

Airbnb Prices in Chicago

Presented by :
Becca, Caitlin, Genni, Sam & Sandhya





About Airbnb

Airbnb is an online marketplace that connects people looking to rent out their homes or apartments with travelers seeking accommodations.

Pricing

Prices on Airbnb are typically set by the hosts who list their properties on the platform. Hosts have the flexibility to set their own prices, which can vary depending on factors such as the location, size, and amenities of the property, as well as seasonal demand and local events. Airbnb also provides hosts with pricing tools and recommendations based on market data to help them set competitive prices. Additionally, guests may see additional fees such as cleaning fees, service fees, and taxes that are added to the base price set by the host.



The Problem

Airbnb listings have a variety of features such as location, bedrooms, number of available beds, bathrooms and reviews that make determining the value of a listing difficult for hosts and consumers.

By analyzing the data, a model can be created to allow for optimal pricing to allow hosts to maximize profits and allow consumers to determine if they are getting a fair price.

Source Data

Data was sourced from the website "Inside Airbnb," "a mission driven project that provides data and advocacy about Airbnb's impact on residential communities... [working] towards a vision where data and information empower communities to understand, decide and control the role of renting residential homes to tourists."

“Chicago.” Insideairbnb.com, insideairbnb.com/chicago.

Exploring The Data

Machine Learning

- A **Random Forest Regression Model** was utilized.
 - This type of model uses an ensemble of decision trees to make predictions about continuous numerical values
- We selected this type of model due to the following advantages:
 - Ability to handle a large number of features and observations
 - Resistance to overfitting
 - Can provide information on feature importance, such as which Airbnb features have the most significant impact on pricing

Model Development

- Predictor Variables: Accommodates, Bedrooms, Price, Minimum Nights, Number of Reviews, Review Scores Rating, Review Scores Accuracy, Review Scores Cleanliness, Review Scores Check-in, Review Scores Communication, Review Scores Location, Review Scores Value, Reviews Per Month, Bathrooms, Property Type
- Binned bathrooms and property type to streamline the data as there were multiple types and many small values
- Price outliers were determined and dropped to decrease skewing

Hypertuning

- Try different values for three settings: 'max_depth', 'min_samples_split', and 'n_estimators'.
- The code will try different values for these settings and train the model multiple times.
- Use cross-validation to make sure the model is good at making predictions on new data.
- Find the best combination of settings found for training the final model.

Original Model

MAE: 64.265512

RMSE: 153.770356

R2 0.384639

Hypertuning

MAE: 63.192465

RMSE: 153.703064

R2 0.385178

Hypertuning Attempt 2

MAE: 63.192465

RMSE: 153.703064

R2 0.385178

Model Results

- For regression – what were your test set performance metrics (e.g., MSE, RMSE)?
- Can you plot the “actual” dependent variable vs. the model’s “predicted” dependent variable?
- From a business perspective, what does your model’s prediction error mean?

Conclusions

- It is difficult to accurately predict Airbnb pricing.
- Despite limitations, clients can still utilize a model like this and especially with further time invested in it, it could become even more accurate and useful in its predictions.
- Additionally, for this purpose, an overprediction is more valuable than an underprediction.

Improvements:

- Improved data collection/organization, or extensive reformatting of long text entries for categorical values

Sources

(2023, March 19). How is Airbnb really being used in and affecting the neighbourhoods of your city? Inside Airbnb. Retrieved April 5, 2023, from <http://insideairbnb.com/get-the-data/>



Questions?
