# A look into Covid-19 spreading through the lodging sector

*Estefania Ruiz Martinez • 30.06.2021*

# TOC

- Better customer experience $\longrightarrow$ better customer satisfaction.
- Online reviews are a type of UGC (User Generated Content) considered as the voice of the customer.
- In the light of the health crisis of Covid-19, has the customer experience been affected?

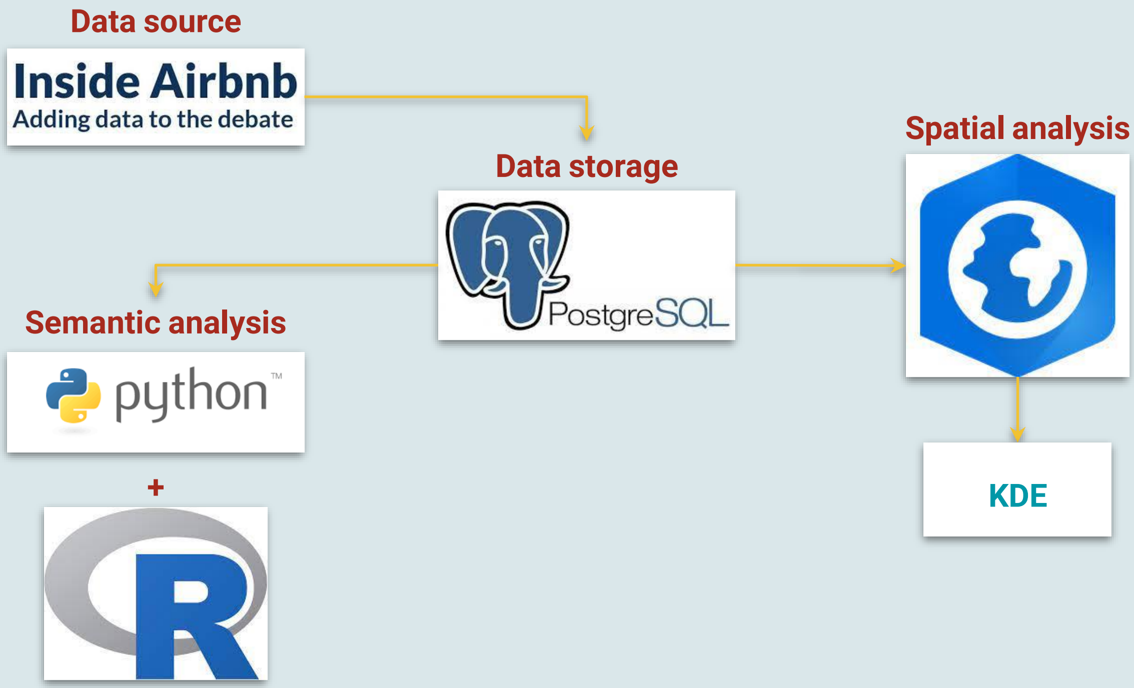
- While academics have already started to explore the behaviour of customers during  the health crisis of Covid-19, none of the studies found have done it with customers of P2P accommodations.
- Up to now, no research found integrates spatial analysis with text mining techniques to reveal new aspects of the customer experience.
- This thesis aims to do so by exploring the behaviour of Airbnb customers in two geographically contrasting cities but as well located in countries highly affected by the health crisis of Covid-19

**How is the experience of Airbnb users after the outbreak of COVID-19?**

- How was the experience of Airbnb users according to the sentiment polarity?
- How was the experience of Airbnb users according to the presence/absence of covid-terms?
- Where did Airbnb users experience positive, neutral, and negative sentiments?
- Where did Airbnb users mention covid-terms?

**To explore the experience of Airbnb users after the outbreak of COVID-19.**

- To classify online reviews according to the sentiment polarity.
- To classify online reviews according to the presence/absence of covid-terms.
- To analyse keywords and their relationship.
- To analyse the spatial distribution of property listings according to sentiment polarity and the presence of covid-terms in property' reviews.


Data source
Inside Airbnb
Adding data to the debate
Data storage
PostgreSQL
Spatial analysis
Semantic analysis
python™
+
R
KDE

1

**Data collection and storage**

## Data selection

- High corona cases
- Touristic cities

**Rio de Janeiro**

**New York**

## Data content

- Property listings, including the geographic coordinates with an error between 0 and 150 m
- Reviews
- Calendar availability

2

**Text preparation**

**First part**

Removal of:

- Automated postings (e.g., “This is an automated posting”)
- Non-English reviews → **Fasttext**
- Duplicated and empty reviews
- Reviews consisting in only two characters, numbers or NaN

**Second part**

- ➢ Tokenization and lowercasing
- ➢ Spelling correction, lemmatization and expansion of contractions
- ➢ Extraction of nouns, adjectives and verbs

Removal of:

- Stopwords, emojis, proper nouns, special characters, numbers, punctuations, extra whitespaces, tabs, and newlines.

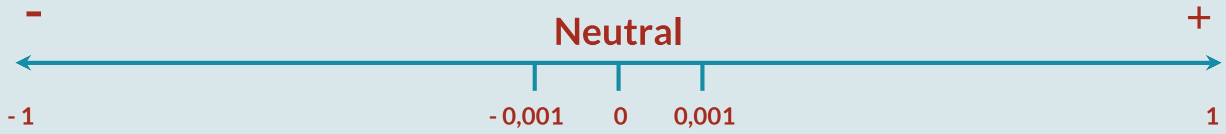
# VADER

## 3

**Text categorization**

*Sentiment analysis*

| | label | review | scores | compound |
|---|---|---|---|---|
| **0** | pos | Stuning even for the non-gamer: This sound tra... | {'neg': 0.088, 'neu': 0.669, 'pos': 0.243, 'co... | 0.9454 |
| **1** | pos | The best soundtrack ever to anything.: I'm rea... | {'neg': 0.018, 'neu': 0.837, 'pos': 0.145, 'co... | 0.8957 |
| **2** | pos | Amazing!: This soundtrack is my favorite music... | {'neg': 0.04, 'neu': 0.692, 'pos': 0.268, 'com... | 0.9858 |
| **3** | pos | Excellent Soundtrack: I truly like this soundt... | {'neg': 0.09, 'neu': 0.615, 'pos': 0.295, 'com... | 0.9814 |
| **4** | pos | Remember, Pull Your Jaw Off The Floor After He... | {'neg': 0.0, 'neu': 0.746, 'pos': 0.254, 'comp... | 0.9781 |



3

**Text categorization**

*Covid terms*

**Includes:**

- Terms that appear very often in texts related the Covid situation (e.g., *mask*, *pandemic*, *lockdown*).
- Terms people use to refer to the name of the virus or the disease and their abbreviations.

