Macro to Combine Listing and Descriptive Statistics Table in One Report
Chenille Lloyd, PharmaNet/i3, Cary, NC

ABSTRACT
A comprehensive presentation of clinical trials data is often desired in order to achieve a greater understanding of the safety and efficacy of a new drug. The most valuable reports are able to provide more than one perspective of the data while maintaining readability. The primary objective of this paper is to generate a combined report of an individual subject listing followed by a descriptive statistics table using SAS/MACRO language and SAS/BASE. A single presentation that concatenates a listing and table is particularly useful for pharmacokinetic analysis studies which require easy identification of outliers and trends in drug concentration data. The macro features additional functionality by allowing the programmer to specify subject exclusions and the number of decimal places used to report results.

MACRO OVERVIEW
This macro is useful for presenting an individual listing and descriptive statistics table in the same report. To successfully use this macro, the following dataset is required: &INDATA which is the input dataset having the variables that will be used to generate the columns of the table.

Input dataset may need to be manipulated before calling the macro so that it is in the correct format of one observation per subject. The macro does not currently include the capability to check the format of the data. The sample data presented in Table 1 represents a typical pharmacokinetic parameter dataset for a Phase 1, three-way cross-over study.

For pharmacokinetic analysis, there will often be requests for particular subjects to be excluded from descriptive summarization. This macro allows the programmer to specify the subject ID of subject(s) to be excluded. Please note that the specified subjects will be excluded from the summary statistics, but will still be presented in the listing.

Summary statistics will be generated for the continuous variables in the input dataset so these variables must be named according to how many decimal places will be presented in the output. Last character of variable name designates decimal place precision, up to 5 decimal places. For example, if we want to see CMAX presented to 2 decimal places, the variable should be named cmax_2. Variable names cannot be greater than characters.

Future enhancements could include formatting results based on significant digits instead of decimal places.

OUR RAW DATA

<table>
<thead>
<tr>
<th>SUBJID</th>
<th>PERIOD</th>
<th>TRT</th>
<th>ANALYTE</th>
<th>AUCINF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1123</td>
<td>1</td>
<td>C</td>
<td>Analyte 1</td>
<td>9774.801</td>
</tr>
<tr>
<td>1123</td>
<td>2</td>
<td>A</td>
<td>Analyte 1</td>
<td>24285.600</td>
</tr>
<tr>
<td>1123</td>
<td>3</td>
<td>B</td>
<td>Analyte 1</td>
<td>24542.120</td>
</tr>
<tr>
<td>1168</td>
<td>1</td>
<td>C</td>
<td>Analyte 1</td>
<td>8934.827</td>
</tr>
<tr>
<td>1168</td>
<td>2</td>
<td>A</td>
<td>Analyte 1</td>
<td>23872.170</td>
</tr>
<tr>
<td>1168</td>
<td>3</td>
<td>B</td>
<td>Analyte 1</td>
<td>21131.390</td>
</tr>
<tr>
<td>1172</td>
<td>1</td>
<td>A</td>
<td>Analyte 1</td>
<td>31553.900</td>
</tr>
<tr>
<td>1172</td>
<td>2</td>
<td>B</td>
<td>Analyte 1</td>
<td>26097.270</td>
</tr>
<tr>
<td>1172</td>
<td>3</td>
<td>C</td>
<td>Analyte 1</td>
<td>10564.280</td>
</tr>
</tbody>
</table>

Table 1: Raw input dataset with one observation per treatment.
Macro to Combine Listing and Descriptive Statistics Table in One Report, continued

<table>
<thead>
<tr>
<th>SUBJIDC</th>
<th>TRTA_1</th>
<th>TRTB_2</th>
<th>TRTC_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1123</td>
<td>24285.600</td>
<td>24542.120</td>
<td>9774.801</td>
</tr>
<tr>
<td>1168</td>
<td>23872.170</td>
<td>21131.390</td>
<td>8934.827</td>
</tr>
<tr>
<td>1172</td>
<td>31553.900</td>
<td>26097.270</td>
<td>10564.280</td>
</tr>
</tbody>
</table>

Table 2: Input dataset (after manipulation) with one observation per subject.

