Using SAS on the UNIX[1]

Running Programs

Unless you are using a graphical interface, you will probably run SAS in batch mode on the UNIX. SAS commands are very much the same as on other platforms. Following is a short SAS program named "sample.sas" that reads compressed raw data and saves a SAS system file. SAS on the UNIX can read raw data that are compressed with either the compress utility (extention of ".Z") or with the gzip utility (extension of ".gz".)   Remember that the UNIX is case-sensitive. When you define data, use upper and lower case characters accurately.

```
* sample.sas ;
options compress = yes linesize = 72 pagesize = 59 ;
filename inz1 PIPE  'zcat /full/path/raw.data.Z' lrecl = 80  ;
libname saslib '/full/path/name' ;
Data saslib.sample ;  infile inz1 ;
input v1 1-3   v2 4-5 ;
proc means ;
```

The filename statement for a non-compressed data set would look like:

```
filename inz1  '/full/path/raw.data' lrecl = 80 ;
```

To run this program, you would type the following:

```
sas sample &
```

For more information about SAS, please see http://staff.washington.edu/glynn/SASINTUN.pdf and for more information about using UNIX, see http://www.albany.edu/csda/classa.pdf .

More about filename statements

A filename statement that would refer to multiple files in a directory could look like:

```
filename inz1 PIPE  'zcat ~/8071/da8071.*gz'  LRECL= 2016 ;
```

There are two important things to note about the above filename statement. First, look at the tilde followed by a slash.   **~/** The tilde refers to the home directory of the user. So, this filename defines files in the directory 8071, which is a subdirectory of the user's home directory. Second, look at the asterisk. The asterisk is used as a wildcard. The above statement defines a filename that includes all files that begin with da8071. and end with .gz. There are many files in this directory (all of the state files for the 1980 STF3A Census files) - such as:

```
da8071.ak02rev.gz
da8071.al01rev.gz
```

The asterisk **da8071.*gz** tells SAS to refer to all files that begin with **da8071** and end with ***gz** no matter what is in between.

For more information about SAS, please see http://staff.washington.edu/glynn/stat.html#SAS