**Excludes:**

- Open compound words
- Ambiguous words

**Total: 88 terms**

**Sources:**

- Scientific articles.
- Study by Lillo (2020) "*COVID-19, the beer flu; Or, the disease of many names*".

## *Keywords and their relationship*

4

**Topic modelling**

Frequent terms
*RTF*

Rare terms
*TF-IDF*

Network diagrams

Edges
Strength of correlation

Nodes
Keywords

Colors
Communities, given by *cluster Louvain algorithm*

5

## Spatial analysis

**Density surface**
*Visualize the spatial distribution of properties*

According to

- **Overall polarity:** *Positive, neutral and negative*
- **Time period:** *Before and after the outbreak of covid*

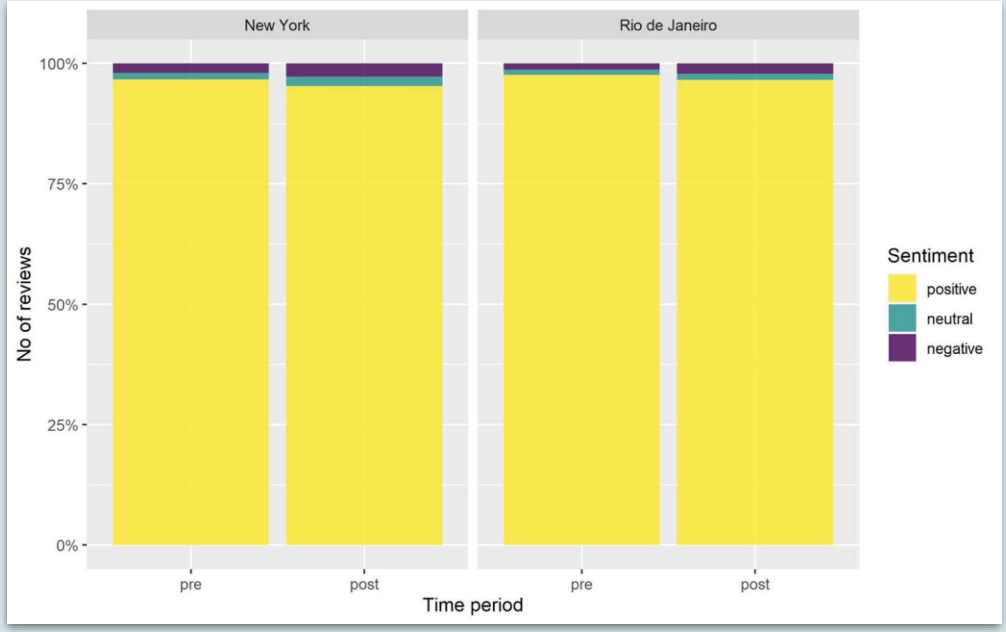
To

- *Identify areas with high intensity positive, neutral and negative sentiments*
- *Identify changes after the outbreak of covid by using* ***Map algebra***

