

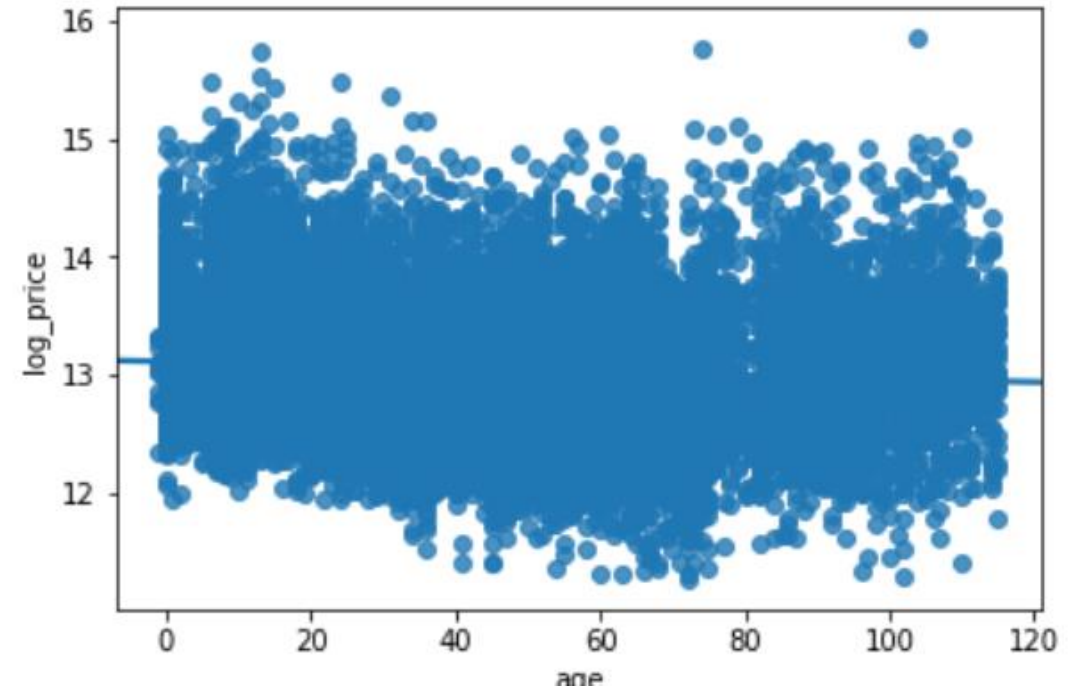
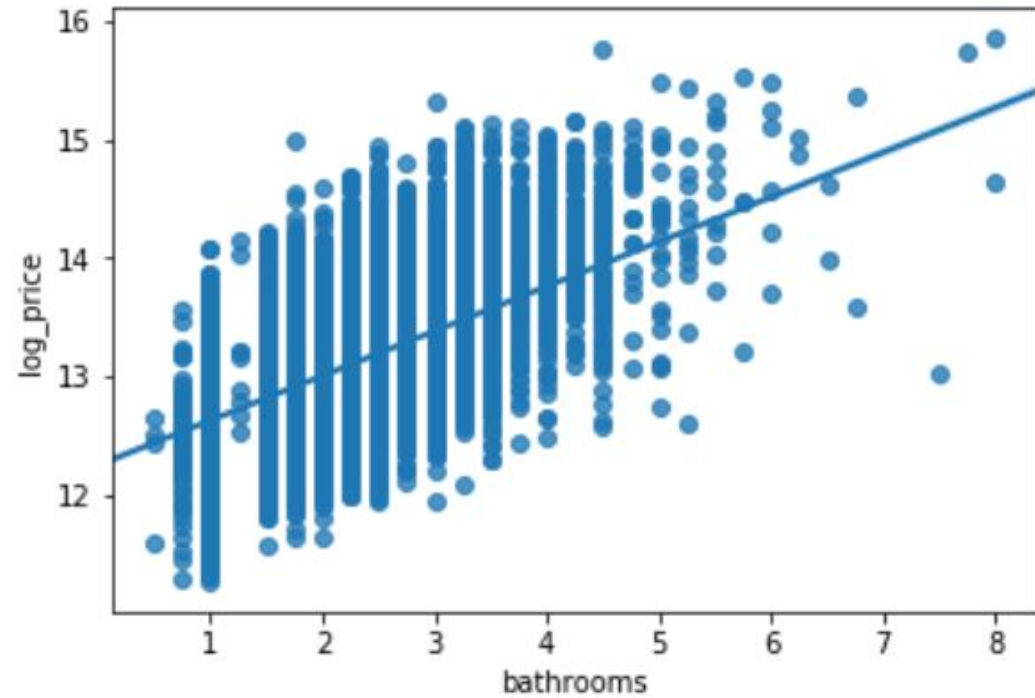
King County House Dataset

A Brief presentation for the analysis of the king count data set.

Feature Selection

- There where total 21 field in the dataset including price.
- Selecting feature based on the effect of the particular variable on the price.

Feature Selection



R Square value

Dep. Variable:	log_price	R-squared:	0.289
Model:	OLS	Adj. R-squared:	0.288
Method:	Least Squares	F-statistic:	1459.
Date:	Mon, 28 Jan 2019	Prob (F-statistic):	0.00
Time:	17:42:47	Log-Likelihood:	-13115.
No. Observations:	21595	AIC:	2.624e+04
Df Residuals:	21588	BIC:	2.630e+04
Df Model:	6		
Covariance Type:	nonrobust		

Train-test-Split

- Train Mean Squared Error:
 - 0.24919533363310273

- Test Mean Squared Error:
 - 0.25438937054929395

Current Model Accuracy

- Base on the Current test and train MES this model has currently 76% to 73% prediction accuracy.