



Building Trust

between hosts and guests



Contents

1	Trust as a currency Introduction	10	Communication is key Evaluating text practices
3	Measuring trust Identifying metrics	16	Recommendations Conclusion
4	What makes a super host? Predicting the ideal host	19	Appendices (A-F)



Trust as a currency



Airbnb is a two-sided marketplace that allows hosts who have spare rooms to accommodate travelers who need a place to stay. Guests pay hosts a fee for their hospitality, but **the most important currency that is exchanged is trust**.

First and foremost, travelers must be comfortable with living in the homes of strangers, often while they are visiting a place they've never been to before.

The trust needed in such a relationship can be difficult to gain— especially through an online platform.

So how do Airbnb hosts do it?



The question

How do hosts build trust?

This report seeks to understand how Airbnb hosts use the platform to build trust with current and potential guests.

Then, it recommends steps that hosts can take to earn more of this crucial currency.

The data



All analyses is based on 2019 data for New York City, which was sourced from Inside Airbnb.¹ The data includes 74 attributes for each of 50,718 listings.

Due to computational constraints, the data was cleaned and subsetted before use. This process is found in Appendix A.

Measuring trust

Superhosts are among the most trusted members of the Airbnb platform.² Hosts with this status have met four requirements that demonstrate their competence, consistency, and character. See table ▶

To confirm that Superhost status was a valuable metric for this report, I performed bivariate analyses between this attribute and hosts' review scores (Appendix B). Superhosts had statistically significant higher mean scores than their not-so-super peers, which suggested that the former group performed better with guests.

Thus, I set **Superhost status as the target variable** in my machine learning models.

Superhost requirement	Trust element ³
Complete at least 10 trips or 3 reservations totalling 100 nights	Competence Ability to do tasks well
Maintain a 90% response rate or higher	Consistency Acts in a reliable way
Maintain a 1% cancellation rate or lower	
Maintain a 4.8 overall review rating	Character Has a good reputation

What makes a super host?

I ran machine learning models to predict how different factors would affect a host's likelihood of being a Superhost— that is, a trusted member of the platform.

First, all variables were considered by their relevance to the topic of trust. For example, “host identity verified” was more appropriate for the model than “number of beds.” Next, possible predictors were tested against the target using bivariate analysis to justify their inclusion in the models (see Appendix C).

In total, sixteen predictors were chosen.
See figure ►



Maximized prediction results

By changing the input values in the Bootstrap Forest model's prediction profiler, the likelihood of an Airbnb host being a Superhost was increased from 30.8% to 69.2%.

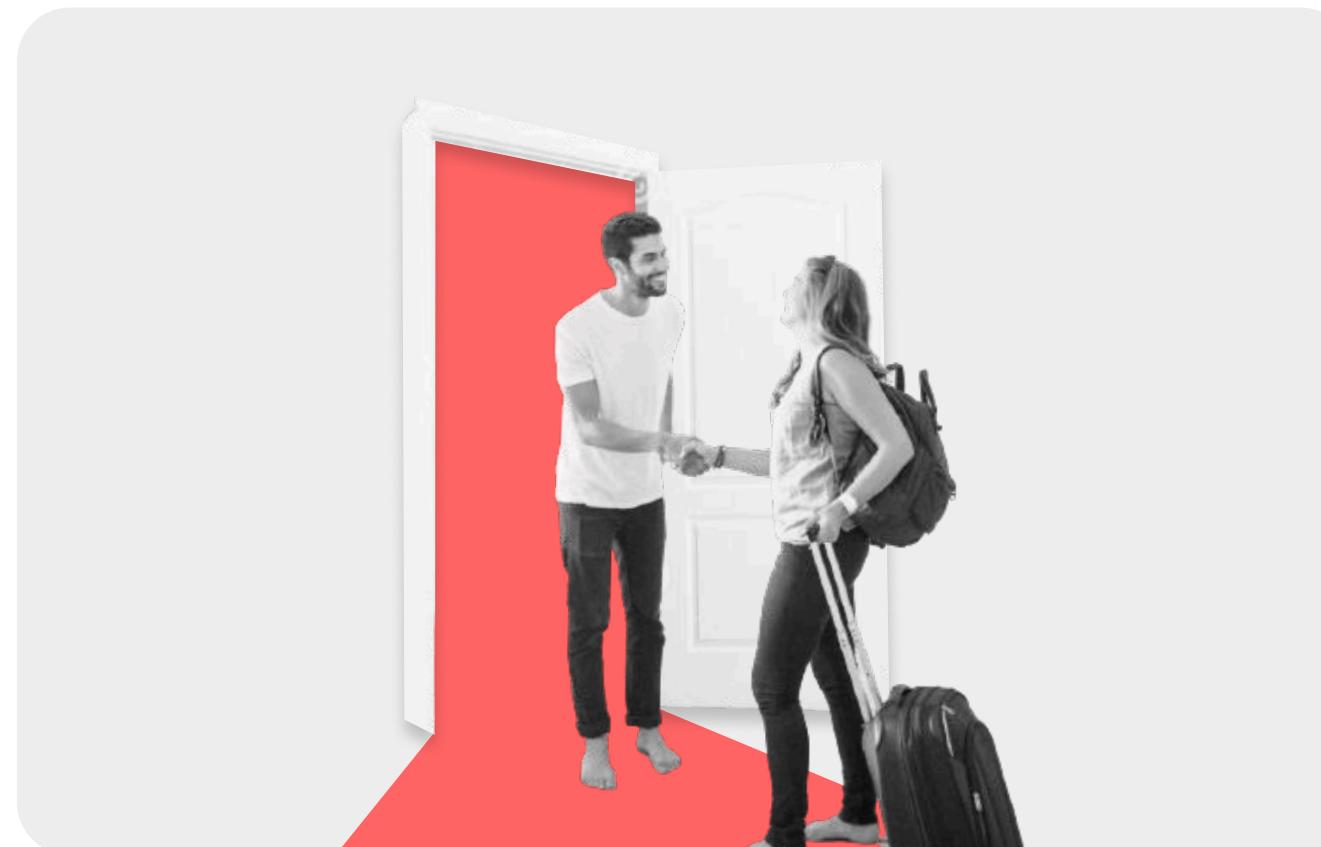


Main findings

The optimized model revealed the following:

1. Guests prefer transparency

Guests want the person greeting them at the door to match the profile that they see on Airbnb. A host who has verified his or her identity and includes a profile photo is 4% more likely to be a Superhost than one who doesn't.



