Determining if there is evidence of "acid" rain:
An example of hypothesis testing for a single mean

This worksheet accompanies the article:

Martinez-Dawson, R. (2003) "Incorporating laboratory experiments in an introductory
     statistics course," *Journal of Statistics Education* [Online], 11(1).
     http://www.amstat.org/publications/jse/v11n1/martinez-dawson.html

Purpose:

The purpose of this experiment is to illustrate the concepts of sampling, statistical
analysis of a small sample for a single mean, and drawing a conclusion based on the
analysis. A 95% confidence interval for the true population mean will be calculated.

We are interested in determining if there is "acid rain" in the Clemson area. Acid rain is
defined as "any rainfall that has a pH value of less than 5.6." For more information
concerning acid rain, see qlink.queensu.ca/~4lrm4/effect2.htm

Materials:

You will need the following materials: pH meter and containers of rain water from the
area.

Procedures:

Rainwater was collected in jars from the Clemson area. Record the pH of each water
specimen in the chart provided.

Water specimen	pH
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Statistical analysis:

Conduct a hypothesis test to determine if acid rain exists in the Clemson area (α =.05).
You will conduct the hypothesis test by hand and computer and compare the results.

1. State the null and alternative hypotheses. Define the parameter(s). (6 pts)

2. Calculate the test statistic. Show all your work. (12 pts)

3. Determine the approximate p-value. Be as precise as possible. (4 pts)

4. What is the decision? Explain. (2 pts)

5. What is the conclusion? (4 pts)

Now conduct the hypothesis test using SAS. The following code (in bold) will perform the hypothesis test using the p-value method. A description of any new code is given to the right of the SAS code.

data acid;
 input ph;
cards;

After the cards statement, type the 10 pH readings. One line for each pH.

proc ttest h0=5.6 alpha=.05;   Computes mean, s, t_{obs}, p-value, and CI's for pH

var ph;
title Type your name here**;**
run;
quit;

Questions based on the SAS output

6. On the printout, what does the value under the "Mean" column represent? (2 pts)

The formula for the test statistic is:

$$t_{obs} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

7. Put the values from the printout correctly into the test statistic formula. (2 pts)

$t_{obs} =$

8. Identify the test statistic **on the printout** by labeling it, t_{obs}. (2 pts)

9. Identify the p-value **on the printout**. (2 pts)

10. Draw the t-distribution and indicate the area that corresponds to the p-value. (4 pts)

11. What is the decision? Explain. (2 pts)

12. Based on the hypothesis test, what is the conclusion? (4 pts)

Compare the results from SAS with the results determined by hand.

13. Name the parameter that the confidence interval created by SAS encloses with 95% confidence. Write the confidence interval in the following notation: (LCL, UCL). (2 pts)

14. Name a factor that may have contributed to variability in pH readings. (2 pts)