



Real estate market analysis at King County with Regression modeling.

Presented by JX Consultant.inc

Why buying a house instead of rent?

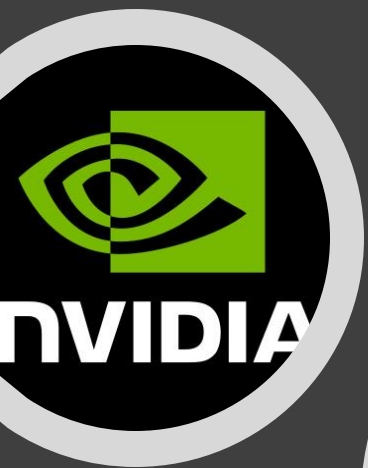
1. With mortgage, monthly payment is similar to the rent fee, especially at place like King County has high price-to-rent ratio .


2. Better living condition than apartment. (Safe neighbour, more space and less interference)

3. Wisely real estate investment will pay back in the future.



World's leading technology
companies' site at Seattle





What features should future
owner consider most?

Top 3 house features to be considered

- Bigger living room
- Waterfront
- South area of King County

OLS Regression Results

Dep. Variable:	price	R-squared:	0.645
Model:	OLS	Adj. R-squared:	0.644
Method:	Least Squares	F-statistic:	2608.
Date:	Sun, 07 Nov 2021	Prob (F-statistic):	0.00
Time:	23:44:52	Log-Likelihood:	-2.3713e+05
No. Observations:	17275	AIC:	4.743e+05
Df Residuals:	17262	BIC:	4.744e+05
Df Model:	12		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	7.273e+05	8123.363	89.534	0.000	7.11e+05	7.43e+05
bedrooms	-5.507e+04	2355.343	-23.381	0.000	-5.97e+04	-5.05e+04
sqft_living	3.008e+05	2688.626	111.895	0.000	2.96e+05	3.06e+05
sqft_lot	4710.1073	2442.934	1.928	0.054	-78.291	9498.506
floors	1.638e+04	2247.447	7.290	0.000	1.2e+04	2.08e+04
waterfront	6.778e+04	1709.595	39.649	0.000	6.44e+04	7.11e+04
condition	2.254e+04	1861.698	12.105	0.000	1.89e+04	2.62e+04
sqft_basement	-1.772e+04	2233.688	-7.935	0.000	-2.21e+04	-1.33e+04
yr_built	-4.09e+04	2336.677	-17.503	0.000	-4.55e+04	-3.63e+04
yr_renovated	1.495e+04	1757.013	8.509	0.000	1.15e+04	1.84e+04
lat	8.91e+04	1756.213	50.733	0.000	8.57e+04	9.25e+04
long	-2.594e+04	1997.640	-12.984	0.000	-2.99e+04	-2.2e+04
sqft_lot15	-1.002e+04	2473.966	-4.050	0.000	-1.49e+04	-5171.427

Omnibus:	12202.438	Durbin-Watson:	1.993
Prob(Omnibus):	0.000	Jarque-Bera (JB):	610660.952
Skew:	2.840	Prob(JB):	0.00
Kurtosis:	31.568	Cond. No.	18.4

What feature should owner care less?

- 1. Bedrooms
- 2. yr_built
- 3. sqft_basement

OLS Regression Results

Dep. Variable:	price	R-squared:	0.645
Model:	OLS	Adj. R-squared:	0.644
Method:	Least Squares	F-statistic:	2608.
Date:	Sun, 07 Nov 2021	Prob (F-statistic):	0.00
Time:	23:44:52	Log-Likelihood:	-2.3713e+05
No. Observations:	17275	AIC:	4.743e+05
Df Residuals:	17262	BIC:	4.744e+05
Df Model:	12		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	7.273e+05	8123.363	89.534	0.000	7.11e+05	7.43e+05
bedrooms	-5.507e+04	2355.343	-23.381	0.000	-5.97e+04	-5.05e+04
sqft_living	3.008e+05	2688.626	111.895	0.000	2.96e+05	3.06e+05
sqft_lot	4710.1073	2442.934	1.928	0.054	-78.291	9498.506
floors	1.638e+04	2247.447	7.290	0.000	1.2e+04	2.08e+04
waterfront	6.778e+04	1709.595	39.649	0.000	6.44e+04	7.11e+04
condition	2.254e+04	1861.698	12.105	0.000	1.89e+04	2.62e+04
sqft_basement	-1.772e+04	2233.688	-7.935	0.000	-2.21e+04	-1.33e+04
yr_built	-4.09e+04	2336.677	-17.503	0.000	-4.55e+04	-3.63e+04
yr_renovated	1.495e+04	1757.013	8.509	0.000	1.15e+04	1.84e+04
lat	8.91e+04	1756.213	50.733	0.000	8.57e+04	9.25e+04
long	-2.594e+04	1997.640	-12.984	0.000	-2.99e+04	-2.2e+04
sqft_lot15	-1.002e+04	2473.966	-4.050	0.000	-1.49e+04	-5171.427

