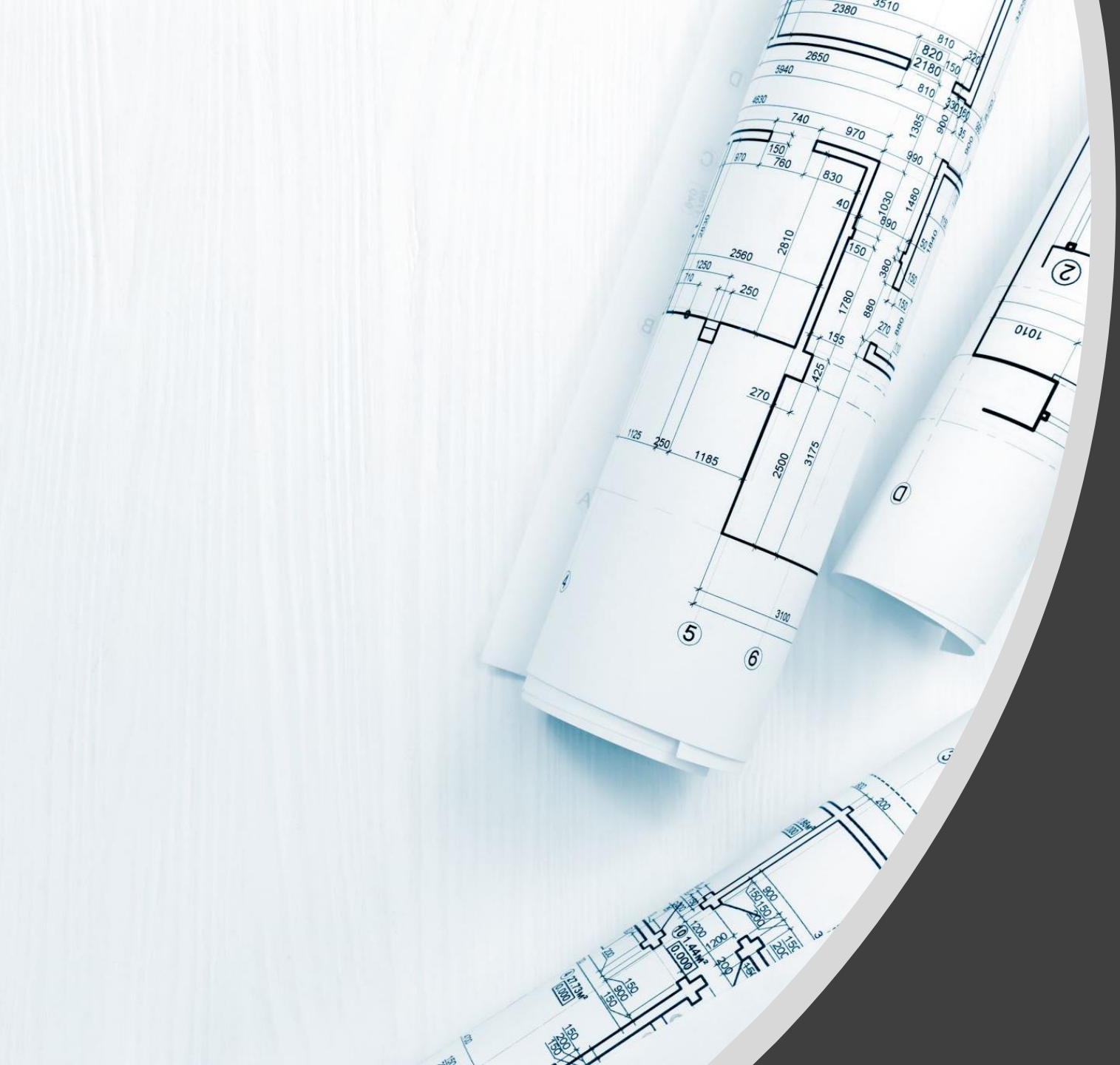


King County Housing Data Regression Project

Johnny Dryman

Overview

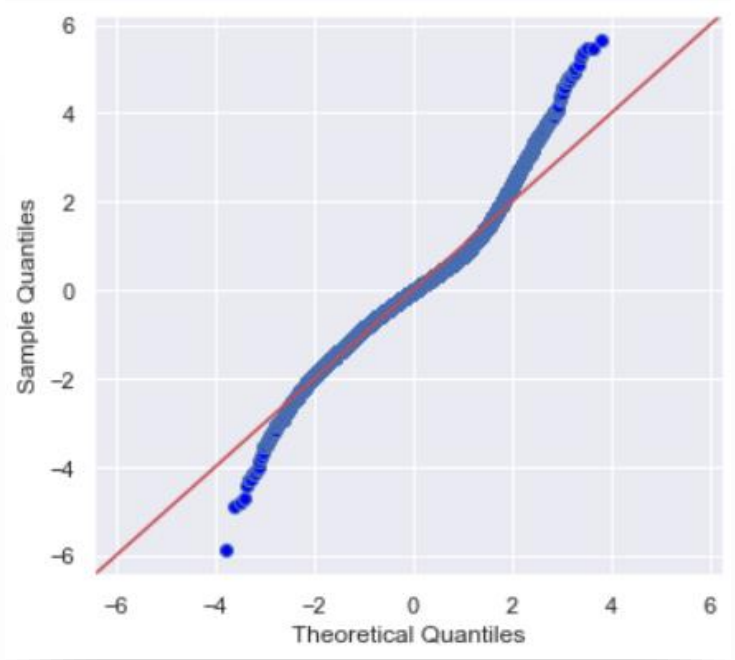
- This project was designed to analyze home and property value in King County (Seattle, WA area)
- Data was obtained from King County home sales between May 2014 – May 2015
- Questions:
  - What quantities and / or qualities are most influential in determining sale price?
  - How can a resident of King County increase the value of their home?

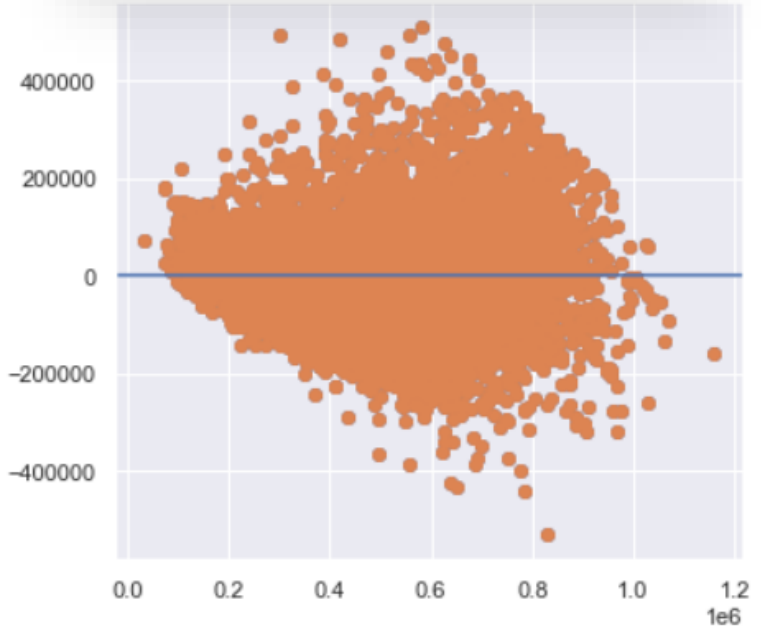
Data

Sale Price	# Floors	Living Area Square Footage	Lot Square Footage	Waterfront (Y/N)
Condition	Year Built	Zip Code	Basement (Y/N)	Renovated (Y/N)

Regression Results

- Model returned an Adjusted R-squared of **0.798**
  - This means our model can explain 79.8% of the variance in price
- QQ plot shows model is mostly normal with fatter tails
- Homoscedasticity check is mostly cone-like
  - This indicates that our model is unbiased





