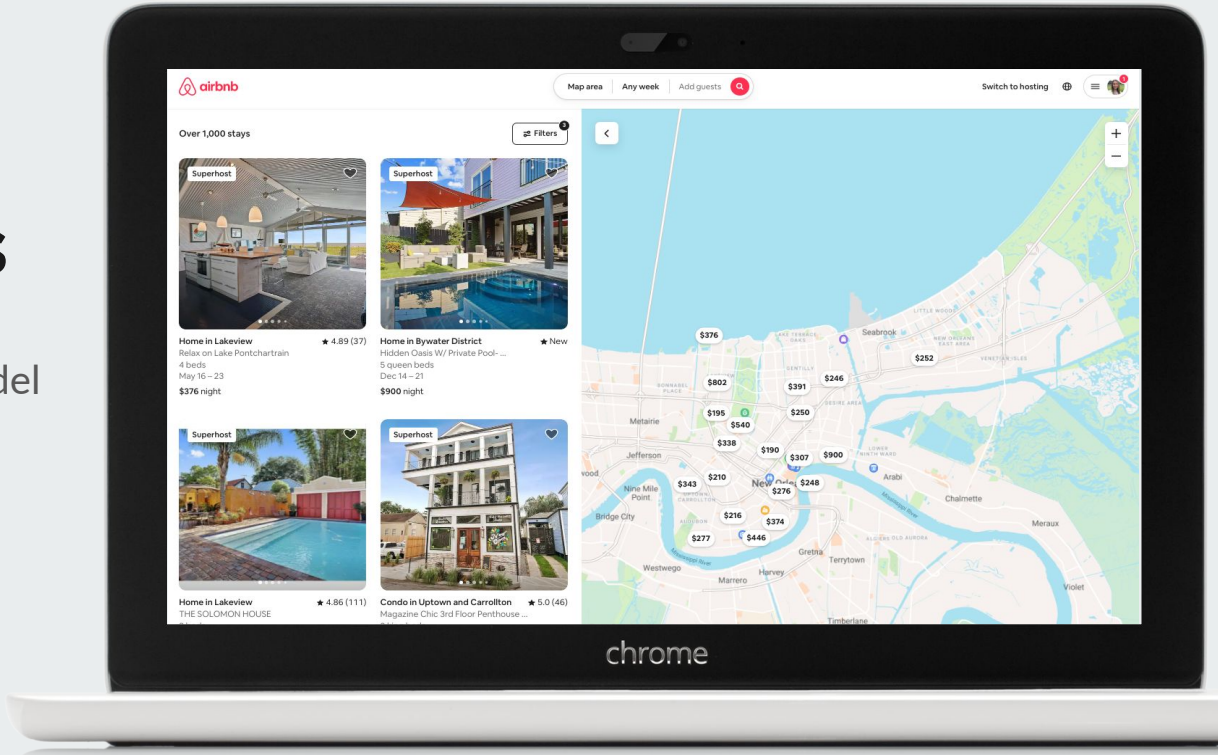
New Orleans Airbnb Analysis

Exploratory Data Analysis and Machine Learning Predictive Model

Author: Rachel Fein



Summary

The goal of this analysis was to gain information and insight on the New Orleans short term rental market. Exploratory data analysis was completed and machine learning was used to develop a predictive model for nightly price.

