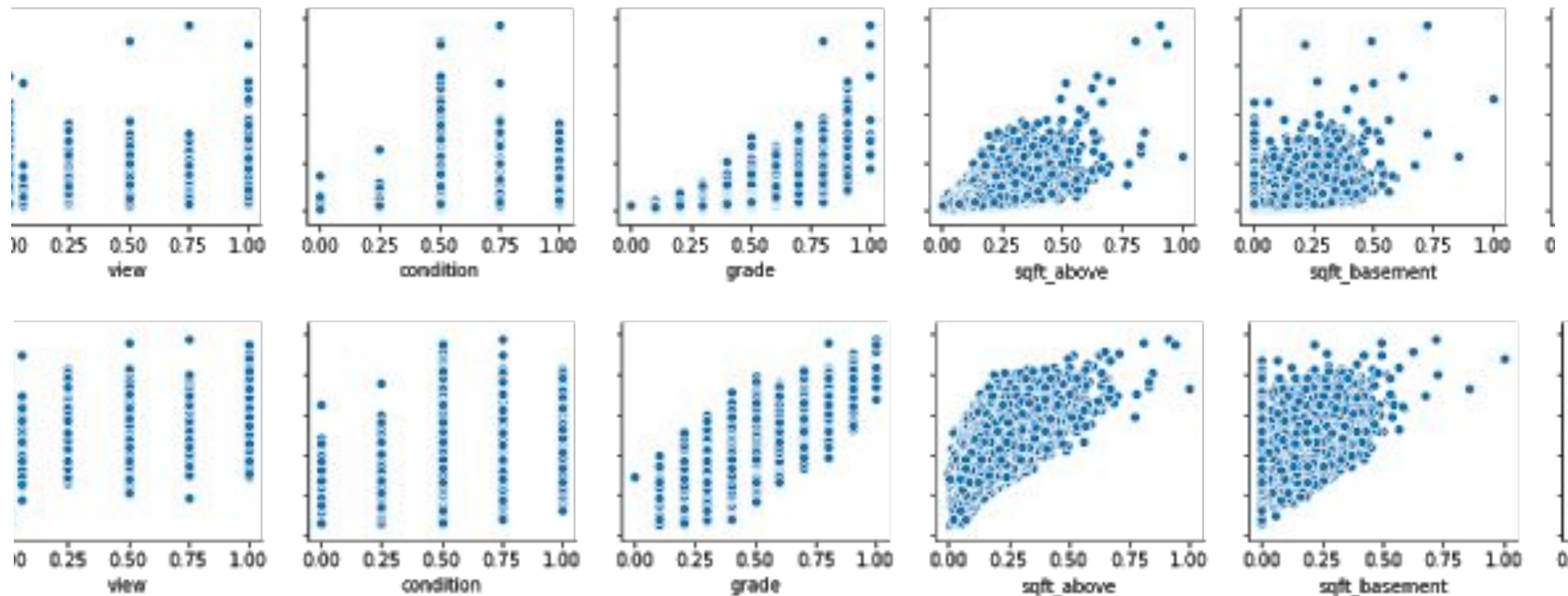


Predicting Housing Prices

