D is for Dynamic, putting Dynamic back into CDISC (A simple macro utilizing PROC SQL which auto-formats continuous variables for lab tables).

Steven Black, SynteractHCR, Carlsbad, CA

ABSTRACT

As CDISC implementation increases, the ability and need for simple dynamic code has also risen. Using PROC SQL and the SAS MACRO language coupled with proper use of ADaM guidelines, generating dynamic code has become much easier. This paper illustrates one way of auto generating the decimal format needed for tables where continuous data is arranged by parameters. This code has been especially helpful with laboratory table values where the decimal formats change between each parameter. Understanding and using these same techniques and framework, additional dynamic code can be created allowing you to put the D back into CDISC.

INTRODUCTION

Recently I came across a situation where the decimal formatting test results needed to be based on the format of the data as it was displayed within the AVAL variable. In the past (non ADaM) datasets I could use the format specified for the specific variable needed, such as systolic blood pressure or height. Now, however, especially with lab data, the variables are stacked one on top of another, and the decimal format varies within each parameter.

Manual Solution: I could look and see for each parameter what the decimal place was and then write some code per each parameter or group those parameters with similar decimal types and hardcode their respective decimal places needed per statistic as defined by the statistical analysis. It would take some time but it can be done.

However, if there is a logically consistent pattern to it, I can macrotize the program and create reusable code. For this example the Statistical Analysis Plan (SAP) had an essential general statement stating: ‘Generally, means, medians and standard deviations will be presented with an increased level of precision as follows: means and medians will be presented to one more decimal place than the raw data, and the standard deviations will be presented to two more decimal places than the raw data.’ This was the key to making this all work, otherwise hardcoding may be a better alternative or additional conditional macro language could be employed.

This paper will be divided up into several steps to illustrate the process and provide the background information needed for each step. The intent of this paper is to provide a succinct example of code and to provide enough information per each step to allow others to understand and then apply this knowledge to meet their own needs.

GETTING STARTED

First I need to know what I want the table to look like in the end. In my case, I knew that I need a table generated for each lab category (hematology, chemistry, etc.) and within each of those lab categories I would have the specific lab parameters (WBC, Basophils, AST, ALT, etc.), separated out by visit and treatment group.

<table>
<thead>
<tr>
<th>Parameter(s)</th>
<th>Data Final</th>
<th>WBC</th>
<th>ALAESA</th>
<th>BETA</th>
</tr>
</thead>
<tbody>
<tr>
<td>WBC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160-360</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Max</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>End of Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160-360</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Max</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Change from Baseline to EOT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160-360</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Max</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Basophils</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160-360</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Max</td>
<td></td>
<td>xx</td>
<td>xx</td>
<td>xx</td>
</tr>
</tbody>
</table>
Second I need to know what data I have to work with. I know that with the ADLB dataset, the is data set up as a very long dataset with each of the laboratory categories and lab parameters set on top of one another, as seen below. I also know that the variable that I am actually analyzing (AVAL) will have a number of different decimal places per each parameter, and that the resulting tables will also have different decimal places or formats per each statistic (mean, median, standard deviation).

