

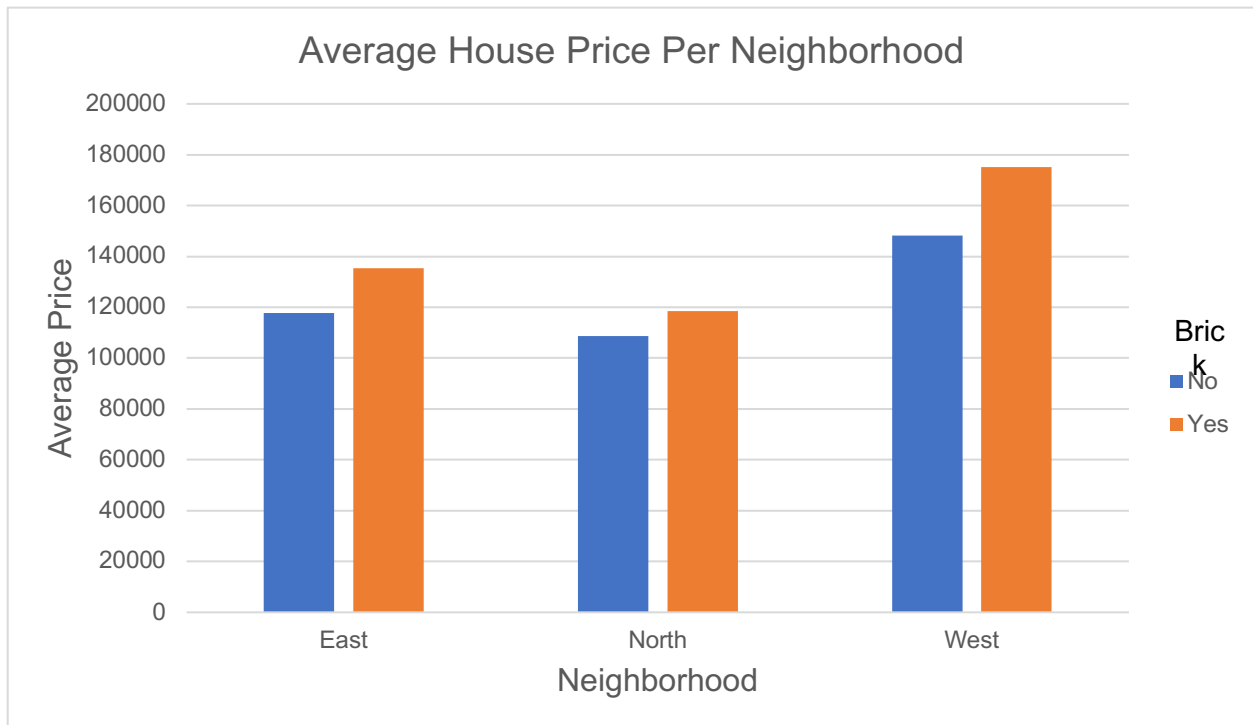
SCM 651
Homework 1

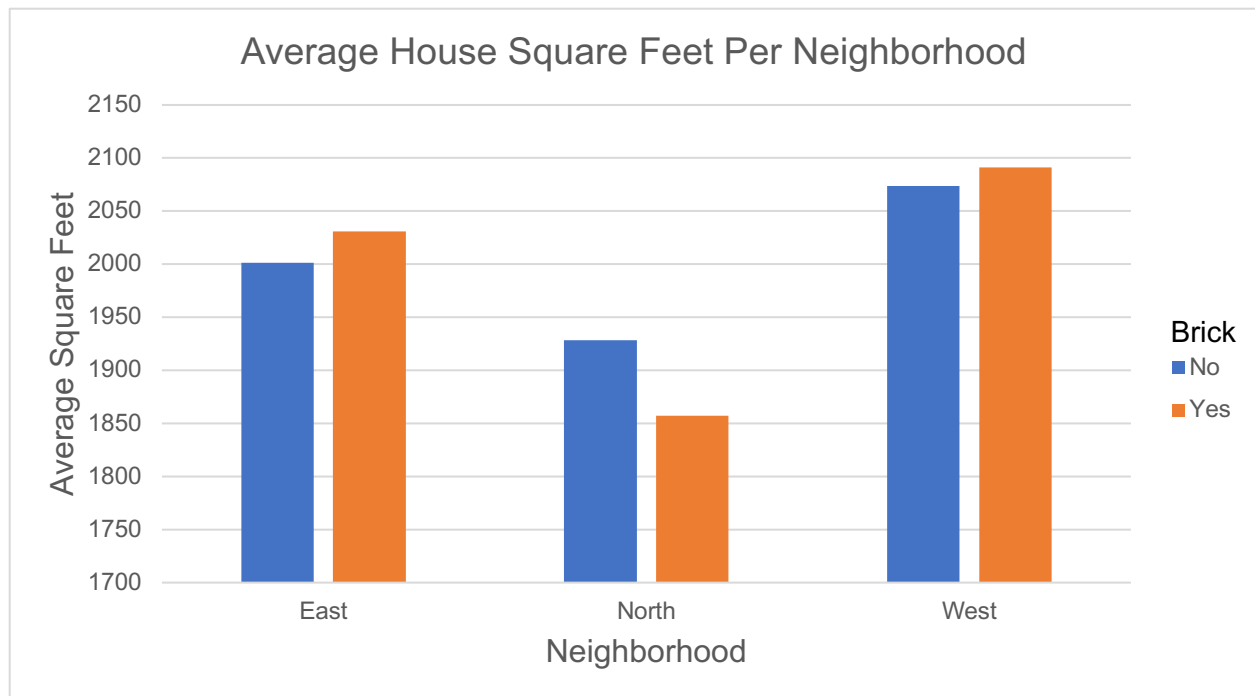
#1 – Pivot Tables (in Excel document)

Average of Price			
Column Labels			
Row Labels	No	Yes	Grand Total
East	117750	135468.4211	125231.1111
North	108583.7838	118457.1429	110154.5455
West	148230.4348	175200	159294.8718
Grand Total	121958.1395	147769.0476	130427.3438

Average of SqFt			
Column Labels			
Row Labels	No	Yes	Grand Total
East	2001.538462	2031.052632	2014
North	1928.108108	1857.142857	1916.818182
West	2073.478261	2091.25	2080.769231
Grand Total	1989.186047	2025	2000.9375

#2 – Pivot Charts (in Excel document)





#3 – Correlation Analysis

Which two variables have the largest magnitude correlation?

SqFt and Price.

Which two variables have the smallest magnitude correlation?

Offers and Bedrooms.

What does the largest magnitude imply if we perform a regression analysis next?

SqFt and Price are related and affect each other. If SqFt goes up, Price goes up, and if SqFt goes down, Price goes down. You would expect to see a significant P value in the regression for SqFt.

Are there any negative correlations?

Offers and Price.

Are these correlations intuitive? If not, why not?

Yes. Normally, the bigger the house, the higher the price. Additionally, the more bedrooms and bathrooms, the larger the square feet of the house, and therefore a higher price. You would expect to see large houses with many bedrooms and bathrooms that have a higher list price in the normal housing market.

Lastly, the number of offers on a house would not necessarily be affected by the price. Many other factors go into making an offer on a house other than just price. So, price and offers having a low and negative correlation makes sense, as well as lower correlation to the other data elements in the provided data set.

#4 – Regression Analysis

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.835573066							
R Square	0.698182349							
Adjusted R Square	0.688367141							
Standard Error	14999.24552							
Observations	128							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	4	64012998276	16003249569	71.13270927	4.43749E-31			
Residual	123	27672216021	224977366					
Total	127	91685214297						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-17347.37695	12724.89631	-1.363262736	0.175289936	-42535.52896	7840.775063	-42535.52896	7840.775063
SqFt	61.8399461	8.263773843	7.48325732	1.20211E-11	45.48231247	78.19757973	45.48231247	78.19757973
Bedrooms	9319.752602	2148.75444	4.33728137	2.97311E-05	5066.424938	13573.08027	5066.424938	13573.08027
Bathrooms	12646.34749	3109.662029	4.066791622	8.44849E-05	6490.962169	18801.7328	6490.962169	18801.7328
Offers	-13601.01141	1324.818659	-10.26631934	3.08843E-18	-16223.40872	-10978.61411	-16223.40872	-10978.61411
Formula		=61.84*(SqFt)+9319.75*(Bedrooms)+12646.35*(Bathrooms)-13601.01*(Offers)-17347.38						

Which variables are statistically significant?

SqFt, Bedrooms, Bathrooms, and Offers

What does each coefficient mean in a real-world sense?

The dollar amount for one unit of the variable. The fixed costs are -\$17,437.38 for essentially just the land. A house with no square footage, no bedrooms, and no bathrooms is an empty plot of land. The coefficient for SqFt is 61.84, meaning \$61.84 per square foot. The coefficient for Bedrooms is 9319.75, meaning \$9,319.75 per bedroom. The coefficient for Bathrooms is 12646.35, meaning \$12,646.35 per bathroom. Lastly, the coefficient for offers is -13601.01, meaning -\$13,601.01 per offer.

Are these coefficients intuitive? If not, why not?

The coefficients for fixed costs, SqFt, Bedrooms, and Bathrooms seem intuitive. There are still expenses for the land even if no house is built, such as loan payments and taxes, so a negative

correlation makes sense for fixed costs. SqFt is at \$61.84, which makes sense when multiplied against the square footage of a house. Bedrooms is at \$9,319.75, which seems to make sense when valuing the cost of each bedroom in a home (more or less adding \$9k per bedroom). Bathrooms is at \$12,646.35, which makes sense because more bathrooms make a house more desirable and bathrooms are expensive to build!

The coefficient for offers does not seem intuitive. Offers is at -\$13,601.01, which would mean the more offers a house has, the more it lowers the price. This does not make sense, since bidding wars often drive up the price of a house.

What does the R-squared mean?

70% of the variability of a house price is explained by the square feet, number of bedrooms, number of bathrooms, and number of offers.

#5 – Prediction Model

	\$	101,462.55	600	1200	1800	2400	3000	3600	4200	4800	5400	6000	6600
	1		59407.56	96511.56	133615.56	170719.56	207823.56	244927.56	282031.56	319135.56	356239.56	393343.56	430447.56
	2		45806.55	82910.55	120014.55	157118.55	194222.55	231326.55	268430.55	305534.55	342638.55	379742.55	416846.55
	3		32205.54	69309.54	106413.54	143517.54	180621.54	217725.54	254829.54	291933.54	329037.54	366141.54	403245.54
# of offers	4		18604.53	55708.53	92812.53	129916.53	167020.53	204124.53	241228.53	278332.53	315436.53	352540.53	389644.53
	5		5003.52	42107.52	79211.52	116315.52	153419.52	190523.52	227627.52	264731.52	301835.52	338939.52	376043.52
	6		-8597.49	28506.51	65610.51	102714.51	139818.51	176922.51	214026.51	251130.51	288234.51	325338.51	362442.51
	7		-22198.5	14905.5	52009.5	89113.5	126217.5	163321.5	200425.5	237529.5	274633.5	311737.5	348841.5
	8		-35799.51	1304.49	38408.49	75512.49	112616.49	149720.49	186824.49	223928.49	261032.49	298136.49	335240.49

	\$	101,462.55	600	1200	1800	2400	3000	3600	4200	4800	5400	6000	6600
	1		27167.05	64271.05	101375.05	138479.05	175583.05	212687.05	249791.05	286895.05	323999.05	361103.05	398207.05
	2		36486.8	73590.8	110694.8	147798.8	184902.8	222006.8	259110.8	296214.8	333318.8	370422.8	407526.8
# bedrooms	3		45806.55	82910.55	120014.55	157118.55	194222.55	231326.55	268430.55	305534.55	342638.55	379742.55	416846.55
	4		55126.3	92230.3	129334.3	166438.3	203542.3	240646.3	277750.3	314854.3	351958.3	389062.3	426166.3
	5		64446.05	101550.05	138654.05	175758.05	212862.05	249966.05	287070.05	324174.05	361278.05	398382.05	435486.05

#6 – Non-intuitive Results

What would explain non-intuitive results in your regression using the data which you were provided?

One thing that might explain the negative correlation between Price and Offers would be offers over a long period of time. Normally, multiple offers at once when a house is first listed drives up the prices with a bidding war. However, if a house is on the market for a long time and has had multiple offers fall through, that normally brings the price of the house down.

What additional data would assist you in explaining the non-intuitive results?

How many days the house was listed on the market before it sold could help explain the Price/Offer negative correlation, as well as how many offers fell through over that period of time.