1

**Sentiment analysis**

*Reviews*




Total reviews
NY: 486,438
RJ: 26,262

Positive reviews
NY: 2%
RJ: 1%

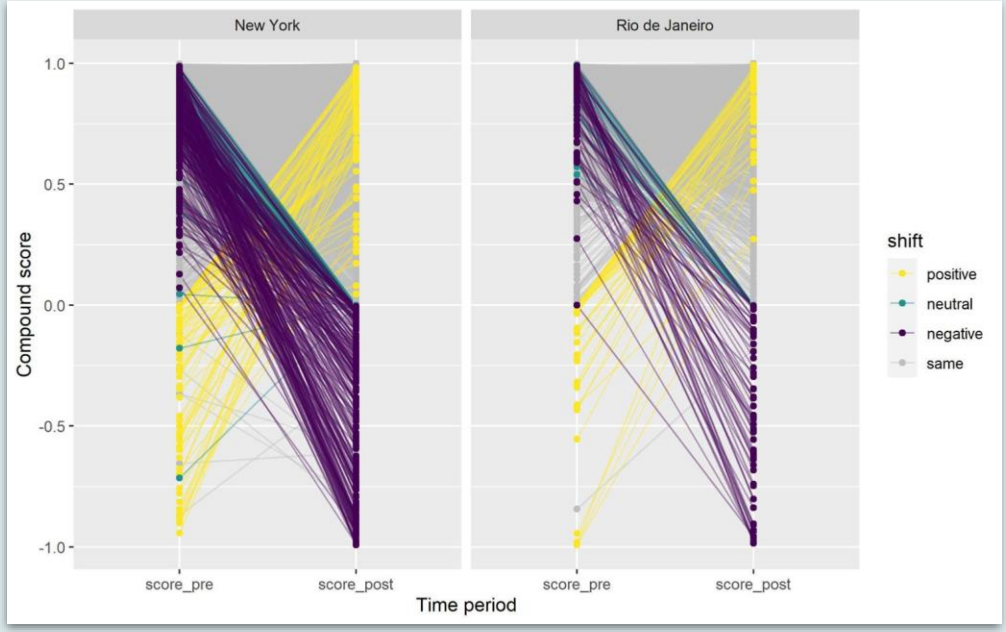
Neutral reviews
NY: 1%
RJ: same amount

Negative reviews
NY: 1%
RJ: 1%

1.1

**Sentiment analysis**

*Property listings*




Total properties
NY: 18,751
RJ: 3,522

Polarity shift
NY: 2,1%
RJ: 3.6%

Positive to neutral
NY: 0,44%
RJ: 0,65%

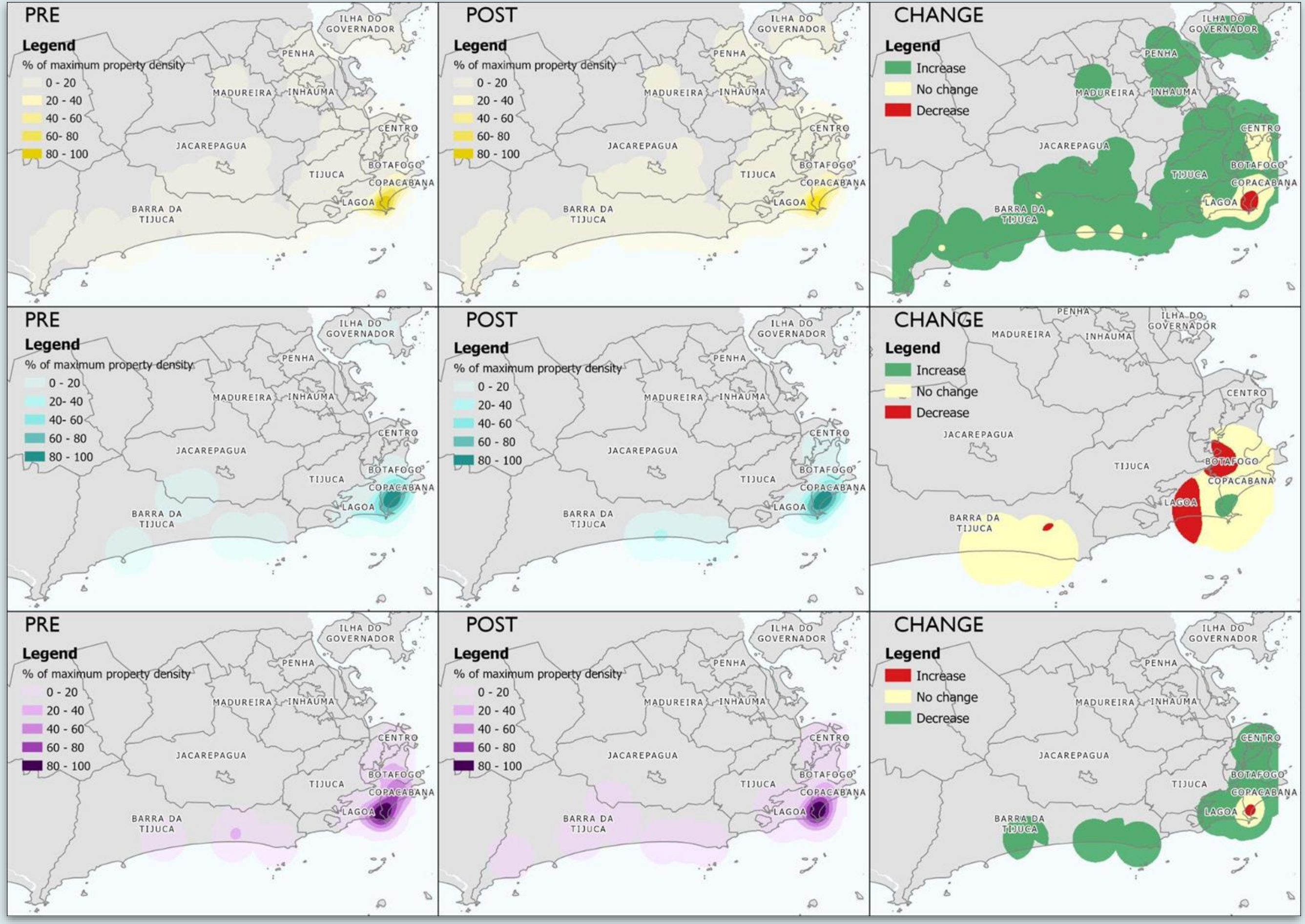
Positive to negative
NY: 1,2%
RJ: 1,7%

1.2

## Sentiment analysis

*Spatial visualization*

**Rio de Janeiro**

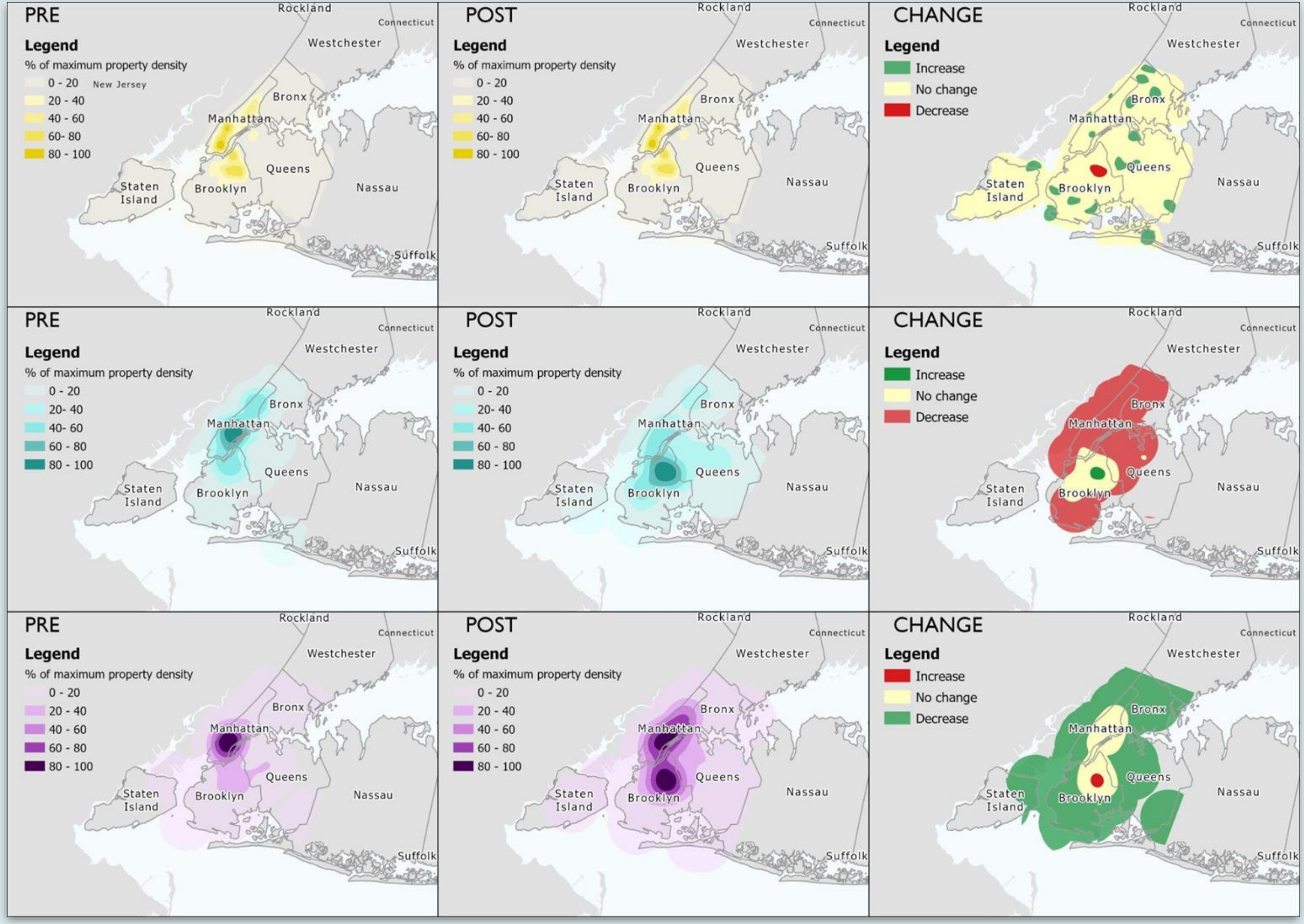


1.2

**Sentiment analysis**

*Spatial visualization*

**New York**



## Pre-covid

**Table 4.** List with the 10 most frequent terms from reviews of properties in RJ before the outbreak of covid.

| | Positive | | | Neutral | | | Negative | | |
|---|---|---|---|---|---|---|---|---|---|
| No. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. |
| 1 | place | 9178 | 1 | place | 24 | 1 | apartment | 160 | 1 |
| 2 | apartment | 8113 | 0,88 | apartment | 24 | 1 | place | 120 | 0,75 |
| 3 | location | 8065 | 0,88 | beach | 24 | 1 | host | 115 | 0,72 |
| 4 | host | 5222 | 0,57 | location | 15 | 0,63 | location | 79 | 0,49 |
| 5 | stay | 4562 | 0,50 | stay | 11 | 0,46 | day | 68 | 0,43 |
| 6 | beach | 4348 | 0,47 | host | 9 | 0,38 | stay | 62 | 0,39 |
| 7 | time | 2587 | 0,28 | room | 8 | 0,33 | room | 62 | 0,39 |
| 8 | view | 2524 | 0,28 | time | 6 | 0,25 | night | 54 | 0,34 |
| 9 | restaurant | 2480 | 0,27 | restaurant | 6 | 0,25 | issue | 37 | 0,23 |
| 10 | room | 1701 | 0,19 | price | 4 | 0,17 | work | 35 | 0,22 |

## Post-covid

**Table 5.** List with the 10 most frequent terms from reviews of properties in RJ after the outbreak of covid.

| No. | Positive | | | Neutral | | | Negative | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. |
| 2 | place | 4525 | 1 | place | 17 | 1 | apartment | 158 | 1 |
| 3 | location | 3874 | 0,86 | beach | 16 | 0,94 | place | 114 | 0,72 |
