Airbnb Text and Data Analysis Report

Executive Summary

This report analyzes Airbnb listing descriptions using text mining in R and data visualization in Tableau. The goal is to uncover business insights regarding pricing, customer sentiment, and listing details. Text mining includes word frequency, sentiment analysis, and topic modeling to identify key trends. Tableau visualizations help explore pricing trends, review scores, and amenity preferences. These insights will help optimize Airbnb listing descriptions, pricing strategies, and customer experience.

Text Mining Insights

Word Frequency Analysis: Most frequent words include 'apartment,' 'kitchen,' and 'walk,' emphasizing room types and location convenience.

Sentiment Analysis: Listings use mostly positive sentiment words, with high trust and anticipation scores.

Topic Modeling (LDA): Five key themes include location, room features, guest experience, amenities, and booking details.

Tableau Visualizations

Price Distribution by Property Type: Villas and Boutique Hotels are the most expensive, while shared rooms are the cheapest.

Average Price by Property Type: Entire homes and boutique hotels have the highest average prices.

Review Scores by Property Type: Boutique hotels, guest suites, and apartments have higher ratings, while shared rooms have lower ratings.

Popular Amenities: Wi-Fi and Kitchen are the most listed amenities, while luxury features are less common.

Dashboard Description

The Tableau dashboard integrates all visualizations, allowing users to interactively explore pricing trends, review distributions, and amenity preferences across different property types.

Key Business Insights

Hosts emphasize location, amenities, and guest experience in their descriptions.

Positive sentiment is associated with higher reviews, while shared accommodations have slightly lower sentiment scores.

Pricing trends show Villas and Boutique Hotels command higher prices, while shared rooms remain budget friendly.

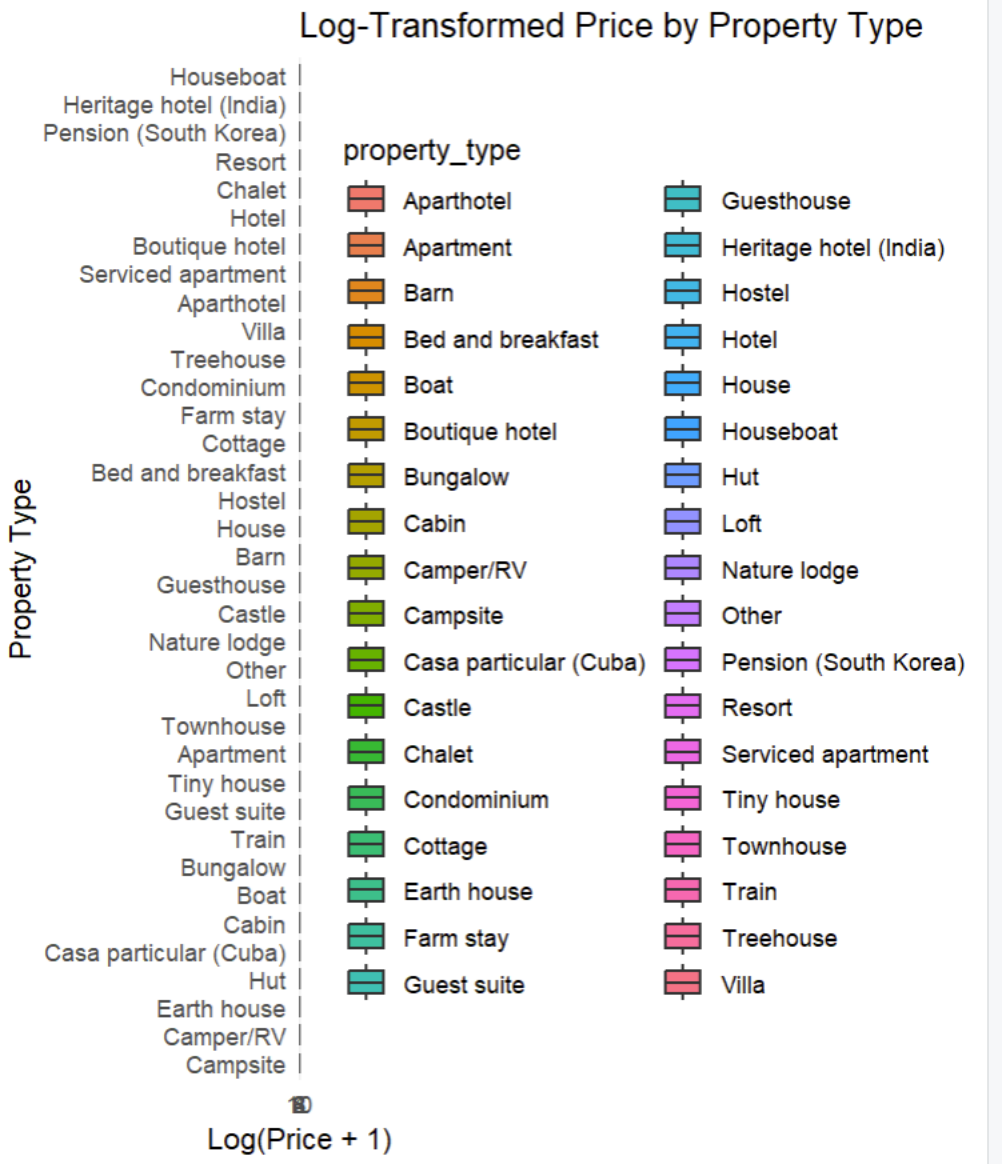
Listings can be optimized using engaging language and promoting high-value amenities.

