Model Card – AirBnB Price Classification

Model Details

- Model developed in Jupyter Notebook using the provided listings.csv for Assignment 6 that has been cleaned according to the Write Up portion of the assignment submission
- Listings.csv includes information regarding Chicago AirBnBs
- Predictions made using a Linear Regression Model via Python's Sklearn library
- Utilized various AirBnB characteristics to predict the pricing of the units

Intended Use

- The intended use of this prediction model is purely academic for CMSC 2510 — Engineering for Ethics, Privacy, and Fairness in Computer Systems
- The model is for educational purposes to learn about Personally Identifiable Information, data cleaning, and the use of multiple variable linear regression models in Python
- Not suitable for industry use

Factors

- This report is based on the given information from Assignment 6 and the data given
- Data has been cleaned based on the knowledge from the course, thus far, and is subject to change
- Evaluations of the pricing of the AirBnBs is based on factors including (after data cleaning and binning) latitude, longitude, room type, and community/Chicago neighborhood

Metrics

- Evaluation of the linear regression output predictions include the coefficients, mean square error, root mean square error, and coefficient of determination
- Together, these evaluation variables indicate the validity and accuracy of the proposed model

Training Data

- The training data used is from listings.csv and is split into training and test data
- Training portion was use for model training

Evaluation Data

- The testing data used is from listings.csv and is split into training and test data
- Testing portion was used for model training

Ethical Considerations

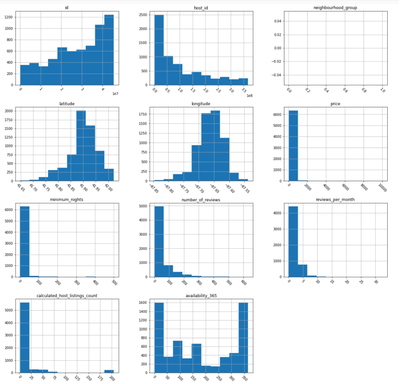
- Within the dataset, information such as the host name and id were given, but this information could be considered Personally Identifiable Information and should be kept private

Caveats and Recommendations

- This dataset and model do not contain information regarding the size (square feet), aesthetic, proximity to tourist attractions, and other information that is pertinent to both the property value and area of the AirBnB that could heavily impact the price of the listing

