King County Home Sales

What Factors Affect Sale Price

Variables

Number of bedrooms

Number of bathrooms

Square footage of the home

Square footage of the lot

Number of floors

Has waterfront view

Number of times viewed

Grade of home based on KC grading system

Square footage of home excluding basement

Square footage of the basement

Year built

Year when house was renovated

Square footage of lot after renovation

Whether or not the home was viewed

Whether or not the home has a basement

Condition of home

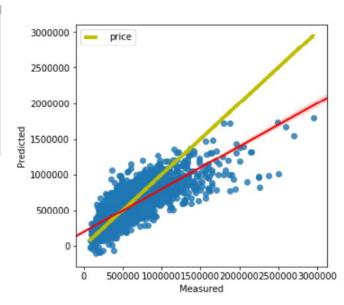Square footage of home after renovation
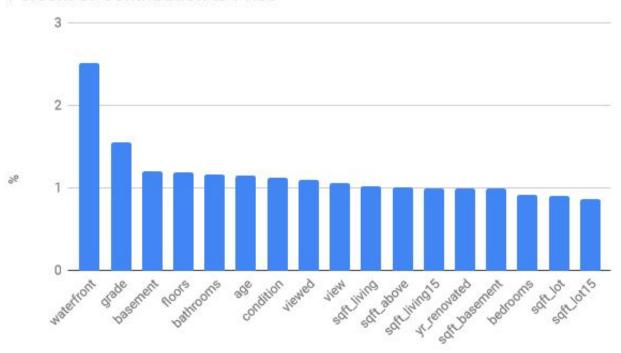
Methodology

OSEMN model

Obtain	Scrub	Explore	Model	iNterpret
• From other location • Query from database or API • Extract from another file • Generate data (e.g. Sensors)	• Filtering lines • Extracting columns or words • Rreplacing values • Handling missing values • Converting formats	• Understanding data • Deriving statistics • Creating visualizations	• Clustering • Classification • Regression • Dimensionality reduction	• Drawing conclusion from data • Evaluating meaning of results • Communicating result

Source: A Taxonomy of Data Science

Model Results

R-squared	% Difference MSE(test vs train)
0.620655	-0.943803

Variable	Coefficient
waterfront	2.512872
grade	1.558873
basement	1.209310
floors	1.196973
bathrooms	1.166954
age	1.160182
condition	1.123023
viewed	1.105606
view	1.069390
sqft_living	1.030210
sqft_above	1.009416
sqft_living15	1.000345
yr_renovated	1.000074
sqft_basement	1.000026
bedrooms	0.923841
sqft_lot	0.909553
sqft_lot15	0.869656

Percent of Contribution to Price

Maximize Sale Price

Focus more on homes with a waterfront view, a high housing grade given by King County, and having or adding a basement.

Focus less on size of basement, number of bedrooms, and size of lot.