House Price Analysis

Background

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

Resources

Use the dataset SCM 651 Homework 1 House Prices spreadsheet.

Assignment

What's due:

Submit a categorization, visualization, correlation, and regression analysis of house prices **before the live class in week 4**. Suggested length is five pages, but should not exceed ten pages, single-spaced, 12-point font.

This is a group assignment; each student should upload a copy of the assignment to the Learning Management System. The paper must be a Microsoft Word document. You should also submit the Excel spreadsheet with the visualization, correlation, and regression analysis. Name the file HW1_Team# where # is your team number. Be sure to include the names of everyone on the team on the first page of the paper. Late assignments will not be accepted. Failure to follow directions will be penalized.

Outline and grading criteria:

1. Develop a categorization of your data using pivot tables. Develop two pivot tables: One pivot table of average price, varying type of construction (brick) and neighborhood as the two dimensions; a second pivot table of average square feet varying type of construction (brick) and neighborhood as the two dimensions (20%)
2. Using the two pivot tables above, generate pivot charts for average price and average square feet by type of construction (brick) and neighborhood (10%)
3. 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? (20%)
4. 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? (25%)
5. Create a spreadsheet prediction of the model. Perform a two-way sensitivity analysis and use conditional formatting to highlight the results. (15%)
6. 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? (10%)

Justify your answers. Provide a snapshot of output from your analysis in your final paper.

House Price Data

ID	unique identifier
Price	price of home in dollars
SqFt	square feet of area of home
Bedrooms	number of bedrooms
Bathrooms	number of bathrooms
Offers	number of offers received on home before sale
Brick	Yes/No on brick construction
Neighborhood	location of home in east, west or north quadrants of city