MACRO PARAMETERS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Mandatory/Optional</th>
</tr>
</thead>
<tbody>
<tr>
<td>INDATA</td>
<td>Name of input dataset that has the variables required for the listing/table</td>
<td>Mandatory</td>
</tr>
<tr>
<td>SUBJVAR</td>
<td>Name of subject identifier</td>
<td>Mandatory</td>
</tr>
<tr>
<td>SUBJEXC</td>
<td>List subject identifier values to be included in listing and excluded from</td>
<td>Optional, if no</td>
</tr>
<tr>
<td></td>
<td>descriptive statistics table; multiple subject IDs to be separated by blank</td>
<td>subjects are</td>
</tr>
<tr>
<td></td>
<td>space</td>
<td>kept blank</td>
</tr>
<tr>
<td>LISTORDN</td>
<td>Name, order and precision of continuous variables from input dataset</td>
<td>Mandatory</td>
</tr>
<tr>
<td>OUTDATA</td>
<td>Name of output dataset including combined listing and table data that can</td>
<td>Mandatory</td>
</tr>
<tr>
<td></td>
<td>be fed directly to proc report</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Macro parameter descriptions

EXAMPLE OF HOW TO CALL MACRO

```sas
%mindstat (INDATA=inf,
          SUBJVAR=subjidc,
          SUBJEXC=1168,
          LISTORDN=trta_3 trtb_2 trtc_1,
          OUTDATA=outpk);
```

The above example macro call will output a SAS dataset named OUTPK that can be used in proc report to generate a customized listing/summary report.

MACRO OUTPUT DATASET

<table>
<thead>
<tr>
<th>CHARVAR</th>
<th>SUBORDER</th>
<th>Statistic</th>
<th># of Decimal Places</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1</td>
<td>Number of subjects</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td>2</td>
<td>Mean</td>
<td>+1</td>
</tr>
<tr>
<td>gm</td>
<td>3</td>
<td>Geometric Mean</td>
<td>+1</td>
</tr>
<tr>
<td>SD</td>
<td>4</td>
<td>Standard Deviation</td>
<td>+2</td>
</tr>
<tr>
<td>SE</td>
<td>5</td>
<td>Standard Error</td>
<td>+2</td>
</tr>
<tr>
<td>cv</td>
<td>6</td>
<td>Coefficient of Variation</td>
<td>2</td>
</tr>
<tr>
<td>Minimum</td>
<td>7</td>
<td>Minimum</td>
<td>Same # of decimal places</td>
</tr>
<tr>
<td>p25</td>
<td>8</td>
<td>25th percentile</td>
<td>+1</td>
</tr>
<tr>
<td>Median</td>
<td>9</td>
<td>Median</td>
<td>+1</td>
</tr>
<tr>
<td>p75</td>
<td>10</td>
<td>75th percentile</td>
<td>+1</td>
</tr>
<tr>
<td>Maximum</td>
<td>11</td>
<td>Maximum</td>
<td>Same # of decimal places</td>
</tr>
</tbody>
</table>

Table 4: Precision and order of each summary statistic in output dataset
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHARVAR</td>
<td>Character variable populated for listing and table output; for listing subject id is presented; for table summary statistics label is presented</td>
</tr>
<tr>
<td>NUMVARx</td>
<td>Character variable populated for listing and table output and formatted to decimal place precision specified in macro call; for listing actual values are presented; for table statistics are presented. X corresponds to order of variables in output dataset</td>
</tr>
<tr>
<td>SUBORDER</td>
<td>Numeric variable populated for table output only; gives order of summary statistics in table</td>
</tr>
<tr>
<td>PAGENUM</td>
<td>Numeric variable populated for listing and table output; flags listing output from table output; can be used to separate listing and table sections in proc report by line or by page</td>
</tr>
</tbody>
</table>

Table 5: Output dataset variables

<table>
<thead>
<tr>
<th>CHARVAR</th>
<th>NUMVAR1</th>
<th>NUMVAR2</th>
<th>NUMVAR3</th>
<th>SUBORDER</th>
<th>PAGENUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1123</td>
<td>24285.6</td>
<td>24542.12</td>
<td>9774.801</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1168</td>
<td>23872.2</td>
<td>21131.39</td>
<td>8934.827</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1172</td>
<td>31553.9</td>
<td>26097.27</td>
<td>10564.280</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>N</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Mean</td>
<td>27919.75</td>
<td>25319.695</td>
<td>10169.5405</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>gm</td>
<td>27682.22</td>
<td>25307.752</td>
<td>10161.8765</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>SD</td>
<td>5139.464</td>
<td>1099.6571</td>
<td>558.24595</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>SE</td>
<td>3634.150</td>
<td>777.5750</td>
<td>394.73950</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>cv</td>
<td>18.41</td>
<td>4.34</td>
<td>5.49</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Minimum</td>
<td>24285.6</td>
<td>24542.12</td>
<td>9774.801</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>p25</td>
<td>24285.60</td>
<td>24542.120</td>
<td>9774.8010</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Median</td>
<td>27919.75</td>
<td>25319.695</td>
<td>10169.5405</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>p75</td>
<td>31553.9</td>
<td>26097.270</td>
<td>10564.2800</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Maximum</td>
<td>31553.9</td>
<td>26097.27</td>
<td>10564.280</td>
<td>11</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6: Output dataset with listing of one observation per subject and summary statistics

