

## Appendix 1: Lasagna Plots Can Be Made in Different (Statistical) Ovens

This appendix serves as an online supplement to “Lasagna Plots Can Be Made in Different (Statistical) Ovens.” We provide SAS code snippets and figures of examples used in the commentary by Dr. Swihart and his colleagues in *Epidemiology*.<sup>1</sup>

The logic that SAS uses to produce a lasagna plot is pretty straightforward. SAS regards a lasagna plot as a non-geographic map: a grid should be created as a “map”; the values of repeated measures for each subject, as represented by colors, will be filled in each corresponding unit on the grid; the labels and reference lines can be added in the last step by annotations.

SAS code to duplicate Figure 1 in the commentary<sup>1</sup>

```
*This is the SAS code to duplicate Figure 1 in Dr. Swihart's
commentary;

*Step 0: create some macro variables for generalizability;
*get the name of the input data;
%let data=lasagna1;
*get the variable name of the patient ID variable;
%let ptid=ptid;
*get number of visit and the variable name of each visit;
%let number_visit=6;
%let visit1=visit1;
%let visit2=visit2;
%let visit3=visit3;
%let visit4=visit4;
%let visit5=visit5;
%let visit6=visit6;
*get number of subjects;
proc sql;
    select count(unique &ptid.) into: subject_number
    from &data.;
quit;
*this macro will be used to transform a wide dataset into a long
dataset;
%macro transpose (input, output);
%do i=1 %to &number_visit.;
    data temp&i.;
        length matchid $20.;
        set &input.;
        keep &ptid. outcome matchid time;
        matchid=left(cat(put(temp, 10.), "_&i."));
        outcome=&&visit&i.;
        time=&i.;
    run;
    proc append base=&output data= temp&i.;
    run;
    proc datasets library=work;
        delete temp&i.;
    run;quit;
%end;
%mend;
```

Three steps to prepare three datasets for lasagna plots.

Step 1: prepare for a grid dataset.

```
data grid_dataset;
  length matchid $ 20;
  do grid_y = 0 to &subject_number. ;
  do grid_x = 0 to &number_visit.;
    matchid=trim(left(put(grid_y,best.)))||'_'||trim(left(put(grid_x,
best.))));
    x=grid_x*(100/&number_visit.)-(100/&number_visit.);
    y=grid_y*(125/&subject_number.)-(125/&subject_number.);

  output;
    x=x+(100/&number_visit.); output;
    y=y+(125/&subject_number.); output;
    x=x-(100/&number_visit.); output;
  end;
end;
run;
```

Step 2: prepare for response datasets in specific sorting and/or clustering.

\*the order of subjects in Figure 1 is based on the order of subject id;

\*descending sort of ID b/c graphing from bottom to top;

```
proc sort data=&data. out=lasagna;
  by descending &ptid. ;
run;
```

\*transpose the dataset from a wide dataset into a long dataset and  
assign matchid so that later can match to the grid;

```
data lasagna;
  set lasagna;
  temp=_n_;
run;
```

```
%transpose(lasagna,response_data1);
```

Step 3: prepare for annotation data sets.

\*annotation dataset for Figure 1;

```
data annol;
input @1 y 3. @5 x 3. @9 rotate 3. @13 angle 2. @16 function
$5. @22 text $7. ;
  xsys="2";
  ysys="2";
  when="a";
  style="'Albany AMT'";
  position='4';
  color='black';
  size=1;
cards;
70 -5 360 90 Label Subject
-4 50 0 0 Label Time
0 0 0 0 Move
0 100 0 0 Draw
125 0 0 0 Move
125 100 0 0 Draw
0 0 0 0 Move
125 0 0 0 Draw
0 100 0 0 Move
```

```
125 100 0 0 Draw  
;run;
```

Step 4: Combine three datasets with the GMAP procedure:

```
filename lasagna 'H:\Lasagna\lasagna_Figure1.gif';  
goptions reset=all dev=JPEG300 gsfname=lasagna xpixels=400  
ypixels=550 ftext='Swiss' htext=0.8;  
  
pattern1 v=s c=CX993404;  
pattern2 v=s c=CXD95F0E;  
pattern3 v=s c=CXFE9929;  
pattern4 v=s c=CXFEC44F;  
pattern5 v=s c=CXFEE391;  
pattern6 v=s c=CXFFFD4;  
  
proc gmap data=response_data1 map=grid_dataset all;  
    id matchid;  
    choro outcome/ nolegend coutline=white anno=anno1 discrete;  
run;quit;
```

