

Name : House Price Analysis
Current Date : 04/21/2020
Homework : 01
Team : SCM Project Group1
Team Members : Alexandra A Rosser, Nicholas Starr, Jared M Mosley & Sathish Kumar Rajendiran

2020-0401 SCM 651 Business Analytics



Table of Contents

| | |
|--|----------|
| Executive Summary | 3 |
| What's Due | 3 |
| Tasks | 3 |
| Pivot Table & Pivot Charts..... | 4 |
| Correlation Analysis | 6 |
| Regression Analysis | 7 |
| Prediction Model using Spread Sheet & 2 -way Sensitive Analysis | 8 |
| Non-Intuitive Results | 9 |
| Conclusion | 9 |

Executive Summary

Using the House Prices data, determine the factors which influence the price of a home.

What's Due

Submit a categorization, visualization, correlation, and regression analysis of house prices.

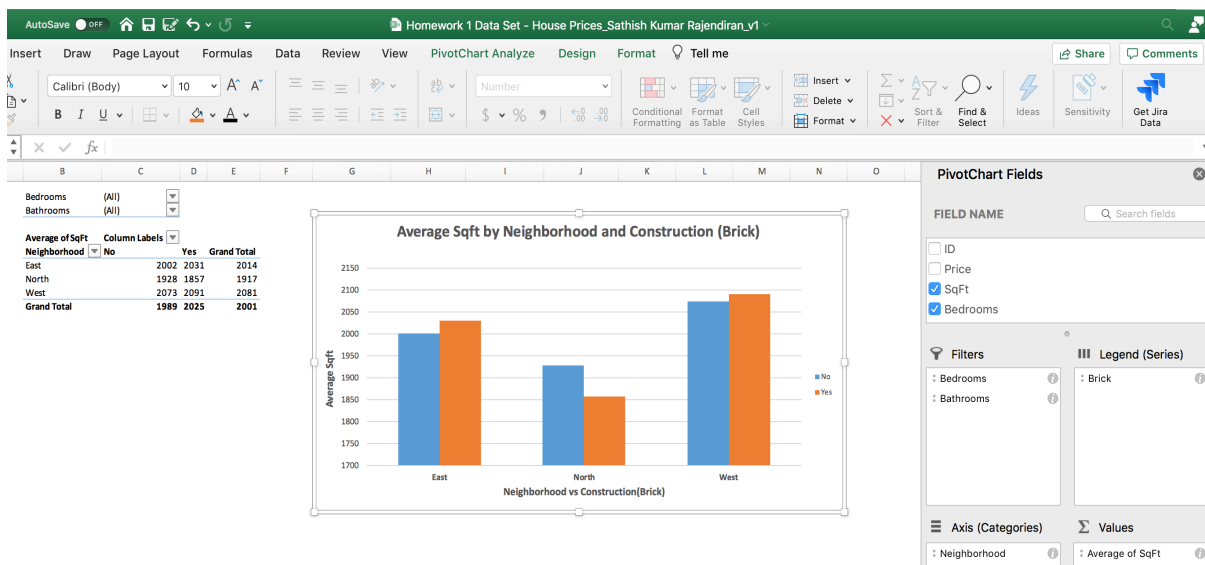
Tasks

This section talks about the overall research tasks to be performed,

- 1 Pivot Tables
 - Develop a categorization of your data using pivot tables. Develop two pivot tables of average price and average square feet by type of construction (brick) and neighborhood
- 2 Pivot Charts
 - Using the two pivot tables above, generate pivot charts for average price and average square feet by type of construction (brick) and neighborhood
- 3 Correlation Analysis
 - Perform a correlation analysis of all quantitative variables except ID. Which two variables have the largest magnitude correlation?
 - Which two variables have the smallest magnitude correlation?
 - What does the largest magnitude imply if we perform a regression analysis next? Are there any negative correlations? Are these correlations intuitive? If not, why not?
- 4 Regression Analysis
 - Perform an initial regression analysis of the quantitative variables excluding the ID. Which variables are statistically significant? What does each coefficient mean in a real-world sense? Are these coefficients intuitive? If not, why not? What does the R- squared mean?
- 5 Prediction Model using Spread sheet
 - Create a spreadsheet prediction of the model. Perform a two-way sensitivity analysis and use conditional formatting to highlight the results.
- 6 Non-Intuitive results
 - What would explain non-intuitive results in your regression using the data which you were provided? What additional data would assist you in explaining the non-intuitive results?

