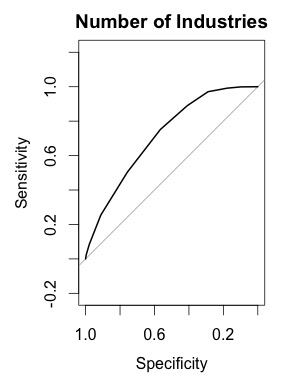
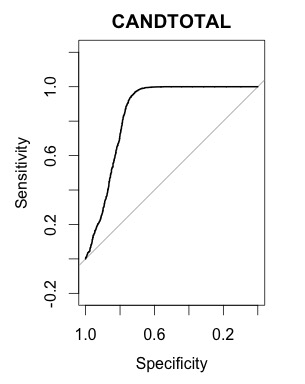
The first Null hypothesis was tested was: the is no difference between the total contributions that an incumbent gets and one that a challenger gets. We used our Merged Data set without outliers. We preformed a student t-test (not pairwise) to test this null hypothesis. We got a p-value of < 2.2e-16 and given that that this a social science analysis, the threshold should be .05. The p-value crosses this threshold and is well in the rejection region, so the p-value of extremely significant. So we reject the null hypothesis in favor of the alternative which is that there is a difference. Given the mean of the two categories of contributions, it is clear that the incumbent has a higher amount of the contributions compared to the challenger. This is also verifiable when compared to the association rules.

The Second Null hypothesis that tested involved a logistic regression model. The idea behind the model is to predict who the winner will be based on the Total amount of money raised and the Number of supporting industries. From the confusion matrix, we know the following. The Precision of the model is 0.8125881, the recall is 0.8662994, is F-measure is 0.8385846. Below are the ROC plots of both variables.



From the ROC curved, it seems like the logistic model predicts many false positives. This could signify that the model maybe be over fitting. This is probably because the model is predicting the values based upon itself.