STAT 440 – Homework 01

Students are encouraged to work together on homework. However, sharing or copying any part of the homework is an infraction of the University's rules on Academic Integrity.

Final submissions must be uploaded to our Compass 2g site on the Homework page. No email, hardcopy, or late submissions will be accepted. Again, NO homework will be accepted after the deadline.

Submitting your work to Compass 2g

You are to submit two (and only two) files for your homework submission.

1. Your SAS program file which should be saved as **HW*n*_*YourNetID*.sas**. For example, my file for the HW01 assignment would be HW01_dunger.sas. All program statements and code should be included in one program file.

2. Your Report file which should be saved as **HWn_*YourNetID*.pdf**. For example, my file for HW01 would be HW01_dunger.pdf . Include all relevant output to address the exercises. Do not include output for every execution of your SAS program.

   For this homework, use ODS to send your results to a Rich Text Format (RTF) file, then open it in Word or some other word processor. Insert your responses to the exercises, then print to a Portable Document Format (PDF) file. The PDF file is the one you will submit.

You have an unlimited number of submissions, but only the last one containing both of these two files will be viewed and graded. Homework submissions must always come as a pair of files, as described above. To submit,

- click on the title of the assignment (e.g., HW01) in Compass,
- attach the two files one at a time (no ZIP or RAR files), then
- click **Submit**. Do not click Save Draft.

1. You will be working with the SAS data file **shoes** which contains worldwide sales figures for a variety of shoe types for an unnamed manufacturer. It is located in the **sashelp** library.

   a. Print the descriptor portion of this SAS data file. Include only the General Information and Variable Attributes tables. Read the results to become familiar with the contents of this data set.
      - Identify the number of variables and observations.
      - Which variables are character and which are numeric?

   b. Print the data portion of **shoes**, but only for sales figures based on stores in Singapore. Which shoe type had the most total sales for this year? Which is sold at the most number of stores?

   c. Create a results table that shows the sum, average, minimum, and maximum of inventory values across all products for every subsidiary location in the Asia region. Which location has the largest total (sum) inventory? Which location has the smallest?

2. For this problem, you will use the SAS Help library's **baseball** dataset.

   a. Print the descriptor portion of this SAS data file. Include only the General Information and Variable Attributes tables. Read the results to become familiar with the contents of this data set.
      - Provide a description of this data set. For example, "Average Northern Hemisphere Temperature from 1856-2000 and eight climate proxies from 1000-2000AD. Data can be used to predict temperatures prior to 1856."
      - Identify the number of variables and observations.
      - How many variables are character and which are numeric?

   b. How many players make more than $2 million? Print a table of all players who make more than $2 million. Include only the players' names, team, position, and salary (in thousands of dollars).

   c. How many players play each Position? Create a frequency table.

   d. How many players are in each unique Division? Create a cross-tabulation to show.
   Note: The variable Div is a single variable whose levels detail each unique division, but because you are specifically asked for a cross-tabulation, you will need to use two other variables instead.

   e. Print a table of all National League players who have been in the major leagues between 2 and 5 years, inclusive. Include only the players' names, team, years in the major leagues, and career home runs.