Moreover, hosts may withhold their listing address until the end of the booking process, but guests worry that this practice is misleading.⁵ They feel safer if they can search for the exact location before booking.

2. Reviews help, up to a point

Hosts who have received feedback from past guests are deemed more trustworthy. Specifically, hosts with around 125 reviews have a 14% greater chance of being a Superhost than those with no reviews.

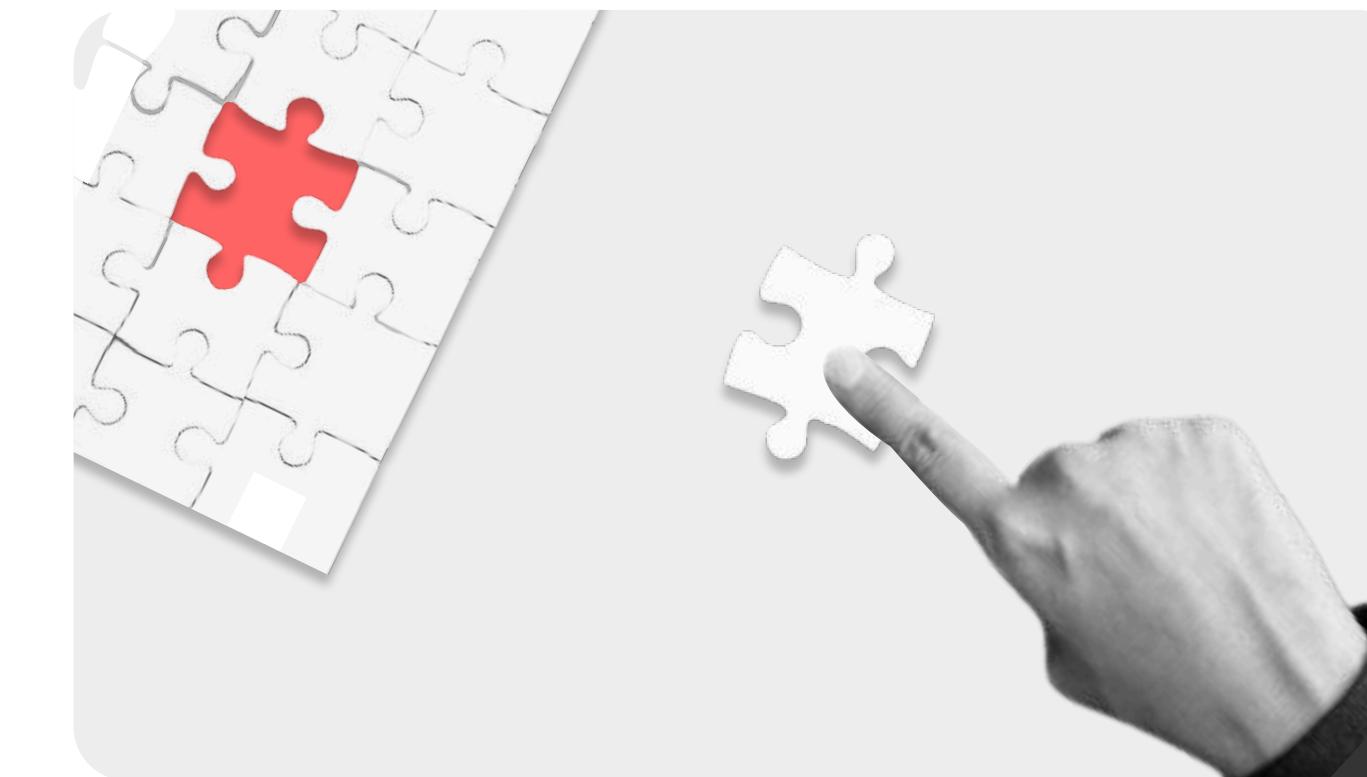
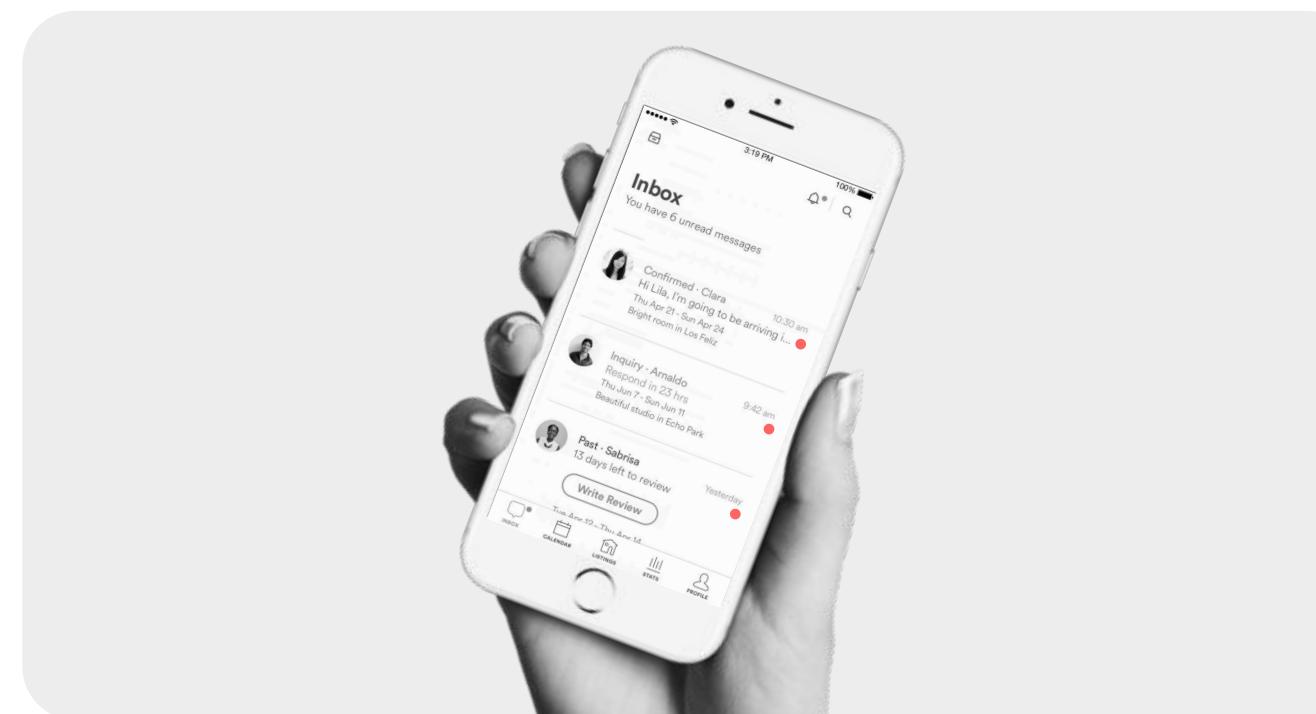
A host's status is less influenced by their review count once they pass that number. However, having too few reviews adversely affects perceived trustworthiness.

