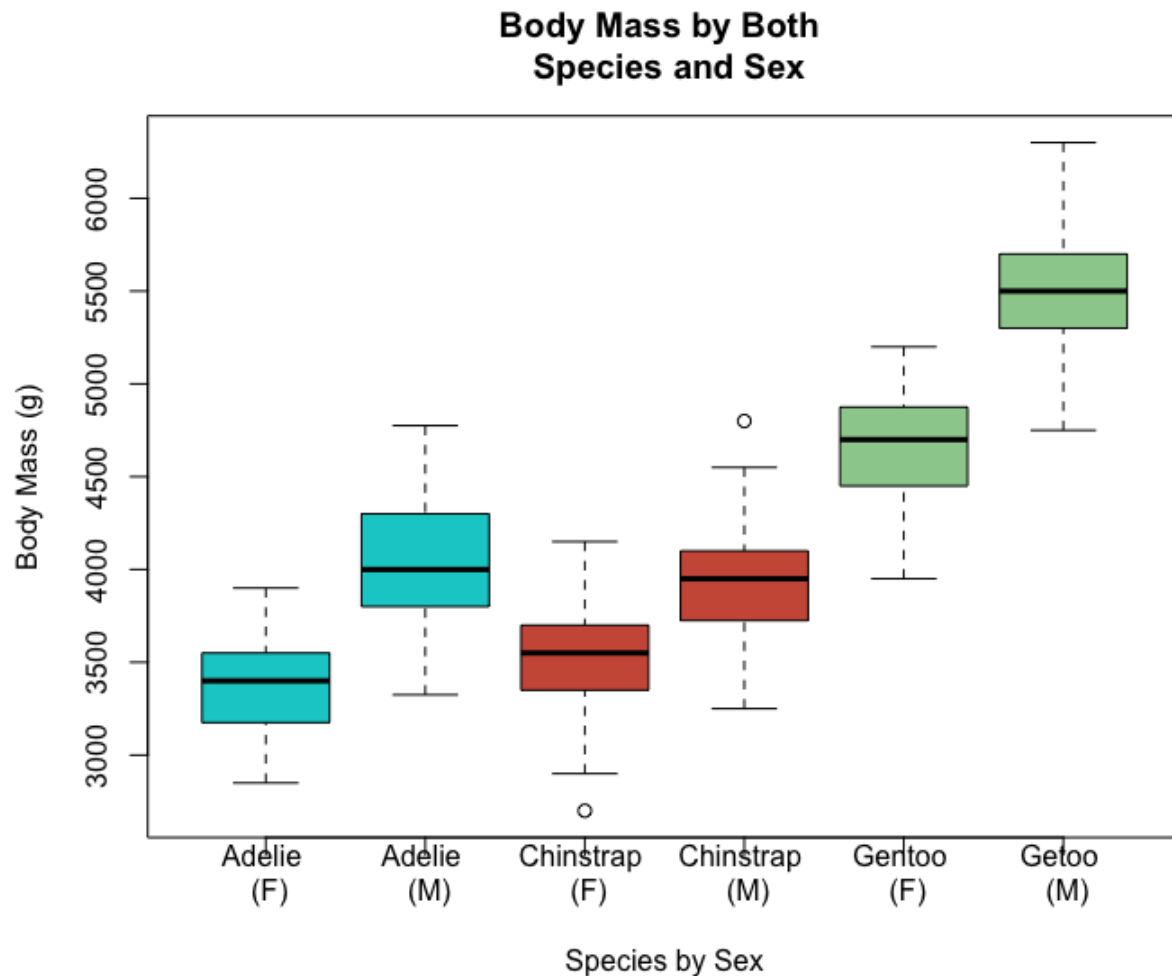Jessica Bonin
Analysis of Environmental Data
Using Models 2 Lecture Assignment
November 28, 2021
Worked with Julia Vineyard

Q1: Re-create the conditional boxplot of penguin body mass conditioned on sex and species.

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

Yes, for each of the species, the male median and box (25th-75th percentile) are not overlapping (in the same body mass range) with the female box. This suggests that the sex plays a significant role in body mass.

Q3: Do you think adding sex to a model that already includes species will improve the model fit? Make sure you justify your answer based on the boxplots and not results of a statistical test.

I think so. The box plot shown in question 1 suggests that sex influences body mass for each of the three species. Having both covariates show a more accurate portrayal of the effects of species and sex on penguin body mass.

Q4: Show the R-code you used to build fit_both.

fit_both=lm(formula=body_mass_g~species*sex,data=penguins)

Q5: What is the base case for the two-way model that includes sex and species?

speciesAdelie (sexFemale)

Q6: 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 and speciesChinstrap

Q7: What is the predicted average mass of female Chinstrap penguins in the interactive model?

3527.21 = 3368.84 + 158.37

Q8: What is the observed average mass of female Chinstrap penguins, calculated from the penguins data?

3527.206

dat_female_chinstrap = subset(dat_chinstrap, sex == "female")
mean(dat_female_chinstrap$body_mass_g)