Predicting AirBNB Prices in Austin, TX

Hank Butler, MS

Data Science Career Track, Springboard, May 2020 Cohort
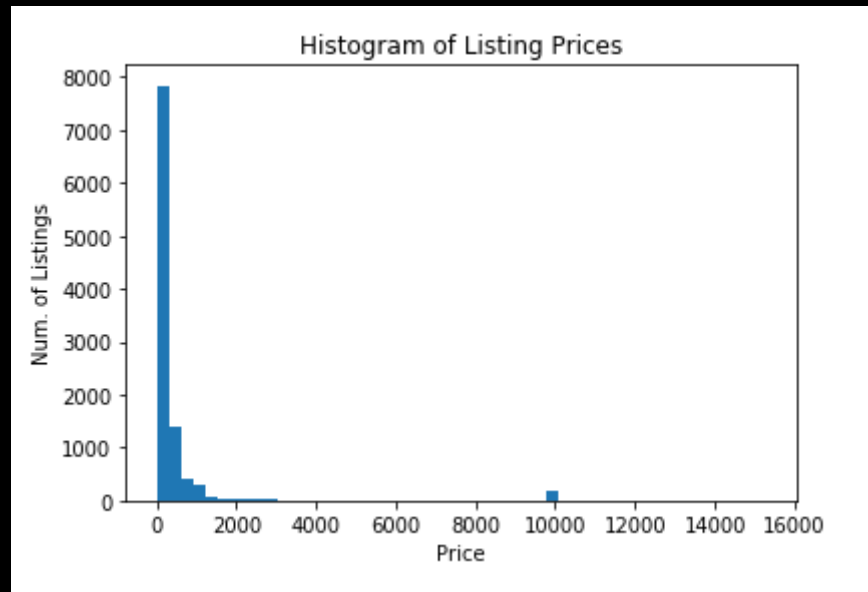
Problem

- Can AirBNB prices be accurately predicted based on their proximity to points of interest in Austin, TX?
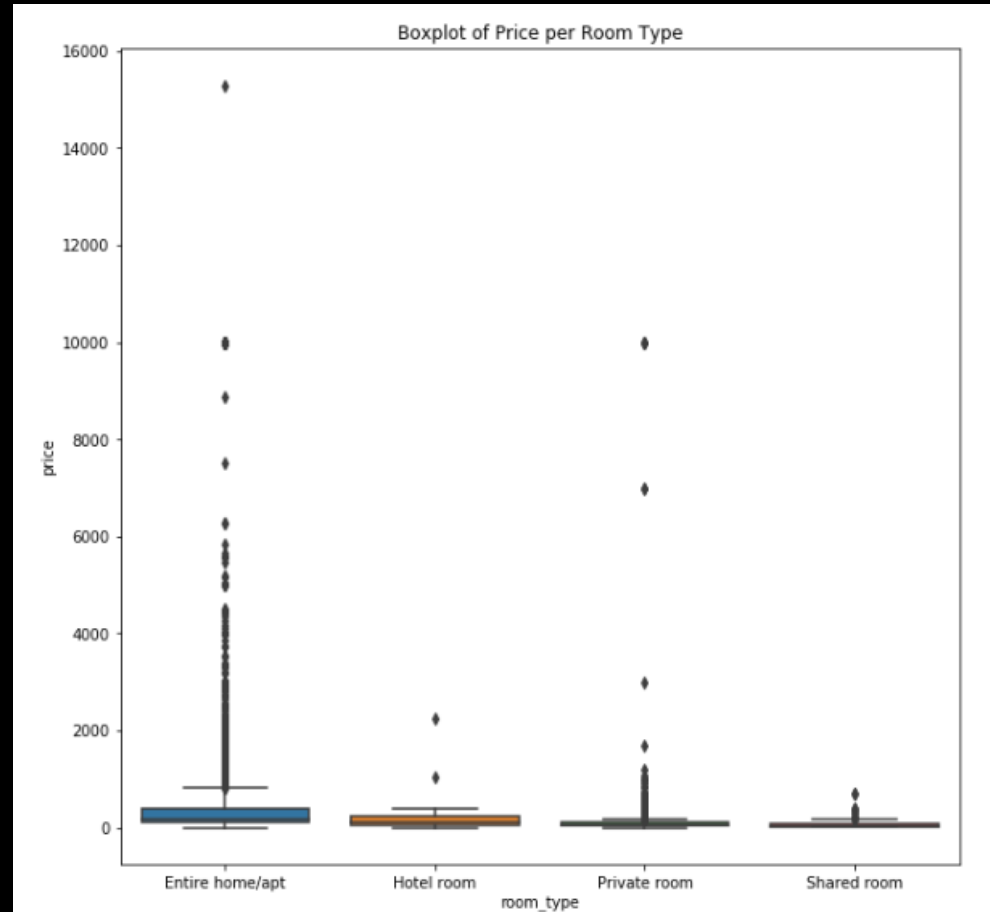
Data Sources

