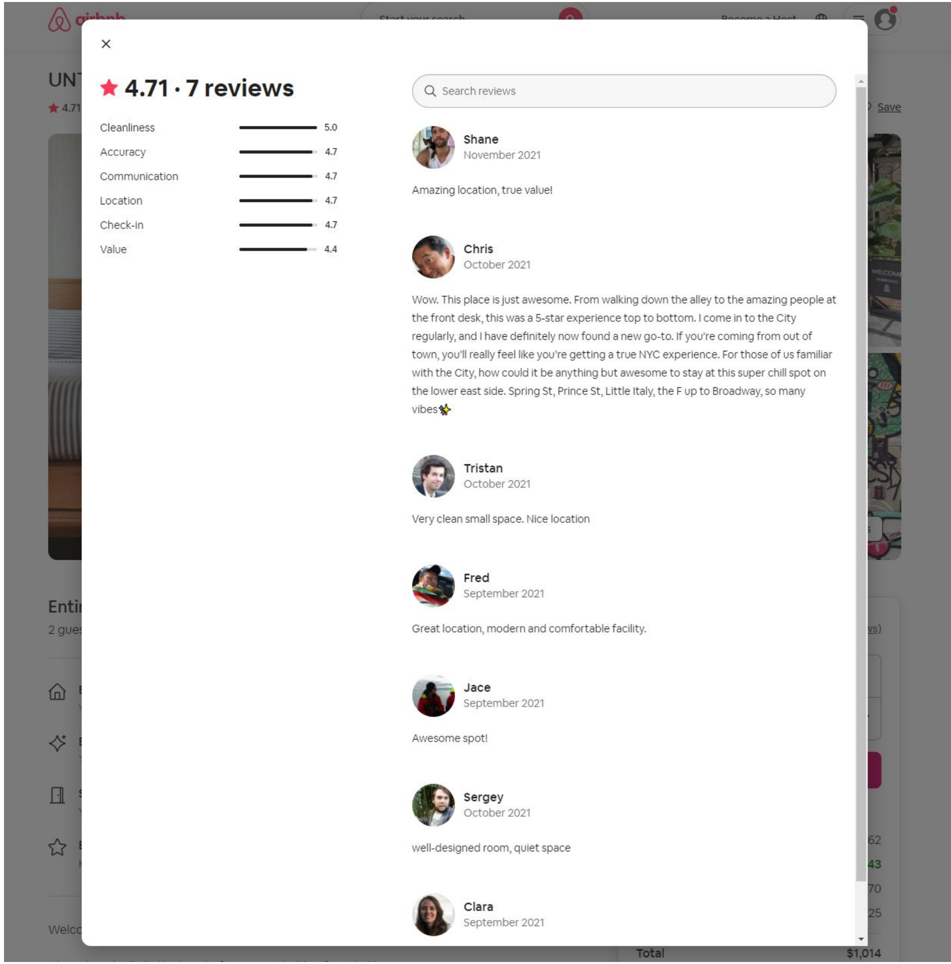
Topic Modelling for Reviews, Sources Affecting Guests’ Positive and Negative Experiences

Gulay Samatli Pac

Why reviews matter?

- Reviews **helps guests** choose their travel plans wisely and enables hosts to open their homes with confidence and **attract guests**.



