

# STATISTICAL TECHNIQUES FOR THE BIOMEDICAL SCIENCES

## BMR 617

### SPRING 2023

#### Assignment 4. Boxplots with ggplot2

In this assignment, you'll examine in more detail what the ggplot2 function `geom_boxplot` is doing. In addition, you'll get some practice reading the help pages for R function.

#### Exploration:

1. Download the mouse metabolic data from Dr. Kim's lab into R with the tidyverse function `read_csv`. The data can be found at <https://denvirlab.marshall.edu/BMR617-2021/data/TH-B6-metabolic.csv>
2. Create a plot of the insulin values, showing *just* the boxplot without the data points overlaid. Your plot should have one box and whisker graph for each of the six groups. Use colors to denote the mouse strain.
3. For the B6 LF group, you should see a point on the graph, as well as the box and whisker plot. What do you think this point represents?
4. In RStudio, click on the "Help" tab in the bottom right pane. In the search bar, search for the `geom_boxplot` function and read the description at the top. Was your prediction in the previous step correct?
5. Read the "Summary Statistics" section of the help. In terms of the data values, what do the boxes represent in the plot? What do the "whiskers" represent?
6. Suppose we wanted to overlay the boxplot with the individual data points, as we did in the Cholesterol plot in the notes? Would the point you observed in step 3 cause a problem with this? Try the plot and see if you are correct.
7. Read the Help for `geom_boxplot` carefully. Somewhere in there it tells you how to suppress the plotting of points. Create a boxplot for the insulin values, without any point plotted.

#### Assignment:

Create a boxplot for the insulin values for B6 and TH mice fed the three diets. Overlay the box plots with the individual data points. You should ensure that no data is duplicated in the plot.

Submit a Word document containing a couple of sentences of context, the boxplot with points overlaid, and a figure legend. The figure legend should explain what is shown in the graph (i.e. the reader should be able to tell from the legend what the boxes, whiskers, and points represent). Also submit the R code used to generate the figure. There is no need to submit the R code you executed during the "Exploration" part of the assignment.