Omnibus:	12202.438	Durbin-Watson:	1.993
-----------------	-----------	-----------------------	-------



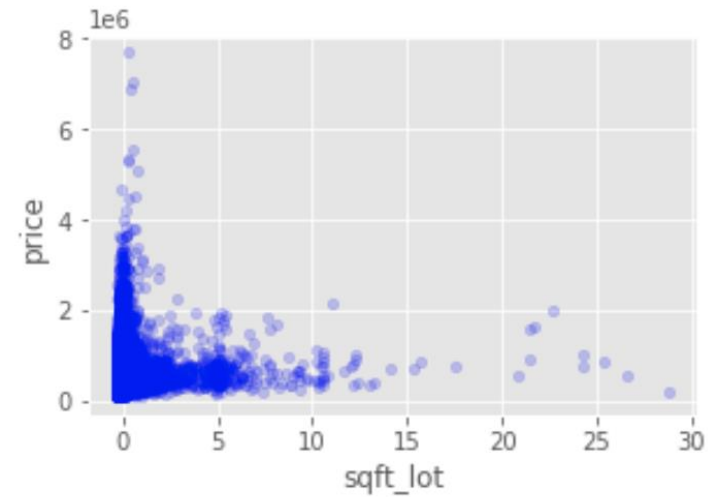
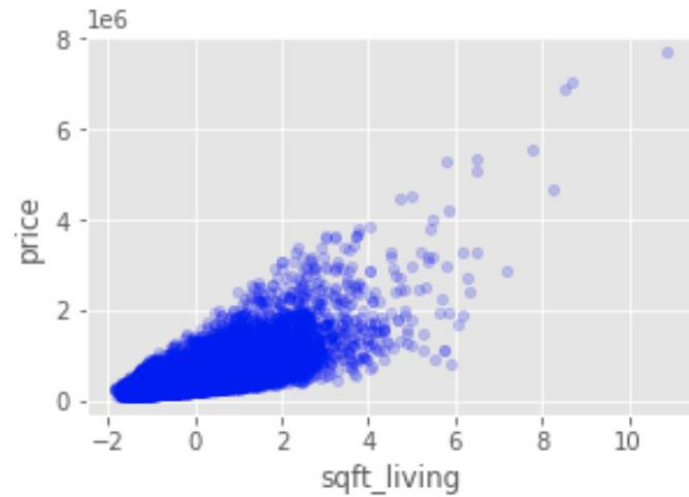
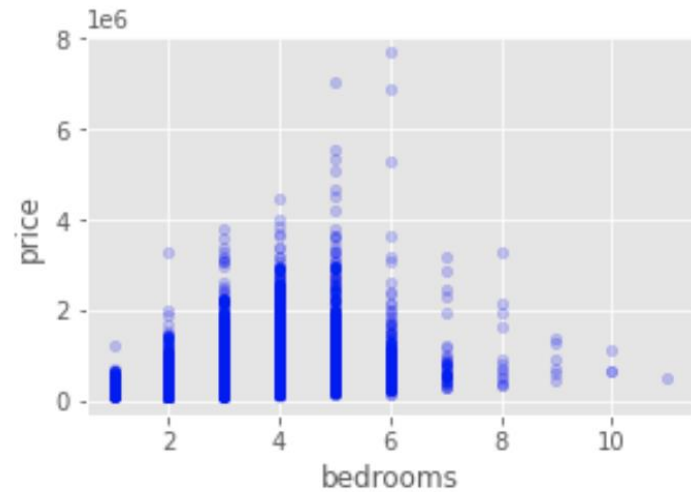
Analysis coefficient among those negative features

According to the `model.summary()`, bedrooms, square feet of basement and year of build are the top 3 negative features.

Reason :

1. House with too many bedrooms are not as hot as house with bedrooms less than 5.
2. Seattle is a place less likely to snow but rich of rain. Basement is damp and dark. Besides, basement like such contains more Radon gas, and Radon is a naturally-occurring radioactive gas that can cause lung cancer.
3. New build house are usually far from waterfront, downtown, coastline and school, those area usually been occupied with older houses.

Should owner buy a house with more rooms than he/she actually needs?



Purchasing strategy

- From the histogram and model summary, we conclude that less bedrooms, more living room sqft and more lot sqft would increase the price. If actual bedrooms need is more than 5, more bedrooms usually get a cheaper house. We speculate this may be because those big houses are less popular and hot than bedrooms number less than 5.

