ST 307: Topic 5 Activity

For this activity you will create a SAS program and upload that program to wolfware. Be sure that your SAS file adheres to the SAS file submission guidelines (available on wolfware).

The data set(s) for this activity are available in the assignment 5 link.

To Do - Write code corresponding to each step below, that is, do not change the code for step 1 to do step 2 (you can copy and paste it so you don't have to retype it, but leave the answer to each step in your program):

1. Create a new library called Week5 where you will save your SAS data sets.

2. In a comment, describe the reason we might use informats for
   a) character variables
   b) numeric data

3. The data set super.txt is available on the assignment link. There are two variables, first and last. Open the file in a program such as and in a comment in your SAS program answer the following:
   a) What is the length of the longest first name?
   b) What is the length of the longest last name?

4. Consider the SAS code below to read this data set in. Assume the path name is correct. What would go wrong when reading in this data? Answer in a comment. If you aren't sure, you can modify the path and run the code if you like!

```
DATA Week5.superfolk;
      INFILE "E:\NCSU classes\ST 307\DataSets\super.txt";
      INPUT First $ Last $;
RUN;
```

5. Modify the DATA step above to correct the issue and successfully read the data set into your library. Hint: Recall the colon (in)format modifier. **Print this data set to the output window**.

6. The data set superPlus.txt has 3 additional variables, Dept, DayOff and TimeIn. In the comments, answer the following.
   a) What appears to be the format of the "DayOff" variable?
   b) What appears to be the format of the "TimeIn" variable?

7. Create a new data step to read in the superPlus.txt data set. You'll need to modify the INPUT statement accordingly. Notice the variable names in the file. (HINT: use DDMMYYw. and TIMEw. informats – see SAS help if needed.) **Print this data set to the output window** - note that SAS stores the dates and times as integers.