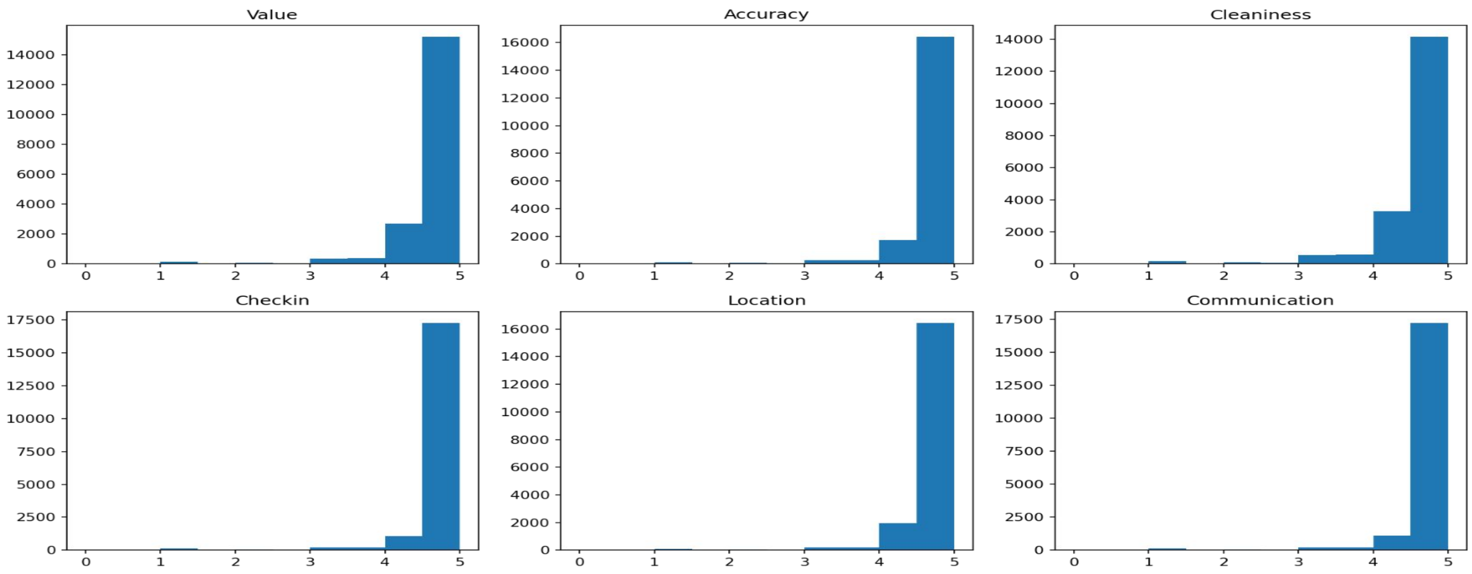
Why reviews matter?

- The first thing to avoid when booking an Airbnb:

  Little or No Reviews

What about ratings?

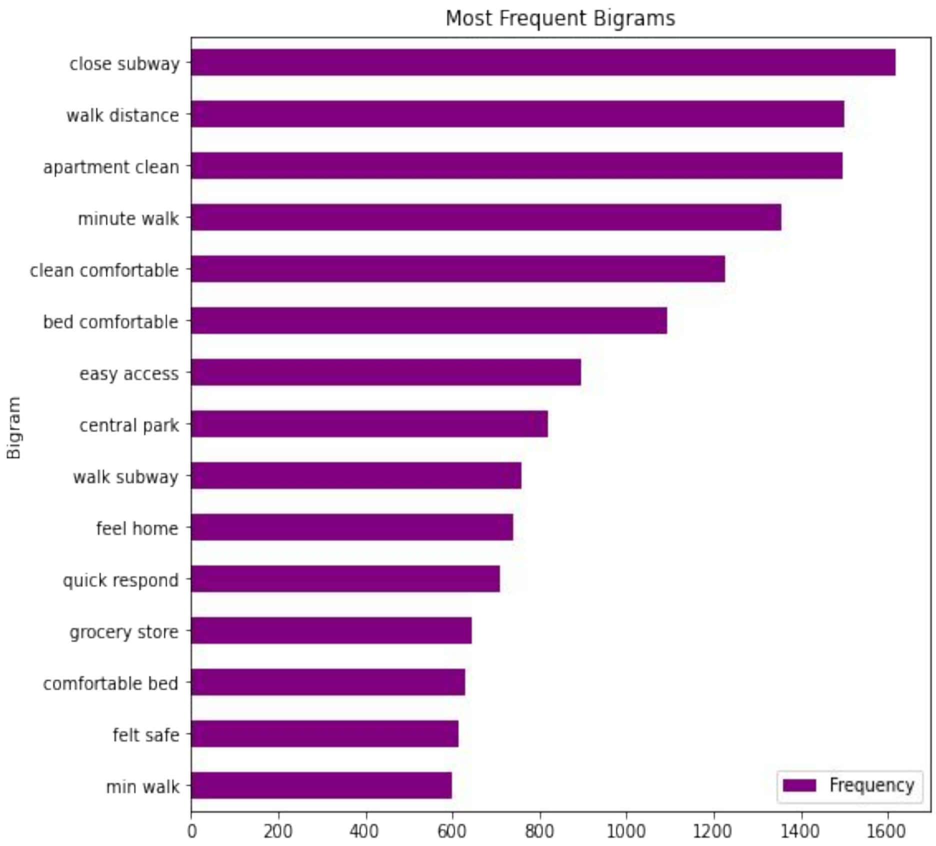
- Guests can provide star ratings with 1–5 stars on specific aspects of their experience.
- The problem:
  - All review scores are highly positive scores.
  - There is no scores less than 4.5 out of 5(see below graphs).

