Map Smart: Design and Build Effective InfoGeographics Using PROC GMAP and Software Intelligence

LeRoy Bessler, Besler & Von Battenberg

Introduction and Acknowledgements

Learn good map design and best use of SAS* and SAS/GRAPH* software for more effective, more efficient exploration and presentation of geographic-keyed data. Macros and programs are reusable and adaptable, by even a new or infrequent user. I cover: best type of map; best area-fill to show response level; effective use of color; readable annotation of states with text, data, rank, and city markers; and dynamic generation of data-appropriate PATTERN statements, legends, and response-range formats.

InfoGeographics* is my name for "statistical mapping", "thematic mapping", "Business Geographics", or the functions of a Geographic Information System (GIS). IG can be done with PROC GMAP, without recourse to a GIS. You can also apply concepts and techniques presented here to use of SAS/GIS* or to building a custom GIS application.

Better than using defaults, or sorting and inspecting responses to "hand pick" ranges (which may be arbitrary anyhow), is Automatic Rationale-based Response Range Assignment (ARbRRA), using "Software Intelligence" (SI). Three rationales are presented, but others also can be automated. Use of SI was demonstrated in other work. See, e.g., LeR. Bessler, "Software Intelligence: Applications That Customize Themselves", in *Proceedings of the Eighteenth Annual SAS Users Group International Conference*, SAS Institute Inc. (Cary, N.C.), 1993.

My general-purpose maps--The Four-Color/Four-Range Map, The Five-Color/Five-Change-Range Map, and The N-Color/N-Cluster Map--are usable for many InfoGeographic applications. *They communicate as much information as possible with a single image.*

(As time permits, the presentation will also include InfoGeographic examples other than those in this paper.)

When annotating areas filled with gray shades, or with dark or intense colors, custom-developed "blanking" provides an inset box of white space, to assure readability. It was first reported on by S. J. Subichin in "Enhanced Useability for Annotation on SAS/GRAPH Maps", in *WISAS Proceedings*, Volume 5, June Issue, WISAS Inc. (South Milwaukee, Wis.), 1993.

I provide adjustments to the USA state-center coordinates in the vendor-supplied SAS/GRAPH data set MAPS.USCENTER. They permit more annotation to be inlaid without crossing the state boundary. The annotation box is as equidistant as possible from all near-points of the boundary.

I am pleased to thank Gary F. Plazyk, who suggested I investigate cluster analysis as a tool to solve The N Color Map Problem.

Why Maps?

"Of all the contrivances hitherto devised for the benefit of geography, the map is the most effective. In the extent and variety of its resources, in rapidity of utterance, in the copiousness and completeness of the information it communicates, in precision, conciseness, perspicuity, in the hold it has upon the memory, in vividness of imagery and power of expression, in convenience of reference, in portability, in the happy combination of so many and such useful qualities, a map has no rival. Everything we say or do has reference to place, and wherever place is concerned a map deserves welcome. There is scarcely one department of knowledge, physical or moral, beyond the sphere of its usefulness; to geography it is indispensable. Modern technology has advanced the process of making maps considerably, and a map still has no rival in its usefulness."

G.B. Geenough
Presidential Address to The Royal Society
London, England
1840

Most SAS user sites have large amounts of data that include geographic unit area designators (in the USA, state code is probably commonest). Though this data can be reported in various tabular formats, a geographic effect (e.g., that of proximity) is not easily revealed without an InfoGeographic. Also, *for presentation, a visual image is more interesting, stimulating, and memorable than a mere listing.*

Why Annotation?

Any map can be supplemented with detail look-up data. Detail can be a full list in key sequence, a ranking report based on response level, or a Top NN List of the NN most signficant (i.e., highest response) geographic unit areas. Better is to inlay all detail (including rank) on the respective geographic areas, with automated annotation. Interactive graphic editing is not suitable for hands-off, production applications. *Most good, supposedly one-shot applications usually end up as ongoing production.*

Just Say "No" to the Designer Drug 3D

Use the straightforward two-dimensional CHOROPLETH map. The 3D alternatives--SURFACE, PRISM, and BLOCK maps--are picturesque, but impractical. SURFACE maps are too vague for serious communication. PRISM and BLOCK maps suffer from the response for some "high" states hiding that for "low" states.

Make It "Easy On the Eyes" With Area Fills

Use of parallel lines or cross-hatching not only yields an ugly image, but also can confuse boundary with area-fill elements.

For some InfoGeographic applications, use of area fills to encode different levels of response is functionally inappropriate. For the use of dot maps or bubble maps, see Plazyk, G. F., "Using the Annotate Facility with Maps: A Tutorial", in *Proceedings of MWSUG '91*, MidWest SAS Users Group (Fox Point, Wis.), 1991.

Figure 1: SAS/GRAPH Defaults Unacceptable

The map in Figure 1 (done with the program in Appendix 3) is an unacceptable map of a real data set, using PATTERN statements, COUTLINE, and otherwise a default invocation of PROC GMAP. The adverse result of accepting SAS/GRAPH default ranges is due to two outliers, 736 and 447; all other values are below 179. Note, also, that default legend text shows range midpoints, instead of the range boundaries which you might expect.