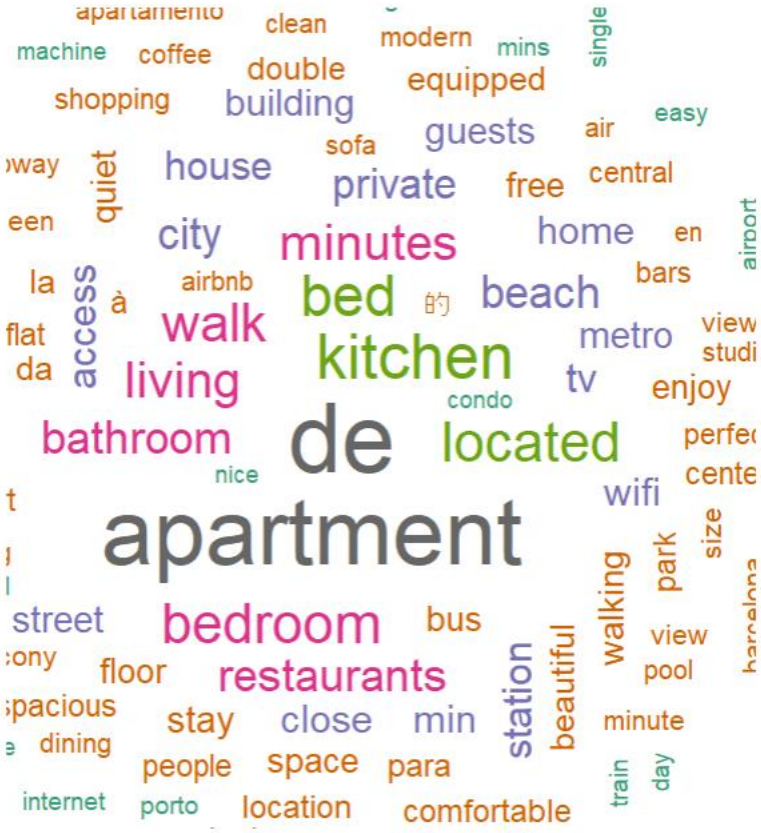
Conclusion

This analysis provides data-driven insights for Airbnb to enhance listings, improve pricing strategies, and optimize guest experiences. The findings from text mining and Tableau visualizations highlight key factors that influence customer preferences and property success.

Appendix

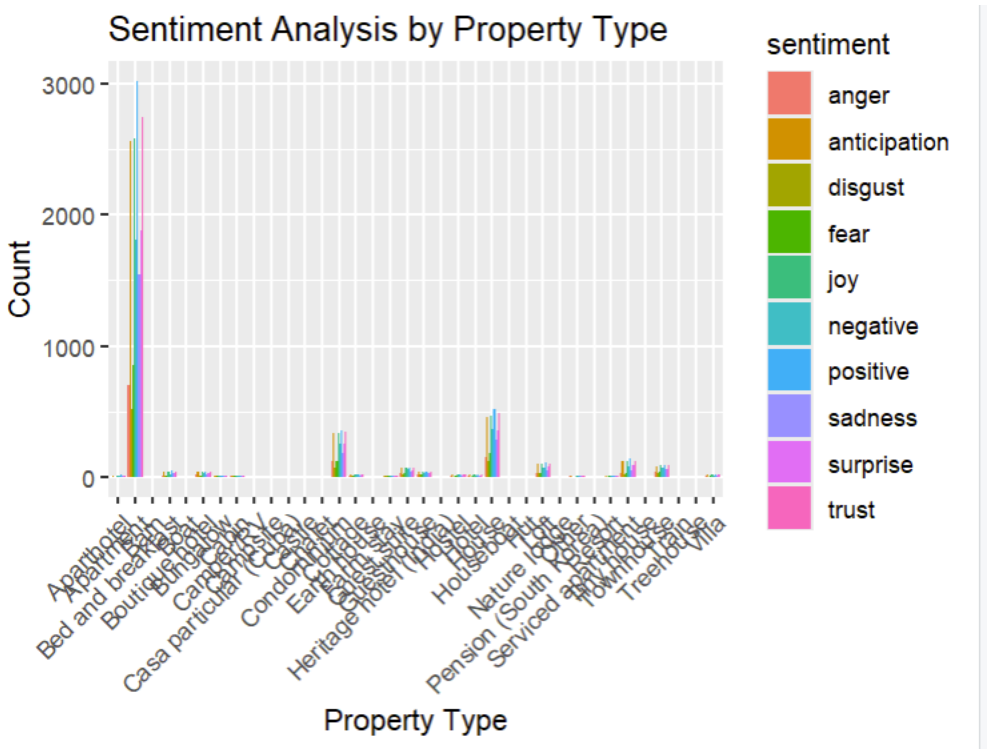
Screenshots of Tableau visualizations and R script used for text analysis are included in the appendix.