Conclusion



- After analysis, my team believe this is the best time getting into valuable house.
- We would suggest :
 - 1. Bigger living room, waterfront, and south area of King County should be considered.
 - 2. For King county, downtown and US west coast is more experience. Don't buy house with basement, and older house usually has more appealing to the buyer.
 - 3. Bedrooms no more than 5 are more popular and hot to trade.

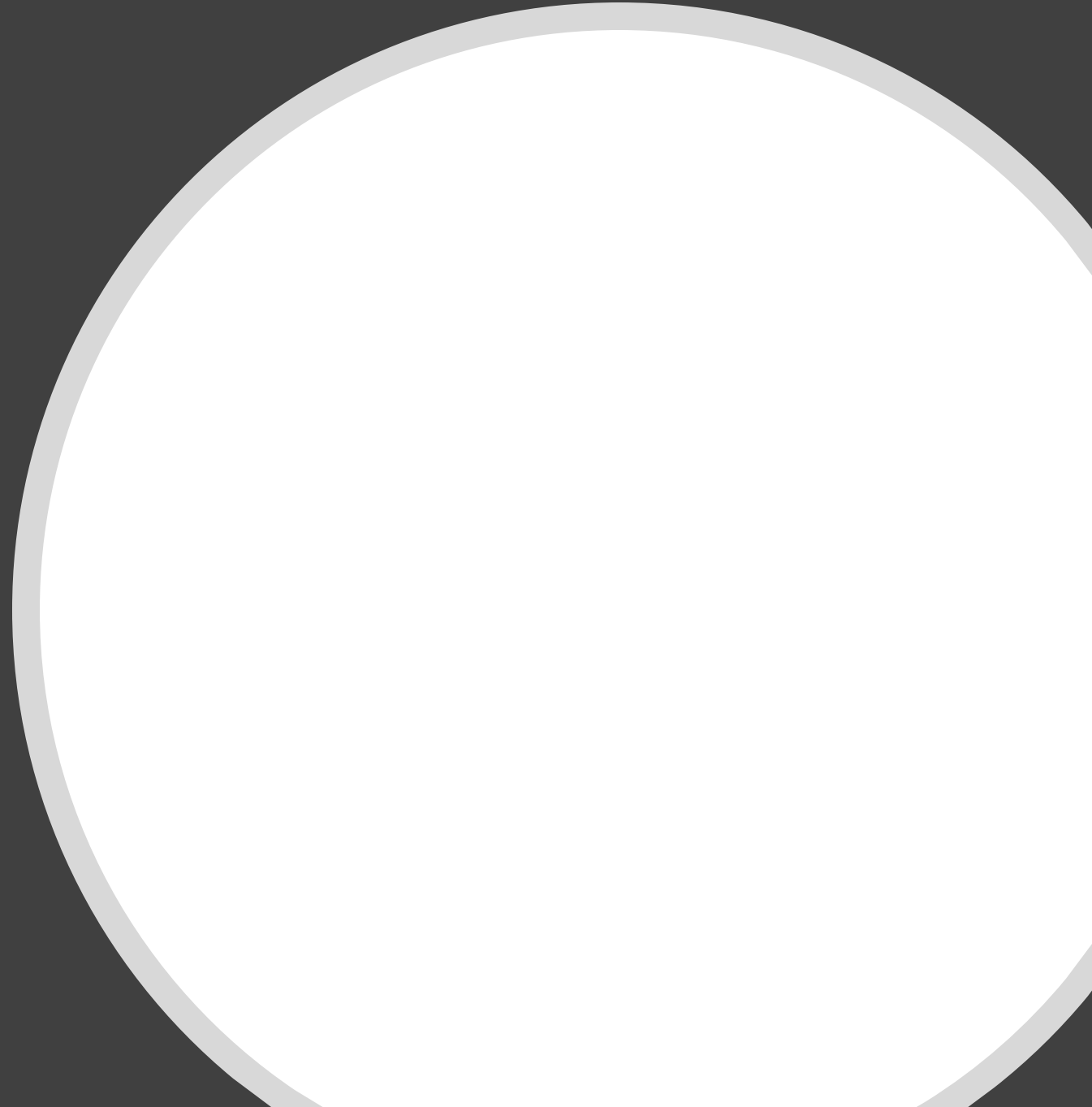
Future Work



- 1. With more time, I would like to look into 'zip' column see if I can find more valuable information.
- 2. For house sale in different months, I want to check if sale price and deal made amount change accordingly. I may need to search for more sales record in different year to get a convincing trend for that.
- 3. There are more features might influence price like world economic environment, Nasdaq Index, pandemic, etc. I want to merge more features to get a better model for the price prediction.



Thank you!



Q&A

- Contact:
- Xu66366@gmail.com
- <https://github.com/Johnxuujj>
- [linkedin.com/in/jiajie-xu-805b84180](https://www.linkedin.com/in/jiajie-xu-805b84180)