| 4 | apartment | 3839 | 0,85 | location | 12 | 0,71 | host | 76 | 0,48 |
| 5 | host | 2505 | 0,55 | host | 10 | 0,59 | location | 67 | 0,42 |
| 6 | stay | 2312 | 0,51 | stay | 9 | 0,53 | day | 61 | 0,39 |
| 7 | beach | 1967 | 0,43 | apartment | 9 | 0,53 | time | 60 | 0,38 |
| 8 | view | 1299 | 0,29 | need | 4 | 0,24 | room | 53 | 0,34 |
| 9 | time | 1263 | 0,28 | view | 3 | 0,18 | night | 49 | 0,31 |
| 10 | restaurant | 1127 | 0,25 | street | 3 | 0,18 | water | 38 | 0,24 |
| No. | room | 829 | 0,18 | restaurant | 3 | 0,18 | people | 36 | 0,23 |

1.3

**Sentiment analysis**

*Frequent terms*

**Rio de Janeiro**

Shared by all:
Place, apartment, location and host.

Unique to neutral
Pre: **price**
Post: **need** and **street**

Unique to negative
Pre: **issue** and **work**
Post: **water** and **people**

1.3

## Sentiment analysis

*Frequent terms*

**New York**

## Pre-covid

**Table 6.** List with the 10 most frequent terms from reviews of properties in NY before the outbreak of covid.

| | Positive | | | Neutral | | | Negative | | |
|---|---|---|---|---|---|---|---|---|---|
| No. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. |
| 1 | **place** | 206130 | 1 | **place** | 1160 | 1 | **place** | 3828 | 1 |
| 2 | **location** | 115551 | 0,56 | **location** | 547 | 0,47 | **room** | 3304 | 0,86 |
| 3 | **stay** | 102795 | 0,50 | **stay** | 536 | 0,46 | **host** | 2339 | 0,61 |
| 4 | **host** | 93211 | 0,45 | **room** | 345 | 0,30 | **apartment** | 2316 | 0,61 |
| 5 | **apartment** | 89161 | 0,43 | **apartment** | 265 | 0,23 | night | 1962 | 0,51 |
| 6 | **room** | 62843 | 0,30 | subway | 240 | 0,21 | **time** | 1553 | 0,41 |
| 7 | **space** | 46827 | 0,23 | **host** | 228 | 0,20 | **stay** | 1523 | 0,40 |
| 8 | **time** | 44286 | 0,21 | station | 216 | 0,19 | bathroom | 1514 | 0,40 |
| 9 | home | 31288 | 0,15 | **time** | 185 | 0,16 | day | 1448 | 0,38 |
| 10 | area | 30974 | 0,15 | **space** | 184 | 0,16 | **location** | 1431 | 0,37 |

