

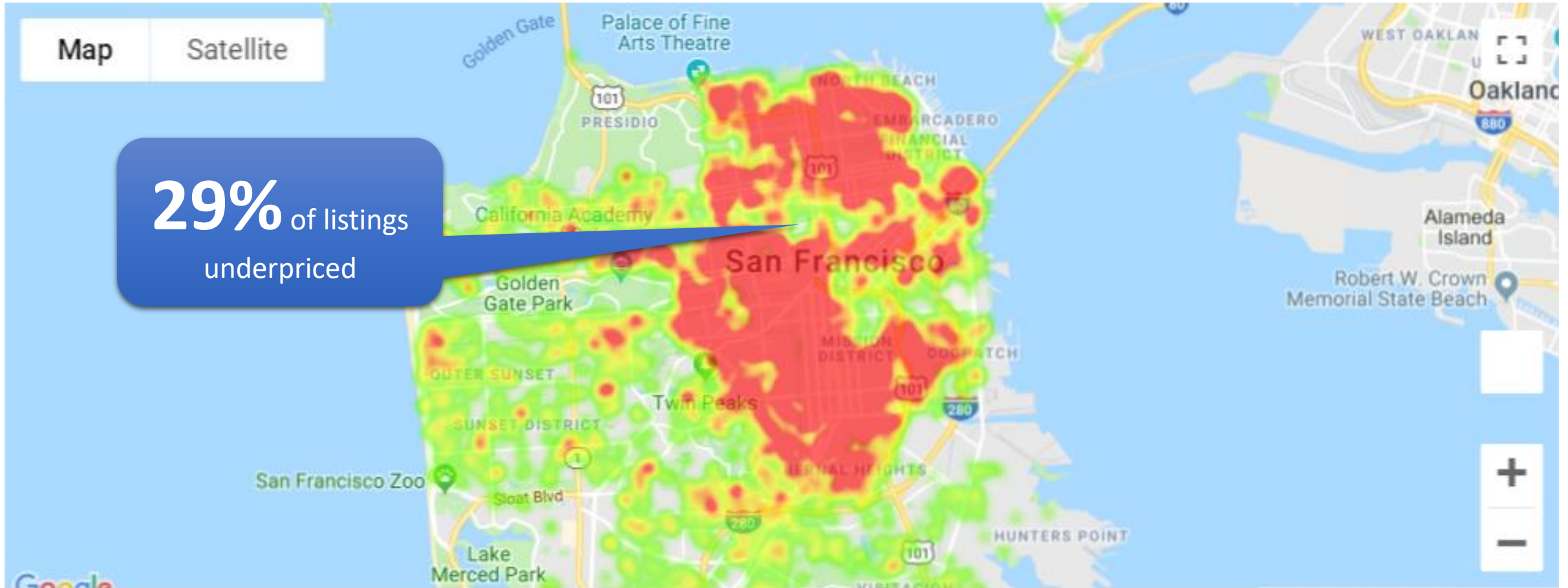


Hosts are unaided by Airbnb to price their properties, while similar properties in an area increase competition for hosts to set a competitive price to attract guests.