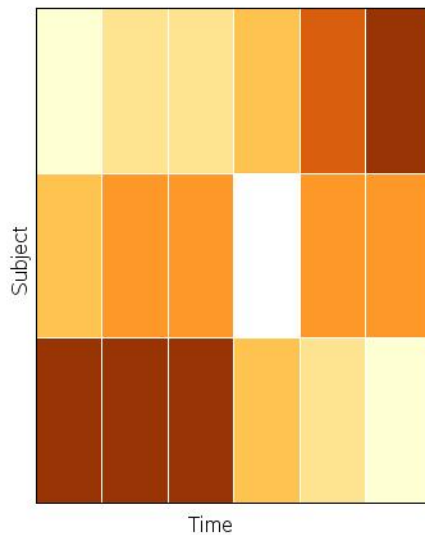


Figure 1 in the commentary<sup>1</sup> duplicated from SAS code

## SAS code to duplicate Figure 2 in the commentary<sup>1</sup>

\*This is the SAS code to duplicate Figure 2 in Dr. Swihart's commentary;

```
*Step 0: create some macro variables for generalizability;
*get the name of the input data;
%let data=lasagna2;
*get the variable name of the patient ID variable;
%let ptid=ptid;
*get number of visit and the variable name of each visit;
%let number_visit=6;
%let visit1=visit1;
%let visit2=visit2;
%let visit3=visit3;
%let visit4=visit4;
%let visit5=visit5;
%let visit6=visit6;
*get number of subjects;
proc sql;
    select count(unique &ptid.) into: subject_number
    from &data.;
quit;
*this macro will be used to transform a wide dataset into a long
dataset;
%macro transpose (input, output);
%do i=1 %to &number_visit.;
    data temp&i.;
        length matchid $20.;
        set &input.;
        keep &ptid. outcome matchid;
        matchid=left(cat(put(temp, 10.), "_&i."));
        outcome=&&visit&i.;
    run;
    proc append base=&output data= temp&i.;
    run;
    proc datasets library=work;
        delete temp&i.;
    run;quit;
%end;
%mend;
```

Three steps to prepare three datasets for lasagna plots.

Step 1: prepare for a grid dataset.

```
data grid_dataset;
    length matchid $ 20;
    do grid_y = 0 to &subject_number. ;
    do grid_x = 0 to &number_visit.;

        matchid=trim(left(put(grid_y,best.)))||'_'||trim(left(put(grid_x,
best.)));

        x=grid_x*(100/&number_visit.)-(100/&number_visit.);
        y=grid_y*(50/&subject_number.)-(50/&subject_number.);
    output;

        x=x+(100/&number_visit.); output;
        y=y+(50/&subject_number.); output;
        x=x-(100/&number_visit.); output;
```

```

end;
end;
run;

```

Step 2: prepare for response datasets in specific sorting and/or clustering.

```

*response dataset for Figure 2A;
*the order of subjects in Figure 2A is based on the order of subject
id;*descending sort of ID b/c graphing from bottom to top;
proc sort data=&data. out=lasagna;
    by descending &ptid.;
run;

```

```

*transpose the dataset from a wide dataset into a long dataset and
assign matchid so that later can match to the grid;

```

```

data lasagna;
    set lasagna;
    temp=_n_;
run;
%transpose (lasagna, response_data2A);

```

```

*response dataset for Figure 2B;
*sort within each row based on value;
%macro sort_within_row;
%do i=1 %to &subject_number.;
    data ptid&i.;
        set &data.;
        where &ptid=&i.;
    run;
    proc transpose data=ptid&i. out=ptid&i.t prefix=_&i ;
    run;
    data ptid&i.t;
        set ptid&i.t;
        if _name_="ptid" or _name_="temp" then delete;
        drop _name_;
    run;
    proc sort data=ptid&i.t;
        by _&i.1;
    run;
    proc transpose data=ptid&i.t out=ptid&i.t2 prefix=visit;
    run;
    proc append base=lasagna data=ptid&i.t2;
    run;
    proc datasets library=work;
        delete ptid&i. ptid&i.t ptid&i.t2;
    run;quit;
%end;
%mend;
%sort_within_row;

