N741: Homework 5

Melinda K. Higgins, PhD.

February 27, 2017

Homework 5 - DUE March 15, 2017

For this homework, continue to work with the "Student Performance" dataset we used in today's class. As part of that exercise, we created and saved the data as "student.RData" which is an R-formatted data file. You can load the data.frame `student` back via this code chunk...

```
load(file = "student.RData")
```

Using this dataset, and today's demos complete the following tasks:

1. Make a table containing the mean and standard deviation of the overall grades (G3) grouped by `school` and `sex`. Give the table a title using the `caption=` option and update the column names with something nice using the `col.names=` option in the `knitr::kable()` command. *HINT: You can list more than 1 variable when using `group_by()` in the `dplyr` package, for 2 variables you would use `dplyr::group_by(var1,var2)`.*
2. Make a table containing the frequencies and relative percentages for Mother's job (the `Mjob` variable in the `student` data.frame). Use the example we did in class to help guide you.
3. Make a regression model (Model 1) for the student's overall grade (`G3`) using `school`, `sex` and `age`. Put the regression model results into a table.
4. Make a second regression model (Model 2) for the student's overall grade (`G3`) using `school`, `sex`, `age`, plus `freetime` and `health`. Put the regression model results into a table.
5. Finally, make a table showing the results from the `anova()` command comparing Model 1 and Model 2 you made above using the example we did in class as a guide.
6. STUDENT CHOICE - pick either a `htmlwidget` from http://gallery.htmlwidgets.org/ or do a "flex-dashboard" using the templates at http://rmarkdown.rstudio.com/flexdashboard/ as a guide.