R bootcamp basic stats and plotting worksheet:

1. Take a look at the 'iris' data set

   Hint: you can use functions such as head(), summary(), unique()

2. Try the following on your own, I will walk you through it soon so don't get too frustrated…
   a. Assign data for the setosa species to one variable named 'setosa' and assign the subset of the data for the virginica species to another variable named 'virginica'.
   b. Plot the sepal lengths of the two species as two histograms.
   c. Now plot the sepal lengths of the two species so that you can compare them.
   d. Is there a statistically significant difference between the sepal lengths of the two species?

3. Plot and compare two other measures on your own....

4. Compare the means of more than two groups.
   a. Plot the sepal length of all three species (**Hint**: you can do this in one figure)
   b. Add a legend with species names to your plot.
   c. Is there a statistically significant difference in sepal length among the three species? (**Hint**: use an ANOVA)
   d. Which species are significantly different (**Hint:** run a post-hoc test)
   e. Add the information from the post-hoc test to your figure

5. Are sepal and petal length correlated?
   a. Plot sepal length against petal length
   b. Are these two measures correlated? (**Hint:** use a Pearson's correlation)
   c. Color code your plot by species, include a legend on the plot.

d. Do species vary in the relationship between sepal and petal length?

6. Examine (plot and stats) the relationship between petal ***length*** and sepal ***width*** and how this relationship differs among the three species.