THE IN-HOUSE TRAINING OF NON-COMPUTER ORIENTED RESEARCH PERSONNEL IN THE USE OF SAS

NEIL M. CASEY
MISSOURI DIVISION OF HEALTH
STATE CENTER FOR HEALTH STATISTICS

By non-computer oriented personnel I mean research analysts who are actively engaged in the collection, analysis and dissemination of health data. All are college graduates, some with advanced degrees, most of whom managed to traverse their entire college careers with little, if any, contact with the computer or computer programming. During the three year period that I have been affiliated with the Missouri Division of Health's Center for Health Statistics, we have had thirteen research analysts working at one time or another. We presently have nine analysts on our staff, seven of them are now able to at least write basic SAS programs, the other two have just joined our staff and will soon begin the training program which I will outline presently.

Let me state at the outset that throughout this paper, I will be referring only to SAS-72. However, I believe that this training program can be adapted to whatever SAS system your installation has. Also, this training program has been developed to meet the specific needs of our own office, but I feel that this program can be modified to meet the specific needs of your own shop.

Orientation to the computer and SAS

As I have just mentioned few of the newly hired research analysts have had any previous contact with the computers. Thus, I usually begin the training program by showing the new analyst some of our shop's previous computer runs, both of a quantitative and qualitative nature. While I am doing this I continually stress the number of cases involved in most of our work. To give you some idea of what I mean, Missouri has over the last 5 years been averaging approximately 70,000 births, 50,000 deaths, 50,000 marriages, and 20,000 divorces, and a new program just over a year old called the Missouri Hospital Discharge System, expects to average over half a million records per year. Thus, the new analyst is faced with a decision. Learn the SAS system or struggle along with pencil and paper while being ever mindful of impending deadlines.

Orientation to SAS will usually be through a short course on SAS given at the University of Missouri-Columbia. We have found that this course gives the analyst a good overview of SAS and it's capabilities. When they complete this short course, or if the new analyst begins his employment at an inconvenient time with relation to M.U.'s short course, I begin my own training program.
First, I instruct my new willing trainee to read the first 70 pages of the SAS manual and to become thoroughly familiar with the statements shown in Figure 1. The needs and requirements of each shop must be weighed in determining which statements to concentrate on.

**Statements:** DATA, INPUT, SET, CARD

**Program:**
- **Assignments:** DELETE, IF-THEN, subsetting IF, PUT, ERROR
- **Operators:** Use of () parenthesis, functions & operators
- **Data Manipulators:** Subsetting, Concatenation, Merging, Macros
- **Data Definers:** COMMENT, TITLE
- **Procedures:** SORT, PRINT

Figure 1

As you can see these are most but not all of the statements found in those first seventy pages. This is part of the focusing in I mentioned earlier concerning making this training program especially suited to our shop.

You will notice in Figure 1 we stress the basic DATA, INPUT, SET and CARD statements. The workings of the expression and assignment statements especially as to the placement of parenthesis (), functions and operators in mathematical formulae, and concentrate on the DELETE, IF-THEN, subsetting IF, PUT, and ERROR statements. We then move on to what our shop calls data manipulators: subsetting, concatenating, merging and macros. Since we often encounter changes in personnel we also stress the COMMENT and TITLE statements in order to have within job documentation and finally the new analyst learns Procedure SORT and Procedure PRINT.

After the new trainee feels he has mastered this section of the manual, I hand him a print-out of a program designed to emphasize the material he has just read. I usually specialize this program to meet the specific needs of the position to which the analyst has been employed. An example of one of these programs is given in Figures 2-7. Let me state here that our printer has only 48 characters so I have had to hand print over some of the erroneous symbols printed out.

Since our office makes frequent use of Macros, you can see in Figure 2 that I immediately set one up before I begin any input of data. I then input data set TEST1A from cards, 3 cards per case and as you can see from the input card, I
have the INPUT statement jump back and forth between cards, I do this to show the dexterity of the SAS INPUT statement. A simple print-out of the new data then follows. After this I create data TEST1 from set TEST1A, and make use of the MACRO and DELETE statement and print out the raw data. The new analyst should be able to explain why data TEST1 is one observation smaller than TEST1A and the sudden arrival of variables C and D in data TEST1.