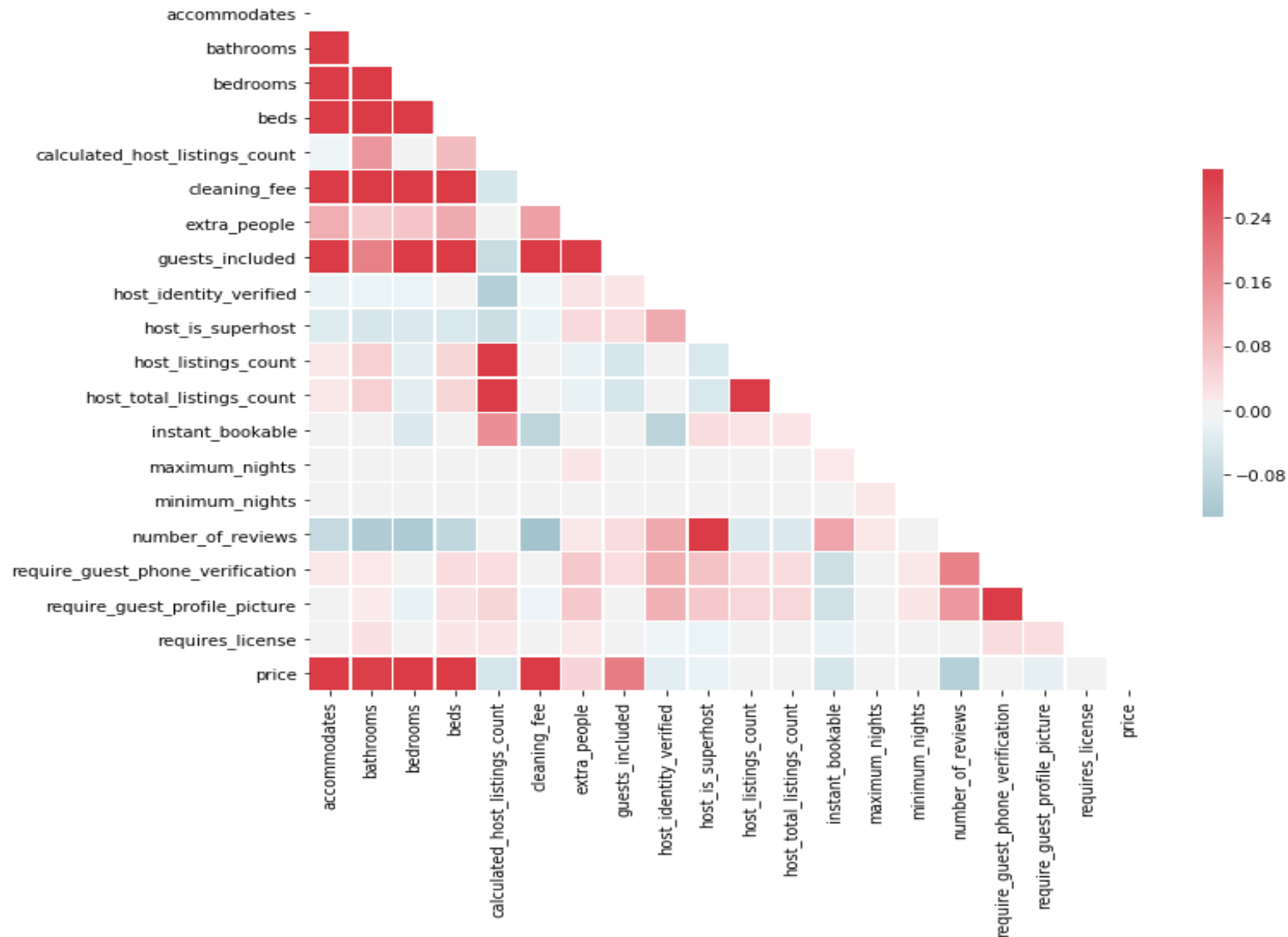
Dataset Characteristics

- Geography – Single city – San Francisco, CA
- Size – Original dataset – 3GB
 - Cleaned dataset – 1 GB
- Details – Listings Characteristics, Reviews and Price
- Source – Open Source for 50+ cities around the globe