<table>
<thead>
<tr>
<th>Category for Lab Test</th>
<th>Parameter Description</th>
<th>Parameter Code</th>
<th>Analysis Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>RDW (%)</td>
<td>14</td>
<td>159</td>
</tr>
<tr>
<td>101</td>
<td>RDW (%)</td>
<td>14</td>
<td>159</td>
</tr>
<tr>
<td>HRAS</td>
<td>HBATC (%)</td>
<td>15</td>
<td>0.56</td>
</tr>
<tr>
<td>HRATC</td>
<td>HBATC (%)</td>
<td>15</td>
<td>0.56</td>
</tr>
<tr>
<td>CHEMISTRY</td>
<td>ALT (U/L)</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>CHEMISTRY</td>
<td>ALT (U/L)</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>CHEMISTRY</td>
<td>ALK Phosphatase (U/L)</td>
<td>18</td>
<td>107</td>
</tr>
<tr>
<td>CHEMISTRY</td>
<td>ALK Phosphatase (U/L)</td>
<td>18</td>
<td>107</td>
</tr>
<tr>
<td>CHEMISTRY</td>
<td>Albumin (g/L)</td>
<td>18</td>
<td>35</td>
</tr>
<tr>
<td>CHEMISTRY</td>
<td>Albumin (g/L)</td>
<td>18</td>
<td>35</td>
</tr>
<tr>
<td>CHEMISTRY</td>
<td>GGT (U/L)</td>
<td>20</td>
<td>54</td>
</tr>
<tr>
<td>USUBJID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBCAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRT01PN or TRT01AN</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Variables that we will use in the macro from our ADLB dataset: PARAMN, AVAL, AVISITN, USUBJID, LBCAT, and TRT01PN or TRT01AN).

Prior to the Macro and PROC SQL, I do a few minor clean-up steps to the ADLB data and call this new dataset _PRE_FREQ.

**STEP 1 – FIND THE DECIMALS WITH PROC SQL**

A brief introduction to PROC SQL: Because of many useful properties of PROC SQL it has become essential to building dynamic SAS programs. Using PROC SQL allows you to simplify and complete many steps within one call. In this paper I will touch on a number of amazing abilities of this procedure but, as in all things in SAS, there is an iceberg of possibilities. PROC SQL uses a bit more English words (underlined) than normal datastep processing.

Let's start simple:

```sas
proc sql;
create table _decil1 as
select paramn, aval
from _pre_freq;
quit;
```

So in this quick step we created a new dataset (or table) and called it _DECIL1, then we keep or selected the variables PARAMN and AVAL, from source dataset _PRE_FREQ. We then end or quit the procedure. Variables are separated by commas and we only put in semi colons when we are done with the statement. This allows us to use multiple statements in one procedure which we will make use of this capability later in the macro.

Next we'll use some SAS® functions on the AVAL variable to capture the length of the decimal places used of for each value.

```sas
create table _decil as
select paramn, aval ,strip(scan(put(aval,best.),2,'.')) as deci
```

I used a PUT statement to convert AVAL to character then used a SCAN function to look for the text after the decimal. In the datastep we would have done something similar with the code “DECI= strip(…” in PROC SQL we use the term ‘as’ instead of the equal sign. The new variable now looks like this

<table>
<thead>
<tr>
<th>Parameter Code</th>
<th>Analysis Value</th>
<th>DECI</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>31</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>32</td>
<td>119473 19473</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>1 00320 03329</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>0.9997 8997</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>1 42076 42076</td>
<td></td>
</tr>
</tbody>
</table>

Next let's create a another variable which captures the length of the variable DECI, and call it LENGTH, but this time we need to add some additional logic and use this newly created variable DECI. To accomplish this we will use a CASE Expression.
D is for Dynamic, putting Dynamic back into CDISC, continued

Using a CASE expression allows us to add ‘when then’ logic to our procedure. In this case, we are calculating the length of the variable DECI but first want to modify the new variable based on the presence or absence of data. Again it’s near English in code: when DECI equals null then set this new variable to 0, else this new variable = length of DECI, at the end we specify what this new variable is to be called again using the ‘AS’ text followed by the new variable name (LENGTH in this case).

