ITEC 620

Fall 2022

**Assignment 3**

Due November 17, 5:15 PM

Please submit your assignment via Canvas, **as a single Word file**.

**For all questions that involve generating R output, you must include both the output itself and the command or sequence of commands used to generate it to receive credit. Include ONLY the commands used to generate that output; do not include entire scripts.**

A real estate agency collects some basic data on recent home sales, which can be found in the “HomePrices” csv file. The variables are:

*Size* – the size of the house, in square feet  
*Age* – the age of the house, in years  
*Foreclosure* – an indicator variable equal to 1 if the house is a foreclosure, and 0 otherwise  
*Crime* – the violent crime rate of the area, in # per year per 100,000 residents  
*Price* – the sale price of the house

Using either Excel or R, run a regression model with *Price* as the dependent variable, using all four of the independent variables provided.

**a)** How much of the variation in Sale Price can be explained by this regression model?

**b)** What does the coefficient on *Size* tell us? (Note that *Size* is measured in square feet.)

**c)** Using this model, predict the price of a house with the following characteristics:  
-1500 square feet  
-50 years old  
-Not a foreclosure  
-300 violent crime rate

**d)** Which of the independent variables are clearly statistically significant?

**e)** Run a regression model using only the independent variables you stated in part d. Does the prediction from part c change? If so, how?