Four Color Map Problems

The Original Four Color Map Problem: *Can you prove that four is the smallest number of colors needed to paint a map so that no two adjacent countries are the same color?*

The Four-Color/Four-Range Map Problem: *If you want to restrict an InfoGeographic to four ranges that span the total range of the response data, how can/should the program automatically specify the ranges?*

Four Color Map Problem Solved, Using Automatic Rationale-based Response Range Assignment

Even without an extreme result as Figure 1, it is better to make a deliberate choice of ranges, based on a rationale. In principle, that requires you to have knowledge of the data distribution. Before creating the map, one can first do a PROC SORT and PROC PRINT, and inspect the data. *However, that is inconvenient, time-consuming, and laborious, and can result in an arbitrary decision anyhow.*

In a prior paper--"Effective and Efficient Information Delivery for Executive Management", in *Proceedings of the Seventeenth Annual SAS Users Group International Conference*, SAS Institute Inc. (Cary, N.C.), 1992)--I emphasized that, typically, a small subset of the observations account for a large majority, or even almost all, of the total response. A Top 10 or Top NN Report (i.e., some one-page-or-less report) usually suffices, often accounting for 80% to 99% of the total response. For the data depicted in Figures 1 and 2, the Top 10 states account for 66.4% of the total response. With 50 states and DC, the Top 10 states are, by definition, always above the 80th percentile.

My favorite percentile is the 50th, i.e., the median. What I call "The Power of the Median" is its *representative* centrality. The influence of outliers suffered by the regrettably popular *average* is absent.

Regardless of the specific choices, it is natural to break up the total range based on percentiles. One can use, e.g., the 20th percentile, median (the 50th percentile), and 80th percentile. The resulting four ranges may be called, e.g., Very Low, Below Median, Above Median, and Very High. Other rationales can be built-in instead. E.g., one might prefer to use the mean and a multiple of the standard deviation to develop ranges. In that connection, consider how you might use Tchebychev's Theorem.

Percentile-based ranges create a talking point for the map. Software defaults or arbitrary breakpoints cannot provide *concept-based defendability.*

The program in Appendix 4 and macros in Appendices 6 and 8 solve The Four Color Map Problem, using SI to do ARbRRA. See the map in Figure 2. The legend displays "trimmed" numeric ranges, instead of text, if you specify LABELTXT=NO. "Trimmed" ranges use actual data values, emphasizing the inter-range separation. Traditional ranges are "contiguous".

Maximal Optimized Annotation, and Blanking

The map in Figure 2 does just about everything one could think of for an InfoGeographic. (OK, everything that *I* could think of.)

Of course, if one also wants, say, Percent of Whole, a four-line annotation could be done with a modification to the ANNOVALU DATA Step in the USANNO3 macro. Four-line annotation would require application of offsets to the state-center y coordinate, and using only POSITIONs '5' and '6'. Then one would need to specify a smaller value for ANNOFONT.

Or one could go the opposite direction, and provide less annotated information. Two-line annotation should use POSITIONs 'B', 'C', 'E', and 'F' instead of the six values used here. Then one could specify a larger value for ANNOFONT.

The POSITION annotate variable is explained on pages 521-524 of *SAS/GRAPH Software Reference, Version 6, First Edition, Volume 1*, SAS Institute Inc. (Cary, N.C.), 1990. The code for DATA Step ANNOVALU in the USANNO3 macro is an improved and extended version of one-line annotation code on page 10 of *SAS/GRAPH Software: Map Data Sets (SAS Technical Report P-196)*, SAS Institute Inc. (Cary, N.C.), 1990.

Annotation with response value has obvious benefit. And since not everyone knows each state name just by shape and relative location on the map, it is "nice to have" that identification. The provision of Rank based on response value reflects my graphic and tabular communication design postulate that most readers and viewers ought to be able to quickly identify *what's important.*

The map uses blanking (i.e., the white boxes) to assure readability, and relies on my recommended adjustments (see Appendix 1) to the state-center coordinates from MAPS.USCENTER.

Also, the map incorporates a star to highlight the conference location. (Instead, the city marker could have been the city name, or "SUGI 14".) See DATA Steps CITYSTAR and ANNODATA in the program in Appendix 4 for how to do this. Appendix 2 is a program to list the GMAP US cities. For a city not listed, pick the nearest city listed, and adjust coordinates. Annotating cities was demonstrated in the manual for Version 5 of SAS/GRAPH. It's included here for completeness--to provide a single, reusable, adaptable model that includes all the techniques likely to be needed for professional-grade InfoGeographics.

I had to develop "boundary-respecting blanking" for the states of Florida, Tennessee, and West Virginia. Strictly rectangular boxes overlap the state boundaries if the annotatable white space is kept at sufficient size. Study the three state-specific WHEN paragraphs in the BOXES DATA Step in the USANNO3 macro to see how adequate annotatable area is provided without white space crossing state lines.