In Figure 3 data TEST1 is sorted and titled and the trainee should be able to easily grasp this SORT. As a final manipulation of data TEST1, I do a selective print in order to acquaint the new analyst with the fact that he can print out his variables in any manner he chooses.

Figure 4 begins with the creation of a new data set from a direct access device, input positions of all the variables are the same as in data TEST1A. I make use of the DELETE and MACRO again. Assignment statements are used along with IF-THEN statements, variable X is created by use of SAS function UNIFORM. ASSIGN, ASSIGN2, LOOK and LOOK1 are used to show the importance of careful placement of parenthesis. ASSIGN and ASSIGN2 are identical with or without parenthesis, however, LOOK and LOOK1 are entirely different. This shows the new analyst that placement of parenthesis in mathematical operations must be done carefully. Since observation #3 has a B value equal to 234 it is listed out with the LIST statement. The data is printed, titled and then sorted, printed and titled.

Figure 5 shows how concatenation is done. Data TEST3 is a concatenation of data TEST1 and data TEST2. The DROP statement is used to delete variable A, however, variable A does not get printed out; the trainee should be able to explain this. The LIST statement doesn't work here because there is no INPUT statement preceding it, again the trainee should catch this fact, two PUT statements are used to show that the literal 'C=' is not necessary. The ERROR statement is also presented. This statement also introduces scientific E notation. Data TEST3 is then printed and in Figure 6 the unsorted data is printed using PROC PRINT and the BY variable sorter. Since the data is unsorted, each time a new value of sorter is encountered a new heading is placed on the printout and if you have 50,000 records that adds up to a lot of wasted paper. Figure 7 is the same PROC PRINT after the data are sorted using the BY variable sorter with the resultant neater, useable printout.

Figure 8 shows the necessity for careful work on merging data sets. Data TEST4 is an attempt to merge data TEST1 and data TEST2 by the unsorted variable C, with the resultant error. Data TEST5 merges data TEST1 and data TEST2 by the
previously sorted variable sorter and as you can see this time the merge succeeds. Also I feel that the values for merge and lastby should be printed so I create merge and last for this purpose.

After the trainee feels he has a good feel for what was on the sample program I set out to teach him the necessary input JCL to accomplish our data requests. You can see from Figure 9 that card input is fairly straight forward but that tape input depends on where the tape was originally created. Is it a DGS or a no label or standard label tape? Which file does the data reside on? etc.. It is my hope that your installation does not give you this problem. I explain to the trainee that we have documented the necessary tape JCL for all tapes currently at our installation and it is only when a new tape is sent down to the tape library that we have encountered problems reading the data file. You may have unique confusing areas which you will need to warn the trainee of and explain to him how to cope with them. After this rocky section I get to the heart of the training program.

I give the trainee an example of a simple data request not too unlike some of our actual requests and ask him to answer it. Such an example is given in Figure 10. As you can see the data needed is ASAP—as soon as possible. It is a simple request asking for a printout of variables listed in a certain order and grouped by county. The biggest error encountered here is to forget to sort by county first with the resultant printout looking somewhat like Figure 6. By the way, these examples are run on small sample tapes I created and not on the 70,000 record master file. Other than this situation the new analyst generally encounters little difficulty on this request and is elated to see some hard printout that he has generated.

The next step is to give the new analyst a second data request to complete. An example of this is shown in Figure 11. It not only asks for a line listing but also requests counts by sex, mean age of mother, percent of illegitimate births in each county by a grouped age of mother, and finally a separate line listing of children born to mother's born in a foreign country. This request is designed to test the trainee's ability to go to the SAS manual and learn the necessary procedures needed to complete it. I generally find the trainee taken aback by this request when he first sees it, but he soon learns to make the necessary commitment to the SAS manual in order to answer it. This request usually takes five to eight runs to complete and it's completion really gives the new analyst a sense of accomplishment.
As a final step in this training program I show the new analyst Figure 12. As you can see this program will destroy the three input tapes. You may have similar pitfalls you need to teach trainees to avoid. I explain this to the analyst and tell him never to name the data set the same as the input DDNAME tape name. Also, this is done to make the trainee aware that we can output SAS data files onto another tape for future recall. This is generally not necessary but was when five years of 70,000 birth records were extracted and concatenated.