San Francisco Dataset



Predictors Correlation Matrix



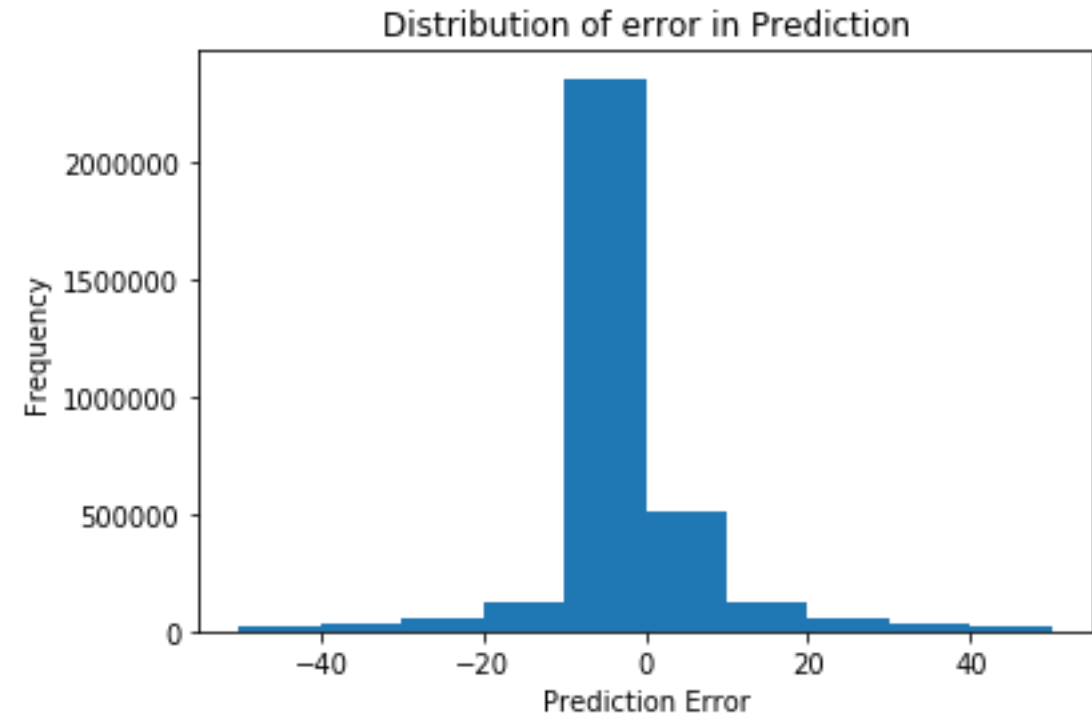
Prediction Error

Price Prediction Model Development:

- 1) Balanced the dataset
- 2) Fit a Random Forest Regressor
- 3) Performed Cross-validation

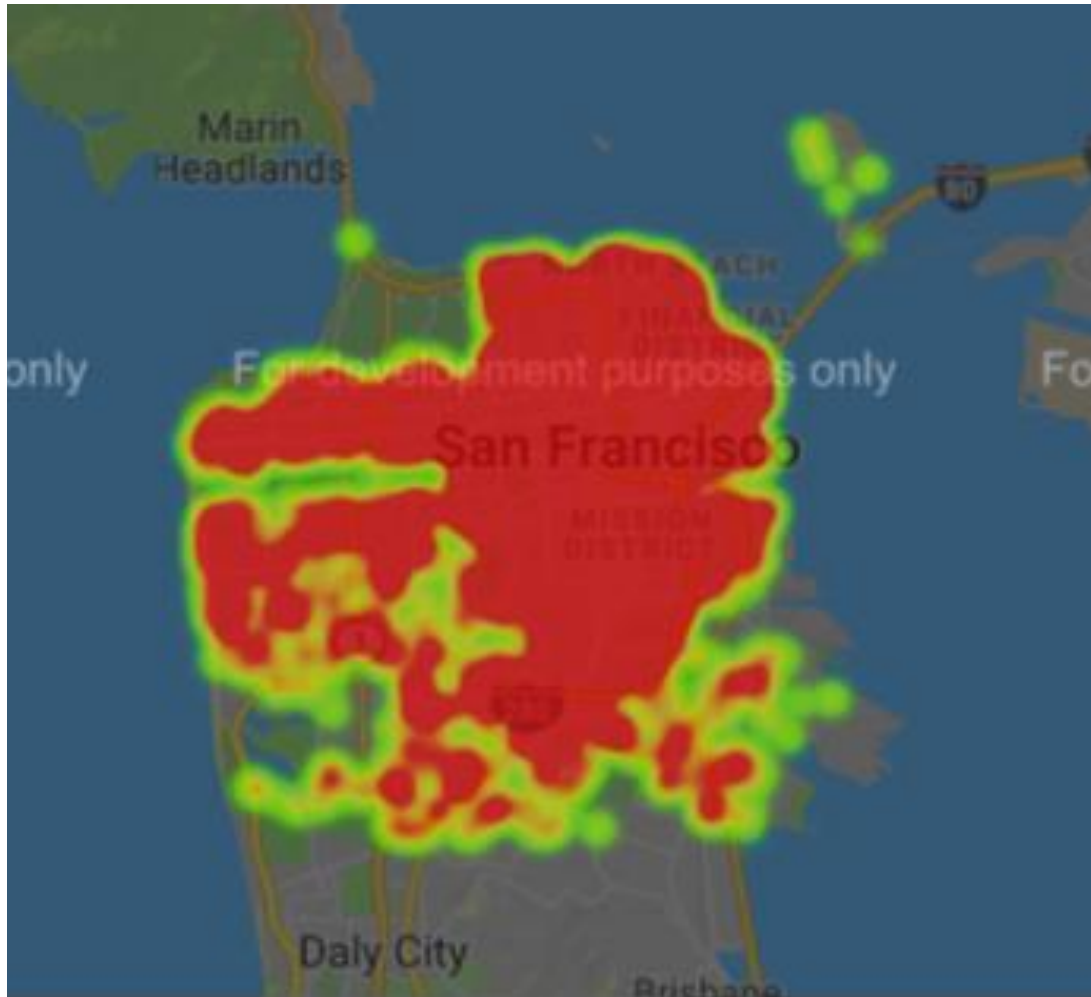
Randon Forest Regressor Model RMSE : \$ 24

