Using_Models_2

Ethan Rutledge

11/27/2021

```
require(palmerpenguins)
```

```
## Loading required package: palmerpenguins
```

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

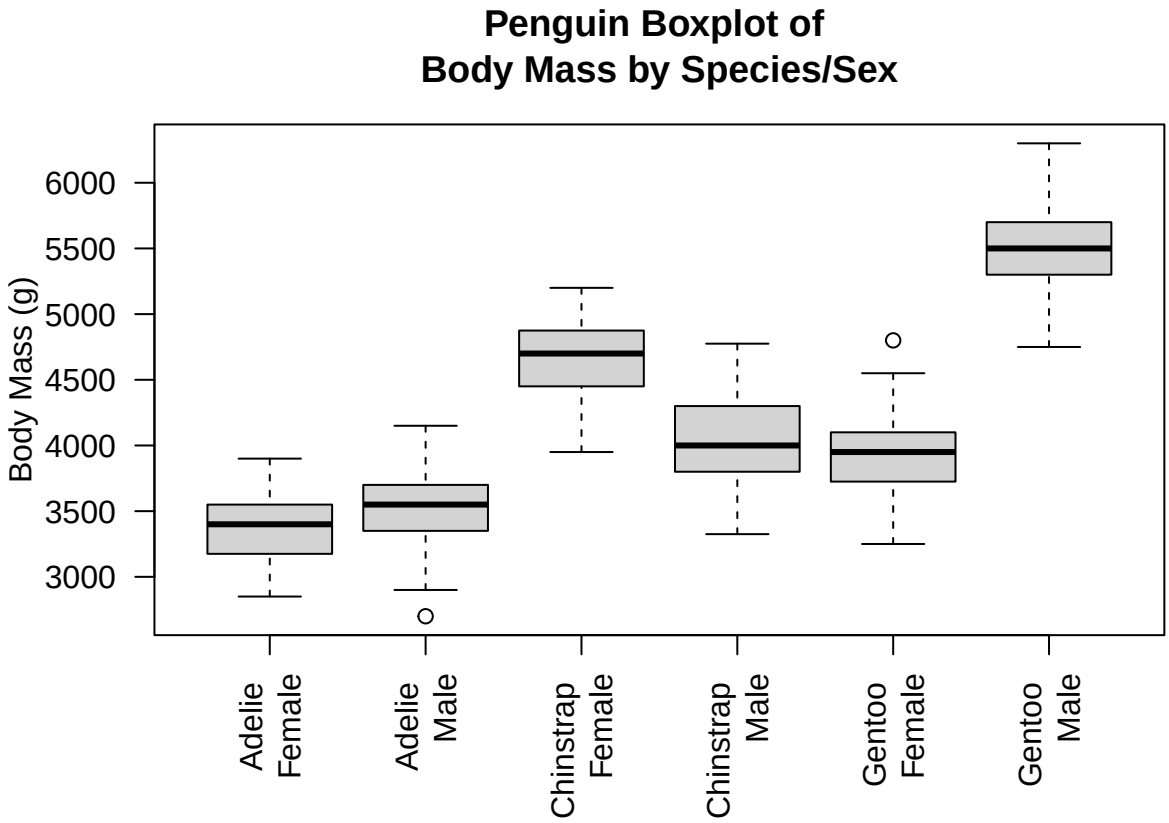
Q2 (2 pts.): Based on the boxplots, do you think male penguins (of any species) are significantly heavier than female penguins? Explain your reasoning. Yes, Gentoo males appear to be significantly heaver than Gentoo females. The error bars/confidence intervals do not overlap.

Q3 (2 pts.): 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 it adds more evidence that there are differences in body mass between species, but we may arrive at that conclusion without sex.

```
boxplot(body_mass_g ~ species * sex, data = penguins, las = 2,
        ylab = "Body Mass (g)",
        xlab = "",
        main = "Penguin Boxplot of \nBody Mass by Species/Sex",
        names = c("Adelie \nFemale", "Adelie \nMale", "Chinstrap \nFemale",
"Chinstrap \nMale","Gentoo \nFemale", "Gentoo \nMale"))
```



Q4 (2 pts.): Show the R-code you used to build fit_both.

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

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

summary(fit_both)
```

```
##
## Call:
## lm(formula = body_mass_g ~ sex * species, data = penguins)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -827.21 -213.97   11.03  206.51  861.03
##
## 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 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 309.4 on 327 degrees of freedom
##   (11 observations deleted due to missingness)
## Multiple R-squared:  0.8546, Adjusted R-squared:  0.8524
## F-statistic: 384.3 on 5 and 327 DF,  p-value: < 2.2e-16
```

```
anova(fit_both)
```

```
## Analysis of Variance Table
##
## Response: body_mass_g
##              Df    Sum Sq  Mean Sq F value    Pr(>F)
## sex           1  38878897 38878897 406.145 < 2.2e-16 ***
## species       2 143401584 71700792 749.016 < 2.2e-16 ***
## sex:species   2   1676557   838278   8.757 0.0001973 ***
## Residuals   327  31302628    95727
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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? sexfemale and specieschinstrap

Q7 (2 pts.): What is the predicted average mass of female Chinstrap penguins in the interactive model? 3527.21

Q8 (2 pts.): What is the observed average mass of female Chinstrap penguins, calcluated from the penguins data? 3527.206

```
aggregate(body_mass_g ~ species * sex, data = penguins, FUN = mean)
```

```
##     species    sex body_mass_g
## 1    Adelie female    3368.836
## 2 Chinstrap female    3527.206
## 3    Gentoo female    4679.741
## 4    Adelie   male    4043.493
## 5 Chinstrap   male    3938.971
## 6    Gentoo   male    5484.836
```