

Housing Price Prediction





Business Problem



Seattle real estate firm noticed that there are no tools that allow their clients to enter information about their home and receive a prediction for their home's sale price. So, the real estate firm has hired us to produce a predictive model that can predict the sale price of a home in King county Washington as accurately as possible.

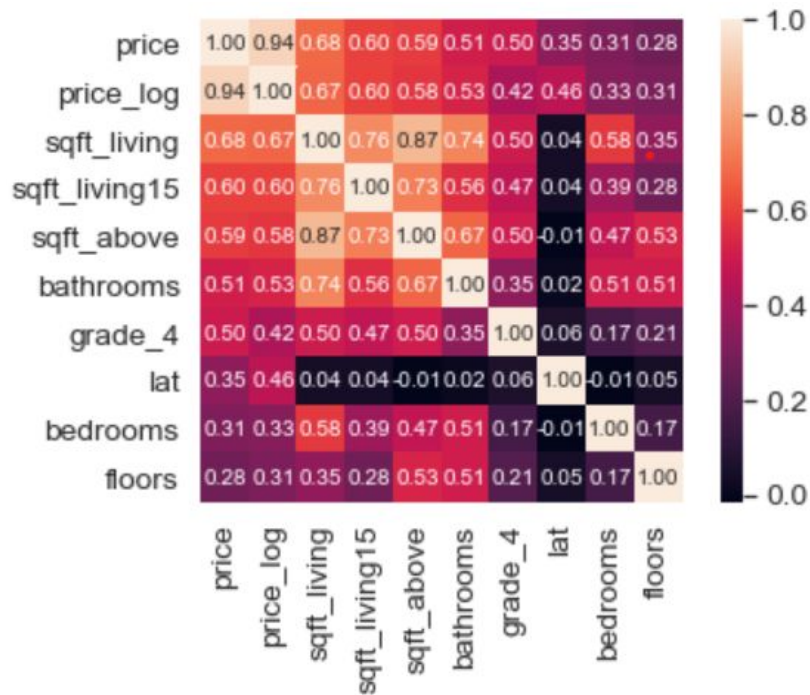


Methods/Data



- Housing Sales from May 2014-May 2015 in King County, Washington
- Contained over 20 variables
- Found best predictors through correlation
- Eliminated outliers
- Narrowed down to 12 variables
- Created a multiple linear regression model to predict price
- Altered what data until most accurate

price	1.000000
sqft_living	0.701917
grade	0.667951
sqft_above	0.605368
sqft_living15	0.585241
bathrooms	0.525906
view	0.395734
bedrooms	0.308787
lat	0.306692
waterfront	0.264306
floors	0.256804
yr_renovated	0.129599
sqft_lot	0.089876
sqft_lot15	0.082845
yr_built	0.053953
condition	0.036056
long	0.022036
id	-0.016772
zipcode	-0.053402



