**Using Multiple Regression to Analyze and Predict Home Prices**

An important problem in real estate is determining how to price homes to be sold. There are so many factors—size, age, and style of the home; number of bedrooms and bathrooms; size of the lot; and so on—which makes setting a price a challenging task. In this project, we will try to help realtors in this task by determining how different characteristics of homes relate to home prices, identifying the key variables in pricing, and building multiple-variable regression models to predict prices based on property characteristics. Our analysis will be based on the Mount Pleasant Real Estate Data (available on AsULearn and [www.hawkeslearning.com](http://www.hawkeslearning.com)). This data set includes information about 245 properties for sale in three communities in the suburban town of Mount Pleasant, South Carolina, in 2017.

**Phase 1: Data Preparation.**

1. Download the Mount Pleasant Real Estate Data and open it with Microsoft Excel.

2. Consider the following variables associated with each property.

Dependent variable:

y: home price.

Independent variables:

x1: number of bedrooms

x2: total number of bathrooms

x3: number of stories

x4: square footage

x5: has fenced yard?

x6: has screened porch?

3. For the qualitative variables x5 and x6, adjust the data such that we have a binary (0/1) indicator. For this, you may have to use an Excel command such as: IF(Z2="Yes",1,0). Here Z2 is an example of a cell containing “Yes” or “No” which is converted to a 0 or 1, respectively. For instance, Z2 could be the second observation of the variable Z.

**Phase 2: Analysis.**

Please enable the Analysis ToolPak add-in under Excel and use the regression tool.

4. Perform a multiple regression analysis such that the predicted home price is obtained from the following equation:

$\hat{y}=b\_{0}+b\_{1}x\_{1}+b\_{2}x\_{2}+b\_{3}x\_{3}+b\_{4}x\_{4}+b\_{5}x\_{5}+b\_{6}x\_{6}.$

Please paste or attach your Excel output here.

5. What is the coefficient of determination, R2, of the regression model? Explain the meaning of this value.

6. Are any variables statistically insignificant predictors of home price at a significance level of α = 0.05? State the p-values of these variables.

7. What is the predicted price for a 2000 square feet 2-story house with 3 bedrooms, 2.5 baths, and a fenced yard, but no screened porch?

8. Apart from the explanatory variables listed above, what are two other factors that could affect home prices? Please note that the factors could be anything and not necessarily in the Mount Pleasant data.