**ECO602**

**Weekly assignment: Using Models 2**

**Justin Maynard, Hazel Ortiz,** [**Joshua Driscoll**](mailto:chilldriscoll@gmail.com)

**Nov 27, 2021**

**Q1 (4 pts.):** Re-create the conditional boxplot of penguin body mass conditioned on sex and species.

boxplot(body\_mass\_g ~ species + sex, xlab = "Sex:Species", ylab = "Body Mass (g)", data = penguins, names = c("Adelie female", "Adelie male", "Chinstrap female", "Chinstrap male", "Gentoo female", "Gentoo male"))

boxplot(formula = body\_mass\_g ~ sex:species,

        data = penguins,

        main = "Boxplot of \nbody mass of male and female \nof each penguin specie",

        names = c("Adelie \nfemale", "Adelie \nmale", "Chinstrap\nfemale", "Chinstrap\nmale",

                  "Gentoo \nfemale", "Gentoo\nmale"))

**Q2 (2 pts.):** Based on the boxplots, do you think male penguins (of any species) are significantly heavier than female penguins? Explain your reasoning.

Given that in the boxplots females have higher averages than males in all cases except Adelie, it appears that males tend to be heavier than females. Also given that there is a larger difference between male and female in Gentoo, we would say that only the Gentoo penguins show a significant difference between males and females.

**Q3 (2 pts.):** Do you think adding sex to a model that already includes species will improve the model fit?

I believe that adding sex to a model that currently includes species will help by distinguishing weight characteristics based on sex. This adds critical data to the equation that will make the distinctions more refined.

**Q4 (2 pts.):** Show the R-code you used to build fit\_both.

fit\_both = lm(body\_mass\_g ~ sex \* species, data = penguins)

**Q5 (2 pts.):** What is the base case for the two-way model that includes sex and species?

The base case in this model would be the Adelie female which is shown as the intercept.

Coefficients:                 Estimate Std. Error t value Pr(>|t|)

(Intercept)               3368.84      36.21  93.030  < 2e-16 \*\*\*

sexmale                    674.66      51.21  13.174  < 2e-16 \*\*\*

speciesChinstrap           158.37      64.24   2.465  0.01420 \*

speciesGentoo             1310.91      54.42  24.088  < 2e-16 \*\*\*

sexmale:speciesChinstrap  -262.89      90.85  -2.894  0.00406 \*\*

sexmale:speciesGentoo      130.44      76.44   1.706  0.08886

**Q6 (2 pts.):** What are the names of the two coefficients (from the first column of the coefficient table) you need to calculate the average mass of female Chinstrap penguins?

(intercept) as 3368.84

speciesChinstrap 158.37

**Q7 (2 pts.):** What is the predicted average mass of female Chinstrap penguins in the interactive model?

Average mass = 3527.21

**Q8 (2 pts.):** What is the observed average mass of female Chinstrap penguins, calculated from the penguins data?

aggregate(body\_mass\_g ~ species + sex, data = penguins, FUN = mean)

The observed average mass for female Chinstrap penguins is 3527.206.