Quantitative Analysis

Histogram Representations of
the Data Being Used in the DataFrame




```
coefficients:
[ 4.61069277e+15  4.61069277e+15  4.61069277e+15  4.61069277e+15
 -9.85258562e+14 -9.85258562e+14 -9.85258562e+14 -9.85258562e+14
  2.52068371e+15  2.52068371e+15  2.52068371e+15  2.52068371e+15
  7.09269848e+15  7.09269848e+15  7.09269848e+15  7.09269848e+15
  7.09269848e+15  7.09269848e+15  7.09269848e+15  7.09269848e+15
  7.09269848e+15]

max price of the air bnb:
10000

mean square error:
125210.58705882353

root mean square error:
353.85108034146725

coefficient of determination:
0.06474711996980087
```

```
predicted values:
[106.  64. 174. 202.  74. 264.  14. 106. 130. 142. 202. 174. 106. 208.
 174. 142. 106.  64. 174. 106.  74. 174. 208.  74. 174.  84. 174. 142.
 130. 174. 208. 208. 174.  84. 174. 130.  74. 182.  24.  74. 174. 156.
  74. 208. 174.  74. 114.  24. 134. 174.  64. 174.  74. 208.  64. 130.
 130.  68. 130. 174.  64. 186. 174. 208. 130. 174.  42.  14.  24. 208.
 208. 208. 174. 174. 174. 114. 174. 202. 174.  96. 174.  74. 106. 174.
 174. 174.  74. 542. 174. 174. 208.  84. 208. 106.  98. 174. 208. 542.
 -34.  64.  74. 208. 208. 106. 156.  92. 106.  64. 208. 208.  74. 142.
 174. 174. 130. 208. 142. 174. 208. 174. 142. 114. 208. 110. 174. 174.
 208. 208. 156. 202.  74. 174. 208. 106. 174.  74. 174. 208. 130. 130.
 174. 174. 174. 174. 130. 114. 130. 174.  64. 130. 264. 130. 130. 174.
 106. 208. 209. 112. 174. 208. 226. 130. 174. 174. 106. 110.  64. 174.
 130. 106. 208. 174. 174. 174. 130. 174. 142.  64.  74. 174. 208.  64.
 130. 174. 106. 174. 142. 208. 202. 208. 174. 208. 136. 130. 174. 174.
 264. 202. 130. 202.  64. 174. 542. 130.  74. 174. 174.  74. 174. 174.
 174. 106. 174. 208. 134. 130. 106.  74. 542. 174. 174. 208. 174. 106.
  64. 174. 208. 204.  88. 208. 114. 174.  24. 106. 174. 114. 106. 208.
 106. 264. 106. -34. 106. 174. 174. 174. 106.  64. 174. 208.  64. 202.
 174. 174. 202. 174. 208. 130. 106. 174. 208. 142.  84.  64. 208. 208.
 134.  64. 174.  46. 106. 142. 106.  88. 174.  74. 174.  64. 130. 106.
 202. 178. 174. 208. 208. 112.  46. 208.  92. 142. 142. 174. 174. 174.
 130. 174. 130. 202. 142. 208. 174. 174. 106. 174. 174. 130. 174.  74.
 208. 174. 142. 174.  64. 174. 110.  92.  74. 130. 106. 174. 208.  88.
 142. 174. 136. 142. 174. 174. 106. 174. 174. 174.  74. 208.  74. 134.
 106. 136. 264. 208. 174. 106. 542. 142. 186. 174.  92. 142. 106. 174.
 174. 130. 264. 150. 208. 208. 106. 174. 208. 142. 142. 174. 106. 106.
  84.  74. 208. 112. 174. 542. 208. 130. 174. 106.  74. 106. 174. 174.
 142. 130. 208. 174. 174. 174. 114. -34. 106. 182.  62. 208. 208. 202.
 186. 130. 202. 208.  74. 174. 106. 142. 264. 208. 130.  92. 208. 208.
 106. 174. 208. 208.  64. 208.  74. 174. 142. 208. 208. 174. 106. 174. 542.
 208. 208. 174. 130. 174.  24. 142.  92. 142. 208. 186.  54. 174. 202.
 130. 208. 136. 204. 174. 142. 156. 208. 174. 106. 174. 264. 174. 208.
 130. 130. 208.  68. 174. 142.  74.  64. 202. 264. 130. 174. 106. 130.
 208. 174. 208. 142. 174. 174.  74.  74. 130. 150. 114. 202. 208. 130.
 142. 508. 208. 202. 130. 174. 208. 174.  74. 136. 120.  24. 208. 174.
   0.  64. 174. 150. 130. 174. 130. 106. 142. 208.  74. 106. 174.  74.
 142. 106. 106. 202. 174. 208. 174. 142. 174. 174. 208.  64. 156. 142.
 264. 142.  74. 264. 156.  62. 174. 208. 174. 142. 142. 208. 174.
 110. 208. 142. 130.  74. 174.  68. 112.  74. 264. 114.  68. 174. 208.
 130. 174. 202. 142. 106. 130. 174. 130. 106.  74. 174. 174. 174.  74.
  84. 174. 174. 134. 174. 174. 142. 106. 142. 208. 120. 208.  74. 106.
  46. 202. 174.  74. 174. 130. 208. 208. 130. 130. 134. 136.  24. 174.
 106. 208.   0. 136. 208. 174. 202. 208. 130.  24. 208.  92. 174. 106.
 106. 142. 130. 506.  74. 110. 110. 174.  88. 174. 174. 174. 542. 208.
  24.  68. 106. 106. 112. 174. 264. 466. 174. 182. 174.  74.  74. 174.
 174. 174. 112. 174. 174. 202. 208.  74.  64. 106. 174. 130. 208. 174.
  46. 174. 208. 208. 130.  24. 208. 130. 106. 208. 142. 208. 130. 106.
 264. 174.  64.  64. 106.  74. 208.  74. 174. 264. 142. 106. 174. 106.
 174. 150. 208. 174. 106. 142. 110.  46. 174. 202.  74.  64. 130. 174.
 208. 174. 142. 208. 142.  64. 142. 174.  92. 174. 208. 142.  92.  64.
 160. 174. 106. 130. 174. 130. 174.  84. 174. 208. 198.  64. 174.  88.
 142. 130. 174. 208. 208. 174.  74. 106.  84. 106. 174. 112. 134. 174.
 174. 202. 130. 174. 174. 208.  84. 142.  64. 208. 174. 130. 208. 208.
  24.  92. 156. 142. 174. 142.  74.  68. 174. 174.  74. 142. 134. 208.
 174. 208. 130. 208.  74. 130. 130. 174. 208. 508.  24. 136. 174. 264.
 106. 136. 130.  74. 208. 142.  74. 142. 208.  74. 174. 202.  64. 142.
 174. 142. 209. 174. 130. 208. 174.  94. 264. 208.  64. 174.  64. 264.
 106. 208. 264. 174. 142. 130.  74. 134. 174. 174.  64. 204. 174. 174.
 174. 174.  64. 208. 202. 174.  64. 174. 110. 142. 208. 174. 182. 106.
 130. 136. 174. 174. 208. 542. 202. 174.  64. 208. 130. 174. 112. 202.
 174. 130.  42. 202.  74. 208. 130. 150. 106. -34.]
```