Main findings

The optimized model revealed the following:

3. Responding sooner is better

A quick and consistent response shows guests that the host values their interest in the listing and takes his or her role seriously. Hosts who reply within a few hours and to nearly all inquiries (a 95% response rate or higher) are 9% more likely to be Superhosts than those who reply after a few days.



4. Page completion matters

When hosts complete their profiles and list the interactions that they typically have with guests, their likelihood of being a Superhost increases by 11%. These communications are optional, but they can help to make the host seem more familiar and approachable.



Host about

Hosted by Min
Joined in September 2018

★ 278 Reviews 🛡️ Identity verified

Hi, I'm Min. I've lived in Soho for 5+ years. I love hosting because it's a great way to meet people from all over! I enjoy all things art, from painting to Broadway to ballet.

During your stay

I'm available to help you during your stay. I love this city and want others to experience it fully. I'm happy to guide you if I'm free and help you make the most of your trip!

Guest interactions

Communication is the key



The model predictions show that hosts' online behaviors, including how often they respond and how much they share, are important determinants of trust. But what exactly do good hosts share?

To understand their communications practices, I applied text mining to hosts' profiles and the interactions section of their listings (i.e. "During your stay"). In the data set, these two columns seemed most directly related to the hosts and, thus, most indicative of their trustworthiness.

Host profile

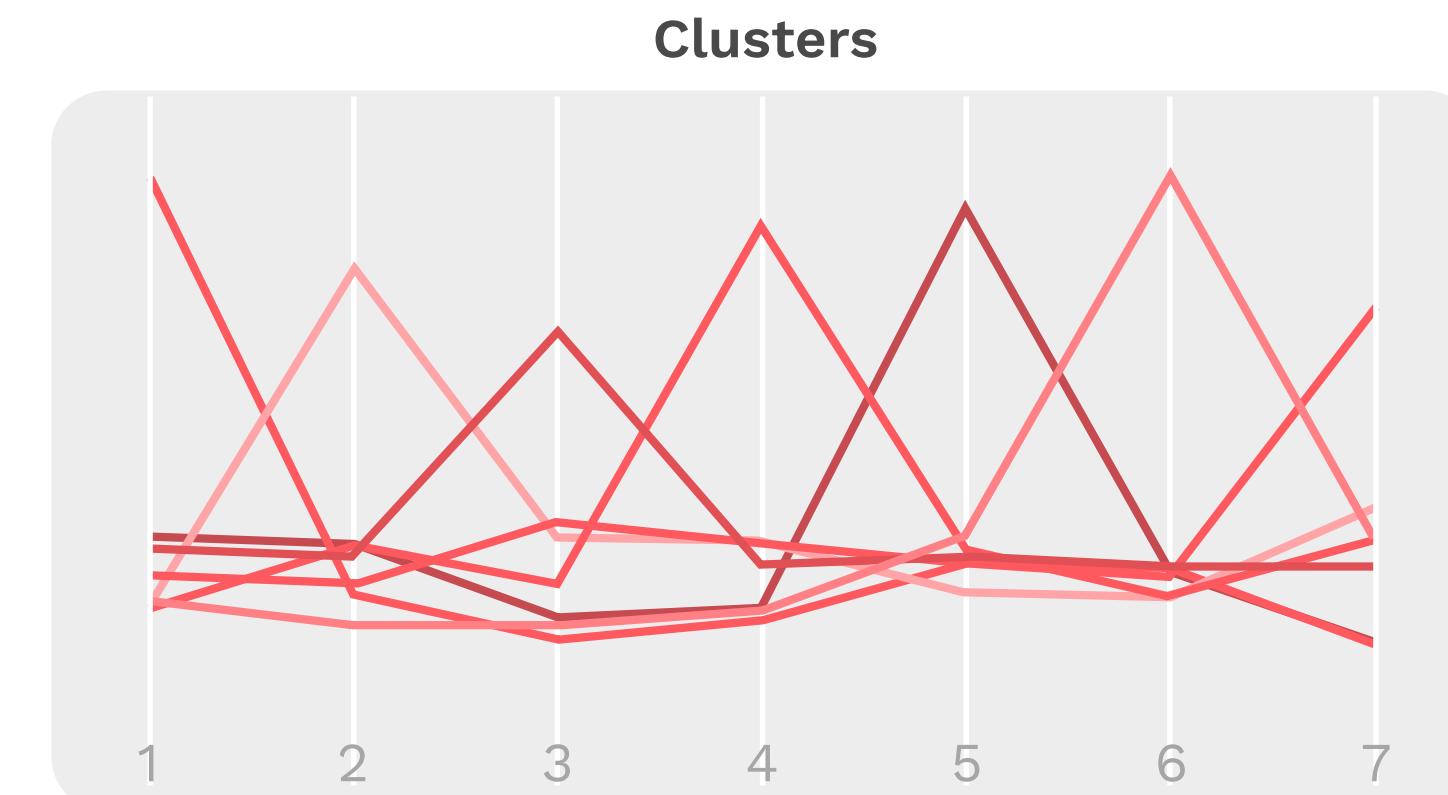
Through LDA, I generated the seven most common topics from hosts' profile bios.



