Multilinear regression assignment explanation

If Rand > 0.3, Test

First regression assumption check:

Normality: To check normality, check the histogram of the dependent variable (Y)

Linearity: To check linearity if there is a linear relationship between y and each x variable.

How to check it: Scatter graphs, if you draw a scatter graph y and x1, if you can fit a line angled downward or upward then it is linear. If it is horizontal, then there is no linear relationship. If it is curve, say y =X12 take X11/2

Correlation: If the correlation coefficient is more than 0.6 with 2 or more variables, we will eliminate that independent variable.

Stepwise approach:

If the p-value e

Once you have final inputs, standardize the training data set. Create regression using standardized values of final inputs. Regression from standardized final data gives variable importance.

How to standardize: Divide each number by the maximum number of the column.

If there is any pattern in the standardized residuals, then the dataset is incorrect/not suitable.

Test sample: remaining 30 % data.

We will use the formula created using final inputs before the standardization gives Y^