Conclusion

At this point the analyst has completed the training program. Generally this program is completed in around three to four weeks. The analyst is then responsible for his own SAS jobs. He can come to me for suggestions and corrections only after he has made an honest attempt to solve his SAS dilemma himself.

The overall result of this training program has been very good. I believe this is because of the nature of the training, i.e., it is individualized to meet the needs of each new analyst and his assigned area of concentration. Real examples using real data are used which gives the analyst a true sense of accomplishment on the job. I encourage other shops to develop their own programs along this line remembering that it can be tailored to meet the needs of each individual within each shop.
CARD INPUT

//JOB CARD
//stepname EXEC SAS
//SAS.SYSIN DD *

TAPE INPUT

//JOB CARD
/*SETUP (xxxxxxx)
or 77xxxxx
//stepname EXEC SAS
//SAS.IN DD UNIT=TAPE, DISP = OLD,
  // LABEL = (1,SL,,IN),
or
  // LABEL = (N,SL,,IN),
or
  // LABEL = (1,NL,,IN),
or
  // LABEL = (2,NL,,IN),
or
  // LABEL = (1,BLP,,IN),
or
  // LABEL = (2,BLP,,IN),
or
// DSN=DGS (Generation #)
or
// VOL=SER=xxxxxxx, DSN=tape label name
or
// VOL=SER=xxxxxxx, DSN=dummy
// DCB=(LRECL=xx, BLKSIZE=xxxx, RECFM=x)

Figure 2
DATA TEST1

4 OBSERVATIONS IN DATA SET TEST1
2 VARIABLES

PROC PRINT, TITLE 'ORIGINAL RAW DATA';

ORIGINAL RAW DATA

OBS A R CUT UP NEW SORTER IDENT

<table>
<thead>
<tr>
<th>OBS</th>
<th>A</th>
<th>R</th>
<th>CUT UP</th>
<th>NEW</th>
<th>SORTER</th>
<th>IDENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>112</td>
<td>34.64</td>
<td>12.3</td>
<td>2</td>
<td>1</td>
<td>A1</td>
</tr>
<tr>
<td>2</td>
<td>214</td>
<td>30.47</td>
<td>17.4</td>
<td>6</td>
<td>1</td>
<td>A1</td>
</tr>
<tr>
<td>3</td>
<td>236</td>
<td>34.68</td>
<td>11.2</td>
<td>6</td>
<td>1</td>
<td>C1</td>
</tr>
<tr>
<td>4</td>
<td>174</td>
<td>45.47</td>
<td>17.3</td>
<td>3</td>
<td>1</td>
<td>B1</td>
</tr>
</tbody>
</table>

Figure 3
**PROC SORT; BY SORTER**

**TITLE "SORTED RAW DATA"**

**SORTED RAW DATA**

**PROC PRINT**

**SORTED RAW DATA**

<table>
<thead>
<tr>
<th>OBS</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ARCO</td>
<td>274</td>
<td>34.56</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>268.58</td>
<td>279.78</td>
</tr>
<tr>
<td>2</td>
<td>ARCO</td>
<td>112</td>
<td>34.56</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>146.56</td>
<td>158.56</td>
</tr>
<tr>
<td>3</td>
<td>ARCO</td>
<td>216</td>
<td>30.42</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>244.42</td>
<td>258.72</td>
</tr>
</tbody>
</table>

**PROC PRINT; VARIABLES A B C D; TITLE "SELECTIVE PRINT"**

**SELECTIVE PRINT**

<table>
<thead>
<tr>
<th>OBS</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ARCO</td>
<td>236</td>
<td>200.59</td>
<td>279.78</td>
</tr>
<tr>
<td>2</td>
<td>ARCO</td>
<td>112</td>
<td>146.56</td>
<td>158.56</td>
</tr>
<tr>
<td>3</td>
<td>ARCO</td>
<td>216</td>
<td>244.42</td>
<td>258.72</td>
</tr>
</tbody>
</table>