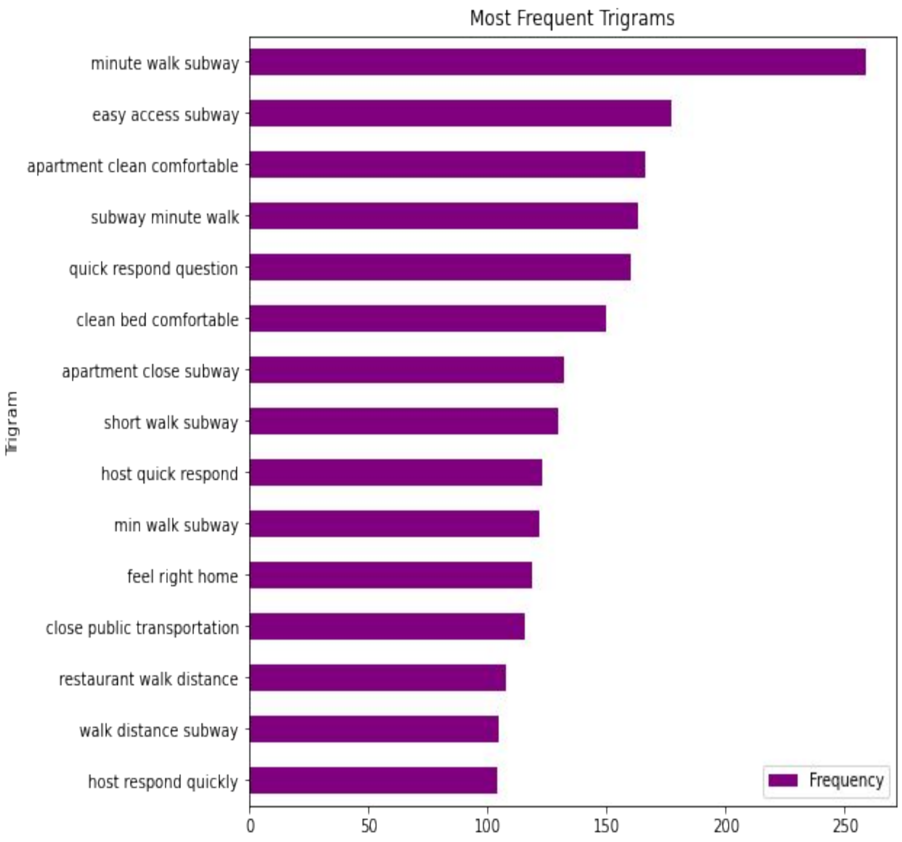


Methodology

Data	Data Cleaning & EDA	Topic Modeling	Final Model
- NYC Listings Dataset, InsideAirbnb - More than 80K Reviews	- Remove numbers, capital letters and punctuations - Eliminate non-English reviews - Lemmatize - Tools: Pandas, Numpy, langdetect, NLTK, Matplotlib, Seaborn	- TF-IDF Vectorizer - CountVectorizer - Models : - NMF - SVD - LDA - CorEx - Tools : sklearn, pyLDAvis	- Count Vectorizer (ngram_range=(1,2), max_df = 0.7, min_df = 10) - LDA - VaderSentimentAnalysis - Tools : sklearn, pyLDAvis