The keyword CALCULATED is needed due to the way PROC SQL process the data. The variable DECI is not contained in the _PRE_FREQ dataset (it is being created within the procedure), so to access this new variable we add the keyword CALCULATED prior to calling the variable. This way PROC SQL knows not to look in the dataset but in the temporary running dataset.

```sql
proc sql;
create table _deci1 as
    select paramn, aval ,strip(scan(put(aval,best.),2,'.')) as deci,
    case when calculated deci='' then 0
    else length(calculated deci)
    end as length
    from _pre_freq;
quit;
```

### STEP 2 – CAPTURE THE MAX LENGTH USED PER PARAMETER

Now we need to find out what the maximum length is per parameter, again we can do this quite efficiently using PROC SQL and do it within the same procedure. We will use the same style of programming to adjust some of the max lengths to fit our needs. This next code would fit between the last semi colon and the quit statement of the previous code segment above.

```sql
create table _deci2 as
    select paramn, max(length) as maxl,
    case when calculated maxl >3 then 4
    else calculated maxl
    end as max_length
    from _deci1
    group by paramn;
```

Here we need to create a new table called _DECI2 where we are selecting the variable PARAMN, and creating a new variable which captures the maximum value of length called MAXL. We again use the CASE expression to create a modified version of that maximum length variable resetting it to 4 if the value is greater than 3, This modification is done as to present the data in a more visually appealing way as it is difficult to fit and view decimal places greater than 4. This new variable is called MAX_LENGTH. We are now pulling the data from _DECI1 and use a ‘group by’ term to capture the max length per each PARAMN. The results of this new table look like this:

<table>
<thead>
<tr>
<th>Parameter (N)</th>
<th>MAX LENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>4</td>
</tr>
<tr>
<td>36</td>
<td>4</td>
</tr>
</tbody>
</table>

### STEP 2 – CAPTURE THE MAX LENGTH USED PER PARAMETER

Now we need to find out what the maximum length is per parameter, again we can do this quite efficiently using PROC SQL and do it within the same procedure. We will use the same style of programming to adjust some of the max lengths to fit our needs. This next code would fit between the last semi colon and the quit statement of the previous code segment above.

```sql
create table _deci2 as
    select paramn, max(length) as maxl,
    case when calculated maxl >3 then 4
    else calculated maxl
    end as max_length
    from _deci1
    group by paramn;
```

Here we need to create a new table called _DECI2 where we are selecting the variable PARAMN, and creating a new variable which captures the maximum value of length called MAXL. We again use the CASE expression to create a modified version of that maximum length variable resetting it to 4 if the value is greater than 3, This modification is done as to present the data in a more visually appealing way as it is difficult to fit and view decimal places greater than 4. This new variable is called MAX_LENGTH. We are now pulling the data from _DECI1 and use a ‘group by’ term to capture the max length per each PARAMN. The results of this new table look like this:

<table>
<thead>
<tr>
<th>MAX LENGTH</th>
<th>Parameter (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>1</td>
<td>31</td>
</tr>
<tr>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
</tr>
</tbody>
</table>

3
STEP 3 – CAPTURE THE NUMBER OF TYPES OF DECIMAL PLACES

In ADLB data there are many parameters and in these past two steps we calculated the max decimal place per parameter but we’ll also need to categorize parameters into types, so we can we can loop through of decimal lengths. To accomplish all this we can again use PROC SQL and keep this code all within the same procedure we have created. We can place this code right before the quit statement and the last semi colon.

\[ \text{select distinct(max_length) into: type1 :: type5} \]
\[ \text{from _deci2;} \]
\[ \%let n_types=&sqlobs; \]

In the above step we are selecting the variable MAX_LENGTH from _DECI2 but this time we only want the DISTINCT values of the variable, similar to performing a “first dot” Or a “nodupkey” statement but it’s not important if it’s the first or the last observation, ensuring that there are no other values like it in the dataset. Secondly we are using PROC SQL’s amazing ability to create MACRO variables. We accomplish this by using the keyword INTO in conjunction with a colon as “INTO:”. We then create our new macro variable names, in this case ranging from 1 to 5. Because of my limiting the number of decimals to 4 I know that I’ll not get above 5, you can also set your max number to 99 or any numerical value. I will then create macro variables type1 – type5. I want to know my highest number that the macro resolves to so I use the temporary variable &SQLOBS and store this in a permanent macro variable called N_TYPES using a %LET statement. The values of these macro variables automatically get pushed out to the output window if data in the output window is not wanted the NOPRINT option can be used, placed next to PROC SQL.

