



Airbnb Ratings Analysis

MKTG 712 Final Project

By Yuna Shu, Bingling Wang, Yixuan Wang, Ji Wu





Introduction



Think of craigslist meets hotels.com, but a lot less crappy

Josh Spear

A cool Alternative to a boring evening in hotel room

Mashable

AirBed & Breakfast is a fun approach to CouchSurfing

Webware



Objective

Business Question:

- What can Airbnb/hosts do to improve customer's homestay experience?

Business Meaning:

- Find out factors that affect review scores
- Help hosts better prepare for the customers
- Help Airbnb improve its brand image and brand quality



Data

Raw dataset

- *Kaggle* (compiled on November 7, 2021)
- The listing activity of Airbnb homestays in New Orleans, Louisiana, as part of the *Inside Airbnb* initiative
- 6028 obs. of 49 variables

Cleaned dataset

- Remove NAs, unrelated variables
- String variables → dummy variables
- 3458 obs. of 29 variables



Analysis Method

Linear Regression

- Find variables that can explain the outcome
- Dependent variable is continuous

DV

- Overall ratings score of the airbnb listing

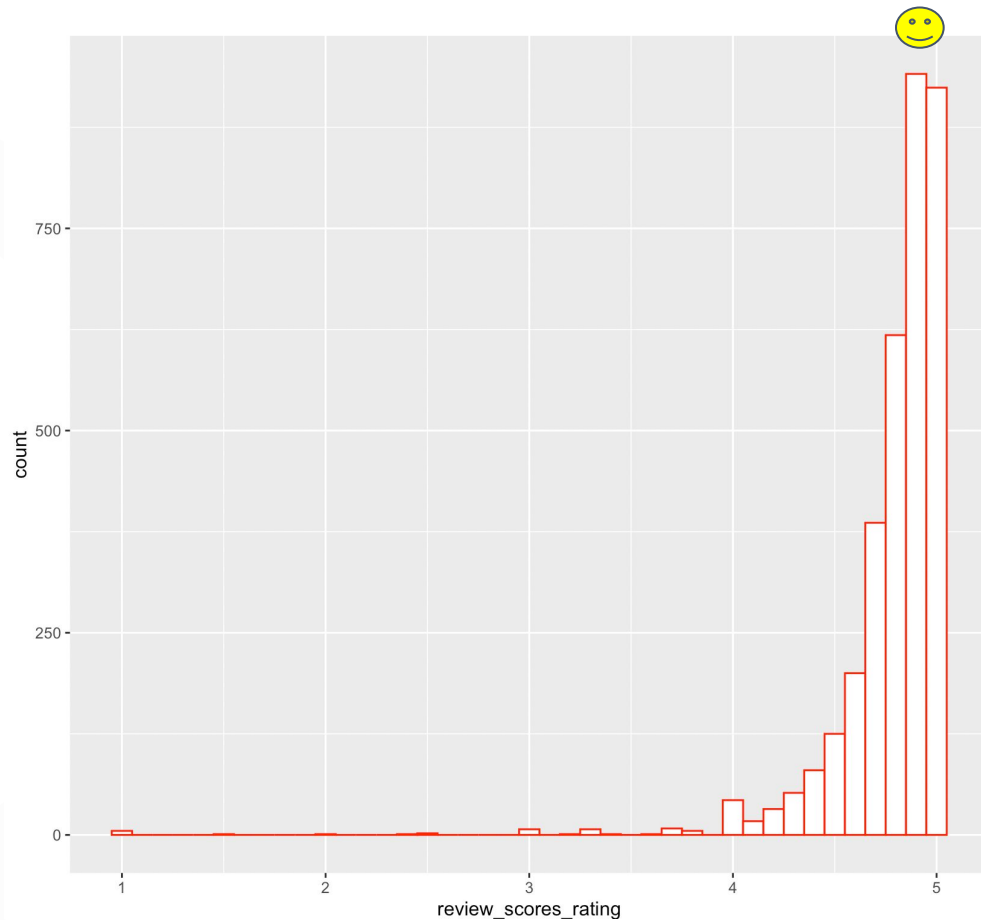
IVs

- Factors that may affect how customers rate the airbnb listings they stayed in



Variables

Dependent variable - *review_scores_rating*



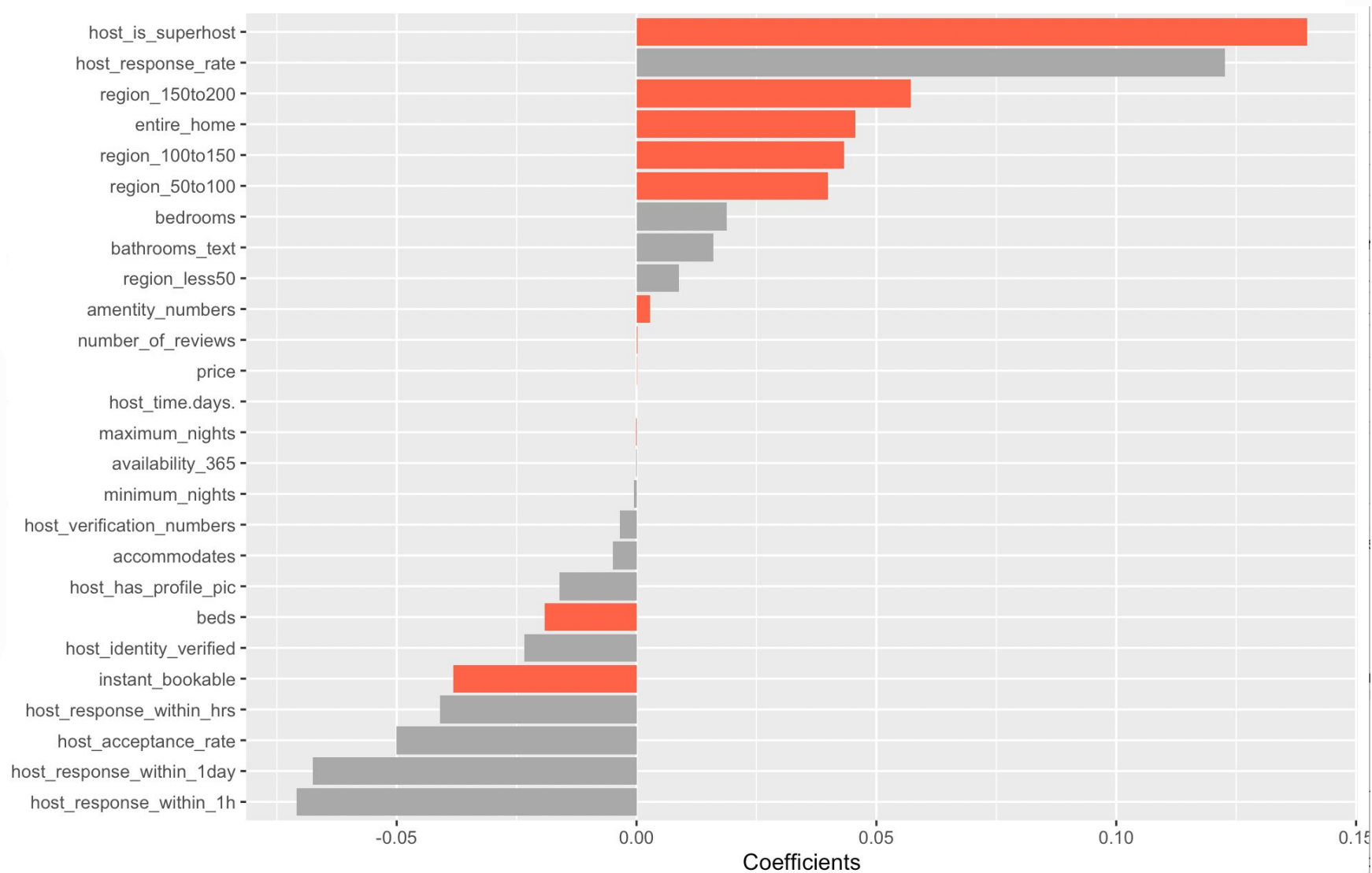
Independent variables

```
host_time.days.  
host_response_rate  
host_acceptance_rate  
host_is_superhost  
host_verification_numbers  
host_has_profile_pic  
host_identity_verified  
accommodates  
bathrooms_text  
bedrooms  
beds  
amentity_numbers  
price  
minimum_nights  
maximum_nights  
availability_365  
number_of_reviews  
instant_bookable  
host_response_within_1h  
host_response_within_hrs  
host_response_within_1day  
entire_home  
region_less50  
region_50to100  
region_100to150  
region_150to200
```

**Dummy variables!
(11)**



Statistical Results





Significant Variables

Variable	Coefficient
number_of_reviews	0.0002
amenity_numbers	0.0029
price	0.0001
maximum_nights	-0.0000
beds	-0.0192
instant_bookable (dummy)	-0.0383
host_is_superhost (dummy)	0.1398
entire_home (dummy)	0.0456
region_50to100 (dummy)	0.0399