The final model was a Random Forest model. In summary this analysis showed:

- The final model had an RMSE of $2.02 on the test set.
- The final model had a MAE of 1.45 on the test set.
- The Exploratory Data Analysis gives a general understanding of the New Orleans' Short term rental market by showing relationships between the target and features.

art by Diane Millsap

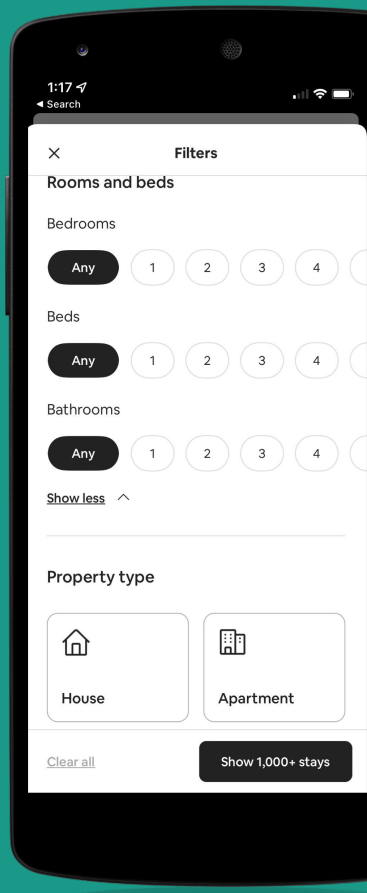
Outline

art by Romare Bearden

The Problem

This analysis was completed for Sarah Alter, a private investor, who is interested in expanding her vacation rental portfolio to New Orleans, LA. Alter is in search of gaining a better understanding of the New Orleans Airbnb market and is in need of information to gain perspective and aid her in finding her next investment property.

Not every investment is a good one. Without a good understanding of the market and nightly rates, the chances of buying a bad investment property can rise.

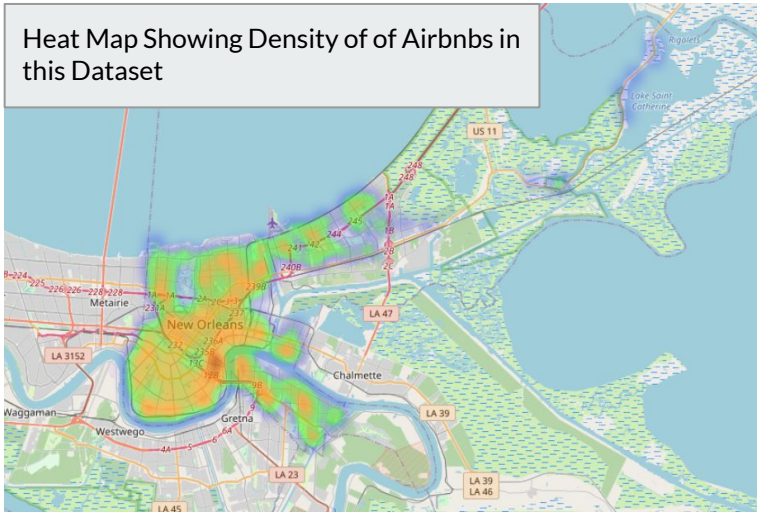


Data

The dataset in this analysis comes from Kaggle, however it is noted the dataset was originally pulled from Inside Airbnb. The dataset was compiled on November 7, 2021 and all time oriented variables in this dataset are based around that date.

Why this dataset is good for solving our problem?

- Contains 6,028 Airbnb Listings, all of which are in New Orleans
- Has features that relate to property features such as number of bedrooms and bathrooms
- Contains features that give insight on each Airbnb's success, such as nightly price and availability
- Limited missing data overall and no missing data for target feature price



Methods

Data Cleaning Methods

- Replace or remove missing data
- Replace or remove illogical data

Preprocessing Methods

- Create Test and Training datasets
- Remove extreme outliers
- Encode categorical data
- Data transformations
- Remove highly correlated features

Modeling Methods

- Create simple baseline model that future models were compared to
- Main performance metric was Root Mean Square Error
- Hyperparameter tuning
- Feature Selection

Results

The goal of this analysis was to gain information about the New Orleans short term rental market using exploratory data analysis and develop a machine learning price predictive model that Alter could use as an aid while searching for an investment property.

Exploratory Data Analysis

The exploratory Data Analysis (EDA) proved to find general relationships among the data to get a grasp on the New Orleans short term rental market. Some highlights from EDA are shown in the following slides.

