Prediction Analysis

I began by performing different prediction techniques using a lot of the variables in the data to predict the sum of duration in risk positions for the cervical axial. These techniques of predicting or clustering the data included multiple regression, PCA, and logistic regression analysis. Overall, it was clear that the multiple regression and PCA analysis didn’t perform very well for the data. It is hard for a few variables to have good prediction capability for the sum of duration in risk positions because there are many factors that go into that. Furthermore, it may be more important to look at the different things that occur while people are at high-risk ranges of motion for extended periods of time like we are currently doing.

Some focused variables that I came up with for the predictions were attendings, non-attendings, type of case at higher risk thresholds, and case length at higher risk thresholds. The multiple regression showcased a low R2 value at around 0.3 and the PCA analysis showcased minimal clustering that could be found in the data. Finally, the logistic regression showcased an accuracy of around 0.70. I performed this using it as predicting above or below the median based on the predictors. I am unsure that any of that analysis will be helpful in this case. But, here are the logistic regression findings.