region_100to150 (dummy)	0.0433
region_150to200 (dummy)	0.0572

Positive significant predictors



- Host is superhost
- Regions with more airbnb listings
- Entire home
- # of amenity

Negative significant predictors



- # of beds
- Instant bookable
- maximum nights



Significant Variables

Variable	Coefficient
number_of_reviews	0.0002
amenity_numbers	0.0029
price	0.0001
maximum_nights	-0.0000
beds	-0.0192
instant_bookable (<i>dummy</i>)	-0.0383
host_is_superhost (<i>dummy</i>)	0.1398
entire_home (<i>dummy</i>)	0.0456
region_50to100 (<i>dummy</i>)	0.0399
region_100to150 (<i>dummy</i>)	0.0433
region_150to200 (<i>dummy</i>)	0.0572

(baseline: region_over200)

Recommendation 1

When evaluating hosts in different neighborhoods, Airbnb should consider the levels of competition in the neighborhood.



Significant Variables

Variable	Coefficient
number_of_reviews	0.0002
amenity_numbers	0.0029
price	0.0001
maximum_nights	-0.0000
beds	-0.0192
instant_bookable (dummy)	-0.0383
host_is_superhost (dummy)	0.1398
entire_home (dummy)	0.0456
region_50to100 (dummy)	0.0399
region_100to150 (dummy)	0.0433
region_150to200 (dummy)	0.0572

Recommendation 2

To maintain the quality of the airbnb listings, Airbnb should develop more strict policies on the instant-book option.



Significant Variables

Variable	Coefficient
number_of_reviews	0.0002
amenity_numbers	0.0029
price	0.0001
maximum_nights	-0.0000
beds	-0.0192
instant_bookable (dummy)	-0.0383
host_is_superhost (dummy)	0.1398
entire_home (dummy)	0.0456
region_50to100 (dummy)	0.0399
region_100to150 (dummy)	0.0433
region_150to200 (dummy)	0.0572

Recommendation 3

Airbnb should ask the hosts to specify what kind of “beds” they are providing.



Applications & Impacts



Host

A guide for them to
improve rating scores



Airbnb

Better manage hosts
to improve customer
experience



Traveller

Critically look at the
rating scores



Limitations and Improvements



Limitation 1: Localization



Apply the results with caution

- Regional differences



Limitation 2: Low R-Square



Include more variables in the model

- Parking
- Travel attractions nearby
- Sound proof
- Check-in/-out convenience