- AirBNB listing data was taken from InsideAirBNB
  - Includes many features
- Points of Interest
  - List of popular places from Wikipedia
  - Coordinates from Google Maps
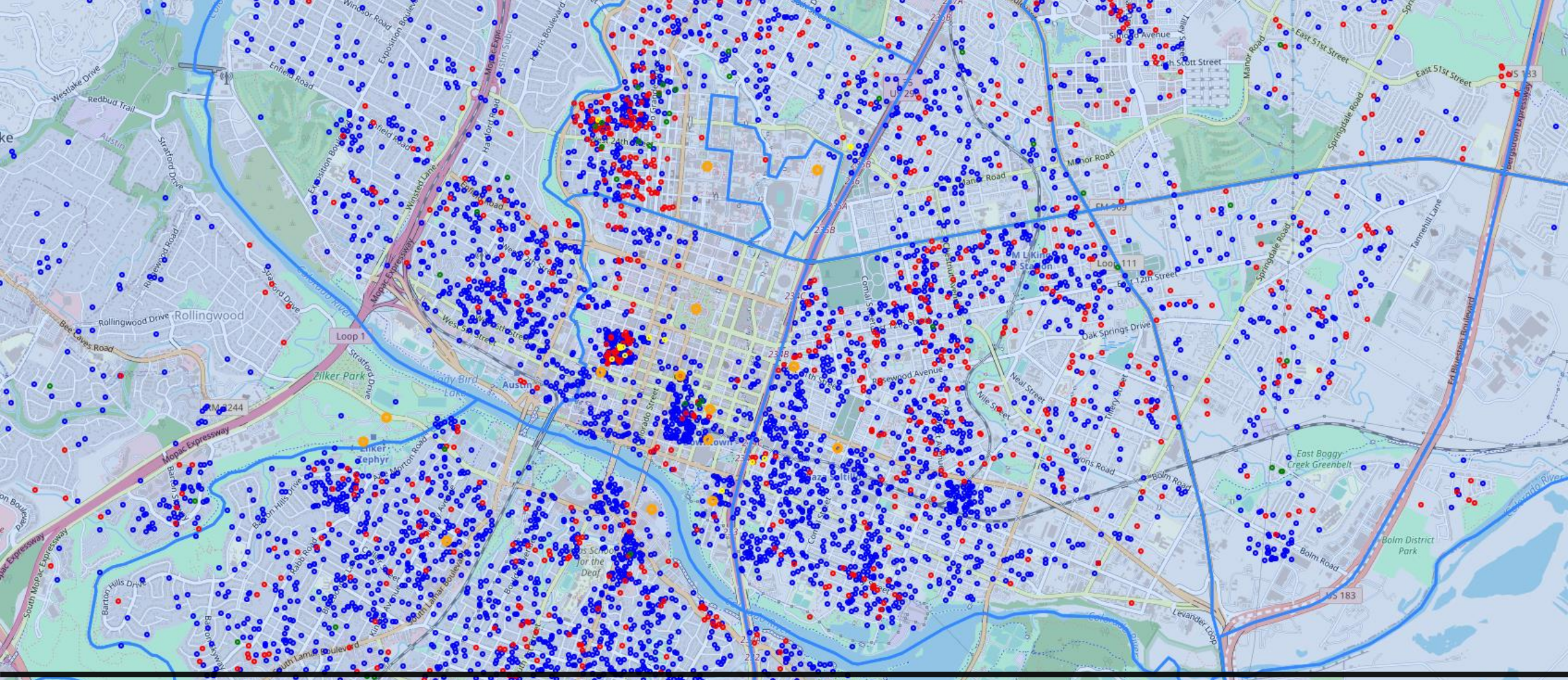
Exploratory Data Analysis

