Lab 4G - Growing trees

**Directions: Record your responses to the lab questions in the spaces provided.**

Why can't we just use a *linear model* to predict whether a passenger on the Titanic survived or not based on their gender?

Write down the labels of the two *branches*.

Write down the labels of the two *leaves*.

Which gender does the model predict will survive?

Where does the plot tell you the number of people that get sorted into each leaf? How do you know?

Where does the plot tell you the number of people that have been sorted *incorrectly* in each leaf?

Mrs. Cumings was a 38 year old female with a 1st class ticket from Cherbourg. Does the model predict that she survived?

Which variable ended up not being used by tree?

How is tree3 different from tree2?

In your own words, explain what the *misclassification rate* is.

Which model (tree1, tree2 or tree3) had the lowest misclassification rate for the titanic\_test data?

Does creating a more complex *classification tree* always lead to better predictions? Why not?