Annotation Without Blanking

As of Release 6.10, blanking is still missing from SAS/GRAPH. Even if someday SAS/GRAPH blanking is provided, the vendor implementation may not offer the flexibility achievable with the custom solution presented here.

Blanking is not needed for area fills that use light colors (e.g., light pink, light yellow, light blue, etc.). But not all devices can render sufficiently light colors. Also, many publications (e.g., *SUGI Proceedings*) do not accept color illustrations.

A very interesting, informative map is one that someone may want to copy. Though color copiers are increasingly available, they are not as widespread, cheap, and fast as black-and-white. Thus, annotated gray-shade maps usually are most practical. For readability, whether of originals or of copies, blanking is always required when annotating in black over gray-shade area fill.

The Five-Color/Five-Change-Range Map Problem

After solving The Four Color Map Problem for SUGI 19, I found myself facing The Five Color Map Problem--which I solved for SUGI 20.

Suppose the responses are positive, negative, and no change. Suppose we want something more interesting than the three obvious response ranges. Tentatively, let's classify the responses as big gains, other gains, no change, big losses, and other losses--necessitating the use of five colors.

For this case study, let's again pick the USA map. Rather than using percentiles, let's distinguish the Ten Best Gains and the Ten Worst Losses. (Please excuse me. Strictly speaking, there can be only one best and only one worst.) Of course, it may happen that there are no gains, fewer than ten gains, or only ten gains; and the same applies to losses.

Area Fill for The Change Map

One might naively pick green and red as natural choices for gains and losses--when increase is good and decrease is bad. However, green and red cannot be distinguished if one suffers from the commonest form of color blindness, and color blindness is not rare (1 out of 14 males has some form of color blindness).

My recommendations are: Blue = Ten Best Gains; Light Blue = Other Gains; Red = Ten Worst Losses; Light Red = Other Losses; and White = No Change.

> Since this paper is published in black-and-white, the area fills have to be disappointingly less informative grey shades. Upon request, the author may be able to provide the map in color.

Best Legend for The Change Map

The best legend for this application will do the following: (a) show the area fill for the Ten Best Gains (Worst Losses), if there are more than ten gains (losses); (b) show the area fill for the Gains (Losses), if there are ten or fewer gains (losses), and will list how many gains (losses) there are; (c) show the area fill for the Other Gains (Losses), if there are more than ten gains (losses), and will list how many Other Gains (Losses) there are; and (d) show the area fill for the Unchanged, if there are any, and will list how many Unchanged there are.

With this design, the legend can have as many as five entries or as few as one entry. It will have entries *only* for the cases manifested by the data, and those entries will provide area-fill sample, state count, and category description. What more could you ask for? Why would you ask for less?

Five Color Map Problem Solved, Using Automatic Rationale-based Response Range Assignment

The program in Appendix 5 and macros in Appendices 7 and 8 solve The Five Color Map Problem. See the map in Figure 3.

Note that the MAPCOL5C macro must be able to handle 14 different cases with its nested PATTERN5 macro. MAPCOL5C can generate 14 different CHORO variable range sets. Its custom legend text must actually support 30 different cases due to, e.g., the possibility of "Gains" vs. "Gain" and "Losses" vs. "Loss". In an automated professional-grade InfoGeographic application (where there is no ad hoc manual editing to update the program to suit the vicissitudes of the data), the category text in the legend must automatically match the plurality or singularity of category count.

A possible future enhancement of the MAPCOL5C macro would allow user specification of legend text (to substitute for the herein hard-coded words "Gain", "Gains", "Loss", "Losses", "Best", "Worst", "Other") as macro parameter assignments.

Why USANNO3 Handles State Codes As It Does

Various FIPxxxxx and STxxxxx SAS functions perform conversion between state FIPS codes, abbreviations, and names.

The input data set contains the two-character state abbreviation. The program converts it to the two-digit state FIPS code, using the STFIPS function. The abbreviation is dropped to produce a minimal data set, but dropping it is not necessary. Subsequent processing, by USANNO3, relies on the FIPS code. But, at some points in the macro, the FIPSTATE SAS function is used in comparisons to identify states which need special handling. Those states are specified by their alphabetic codes.

Conversion back to alphabetic state codes which were initially present, converted from, and then (unnecessarily) dropped seems inefficient, and/or clumsy, and/or unjustified. Well, not only may it be the case that more typical input data sets might contain only the FIPS code, but also PROC GMAP and its map data sets identify states only by FIPS code. Hence the macro's focus on FIPS code. However, when testing observations for specific states, the macro is more intelligible if it identifies those states by their alphabetic codes. In fact, the FIPNAME or FIPNAMEL SAS function could instead be used to test with state names.

The N Color Map Problem & Its ARbRRA Solution

After solving The Four Color Map Problem and The Five Color Map Problem, I was still dissatisfied by lack of generality in these ways of presenting geo-based information.

