STAT 512 Section 2 Assignment # 4

See Canvas calendar for due date.
44 points total, **4 points** per question unless otherwise noted.

Question 1 through 6 (FisherAlpha): Ott & Longnecker Problem 16.23 describes a study original published in Pyke (2001). Researchers studied the floristic composition of lowland tropical forest in the watershed of the Panama Canal. For this group of question we will work on fitting a model to explain **FisherAlpha (Y)** using **Age and Ppt** as predictors. The following characteristics were measured on 45 plots:

FisherAlpha: a biodiversity index
Age: 1 = secondary forest, 2 = mature secondary, 3 = old growth, primary forest
Ppt: annual precipitation (mm)

Note that Age should be defined as.factor in R.

1. Construct a scatterplot of FisherAlpha (Y) vs Ppt (X) for all Age groups on the same plot. Overlay a fitted regression line for each Age group. (**2 pts**)
2. Fit the ANCOVA model WITH interaction. Include the Type 3 ANOVA table in your assignment. What can we conclude about differences between the slopes for the Age groups? Briefly justify your response.

Note: Continue using the ANCOVA WITH interaction model for questions 3-6.

3. Consider the diagnostic plots (Resids vs Fitted and QQplot of residuals). You do not need to include these plots in your assignment, but briefly discuss your findings.
4. For each Age group, provide the estimated intercept, slope (corresponding to Ppt) and p-value corresponding to a test of the slope. (**6 pts**)
5. Calculate Tukey adjusted pairwise comparisons of the slopes. What can we conclude about differences between the slopes for the Age groups? Briefly justify your response.
6. Calculate emmeans for the Age groups at (A) Ppt = 2500 and (B) Ppt = 3000.

Questions 7 and 8 (FisherAlpha continued): Use the ANCOVA WITH interaction model above as the "full" model. But our goal is to choose a model that predicts FisherAlpha.

7. Based on a backwards elimination approach, which model is preferred? Briefly justify your response. Use $\alpha = 0.05$.
8. Based on **AIC** criteria, which model is preferred? Briefly justify your response.
   Hint: Use dredge() from MuMIn. Use code something like this:

```
library(MuMIn)
options(na.action = "na.fail")
dredge(FullModel, rank = "AIC")
```

Question 9 through 11: Return to the Body Fat data from Assign#2. The data is available from Canvas as "BodyFat.csv". With 3 predictors, there are 8 possible models. Which model would you choose? To identify the model, just state which predictors are included.

9. Choose a model using "backwards elimination" (hypothesis testing) approach. Use $\alpha = 0.05$. No need to discuss, just state your final model.
10. Choose a model using "forward selection" (hypothesis testing) approach. Use $\alpha = 0.05$. No need to discuss, just state your final model.
11. Choose a model using **AICC**. Hint: Use dredge() from MuMIn. No need to discuss, just state your final model.
    Use code something like this:

```
library(MuMIn)
FullModel <- lm(BodyFat ~ ., data = InData)
options(na.action = "na.fail")
dredge(FullModel)
```