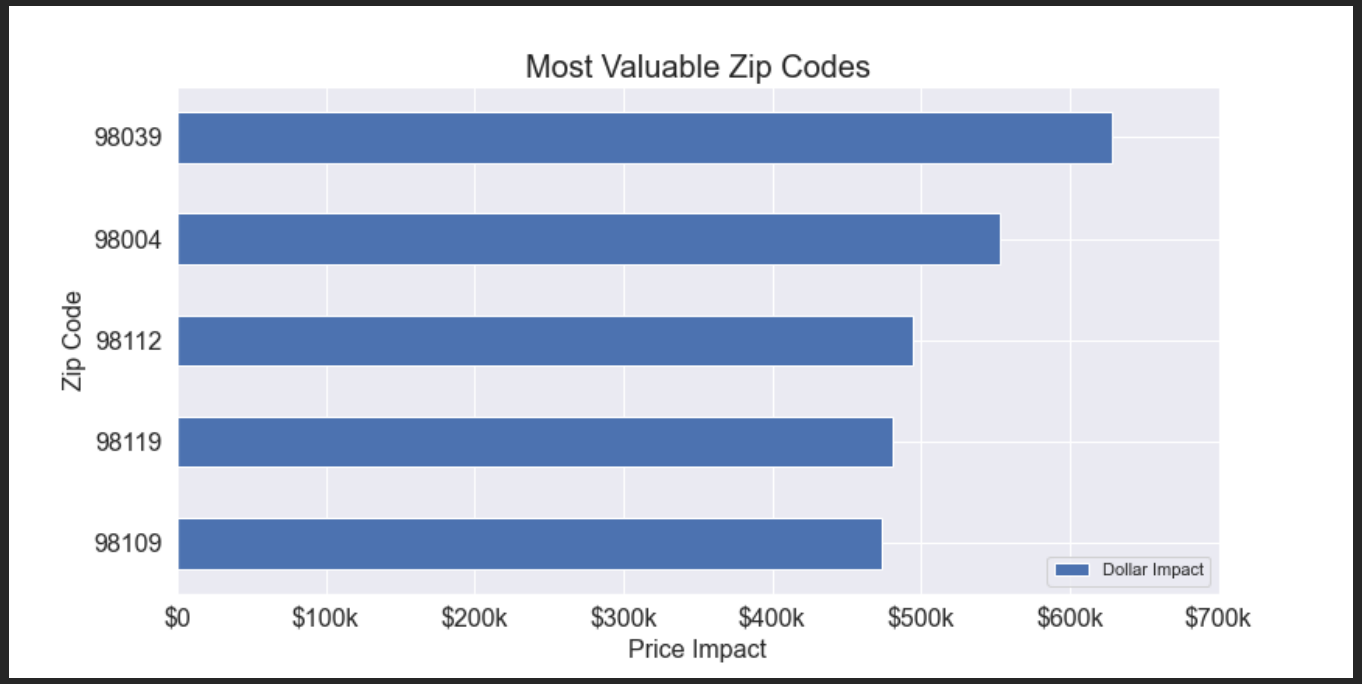
Findings

Primary Price Drivers

- **Living Square Feet** - $159 / sqft
- **Zip Codes** - top 5 add $473-628k
- **Waterfront** - $338k
- **Lot Square Feet** - $3.45 / sqft
- **Basement** - $23,634 penalty
- **Renovated** - 42,260 bonus
- **Condition** – between $0 and $174k