If there is nothing inherent in the nature of the responses (e.g., they are not signed numbers), nor in your or your audience's preference as to how to show or see the information, then you have no guide for choice of response ranges. Your only recourse is to inspect the data to find natural groups of responses. Intuitively, "natural" groups or classes are distinguishable by sufficient separation between their ranges. Choice of the *number* N of ranges is still arbitrary, but at least the boundaries of well separated ranges can be justified.

An Automated Rationale-based tool that reveals those natural ranges, for any given N, is PROC CLUSTER. It offers eleven statistical methods for cluster analysis. The one used here is METHOD=CENTROID. It copes with outliers well. I can't say which is the "best" clustering method. For me, reasonability of results is the measure of adequacy of any method used. However, my macro MAPCOLNC (see Appendix 10) permits user specification of METHOD=. The macro builds PATTERN statements, legend text, and the response-range format.

Macro MAPCOLNC requires specification of the number N of ranges (i.e., clusters), where maximum N is 7. (The macro could be enhanced to support N greater than 7.) The user may request default colored area fills, may specify other colors, or may accept default gray-shade area fills. The gray shades are selected to be maximally distinguishable, given the number required.

Distinguishability of gray shades is always a potential problem, either on the original, or after photocopying. That's why the macro supports a maximum of only 7 ranges. With the printer I used for Figure 4, the five-gray-shade map has no distinguishability problem, the six-gray-shade map is marginal, and the seven-gray-shade map is hard to interpret.

If you use color, you can safely revise the macro to accommodate N greater than 7. But then you need a strategy to pick colors, unless you use one hue and vary the lightness--in which case you will have a distinguishability problem for sufficiently large N.

See Figure 4 for maps obtained for various numbers of clusters (ranges). For simplicity, annotation is omitted. See Appendix 9 for the program used to produce The Seven Color Map.

Notices

Author

Dr. LeRoy Bessler
Besler & Von Battenberg
P.O. Box 96
Milwaukee, WI 53201-0096, USA
Telephone: 414-351-6748

```
01 N -0.0037 +0.0000 AL
02 N +0.0050 +0.0000 AK
04 N -0.0030 +0.0000 AZ
05 N -0.0025 +0.0037 AR
08 N +0.0000 +0.0020 CO
09 Y -0.0140 -0.0050 CT
10 Y -0.0025 -0.0060 DE
11 Y -0.0100 -0.0200 DC
12 N +0.0045 -0.0060 FL
15 N +0.0050 +0.0250 HI
16 N -0.0050 +0.0000 ID
17 N -0.0015 +0.0050 IL
18 N -0.0013 +0.0000 IN
19 N +0.0000 -0.0020 IA
20 N +0.0000 -0.0015 KS
21 N +0.0000 +0.0035 KY
22 N -0.0010 +0.0080 LA
23 N -0.0030 +0.0060 ME
24 Y +0.0250 -0.0200 MD
25 Y -0.0050 +0.0015 MA
26 N +0.0010 -0.0030 MI
28 N -0.0013 +0.0000 MS
29 N -0.0037 +0.0000 MO
30 N +0.0070 -0.0005 MT
33 Y -0.0010 +0.0150 NH
34 Y +0.0275 -0.0200 NJ
36 N +0.0025 +0.0015 NY
37 N +0.0000 +0.0035 NC
38 N +0.0000 +0.0020 ND
39 N -0.0050 +0.0025 OH
42 N +0.0000 +0.0045 PA
44 Y +0.0000 -0.0100 RI
45 N +0.0055 +0.0015 SC
46 N +0.0000 +0.0020 SD
47 N +0.0000 +0.0025 TN
50 Y -0.0500 +0.0400 VT
51 N +0.0000 -0.0015 VA
53 N +0.0100 +0.0000 WA
54 N -0.0020 +0.0010 WV
55 N +0.0010 +0.0000 WI
56 N -0.0050 +0.0020 WY
```

Appendix 1: Adjustments for MAPS.USCENTER

```
DATA USCITIES;
SET MAPS.USCITY;
STATENAM = FIPNAMEL(STATE);
IF CAPITAL = 'N' THEN CAPITAL = ' ';
RUN;
PROC PRINT DATA=USCITIES UNIFORM LABEL;
PAGEBY STATE;
BY STATE STATENAM;
ID STATE STATENAM;
VAR CITY X Y CAPITAL POP;
LABEL STATE='ID';
LABEL STATENAM='State';
LABEL CAPITAL='Capital?';
TITLE 'SAS/GRAPH MAPS.USCITY File';
RUN;
```

Appendix 2: Program to List GMAP US Cities

```
 /* goptions statements */

DATA INDATA;
INFILE I;
INPUT STATEABB $ ATTEND;
STATE = STFIPS(STATEABB);
RUN;

 /* pattern statements */
 /* title statements */
PROC GMAP DATA=INDATA MAP=MAPS.US ALL;
ID STATE;
CHORO ATTEND / COUTLINE=BLACK;
RUN;
```

