Case Study 1:

Dylan Scott

In this assignment I was tasked with investigating critical temp of a list of chemical compounds. The goal would be to build a model that could predict the critical temperature of any new material presented to the model given its features provided. The data was provided in the form of two sheets. ‘unique\_m’ was a one hot encoded sheet noting all the elements that could be found within the compound. ‘train’ listed all the other features for the compound. After some initial investigation I found it best to not use ‘unique\_m’ and solely focus on the features provided in ‘train’.

Upon investigating the data I found that there was no null vales. In ‘train’ there was \_\_\_\_ number of values with \_\_\_\_ number of features. No categorical data found outside the ‘unique\_m’ document. As expected I went ahead and isolated my target variable to ‘crit temp’ and removed it from data set. Next, I standard scaled the data to help with normalizing my features.