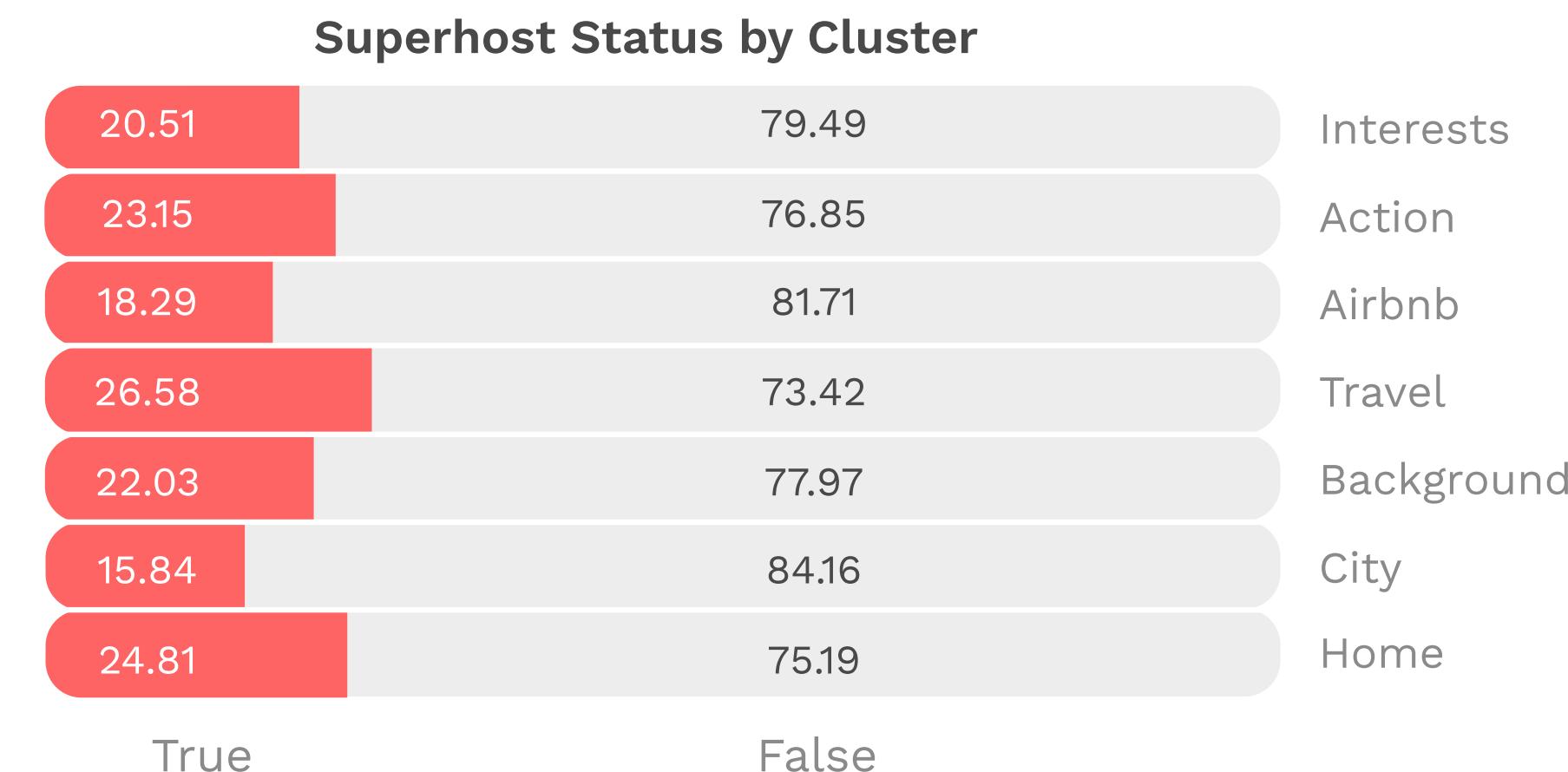
Then, I clustered the data into seven groups corresponding to the topic modeling outputs.

Graphing these clusters by the Superhost metric revealed a statistically significant difference in a host's status **based on the contents of his or her profile**. Among the hosts who wrote about their own travels, 26.6% were Superhosts, which was the greatest proportion out of any cluster. In comparison, only 15.8% of hosts who wrote about New York City were Superhosts.

This difference in outcomes suggests that hosts who use their profiles to share some part of themselves rather than to describe the location that guests are coming to visit are more likely to be trusted individuals.



See Appendix F for cluster means.