```

```

*the order of subjects in Figure 2B is based on the order of subject
id;*descending sort of ID b/c graphing from bottom to top;
proc sort data= lasagna;
    by descending _name_;
run;

```

```

*transpose the dataset from a wide dataset into a long dataset and
assign matchid so that later can match to the grid;
data lasagna;
    set lasagna;
    temp=_n_;
run;
%transpose (lasagna, response_data2B);

*response dataset for Figure 2C;
*re order the subjects;
data lasagna;
    set &data.;
    if label="T1" then group=1;
    else if label="T2" then group=1;
    else if label="P1" then group=2;
    else if label="P2" then group=2;
run;
proc sort data=lasagna;
by descending group descending label;
run;

*transpose the dataset from a wide dataset into a long dataset and
assign matchid so that later can match to the grid;
data lasagna;
    set lasagna;
    temp=_n_;
run;
%transpose (lasagna, response_data2C);

*response dataset for Figure 2D;
*sort and transpose the dataset from a wide dataset into a long dataset
and assign matchid so that later can match to the grid;
%macro transpose2D (input,output);
%do i=1 %to &number_visit.;
    data temp&i.;
        set &input.;
        keep &ptid. outcome;
        outcome=&&visit&i.;
    run;
    proc sort data=temp&i.;
        by outcome;
    run;
    data temp&i.;
        length matchid $20.;
        set temp&i.;
        matchid=left(cat(put(_n_, 10.), "_&i."));
    run;
    proc append base=&output data=temp&i.;
    run;
    proc datasets library=work;
        delete temp&i.;
    run;quit;
%end;
%mend;
%transpose2D(lasagna2,response_data2D);

```

```

*response dataset for Figure 2E;
*sort and transpose the dataset from a wide dataset into a long
dataset;
%macro transpose2E (input,output);
%do i=1 %to &number_visit.;
    data temp&i.;
        set &data.;
        keep &ptid. outcome time;
        outcome= &&visit&i.;
        time=&i.;
    run;
    proc sort data= temp&i.;
        by outcome;
    run;
    proc append base=temp data= temp&i.;
        run;
    proc datasets library=work;
        delete temp&i.;
    run;quit;
%end;
%mend;
%transpose2E(lasagna2,temp);

*calculate number of orange in each column and sort based on number of
orange;
proc sql;
    create table temp2
    as select *, CASE
                                WHEN outcome=1 THEN 1
                                ELSE 0
                            END AS flag
        , sum(calculated flag) as orange
    from temp
    group by time;
quit;

proc sort data=temp2;
    by descending orange time outcome;
run;

*assign matchid so that later can match to the grid;
data temp2;
    set temp2;
    key=_n_;
run;

data temp3;
    do grid_x = 1 to &number_visit.;
        do grid_y = 1 to &subject_number. ;

            matchid=trim(left(put(grid_y,best.)))||'_'||trim(left(put(grid_x,
best.))); output;
        end;
    end;
run;

data temp3;

```

```

        set temp3;
        key=_n_;
run;

proc sql;
    create table response_data2E
    as select a.outcome,b.matchid
    from temp2 as a inner join temp3 as b
    on a.key=b.key;
quit;

```

Step 3: prepare for annotation data sets.

\*annotation dataset for Figure 2A;

```

data anno2A;
input @1 y 3.      @5 x 3.      @9 rotate 3. @13 angle 2. @16 function
$5. text $30. ;
    xsys="2";
    ysys="2";
    when="a";
    style="'Albany AMT'";
    position='4';
    color='black';
    size=2;
cards;
54 45      0 0 Label A  Initial Lasagna Plot
44 -5      0 0 Label P1
31 -5      0 0 Label T1
19 -5      0 0 Label P2
6  -5      0 0 Label T2
-4 8       0 0 Label 1
-4 25      0 0 Label 2
-4 42      0 0 Label 3
-4 59      0 0 Label 4
-4 76      0 0 Label 5
-4 93      0 0 Label 6
0 0        0 0 Move
0 100      0 0 Draw
50 0       0 0 Move
50 100     0 0 Draw
0 0        0 0 Move
50 0       0 0 Draw
0 100      0 0 Move
50 100     0 0 Draw
;
run;