Appendix 3: Program for Default Map (Figure 1)

```
 /* goptions statements */

DATA INDATA;
INFILE I;
INPUT STATEABB $ ATTEND;
STATE = STFIPS(STATEABB);
DROP STATEABB;
RUN;

%MAPCOL4R(DATA=INDATA,
          CHOROVAR=ATTEND,
          CHORINCR=1,
          PCTLLOW=20,
          PCTLMID=50,
          PCTLHIGH=80,
          LABELTXT=YES, /* same as default */
          LABVLOW=Very Low,
          LABBELOW=Low,
          LABABOVE=High,
          LABVHIGH=Very High,
          COLVLOW=GRAYCC,
          COLBELOW=GRAY99,
          COLABOVE=GRAY66,
          COLVHIGH=GRAY33)
RUN;

%USANNO3(DATA=INDATA,
         ANNOVAR2=ATTEND,
         ANNOVAR3=@@RANK@@, /* same as default */
         OUTANNO=ANNODATA,
         NOBOXSTA=, /* if use C=WHITE, dataset of state IDs */
         ANNOFONT=NONE,      /* same as default */
         ANNOSIZE=0.90,      /* reduce height if DOWN > 1 */
         OUTMAP=MBC_US,
         BLANK=YES)
RUN;

DATA  CITYSTAR; /* marker for city location */
SET  MAPS.USCITY;
IF  CITY = 'San Francisco'; /* get coordinates for SF */
XSYS = '2'; /* do not change */
YSYS = '2'; /* do not change */
WHEN = 'A'; /* do not change */
FUNCTION = 'LABEL';
STYLE = 'MARKER'; /* special font of markers */
SIZE = 1.00; /* adjust as desired */
COLOR = 'BLACK'; /* change, if desired */
TEXT = 'V'; /* this marker is a star */
POSITION = '5'; /* adjust for limited reposition */
X = X - 0.005; /* shift left for visibility */
RUN;

DATA  ANNODATA; /* append city marker to anno */
SET  ANNODATA  CITYSTAR;
RUN;

 /* title & footnote statements */
LEGEND1 LABEL=NONE VALUE=(F=CENTX) SHAPE=BAR(3.4,0.8) DOWN=1;
PROC GMAP DATA=INDATA MAP=MBC_US ALL;
     ID STATE;
     CHORO ATTEND /
     ANNO=ANNODATA
     LEGEND=LEGEND1
     DISCRETE
     MISSING
     COUTLINE=BLACK;
FORMAT ATTEND FMTCOL4R.;
RUN;
```

Appendix 4: Program for Four Color Map (Figure 2)

```
 /* goptions statements */

DATA INDATA;
INFILE I;
INPUT @1 STATEABB $2. @60 SHARE $4. @71 SHRCHG $CHAR5.;
STATE = STFIPS(STATEABB);
LENGTH NSHRCHG 5.;
NSHRCHG=SHRCHG;
SHRCHG=LEFT(SHRCHG);
IF SUBSTR(SHRCHG,1,1) ¬= '-' AND NSHRCHG ¬= 0
     THEN SHRCHG= '+' || SHRCHG;
DROP STATEABB;
RUN;

%MAPCOL5C(DATA=INDATA,
          OUT=TOGMAP,
          RESPONSE=NSHRCHG,
          CHOROVAR=GROUP,
          TOPCOUNT=10,        /* same as default */
          CTOPGAIN=GRAY33,
          COTHGAIN=GRAY66,
          CNOCHANG=GRAYFF,
          COTHLOSS=GRAY99,
          CTOPLOSS=GRAYCC)
RUN;

%USANNO3(DATA=INDATA,
         ANNOVAR2=SHARE,
         ANNOVAR3=SHRCHG,
         OUTANNO=ANNODATA,
         NOBOXSTA=NOCHANG, /* because CNOCHANG=WHITE */
         ANNOFONT=NONE,    /* same as default */
         ANNOSIZE=0.90,
         OUTMAP=MBC_US,
         BLANK=YES);
RUN;

 /* title & footnote statements */
LEGEND1 LABEL=NONE VALUE=(F=CENTX) SHAPE=BAR(3.4,0.8) DOWN=1;
PROC GMAP DATA=TOGMAP MAP=MBC_US ALL;
     ID STATE;
     CHORO GROUP /
     ANNO=ANNODATA
     LEGEND=LEGEND1
     DISCRETE
     MISSING
     COUTLINE=BLACK;
FORMAT GROUP FMTCOL5C.;
RUN;
```