## Post-covid

**Table 7**. List with the 10 most frequent terms from reviews of properties in NY after the outbreak of covid.

| | Positive | | | Neutral | | | Negative | | |
|---|---|---|---|---|---|---|---|---|---|
| No. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. |
| 1 | **place** | 70122 | 1 | **place** | 486 | 1 | **place** | 1930 | 1 |
| 2 | **stay** | 37576 | 0,54 | **stay** | 302 | 0,62 | **room** | 1755 | 0,91 |
| 3 | **location** | 34874 | 0,50 | **location** | 220 | 0,45 | **host** | 1329 | 0,69 |
| 4 | **host** | 29807 | 0,43 | **room** | 175 | 0,36 | **apartment** | 1047 | 0,54 |
| 5 | **apartment** | 24587 | 0,35 | **space** | 112 | 0,23 | **night** | 911 | 0,47 |
| 6 | **room** | 17956 | 0,26 | **host** | 102 | 0,21 | **stay** | 870 | 0,45 |
| 7 | **space** | 17153 | 0,24 | **time** | 90 | 0,19 | **time** | 863 | 0,45 |
| 8 | **time** | 13749 | 0,20 | **apartment** | 86 | 0,18 | **day** | 861 | 0,45 |
| 9 | **home** | 11197 | 0,16 | book | 75 | 0,15 | **bathroom** | 632 | 0,33 |
| 10 | area | 8943 | 0,13 | **home** | 61 | 0,13 | **door** | 515 | 0,27 |

Shared by all:
Place, apartment, host, room, time and stay.

Unique to neutral
Pre: **subway** and **station**
Post: **book**

Unique to negative
Post: **door**

1.3

**Sentiment analysis**

*Frequent terms*

**Comparison**

## Keywords shared

Shared by all:
**Place, apartment, and host.**

Shared by all positive:
**Place, apartment, location, host, room, time and stay.**

Shared by all neutral:
**Place, apartment, location, host and stay.**

Shared by all negative:
**Place, apartment, room, host, *day* and *night*.**

1.4

# Sentiment analysis

*Rare terms*

## Pre-covid

| | Positive | | Neutral | | Negative | |
|---|---|---|---|---|---|---|
| No. | Term | Score | Term | Score | Term | Score |
| 1 | **year** | 1 | **year** | 1 | sheet | 1 |
| 2 | word | 1 | xbox | 1 | review | 1 |
| 3 | wish | 1 | **time** | 1 | renter | 1 |
| 4 | wifi | 1 | thumb | 1 | pay | 1 |
| 5 | wife | 1 | stay | 1 | block | 1 |
| 6 | welcome | 1 | star | 1 | shop | 0,93 |
| 7 | wait | 1 | saucer | 1 | spot | 0,92 |
| 8 | **view** | 1 | rockstar | 1 | **view** | 0,87 |
| 9 | value | 1 | right | 1 | experience | 0,84 |
| 10 | treat | 1 | reply | 1 | clean | 0,83 |

| | Positive | | Neutral | | Negative | |
|---|---|---|---|---|---|---|
| No. | Term | Score | Term | Score | Term | Score |
| 1 | zone | 1 | way | 1 | train | 1 |
| 2 | worth | 1 | **visit** | 1 | thanks | 1 |
| 3 | worry | 1 | think | 1 | television | 1 |
| 4 | woman | 1 | term | 1 | **stay** | 1 |
| 5 | welcome | 1 | subway | 1 | pay | 1 |
| 6 | want | 1 | **stay** | 1 | neighborhood | 0,93 |
| 7 | walk | 1 | star | 1 | need | 0,92 |
| 8 | **visit** | 1 | spot | 1 | host | 0,87 |
| 9 | **value** | 1 | room | 1 | complaint | 0,84 |
| 10 | **trip** | 1 | review | 1 | comment | 0,83 |

## Post-covid

| | Positive | | Neutral | | Negative | |
|---|---|---|---|---|---|---|
| No. | Term | Score | Term | Score | Term | Score |
| 1 | wow | 1 | **view** | 1 | reservation | 1 |
| 2 | worth | 1 | **time** | 1 | problem | 1 |
| 3 | work | 1 | supermarket | 1 | comment | 1 |
| 4 | wait | 1 | street | 1 | cold | 0,93 |
| 5 | visit | 1 | stay | 1 | doubt | 0,82 |
| 6 | **view** | 1 | star | 1 | mosquito | 0,81 |
| 7 | value | 1 | sim | 1 | guest | 0,80 |
| 8 | **time** | 1 | show | 1 | frill | 0,78 |
| 9 | think | 1 | renovation | 1 | change | 0,77 |
| 10 | thanks | 1 | place | 1 | truth | 0,77 |

| | Positive | | Neutral | | Negative | |
|---|---|---|---|---|---|---|
| No. | Term | Score | Term | Score | Term | Score |
| 1 | year | 1 | wifi | 1 | **worth** | 1 |
| 2 | **worth** | 1 | **visit** | 1 | thanks | 1 |
| 3 | window | 1 | time | 1 | **stay** | 1 |
| 4 | welcome | 1 | **stay** | 1 | **room** | 0,93 |
| 5 | **visit** | 1 | star | 1 | refund | 0,82 |
| 6 | view | 1 | spot | 1 | place | 0,81 |
| 7 | vibe | 1 | smoking | 1 | money | 0,80 |
| 8 | value | 1 | service | 1 | host | 0,78 |
| 9 | use | 1 | **room** | 1 | condition | 0,77 |
| 10 | trip | 1 | review | 1 | comment | 0,77 |