Reviews: Most common words in reviews





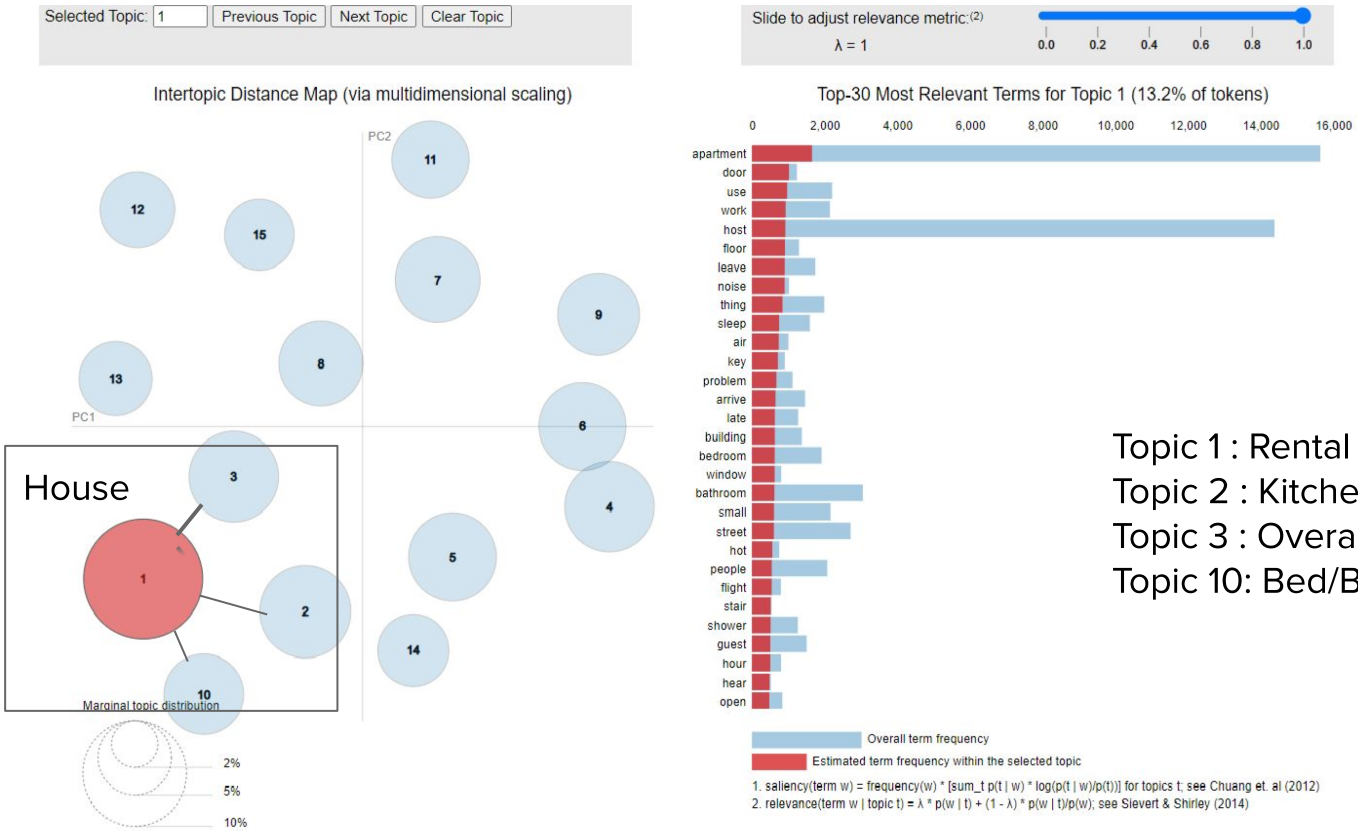
Topic Modelling

- Vectorizer :
    - CountVectorizer
- Topic Modeler :
    - LDA
- Number of Topic 15