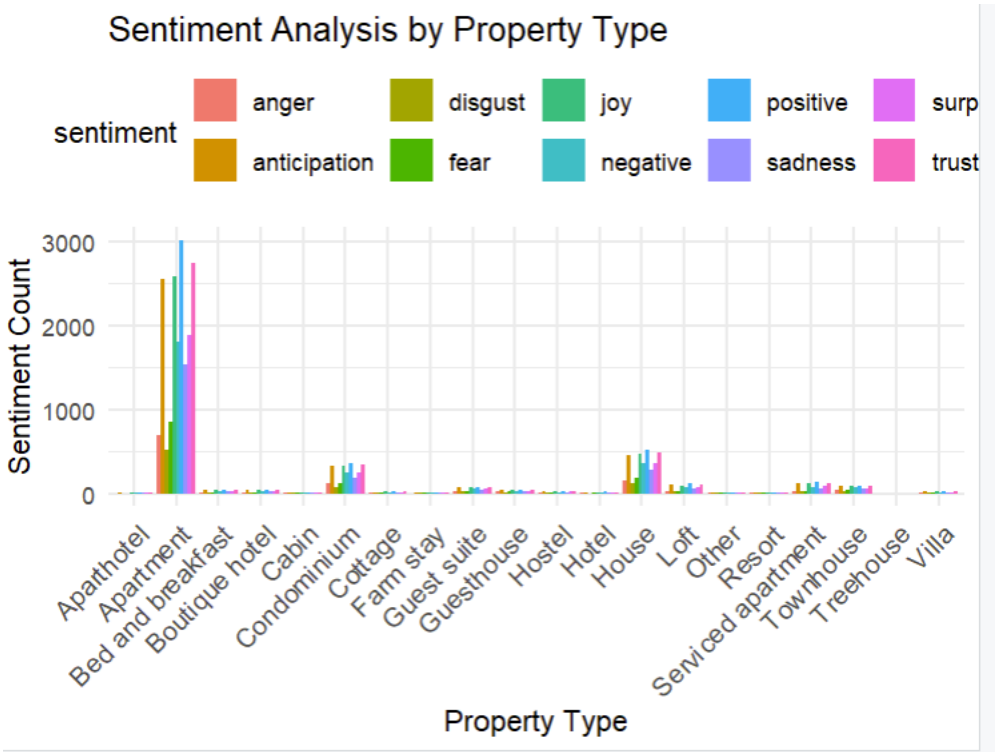

Log-Transformed Price by Property Type
Houseboat
Heritage hotel (India)
Pension (South Korea)
Resort
Chalet
Hotel
Boutique hotel
Serviced apartment
Aparthotel
Villa
Treehouse
Condominium
Farm stay
Cottage
Bed and breakfast
Hostel
House
Barn
Guesthouse
Castle
Nature lodge
Other
Loft
Townhouse
Apartment
Tiny house
Guest suite
Train
Bungalow
Boat
Cabin
Casa particular (Cuba)
Hut
Earth house
Camper/RV
Campsite
Property Type
property_type
Aparthotel
Apartment
Barn
Bed and breakfast
Boat
Boutique hotel
Bungalow
Cabin
Camper/RV
Campsite
Casa particular (Cuba)
Castle
Chalet
Condominium
Cottage
Earth house
Farm stay
Guest suite
Guesthouse
Heritage hotel (India)
Hostel
Hotel
House
Houseboat
Hut
Loft
Nature lodge
Other
Pension (South Korea)
Resort
Serviced apartment
Tiny house
Townhouse
Train
Treehouse
Villa
Log(Price + 1)

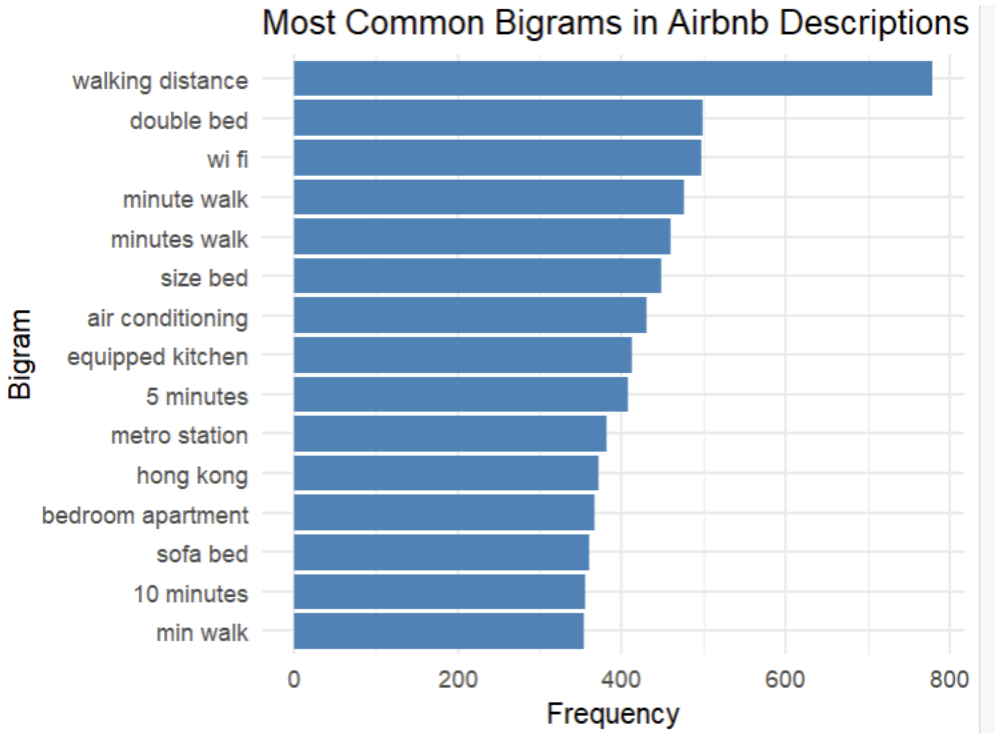

apartamento
clean
modern
mins
single
machine
coffee
double
equipped
shopping
building
guests
air
easy
sofa
house
private
free
central
quiet
city
minutes
home
en
airport
bars
la
access
à
airbnb
bed
beach
walk
metro
view
studi
flat
da
living
kitchen
tv
enjoy
condo
bathroom
de
located
perfe
cente
nice
wifi
apartment
park
size
walking
street
bedroom
bus
view
floor
restaurants
pool
station
beautiful
spacious
stay
close
min
minute
dining
people
space
para
train
day
internet
porto
location
comfortable

