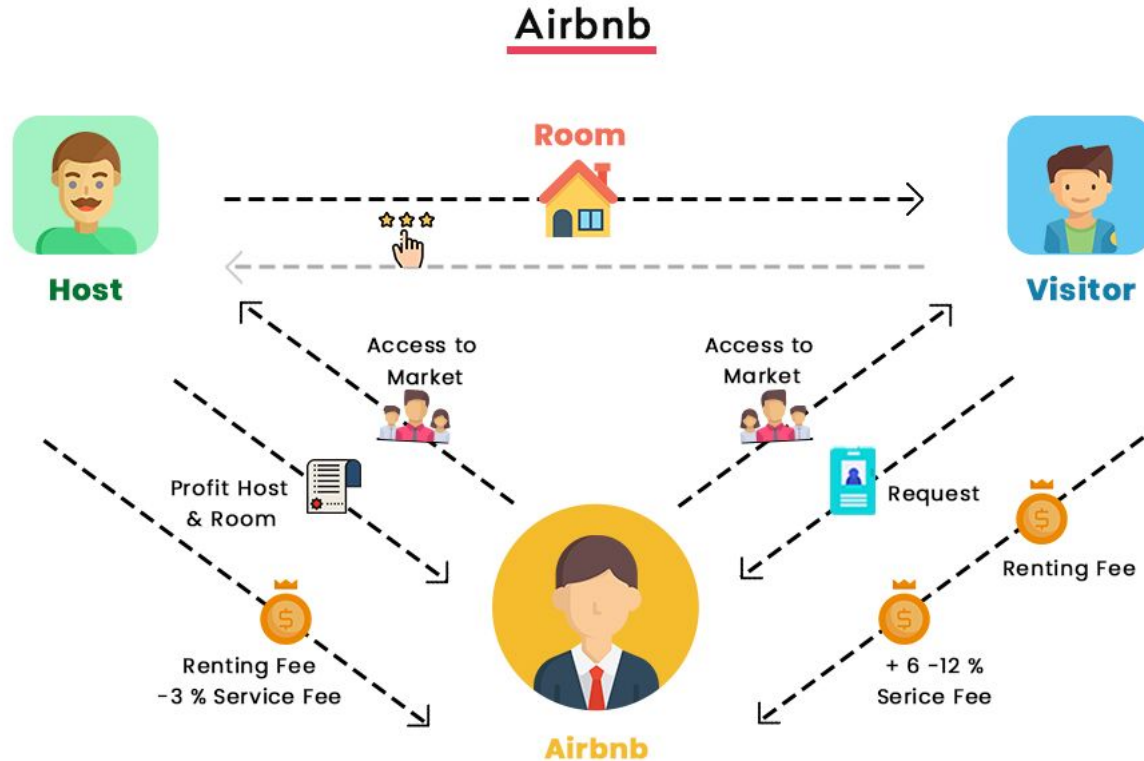


# Feature Importance Scaled Data





# Recommendation System



# Recommendation System Output

\*\*\*\*\*

\* Highland Park :

\*\*\*\*\*

neighbourhood	Highland Park
latitude	47.53183
longitude	-122.35404
room_type	Private room
price	45
minimum_nights	1 night
availability_365	47
host_id	13701143
income	14310

Name: 956, dtype: object

---

neighbourhood	Highland Park
latitude	47.52253
longitude	-122.35399
room_type	Private room
price	55
minimum_nights	2 nights
availability_365	333
host_id	3264544
income	1760

Name: 4155, dtype: object

---

neighbourhood	Highland Park
latitude	47.52309
longitude	-122.35201
room_type	Private room
price	55
minimum_nights	2 nights
availability_365	333
host_id	3264544
income	1760

Name: 4156, dtype: object

\*\*\*\*\*

\* Olympic Hills :

\*\*\*\*\*

neighbourhood	Olympic Hills
latitude	47.72422
longitude	-122.30482
room_type	Private room
price	57
minimum_nights	2 nights
availability_365	65
host_id	86433687
income	17100

