

What makes an apartment great?

\$3,795 / 2br - 950ft² - Spacious 2BR,1BA PacHeights, washer dryer, Lafayette pk (pacific heights)



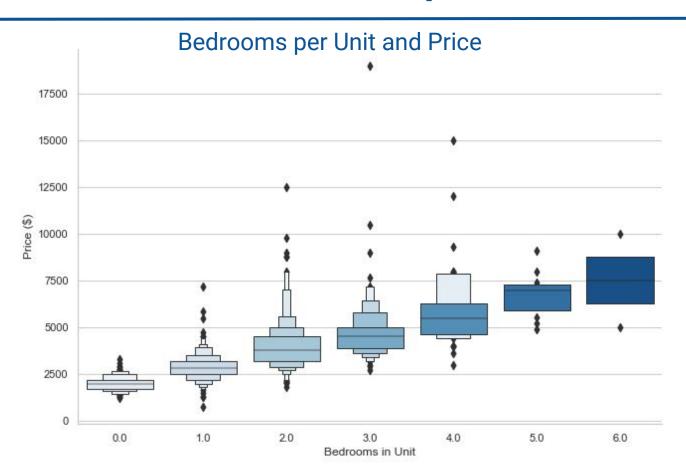
Available now, nice, spacious, immaculate, FURNSHED flat (will consider unfurnishing) in best location. Ideal home for families, travel nurses, vacationers, corporate stays; two large bedrooms, bathroom with separate bathtub and shower, living Room, dining area, eat-in Kitchen, granite counters, breakfast nook, dishwasher, microwave, gas stove, office area, two large walk- in closets .Washer -dryer in this apt unit.



Strong relationship between unit size and price...

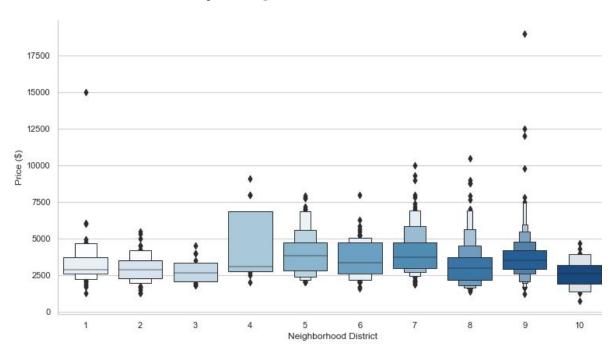


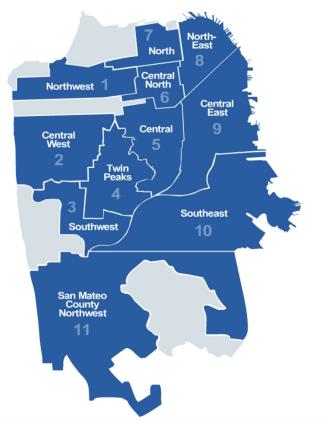
...and between bedrooms and price



Location can affect high-end of price range

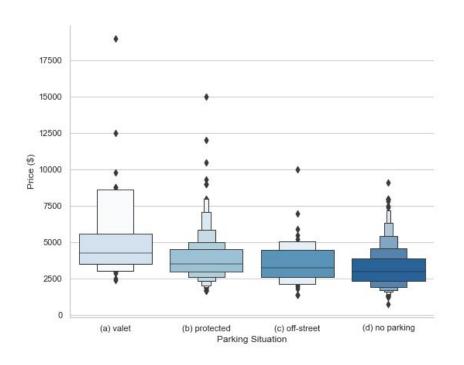




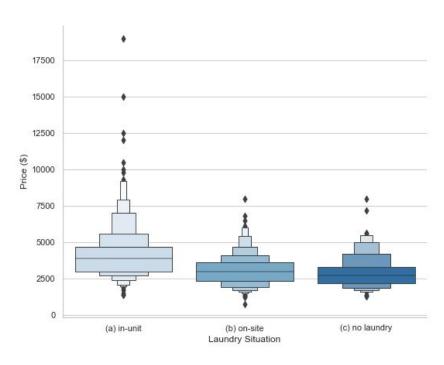


Some amenities command a premium

Price by Parking Situation



Price by Laundry Facilities



Model construction and selection

Feature Selection

Engineering

Model Selection

Continuous Variables

→ Unit size (sqft)

Discrete Variables

- → Bedrooms
- → Bathrooms

Ordinal Variables

- → Parking
- → Laundry

Categorical

- → Pets
- → Neighborhood

Polynomial Features

→ Degree = 2

Regularization methods tried

- \rightarrow Ridge ($\alpha = ~8$)
- \rightarrow Lasso ($\alpha = \sim 6$)

