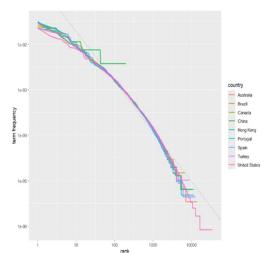
#### **Executive Summary**

Airbnb descriptions in this analysis are related to the amenities and the city places around its location. Across the nine countries present, essential words are consistently related to the cities and tourist attractions in them. Reviews showed positive outcomes are related to cleanliness and good guest experiences as driving causes, while noise is related to negative reviews. Feelings of trust, anticipation, and joy indicate overall satisfaction with the accommodation on the platform.

Data Overview: The data overview charts are in the appendix a) to c). The leading country in frequency is the USA, with 21%; other countries share a similar amount of around 10%. Most of these countries only have one city associated, but the distribution of the cities is more equitable, with most of them having around 9 and 11%. Hong Kong has the highest prices on average at

over 800\$, while Portugal is the most affordable at less than 100\$, as there are various countries with other languages, with the top 4 most predominant languages. Both texts in descriptions and reviews follow Zipf's law, showing that they are written in natural language except China because of the vocabulary they hold.



#### Bigrams and Trigrams of descriptions: Word

associations in the bigram map talk about places like swimming pools, Central Park, Sagrada Familia, or the city center. However, words like number 1 or 2 are also associated with spaces inside the accommodation, and other numbers of multiples of 5, like 15 or 10, are associated with minutes or walks. In the trigram, the associations are similar but with an interesting one regarding coffee, restaurants, and art. This shows that descriptions go around the nearby accommodation, activities, and places.

**Correlations of terms**: when comparing by country like in the appendix i) USA has more uniqueness towards shore or beach related words. In contrast, other countries have more public

transport words related.

Looking at the correlations

by property type:

countryside retreats and lofts

are closely associated with

terms like "acre" and

"acres," hinting at expansive

oons

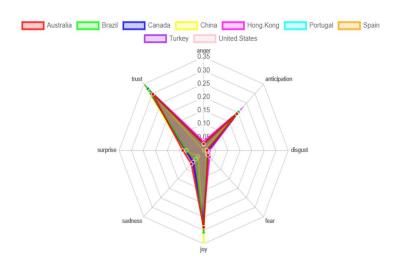
oon

Condominiums stand out with words like "complex" and "condo," emphasizing community features, while apartments are linked to urban descriptors like "city" and "barrio." Houses highlight outdoor amenities such as "bbq" and "deck," contrasting with apartments' urban references like "Hong" and "Catalunya." Serviced apartments blend hotel-like amenities with apartment setups, sharing attributes like "pool" and "hotel" with apartments serviced apartment type is focused on the middle. Still, with words like pool or hotel or ama more linked to it than apartment, which does not seem to have a higher correlation in the shared words. both share a powerful word like beautiful.

#### **Sentiment Analysis of Reviews:**

To get a big picture of the sentiments in the reviews, it is better to compare the three more used lexicons, NRC, Bing and AFINN, as in Appendix j).

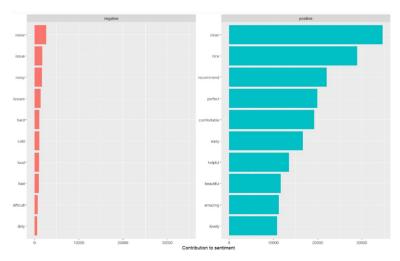
Overall, the reviews in this



analysis are positive across all of the lexicons. However, with the NRC lexicon, there are three

dominant feelings for all countries: trust, Anticipation, and Joy, which shows that users are satisfied with the accommodations in the data overall.

On the other hand, with the Bing lexicon, reviews of negative and positive words were plotted.



This plot hints that noise-related words negatively affect the reviews, while cleanliness is essential for the accommodation to have a positive review.

LDA of reviews by country: Most of the reviews in the top 5 countries are related to the positive description of the accommodation, while the reviews of Turkey are more related to places. In the cases of Australia, the USA, Spain or Canada, the accommodation is where the user spends more time, and it even could be the case that their experience in the accommodation is more impactful while in Turkey or Brazil, the places around are more impactful in their experience places.

Conclusions: Airbnb descriptions are keen on describing the accommodation details while mentioning activities to do and places to visit nearby. The company shown check on the Cleanliness of the accommodation as well as the noise around it on their platform as this are leading drivers for positive and negative reviews.

