**Assignment Problems: Day 6**

The following problems require access to the file named ‘Dataset\_Day6.csv’, provided with this assignment. (**Use Random State = 50**)

Problems to solve –

1. Treat the data to replace all missing data with median/ mode (whichever applicable), and remove all rows with outliers. If data volume is reducing by more than 30% then only remove for price.
2. Use One Hot Encoding to encode all character variables.
3. Split the data into 80% training and 20% testing data. Then, create a multiple linear regression model with target variable as ‘Price’.
   1. Print the model performance metrics. R2, adjusted R2, MAE
4. Repeat the above process for a Ridge Regression and show from the new evaluation metrics if there is any improvement in the model performance?
5. Also do above for Lasso Regression (research code online)