```

\*annotation dataset for Figure 2B;

```

data anno2B;
input @1 y 3.      @5 x 3.      @9 rotate 3. @13 angle 2. @16 function
$5. text $30. ;
    xsys="2";
    ysys="2";
    when="a";
    style="'Albany AMT'";
    position='4';
    color='black';
    size=2;

```



```

cards;
54 58 0 0 Label B Within-row sorting of (A)
44 -5 0 0 Label P1
31 -5 0 0 Label T1
19 -5 0 0 Label P2
6 -5 0 0 Label T2
-4 19 0 0 Label 1/6
-4 36 0 0 Label 1/3
-4 53 0 0 Label 1/2
-4 70 0 0 Label 2/3
-4 87 0 0 Label 5/6
0 0 0 0 Move
0 100 0 0 Draw
50 0 0 0 Move
50 100 0 0 Draw
0 0 0 0 Move
50 0 0 0 Draw
0 100 0 0 Move
50 100 0 0 Draw
;run;

*annotation dataset for Figure 2C;
data anno2C;
input @1 y 3. @5 x 3. @9 rotate 3. @13 angle 2. @16 function
$5. text $30. ;
    xsys="2";
    ysys="2";
    when="a";
    style="'Albany AMT'";
    position='4';
    color='black';
    size=2;
cards;
54 56 0 0 Label C Entire-row sorting of (A)
44 -5 0 0 Label T1
31 -5 0 0 Label T2
19 -5 0 0 Label P1
6 -5 0 0 Label P2
-4 8 0 0 Label 1
-4 25 0 0 Label 2
-4 42 0 0 Label 3
-4 59 0 0 Label 4
-4 76 0 0 Label 5
-4 93 0 0 Label 6
0 0 0 0 Move
0 100 0 0 Draw
50 0 0 0 Move
50 100 0 0 Draw
0 0 0 0 Move
50 0 0 0 Draw
0 100 0 0 Move
50 100 0 0 Draw
;
run;

*annotation dataset for Figure 2D;
data anno2D;

```

```

input @1 y 3.      @5 x 3.      @9 rotate 3. @13 angle 2. @16 function
$5. text $35. ;
    xsys="2";
    ysys="2";
    when="a";
    style="'Albany AMT'";
    position='4';
    color='black';
    size=2;
cards;
54 64      0 0 Label D Within-column sorting of (C)
38 -5      0 0 Label 3/4
25 -5      0 0 Label 1/2
13 -5      0 0 Label 1/4
-4 8       0 0 Label 1
-4 25      0 0 Label 2
-4 42      0 0 Label 3
-4 59      0 0 Label 4
-4 76      0 0 Label 5
-4 93      0 0 Label 6
0 0        0 0 Move
0 100      0 0 Draw
50 0       0 0 Move
50 100     0 0 Draw
0 0        0 0 Move
50 0       0 0 Draw
0 100      0 0 Move
50 100     0 0 Draw
;
run;

*annotation dataset for Figure 2E;
data anno;
    length text $35.;
    set temp2;
*the data set temp2 contains the correct order of the data produced in
the response dataset step;
    where ptid=1;
    text=left(put(time,2.));
    y=-4;
    rotate=0;
    angle=0;
    function="Label";
    if _n_=1 then x=8;
    if _n_=2 then x=25;
    if _n_=3 then x=42;
    if _n_=4 then x=59;
    if _n_=5 then x=76;
    if _n_=6 then x=93;
run;

data anno2;
input @1 y 3.      @5 x 3.      @9 rotate 3. @13 angle 2. @16 function
$5. text $35. ;
cards;
54 60      0 0 Label E Entire-column sorting of (D)
38 -5      0 0 Label 3/4

