

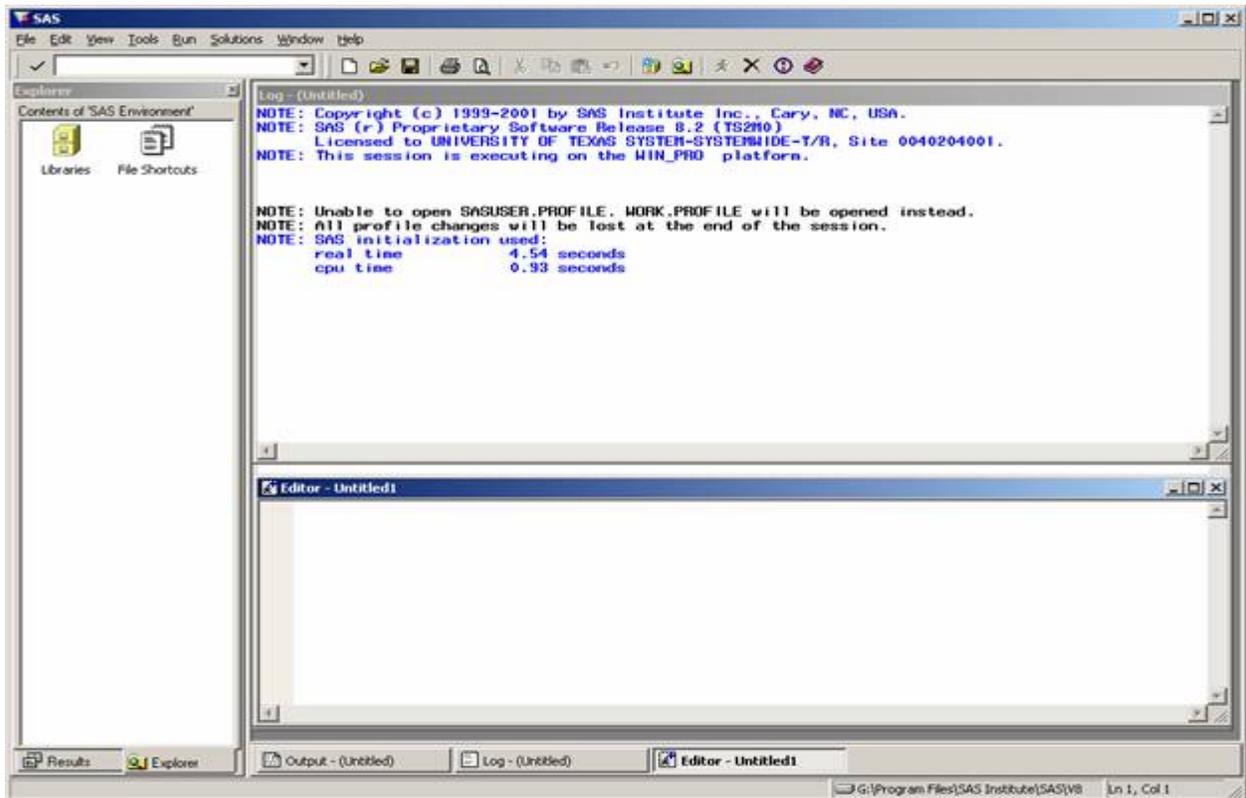
A Really Quick Introduction to SAS¹

SAS is a software package used for conducting statistical analysis, manipulating data, and generating tables and graphs that summarize data.

Navigating SAS for Windows

The SAS System for Windows consists of the 4 main windows

1. Results and Explorer window
 - a. Explorer pane is browsing tool for SAS libraries
 - b. Results pane shows a tree-like summary of the output window. You can select, delete, or edit the output before printing, saving, or copying the results
2. Enhanced Program Editor window - Used to create, edit and execute SAS programs
3. Output Window - Displays output from SAS program. Can view, print, copy or save information in the output window
4. Log Window - Reports on progress of SAS procedures (BLUE). Displays error messages (RED) and warnings (GREEN).



¹ Some of the material for this introduction was derived and adapted from SAS I: Getting Started created by Information Technology Services of the University of Texas and available at www.utexas.edu/its/rc/stat/sas/SASI/sasgettingstarted.html.

General Information

- All statement lines must end with a semi-colon.
- Comments are indicated in 2 ways.
 1. Start line with * and end with a semi-colon
 2. Enclose with as */* put comment text here */*
- End all procedures with RUN.