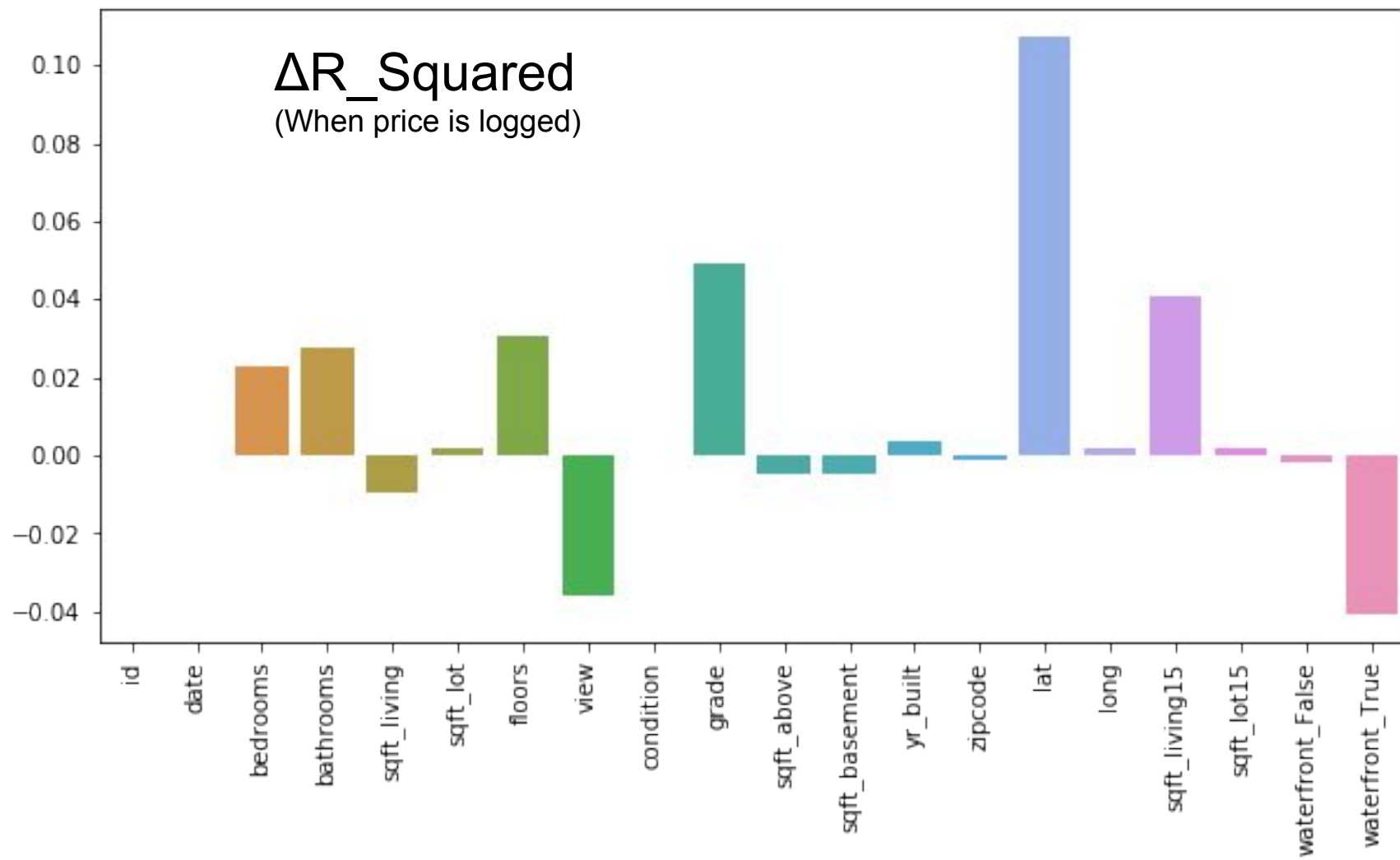
Using the Kings Count House Sales dataset