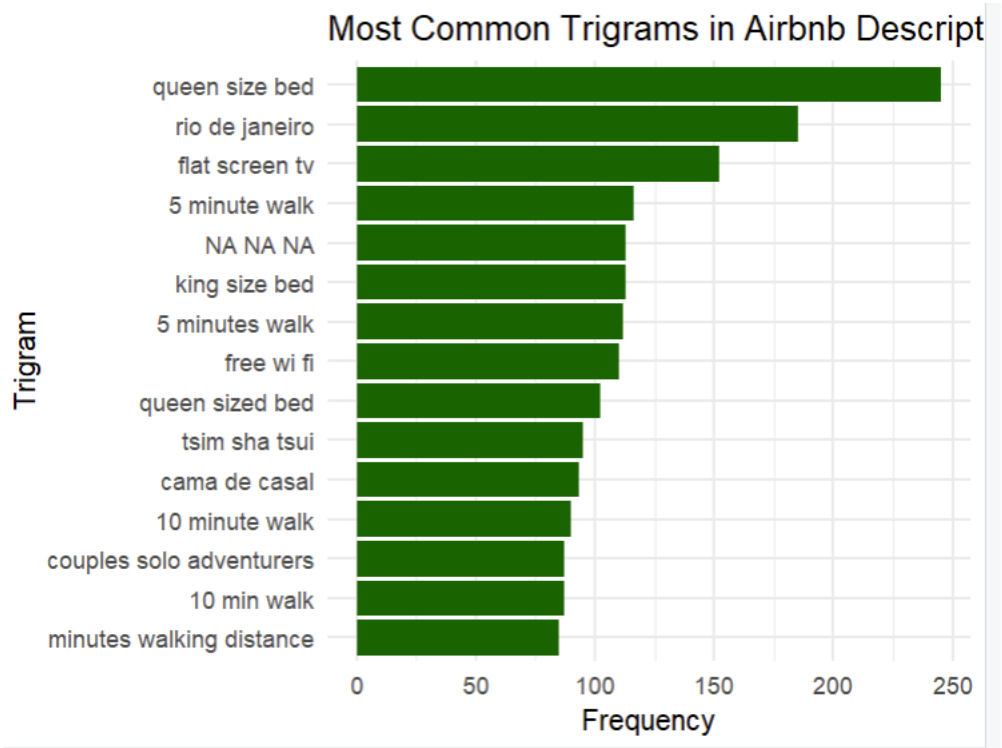

comfortable
spacious
bedrooms
location
kitchen
barcelona
studio
space
central
size
bedroom
distanc
ean
guest
quare
stay
perfect
flat
free
beach
le
garden
quiet
ocean
park
machine
house
heart
city
easy
la
para
à
shops
dining
en
guests
mins
的
min
street
air
nice
floor
da
home
bnb
bus
neighborhood
hed
single
airport
restaurants
pool
metro
close
minute
view
equipped
day
wifi
train
cozy
internet
shopping
shower
tv
modern
apartamento
local
de
porto
alcony
queen
building
towels
access
double
hidde
d

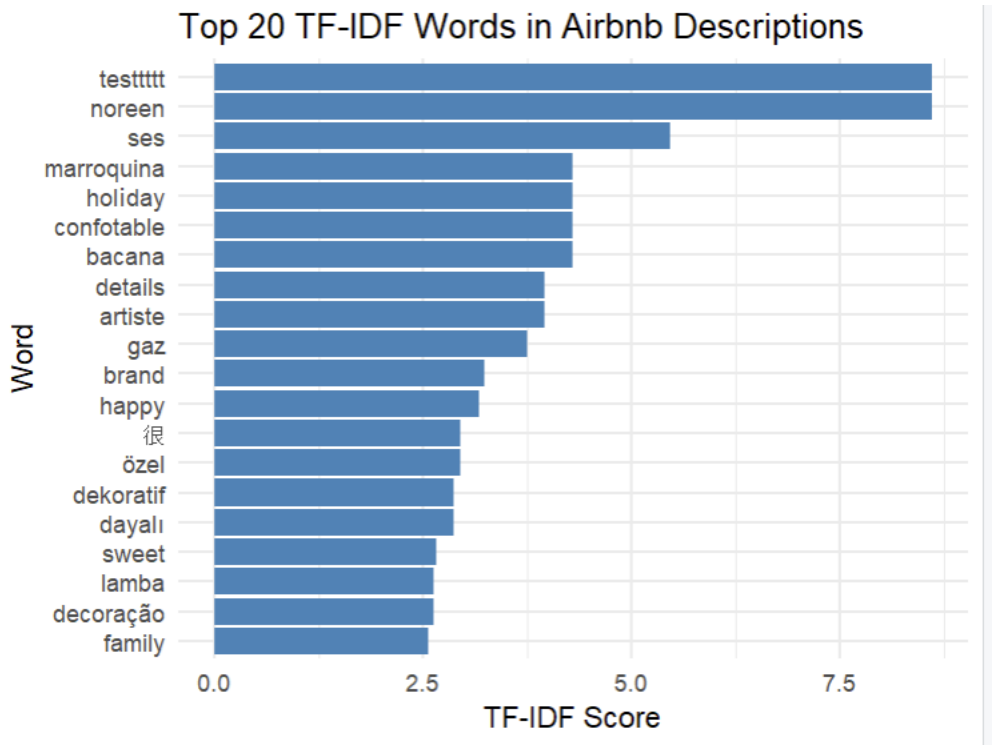

neighborhood
day
balcony
comfortable
garden
single
views
location
renovated
modern
quiet
view
building
distance
airport
spacious
street
guests
nb
bathroom
bars
mins
bus
en
enjoy
floor
studio
bedroom
apartment
home
da
à
beach
hidden
minute
close
heart
bed
wifi
walk
tv
metro
air
la
de
access
minutes
house
kitchen
private
para
min
located
station
stay
living
city
center
nice
flat
shops
rking
的
walking
park
a
restaurants
internet
double
space
free
equipped
size
people
pool
sofa