Getting Data into SAS

There are 3 ways to get data into SAS: 1) Read in-stream data, 2) Use INFILE statement, and 3) Import the data from Excel.

Read In-stream Data

The data area typed directly into the SAS program. This approach works well for small datasets.

```
DATA HTWT;  
    INPUT SUBJECT GENDER $ HEIGHT WEIGHT;  
DATALINES;  
1 M 68.5 155  
2 F 61.2 99  
3 F 63.0 115  
4 M 70.0 205  
5 M 68.6 170  
6 F 65.1 125  
7 M 72.4 220  
;  
  
PROC PRINT DATA=HTWT;
```

The DATA statement indicates that a data set called HTWT is to be created. The INPUT statement names the variables in the dataset and indicates the type of variable. The \$ following GENDER signifies that it is a categorical variable. The DATALINES statement indicates that the data to be read follows. The actual data is included in the program following the DATALINES statement and ending with a semi-colon.

To save your program, go to File... Save and save the program with a *.sas extension. You also can create programs in any text editor.

Use the INFILE Statement

The INFILE statement allows SAS to access ASCII or text files. The INFILE command reads from the file HTWT.txt. The INPUT command names the variables as before.

```
DATA HTWT;  
    INFILE 'c:\HTWT.txt';  
    INPUT SUBJECT GENDER $ HEIGHT WEIGHT;
```

SAS Import Wizard

The SAS Import wizard can be used to access spreadsheets (Excel, Lotus) and database (Access) files. It is most convenient if the variable names are in the first line of the Excel

spreadsheet and comply with SAS naming conventions – no more than 8 characters long, no spaces in the middle, and start with a letter.

Go to File...Import Data

Select file type then select Next

Enter location of the file. Can use the Browse button. Use the Options button and make sure “Column names in first row” is checked if your spreadsheet contains variable names. Select Next.

Indicate the library (usually WORK) and the member – This is just a name to call the file while working with it in SAS. Select NEXT if you want the PROC IMPORT statements saved for later use or FINISH to import the data.

The imported file is can be viewed using the Explorer window. Once the data is imported you can refer to the data set by its name (e.g., DATA=HTWT).

A Short SAS Program

SAS programs can be typed directly into the Program Editor or created in any text editor and copied into the Program Editor. To run the program, select the  button.

```
TITLE 'A SHORT SAS PROGRAM';
OPTIONS LS=72;

/* Create data file with height and weight data*/

DATA HTWT;
    INPUT SUBJECT GENDER $ HEIGHT WEIGHT;
DATALINES;
1 M 68.5 155
2 F 61.2 99
3 F 63.0 115
4 M 70.0 205
5 M 68.6 170
6 F 65.1 125
7 M 72.4 220
;

/* Create a new categorical variable for height*/

DATA HTWT;
    SET work.HTWT;
    IF HEIGHT < 68 THEN STATURE='Short';
    IF HEIGHT >= 68 THEN STATURE='Tall';
RUN;

* Print the data file HTWT;

PROC PRINT DATA=HTWT;
    TITLE 'HEIGHT AND WEIGHT DATA';
RUN;

* Sorts the data by gender. Some procedures require sorted data;

PROC SORT DATA=HTWT OUT=sorted;
```

```

        BY GENDER;
RUN;

* Generates descriptive statistics;

PROC MEANS DATA=sorted;
    TITLE 'SIMPLE DESCRIPTIVE STATISTICS FOR HTWT';
    BY GENDER;
RUN;

* Generates frequency 2x2 frequency table;

PROC FREQ DATA=HTWT;
    TABLES GENDER*STATURE;
RUN;

```

For Practice

Download the data file marten.txt from the class website. This data set consists of temperature data of different types of resting sites used by the American marten (a weasel like animal) during periods of deep snow and shallow snow. The first column is the type of structure used for resting:

Logs – the marten rested under a decaying log

Tree – the marten rested in a hole in a tree

Rocks – the marten rested in a cavity in a pile of rocks.

The second column is the temperature difference between the temperature inside the structure and outside the structure in degrees Celsius. A positive number indicates that it was warmer inside the structure than outside. The third column indicates whether the snow was deep or shallow.

Write SAS statements to:

1. Bring the data into SAS using any method you like.
2. Create a new categorical variable for temperature differences greater than 0 and temperature differences less than or equal to 0
3. Print the data.
4. Sort the data by one of the variables.
5. Generate descriptive statistics for temperature differences by structure type or snow depth.
6. Generate frequency tables for the new categorical variable for temperature differences and structure type and snow depth