Ridge scored most consistent between training and validation testing

- \rightarrow R²
- → RMSE
- → MAE

Lasso performed consistently on R², but showed greater variance between training and validation scoring on RMSE and MAE

Model results

R² Score

→ Train: 81%

→ Test: 77%

RMSE

→ Train: \$451

→ Test: \$477

→ +5.7%

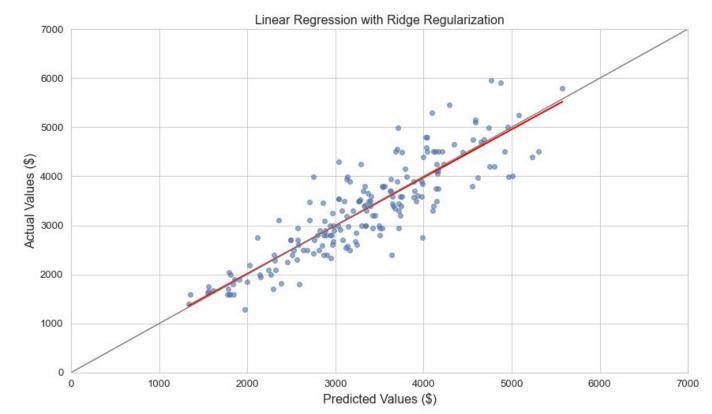
MAE

→ Train: \$341

→ Test: \$366

→ +7.3%





Model evaluation

Tolerance of price variation is *relative* to apartment price and *subjective* to budget

- → ~\$400 can be substantial for one person
- → but split between two or more people may be okay

Model does a poor job of predicting for units > \$4,000 / month

- → Residuals plot¹ shows greater variance above \$4k
- → Most SF apartments in sample fall between \$2,500 and \$4,000 per month²

¹Appendix A - Residuals plot of predicted values

²Appendix B - Distribution of rental prices in sample

Final thoughts

There could be other factors influencing rents

- → Rent history, control
- → Other amenities not standard in Craigslist posts (fitness facilities, common spaces, etc.)
- → Other environmental or economic factors¹

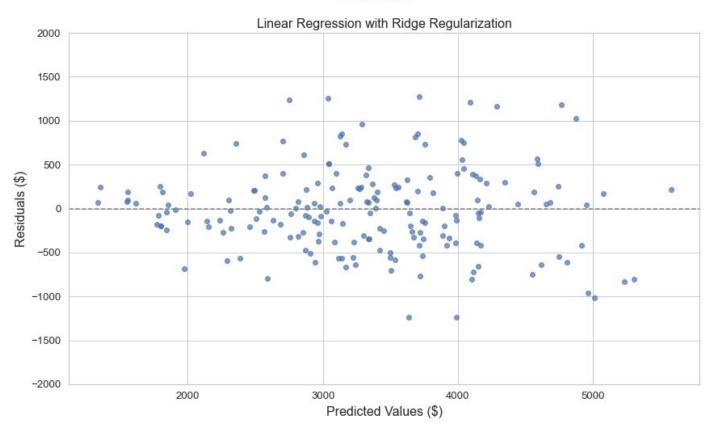
Suggestions for future work

- → Sample again when COVID's influence subsides
- → Increase scope of scraping methods
 - Search text for key words like 'gym' or 'backyard' as additional amenities
 - Find information on security deposit and/or lease requirements
 - Implement image processing tools to evaluate apartment quality from photos
- → Scam post identification

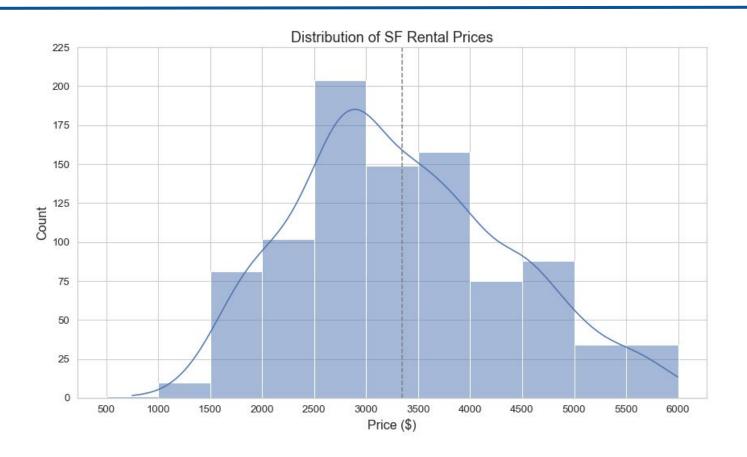


APPENDIX A - Residuals plot of predicted values





APPENDIX B - Distribution of rental prices



APPENDIX C - Declining SF Rents

San Francisco Rent Growth Over Past 12 Months

