

Troubleshooting and Finding Descriptive Statistics

1 Cards or Datalines

During the last lab we explored how to get data set into SAS using external files. The other way of getting data into SAS is using a cards or datalines statement. For example;

```
data c;
input race sex hair $;
cards;
4 1 black
2 1 brown
3 0 blonde
3 1 brown
2 1 brown
1 0 black
4 1 brown
4 1 brown
1 0 blonde
2 0 black
3 1 black
4 0 brown
3 1 black
6 0 brown
;

proc print data=c;
run;

proc freq;
tables hair;
run;
```

2 Troubleshooting

2.1 Correctly Identifying Your Data

Problem: Trying to use a proc on a data set that does not exist. For example:

```
data b;
input sex height;
cards;
1 72
0 70
1 69
0 62
0 59
```

```

0 67
1 70
1 69
;

```

```

proc means data=lab1;
var height;
run;

```

Notice that the data name used in the proc means statement does not match the name of the data set created. This code will result in the data set 'b' being entered, but no proc means statement will be done. So, either make sure that your data= is correct or don't put it there at all. For example notice in the following that there is no data= statements.

```

OPTIONS LS=80 NOCENTER pagno=1;
DATA ONE ;
array lamda (2, 6) lamda11-lamda16 lamda21-lamda26;
input ITER NOBS lamda11-lamda16 lamda21-lamda26 ;
CARDS;
1 20 .1 .2 .4 .8 .8 .8 .15 .3 .6 1.2 1.2 1.2
;

```

```

DATA TWO; SET ONE; array lamda (2, 6) lamda11-lamda16
lamda21-lamda26; seed1=136121; * you can change this to get a
different set of data; do i=1 to nob;
    do J=1 to 2;
        do k=1 to 6;
Y1=-LOG(ranuni(seed1))/LAMDA{j,k}*abs(J-2)+(J-1)*
sqrt(-LOG(ranuni(seed1))/LAMDA(j,k)) ;
DEATH1=0;
C1=1000;
IF C1 GT Y1 THEN DEATH1=1;
T11=DEATH1*Y1+(1-DEATH1)*C1;
agecat=j;
dose =k;
hazard=LAMDA{j,k}*abs(J-2)+(J-1)* LAMDA{j,k}*2*t11;
OUTPUT;
end;
END;
end;
run;

```

```

data three; set two ;
if dose ge 5 and agecat=2 then dose=4;
if dose gt 4 then delete;
agecat=agecat-1; * this is just to make agecat 0,1 instead of 1,2;
agitates=agecat*dose;
dlogist=0;

```

```

t1new=t11;
if t11 gt 2 then do;
dlogist=1;
t1new=2;
end;
logt11=log(t11);
run;

proc sort; by agecat dose;

proc means; var t11; by agecat dose;
run;

proc lifetest plots=(s,lls)notable;
time t11*death1(0);
strata agecat dose;
run;

proc phreg;
model t11*death1(0)=dose chkdose;
chkdose=t11*dose;
test1: test dose + .5*chkdose=0,dose + 2*chkdose=0,dose + 4*chkdose=0/print;
run;

```

The data= statement is only for using a proc on a data set that was *not* the last one read in. This is often the case using permanent SAS files.

2.2 Running...

As seen by many participants in the previous lab, SAS can get stuck in a loop on a line. When this happens SAS will continue to try and submit a line until submit is hit again, in which SAS will skip over the line, but still may not run correctly. For example:

```

data examp;
input sex height;
cards;
1 72
0 70
1 69
;
proc means data=examp;
var sex;
run

```

The best option when Running... is come across is to click the circled **!** select option **2. Cancel Submitted Statements** then select yes. This will tell SAS to stop, then any correction can be made.

If when running SAS you cannot figure out the problem you can always re-launch SAS and start from scratch.

2.3 Title Line

The title line in fact has two places that it should go in the code. The title line can go after a data or proc statement is completed, or on the first line. Here are some examples:

```
data b;
input sex height;
cards;
1 72
0 70
1 69
;
run;
title 'howdy';
proc means data=b;
var height;
run;
```

or

```
title 'howdy';
data b;
input sex height;
cards;
1 72
0 70
1 69
;
run;
proc means data=b;
var height;
run;
```

3 Getting Descriptive Statistics from Temporary and Permanent Data Sets

3.1 Temporary Files

There are many options when using proc means. In it's most general form the official input for proc means is:

```
PROC MEANS <option(s)> <statistic-keyword(s)>;
  VAR variables;
  BY variables;
  CLASS variables;
```

```
FREQ variables;  
OUTPUT OUT=SASdataset keyword=names ...;
```

There are many keywords that can go in the 'option(s)' portion, they range from fixing the number of decimal points used, to setting the alpha on a t-test. The different statistical keywords that can be used are: CLM, RANGE, CSS, SKEWNESS—SKEW, CV, STDDEV—STD, KURTOSIS—KURT, STDERR, LCLM, SUM, MAX, SUMWGT, MEAN, UCLM, MIN, USS, N, VAR, NMISS, MEDIAN—P50, Q3—P75, P1, P90, P5, P95, P10, P99, Q1—P25, QRANGE, PROBT and T.

There is not much difference in the BY and CLASS statements. The main difference being that when CLASS is used the variables do not have to be sorted. For this reason I recommend using CLASS.

3.2 Permanent Files

3.2.1 Using the Editor

For some review permanent SAS files are created by the following code:

```
proc import datafile=P:\mclainexcel.xls out=excel replace;  
getnames=yes;  
run;  
libname abc P:\;  
data abc.perm;  
set excel;  
run;
```

Which will create a permanent SAS file in your P directory from the excel file 'mclainexcel'.

After the permanent SAS file is created it can be recalled using the following:

```
libname usc P:\;  
proc print data=usc.perm;  
run;  
proc means data=usc.perm;  
var Height;  
run;
```

Another way to load a permanent SAS file is by double clicking on the file, try this on the permanent SAS file in Blackboard. Once in SAS notice the name of the SAS file in the upper left corner, this will be the name you will use execute a proc.

In this next example 'Hanes' has been double clicked and the name TMP1.hanes has been assigned to it. Now proc means can be used on the variable 'sysbp' which is the variable name for 'Average Systolic BP' using the following:

```
proc means data=tmp1.hanes;  
var sysbp;  
run;
```

3.2.2 Using the Point and Click

While in the Editor Window click the pull down menu 'Solutions', then click the first option 'Analysis', and finally choose the 'Analyst' option. This should bring you to a window with a spreadsheet. Now load in your permanent SAS using the File pull down menu and Open. Now find you permanent SAS file in the directory, click on it, and select open.

Most of the descriptive statistics in proc means will be under Statistics→Descriptive→Summary Statistics.