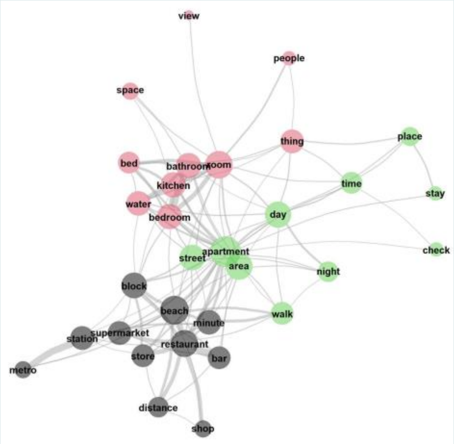
1.5

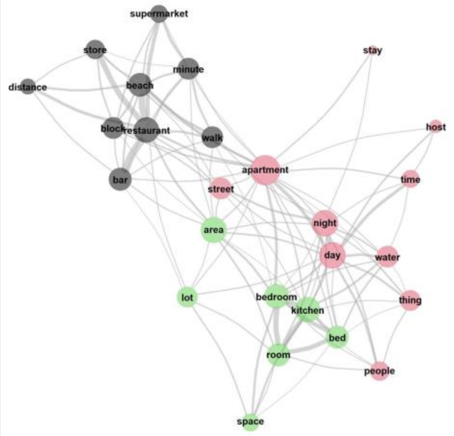
## Sentiment analysis

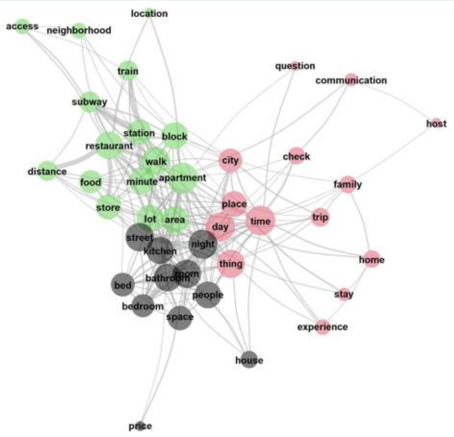
*Network diagrams*

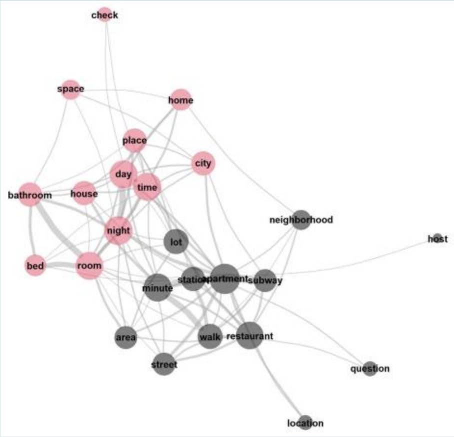
## Positive reviews

| | Pre-covid | Post-covid |
|---|---|---|



## Pre-covid

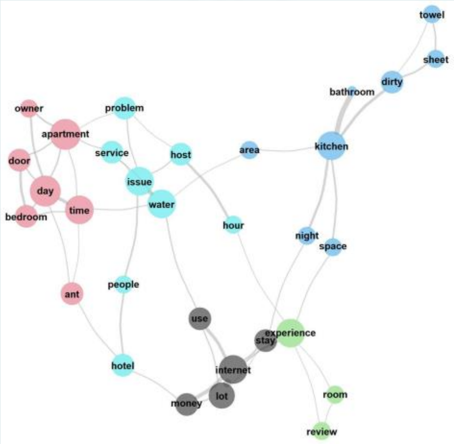
## Post-covid

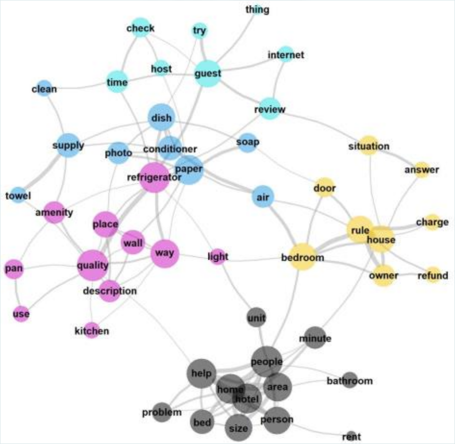
1.5

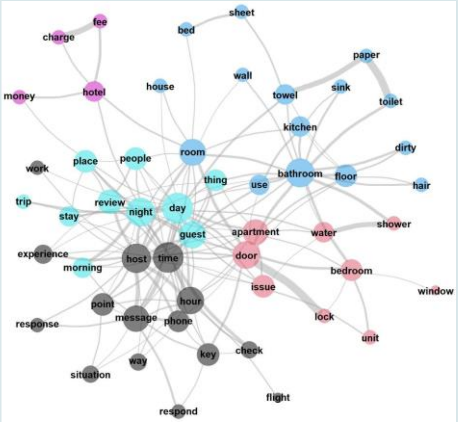
# Sentiment analysis

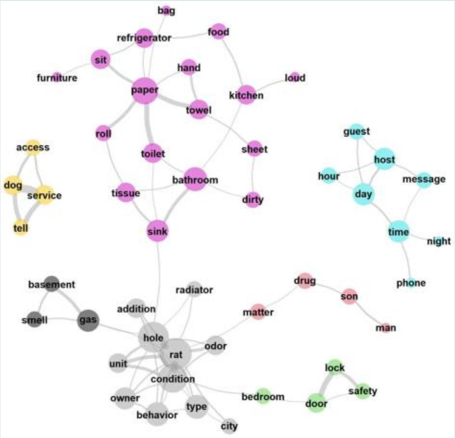
*Network diagrams*

## Negative reviews



2

**Covid reviews**

*Sentiment analysis*

**Table 12.** Frequency of covid terms

| | Rio de Janeiro | | New York | |
|---|---|---|---|---|
| Covid-term | Count | Percent (%) | Count | Percent (%) |
| corona | 16 | 15 | 58 | 4 |
| coronavirus | 13 | 12 | 99 | 7 |
| covid | 14 | 13 | 319 | 23 |
| lockdown | 4 | 4 | 59 | 4 |
| mask | 9 | 8 | 132 | 10 |
| pandemic | 34 | 31 | 579 | 42 |
| quarantine | 20 | 18 | 136 | 10 |
| Total | 110 | 100 | 1382 | 100 |