**Figure 4**
DATA TEST2: INPUT NONAME=IN A 1-4 B A-R CUT UP 9-12 Z

IF R=234 THEN LIST;
IF A=*ARC01 THEN DELETE;
AND

SUR=C-D;
X=UNIEND(0);
IF (X>2) THEN Y=1;
IF A<20 THEN TRIG=0(Y);
ASSIGN=ARC01/NEW+(SUR-R);
ASSIGN2=ARC01/NEW+SUB-R;
LOOK=LOOK23;
LOOK1=LOOK22;

OBSERVATION 3 ARC01234445A
A=112
B=02

3 OBSERVATIONS IN DATA SET TEST2　　　16 VARIABLES

PROC PRINT; TITLE 'UNSORTED TEST2 DATA';

UNSORTED TEST2 DATA

<table>
<thead>
<tr>
<th>OBS</th>
<th>A.</th>
<th>B.</th>
<th>CUT UP</th>
<th>NEW</th>
<th>SORTER</th>
<th>IDENT</th>
<th>C.</th>
<th>D.</th>
<th>SUR</th>
<th>X</th>
<th>Y</th>
<th>TRIG</th>
<th>ASSIGN</th>
<th>ASSIGN2</th>
<th>LOOK</th>
<th>LOOK1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ARC0112</td>
<td>34.56</td>
<td>12.3</td>
<td>2</td>
<td>A2</td>
<td>146.25</td>
<td>158.86</td>
<td>-12.3</td>
<td>6242765</td>
<td>0</td>
<td>1</td>
<td>2211879</td>
<td>211879</td>
<td>5.324985E+08</td>
<td>1.422985E+13</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ARC01216</td>
<td>30.62</td>
<td>12.3</td>
<td>4</td>
<td>A2</td>
<td>206.62</td>
<td>258.72</td>
<td>-12.3</td>
<td>653460</td>
<td>0</td>
<td>1</td>
<td>119350</td>
<td>211879</td>
<td>3.227085E+09</td>
<td>1.607962E+14</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ARC01234</td>
<td>34.56</td>
<td>11.2</td>
<td>1</td>
<td>C2</td>
<td>260.58</td>
<td>279.78</td>
<td>-11.2</td>
<td>6004760</td>
<td>1</td>
<td>1</td>
<td>1560713</td>
<td>1560713</td>
<td>4.533457E+09</td>
<td>1.422985E+14</td>
<td></td>
</tr>
</tbody>
</table>

PROC PRINT; TITLE 'SORTED TEST2 DATA';

SORTED TEST2 DATA

<table>
<thead>
<tr>
<th>OBS</th>
<th>A.</th>
<th>B.</th>
<th>CUT UP</th>
<th>NEW</th>
<th>SORTER</th>
<th>IDENT</th>
<th>C.</th>
<th>D.</th>
<th>SUR</th>
<th>X</th>
<th>Y</th>
<th>TRIG</th>
<th>ASSIGN</th>
<th>ASSIGN2</th>
<th>LOOK</th>
<th>LOOK1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ARC01234</td>
<td>34.56</td>
<td>11.2</td>
<td>1</td>
<td>C2</td>
<td>260.58</td>
<td>279.78</td>
<td>-11.2</td>
<td>6004760</td>
<td>1</td>
<td>1</td>
<td>1560713</td>
<td>1560713</td>
<td>4.533457E+09</td>
<td>7.429235E+14</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ARC0112</td>
<td>34.56</td>
<td>12.3</td>
<td>2</td>
<td>A2</td>
<td>144.46</td>
<td>168.86</td>
<td>-12.3</td>
<td>612276</td>
<td>0</td>
<td>1</td>
<td>2211879</td>
<td>211879</td>
<td>3.227085E+09</td>
<td>1.607962E+14</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ARC01216</td>
<td>30.62</td>
<td>12.3</td>
<td>4</td>
<td>A2</td>
<td>246.62</td>
<td>258.72</td>
<td>-12.3</td>
<td>693460</td>
<td>0</td>
<td>1</td>
<td>119350</td>
<td>211879</td>
<td>3.227085E+09</td>
<td>1.607962E+14</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5
**SORTED FILE DATA**

```
DATA TEST(-j-SET-TEST);
PROC SORT DATA=TEST;
```

```
CUT A LIST \*;\*;
```

```
R = 2.24880E+62
```

```
W = 2.24880E+62
```

```
LIMIT POSSIBLE Y = 0.000000000000
```

```
NOT POSSIBLE Y = 0.000000000000
```