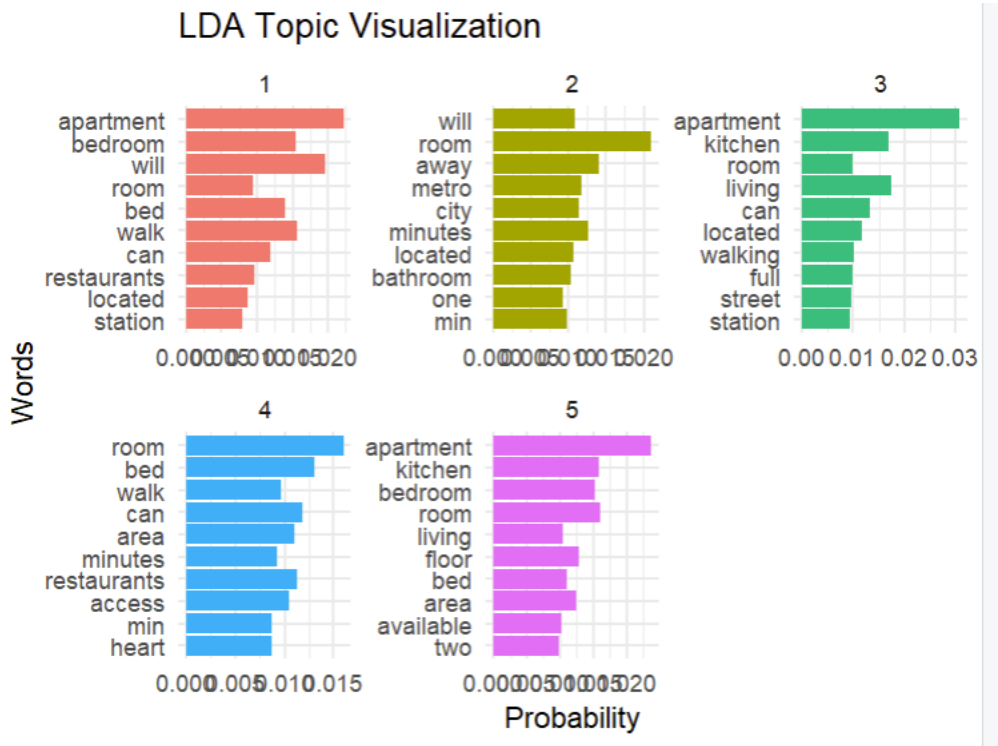

Sentiment Analysis by Property Type
sentiment
anger
anticipation
disgust
fear
joy
negative
positive
sadness
surprise
trust
Count
3000
2000
1000
0
Property Type

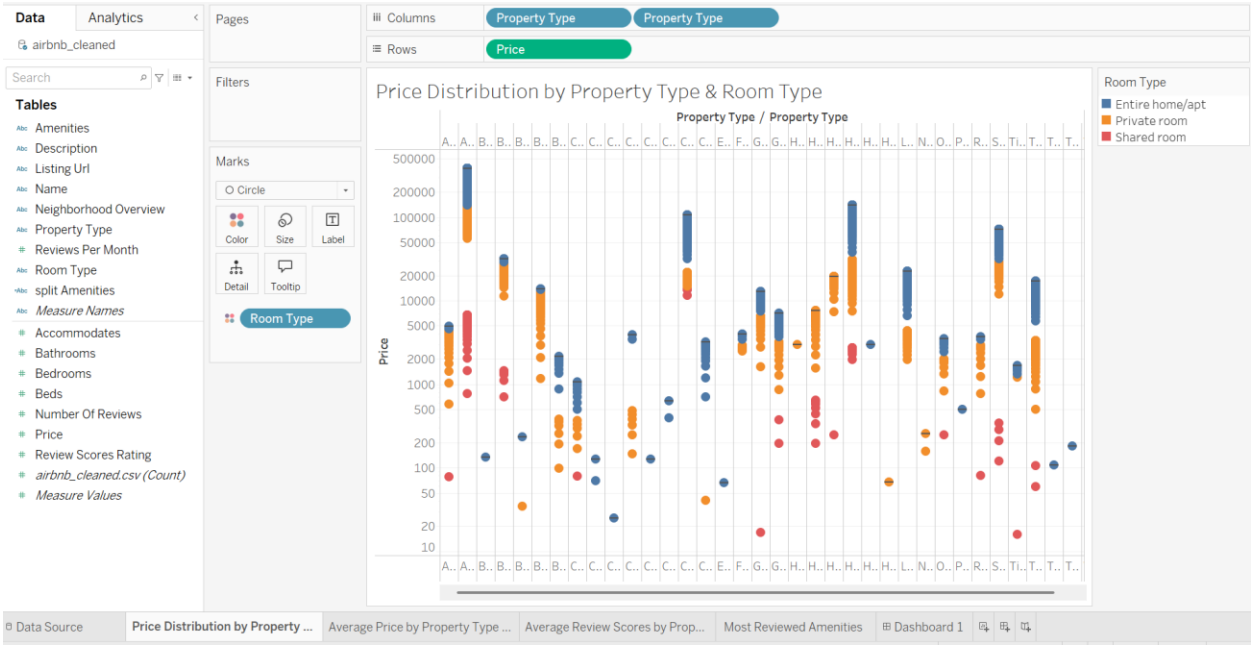

Sentiment Analysis by Property Type
sentiment
anger
anticipation
disgust
fear
joy
negative
positive
sadness
surp
trust
Sentiment Count
3000
2000
1000
0
Apartment
Bed and breakfast
Boutique hotel
Cabin
Condominium
Cottage
Farm stay
Guest suite
Guesthouse
Hostel
Hotel
House
Loft
Other
Resort
Serviced apartment
Townhouse
Treehouse
Villa
Property Type