Topics

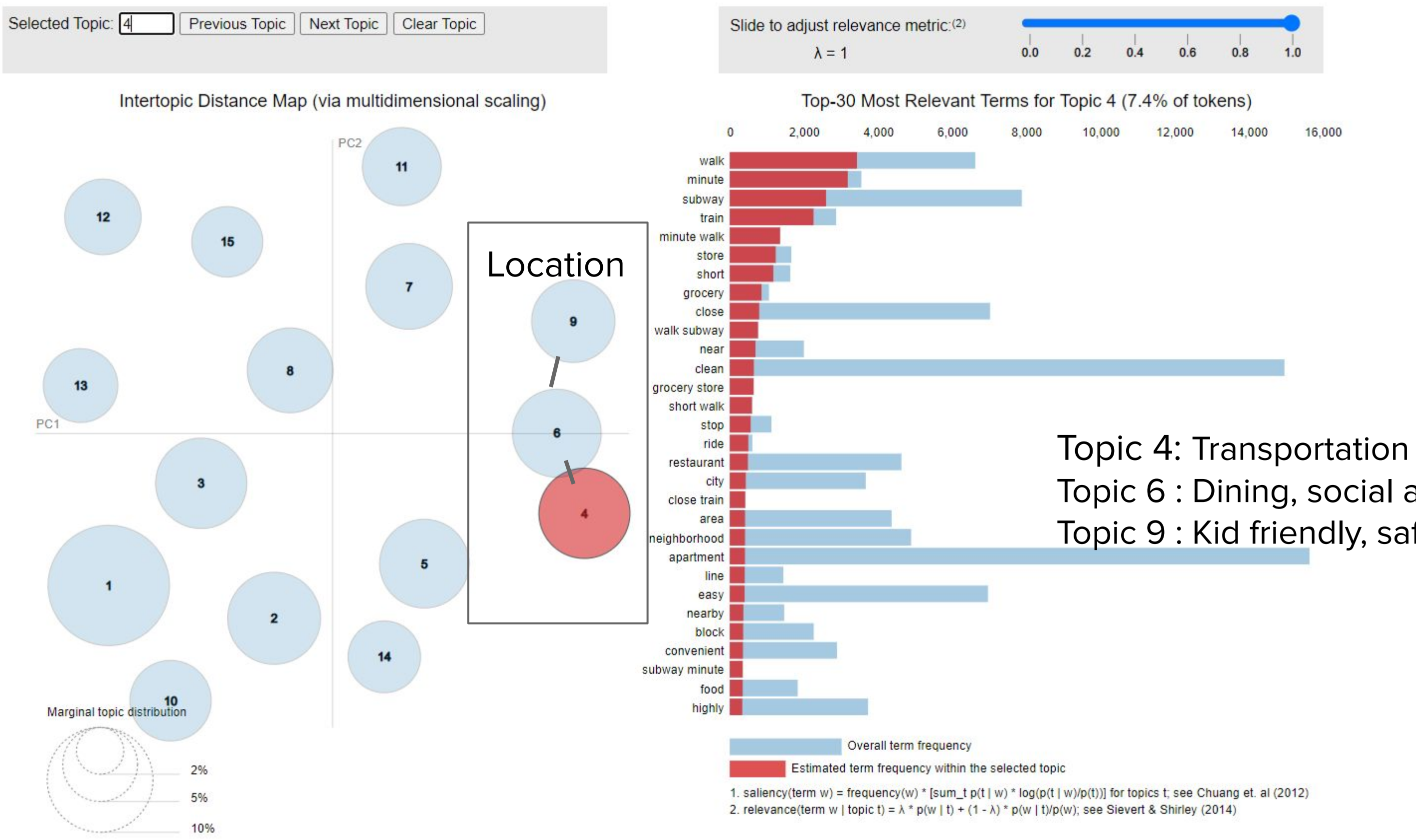
🏠 Rental interior issues
🍳 Kitchen Experience
Overall Airbnb Experience
🚇 Neighborhood- accessibility to transportation
👫 Home-like comfort/experience
🍷 Neighborhood/ accessibility to dining, social attractions
🧼 Cleanliness
👩‍💼 Host-hospitality
🚸 Location- safety/family friendly
🛏️🛀Bed/Bathroom
😄 Overall trip experience
📲 Host-responsiveness
🔑 Convenience (check in/out, comfort, hotel like)
💰 Comfort/Value
🏡 Listing Accuracy

