BUAN 6356 (Johnston)
Homework 4B (20220925)
Due: 2 October 2022 (11:59PM)

This part of the assignment is about building logistic regression models, calculating predicted values, classifying outcomes, and summarizing the classification results. Both testing and training subsets of the "wk" data from HW 4A will be used.

The initial ctree() model is required. Two glm() models are to be considered:

- default~student+balance+income (Common Slopes)

- default~student*(balance+income) (Individual Group Slopes)

You will be directed to report the model summary information for each model as well as summary information about both the training and testing data subsets. You will then need to classify the observations based on predictions and report on the accuracy of these classifications. You also may be asked about risk multipliers for individual coefficients from either model.

You will want to reference 01f_classification_titanic for examples of the type of classification summary results you may be asked about.

You may submit this assignment as many times as needed until you get full credit.

Note: "risk multiplers" are obtained from individual covariate (Independent variables, predictors) coefficient values (estimates) glm() logit models (family=binomial) through *exp*(<estimate>). These show the multiplicative effect of a unit increase of an individual variable. Thus, when the covariate is an indicator the "risk multiplier" show the effect of the indicated situation (presence or absence of a condition). Sorry about all the expressions in parenthesis but there are many different labels applied to individual concepts in linear models.

Hint: the function (tidy) from the "broom" package returns model parameter estimates in a "tibble" which is an update of the data.frame idea. From that point, calculation of the "risk multiplier" can be vectorized for ease of use. You should experiment with this in your code.

Hint: Look at the summary() function as applied to a vector. Remember that the 1st quartile is also the 25th percentile, the 2nd quartile is the 50th percentile, and the 3rd quartile is the 75th percentile.