**Question 16.1**

# Using the GermanCredit data set germancredit.txt from <http://archive.ics.uci.edu/ml/machine-learning-databases/statlog/german> / (description at <http://archive.ics.uci.edu/ml/datasets/Statlog+%28German+Credit+Data%29> ), use the *xgboost* gradient boosting algorithm to find a good classifier. Use R’s *xgboost* function and set the objective to “binary:logistic.”

* Unlike *glm*, *xgboost* will not create categorical columns for you. One way to approach this is to use *dummy\_cols* from the *fastDummies* package, which will take all your categorical columns and one-hot-encode them.
* *xgboost* requires the response variable to be 0s and 1s. The data set uses 1s and 2s instead, so after reading in the data, your R code will need to make that change.

1. Compare your xgboost classifier to your logistic regression classifier from Question 10.3. How does model performance vary over key metrics (accuracy, sensitivity, specificity, etc.)?