Topic Visualization



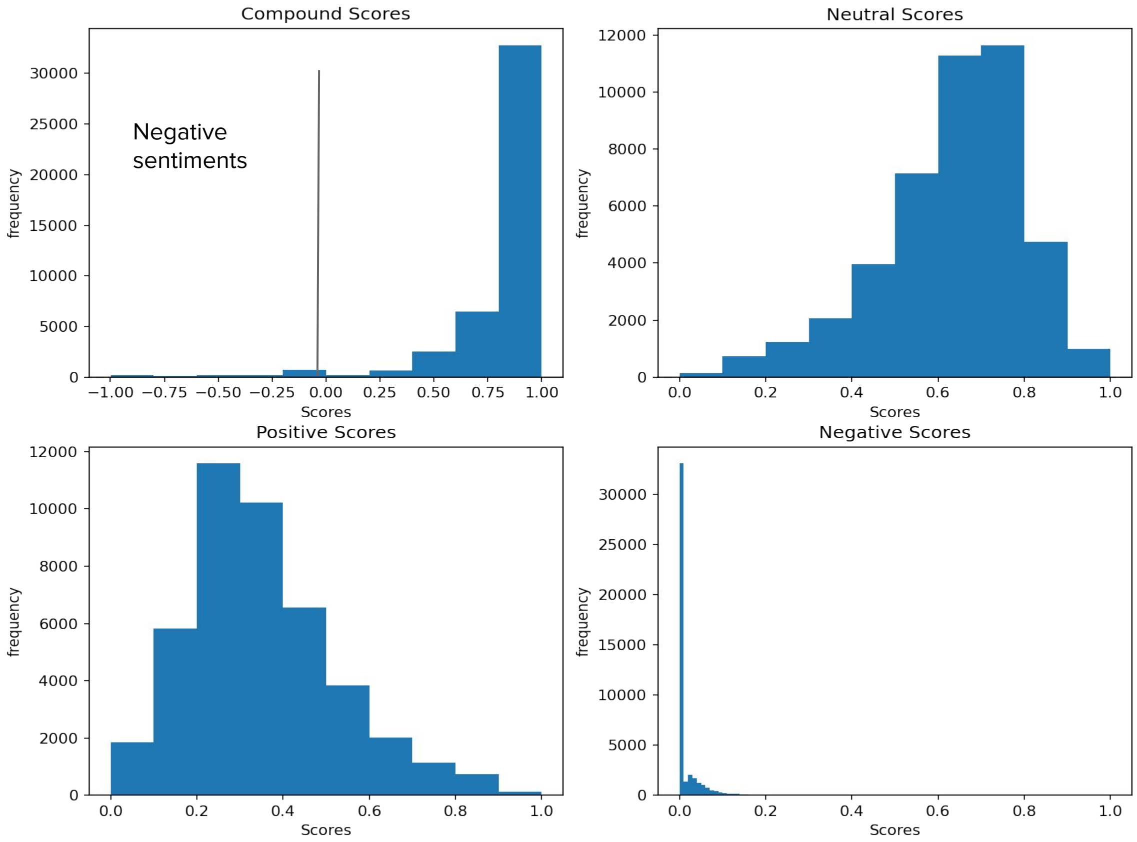
Topic 1 : Rental interior
Topic 2 : Kitchen
Topic 3 : Overall experience
Topic 10: Bed/Bathroom

Topic Visualization



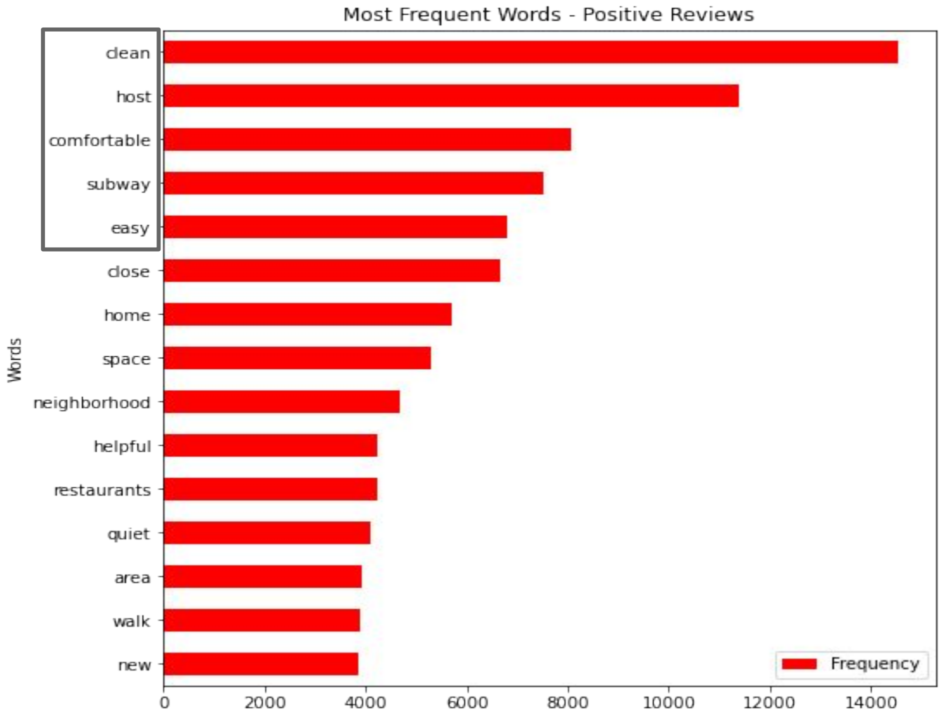
Topic 4: Transportation
Topic 6 : Dining, social attractions
Topic 9 : Kid friendly, safety