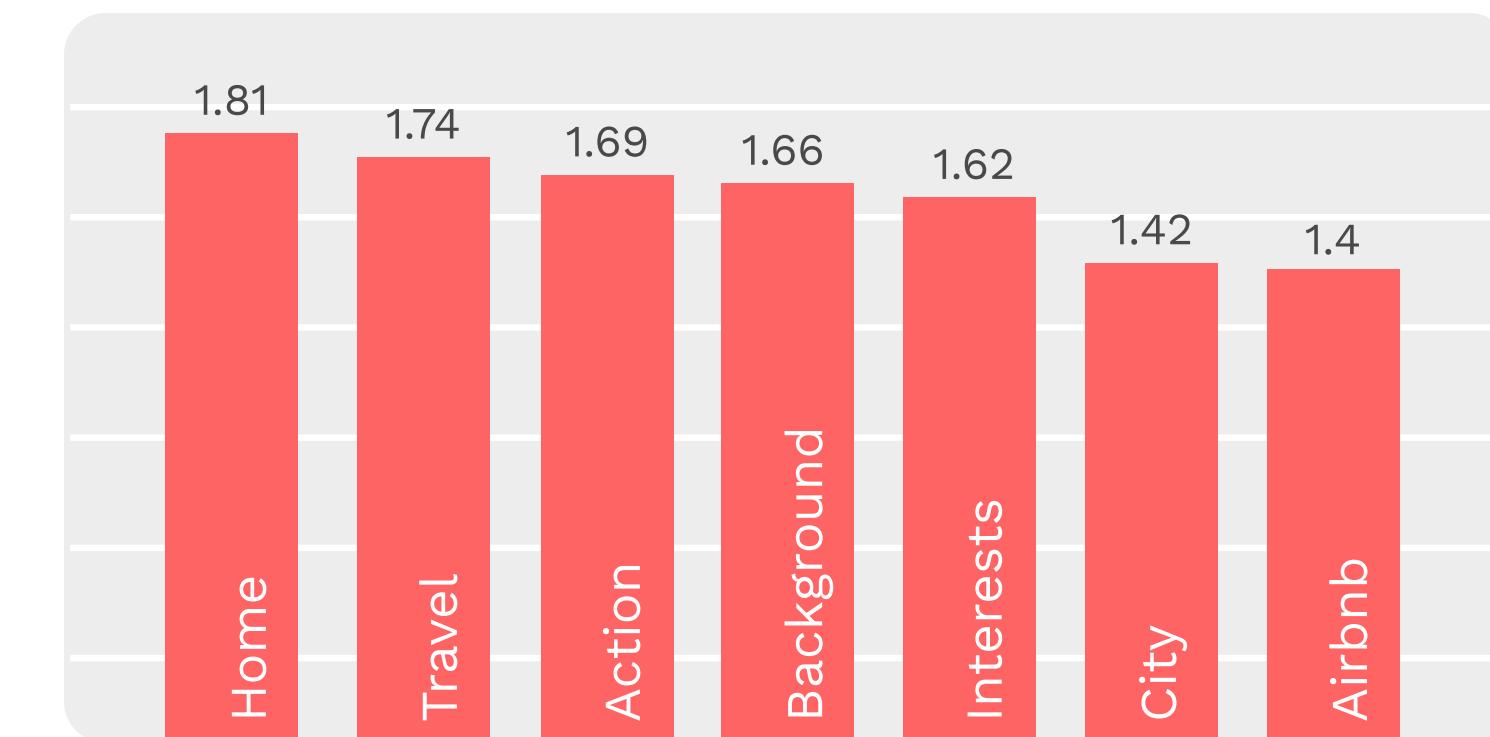


Analyses of clusters against average reviews per month and against total review count further suggested that hosts' performance was linked to their profiles.

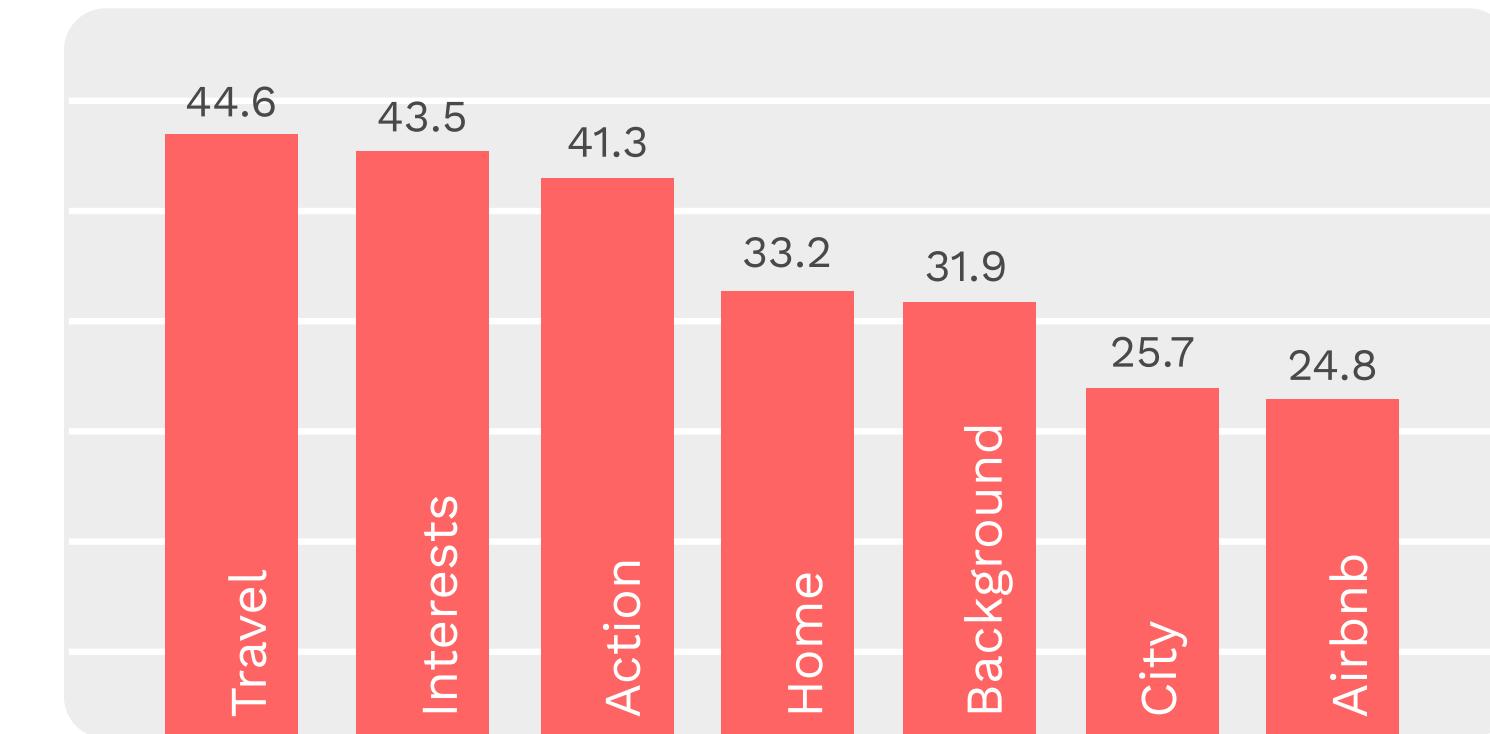
Hosts with bios that belonged to the travel cluster scored 0.3 average monthly reviews more than hosts with bios in the city cluster. Moreover hosts who shared their personal interests had a statistically significant higher review count than hosts who kept to the standard topic of Airbnb. On average, the former group had 18.7 more reviews than the latter, suggesting **greater demand for their listings on the market**.

Furthermore, hosts who wrote about any of the seven topics performed better than hosts who left their profiles empty.