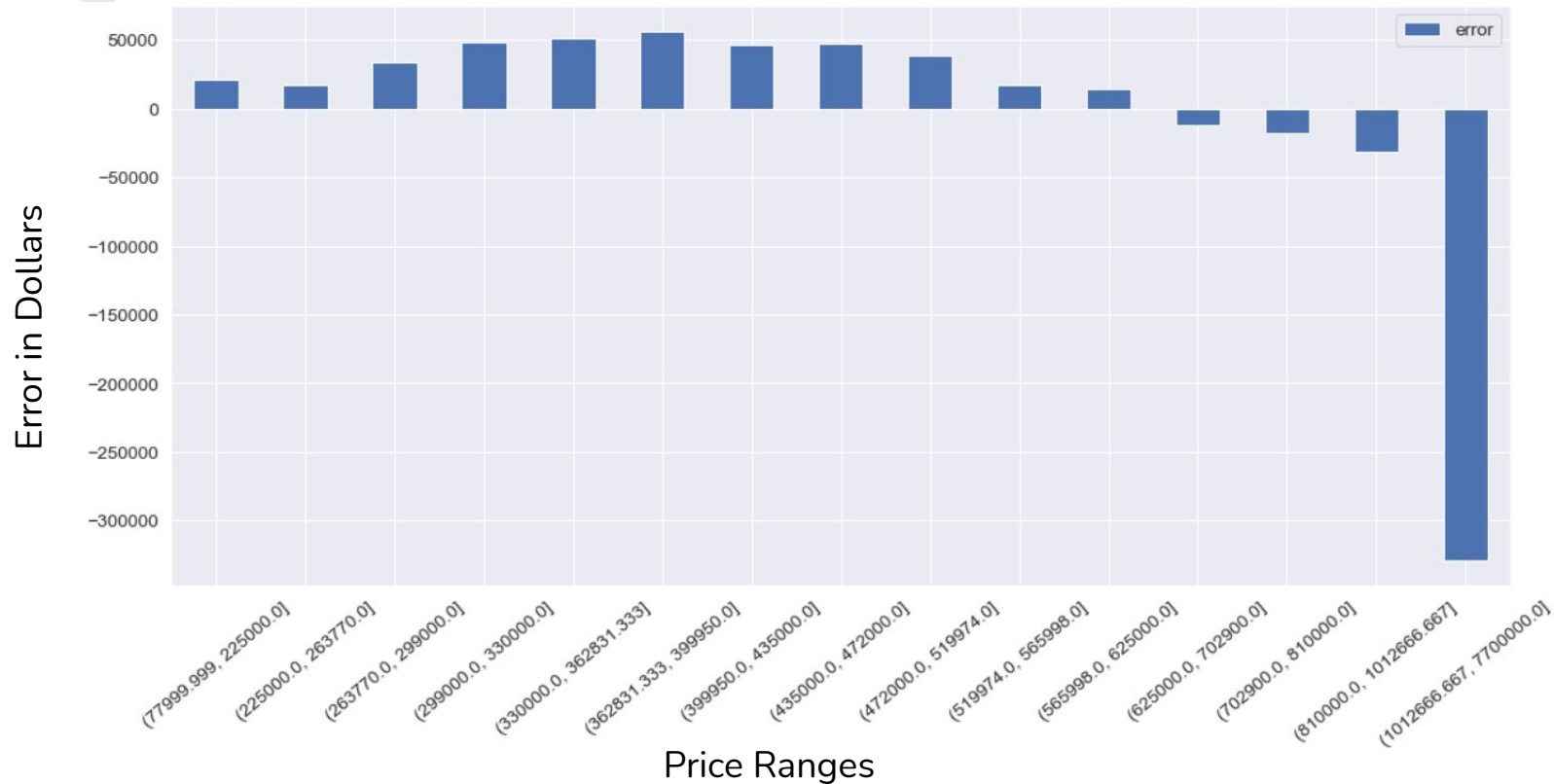




Model 1

Factors: #bedrooms, #bathrooms, square footage of living space, # floors, waterfront, Condition of the house, square feet above ground, latitude, square footage of the 15 nearest houses, grade of house