CONCLUSION

This paper has described how a macro can combine listing and table output into one report. This combined report allows for more efficient cross-checking of data points and contributes to a more automated programming process overall. The output dataset generated by the macro easily feeds into the SAS report procedure where the report formatting can be customized for the particular clinical trial.

REFERENCE

- SAS/BASE Software: Version 9, SAS® Institute Inc., Cary NC

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Chenille Lloyd
Enterprise: PharmaNet/i3
Address: 1001 Winstead Drive, Suite 200
City, State ZIP: Cary, NC 27513
Work Phone: 919-677-5985
Email: clloyd@pharmanet-i3.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration. Other brand and product names are trademarks of their respective companies.
%macro mindstat (INDATA=, SUBJVAR=, SUBJEXC=, LISTORDN=, OUTDATA=);

***********************************************************************;
* LISTING                                                             *
***********************************************************************;

** Count number of numeric variables;
%let i=1;
%do %while (%scan(&listordn,&i) ne %str());
  %let _var&i = %scan(&listordn,&i);
  %let i = %eval(&i + 1);
%end;
%let pnum = %eval(&i - 1);

data list1;
  set &indata;
  keep &subjvar &listordn;
run;

data LISTOUT;
  length charvar $20;
  ** Create column(s) for numeric vars;
  %do j = 1 %to &pnum;
    length numvar&j $20 dec&j $8;
  %end;
  set list1;
  charvar = compress(&subjvar);
  drop &subjvar;
  ** Display numeric vars with correct number of decimal places;
  %do j = 1 %to &pnum;
    call vname(%scan(&listordn,&j),dec&j);
    if substr(right(dec&j),8,1) = '1' then numvar&j = put(%scan(&listordn,&j),8.1);
    else if substr(right(dec&j),8,1) = '2' then numvar&j = put(%scan(&listordn,&j),9.2);
    else if substr(right(dec&j),8,1) = '3' then numvar&j = put(%scan(&listordn,&j),10.3);
    else if substr(right(dec&j),8,1) = '4' then numvar&j = put(%scan(&listordn,&j),11.4);
    else if substr(right(dec&j),8,1) = '5' then numvar&j = put(%scan(&listordn,&j),12.5);
    else if substr(right(dec&j),8,1) not in ('1','2','3','4','5') then numvar&j = put(%scan(&listordn,&j),6.);
    drop dec&j %scan(&listordn,&j);
  %end;
run;

***********************************************************************;
* TABLE                                                                *
***********************************************************************;

** Count number of subjects excluded from stats;
%let i = 1;
%do %while (%scan(&subjexc,&i) ne %str());
  %let _var&i = %scan(&subjexc,&i);
  %let i = %eval(&i + 1);
**Exclude specified subjects from stats only:**
%let enum = %eval(&i - 1);

**Calculate stats:**
proc transpose data=&indata out=stat1(rename=(COL1=val _NAME_=parname));
   by &subjvar;
   var &listordn;
run;

data temp;
   set stat1;
   if val > 0 then ln=log(val);
run;

proc means data=stat1 nway missing noprint;
   class parname;
   var val;
   output out = stat2
      n = N
      mean = Mean
      median = Median
      min = Minimum
      max = Maximum
      std = SD
      stderr = SE
      cv = cv
      p25 = p25
      p75 = p75;
run;

proc means data=temp nway missing noprint;
   class parname;
   var ln;
   output out = gm1
      mean = gmean;
run;

data gm1;
   set gm1;
   gm = exp(gmean);
   keep parname gm;
run;

data stat3;
   merge stat2
gm1;
   by parname;
run;