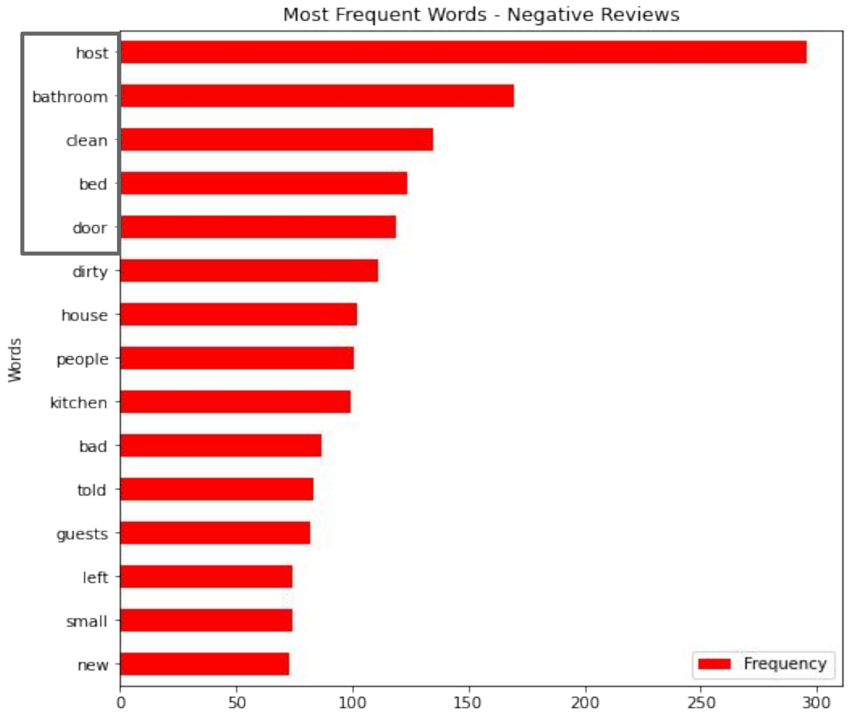
Reviews: Sentiment Analysis/ VaderSentiment



- Most guest had pleasant experiences during their Airbnb stays.

Positive vs Negative Reviews





Host interaction, Cleanliness ➔ Overall experience

Insights for Hosts

Common topics in reviews

- The condition of the house
  - Kitchen, bedrooms, and bathrooms
- Cleanliness
- Location
- Host responsiveness and hospitality

Next Steps

- Cluster analysis to group topics
- Use keywords as a filtering options for the search tool
- Incorporate reviews as a feature to predict occupancy rates