Appendix 5: Program for Five Color Map (Figure 3)

```
%MACRO MAPCOL4R(DATA=,
                CHOROVAR=,
                CHORINCR=1,
                PCTLLOW=20,
                PCTLMID=50,
                PCTLHIGH=80,
                LABELTXT=YES,
                LABVLOW='Very Low',
                LABBELOW='Below Median',
                LABABOVE='Above Median',
                LABVHIGH='Very High',
                COLVLOW=,
                COLBELOW=,
                COLABOVE=,
                COLVHIGH=);

PROC UNIVARIATE DATA=&DATA NOPRINT PCTLDEF=2;
VAR &CHOROVAR;  /* get min, max, percentiles */
OUTPUT OUT=STATS MIN=MIN MAX=MAX PCTLPRE=PCTL
                                 PCTLPTS=&PCTLLOW
                                         &PCTLMID
                                         &PCTLHIGH;
RUN;

DATA _NULL_; /* pass range bounds as global variables */
SET STATS;
%GLOBAL MIN B1 B2 B3 B4 B5 B6 MAX;
CALL SYMPUT('MIN',TRIM(LEFT(MIN                         )));
CALL SYMPUT('B1' ,TRIM(LEFT(PCTL&PCTLLOW                )));
CALL SYMPUT('B2' ,TRIM(LEFT(PCTL&PCTLLOW  + &CHORINCR)));
CALL SYMPUT('B3' ,TRIM(LEFT(PCTL&PCTLMID                )));
CALL SYMPUT('B4' ,TRIM(LEFT(PCTL&PCTLMID  + &CHORINCR)));
CALL SYMPUT('B5' ,TRIM(LEFT(PCTL&PCTLHIGH               )));
CALL SYMPUT('B6' ,TRIM(LEFT(PCTL&PCTLHIGH + &CHORINCR)));
CALL SYMPUT('MAX',TRIM(LEFT(MAX                         )));
RUN;

PROC FORMAT; /* format for response ranges & legend text */
%IF %UPCASE(&LABELTXT) EQ YES %THEN %DO;
VALUE FMTCOL4R &MIN - &B1  = "&LABVLOW"
               &B2 - &B3   = "&LABBELOW"
               &B4 - &B5   = "&LABABOVE"
               &B6 - &MAX  = "&LABVHIGH";
                                  %END;
%ELSE %DO;
VALUE FMTCOL4R &MIN - &B1  = "&MIN - &B1"
               &B2 - &B3   = "&B2 - &B3"
               &B4 - &B5   = "&B4 - &B5"
               &B6 - &MAX  = "&B6 - &MAX";
      %END;
RUN;

PATTERN1 V=MSOLID C=&COLVLOW ;
PATTERN2 V=MSOLID C=&COLBELOW;
PATTERN3 V=MSOLID C=&COLABOVE;
PATTERN4 V=MSOLID C=&COLVHIGH;

%MEND MAPCOL4R;
```