```

```

25  -5      0   0  Label 1/2
13  -5      0   0  Label 1/4
0   0   0   0  Move
0   100  0   0  Draw
50  0   0   0  Move
50  100  0   0  Draw
0   0   0   0  Move
50  0   0   0  Draw
0   100  0   0  Move
50  100  0   0  Draw
;run;

```

```

data anno2E;
set anno anno2;
    xsys="2";
    ysys="2";
    when="a";
    style="'Albany AMT';
    position='4';
    color='black';
    size=2;
run;

```

Step 4: Combine three datasets with the GMAP procedure.

```

goptions reset=all dev=JPEG300 gsfname=lasagna xpixels=400
ypixels=250 ftext='Swiss' htext=0.8;
pattern1 v=s c=CXD73027;
pattern3 v=s c=CX762A83;
pattern2 v=s c=CXAF8DC3;
filename lasagna 'H:\Lasagna\lasagna_graph2A.gif';
proc gmap data=response_data2A map= grid_dataset all;
    id matchid;
    choro outcome/ nolegend coutline=white anno=anno2A discrete;
run;quit;

filename lasagna 'H:\Lasagna\lasagna_graph2B.gif';
proc gmap data=response_data2B map= grid_dataset all;
    id matchid;
    choro outcome/ nolegend coutline=white anno=anno2B discrete;
run;quit;

filename lasagna 'H:\Lasagna\lasagna_graph2C.gif';
proc gmap data=response_data2C map= grid_dataset all;
    id matchid;
    choro outcome/ nolegend coutline=white anno=anno2C discrete;
run;quit;

filename lasagna 'H:\Lasagna\lasagna_graph2D.gif';
proc gmap data=response_data2D map= grid_dataset all;
    id matchid;
    choro outcome/ nolegend coutline=white anno=anno2D discrete;
run;quit;

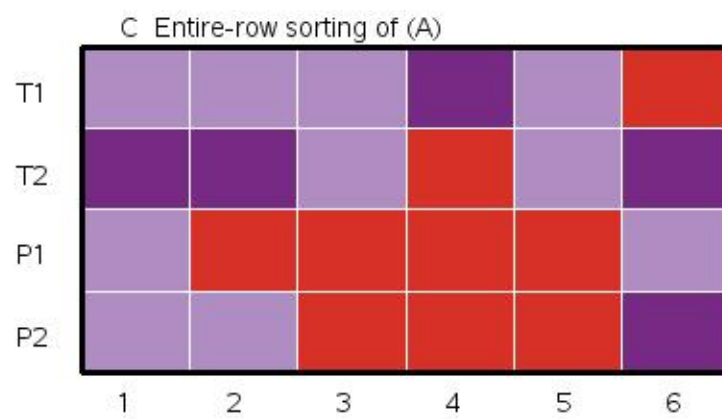
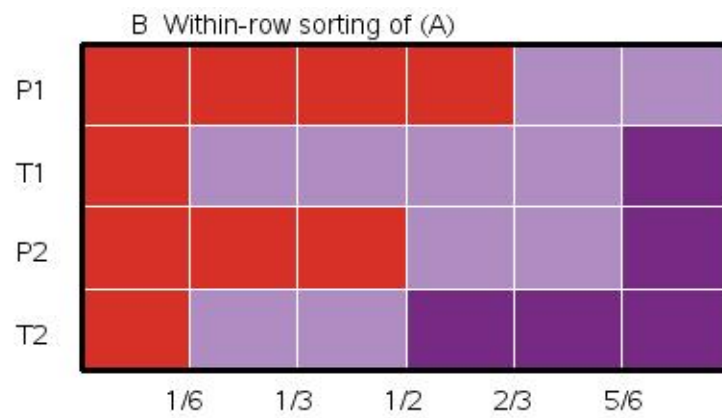
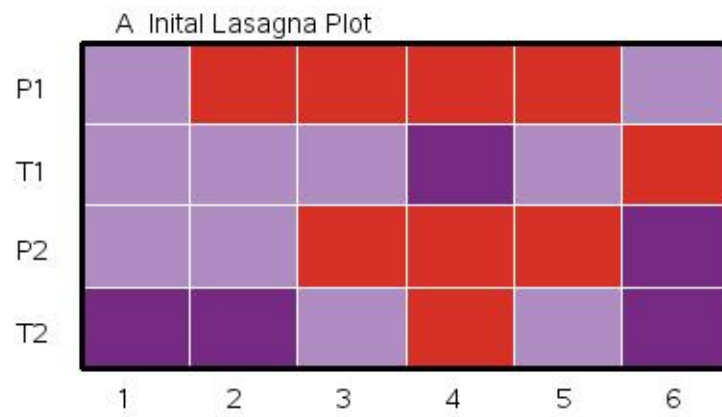
filename lasagna 'H:\Lasagna\lasagna_graph2E.gif';
proc gmap data=response_data2E map= grid_dataset all;
    id matchid;

