# Executive summary

## Business Problem

The market for short-term rental properties is complex and identifying key factors that drive potential renters' preferences and significantly impact ratings is essential for maximizing investment returns. We are looking at Austin, TX as we aim to uncover the pivotal features that distinguish high-performing rental properties from others. Through this process, we aim to assist our real estate investor in making data-driven decisions when considering a new investment property to list on Airbnb.

## Overview of the Airbnb data

We were able to get a data set containing information of almost 6000 Airbnb rental units in Austin, TX. For each property we had around 9 descriptive columns containing guest summaries on the rental property or additional information about the living space. Also, there were columns that would include the type of property, the bed type, to the amenities available at the property. Other columns were numeric and contained information on the number of beds, bathrooms, square footage, and price. The latter half were guest reviews rating their experience from cleanliness to host communication, and the overall experience at staying in that rental property.

With large amounts of data comes data cleaning to allow for the development of our models. We decided not to remove any numeric columns or rows that had more than 10% null values from the dataset to prevent losing valuable information. The best method was to impute the mean for all numerical columns. However, at our attempt to enter “unknown” for categorical columns, we unfortunately were led to a more problematic path, so we decided to drop those instead.

## Overview of Applied Techniques

Since we are interested in the rental market in Austin, linear regression helps us figure out how a specific factor, like size of a house or availability, affects its price. Essentially, it uses past data to estimate how much a house might cost based on a feature. So, utilizing this powerful technique helped identify key factors that contribute to nightly rental price based on the dataset features.

Now, for questions that have only two values of “yes” or “no,” like will this Airbnb be booked? Logistic regression is the best technique to apply. It will help make predictions by looking at previous rental properties that did get booked and looking at their features to estimate the probability that a new rental property will be booked in the future. We used a logistic regression with an L1 regularization to help make a more accurate model. In doing so, we could investigate what key factors drive a rental property to be booked and look at the reviews to see how important they are.

Lastly, we used K means cluster to segment the Airbnb properties in the Austin area into different groups. By identifying these groups, we could investigate which segment commands the highest nightly rates, which receives the best ratings, and overall, most units booked.

## Findings

For nightly rental price we found that the host response time is important as it could decrease rental nightly price. Moreover, shared rooms or private rooms decrease prices since it is not an entire house/apartment. The more obvious factors that increase prices are the number of bedrooms, bathrooms, and the square footage. In addition, having a cancellation policy of super strict where a guest can cancel 30 days prior to check-in for 50% of refund also increases nightly rental price. Less obvious factors were the neighborhoods where these rental properties are located. East Downtown, South First, Old West Austin, Zilker, Travis Heights, Downtown, Boulding Creek, South Congress and Holly, all positively impact nightly rental prices. Properties located near popular tourist attractions, entertainment venues, or major landmarks tend to command higher prices due to their convenience and accessibility.

Our logistic model provided this chart indicating all the important features of keeping a unit booked in Austin [Figure 1]. The leading factors are the availability of the property over the next two months and the quality of reviews. The number of days the property remains available within a specified timeframe directly impacts its attractiveness to potential guests. An increased availability provides more opportunities for bookings, catering to varying schedules and preferences.

Moreover, positive reviews significantly contribute to a property's credibility and desirability among prospective renters. Favorable feedback serves as a powerful marketing tool, instilling trust and confidence in potential guests, thereby increasing the likelihood of consistent bookings. Combining availability and positive reviews enhances the unit's marketability and strengthens its position in Austin's competitive rental landscape, contributing significantly to its sustained occupancy and success.

Lastly our K means clustering technique created four segmented groups.

1. **Group 1:** Typically, hosts in this group have an average response rate with a moderate presence of super hosts. This is a group that might be new to the short-term rental market since they have a lower number of listings. These listings accommodate an average of 3 to 4 guests, with around 1-2 bedrooms and bathrooms. They also have a high number of reviews and high review scores across various categories but receive the lowest number of bookings, which might be because of their low visibility in the market since they have a low number of listings.
2. **Group 2:** Listings in this cluster have slightly larger accommodations than previous groups and slightly more spacious properties. Hosts maintain a high response rate and have an above-average number of listings. They also possess good verification and profile completion. Despite slightly lower ratings these properties receive a significant number of bookings, indicating a balance between moderate ratings and higher demand, possibly due to other factors like location or pricing strategy.
3. **Group 3:**

Hosts in this group have larger high-end listings which means that they have the highest nightly rates out of all other groups. They also have a relatively higher response rate and a substantial number of listings. This group represents larger, upscale listings accommodating around 9 guests on average with more bedrooms and bathrooms. They receive a low number of reviews, but maintain exceptional ratings across various review categories, suggesting high guest satisfaction. Despite the outstanding ratings, these listings have a moderate booking rate, possibly due to their larger size and higher pricing, limiting the potential audience.

1. **Group 4:**

These listings have a similar accommodation like in previous groups. Hosts in this cluster have a moderate number of listings but tend to have lower response rates compared to other clusters. They receive relatively lower ratings across different review categories. Despite moderate capacity and reasonable booking rates, the lower ratings might impact on the overall demand for these properties, affecting their booking frequency compared to clusters with higher ratings.

## Conclusion and Recommendations

The findings highlighted a variety of factors impacting rental prices and booking probabilities. Host response time, property type, amenities, and neighborhood location emerged as determinants of nightly rental prices. Availability, quality of reviews, and positive guest feedback significantly influenced the likelihood of a unit being booked, emphasizing the importance of maintaining high standards and ensuring positive guest experiences.

The segmentation analysis revealed four distinct groups within the Austin rental market, each with its unique characteristics and performance metrics. These groups shed light on different strategies and market positions within the short-term rental landscape showcasing the successes and the lost opportunities per group.

**Recommendations:**

*Optimize Quality of Reviews:* We encourage high response rates as a host to maintain exceptional guest experiences and increase the probability of consistent bookings. Implementing measures to improve communication and ensuring a high standard of cleanliness, amenities, and overall guest satisfaction are vital. That said, ask guests to leave a review after every booking to increase credibility and to ensure that your visibility within the market is high.

*Location and Neighborhood Insights*: Emphasize proximity to attractions, landmarks, and entertainment venues in property descriptions to attract more guests.

*Continuous Adaptation*: Regularly review and analyze market trends, competitor behavior, and guest preferences to adapt pricing strategies accordingly. Embrace flexibility and agility in responding to market shifts and changes in guest demands.

Leveraging the insights gained from data analysis and implementing strategic decisions based on these findings will empower informed decision to enhance property performance and maximize returns in the competitive Austin short-term rental market.

## Appendix

**Figure 1:**

A screenshot of a graph

Description automatically generated