Output Window Result of MAX_LENGTH

| 0 | 1 | 2 | 3 | 4 |

STEP 4 – GROUPING THE PARAMETERS INTO DECIMAL TYPES

Now that we know the range of decimals available and we have them stored as macro variables we can create the decimal formats for all grouped parameters at the same time. To do this we will further utilize the power of PROC SQL and the MACRO language. We will also utilize a macro DO loop to iterate through each of the decimal types. Macro DO loops need to be run within a macro, so we will place our ever growing single PROC SQL procedure within a macro called DYNAMIC.

Basic MACRO principles: %MACRO MACRO_NAME starts a macro, %MEND ends the macro, %MACRO_NAME runs/executes the macro. Generally macro variables created within the macro only live within that specific macro (local) unless otherwise specified using %global option. Macro DO loops have % before each key word (%do, %end, %to, %then, %else). A %DO needs a corresponding %END statement. Macro variables that are being resolved more than once need multiple ampersands.

%macro dynamic;
\text{proc sql;}\text{ create table _deci1 as} \ldots \text{ %let n_types=&sqlobs;} \text{ %do x= 1 %to &n_types;}\text{ select paramn into: mparams&x separated by ',',} \text{ from _deci2 having max_length = &&type&x;} \text{ %end;} \text{ %quit;} %mend;

I use a macro DO loop to move through each of the types of decimal places found using the &n_types macro variable created in the step above and using X as my macro counter variable. Within the loop I use the similar PROC SQL code for selecting the variables wanted and creating a new macro variable mparams&x which will be equal to mparams1 –mparams5 once resolved. I also use the SEPARATED BY statement to allow me to capture all of the
values of the PARAMN variable into one macro term. I further specify that I want these values separated by a comma. I establish where I am pulling my data from and also incorporate a HAVING statement. This HAVING statement allows for the limiting of groups of data based on the logic used. The WHERE clause works on individual rows of data, but in our example we need the PARAM to be grouped into groups based on the max length, so the HAVING clause works perfect.

Per each loop the macro variable &type&x is resolved twice, for x=1 first it is resolved from &type&x to &type1 then &type1 resolves to 0. When the loop completes, the macro will end and the new mparams1 - mparams5 macro variables will contain all the parameters per decimal group and be available to use.

**STEP 5 – CREATE STATISTICAL OUTPUT**

Once we have created these macro variables based off of the data in the ADLB dataset we can now calculate the statistics needed, per the SAP. This can be done using PROC UNIVARIATE, PROC MEANS, PROC TTEST, even PROC SQL! Once data from the procedure has been output we can create dynamic formats. Below is a basic example of a PROC UNIVARIATE outputting to a new dataset _STAT_V1. Note that the data had been previously sorted by PARAMN,TXGROUP and AVISITN.

```
proc univariate data=_pre_freq noprint;
class txgroup avisitn;
var aval;
output n=n mean=mean median=median std=std min=min max=max out=_stat_v1;
by paramn;
where avisitn in (2,10);
run;
```