```

```

choro outcome/ nolegend coutline=white anno=anno2E discrete;
run;quit;

```



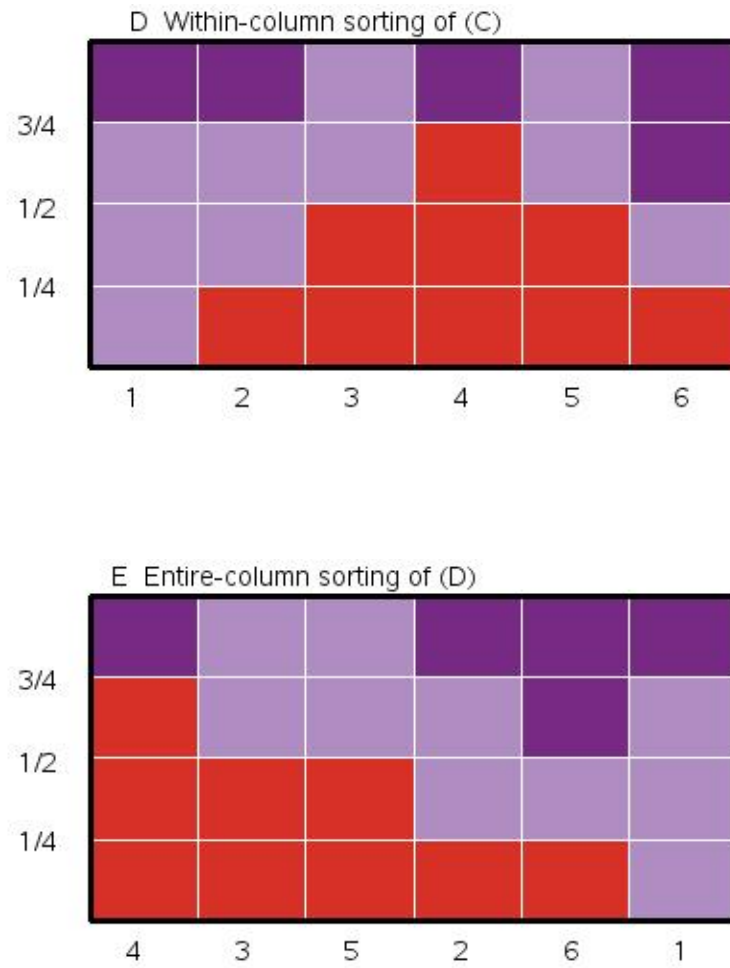


Figure 2 in the commentary<sup>1</sup> duplicated from SAS code

## SAS code to duplicate Figure 3 in the commentary<sup>1</sup>

```
*This is the SAS code to duplicate Figure 3 in Dr. Swihart's
commentary;
*Since I do not have the exact data, the code only shows the method of
clustering and sorting;

*Step 0: create some macro variables for generalizability;
*get the name of the input data;
%let data=lasagna3;
*get the name of cluster variable;
%let cluster=cluster;
*get the variable name of the patient ID variable;
%let ptid=ptid;
*get number of visit we are going to plot and the variable name of each
visit;
%let number_visit=4;
%let visit1=visit1;
%let visit2=visit2;
%let visit3=visit3;
%let visit4=visit4;
*get number of subjects;
proc sql;
    select count(unique &ptid.) into: subject_number
    from &data.;
quit;
*this macro will be used to transform a wide dataset into a long
dataset;
%macro transpose (input, output);
%do i=1 %to &number_visit.;
    data temp&i.;
        length matchid $20.;
        set &input.;
        keep &ptid. visit matchid time;
        matchid=left(cat(put(temp, 10.), "_&i."));
        visit=&&visit&i.;
        time=&i.;
    run;
    proc append base=&output data=temp&i.;
    run;
    proc datasets library=work;
        delete temp&i.;
    run;quit;
%end;
%mend;
```

Three steps to prepare three datasets for lasagna plots.