Appendix 6: MAPCOL4R Macro

```
%MACRO MAPCOL5C(DATA=,OUT=,RESPONSE=,CHOROVAR=,
                TOPCOUNT=10,
                CTOPGAIN=,COTHGAIN=,CNOCHANG=,
                CTOPLOSS=,COTHLOSS=);

%MACRO PATTERN5(TGCOL=,
                OGCOL=,
                NCCOL=,
                TLCOL=,
                OLCOL=);
%IF %EVAL(&COUNTOG)>0
AND %EVAL(&COUNTOL)>0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 12345 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&OGCOL;
  PATTERN3 V=MSOLID C=&NCCOL;
  PATTERN4 V=MSOLID C=&OLCOL;
  PATTERN5 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTOG)>0
AND %EVAL(&COUNTOL)>0
AND %EVAL(&COUNTNC)=0 %THEN   %DO; /* case 1245 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&OGCOL;
  PATTERN3 V=MSOLID C=&OLCOL;
  PATTERN4 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTOG)>0
AND %EVAL(&COUNTOL)=0
AND %EVAL(&COUNTTL)>0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 1235 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&OGCOL;
  PATTERN3 V=MSOLID C=&NCCOL;
  PATTERN4 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTOG)>0
AND %EVAL(&COUNTOL)=0
AND %EVAL(&COUNTTL)>0
AND %EVAL(&COUNTNC)=0 %THEN   %DO; /* case 125 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&OGCOL;
  PATTERN3 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTOG)>0
AND %EVAL(&COUNTTL)=0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 123 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&OGCOL;
  PATTERN3 V=MSOLID C=&NCCOL; %END;
%ELSE
%IF %EVAL(&COUNTOG)>0
AND %EVAL(&COUNTTL)=0
AND %EVAL(&COUNTNC)=0 %THEN   %DO; /* case 12 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&OGCOL; %END;
%ELSE
%IF %EVAL(&COUNTTG)>0
AND %EVAL(&COUNTOG)=0
AND %EVAL(&COUNTOL)>0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 1345 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&NCCOL;
  PATTERN3 V=MSOLID C=&OLCOL;
  PATTERN4 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTTG)>0
AND %EVAL(&COUNTOG)=0
AND %EVAL(&COUNTOL)>0
AND %EVAL(&COUNTNC)=0 %THEN   %DO; /* case 145 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&OLCOL;
  PATTERN3 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTTG)>0
AND %EVAL(&COUNTOG)=0
AND %EVAL(&COUNTOL)=0
AND %EVAL(&COUNTTL)>0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 135 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&NCCOL;
  PATTERN3 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTTG)>0
AND %EVAL(&COUNTOG)=0
AND %EVAL(&COUNTTL)=0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 13 */
  PATTERN1 V=MSOLID C=&TGCOL;
  PATTERN2 V=MSOLID C=&NCCOL; %END;
%ELSE
%IF %EVAL(&COUNTTG)=0
AND %EVAL(&COUNTOL)>0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 345 */
  PATTERN1 V=MSOLID C=&NCCOL;
  PATTERN2 V=MSOLID C=&OLCOL;
  PATTERN3 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTTG)=0
AND %EVAL(&COUNTOL)>0
AND %EVAL(&COUNTNC)=0 %THEN   %DO; /* case 45 */
  PATTERN1 V=MSOLID C=&OLCOL;
  PATTERN2 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTTG)=0
AND %EVAL(&COUNTOL)=0
AND %EVAL(&COUNTTL)>0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 35 */
  PATTERN1 V=MSOLID C=&NCCOL;
  PATTERN2 V=MSOLID C=&TLCOL; %END;
%ELSE
%IF %EVAL(&COUNTTG)=0
AND %EVAL(&COUNTTL)=0
AND %EVAL(&COUNTNC)>0 %THEN   %DO; /* case 3 */
  PATTERN1 V=MSOLID C=&NCCOL; %END;
%MEND  PATTERN5;

PROC SORT DATA=&DATA OUT=SORTED; BY DESCENDING &RESPONSE;
RUN;
DATA TOPGAIN OTHGAIN NOCHANG OTHER;
RETAIN COUNTTG COUNTOG COUNTNC 0;
IF _N_=1 THEN DO;
CALL SYMPUT('COUNTTG',TRIM(LEFT(COUNTTG))); /* group 1 */
CALL SYMPUT('COUNTOG',TRIM(LEFT(COUNTOG))); /* group 2 */
CALL SYMPUT('COUNTNC',TRIM(LEFT(COUNTNC))); /* group 3 */
CALL SYMPUT('TOPGTEXT',' ');
CALL SYMPUT('OTHGTEXT',' ');
              END;
SET SORTED END=LAST;
IF _N_ <= &TOPCOUNT AND &RESPONSE > 0 THEN DO;
  &CHOROVAR=1;
  COUNTTG=COUNTTG+1;
  OUTPUT TOPGAIN;                              END;
ELSE
IF &RESPONSE > 0 THEN DO;
  &CHOROVAR=2;
  COUNTOG=COUNTOG+1;
  OUTPUT OTHGAIN;      END;
ELSE
IF &RESPONSE = 0 THEN DO;
  &CHOROVAR=3;
  COUNTNC=COUNTNC+1;
  OUTPUT NOCHANG;      END;
ELSE                   DO;
  OUTPUT OTHER;        END;
IF LAST THEN DO;
  CALL SYMPUT('COUNTTG',TRIM(LEFT(COUNTTG)));
  CALL SYMPUT('COUNTOG',TRIM(LEFT(COUNTOG)));
  CALL SYMPUT('COUNTNC',TRIM(LEFT(COUNTNC)));
  IF COUNTTG  = 1
    THEN CALL SYMPUT('TOPGTEXT','Gain');
  ELSE
  IF COUNTOG = 0
    THEN CALL SYMPUT('TOPGTEXT','Gains');
  ELSE
    CALL SYMPUT('TOPGTEXT','Best Gains');
    IF COUNTOG  = 1
      THEN CALL SYMPUT('OTHGTEXT','Other Gain');
    ELSE
      CALL SYMPUT('OTHGTEXT','Other Gains');
             END;
RUN;
PROC SORT DATA=OTHER OUT=SORTED2; BY &RESPONSE;
RUN;
DATA OTHLOSS TOPLOSS;
RETAIN COUNTOL COUNTTL 0;
IF _N_=1 THEN DO;
CALL SYMPUT('COUNTOL',TRIM(LEFT(COUNTOL))); /* group 4 */
CALL SYMPUT('COUNTTL',TRIM(LEFT(COUNTTL))); /* group 5 */
CALL SYMPUT('TOPLTEXT',' ');
CALL SYMPUT('OTHLTEXT',' ');
              END;
SET SORTED2 END=LAST ;
IF _N_ <= &TOPCOUNT THEN DO;
  &CHOROVAR=5;
  COUNTTL=COUNTTL+1;
  OUTPUT TOPLOSS;         END;
ELSE                      DO;
  &CHOROVAR=4;
  COUNTOL=COUNTOL+1;
  OUTPUT OTHLOSS;         END;
IF LAST THEN DO;
  CALL SYMPUT('COUNTOL',TRIM(LEFT(COUNTOL)));
  CALL SYMPUT('COUNTTL',TRIM(LEFT(COUNTTL)));
  IF COUNTTL  = 1
    THEN CALL SYMPUT('TOPLTEXT','Loss');
  ELSE
  IF COUNTOL = 0
    THEN CALL SYMPUT('TOPLTEXT','Losses');
  ELSE
    CALL SYMPUT('TOPLTEXT','Worst Losses');
    IF COUNTOL  = 1
      THEN CALL SYMPUT('OTHLTEXT','Other Loss');
    ELSE
      CALL SYMPUT('OTHLTEXT','Other Losses');
             END;
RUN;
%PATTERN5(TGCOL=&CTOPGAIN,
          OGCOL=&COTHGAIN,
          NCCOL=&CNOCHANG,
          TLCOL=&CTOPLOSS,
          OLCOL=&COTHLOSS);
RUN;
PROC FORMAT;
VALUE FMTCOL5C
     1="&COUNTTG &TOPGTEXT"
     2="&COUNTOG &OTHGTEXT"
     3="&COUNTNC Unchanged"
     4="&COUNTOL &OTHLTEXT"
     5="&COUNTTL &TOPLTEXT"
 ;
RUN;
DATA &OUT; SET TOPGAIN OTHGAIN NOCHANG OTHLOSS TOPLOSS;
RUN;
%MEND MAPCOL5C;
```

