A Model to Set Rents in San Francisco

Amy Sillman
Metis Project 2
January 22, 2021

Introduction

- Residential rental market in San Francisco is changing due to the COVID-19 pandemic
- Setting a monthly rent amount may be challenging
- Develop a model to set a reasonable rent for a San Francisco apartment, based on apartment attributes
- Data source: Craigslist San Francisco apartment search
- Features: Number of bedrooms, number of bathrooms, square footage
- Target: Monthly rent

Process

Acquisition

Craigslist Search

Primary Scrape (listing of URLs)

Secondary Scrape (apartment pages)

Cleaning

Convert bedroom NaNs

Bathrooms

Null rent

Numeric data types

Duplicates!

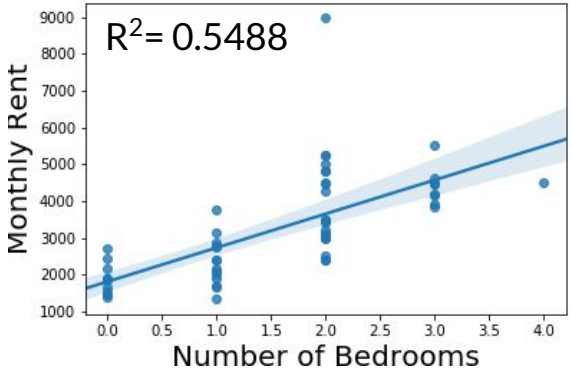
EDA, Analysis

Seaborn pairplot

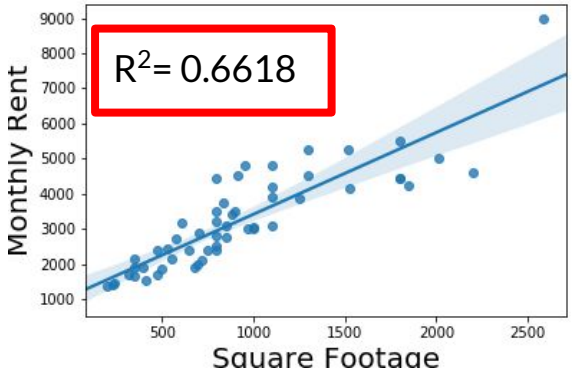
Linear Regression Analysis

Model Training

- 54 unique records that included the square footage feature
- Train-test split was 0.67/0.33
- Models were trained on the training data, and then tested on the testing data

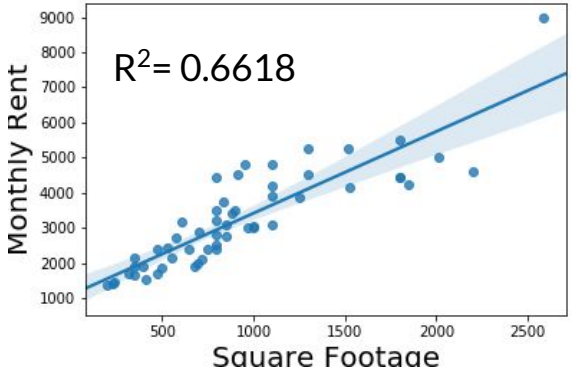


R^2= 0.2944
(test data)

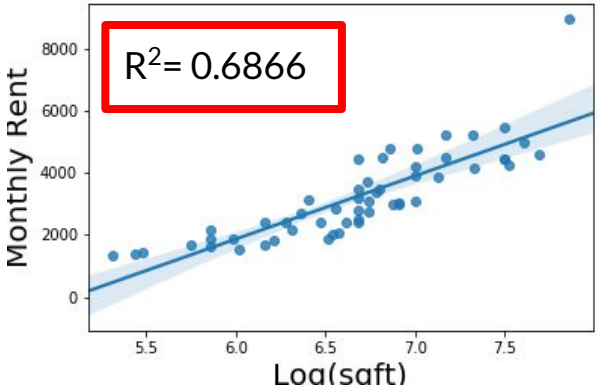


R^2= 0.8160
(test data)

Log(sqft) results in a better fit



R^2= 0.8160
(test data)



R^2= 0.6873
(test data)

Next Steps

- Troubleshoot initial scraping code
- Find additional features that may affect rental price
- Break data down by neighborhood; better recommendations may be possible

Thank you!

- **Metis staff, especially Lucy and Tina**
- **My project group: Emma-Claire McCarthy and Sara Zong**

Diagnosing the duplicates issue