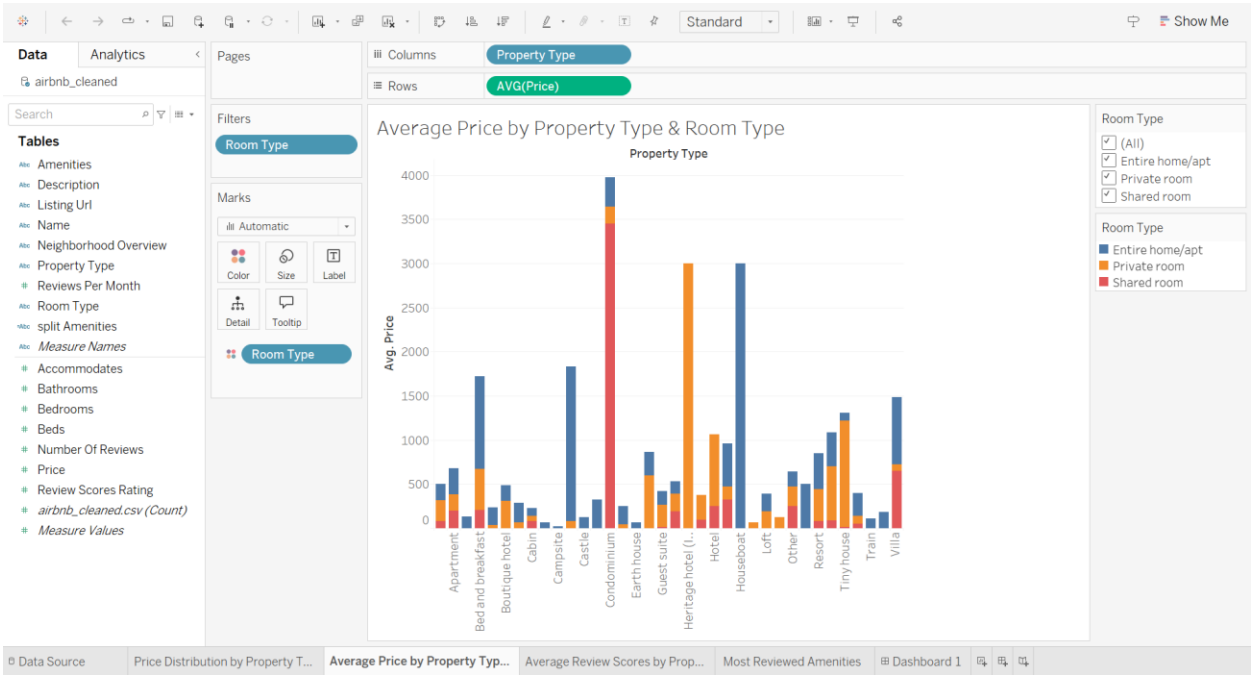

Most Common Bigrams in Airbnb Descriptions
walking distance
double bed
wi fi
minute walk
minutes walk
size bed
air conditioning
equipped kitchen
5 minutes
metro station
hong kong
bedroom apartment
sofa bed
10 minutes
min walk
Bigram
0
200
400
600
800
Frequency

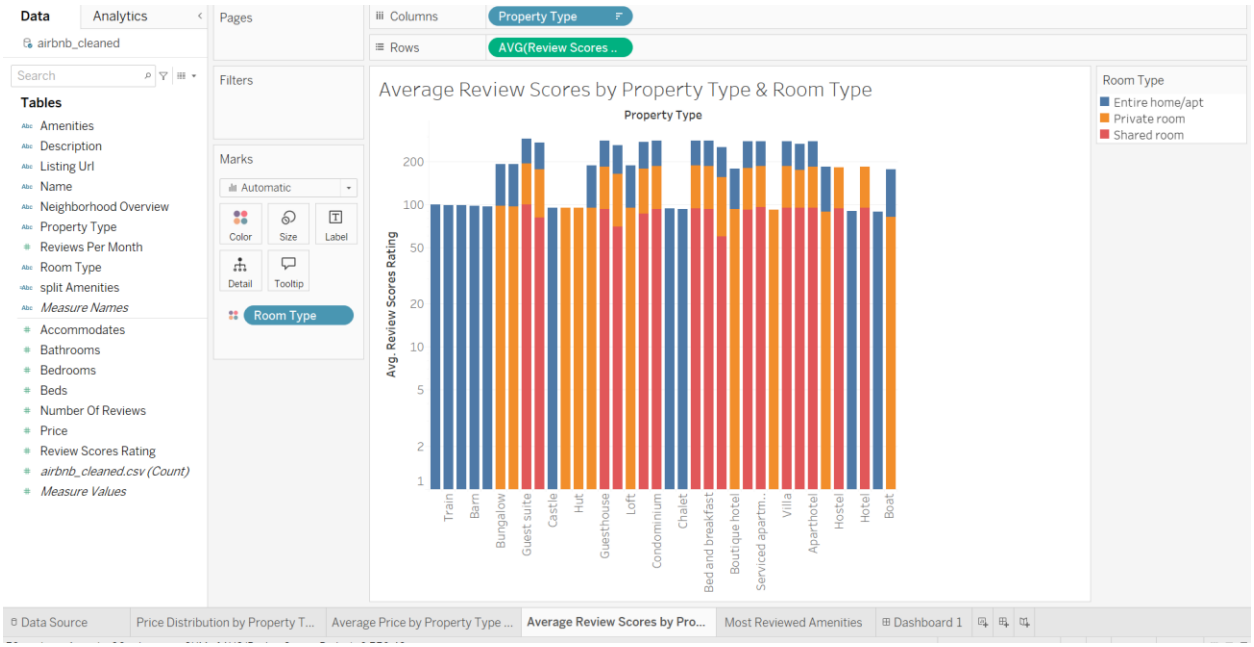

Most Common Trigrams in Airbnb Descript
queen size bed
rio de janeiro
flat screen tv
5 minute walk
NA NA NA
king size bed
5 minutes walk
free wi fi
queen sized bed
tsim sha tsui
cama de casal
10 minute walk
couples solo adventurers
10 min walk
minutes walking distance
Trigram
0
50
100
150
200
250
Frequency

