1. Do Computational Exercise 12 from Chapter 2 on marijuana use during pregnancy.

2. Do Computational Exercise 20 from Chapter 2 on the relationship between a fish oil diet and blood pressure.

3. We worked through a handout in class on comparing two means. The first example involved analysis of the Bumpus data on the humerus lengths of sparrows. The analysis that we did can be handled easily by SAS. Download the program *sparrow.sas* from the course web site. Examine the contents of this file and run *sparrow.sas* to answer the following questions.

    (a) The data contains two columns representing two variables. The second column consists of 1s and 2s. What do these numbers represent? (Hint: Read the comments in *sparrow.sas*.)

    (b) After reading the data the program executes a *proc univariate plot;* command. The *var humerus;* command tells sas to run the *univariate* procedure on the variable *humerus*. Explain what the *by status;* command does by studying the output of the program.

    (c) Use the output from the *univariate* procedure to compute a 95% confidence interval for the mean humerus lengths of sparrows who would have survived the storm had they been exposed. The *univariate* procedure does not compute the interval for you, but it does compute the summary statistics that you need.

    (d) The *univariate* procedure provides side-by-side boxplots that allow a visual comparison of the data in the two groups. You made similar boxplots for this data in the first homework assignment. There is a + that appears in each boxplot. Can you determine what the + represents?

    (e) There is a 0 that appears past the lower endpoint of each boxplot. In both cases this 0 represents the smallest observation in the data set. SAS does not extend the boxplot to include any observations that are more than 1.5 times the interquartile range above the upper quartile or more than 1.5 times the interquartile range below the lower quartile. The interquartile range is $Q3 - Q1$.

        i. Find the interquartile range for the sparrows that survived.
        ii. Find $Q1 - 1.5 * (Q3 - Q1)$ and $Q3 + 1.5 * (Q3 - Q1)$ for the sparrows that survived.
        iii. Verify that the smallest humerus length of a sparrow that survived is less than $Q1 - 1.5 * (Q3 - Q1)$.
        iv. Verify that no other humerus length for a sparrow that survived falls outside the limits computed in part (ii).

    (f) The next set of commands begins with *proc ttest;*. The procedure *ttest* will conduct a ttest to compare the means of the two groups specified by the variable named in the *class* statement. State the degrees of freedom, test statistic, and p-value for the test of $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$ assuming the standard deviations in the two groups are equal. (Here μ_1 is the mean humerus length of sparrows who would have died had they been exposed to the storm, and μ_2 is the mean humerus length of sparrows who would have survived the storm.)

    (g) Provide a 95% confidence interval for $\mu_2 - \mu_1$. This interval is provided by the *proc ttest* commands. The *ttest* procedure provides many confidence intervals, so you may have to do a little work to verify that you have found the right one.

4. Suppose 6 pairs of hogs are used in an experiment to determine if a dietary supplement can increase weight gain. Each pair of hogs are full siblings. One sibling is randomly assigned to receive the dietary supplement. The other sibling receives a supplement that does not contain the active ingredient believed to affect weight gain. The gains in pounds are provided below.

Sib Pair	Treatment Weight Gain	Placebo Weight Gain
1	70.4	69.3
2	38.2	37.2
3	59.7	58.7
4	40.3	39.3
5	93.5	92.5
6	50.7	49.6

Note: You don't have to use SAS to answer the next two problems, but I think you will save yourself some time and tedious work if you do use SAS. You should be able to develop the necessary code by examining programs such as *sparrow.sas* and *trees.sas*.

(a) Is there significant evidence that the supplement increases weight gain? First analyze the data by examining the mean of the differences in weight gain for each sib pair. Compute the test statistic, p-value, and state a conclusion. (See *trees.sas* for some sample SAS code that may be helpful.)

(b) Now analyze the data by assuming that the two columns of data represent independent samples from two populations (hogs treated with the supplement and hogs treated with the placebo). Compute the test statistic, p-value, and state a conclusion. (See *sparrow.sas* for some sample SAS code that may be helpful.)

(c) Which analysis of the data is correct? Explain briefly.

(d) Can you explain what features of the data cause the results of the two analyses to differ?