Predictive Model

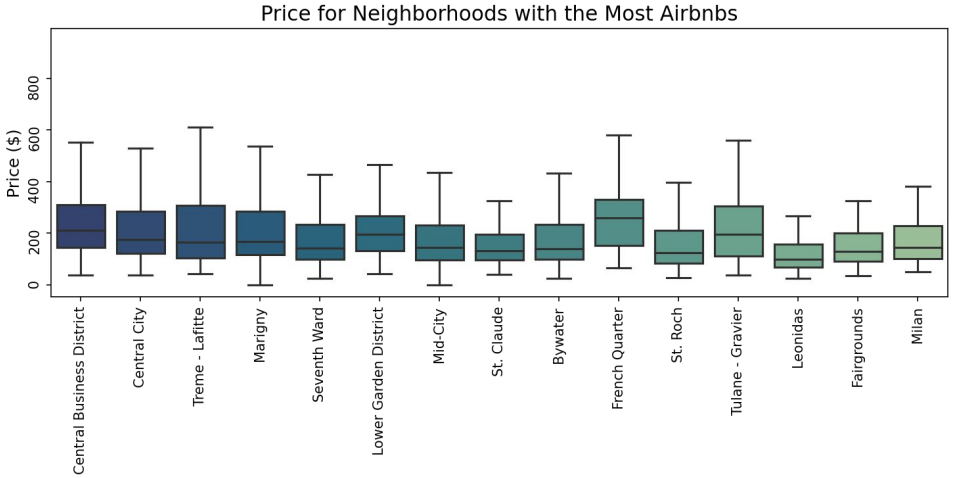
The final model in this analysis was the model with the lowest RMSE. The final model was a Random Forest model that went through hyperparameter and feature selection.

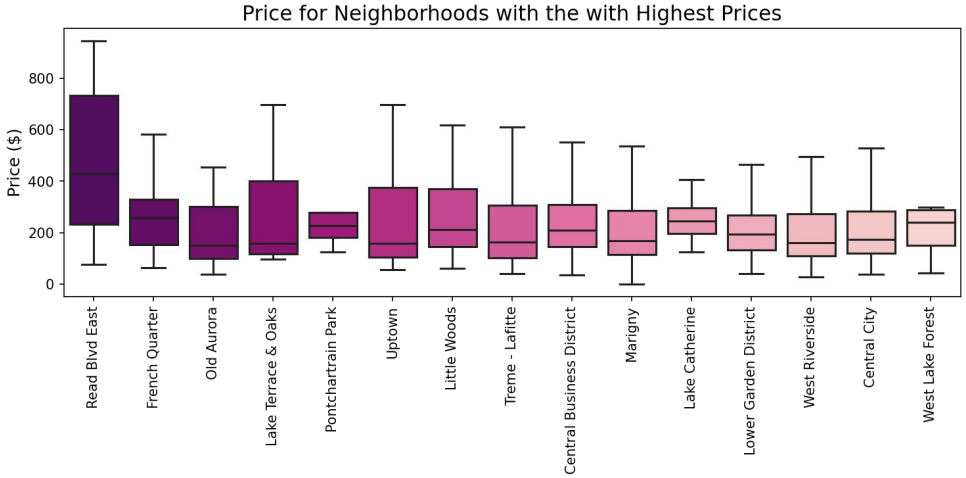
The machine learning analysis showed that there is enough good data to develop a predictive model that correctly predicts the nightly rate within $2.02. For perspective the baseline model had a RMSE of $3.18. The final model had an MAE of 1.45 on the test set.

Data Exploration

Knowing where the most densely populated areas are for airbnbs give a lead to the best neighborhoods to purchase an Airbnb.