Less Significant Features

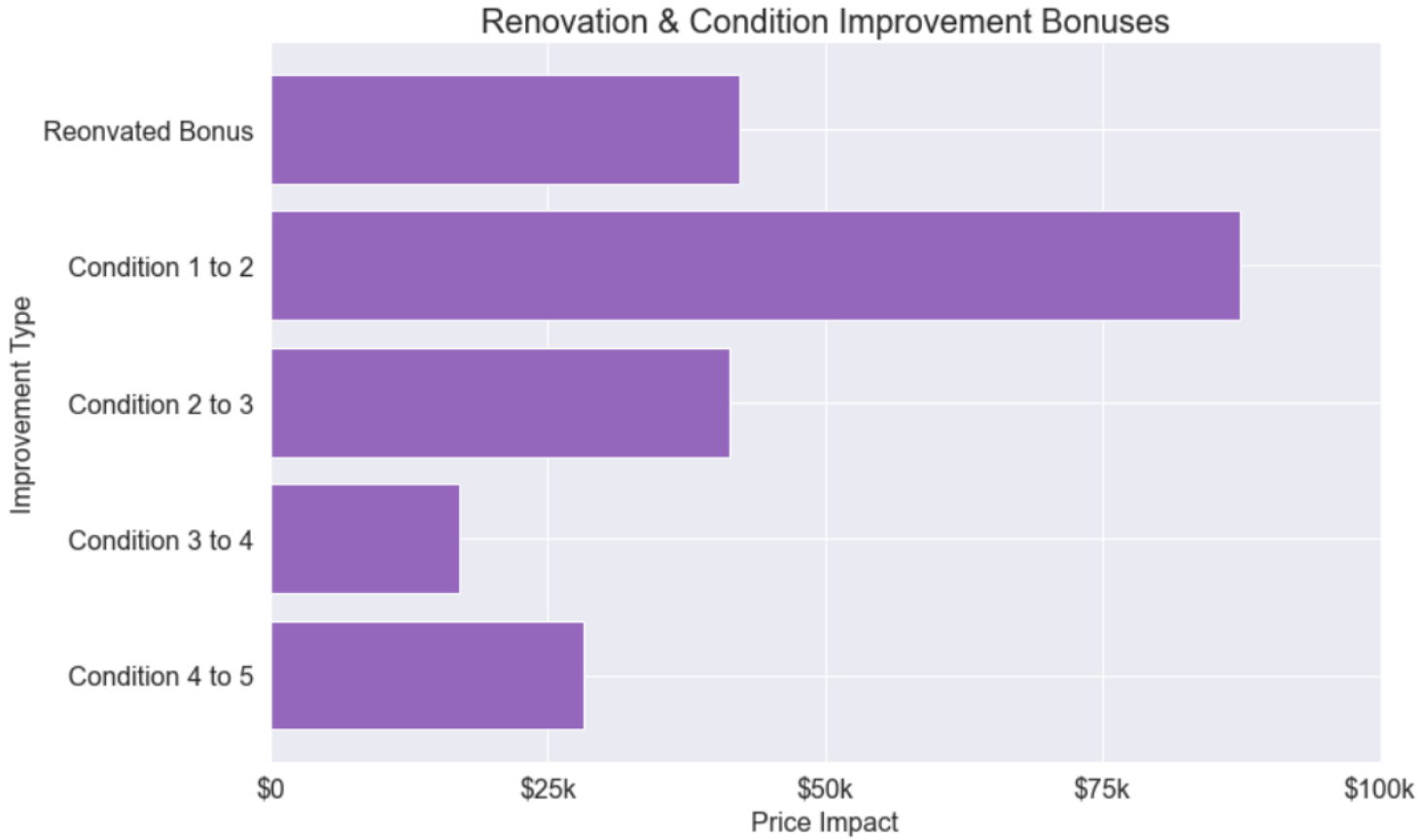
- **# Floors** – penalty or bonus vary
- **Age** - $168 penalty per year



Zip Codes Most / Least Valuable

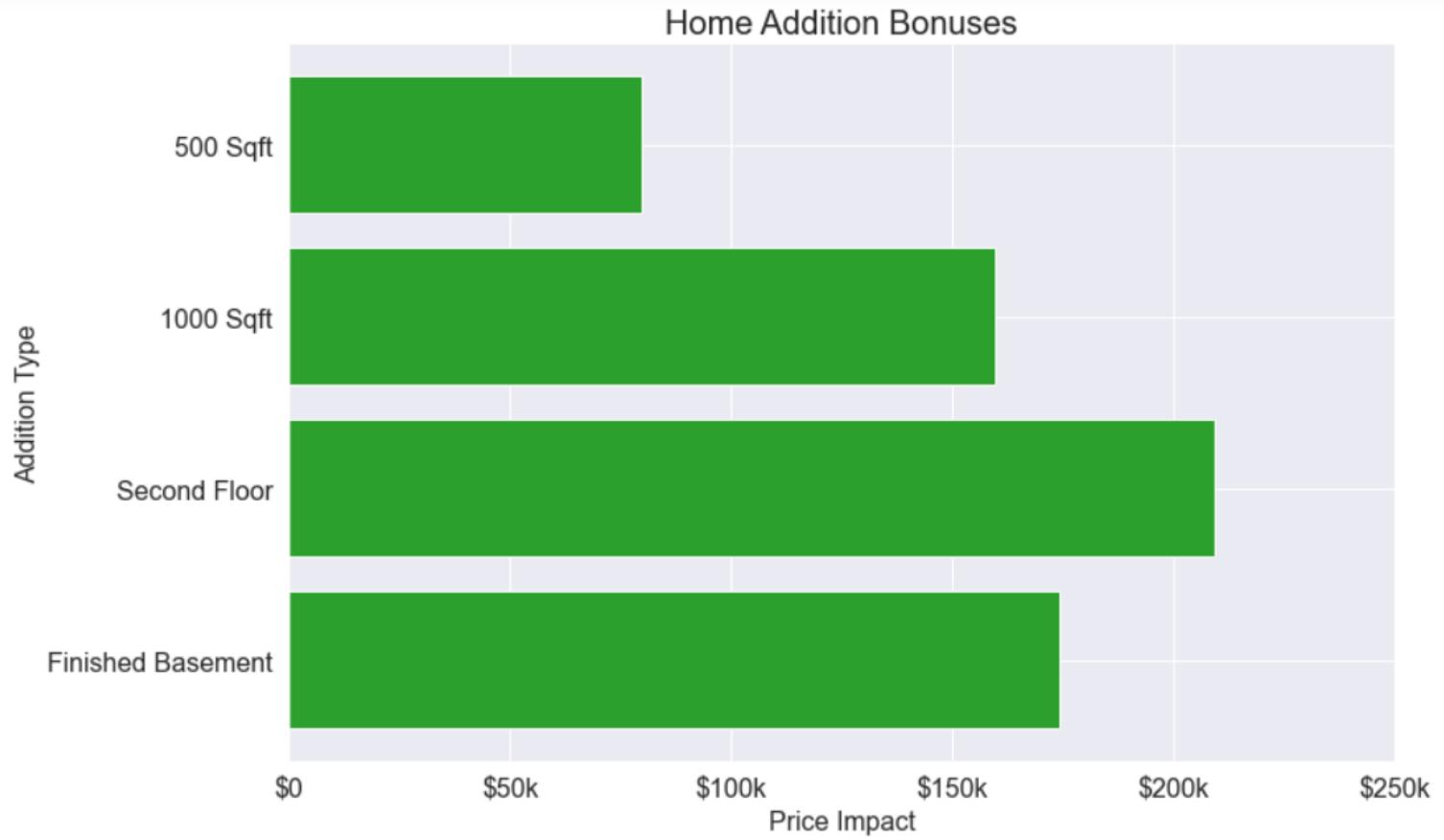
- Top 5 Zip Codes
  - Add $473-628k to value
  - Located in metro area (Seattle, Bellevue, Mercer Island)
  - Closer to water
- Bottom 5 Zip Codes
  - Range from $5k penalty to $10k bonus
  - Located in southern King County, Kent area
  - Landlocked