**A. OBSERVATIONS IN DATA SET TESTS**

**B. VARIABLES**

```
PROC PRINT; TITLE "UNSORTED CONCATENATED DATA SETS";
```

**UNSORTED CONCATENATED DATA SETS**

<table>
<thead>
<tr>
<th>OBS</th>
<th>A</th>
<th>B</th>
<th>CUTF</th>
<th>NEW</th>
<th>SORTER</th>
<th>INIT</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ARC</td>
<td>244</td>
<td>34.54</td>
<td>12.3</td>
<td>1</td>
<td>51</td>
<td>274.28</td>
<td>274.28</td>
</tr>
<tr>
<td>2</td>
<td>ARC</td>
<td>117</td>
<td>34.54</td>
<td>12.3</td>
<td>2</td>
<td>51</td>
<td>144.54</td>
<td>159.78</td>
</tr>
<tr>
<td>3</td>
<td>ARC</td>
<td>216</td>
<td>34.54</td>
<td>12.3</td>
<td>3</td>
<td>52</td>
<td>259.72</td>
<td>259.72</td>
</tr>
<tr>
<td>4</td>
<td>ARC</td>
<td>216</td>
<td>34.54</td>
<td>12.3</td>
<td>1</td>
<td>52</td>
<td>259.72</td>
<td>259.72</td>
</tr>
<tr>
<td>5</td>
<td>ARC</td>
<td>117</td>
<td>34.54</td>
<td>12.3</td>
<td>2</td>
<td>52</td>
<td>144.54</td>
<td>159.78</td>
</tr>
<tr>
<td>6</td>
<td>ARC</td>
<td>216</td>
<td>34.54</td>
<td>12.3</td>
<td>4</td>
<td>52</td>
<td>259.72</td>
<td>259.72</td>
</tr>
</tbody>
</table>

Figure 6
### UNSORTED CONCATENATED DATA SETS

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACMO</td>
<td>234</td>
<td>54.58</td>
<td>13.7</td>
<td>C1</td>
</tr>
</tbody>
</table>

### SORTER=2

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACMO</td>
<td>112</td>
<td>345.8</td>
<td>12.3</td>
<td>A1</td>
</tr>
</tbody>
</table>

### SORTER=4

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACMO</td>
<td>216</td>
<td>40.42</td>
<td>12.3</td>
<td>B1</td>
</tr>
</tbody>
</table>

### SORTER=1

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACMO</td>
<td>234</td>
<td>34.58</td>
<td>11.2</td>
<td>C2</td>
</tr>
</tbody>
</table>

### SORTER=2

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACMO</td>
<td>112</td>
<td>345.8</td>
<td>12.3</td>
<td>A2</td>
</tr>
</tbody>
</table>

### SORTER=4

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACMO</td>
<td>216</td>
<td>40.42</td>
<td>12.3</td>
<td>B2</td>
</tr>
</tbody>
</table>

---

Figure 7
### PROC. SORT; BY SORTER;

#### INSORTED CONCATENATED DATA SETS

<table>
<thead>
<tr>
<th>OBS</th>
<th>A</th>
<th>B</th>
<th>CUT UP</th>
<th>NEW IDENT</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACOO</td>
<td>234</td>
<td>34.54</td>
<td>11.2</td>
<td>C1</td>
<td>23.858</td>
</tr>
<tr>
<td>2</td>
<td>ACOO</td>
<td>234</td>
<td>34.54</td>
<td>11.2</td>
<td>C2</td>
<td>23.858</td>
</tr>
</tbody>
</table>

#### SORTER=2

<table>
<thead>
<tr>
<th>OBS</th>
<th>A</th>
<th>B</th>
<th>CUT UP</th>
<th>NEW IDENT</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACOO</td>
<td>112</td>
<td>34.56</td>
<td>12.3</td>
<td>A1</td>
<td>144.54</td>
</tr>
<tr>
<td>2</td>
<td>ACOO</td>
<td>112</td>
<td>34.56</td>
<td>12.3</td>
<td>A2</td>
<td>144.54</td>
</tr>
</tbody>
</table>