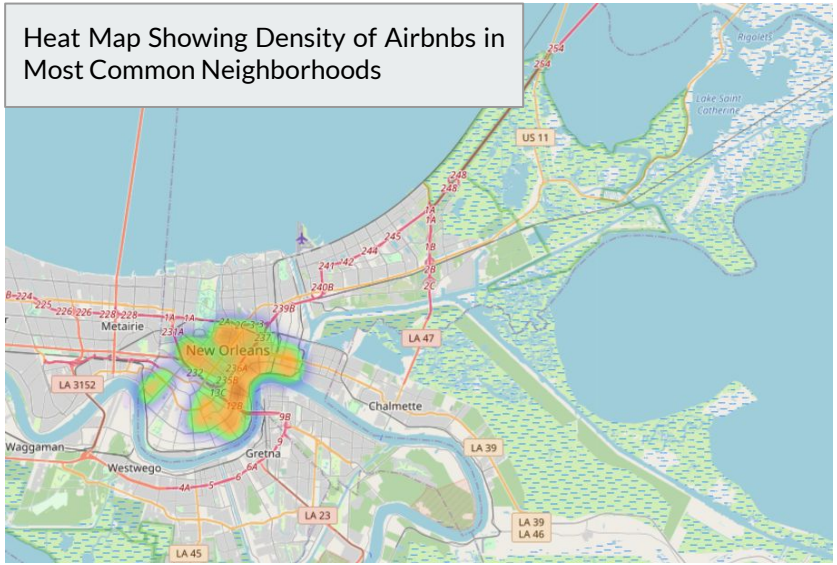
Knowing where the highest priced Airbnbs is also important as higher prices can mean higher profits.



