

Modeling the Prices of Homes

King County, 2014-2015

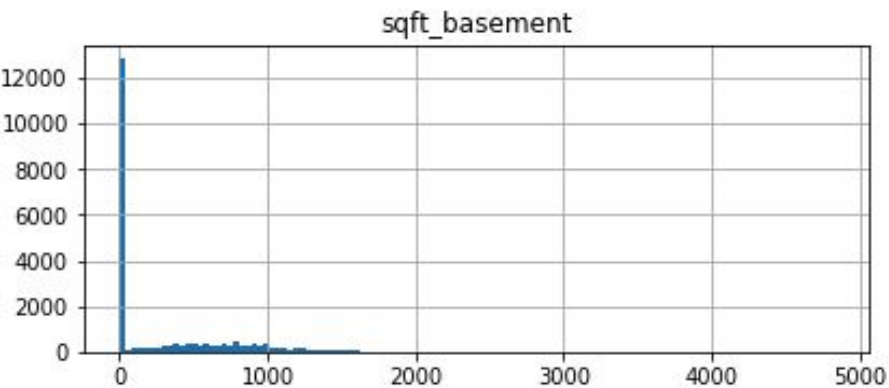


Business Value

Why should we model home prices?

KNOWLEDGE

PROFITABILITY



OLS Regression Results

Dep. Variable:	log_price	R-squared:	0.529
Model:	OLS	Adj. R-squared:	0.529
Method:	Least Squares	F-statistic:	5925.
Date:	Fri, 07 Feb 2020	Prob (F-statistic):	0.00
Time:	13:31:16	Log-Likelihood:	-21970.
No. Observations:	21082	AIC:	4.395e+04
Df Residuals:	21077	BIC:	4.399e+04
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	1.816e-14	0.005	3.84e-12	1.000	-0.009	0.009
bathrooms	0.1052	0.007	15.196	0.000	0.092	0.119
grade	0.4876	0.008	62.379	0.000	0.472	0.503
log_sqft_above	-0.0011	0.008	-0.135	0.893	-0.017	0.015
log_sqft_living15	0.2125	0.007	29.655	0.000	0.198	0.227

Omnibus:	87.780	Durbin-Watson:	1.970
Prob(Omnibus):	0.000	Jarque-Bera (JB):	87.632
Skew:	0.149	Prob(JB):	9.35e-20
Kurtosis:	2.895	Cond. No.	3.53

Business Recommendations

Apply Models

OLS Regression Results

Dep. Variable: log_price R-squared: 0.775

		coef	corr	tier
	waterfront_yes	0.476850	+	A
	grade	0.367568	+	A
	lat	0.363077	+	A
	waterfront_no	0.291357	-	B
	yr_built	0.209006	-	B
	const	0.185493	+	C
	renovated_yes	0.167186	+	C
	basement_yes	0.164793	+	C
	log_sqft_living	0.143362	+	C
	log_sqft_above	0.135246	+	C
	log_sqft_living15	0.131301	+	C

Prob

Skew:	Omnibus: 393.905	Durbin-Watson: 1.991
	Prob(Omnibus): 0.000	Jarque-Bera (JB): 756.051
Kurtosis:	Skew: 0.107	Prob(JB): 6.69e-165
	Kurtosis: 3.903	Cond. No. 3.02e+16

Business Recommendations

Mind the Season

Spring Sales

Number of homes sales: 5222

Average price: \$ 543266.4

Summer Sales

Number of homes sales: 6654

Average price: \$ 550992.05

Fall Sales

Number of homes sales: 5442

Average price: \$ 534689.17

Winter Sales

Number of homes sales: 3764

Average price: \$ 525097.84

Business

Zipcode 98039 has the highest average
Zipcode 98002 has the lowest average
Zipcode 98103 has the most homes sold
Zipcode 98039 has the most least sold

	price	homes_sold	target_metric
zipcode			
98042	311786.979630	540	0.001732
98023	283932.361570	484	0.001705
98038	366476.504378	571	0.001558
98133	386460.070393	483	0.001250
98058	352936.913832	441	0.001250
98001	281478.769231	351	0.001247
98118	418176.794769	497	0.001188
98168	241543.609195	261	0.001081
98106	319821.480243	329	0.001029
98155	424395.310345	435	0.001025

Future Work

Improve, Understand, and Apply

Experiment with models

More robust model?

Other areas and time periods

Understand Model

Thank You

Questions or Comments?

Contact Me!

John Nguyen

Email: jdnguyen91@gmail.com
