

Lab 1 Homework Template

Your Name Here

2020-08-25

1. Lab 1a. Categorical data

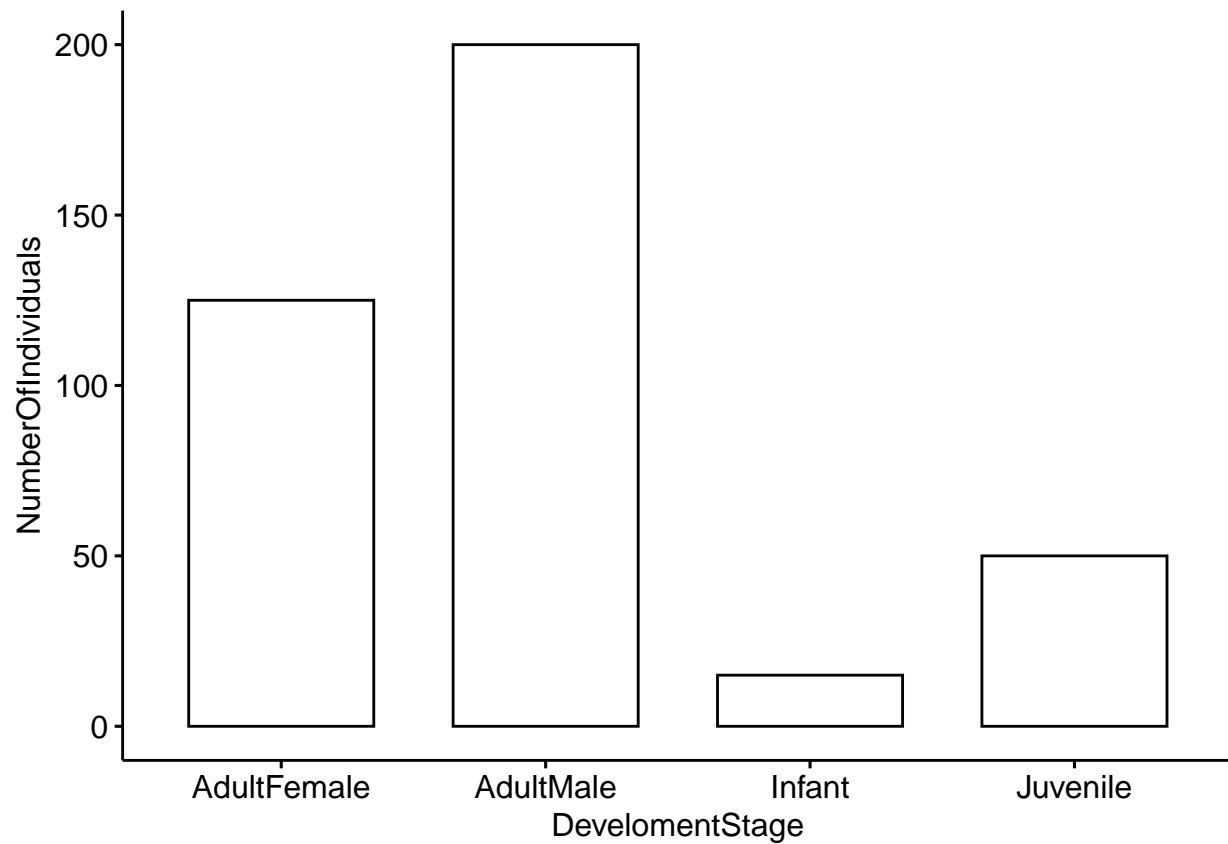
Question 1. How would you interpret this figure? Which category has the most individuals, and which has the least?

ADD YOUR ANSWER HERE

Question 2a. This is a simulated population, but what do you think the ratio of males to females would mean for male-male competition?

ADD YOUR ANSWER HERE

Question 2b. Modify the code so that our simulated deer population has a more even ratio of males to females.



