Q SCI 482 Assignment 4 (100 points)

Your name: _____

Scenario: are blue rockfish actually two species?

Blue Rockfish (*Sebastes mystinus*) was the most important commercial fish species in California in the 19th century, and remains a key recreational species in the area. The *Sebastes* genus in the eastern Pacific Ocean is incredibly diverse, including almost 100 species, many of commercial importance, and all with similar body shapes. Species identification is a problem. A new analysis (Frable *et al*. 2015) suggests that this species might actually be two separate species, the Blue Rockfish and a new species they have named the Deacon Rockfish (*Sebastes diaconus*). This determination was made partly on skin coloration, with the Deacon Rockfish having a more solid color (top photo on the right) and the Blue Rockfish a more mottled color (bottom photo); partly on genetics; and partly on morphometrics (body measurements). If this reclassification is correct, not only will further work be needed to figure out where the two species overlap, but all of the historical catches need to be separated, and the fisheries stock assessments rerun to ensure that catch limits on both species are sustainable. In this assignment, we will focus on five body measurements contained in the "`Sebastes_diaconus_data.xlsx`" file in the assignment folder, and analyze these data to see if they support the separation of Blue Rockfish into two separate species. Data for analysis will be restricted to fish of lengths 150 to 200 mm standard length, and each of the five body measurements from Dr. Frable have been converted into percent of standard body length in sheet "`Data for QSCI 482 assignment`". They can be assumed to come from normal distributions.

Figure 1 Top: Deacon Rockfish, bottom: Blue Rockfish.

Q1: Converting Excel sheet into .csv file (10 points)

Browse all of the sheets in the Excel workbook to familiarize yourself with the data, and look at the pictures of what each measurement represents. Create an R project for Assignment 4, start a new .R script file, and save it. Follow the instructions from the Lab 4 handout to:

(a) Create a .csv file (suggestion "`Sebastes_diaconus_data.csv`") from the appropriate sheet in Excel (just the data, not any accompanying analysis) and load it into R. Examine the read-in data to ensure that text columns remain as text, and that there are not any extra columns or rows in the data. Upload the .csv file to Canvas as part of your solution. [5 points]

(b) Write R code to extract the data for the caudal peduncle of Deacon Rockfish from the .csv file and save it to a variable. [5 points]

Q2: Caudal peduncle comparison (35 points)

The variance for the two putative species is quite similar for the caudal peduncle data. Run a statistical test to test if the means of caudal peduncle length (data are percentages of standard length) for the two samples of the putative species are equal, assuming that variances are equal.

(a) Run the test the long way around (calculating the appropriate test statistic, the *p*-value, critical values of the test statistic, and reaching conclusions for $\alpha = 0.05$) [30 points]
(b) Run the test using a built-in statistical test in R, reporting the key results [5 points]

The two tests should match up.

Q3: Anal spine I comparison (35 points)

The variance of anal spine I is much greater for Blue Rockfish than Deacon Rockfish. Run a statistical test to test if the means of anal spine I lengths (data are percentages of standard length) for the two samples of the putative species are equal, assuming that variances are not equal.

(a) Run the test the long way around (calculating the appropriate test statistic, the p-value, critical values of the test statistic, and reaching conclusions for $\alpha = 0.05$) [30 points]
(b) Run the test using a built-in statistical test in R, reporting the key results [5 points]

The two tests should match up.

Q4: Comparing the other measurements (15 points)

Now run statistical tests on the other measurements using the appropriate built-in statistical tests in R. You will need to decide if the variances are similar or not. Report the key results for differences in the means for:

(a) The symphyseal knob [5 points]
(b) The anal spine II [5 points]
(c) The anal spine III [5 points]

Q5: Conclusions (5 points)

What do you conclude from these tests and the data: should blue rockfish have been separated into two species or not? Explain your reasoning.

R functions and constants that might be useful for this assignment

```
mean()      #mean of a vector of data
pt()        #given the tvalue and d.f., returns the area to the left (CDF)
qt()        #given CDF probability and d.f., returns value t from t-distribution
read.csv() #read in a CSV file, e.g. xdata <- read.csv(file="values.csv")
t.test()    #runs a t-test, either one-sample or two-sample on a set of data
var()       #returns the sample variance of values in a vector, e.g. var(vector)
```

Reference

Frable BW, Wagman DW, Frierson TN, Aguilar A, Sidlauskas BL (2015) A new species of *Sebastes* (Scorpaeniformes: Sebastidae) from the northeastern Pacific, with a redescription of the blue rockfish, *S. mystinus* (Jordan and Gilbert, 1881). Fishery Bulletin 113:355-377.