Bounty from King County

Aroa Gomez & Titus Bridgwood

Purpose of our analysis: New seller price estimation

The purpose of our study is to find a way to help new sellers with an estimate of their property value in King County.

For that we will perform a data study on properties sold between 2014 - 2015

We will try to help our potential new customers (sellers) to understand the conditions of the market and the potential value added of renovations and condition improvement.

Housing overview:

A quick look on our housing population

We are looking at a sample of over 21 thousand houses sold over the course of 2014-15.

What our average house looks like:

- 2,000 sqft
- 3 bedrooms
- 2 bathrooms

What our standard (average) house costs:

- USD 540,000

Is this information enough to help our sellers?

Can we help new sellers to estimate how much they could sell their house for?

Features that determine cost:

- Living space
- Grading
- Presence of a Basement
- Number of Floors
- Living space of the nearest 15 neighbors

Price Estimation Considerations

Our model tells us that these features can explain 60% of the variation in the price of the properties in the dataset.

New instances of these features would be enough to account for 60% of the estimation of a new house price.

The other 40% - tricky variables like neighborhood, nearby schools or even the amount of light and air a room gets.

Example Price Estimation

Input:

- Grade 7
- 2100 Square foot of living space
- No basement
- Nearest 15 neighbors have an average of 1640 square foot
- 2 floors

Base Estimation:

- ~$390,000

Findings and recommendations

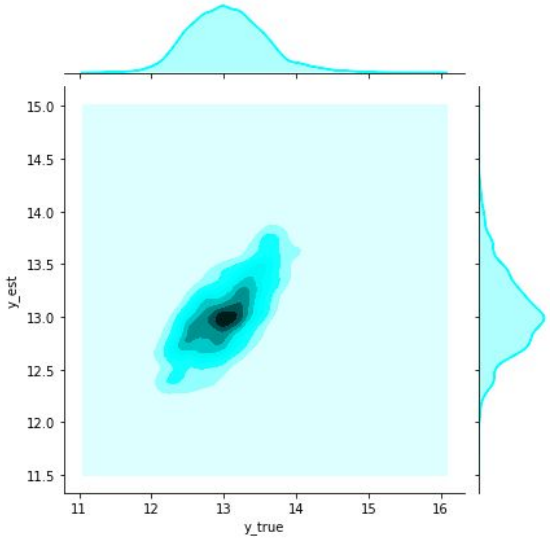
Should you make any changes?

- An increase of 10% in square footage results in a 2.5% increase in price.
- Each extra floor increases the average house price 3%.
- The presence of a basement in a house increases the average house price 15%
- On average an improvement in condition from grades 2 to 3 could result in a price increase of around $82 per square foot.
- An improvement from condition 3 to 4 does not appear to have any effect on house price, we therefore would not recommend the necessary renovations.
- How long ago a renovation occurred appears to make no difference to price.

Thank You

Q&A

Our Model



OLS Regression Results

Dep. Variable:	price_log	**R-squared:**	0.567
Model:	OLS	**Adj. R-squared:**	0.567
Method:	Least Squares	**F-statistic:**	5648.
Date:	Tue, 22 Oct 2019	**Prob (F-statistic):**	0.00
Time:	14:56:15	**Log-Likelihood:**	-7760.2
No. Observations:	21596	**AIC:**	1.553e+04
Df Residuals:	21590	**BIC:**	1.558e+04
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	7.8729	0.066	120.192	0.000	7.744	8.001
grade	0.1954	0.003	58.877	0.000	0.189	0.202
sqft_living_log	0.2546	0.010	24.929	0.000	0.235	0.275
sqft_basement_dummy	0.1502	0.005	27.446	0.000	0.139	0.161
sqft_living15_log	0.2189	0.012	18.916	0.000	0.196	0.242
floors	0.0323	0.005	6.102	0.000	0.022	0.043

Omnibus:	72.012	**Durbin-Watson:**	1.993
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	72.676
Skew:	0.141	**Prob(JB):**	1.65e-16
Kurtosis:	2.966	**Cond. No.**	372.

Log interpretation:

We will use `sqft_living` and `floors` to represent how the values would be calculated.

For a 10% increase in sqft of living, all other variables remaining equal (or unchanged) we would expect the price to increase by 2.5% where 1.1 is the 10% increase (1+0.1) and exp 0.25 is our beta result, then we extract 1 and multiply by 100 in order to get the percentage change [(((1.1^0.25)-1)*100]

Our floor variable, implies that one unit increase (i.e one extra floor) will translate in average into a 3% price increase (exp(0.0323)) where 0.0323 is the beta value

