ASSIGNMENT 11 - LINEAR REGRESSION IN R

IS 605 FUNDAMENTALS OF COMPUTATIONAL MATHEMATICS - 2014

Using R's *lm* function, perform regression analysis and measure the significance of the independent variables for the following two data sets. In the first case, you are evaluating the statement that we hear that Maximum Heart Rate of a person is related to their age by the following equation:

$$MaxHR = 220 - Age$$

You have been given the following sample:

Age	18	23	25	35	65	54	34	56	72	19	23	42	18	39	37
MaxHR	202	186	187	180	156	169	174	172	153	199	193	174	198	183	178

Perform a linear regression analysis fitting the Max Heart Rate to Age using the *lm* function in R. What is the resulting equation? Is the effect of Age on Max HR significant? What is the significance level? Please also plot the fitted relationship between Max HR and Age.

Using the *Auto* data set from Assignment 5 (also attached here) perform a Linear Regression analysis using *mpg* as the dependent variable and the other 4 (*displacement, horsepower, weight, acceleration*) as independent variables. What is the final linear regression fit equation? Which of the 4 independent variables have a significant impact on *mpg*? What are their corresponding significance levels? What are the standard errors on each of the coefficients? Please perform this experiment in two ways. First take any random 40 data points from the entire auto data sample and perform the linear regression fit and measure the 95% confidence intervals. Then, take the entire data set (all 392 points) and perform linear regression and measure the 95% confidence intervals. Please report the resulting fit equation, their significance values and confidence intervals for each of the two runs.

Please submit an R-markdown file documenting your experiments. Your submission should include the final linear fits, and their corresponding significance levels. In addition, you should clearly state what you concluded from looking at the fit and their significance levels.