Accuracy: average error of \$208,440.52





Model 2

Changes: Dropped outliers

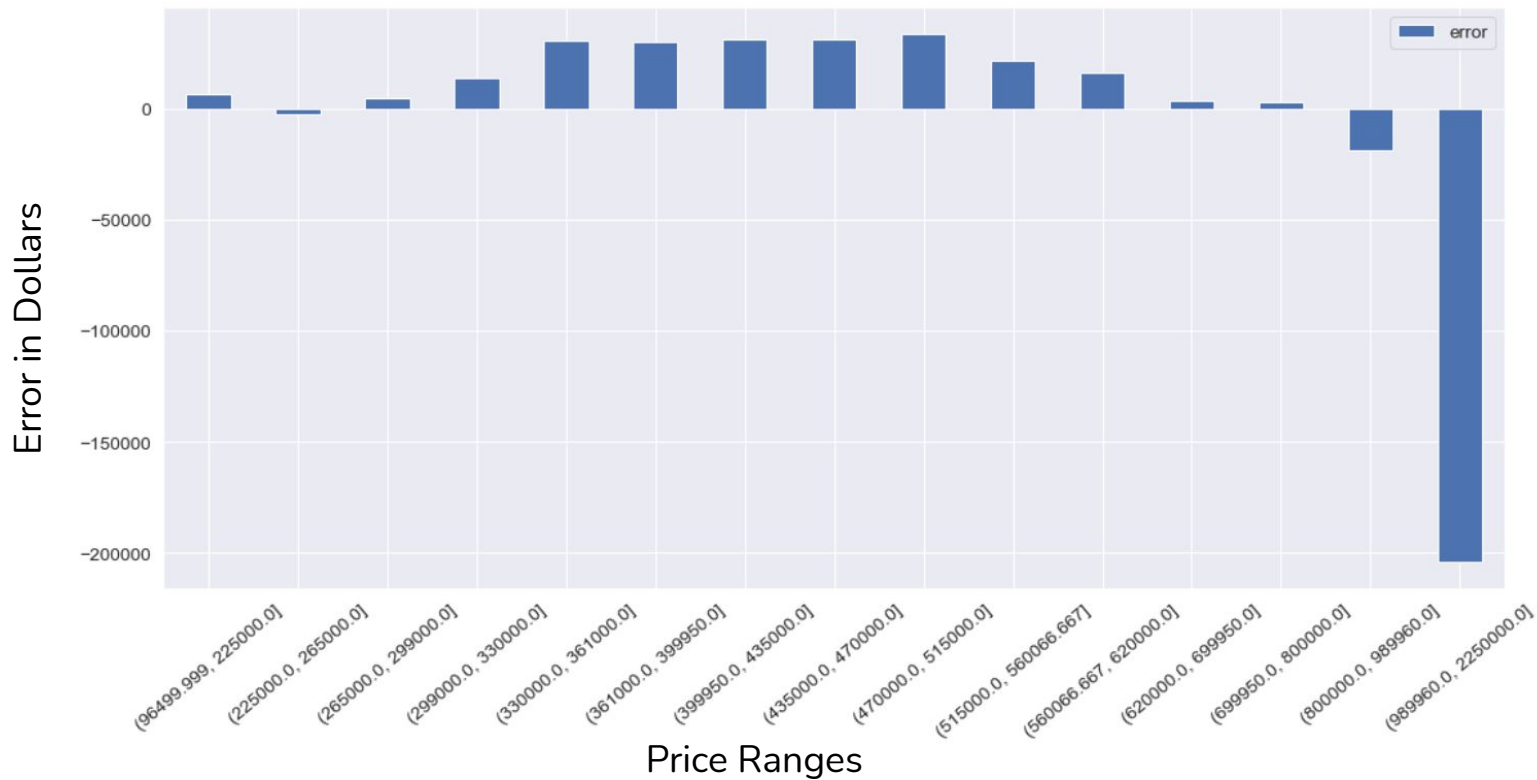
Accuracy: average error of \$165,753.05



Model 3

Changes: Factored in Zip Code

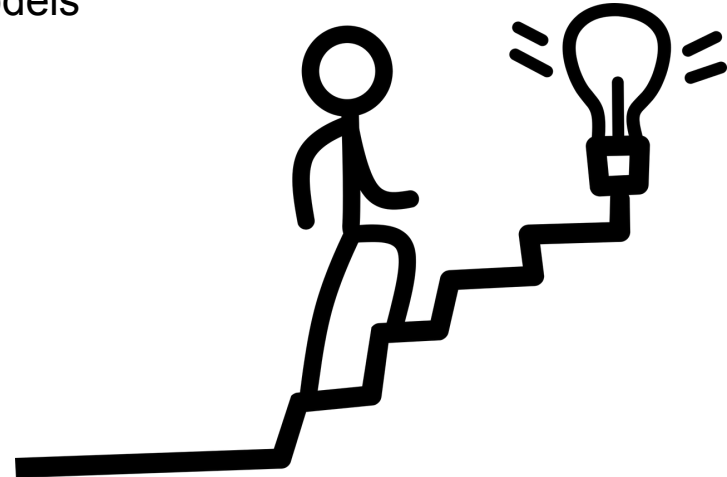
Accuracy: average error of \$124,774.52





Conclusion and Future Improvements

- Model accurate within ~\$50,000 up to ~\$1.1 million
- Significantly undervalues the most expensive houses
- Would highly recommend for the lower ranges
- Must Identify variables that impact the priciest houses
- Possibly create more advanced predictive models



Thank You



**For any additional questions,
please contact us via email**

Email:

Qiwen.ou0721@gmail.com

Robertoliver2000@gmail.com