Exploratory Data Analysis

Points of Interest

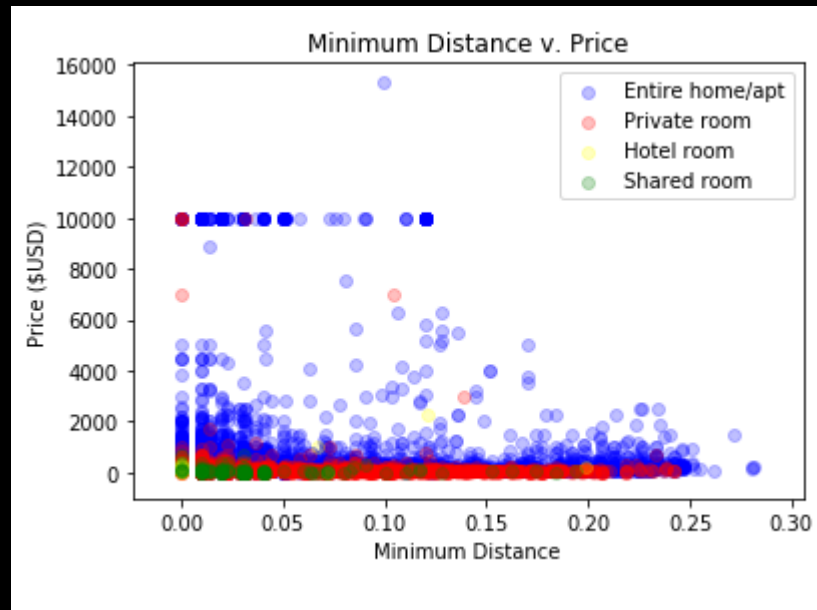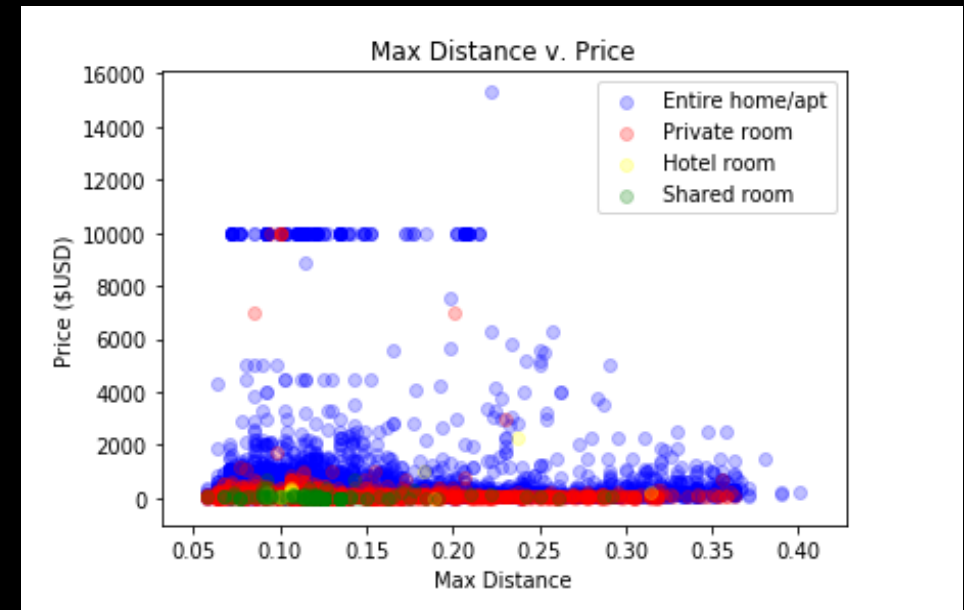| Name	Full_Address
utexas	110 Inner Campus Drive, Austin, TX 78705, USA
austin_convention_center	500 E Cesar Chavez St, Austin, TX 78701, USA
zilker_park	2100 Barton Springs Rd, Austin, TX 78704, USA
alamo_drafthouse	1120 S Lamar Blvd, Austin, TX 78704, USA
paramount_theater	713 Congress Ave, Austin, TX 78701, USA
tx_state_capitol	1100 Congress Ave, Austin, TX 78701, USA
lbj_museum	2313 Red River St, Austin, TX 78705, USA
barton_springs_pool	2131 William Barton Dr, Austin, TX 78746, USA
airport	3600 Presidential Blvd, Austin, TX 78719, USA
bat_observation_center	305 S Congress Ave, Austin, TX 78704, USA
rainey_street	82 Rainey St, Austin, TX 78701, USA
franklin_bbq	900 E 11th St, Austin, TX 78702, USA
east_6th_st	407 E 6th St, Austin, TX 78701, USA
west_6th_st	613 W 6th St, Austin, TX 78701, USA
far_east_6th_st	1502 E 6th St, Austin, TX 78702, USA

Map of Points of Interest and Listings

Listing distance to points of interest

Minimum Distance

Maximum Distance

Modeling

- Linear Models
  - Started with linear models, but quickly abandoned
- Tree Based Models
  - Random Forest Regression
  - Gradient Boosting Regression
  - Extra Trees Regression

Model Performance

- Overall weak performance from models
- Overfitting
  - Tried to combat with adjustment of hyperparameters
    - Min_samples_leaf, n_estimators, max_depth
  - Removal of unnecessary features
    - Feature_importances_ attribute
- Best performing models
  - Random Forest Regression
    - Training set score: 0.804
    - Test set score: 0.345
  - Gradient Boosting Regression

Business Impact

- Now know that distance to points of interest does have an impact on a listing's price.
  - Can ensure that equidistant listings with similar features (e.g. bedrooms and bathrooms) can be priced similarly
- Ensure that new and existing listings can be priced effectively to maximize profit for AirBNB and host.
- Shows one factor in price, further exploration into other factors that could have more explanatory impact.

How to improve

- Other data points that could explain price more thoroughly
  - Square footage
  - Amount hosts spend on rent/mortgage