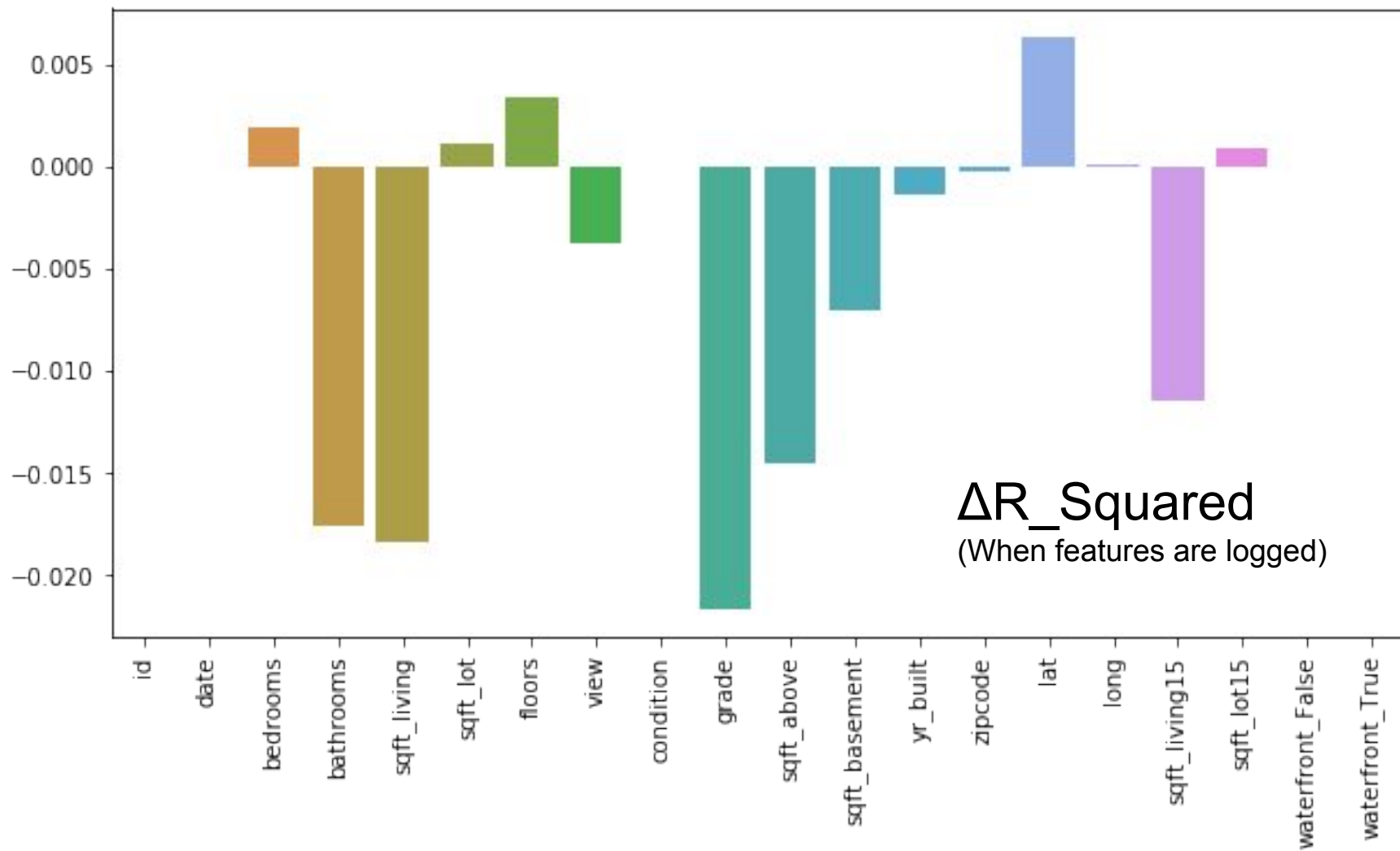
Price Logged



$\Delta R_Squared$

(When price is logged)





“Feature Engineering”

- Base^Feature
 - Price Logged/Not Logged
- Feature^Base
 - Price Logged/Not Logged
- For all features
- For all integer base between 0 and 100

```
[('yr_built', 0.02072550838163889, 6, 'feature**base'),  
 ('yr_built', 0.02072550838163889, 6, 'base**feature'),  
 ('zipcode', 0.007362707072450347, 4, 'feature**base'),  
 ('zipcode', 0.007362707072450347, 4, 'base**feature'),  
 ('sqft_basement', 0.004261847058947277, 2,  
  'feature**base'),  
 ('sqft_basement', 0.004261847058947277, 2,  
  'base**feature')]
```

```
('yr_built',  
0.02072550838163889,  
6,  
'feature**base')
```

Work for nothing...

OLS Regression Results

Dep. Variable:	price	R-squared:	0.771
Model:	OLS	Adj. R-squared:	0.771
Method:	Least Squares	F-statistic:	4548.
Date:	Wed, 08 May 2019	Prob (F-statistic):	0.00
Time:	11:59:29	Log-Likelihood:	-862.20
No. Observations:	21597	AIC:	1758.
Df Residuals:	21580	BIC:	1894.
Df Model:	16		
Covariance Type:	nonrobust		

Omnibus:	373.821	Durbin-Watson:	1.986
Prob(Omnibus):	0.000	Jarque-Bera (JB):	756.828
Skew:	-0.037	Prob(JB):	4.54e-165
Kurtosis:	3.914	Cond. No.	119.

	coef	std err	t	P> t	[0.025	0.975]
const	11.3711	0.016	730.526	0.000	11.341	11.402
date	0.0760	0.006	12.813	0.000	0.064	0.088
bedrooms	-0.4209	0.076	-5.531	0.000	-0.570	-0.272
bathrooms	0.5710	0.030	19.013	0.000	0.512	0.630
sqft_living	1.8571	0.055	33.617	0.000	1.749	1.965
sqft_lot	0.7475	0.099	7.552	0.000	0.553	0.941
floors	0.1805	0.010	17.889	0.000	0.161	0.200
view	0.2544	0.010	24.255	0.000	0.234	0.275
condition	0.2468	0.012	21.302	0.000	0.224	0.270
grade	1.6008	0.027	59.822	0.000	1.548	1.653
yr_built	-0.4205	0.010	-42.393	0.000	-0.440	-0.401
zipcode	-0.1292	0.008	-15.806	0.000	-0.145	-0.113
lat	0.8719	0.008	105.087	0.000	0.856	0.888
long	-0.1951	0.020	-9.999	0.000	-0.233	-0.157
sqft_living15	0.5509	0.025	22.205	0.000	0.502	0.600
sqft_lot15	-0.2293	0.080	-2.870	0.004	-0.386	-0.073
waterfront_True	0.3820	0.023	16.848	0.000	0.338	0.426