#### SORTER=6

<table>
<thead>
<tr>
<th>OBS</th>
<th>A</th>
<th>B</th>
<th>CUT UP</th>
<th>NEW IDENT</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACOO</td>
<td>216</td>
<td>30.62</td>
<td>12.3</td>
<td>B1</td>
<td>244.62</td>
</tr>
<tr>
<td>2</td>
<td>ACOO</td>
<td>216</td>
<td>30.62</td>
<td>12.3</td>
<td>B2</td>
<td>244.62</td>
</tr>
</tbody>
</table>

---

Figure 8
### Sorted Concatenated Data Sets

**Data Tests:** Merge Tests 1, Tests 2
**Merge Order:** Last, Last-First

1 observation in data set TESTS
10 variables

**Proc Print, Title: Unsorted Merged Data Sets**

<table>
<thead>
<tr>
<th>DATA SET</th>
<th>C</th>
<th>D</th>
<th>X Y TRIG</th>
<th>ASSIGN ASSIGN</th>
<th>LOOK</th>
<th>LOOK</th>
<th>HEAD LAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>TESTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

### Sorted Merged Data Sets

**Data Tests:** Merge Tests 1, Tests 2
**Merge Order:** Last, Last-First

3 observations in data set TESTS
10 variables

**Proc Print, Title: Unsorted Merged Data Sets**

<table>
<thead>
<tr>
<th>DATA SET</th>
<th>C</th>
<th>D</th>
<th>X Y TRIG</th>
<th>ASSIGN ASSIGN</th>
<th>LOOK</th>
<th>LOOK</th>
<th>HEAD LAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>TESTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- Observations are not sorted in data set TESTS.
- Merged data sets are sorted.

**Sorted Merged Data Sets**
MISSOURI CENTER FOR HEALTH STATISTICS

DATA REQUEST

Requested by: Dr. Painkiller

Address: Section of Pain Relief

Telephone Number: 123 - 4567

Date Requested: 1/27/76

Date Needed: ASAP

Information Needed: (Give specific details on items, definitions, tables, etc.)

List of names and addresses of children born in 1974 by county, sorted by last name, suffix, first name, and middle initial.

Intended Use of Information:

Will use in mail survey to determine pain tolerance in 2 year olds.

Request Reviewed: ____________________________ Date: 1/27/76

Person Assigned: ____________________________ Date: ____________________________

Information Provided by: Phone □
Letter □

Published Material: Yes □
No □

Figure 10
MISSOURI CENTER FOR HEALTH STATISTICS

DATA REQUEST

Requested by: Dr. Painkiller
Date Requested: 2/27/76
Address: Section of Pain Relief
Date Needed: ASAP
Telephone Number: 123 - 4567

Information Needed: (Give specific details on items, definitions, tables, etc.)

List of names and addresses of children born in 1974. By county, sorted by last name, suffix, first name, middle initial. For each county would also like count of males and females, mean age of mother and percent of illegitimate births by age of mother grouped (<19, 20-24, 25-29, 30-34, 35-39, >40) and separate list of children of foreign born mothers.

Intended Use of Information:

Will use in conjunction with mail survey to determine pain tolerance of 2 year olds.

Request Reviewed: ____________________________ Date: 2/27/76
Person Assigned: ____________________________ Date: ____________

Information Provided by: Phone [ ] Letter [ ]
Published Material: Yes [ ] No [ ]

Figure 11
//JOB CARD
//SETUP (001234) (001235) (001236) (SCRATCH)
//stepname EXEC SAS
//SAS.IN1 DD UNIT = TAPE, VOL = SER = 001234, --
//SAS.IN2 DD UNIT = TAPE, VOL = SER = 001235, --
//SAS.IN3 DD UNIT = TAPE, VOL = SER = 001236, --
//SAS.OUT DD UNIT = TAPE, VOL = SER = SCRATCH, --
//SAS.SYSIN DD *

DATA IN1; INPUT DDNAME = IN1 ---
IF RACE = 3;
DATA IN2; INPUT DDNAME = IN2 ---
IF RACE = 3;
DATA IN3; INPUT DDNAME = IN3 ---
IF RACE = 3;
DATA OUT; SET IN1; SET IN2; SET IN3;
/

Figure 12