Average Reviews per Month



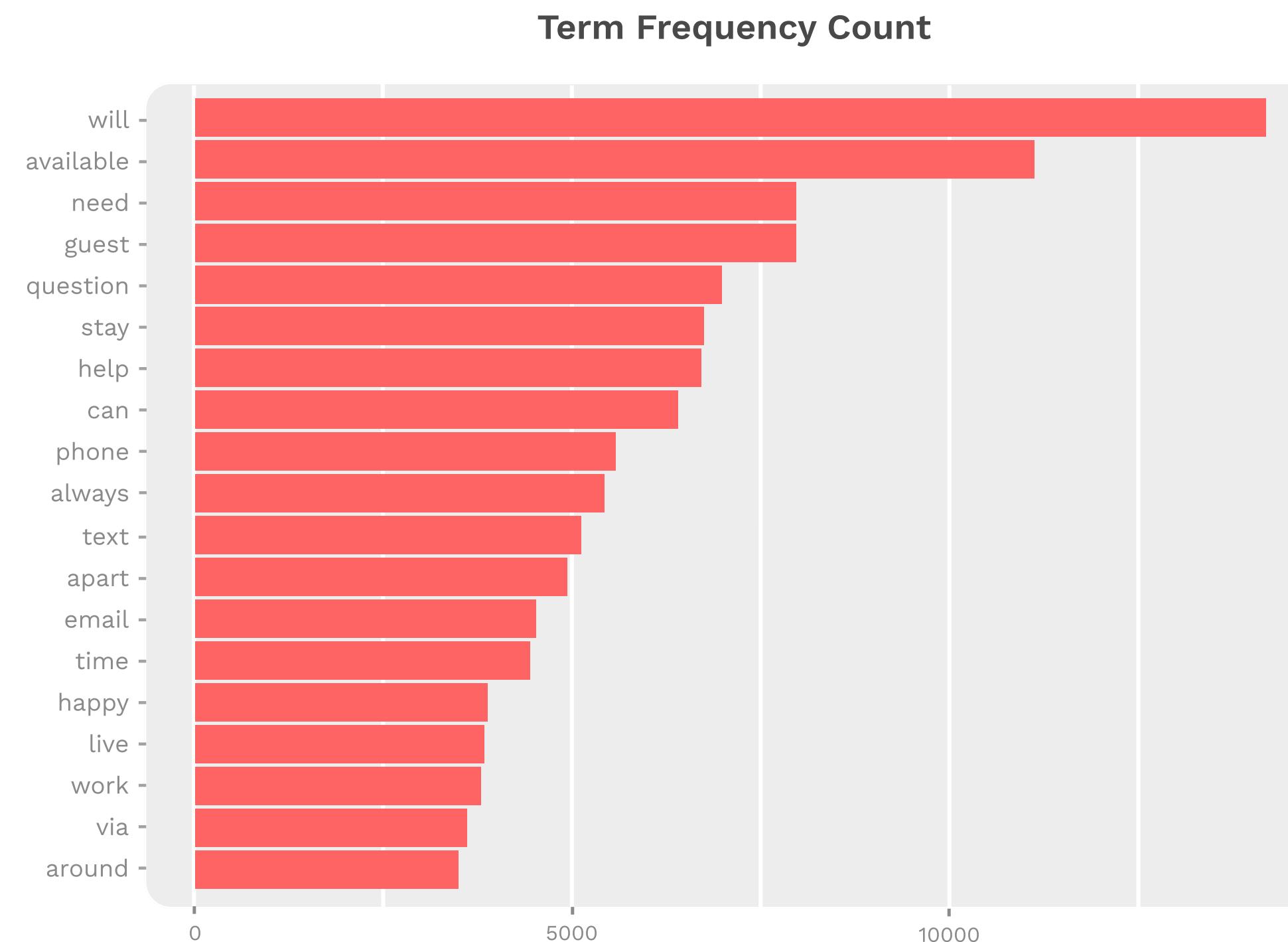
Average Total Reviews



Guest interactions

The text in this section was evaluated using a term frequency count. The two most popular terms, “will” and “available,” were used by hosts over 10,000 times. Other terms such as “help,” “need,” and “question,” each surfaced more than 5,000 times.

Taking a random sample of the data and looking at the contexts in which these words were used, they often revealed hosts’ desire to help their guests during the stay. For example, one host wrote, “I will always be available in the house to answer any questions you might have.”



Subsequent analyses revealed a statistically significant difference between the binary version of the interactions variable and the review scores. Hosts who had completed the interactions section of the listing page (where level = 1) consistently achieved higher mean scores across these four categories.

A Qualification. I should note that the review scores were skewed toward the positive end, so the differences in means were not large enough to be of practical significance.

However, hosts who describe possible guest interactions may still benefit from an increase in trust that can not be captured by the data (for example, improved in-person exchanges with their guests).