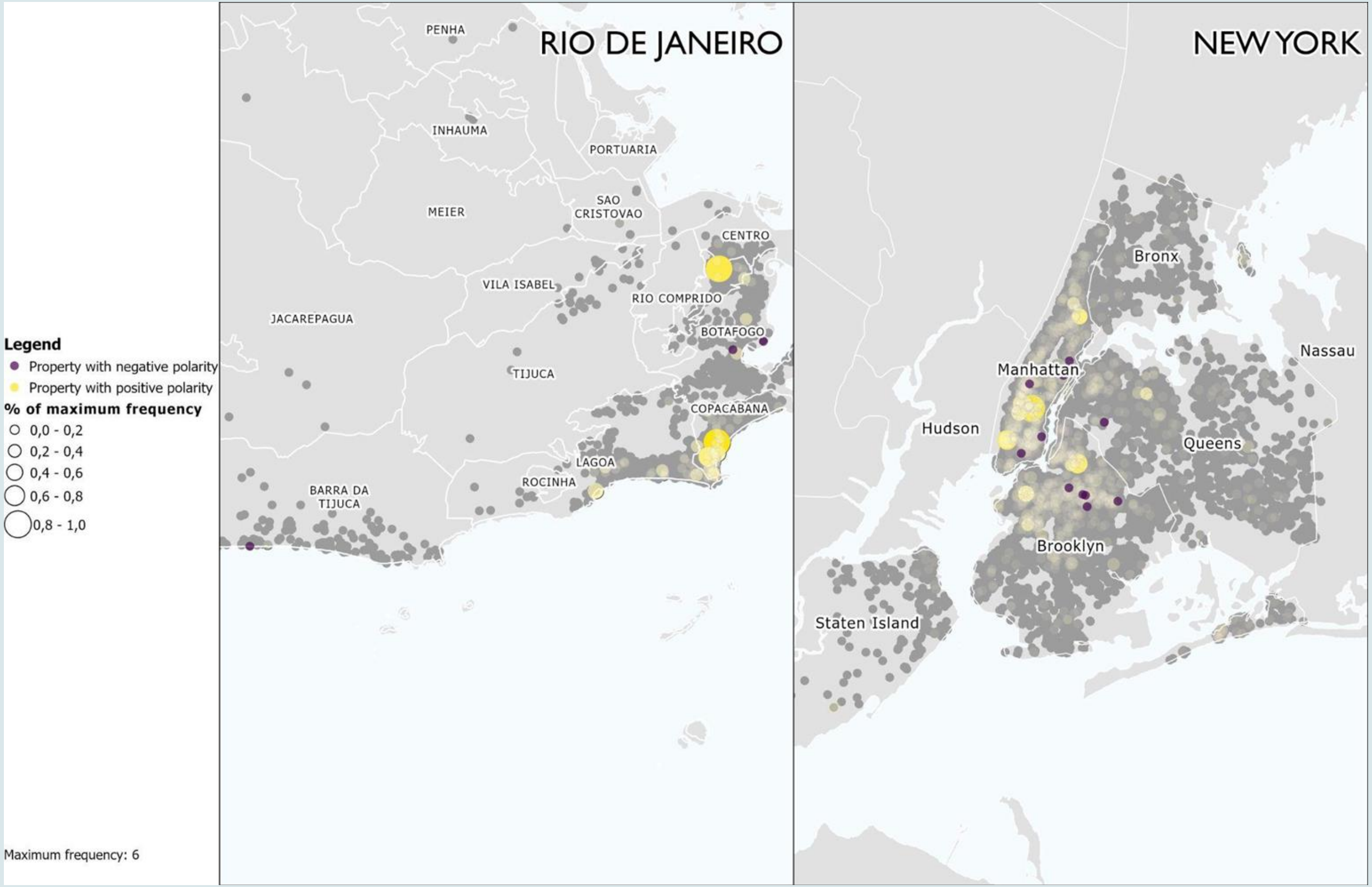
**Table 13.** Number and percentage of covid reviews per sentiment polarity from properties in RJ and NY.

| | Rio de Janeiro | | New York | |
|---|---|---|---|---|
| Sentiment | Count | Percent (%) | Count | Percent (%) |
| Positive | 74 | 87 | 1098 | 89,3 |
| Neutral | | | 5 | 0,4 |
| Negative | 11 | 13 | 126 | 10,3 |
| Total | 85 | 100 | 1229 | 100 |

2.1

## Covid reviews

*Spatial visualization*



| | Positive | | | Negative | | |
|---|---|---|---|---|---|---|
| No. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. |
| 1 | **place** | 72 | 1 | **apartment** | 14 | 1 |
| 2 | **apartment** | 55 | 0,76 | **host** | 13 | 0,93 |
| 3 | stay | 31 | 0,43 | owner | 11 | 0,79 |
| 4 | location | 27 | 0,38 | **place** | 10 | 0,71 |
| 5 | **host** | 26 | 0,36 | house | 9 | 0,64 |
| 6 | beach | 23 | 0,32 | air | 9 | 0,64 |
| 7 | day | 17 | 0,24 | people | 8 | 0,57 |
| 8 | **time** | 16 | 0,22 | rule | 7 | 0,50 |
| 9 | month | 15 | 0,21 | refund | 7 | 0,50 |
| 10 | view | 15 | 0,21 | **time** | 7 | 0,50 |

2.2

**Covid reviews**

*Frequent terms*

| | Positive | | | Neutral | | | Negative | | |
|---|---|---|---|---|---|---|---|---|---|
| No. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. | Term | Freq. | R. Freq. |
| 1 | **place** | 1038 | 1 | bed | 3 | 1 | **host** | 104 | 1 |
| 2 | **apartment** | 582 | 0,56 | **room** | 3 | 1 | **room** | 96 | 0,92 |
| 3 | **stay** | 515 | 0,50 | ***covid*** | 3 | 1 | **place** | 96 | 0,92 |
| 4 | **host** | 478 | 0,46 | **day** | 2 | 0,67 | **day** | 90 | 0,87 |
| 5 | **time** | 473 | 0,46 | king | 2 | 0,67 | **time** | 76 | 0,73 |
| 6 | location | 368 | 0,35 | cabin | 1 | 0,33 | **apartment** | 76 | 0,73 |
| 7 | **room** | 353 | 0,34 | ceiling | 1 | 0,33 | **stay** | 60 | 0,58 |
| 8 | space | 341 | 0,33 | hostel | 1 | 0,33 | night | 53 | 0,51 |
| 9 | home | 280 | 0,27 | dorm | 1 | 0,33 | people | 46 | 0,44 |
| 10 | **day** | 263 | 0,25 | plug | 1 | 0,33 | issue | 37 | 0,36 |