#### Tableau Dashboard

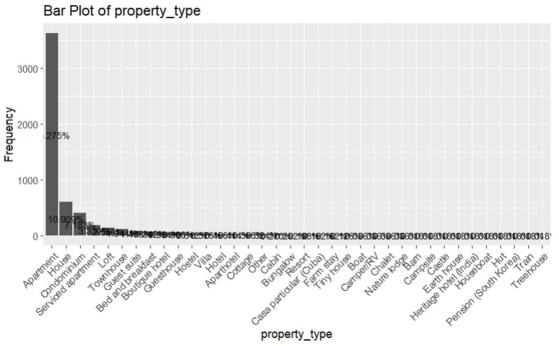
The dashboard offers an overall view of the data regarding property type and prices. It shows the distribution by location in the first view of the storyline, the charts showing the top 10 positive and negative words and finally, a look at property types by city Preview in Appendix k). Link to the Dashboard in Tableau Public:

https://public.tableau.com/app/profile/jose.romainville.solis/viz/dashboard\_17106317692710/Air

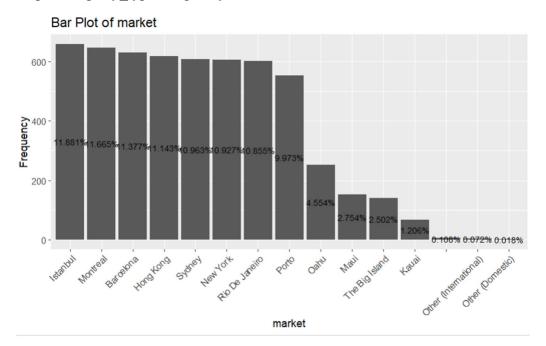
BnbCommentsDescriptions?publish=yes

## Appendix

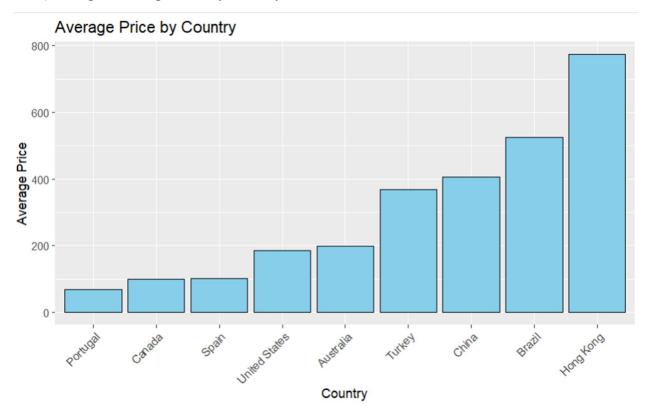
## a) Bar plot Country Frequency



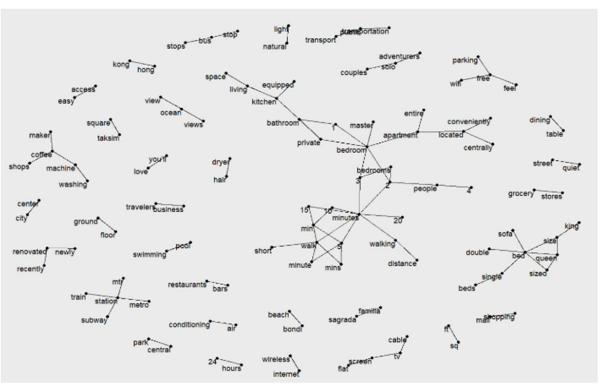
### b) Bar plot Property\_type Frequency



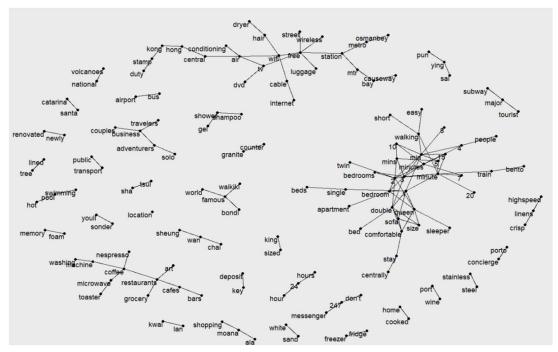
### c) Bar plot Average Price by Country