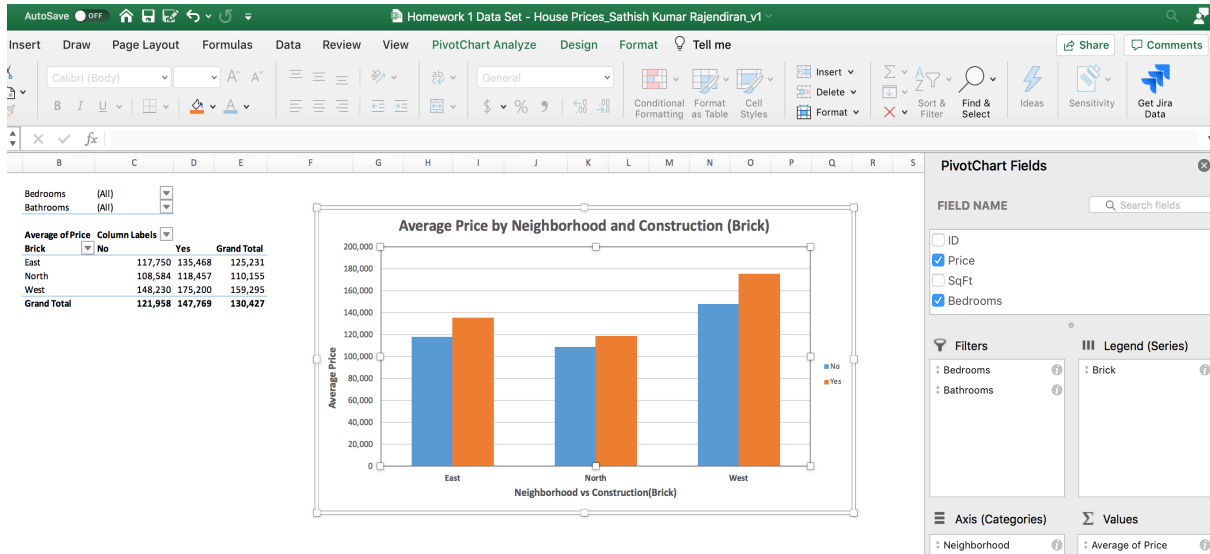
Pivot Table & Pivot Charts

- Average Sqft by Neighborhood and Construction material (Brick)



- Above picture shows pivot table & chart analysis on average sqft by neighborhood and building material (Brick)
- Pivot table contains
 - Column value having construction material (Brick) value as “Yes” or “No”
 - Row value having neighborhood type (East, North & West)
 - Measure as average Sqft value
 - Optionally number of bedrooms and bathrooms are added to filters
- Pivot Chart contains
 - X- Axis having neighborhood vs construction material (Bricks)
 - Legend color Orange → construction material used as Brick
 - Legend color Blue → construction material used non-Brick
 - Y-Axis having average Sqft

- Average Price by Neighborhood and Construction material (Brick)



- Above picture shows pivot table & chart analysis on average price by neighborhood and building material (Brick)
- Pivot Table contains
 - Column value having Construction material (Brick) value as “Yes” or “No”
 - Row value having Neighborhood type (East, North & West)
 - Measure as average price value
 - Optionally number of bedrooms and bathrooms are added to filters
- Pivot Chart contains
 - X- Axis having neighborhood vs construction material (Bricks)
 - Legend color Orange → construction material used as Brick
 - Legend color Blue → construction material used non-Brick
 - Y-Axis having average price
- Average price by neighborhood and construction material (Brick)

Correlation Analysis

| | <i>Price</i> | <i>SqFt</i> | <i>Bedrooms</i> | <i>Bathrooms</i> | <i>Offers</i> |
|-----------|--------------|-------------|-----------------|------------------|---------------|
| Price | 1 | | | | |
| SqFt | 0.55298224 | 1 | | | |
| Bedrooms | 0.52592606 | 0.48380711 | 1 | | |
| Bathrooms | 0.52325776 | 0.5227453 | 0.41455596 | 1 | |
| Offers | -0.3136359 | 0.33692335 | 0.11427061 | 0.1437934 | 1 |

Based on the data sample, picture above shows the correlation analysis among the variables.

The highest magnitude correlation is between square footage and price, followed closely by bedrooms and price (with bathrooms and price not far behind). Bedrooms and offers have the smallest magnitude of correlation, followed by bathrooms and offers. There is one negative correlation between offers and price. We find the price to square footage correlations to be very intuitive as well as price to bedrooms and bathroom. It also makes sense that square footage has a moderately strong correlation to bedrooms and bathrooms. Lastly, it is intuitive that as price decreases, number of offers increase, and vice-versa.

Regression Analysis

AutoSave

OFF

Homework 1 Data Set - House Prices_Sathish Kumar Rajendiran_v1

Insert

Draw

Page Layout

Formulas

Data

Review

View

Tell me

Calibri (Body)

11

A^A

ab

General

Based on the data sample, picture above shows the regression analysis among the variables.

- Level of significance shows the model is 99.99% statistically significant.
- P-value also suggests all variable have the strongest confidence that level of error is lower than 0.05
- Coefficient values suggests predicted values will be 61.84 times Sqft, 9319.75 times bedroom quantity, 12646.35 times bathroom quantity with fixed cost (intercept) -17347.34
- Adjusted R-Square shows the output value is 68.84% affected by the variables. Yet, 32% deviates from having perfect model. Suggests there is potential improvement needed.

[illegible]

- Formula for Prediction goes as

$$\text{Prediction Price} = \text{Intercept} + \text{Sqft} * \text{Coeff. Sqft} + \text{Bedrooms} * \text{Coeff. Bedrooms} + \text{Bathrooms} * \text{Coeff. Bathrooms} + \text{offers} * \text{Coeff. Offers}$$
- 2-way analysis is done between Sqft and number of bedrooms.
- Conditional formatting is applied with graded color scale.

Non-Intuitive Results

As referenced in question 4, offers may initially seem non-intuitive, but actually represents a home being undervalued. When the amount of offers increase, the price decreases, implying that there is higher demand at that price level. We believe the intercept to be non-intuitive. If all else were equal to 0, you essentially are left with a lot/property and the regression implies this is worth -\$17347.

Additional data could include the neighborhood located in, construction type, the school district scores, the level of crime in the area, the level of growth in the area, age of the home, tax rates, and more.

Conclusion

Overall, house price analysis exercise helped us explore pivot table/chart, data analysis tools (Correlation, regression analysis), conditional formatting & 2-way data analysis excel features along with prediction, non-intuitive results and statistical modeling concepts.