Step 1: prepare for a grid dataset.

```
data grid_dataset;
    length matchid $ 20;
    do grid_y = 0 to &subject_number. ;
    do grid_x = 0 to &number_visit.;

        matchid=trim(left(put(grid_y,best.)))||'_'||trim(left(put(grid_x,
best.)));
        x=grid_x*(100/&number_visit.)-(100/&number_visit.);
```

```

y=grid_y*(100/&subject_number.)-(100/&subject_number.);
output;
x=x+(100/&number_visit.); output;
y=y+(100/&subject_number.); output;
x=x-(100/&number_visit.); output;
end;
end;
run;

```

Step 2: prepare for response datasets in specific sorting and/or clustering.

\* sort the data first by cluster than by value of each visit;

```

proc sort data=&data. out=lasagna;
  by &cluster &visit1 &visit2 &visit3 &visit4 ;
run;

```

```

data lasagna;
  set lasagna;
  temp=_n_;
run;

```

```
%transpose(lasagna, response_data3);
```

\*categorize the values of each visit into five categories to match five colors;

```

data response_data3;
  set response_data3;
  if visit=. then visitc=.;
  else if visit lt 30 then visitc=1;
  else if visit lt 60 then visitc=2;
  else if visit lt 90 then visitc=3;
  else if visit lt 120 then visitc=4;
  else visitc=5;
run;

```

Step 3: prepare for annotation data sets.

\*annotation dataset for Figure 3;

```

data anno;
input @1 y 3. @5 x 3. @9 rotate 3. @13 angle 2. @16 size 3.1
@20 function $5. @26 text $70. ;
cards;
97 -10 360 90 0.7 Label Subjects(Intervention group below, Placebo
group above black line)
40 -5 360 90 0.7 Label Intervention Group
90 -5 360 90 0.7 Label Placebo Group
105 90 0 0 1.0 Label Lasagna Plot: 1000 subjects, 4 visits
-1 13 0 0 1.0 Label 1
-1 38 0 0 1.0 Label 2
-1 63 0 0 1.0 Label 3
-1 88 0 0 1.0 Label 4
-5 60 0 0 1.0 Label Visit #
0 0 0 0 1.0 Move
0 100 0 0 1.0 Draw
100 0 0 0 1.0 Move
100 100 0 0 1.0 Draw
0 0 0 0 1.0 Move
100 0 0 0 1.0 Draw
0 100 0 0 1.0 Move

```

```
100 100 0 0 1.0 Draw
;run;
```

```
*calculate the y position for the cluster line;
```

```
proc sql;
    create table line as
    select count(unique &ptid.)*(100/&subject_number.) as y
    from lasagna
    where &cluster. =1;
quit;
```

```
*create a annotation dataset for the cluster line;
```

```
data line;
    set line;
    x=-2;function="Move";output;
    x=102;function="Draw";size=4;output;
run;
```

```
data anno3;
    set anno line;
    xsys="2";
    ysys="2";
    when="a";
    style="'Swiss'";
    position='4';
    color='black';
run;
```

```
Step 4: Combine three datasets with the GMAP procedure:
```

```
filename lasagna 'H:\Lasagna\lasagna_cluster.gif';
goptions reset=all dev=JPEG300 gsfname= lasagna xpixels=5000
ypixels=5000 ftext='Swiss' htext=0.8;
pattern5 v=s c=CX006D2C;
pattern4 v=s c=CX238B45;
pattern3 v=s c=CX41AB5D;
pattern2 v=s c=CX74C476;
pattern1 v=s c=CXA1D99B;
legend1 POSITION=(bottom right OUTSIDE) mode=PROTECT across=5
shape=bar(2,2)pct value=(height=0.5 "0-30" "30-60" "60-90" "90-120"
"120-150") label=(height=0.5 "Test Score:");
proc gmap data=response_data3 map=grid_dataset all;
    id matchid;
    choro visitC/ legend=legend1 coutline=white anno=anno3 discrete;
run;quit;
```