### Resulting Data from Statistical Output

<table>
<thead>
<tr>
<th>Parameter (N)</th>
<th>TXGROUP</th>
<th>Analysis Tempers</th>
<th>number of nominating values</th>
<th>AVAL</th>
<th>the mean, AVAL</th>
<th>the standard deviation, AVAL</th>
<th>the largest value, AVAL</th>
<th>the median, AVAL</th>
<th>the smallest value, AVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>38</td>
<td>0.2506316</td>
<td>2.021738153</td>
<td>12.9</td>
<td>6.2</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>10</td>
<td>36</td>
<td>0.5422682</td>
<td>2.406480148</td>
<td>14.6</td>
<td>8.15</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>27</td>
<td>7.9404286</td>
<td>1.060641534</td>
<td>12.1</td>
<td>7.0</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>10</td>
<td>25</td>
<td>1.25842387</td>
<td>1.258423872</td>
<td>11.9</td>
<td>7.7</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>20</td>
<td>0.4268442</td>
<td>0.4268442</td>
<td>2.6</td>
<td>0.64</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>10</td>
<td>36</td>
<td>0.4659538</td>
<td>0.4659538</td>
<td>9.8</td>
<td>0.64</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>37</td>
<td>0.5622733</td>
<td>0.5622733</td>
<td>1.5</td>
<td>0.63</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>10</td>
<td>35</td>
<td>0.2914285</td>
<td>0.2914285</td>
<td>1</td>
<td>0.63</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>38</td>
<td>2.0068316</td>
<td>2.121549003</td>
<td>5.4</td>
<td>1.18</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>10</td>
<td>36</td>
<td>2.3388189</td>
<td>1.873878653</td>
<td>7.6</td>
<td>1.9</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>37</td>
<td>2.5856404</td>
<td>2.460867789</td>
<td>12.3</td>
<td>1.66</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>10</td>
<td>35</td>
<td>2.3814285</td>
<td>2.145012356</td>
<td>19.6</td>
<td>1.7</td>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>

**STEP 6 – FINALIZE THE FORMATS FROM STATISTICAL OUTPUT**

In this last step we will set the statistical output datasets together and then create needed variables in proper order so that we can later transpose them as per the table shell. Much of this is dependent upon the style of the programmer however the basic principles remain. In my tables I create a Y_AXIS and X_AXIS variable these variables assist in how I transpose the data later. Y_AXIS will determine the order of my rows and X_AXIS determines the order of my columns. Below is the basic code used:

```
data _means (keep=paramn avisitn y_axis display x_axis); length display $20;
set _stat_v1 (in=a);
*** create x_axis ***;
x_axis=txgroup;

*** create y_axis for var1 ***;
y_axis=10;
display=compress(put(n,best.));
output;
y_axis=11;
display= strip(put(mean,10.1))||" ("||strip(put(std,10.2))||")";
output;
y_axis=12;
display=compress(put(median,10.1));
output;
y_axis=13;
```

5
D is for Dynamic, putting Dynamic back into CDISC, continued

```
display= strip(put(min,10.0))||", "||strip(put(max,10.0));
output;
run;
```

In this step I create the X_AXIS and Y_AXIS variables. The display variable contains the formatted values of the various statistics needed with a preset length of 20. These variables are output to the _MEANS dataset.

Now let’s add some dynamics: First we will need to loop through each of the decimal types using a macro DO loop, then we will limit the processing to only those parameters that meet that decimal type. Finally we will modify the formats used in the PUT statements based on the type used. All this code would be placed inside the %dynamic macro.

```
data _means; length display $20;
set _...
%do x= 1 %to &n_types;
  if paramn in (&&mparams&x) then do;
    y_axis=10;
    display=compress(put(n,best.));
    output;
    y_axis=11;
    display=strip(put(mean,10.%eval(&&type&x+1)))||" (||
        strip(put(std,10.%eval(&&type&x+2)))||")";
    output;
    y_axis=12;
    display=compress(put(median,10.%eval(&&type&x+1)));
    output;
    y_axis=13;
    display=strip(put(min,10.&&type&x))||", "|| strip(put(max,10.&&type&x));
    output;
  end;
%end;
run;
%mend;
%dynamic;
```

We use the macro variable &&mparams&x which equal a list of the parameter values per type of decimal count. Then using the values for &&type&x (values:0-4) we can determine the decimal places for each of the put statements. Using the SAP as a guide, we place the mean and median decimal place up one from the original data, standard deviation (STD) is +2, whereas the min and max equal the value found in the data. Once the macro loops through each type all parameters will have the correct format applied and we can move on to transposing the data.