Name: 4486, dtype: object

---

neighbourhood	Olympic Hills
latitude	47.72575
longitude	-122.30467
room_type	Private room
price	42
minimum_nights	1 night
availability_365	78
host_id	86433687
income	12054

Name: 4583, dtype: object

---

neighbourhood	Olympic Hills
latitude	47.72865
longitude	-122.29914
room_type	Private room
price	70
minimum_nights	2 nights
availability_365	295
host_id	53169216
income	4900

Name: 458, dtype: object

\*\*\*\*\*

\* Crown Hill :

\*\*\*\*\*

neighbourhood	Crown Hill
latitude	47.7019
longitude	-122.36781
room_type	Private room
price	50
minimum_nights	30 nights
availability_365	365
host_id	2438665
income	0

Name: 51, dtype: object

---

neighbourhood	Crown Hill
latitude	47.70253
longitude	-122.37196
room_type	Private room
price	52
minimum_nights	11 to 29 nights
availability_365	328
host_id	20765151
income	1924

Name: 316, dtype: object

---

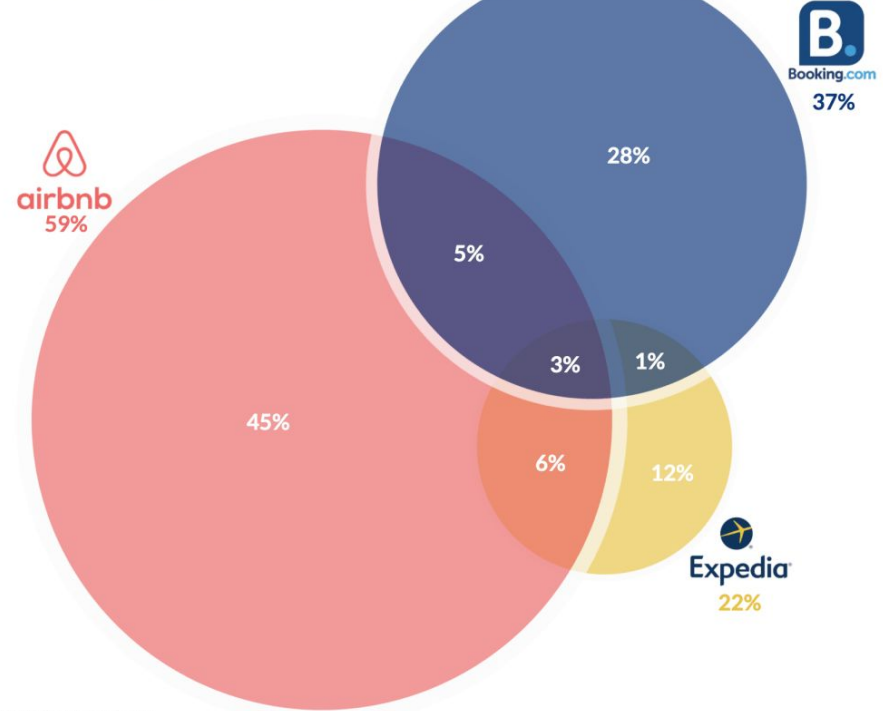
neighbourhood	Crown Hill
latitude	47.69608
longitude	-122.3685
room_type	Private room
price	46
minimum_nights	10 or less nights
availability_365	68
host_id	172036746
income	13662

Name: 3629, dtype: object

# Future Analysis

- Incorporate Text Data:
  - Reviews
  - Listings Names
- Incorporate Datetime Data:
  - Seasonality:
    - Availability
    - Price
    - Income
- Further Feature Engineering
- Beyond Seattle, WA
- Beyond Airbnb
- Improve Recommendation System

OTA listing share & overlap



Source: seetransparent.com  
Chart includes entire home listings only  
Expedia data includes VRBO brand.