Interaction by Accuracy Score

Level	Mean
0	9.50
1	9.65
t-Test	
Difference	0.14
Prob > t	<.0001

Interaction by Checkin Score

Level	Mean
0	9.67
1	9.79
t-Test	
Difference	0.12
Prob > t	<.0001

Interaction by Communication Score

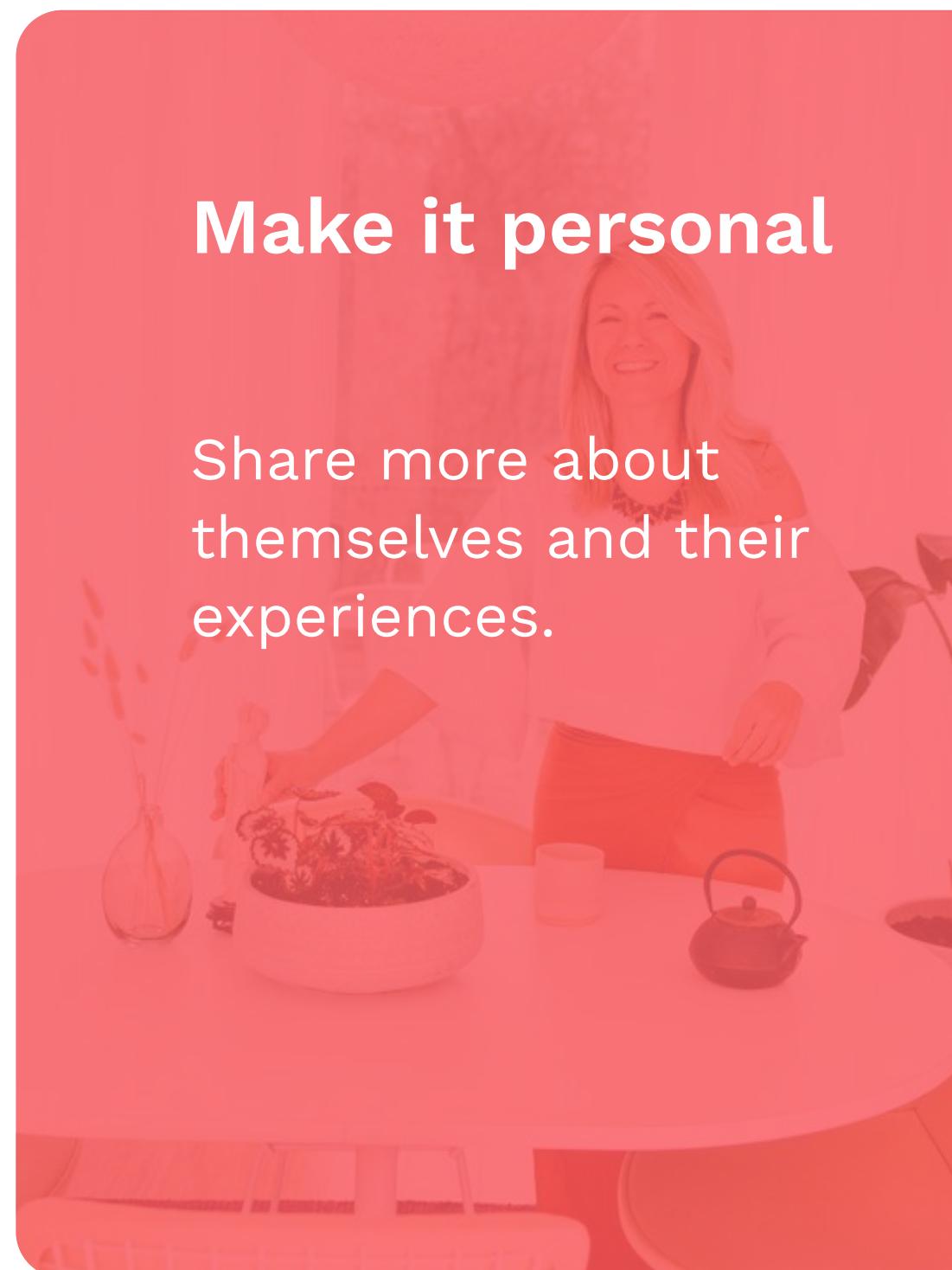
Level	Mean
0	9.71
1	9.82
t-Test	
Difference	0.11
Prob > t	<.0001

Interaction by Overall Rating

Level	Mean
0	93.03
1	94.33
t-Test	
Difference	1.30
Prob > t	<.0001

Recommendations

Given our findings, Airbnb hosts should follow three rules to build relationships of trust with their current and potential guests:



Make it personal

Share more about themselves and their experiences.

Hosts should let their personalities shine on their profile page. The topic modeling results showed that hosts who keep strictly to business tend to receive fewer reviews each month and over the entire year.

The typical Airbnb guest seeks an intimate travel experience. They want to feel closer to their environment and the people they're living with. Hosts who **embrace their unique identities as locals and fellow travelers** can appeal to this desire for connection.

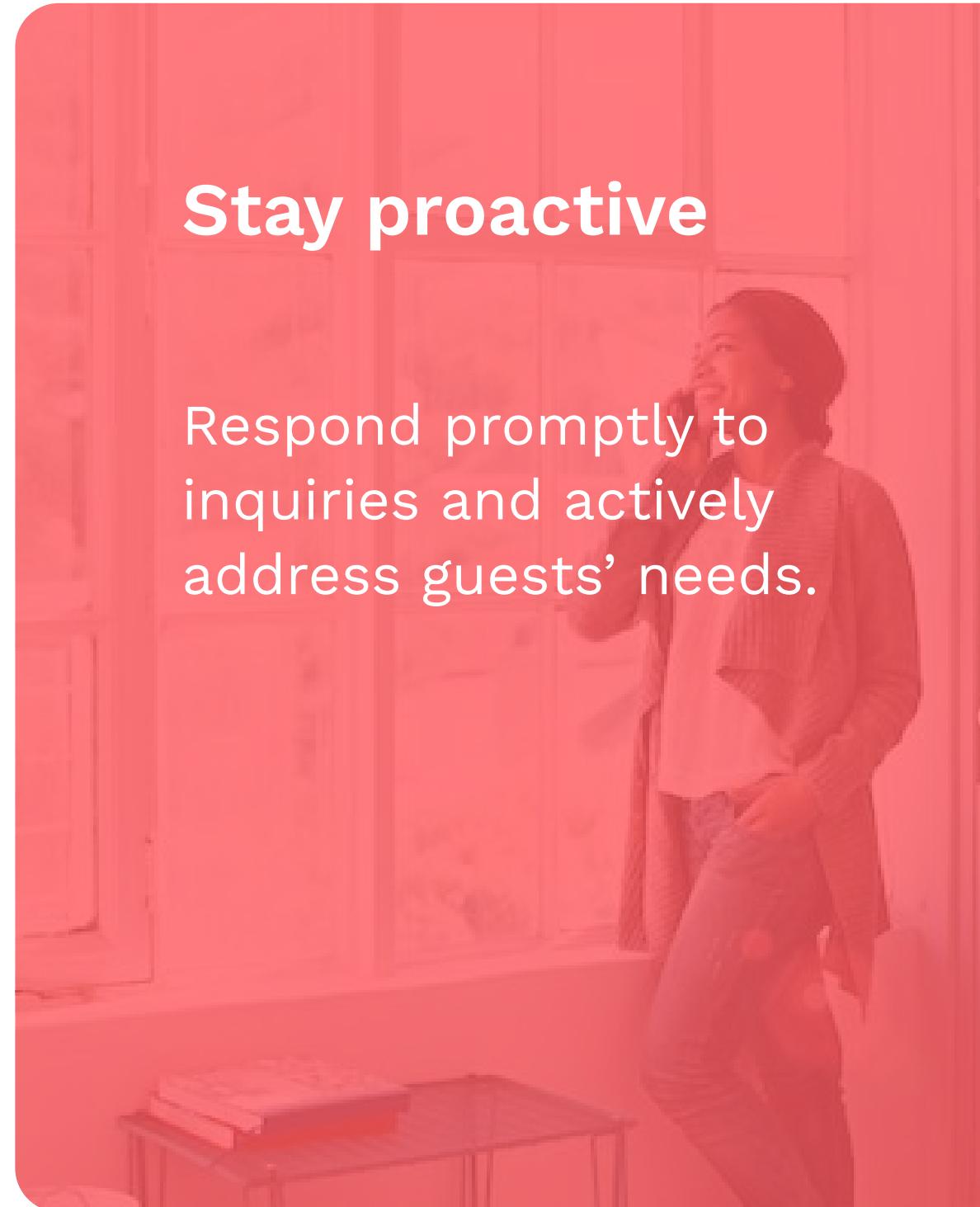


Recommendations

Given our findings, Airbnb hosts should follow three rules to build relationships of trust with their current and potential guests:

Stay proactive

Respond promptly to inquiries and actively address guests' needs.