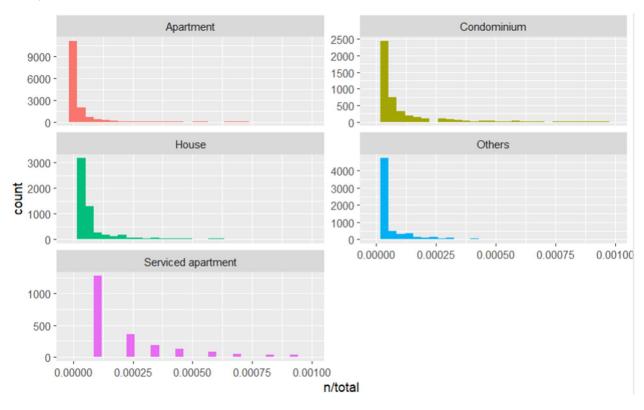
### d) Bigrams map



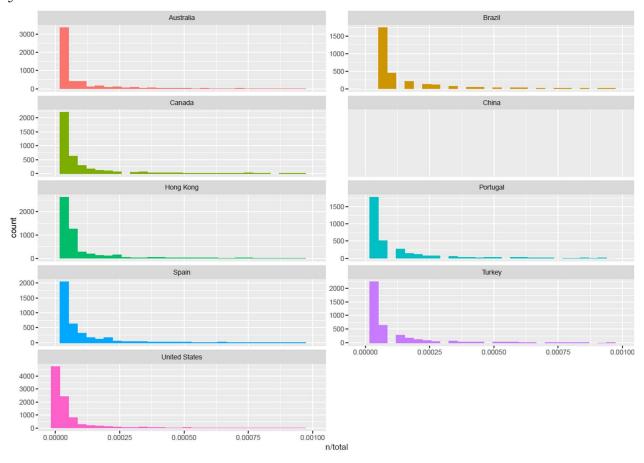
e) Trigram Map



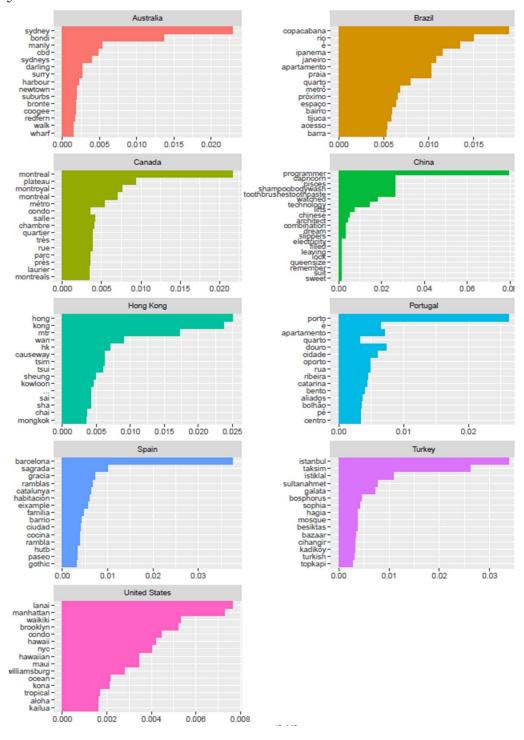
### f) Term frequency by Property Type



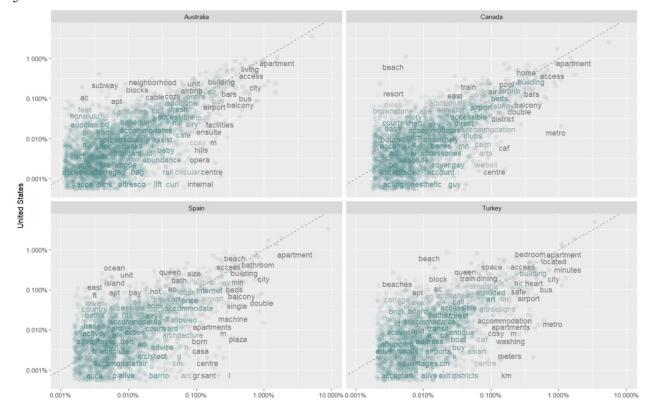
### g) Term frequency by country



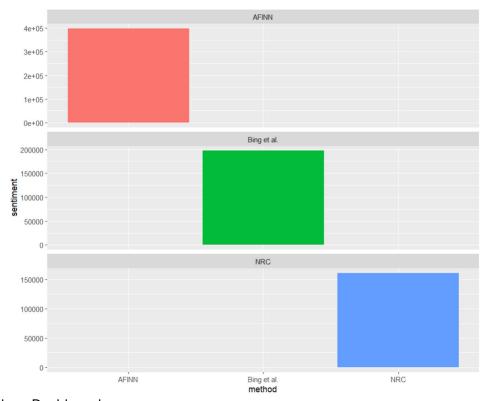
h) TDF-IDF



i) Correlations of Descriptions by Country



#### j) Sentiment analysis with 3 lexicons



#### k) Tableau Dashboard