Recommendation:
Renovate and Improve / Maintain Condition



- Renovating to improve condition will provide $42k bonus
- If the renovation improves the condition, additional bonus will be applied
  - Condition 1 to 2: + $87,360
  - Condition 2 to 3: + $41,455
  - Condition 3 to 4: + $17,031
  - Condition 4 to 5: + $28,288
- Invest in regular maintenance to avoid condition deterioration penalty

Recommendation:
Add Living Square Footage through Construction



- Each additional square foot will add $159 to the home value
  - **500 sqft: $79,740**
  - **1000 sqft: $159,480**
- Building a second floor (approx. 1240 sqft)
  - 1240 sqft: $197,755
  - 2nd floor bonus: $11,448
  - **Total: $209,204**
- Finishing a basement:
  - 1240 sqft: $197,755
  - Basement penalty: (-) $23,634
  - **Total: $174,121**

Conclusions

- Living Square Footage is most significant factor in home price
- Zip Code is a primary price driver
  - Houses in city center and near water have higher value
  - Landlocked houses further from Seattle (especially in the Kent area) have less value
- Recommendations:
  - Add living square footage via extension, finished basement, or second floor
  - Renovate
  - Improve / maintain condition

Next Steps

- Implement Latitude, Longitude, Year Renovated, and Living & Lot Square Footage for closest 15 neighbors
- Develop heatmap to refine geographic understanding
- Normalize features to improve predictive quality
- Create dynamic splitting functionality to run model on filtered datasets
    - Example: how specifically could the owner of a 2 story, 4 bedroom house in Bellevue improve their home value?

Thank you for your time!

Please feel free to ask any questions.