Questions

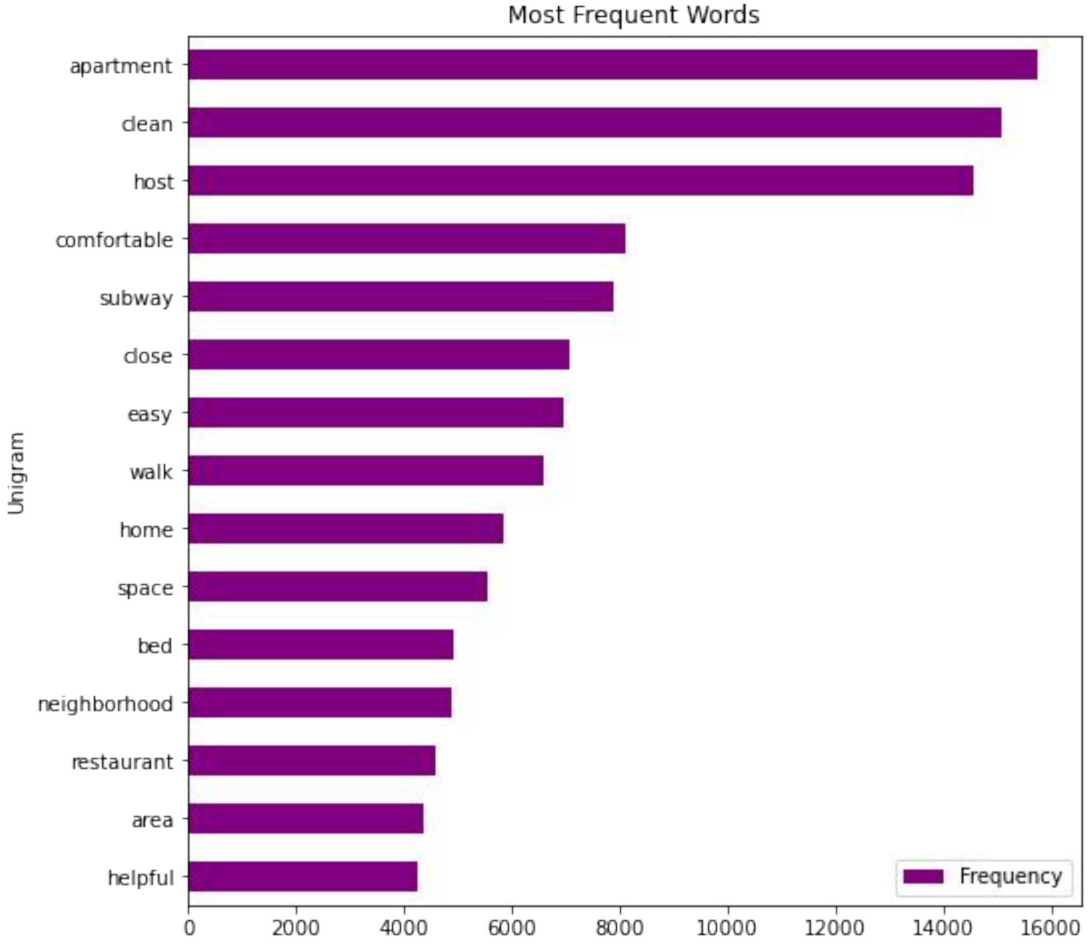
Thank you!

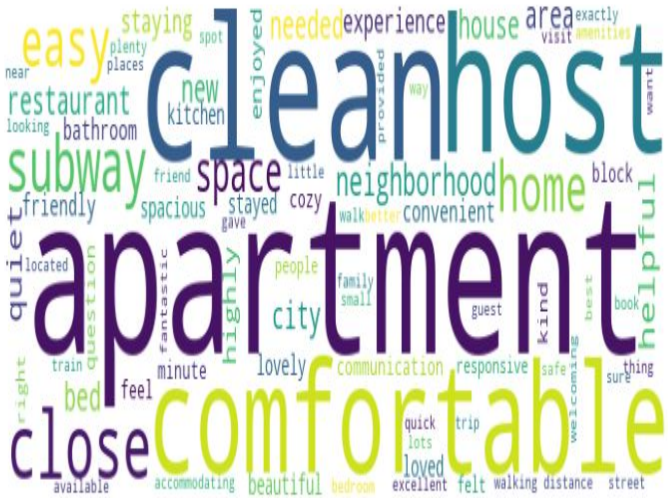
Appendix

Review Topics

1 Apartment interior issues
2 Kitchen Experience
3 Airbnb Experience
4 Neighborhood- accessibility to transportation
5 Home-like comfort/experience
6 Neighborhood/ accessibility to social attractions
7 Cleanliness
8 Host-hospitality
9 Location- safety/family friendly
10 Bed/Bathroom
11 Overall trip experience
12 Host-responsiveness
13 Convenience (check in/out, comfort, hotel like)
14 Comfort/Value
15 Listing Accuracy

Reviews: Most common words in reviews





Negative Topics:

Topic 1 - Inaccurate listings
Topic 2 - Check-in/out
Topic 3 - Bed/Bathroom
Topic 4 - Dirtiness and smell
Topic 5 - Uncomfortable sleep conditions
Topic 6 - Location
Topic 7 - Poor house maintenance
Topic 8 - Noise
Topic 9 - Hot Water/Heater
Topic 10 - Dirtiness
Topic 11 - Location
Topic 12 - Location general
Topic 13 - Value

Positive Topics Topics: