

SCM 651 Business Analytics

Team 4: Seth J Rissmiller, Frederick B Lamarca III, and Maria Ng

HW#1: House Price Analysis

1. Pivot Table

Generate pivot tables to showcase average house price and square feet by type of construction and neighborhood. Type of construction indicates no or yes to house build of brick. Type of neighborhood consists of East, North, and West.

Table 1: Average of Price by Type of Construction and Neighborhood

| Row Labels | Average of Price |
|--------------------|---------------------|
| ▣ East | \$125,231.11 |
| No | \$117,750.00 |
| Yes | \$135,468.42 |
| ▣ North | \$110,154.55 |
| No | \$108,583.78 |
| Yes | \$118,457.14 |
| ▣ West | \$159,294.87 |
| No | \$148,230.43 |
| Yes | \$175,200.00 |
| Grand Total | \$130,427.34 |

Table 2: Average of Square Feet by Type of Construction and Neighborhood

| Row Labels | Average of SqFt |
|--------------------|-----------------|
| ▣ East | 2014.00 |
| No | 2001.54 |
| Yes | 2031.05 |
| ▣ North | 1916.82 |
| No | 1928.11 |
| Yes | 1857.14 |
| ▣ West | 2080.77 |
| No | 2073.48 |
| Yes | 2091.25 |
| Grand Total | 2000.94 |

2. Pivot Chart

Develop two pivot charts (Figures 1 and 2) to display average house price and square feet by construction in brick and neighborhood.

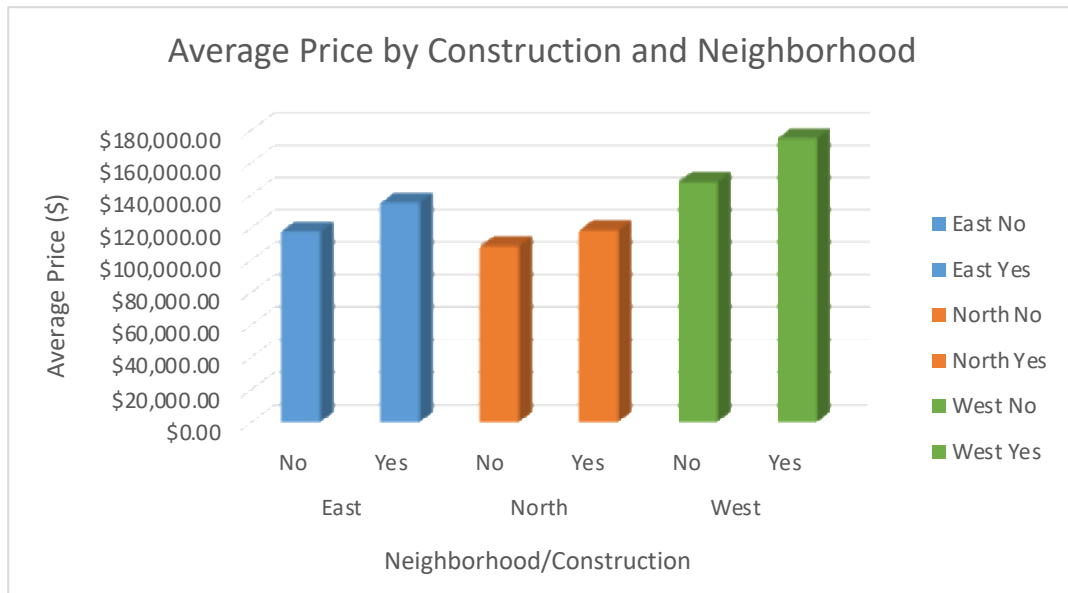


Figure 1: Average Price by Construction in Brick and Neighborhood

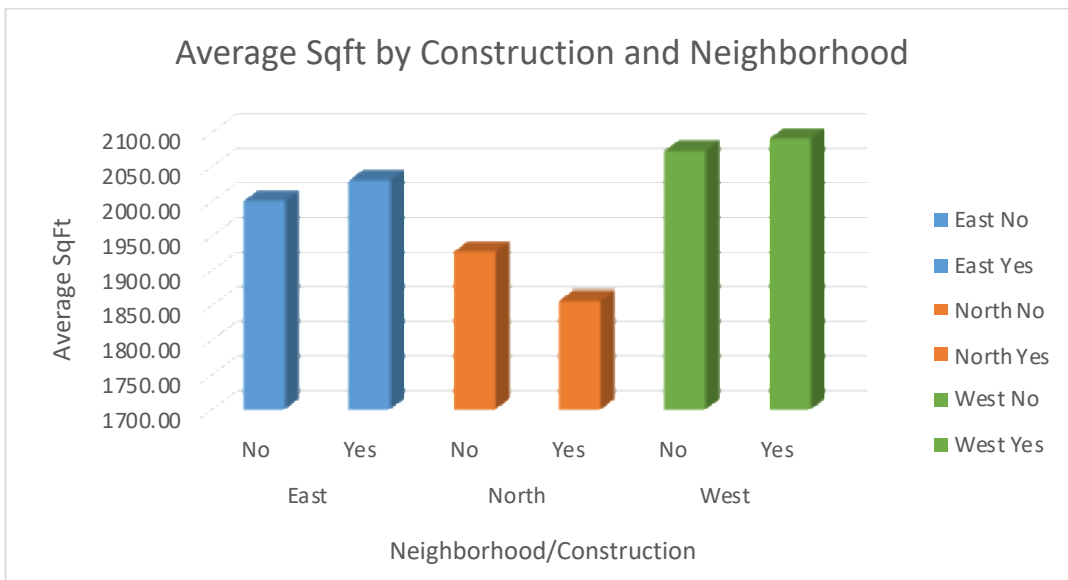


Figure 2: Average Square Feet by Construction in Brick and Neighborhood

3. Correlation Analysis

A correlation analysis performed on price, square feet, bedroom, bathrooms, and offers to determine the magnitude and direction of the correlation. Two variables with the largest magnitude correlation are square feet and house prices. Two variables with the smallest magnitude correlation are bedrooms and offers. If we were to plot out regression, we would expect square feet and price to have a strong correlation and be statistically significant. There is a negative correlation, which is price and offers. The correlation is intuitive, which implies the more the square feet, the higher the price. The higher the price, the less likely to have any offers.

Table 3: Correlation Analysis

| | Price | SqFt | Bedrooms | Bathrooms | Offers |
|-----------|----------|----------|----------|-----------|--------|
| Price | 1 | | | | |
| SqFt | 0.552982 | 1 | | | |
| Bedrooms | 0.525926 | 0.483807 | 1 | | |
| Bathrooms | 0.523258 | 0.522745 | 0.414556 | 1 | |
| Offers | -0.31364 | 0.336923 | 0.114271 | 0.143793 | 1 |

4. Regression Analysis

In a real word sense each coefficient is the relative unit cost for respective variable. The coefficients for Sqft, Bedrooms, and Bathrooms are intuitive. Example, 1 Sqft is equal to \$61.84 and 1 Bedroom is equal to \$9,320. The coefficients for offers is not intuitive because this value is negative. It would not make sense to have a negative unit cost associated for each offer. R-square means that ~70% of data is explained in model.

SUMMARY OUTPUT

| Regression Statistics | |
|-----------------------|------------|
| Multiple R | 0.83557307 |
| R Square | 0.69818235 |
| Adjusted R Squ | 0.68836714 |
| Standard Error | 14999.2455 |
| Observations | 128 |

ANOVA

| | df | SS | MS | F | Significance F |
|------------|-----|-------------|-------------|-----------|----------------|
| Regression | 4 | 64012998276 | 16003249569 | 71.132709 | 4.4375E-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 | -17347.3769 | 12724.89631 | -1.363262736 | 1.75E-01 | -42535.529 | 7840.7751 | -42535.53 | 7840.7751 |
| SqFt | 61.8399461 | 8.263773843 | 7.48325732 | 1.20E-11 | 45.4823125 | 78.19758 | 45.48231 | 78.19758 |
| Bedrooms | 9319.7526 | 2148.75444 | 4.33728137 | 2.97E-05 | 5066.42494 | 13573.08 | 5066.425 | 13573.08 |
| Bathrooms | 12646.3475 | 3109.662029 | 4.066791622 | 8.45E-05 | 6490.96217 | 18801.733 | 6490.962 | 18801.733 |
| Offers | -13601.0114 | 1324.818659 | -10.26631934 | 3.09E-18 | -16223.4087 | -10978.61 | -16223.41 | -10978.61 |

Figure 3: Regression analysis for house price.

5. Two-Way Sensitivity Analysis

Table 4: Dynamic table for regression analysis equation.

| | Coefficients | Input | Coefficients * Input |
|-----------|--------------|-------|----------------------|
| Intercept | -17347.3769 | 1 | -17347.37695 |
| SqFt | 61.8399461 | 2000 | 123679.8922 |
| Bedrooms | 9319.7526 | 3 | 27959.25781 |
| Bathrooms | 12646.3475 | 3 | 37939.04246 |
| Offers | -13601.0114 | 5 | -68005.05706 |
| | | | \$ 104,225.76 |

Table 5: Two-Way Sensitivity Analysis with conditional formatting for SqFt and Bedrooms.

| | | SqFt | | | | | | | | | | |
|----------|---|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------|
| | | \$104,226 | 1500 | 1600 | 1700 | 1800 | 1900 | 2000 | 2100 | 2200 | 2300 | 2400 |
| Bedrooms | 2 | \$ 63,986 | \$ 70,170 | \$ 76,354 | \$ 82,538 | \$ 88,722 | \$ 94,906 | \$ 101,090 | \$ 107,274 | \$ 113,458 | \$ 119,642 | |
| | 3 | \$ 73,306 | \$ 79,490 | \$ 85,674 | \$ 91,858 | \$ 98,042 | \$ 104,226 | \$ 110,410 | \$ 116,594 | \$ 122,778 | \$ 128,962 | |
| | 4 | \$ 82,626 | \$ 88,810 | \$ 94,994 | \$ 101,178 | \$ 107,362 | \$ 113,546 | \$ 119,730 | \$ 125,914 | \$ 132,097 | \$ 138,281 | |
| | 5 | \$ 91,945 | \$ 98,129 | \$ 104,313 | \$ 110,497 | \$ 116,681 | \$ 122,865 | \$ 129,049 | \$ 135,233 | \$ 141,417 | \$ 147,601 | |

Green = Lower Price, **Red** = Higher Price

6. There is a negative correlation (indirectly proportional) between offers and price. This implies as price increase the number of offers decreases. Our regression indicates a negative cost would be associated with more offers. This contradicts the correlation analysis because it implies there will be a negative cost when there are more offers and regression analysis reflects there will be more offers when price is low. Even though P-values are lower than alpha for all variables, implying they are statistically significant the x variables of offers is not a true driving factor for price. It is more of a by-product of price. For example, if we were trying to sell a home for a higher price, we could positively influence the price by adding more sqft, bedrooms, and bathrooms. The same methodology cannot be used for offers. Using a two-way sensitizing analysis for sqft and offers one can see as sqft increases and offers stays constant, price increases. This is a great example of sqft being one of the true driving factor of price and offers not impacting the price.

Table 6: Supporting Two-Way Sensitivity Analysis with conditional formatting for Sqft and Offers.

| | | SqFt | | | | | | | | | | |
|-------|---|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | | \$ 104,226 | 1500 | 1600 | 1700 | 1800 | 1900 | 2000 | 2100 | 2200 | 2300 | 2400 |
| Offer | 1 | | \$127,710 | \$133,894 | \$140,078 | \$146,262 | \$152,446 | \$158,630 | \$164,814 | \$170,998 | \$177,182 | \$183,366 |
| | 2 | | \$114,109 | \$120,293 | \$126,477 | \$132,661 | \$138,845 | \$145,029 | \$151,213 | \$157,397 | \$163,581 | \$169,765 |
| | 3 | | \$100,508 | \$106,692 | \$112,876 | \$119,060 | \$125,244 | \$131,428 | \$137,612 | \$143,796 | \$149,980 | \$156,164 |
| | 4 | | \$86,907 | \$93,091 | \$99,275 | \$105,459 | \$111,643 | \$117,827 | \$124,011 | \$130,195 | \$136,379 | \$142,563 |
| | 5 | | \$73,306 | \$79,490 | \$85,674 | \$91,858 | \$98,042 | \$104,226 | \$110,410 | \$116,594 | \$122,778 | \$128,962 |
| | 6 | | \$59,705 | \$65,889 | \$72,073 | \$78,257 | \$84,441 | \$90,625 | \$96,809 | \$102,993 | \$109,177 | \$115,361 |