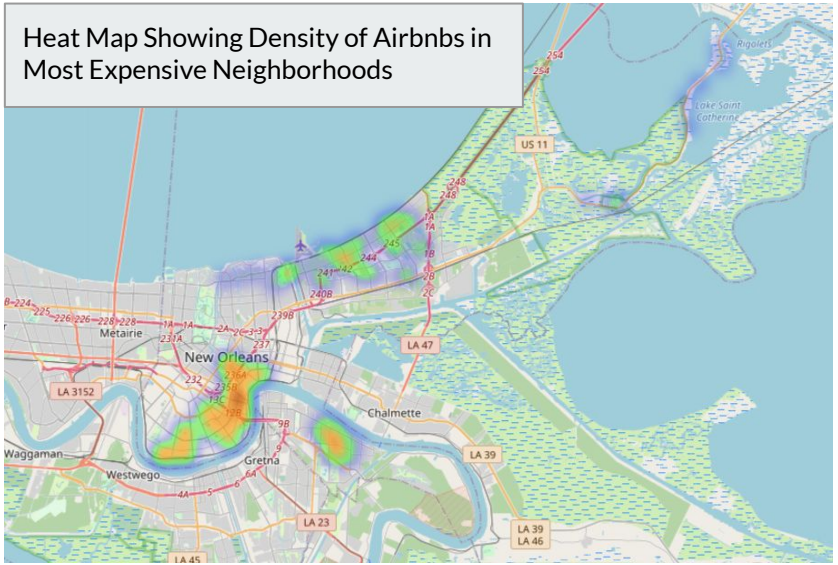


Data Exploration

Heat Map Showing Density of Airbnbs in Most Common Neighborhoods

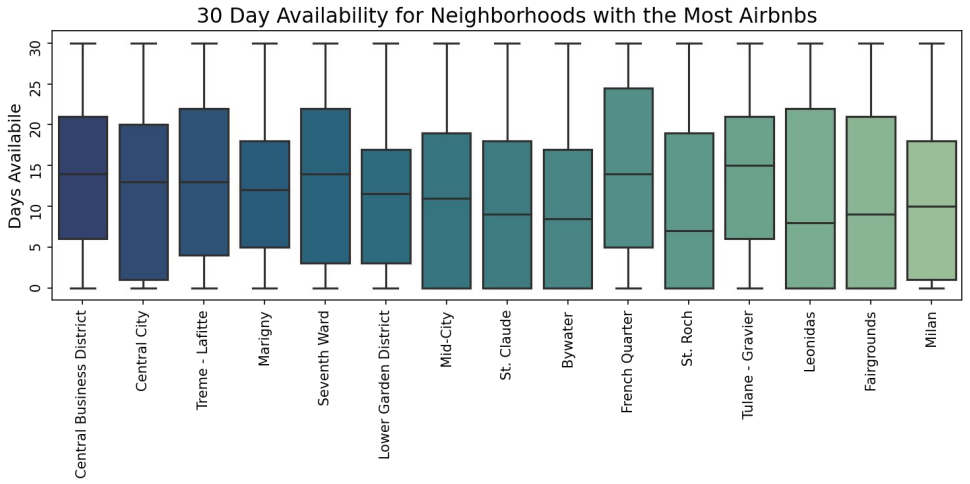


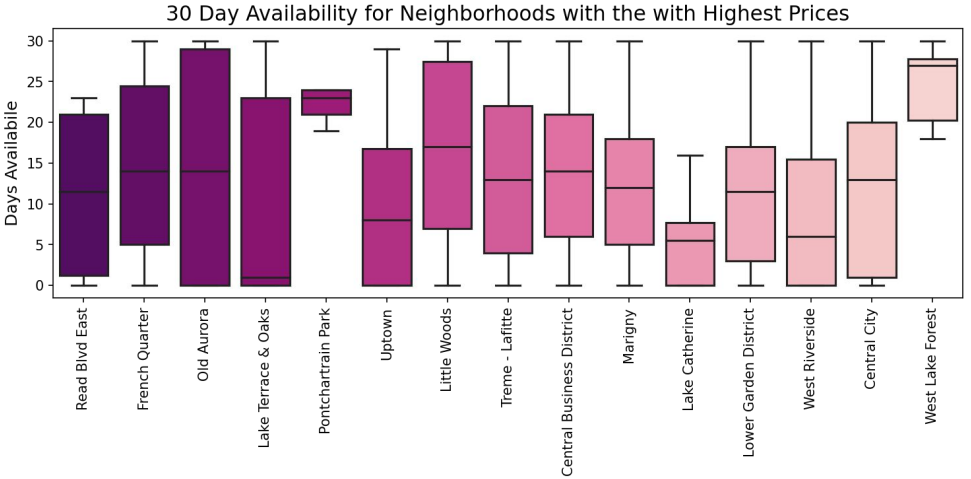
Heat Map Showing Density of Airbnbs in Most Expensive Neighborhoods



Data Exploration

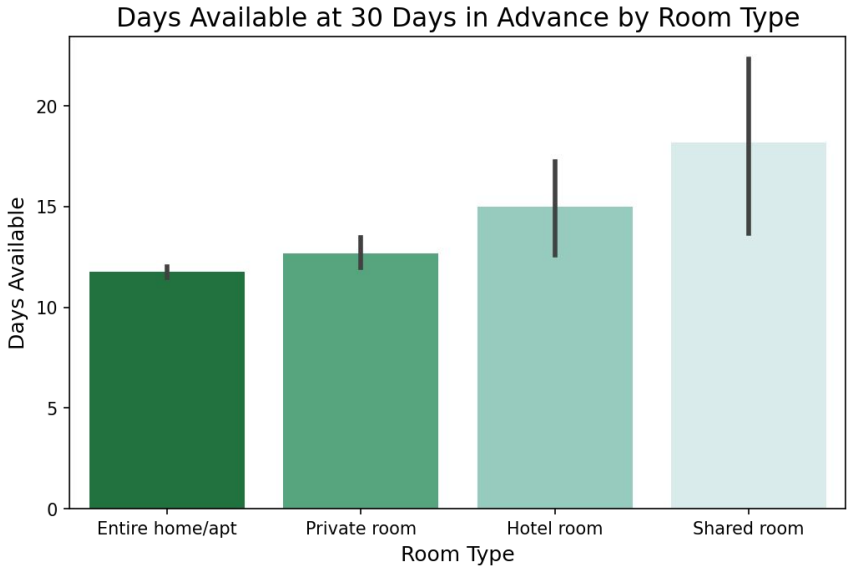
Having little availability at 30 days out is a good sign that the Neighbourhood is highly sought after by tourists and good business for the Airbnb owners.