Model	RMSE (USD)	Improvement in Accuracy
Linear Regression	113	78%
Baseline	174	86%



Cross-Validation Results HeatMap

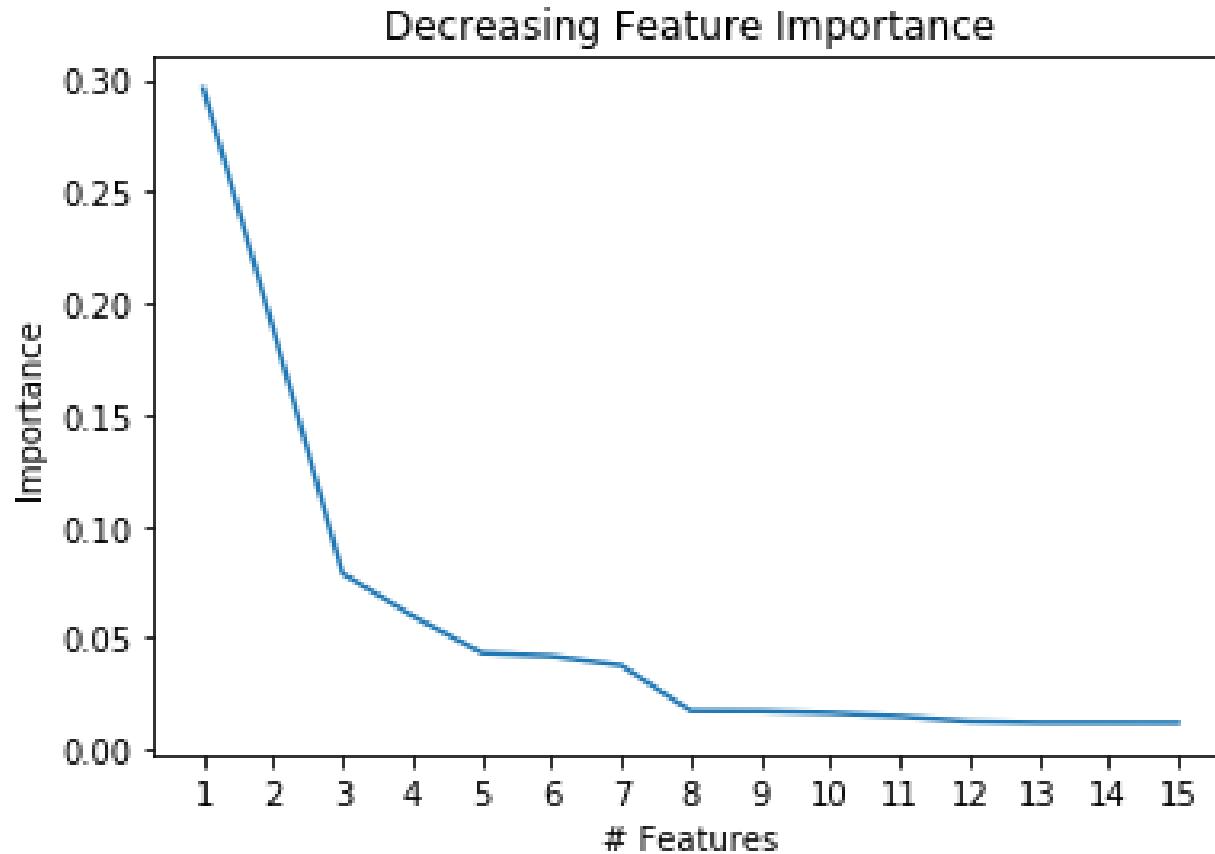
Observed Price HeatMap



Predicted Price Heatmap



Feature Importance



Features Identified

cleaning_fee	0.296059
bedrooms	0.188901
minimum_nights	0.078936
number_of_reviews	0.059952
maximum_nights	0.043266
extra_people	0.041918
accommodates	0.037742
Apartment	0.017380