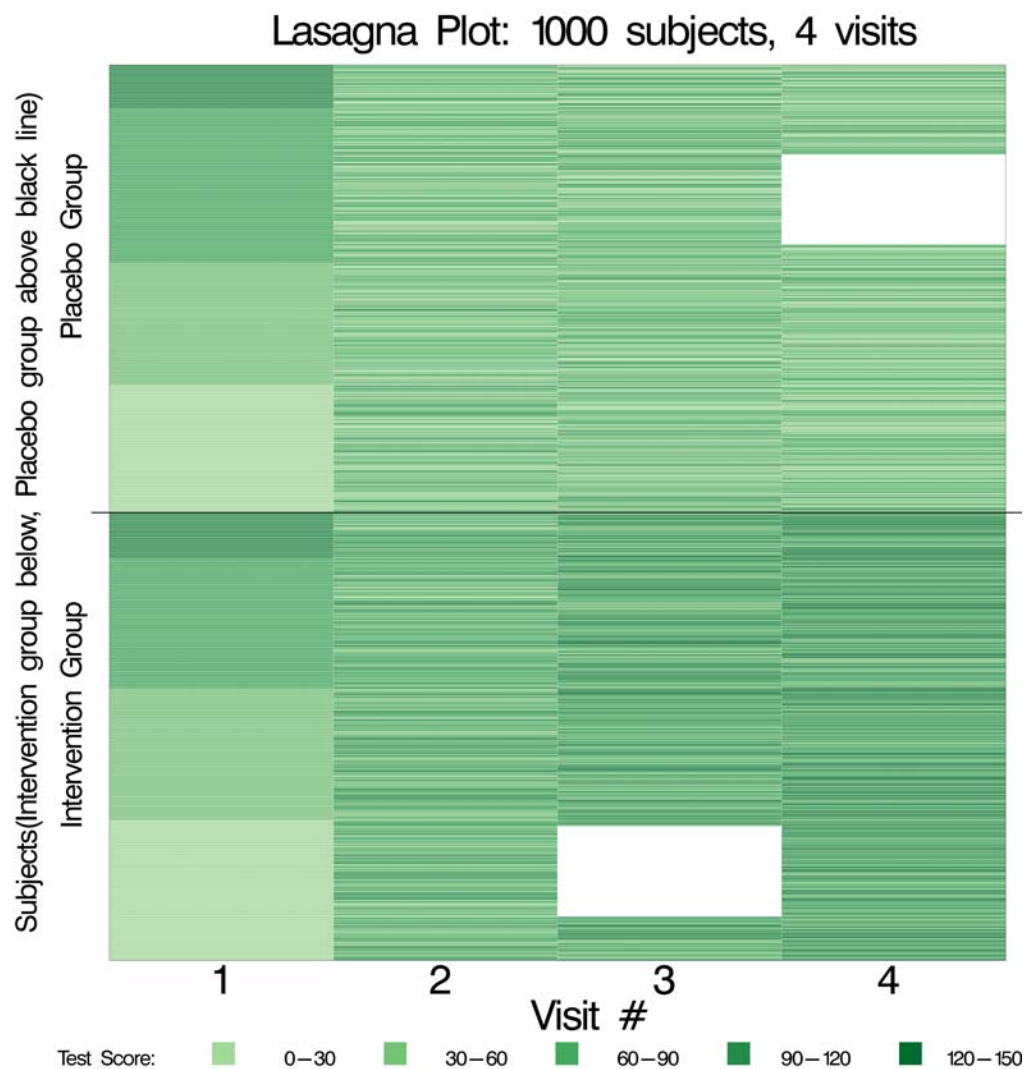


Figure 3: Lasagna plot with sorting and clustering for a large sample duplicated from SAS code

If you have any question or improvement about the SAS code, please contact Hongyuan Gao (email: [gao.h@ghc.org](mailto:gao.h@ghc.org)).

#### References

1. Swihart BJ, Caffo B, James BD, et al. Lasagna plots: a saucy alternative to spaghetti plots. *Epidemiology*. 2010;**21**(5):621-5.