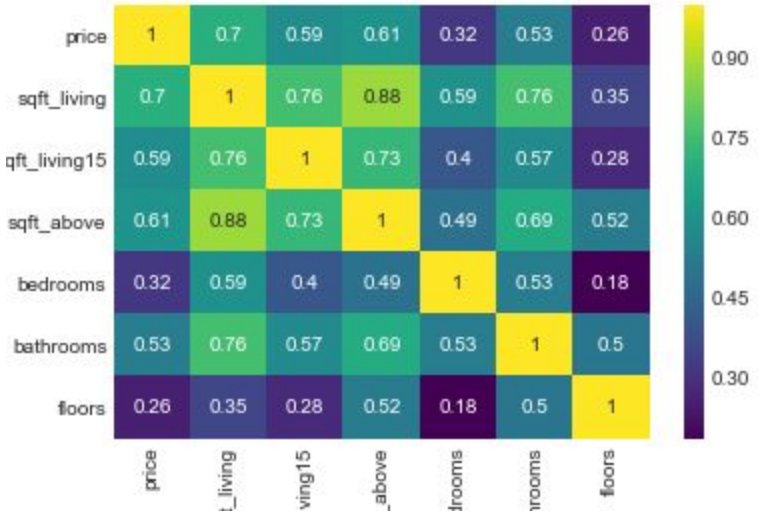
Some Math:

$\log(\text{price}(n2)) - \log(\text{price}(n1)) = \beta * [\log(\text{sqft2}) - \log(\text{sqft1})]$

$\log(\text{price}(n2)/\log(\text{price}(n1) = \beta * [\log(\text{sqft2}/\text{sqft1})]$

$\text{price}(n2)/\text{price}(n1) = (\text{sqft2}/\text{sqft1})^{\beta}$

Looking for correlations

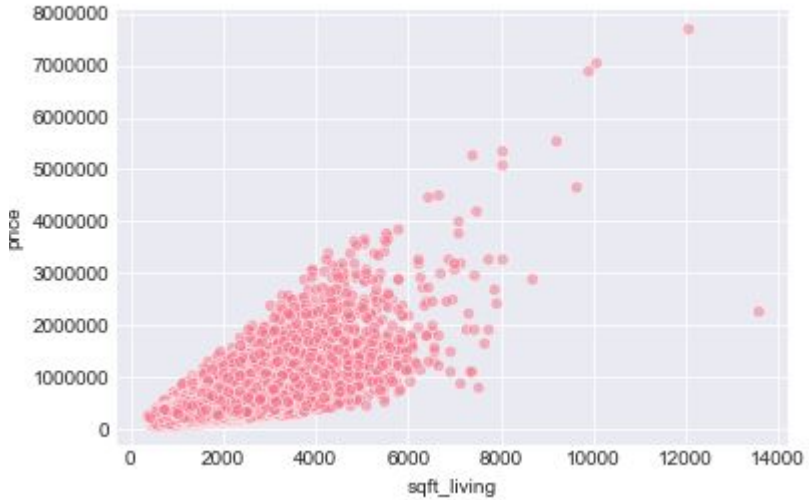


Can living space predict price?

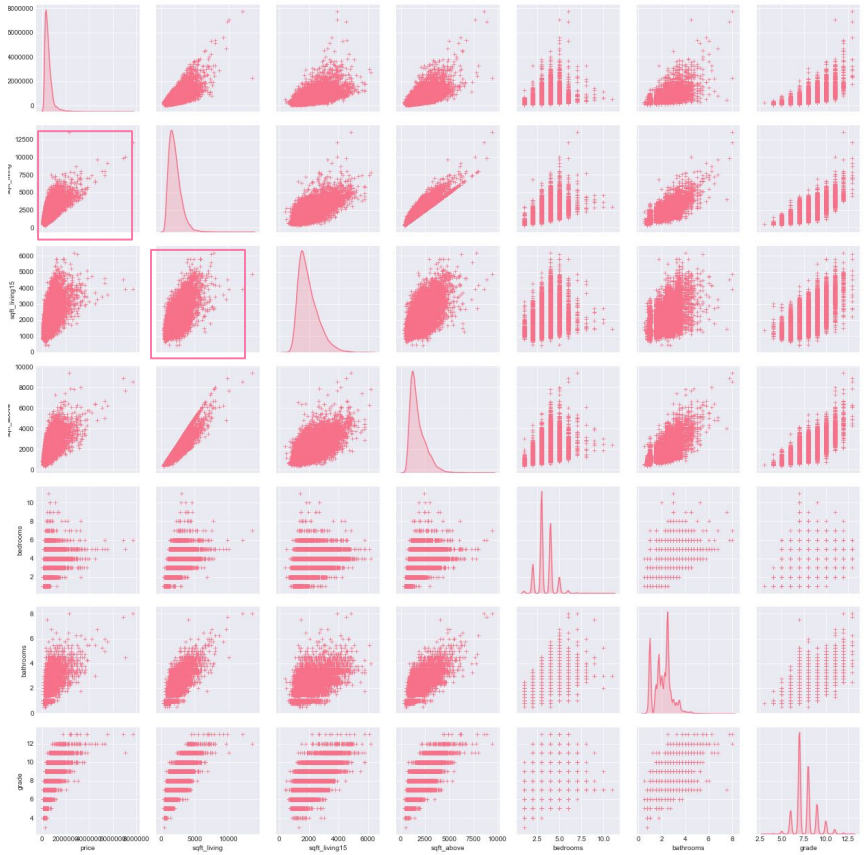
- There is a large standard variation for square footage.

  STD = 918.0ft2

- Square footage has a strong correlation with price.



Avoiding Multicollinearity



The condition jump