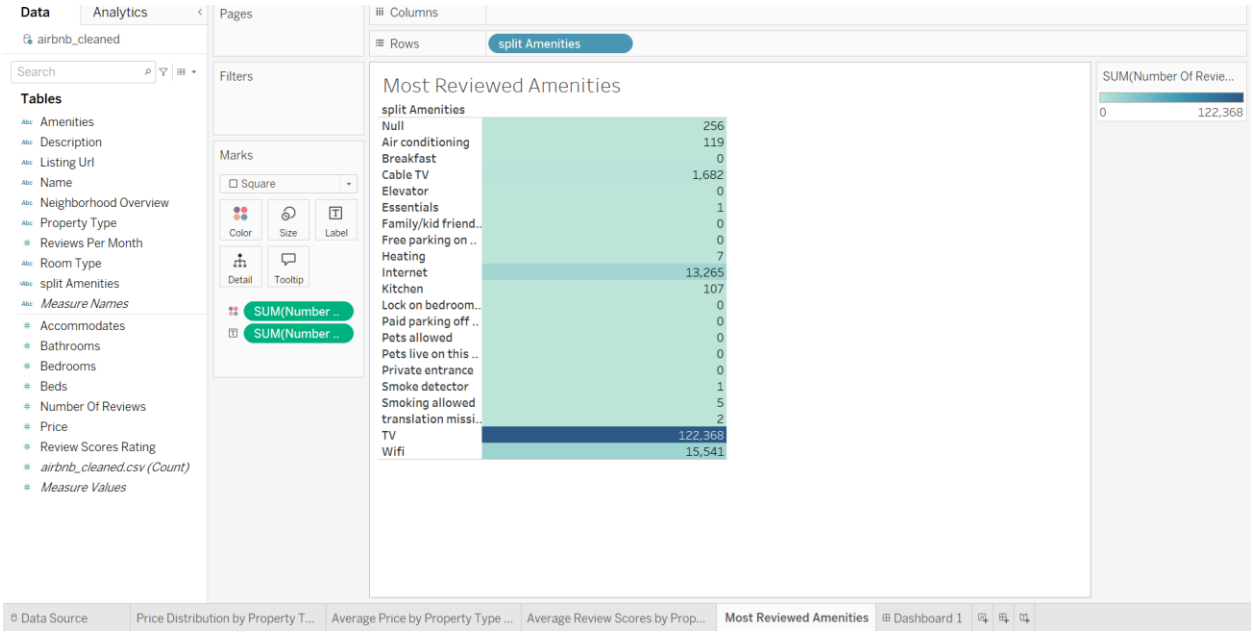

Top 20 TF-IDF Words in Airbnb Descriptions
testttt
noreen
ses
marroquina
holiday
confotable
bacana
details
artiste
gaz
brand
happy
很
özel
dekoratif
dayalı
sweet
lamba
decoração
family
Word
0.0
2.5
5.0
7.5
TF-IDF Score