**STEP 7 - MAKE IT DYNAMIC**

Throughout the program we have referenced different dataset names and variables names that we can now turn into macro variables and specify them as parameters in a macro call. We have already done much of the work when we created the %DYNAMIC Macro. We will now create with the macro parameters at the top and bottom of the macro.

```
%macro dynamic (orig_data, stat_data, out_data, txgroup)
...
%mend;
%dynamic (_pre_freq, %str(_stat_v1 (in=a)), _means, trt01pn);
```

In this example we are creating a new macro variable called ORIG_DATA and setting the value of this new parameter to _PRE_FREQ. In some cases character strings such as ‘=,%,’ may be wanted within the macro parameter, using %STR option will mask these character strings during compilation of a macro. Next I’ll replace those specific parameters with the macro variable assigned. Now when I run the %DYNAMIC macro I have the ability to modify the parameters to my own preference and still be able to run it.

If further values need to be run such change from baseline we can simply replace the AVAL variable with a macro.
CONCLUSION

This paper presents a useful macro for dynamically creating formats for lab tables based off of the data found in the AVAL variable per parameter. It also includes specific information regarding the benefits of using PROC SQL combined with the MACRO language. By using the information and SAS code in this paper it is hoped that many other dynamic applications can be created.

REFERENCES


ACKNOWLEDGMENTS

Many thanks to Priya Venkata and Kurtis Voris for their willingness to listen, and provide valuable insight and feedback.

CONTACT INFORMATION

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

Steve Black
SynteractHCR
5759 Fleet St. Suite 100
Carlsbad, CA 92081
760-268-8015
steven.c.black@gmail.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.

APPENDIX

Full code used:

```
%macro dynamic (orig_data, stat_data, out_data, txgroup);
```
*** use sql to get max decimal places per paramn ***;

proc sql;

*** capture length of data after decimal place ***;

create table _deci1 as
select paramn, aval ,strip(scan(put(aval,best.,2,'.')) as deci,
case when calculated deci='' then 0
else length(calculated deci)
end as length
from &orig_data;

*** capture max length per paramn if >3 then set to 4 ***;

create table _deci2
as select max(length) as maxl, paramn,
case when calculated maxl > 3 then 4
else calculated maxl
end as max_length
from _deci1
group by paramn;

*** capture number of types of decimal places ***;

select distinct(max_length) into: type1 -- type5
from _deci2;
%let n_types=&sqlobs;

*** macrotize paramns per decimal place type ***;

%do x= 1 %to &n_types;
   select paramn into: mparams&x separated by ','
   from _deci2 having max_length =&&type&x;
%end;

quit;

*** optional to have within the macro ***;

proc sort data=&orig_data;
by paramn &txgroup avisitn;
run;

proc univariate data=&orig_data noprinit;
class &txgroup avisitn;
var aval;
output n=n mean=mean median=median std=std min=min max=max out=&stat_data;
by paramn;
where avisitn in (2,10);
run;

*** finalize statistical data ***;

data &out_data; length display $20 y_axis x_axis 3;
set &stat_data;
x_axis=&txgroup;

%do x= 1 %to &n_types;
   if paramn in (&mparams&x) then do;
      y_axis=10;
      display=compress(put(n, best.));
      output;
      y_axis=11;
      display=strip(put(mean, 10.%eval(&&type&x+1)) || "(" ||
                     strip(put(std, 10.%eval(&&type&x+2)) || "))" );
      output;
      y_axis=12;
      display=compress(put(median, 10.%eval(&&type&x+1)));
      output;
      y_axis=13;
      display=strip(put(min, 10.&type&x)) || ", " ||
                     strip(put(max, 10.&type&x));
      output;
   end;
%end;
run;

%mend;
%dynamic (_pre_freq, _stat_v1, _means, trt01pn);