Appendix

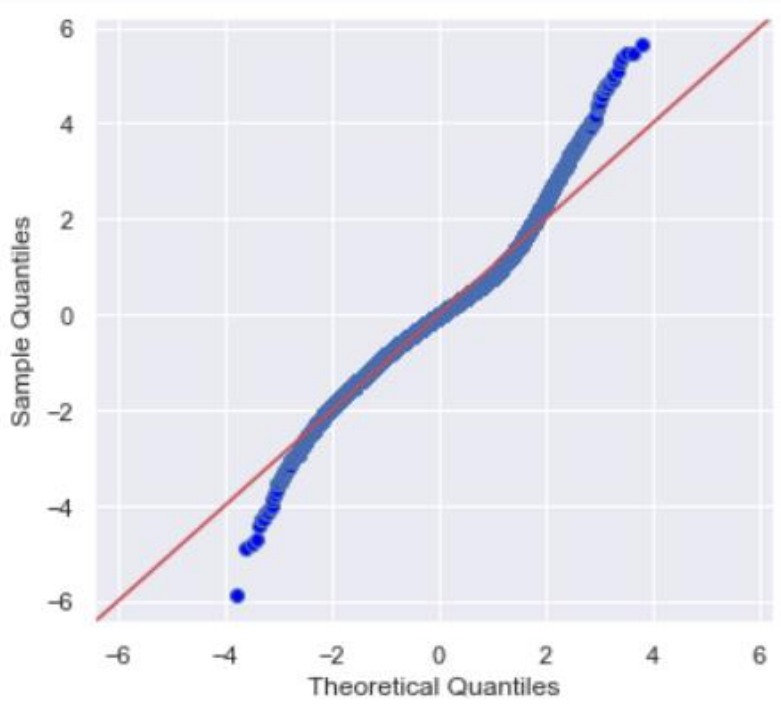
Unused Data

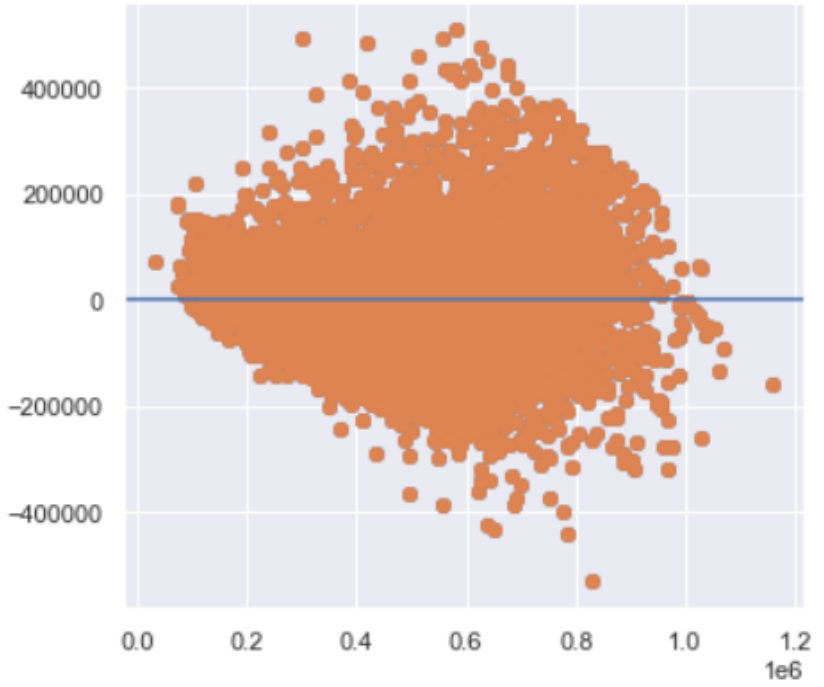
- Sale Date
- # Bedrooms
- # Bathrooms
- View
- Grade
- Basement Square Footage
- Non-Basement Square Footage
- Year Renovated
- Latitude
- Longitude
- Living Square Footage – 15 Nearest Neighbors
- Lot Square Footage – 15 Nearest Neighbors

Regression Model

OLS Regression Results

Dep. Variable:	price	**R-squared:**	0.799
Model:	OLS	**Adj. R-squared:**	0.798
Method:	Least Squares	**F-statistic:**	617.7
Date:	Wed, 21 Apr 2021	**Prob (F-statistic):**	0.00
Time:	16:52:19	**Log-Likelihood:**	-1.6841e+05
No. Observations:	13128	**AIC:**	3.370e+05
Df Residuals:	13043	**BIC:**	3.376e+05
Df Model:	84		
Covariance Type:	nonrobust		





P-Value Rankings

	Variable	P_Value
1	sqft_living	0.0000000000
35	C(waterfront)[T.1.0]	0.0000000000
48	C(has_basement)[T.1]	0.0000000000
49	sqft_lot	0.0000000000
51	C(renovated)[T.1]	0.0000000000
54	Intercept	0.0000000000
55	C(condition)[T.5]	0.0000000000
58	C(condition)[T.4]	0.0000000085
62	C(condition)[T.3]	0.0000003648
66	C(floors)[T.2.0]	0.0000224120
68	C(floors)[T.3.0]	0.0001514949
69	age	0.0004193930
70	C(condition)[T.2]	0.0011577263
71	C(floors)[T.1.5]	0.0044800397
78	C(floors)[T.3.5]	0.3210375106
83	C(floors)[T.2.5]	0.9151773929