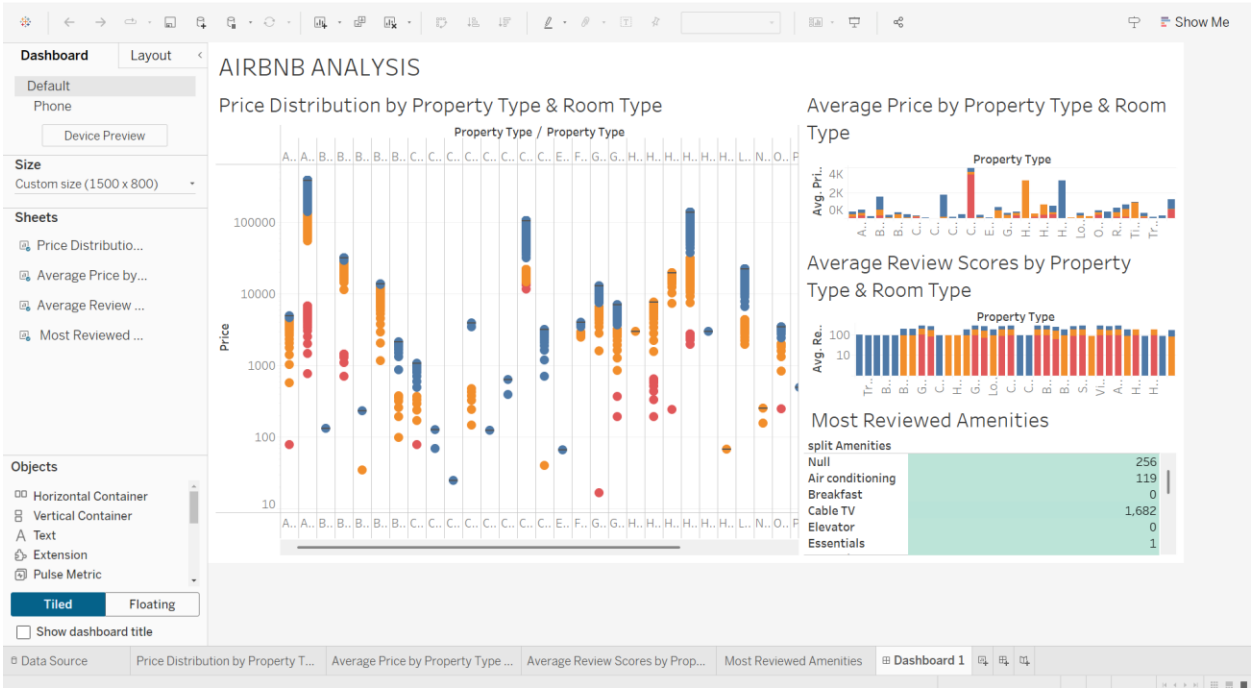

LDA Topic Visualization
1
apartment
bedroom
will
room
bed
walk
can
restaurants
located
station
2
will
room
away
metro
city
minutes
located
bathroom
one
min
3
apartment
kitchen
room
living
can
located
walking
full
street
station
0.00 0.01 0.02 0.03
4
room
bed
walk
can
area
minutes
restaurants
access
min
heart
0.000 0.005 0.010 0.015
5
apartment
kitchen
bedroom
room
living
floor
bed
area
available
two
Words
Probability


Data
Analytics
airbnb_cleaned
Search
Tables
Amenities
Description
Listing Url
Name
Neighborhood Overview
Property Type
Reviews Per Month
Room Type
split Amenities
Measure Names
Accommodates
Bathrooms
Bedrooms
Beds
Number Of Reviews
Price
Review Scores Rating
airbnb_cleaned.csv (Count)
Measure Values
Pages
Filters
Marks
Circle
Color
Size
Label
Detail
Tooltip
Room Type
Columns
Property Type
Property Type
Rows
Price
Price Distribution by Property Type & Room Type
Property Type / Property Type
Price
500000
200000
100000
50000
20000
10000
5000
2000
1000
500
200
100
50
20
10
Room Type
Entire home/apt
Private room
Shared room
Data Source
Price Distribution by Property ...
Average Price by Property Type ...
Average Review Scores by Prop...
Most Reviewed Amenities
Dashboard 1


Standard
Show Me
Data
Analytics
airbnb_cleaned
Search
Tables
Amenities
Description
Listing Url
Name
Neighborhood Overview
Property Type
Reviews Per Month
Room Type
split Amenities
Measure Names
Accommodates
Bathrooms
Bedrooms
Beds
Number Of Reviews
Price
Review Scores Rating
airbnb_cleaned.csv (Count)
Measure Values
Pages
Filters
Room Type
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Room Type
Columns
Property Type
Rows
AVG(Price)
Average Price by Property Type & Room Type
Property Type
Avg. Price
4000
3500
3000
2500
2000
1500
1000
500
0
Apartment
Bed and breakfast
Boutique hotel
Cabin
Campsite
Castle
Condominium
Earth house
Guest suite
Heritage hotel (i...
Hotel
Houseboat
Loft
Other
Resort
Tiny house
Train
Villa
Room Type
(All)
Entire home/apt
Private room
Shared room
Room Type
Entire home/apt
Private room
Shared room
Data Source
Price Distribution by Property T...
Average Price by Property Typ...
Average Review Scores by Prop...
Most Reviewed Amenities
Dashboard 1


Data
Analytics
airbnb_cleaned
Search
Tables
Amenities
Description
Listing Url
Name
Neighborhood Overview
Property Type
Reviews Per Month
Room Type
split Amenities
Measure Names
Accommodates
Bathrooms
Bedrooms
Beds
Number Of Reviews
Price
Review Scores Rating
airbnb_cleaned.csv (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Room Type
Columns
Property Type
Rows
AVG(Review Scores ..
Average Review Scores by Property Type & Room Type
Property Type
Avg. Review Scores Rating
200
100
50
20
10
5
2
1
Train
Barn
Bungalow
Guest suite
Castle
Hut
Guesthouse
Loft
Condominium
Chalet
Bed and breakfast
Boutique hotel
Serviced apartm..
Villa
Aparthotel
Hostel
Hotel
Boat
Room Type
Entire home/apt
Private room
Shared room
Data Source
Price Distribution by Property T...
Average Price by Property Type ...
Average Review Scores by Pro...
Most Reviewed Amenities
Dashboard 1


Data
Analytics
airbnb_cleaned
Search
Tables
Amenities
Description
Listing Url
Name
Neighborhood Overview
Property Type
Reviews Per Month
Room Type
split Amenities
Measure Names
Accommodates
Bathrooms
Bedrooms
Beds
Number Of Reviews
Price
Review Scores Rating
airbnb_cleaned.csv (Count)
Measure Values
Pages
Filters
Marks
Square
Color
Size
Label
Detail
Tooltip
SUM(Number ..
SUM(Number ..
Columns
Rows
split Amenities
Most Reviewed Amenities
split Amenities
Null 256
Air conditioning 119
Breakfast 0
Cable TV 1,682
Elevator 0
Essentials 1
Family/kid friend.. 0
Free parking on .. 0
Heating 7
Internet 13,265
Kitchen 107
Lock on bedroom.. 0
Paid parking off .. 0
Pets allowed 0
Pets live on this .. 0
Private entrance 0
Smoke detector 1
Smoking allowed 5
translation missi.. 2
TV 122,368
Wifi 15,541
SUM(Number Of Revie...
0
122,368
Data Source
Price Distribution by Property T...
Average Price by Property Type ...
Average Review Scores by Prop...
Most Reviewed Amenities
Dashboard 1



References

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. Computational Intelligence, 29(3), 436-465. https://doi.org/10.1111/j.1467-8640.2012.00460.x