## 2.3

## Covid reviews

*Rare terms*

| | Positive | | Negative | |
|---|---|---|---|---|
| No. | Term | Score | Term | Score |
| 1 | wait | 0,75 | ***covid*** | 0,476 |
| 2 | condo | 0,73 | ***quarantine*** | 0,456 |
| 3 | serve | 0,72 | home | 0,456 |
| 4 | house | 0,69 | friend | 0,456 |
| 5 | charm | 0,67 | apartment | 0,451 |
| 6 | begin | 0,65 | check | 0,418 |
| 7 | help | 0,64 | travel | 0,407 |
| 8 | deck | 0,58 | value | 0,391 |
| 9 | window | 0,57 | iron | 0,391 |
| 10 | studio | 0,57 | ice | 0,391 |

| | Positive | | Neutral | | Negative | |
|---|---|---|---|---|---|---|
| No. | Term | Score | Term | Score | Term | Score |
| 1 | home | 1 | stay | 0,84 | ***covid*** | 1 |
| 2 | eats | 1 | day | 0,65 | property | 0,79 |
| 3 | birthday | 1 | king | 0,57 | cancellation | 0,78 |
| 4 | cleaner | 0,96 | ***covid*** | 0,54 | scruple | 0,76 |
| 5 | value | 0,94 | room | 0,45 | crew | 0,74 |
| 6 | ***coronavirus*** | 0,93 | bed | 0,45 | conference | 0,70 |
| 7 | guideline | 0,92 | sleep | 0,35 | unit | 0,69 |
| 8 | covid | 0,90 | plug | 0,35 | people | 0,67 |
| 9 | thanks | 0,89 | hostel | 0,35 | rental | 0,65 |
| 10 | umbrella | 0,86 | eye | 0,35 | host | 0,64 |

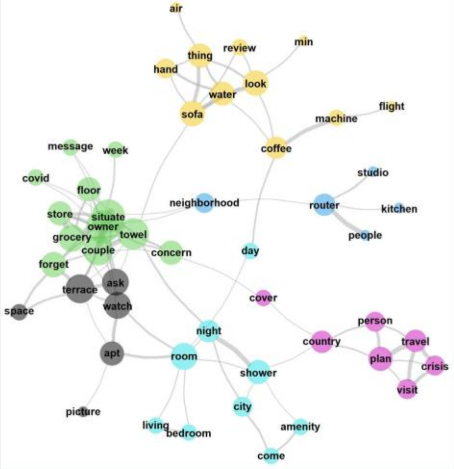
# 2.4

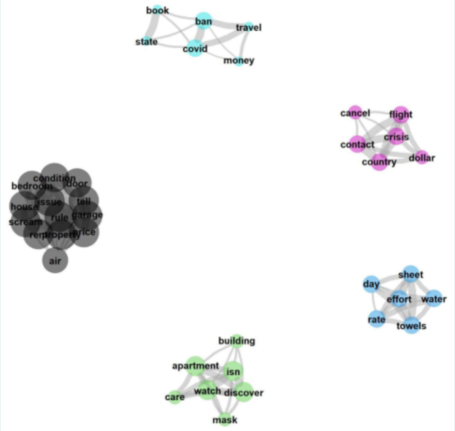
## Covid reviews

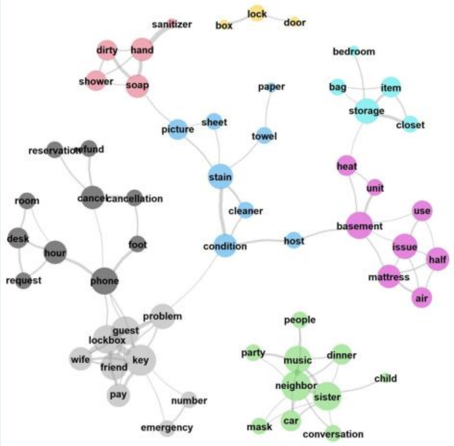
*Network diagrams*

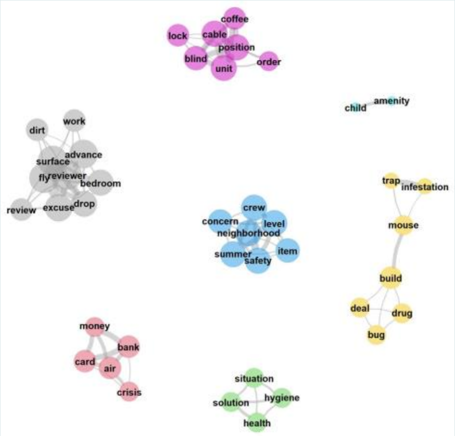
### Positive reviews

### Negative reviews



- Not enough evidence to claim that after the outbreak of covid there was a significant change in the experience of Airbnb users in RJ and NY.
- Nevertheless, topics from covid reviews suggest that users experienced situations related with the health crisis (e.g., use of mask and hand sanitizer, etc).
- Further analyses are required to verify whether the different situations experienced after the outbreak can be linked to the Covid crisis.
- On the other hand, it was possible to visualize changes in the concentration of properties associated with overall positive, neutral and negative user experiences.
- This paper can be useful in the lodging industry to uncover other aspects of the user experience, for instance, future studies could explore the role of location of lodging places in the customer behaviour during health crises.

# References


Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014, May, 216–225.

Hwang, J., & Seo, S. (2016). A critical review of research on customer experience management: Theoretical, methodological and cultural perspectives. International Journal of Contemporary Hospitality Management, 28(10), 2218–2246. https://doi.org/10.1108/IJCHM-04-2015-0192

Lillo, A. (2020). COVID-19, the beer flu; Or, the disease of many names. Lebende Sprachen, 49(5), 411–438. https://doi.org/10.1515/les-2020-0021

*Thank you for your attention.*