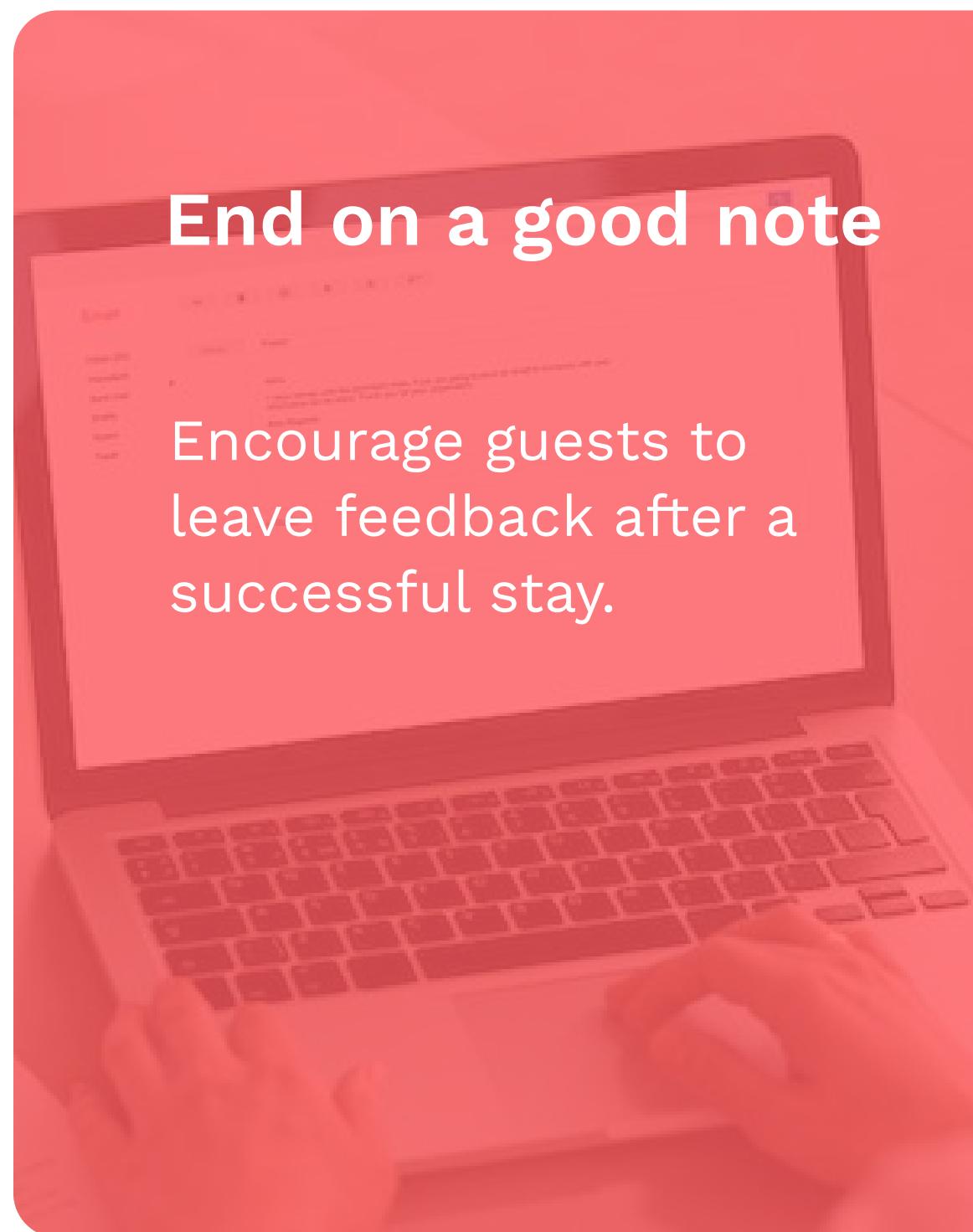
The machine learning predictions revealed that hosts' response rates and times impact their trustworthiness. Moreover, hosts who fail to communicate their accessibility during the stay receive lower review scores.

Travelers want a host whom they know can be reached during the moments that matter. Hosts can demonstrate their reliability through **consistent engagement** online and by clearly outlining potential interactions on their listings pages.



Recommendations

Given our findings, Airbnb hosts should follow three rules to build relationships of trust with their current and potential guests:



Finally, the model predictions showed that the number and frequency of reviews affect a host's credibility. On average, only 70% of guests will leave feedback.⁶

Given this information, hosts should **follow up with guests** at the end of every stay. Sending a simple thank you message can be a great way for hosts to express appreciation for their time together and prompt the guest to reciprocate with a positive review.



References

- 1 Inside Airbnb**
<http://insideairbnb.com/>
- 2 What is a Superhost?**
<https://www.airbnb.com/help/article/828/what-is-a-superhost>
- 3 Elements of Trust**
<https://hbr.org/2019/02/the-3-elements-of-trust>
- 4 Should hosts reveal their exact location?**
<https://community.withairbnb.com/t5/Hosting/exact-location-revealing/td-p/5892>
- 5 Listing location not accurate**
<https://community.withairbnb.com/t5/Hosting/Location-on-map-not-accurate/td-p/21764>
- 6 Host review rate**
<https://community.withairbnb.com/t5/Help/Review-Rate/td-p/278361>

Acknowledgments

I completed this report with my learnings from Data Driven Decision Making, a course that I took in the Spring of 2021. Professor Xian Liu offered a wonderful introduction to all of the tools that I used, and she has inspired me to dive deeper into the world of machine learning. I would also like to acknowledge two classmates, Karen Carillo and Paola Arguello, with whom I worked on the final project for the course. The topic that we explored, of building trust between Airbnb hosts and guests, is the same one that I used in this report; however, I have chosen to take a different approach to the problem, running my models and analyses with different metrics and predictors.

To clarify, all content in this report, including the topic of exploration, is my own.