STAT 517: Final Exam

All data sets you need for the exam can be found as separate worksheets in the Excel workbook; import them as needed. When importing Stage and WQ for Question 4, be sure to de-select the DATE format under Worksheet Options in the SAS Import Wizard.

1. Read the Excel worksheet Q1 into WORK data set Q1; this is the transposed fish data from Homework 4. The The following commands construct a two-way table of fish count by Lake and Month.

```
PROC FREQ DATA=Q1OUT;
WEIGHT COUNT;
TABLE LAKE*MONTH/CHISQ;
RUN;
```

   (a) Use `ODS TRACE` and `ODS SELECT` to identify and save the table of chi-squared test statistics.

   (b) Using the output data set created by `ODS SELECT`, print only the results for the chi-squared test statistic and the likelihood ratio chi-squared test statistic, including the name of the test, its degrees of freedom, its test statistic and its p-value. Use attractive labels.

2. Import the SES worksheet for this problem.

   (a) Use `%LET` to specify a single dependent variable and a single independent variable to be input into `PROC REG`. Regress the math score (dependent variable) on the social studies score (independent variable).

   (b) Use `PROC MEANS` to find and store the median reading score.

   (c) Use `CALL SYMPUT` to print the records of all individuals with a reading score greater than the median reading score.

3. Return to the fish data from Question 1. In all cases, print the SQL tables.

   (a) Import the data into SAS (unless you have done so already), then read it into a SQL table.

   (b) Order the data by lake and month. Comment on the ordering.

   (c) Use `PROC SQL` commands to compute the sum of counts by Lake.

   (d) In `PROC SQL`, select only the records for Largemouth Bass and save them in a new table called Largemouth.

   (e) *Grad students* Produce a table that includes the high count over all three fish species for each lake and month combination. The final table should include lake, month, fish species with highest count, and high count.

4. Read in both data worksheets from the Excel file and save them under the names Stage and WQ.

(a) Construct an inner join in PROC SQL on River and Time; generate the same output table in SAS with a match merge. Comment.

(b) Construct a right join in PROC SQL on River and Time. Comment.

(c) Grad students should generate a full join in PROC SQL. Print the temperature in Fahrenheit rather than degrees centigrade, and be sure to fix any formatting problems in the output. Comment.