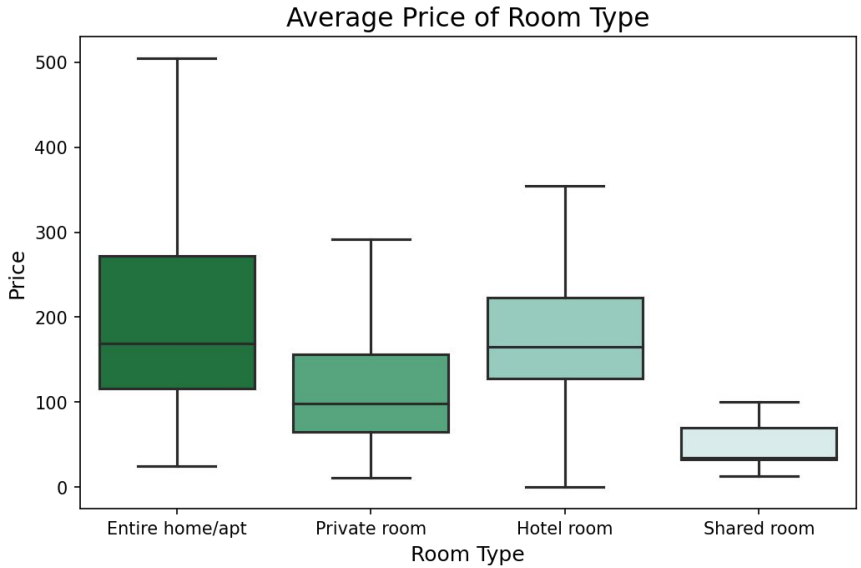




Data Exploration

To get a better understanding of what type of property is more valued by travelers we see which property type has the lowest availability and highest average price.





Conclusion

The final model was a Random Forest model. In summary this analysis showed:

- The final model had an RMSE of $2.02 on the test set.
- The final model had a MAE of 1.45 on the test set.
- The Exploratory Data Analysis shows to be informative in getting an general understanding of the New Orleans' Short term rental market to aid Alter in her property search. Such as which neighborhoods have the highest priced airbnbs and the lowest availability.

Next Steps:

Adding more continuous & relevant variables.

Take into consideration seasonality.

Use what was gathered from EDA to further investigate specific neighborhoods.

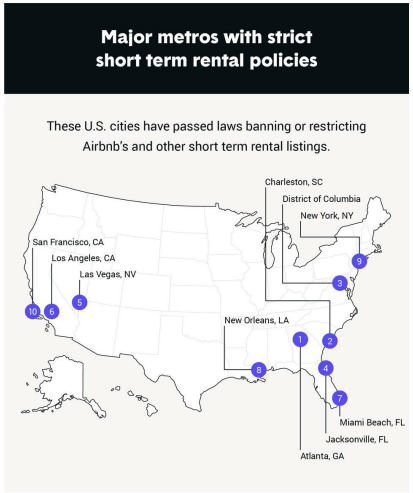


Image from Zebra.com

Thank you!

Email: rachellfein@gmail.com
Github: @rachelfein
LinkedIn: linkedin.com/in/rachellfein/

References

D.K. Shifflet & Associates: New Orleans Tourism Report

Airbnb Stats

Data downloaded from Kaggle: New Orleans Airbnb Listings and Reviews

Data originally sourced from Inside Airbnb