proc sort data=stat3;
   by parname;
run;

proc transpose data=stat3 out=stat4;
   by parname;
   var N Mean Median Minimum Maximum SD SE CV P25 P75 GM;
run;

data stat5;
   length statname $8 statval $20;
   set stat4;
   statname = _name_; 
   _name_ = upcase(_name_);
   select (_name_);
      when ('N') suborder = 1;
      when ('MEAN') suborder = 2;
      when ('GM') suborder = 3;
      when ('SD') suborder = 4;
      when ('SE') suborder = 5;
      when ('CV') suborder = 6;
      when ('MINIMUM') suborder = 7;
      when ('P25') suborder = 8;
      when ('MEDIAN') suborder = 9;
      when ('P75') suborder = 10;
      when ('MAXIMUM') suborder = 11;
   otherwise; end;
   if substr(right(parname),8,1) = '1' then do;
      if suborder = 1 then statval = put(col1,6.);
      else if suborder in (7,11) then statval = put(col1,8.1);
      else if suborder in (2,3,8,9,10) then statval = put(col1,9.2);
      else if suborder in (4,5) then statval = put(col1,10.3);
      else if suborder in (6) then statval = put(col1,9.2);
   end;
   else if substr(right(parname),8,1) = '2' then do;
      if suborder = 1 then statval = put(col1,6.);
      else if suborder in (7,11) then statval = put(col1,9.2);
      else if suborder in (2,3,8,9,10) then statval = put(col1,10.3);
      else if suborder in (4,5) then statval = put(col1,11.4);
      else if suborder in (6) then statval = put(col1,9.2);
   end;
   else if substr(right(parname),8,1) = '3' then do;
      if suborder = 1 then statval = put(col1,6.);
      else if suborder in (7,11) then statval = put(col1,10.3);
      else if suborder in (2,3,8,9,10) then statval = put(col1,11.4);
      else if suborder in (4,5) then statval = put(col1,12.5);
      else if suborder in (6) then statval = put(col1,9.2);
   end;
   else if substr(right(parname),8,1) = '4' then do;
      if suborder = 1 then statval = put(col1,6.);
      else if suborder in (7,11) then statval = put(col1,11.4);
      else if suborder in (2,3,8,9,10) then statval = put(col1,12.5);
      else if suborder in (4,5) then statval = put(col1,13.6);
      else if suborder in (6) then statval = put(col1,9.2);
   end;
   else if substr(right(parname),8,1) = '5' then do;
      if suborder = 1 then statval = put(col1,8.);
      else if suborder in (7,11) then statval = put(col1,12.5);
      else if suborder in (2,3,8,9,10) then statval = put(col1,13.6);
      else if suborder in (4,5) then statval = put(col1,14.7);
else if suborder in (6) then statval = put(col1,9.2);
end;
else if substr(right(parname),8,1) not in ( '1','2','3','4','5') then do;
   if suborder = 1 then statval = put(col1,6.);
   else if suborder in (7,11) then statval = put(col1,6.);
   else if suborder in (2,3,8,9,10) then statval = put(col1,8.1);
   else if suborder in (4,5) then statval = put(col1,9.2);
   else if suborder in (6) then statval = put(col1,9.2);
end;
if statval = ' ' then statval = '   ';
keep parname statname statval COL1 suborder;
run;

proc sort data=stat5;
   by suborder statname;
run;

proc transpose data=stat5 out=stat6(drop=_name_);
   by suborder statname;
   id parname;
   var statval;
run;

data STATOUT;
   length charvar $20;
   * Create column(s) for numeric vars;
   %do j = 1 %to &pnum;
      length numvar&j $20;
   %end;
   set stat6;
   * Stat names will always go under subject identifier;
   charvar = statname;
   * Assign parameter stat values to variables in specified order;
   %do j = 1 %to &pnum;
      numvar&j = %scan(&listordn,&j);
   %end;
   drop statname &listordn;
run;

** Add blank line between listing and stats;
data BLANK;
   length charvar $20;
   %do j = 1 %to &pnum;
      length numvar&j $20;
   %end;
   charvar = ' ';
   %do j = 1 %to &pnum;
      numvar&j = ' ';
   %end;
   output;
run;

** Combine listing and stats;
data &outdata;
   set listout (in=a)
      blank (in=b)
      statout (in=c);
   if a or b then pagenum=1;
   else if c then pagenum=2;
run;
%mend mindstat;