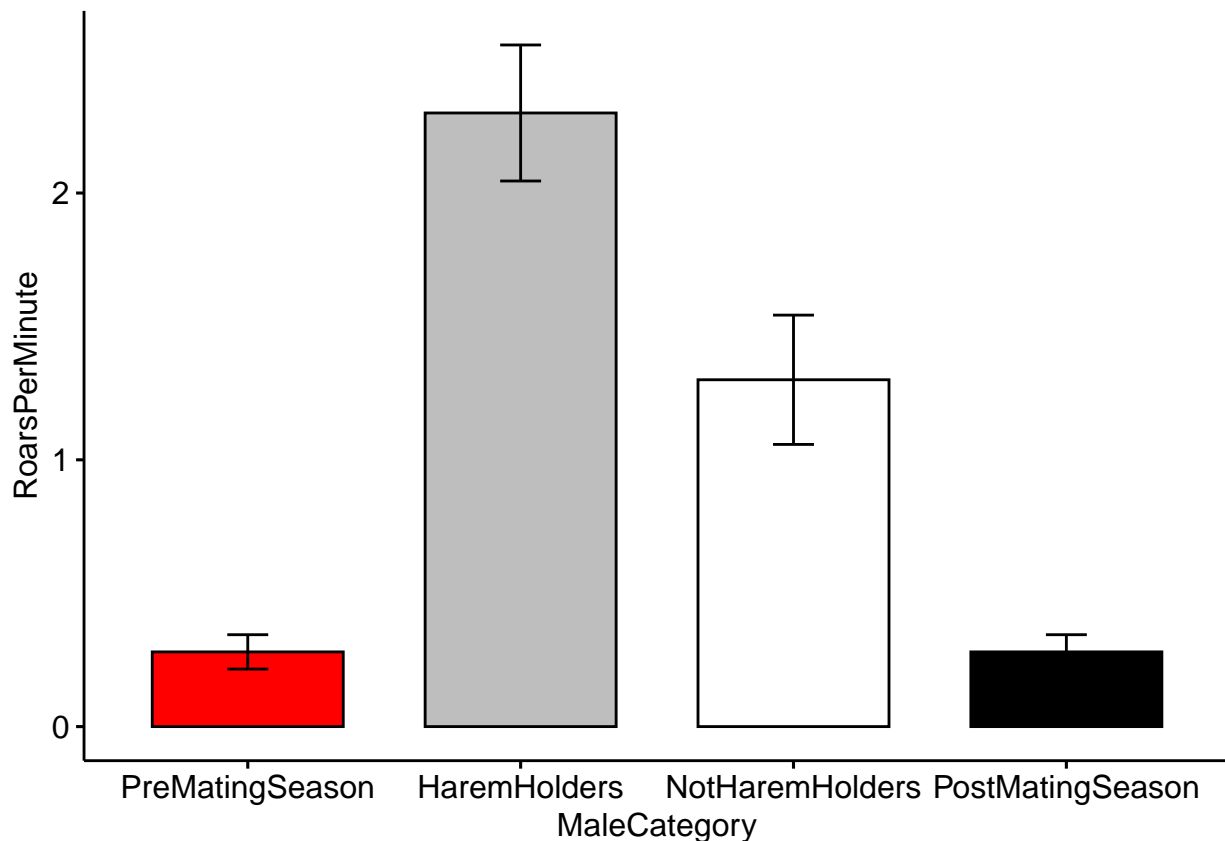
Lab 1b. Categorical and continuous data

Question 3. How do you interpret the barplot figure?

ADD YOUR ANSWER HERE

Question 4. Visit this site (<http://www.stat.columbia.edu/~tzheng/files/Rcolor.pdf>) and change the colors of the plot.

Hint: change `'palette = c('red','gray','white','black')` to `'palette = c('color1','color2','color3','color4')`.



Lab 1c. Categorical and continuous data

Question 5a. What happens when you change the correlation coefficient to a much smaller number and re-run the code?

ADD YOUR ANSWER HERE

Question 5b. What about when you change it to a much bigger number?

ADD YOUR ANSWER HERE

Lab 1d. Multivariate data

Question 6. In your own words, explain why we would want to use PCA for data visualization? What can we learn about the data when it is visualized this way?

ADD YOUR ANSWER HERE