Appendix 7: MAPCOL5C Macro

```
%MACRO USANNO3(NOBOXSTA=,
               DATA=,
               ANNOVAR2=,
               ANNOVAR3=@@RANK@@,
               OUTANNO=,
               ANNOFONT=NONE,
               ANNOSIZE=1.00,
               OUTMAP=,
               BLANK=);

PROC  SORT  DATA=&DATA  OUT=VALUES; BY  STATE;
RUN;

%IF  &ANNOVAR3 NE @@RANK@@  %THEN  %GOTO  PASTRANK;

PROC  SORT  DATA=&DATA  OUT=RANKED;
BY  DESCENDING  &ANNOVAR2;
RUN;

DATA  PUTRANK;
SET  RANKED;
FORMAT RANK $2.;
RANK = _N_;
RUN;

PROC  SORT  DATA=PUTRANK  OUT=RANKS; BY  STATE;
RUN;

DATA  VALUES;
MERGE  VALUES  RANKS; BY STATE;
RUN;

%PASTRANK:
  /* vendor-provided coordinates of state centers: raw data */
PROC SORT DATA=MAPS.USCENTER OUT=CENTERAW;
BY  STATE  DESCENDING  OCEAN;
RUN;

  /* user-developed adjustments - see CARDS in paper */
DATA  CHANGES;
INFILE 'G108.SP.SAS.MBCPGMS(USCENFIX)';
INPUT  STATE  OCEAN  $  CHANGEX  CHANGEY;
RUN;

PROC  SORT  DATA=CHANGES  OUT=CHANGES;
BY  STATE  DESCENDING  OCEAN;
RUN;

  /* shift coordinates as user-recommended */
DATA  CENTERS(KEEP=STATE  OCEAN  X  Y);
MERGE  CENTERAW  CHANGES(IN=CHANGE);
BY  STATE  DESCENDING  OCEAN;
IF  CHANGE  THEN  DO;
  X = X + CHANGEX;
  Y = Y + CHANGEY;
END; /* these centers yield maximum blanked area for */
RUN; /* annotation, equidistant from all boundaries */

  /* merge ANNOVAR2 values & state centers */
DATA  MERGED;
MERGE  VALUES(IN=INVALUES) CENTERS; BY  STATE;
IF  INVALUES;
RUN;

  /* this DATA Step improved & extended version of */
  /* code on Page 10 of SAS Technical Report P-196 */
DATA  &OUTANNO;
SET  MERGED;
RETAIN  FLAG;
LENGTH  COLOR  FUNCTION  STYLE  $  8;
IF  _N_ = 1  THEN  FLAG = 0;
XSYS = '2'; /* do not change */
YSYS = '2'; /* do not change */
WHEN = 'A'; /* do not change */
COLOR = 'BLACK';
IF  FLAG = 0  THEN  DO;
  FUNCTION = 'LABEL'; /* this anno record for label */
  STYLE = "&ANNOFONT"; /* font */
  SIZE = &ANNOSIZE; /* fontsize */
  TEXT = FIPSTATE(STATE); /* state abbreviation */
  IF  OCEAN = 'N'  OR  FIPSTATE(STATE) = 'VT'
    THEN  POSITION = '2'; /* 1 cell above center */
  ELSE  POSITION = '3'; /* also justify left */
  OUTPUT;
  TEXT = LEFT(&ANNOVAR2);
  IF  OCEAN = 'N'  OR  FIPSTATE(STATE) = 'VT'
    THEN  POSITION = '5'; /* locate at center */
  ELSE  POSITION = '6'; /* also justify left */
  OUTPUT;
%IF  &ANNOVAR3 NE @@RANK@@  %THEN  %DO;
  TEXT = LEFT(&ANNOVAR3);
                                   %END;
%ELSE                              %DO;
  TEXT = 'R=' || TRIM(LEFT(RANK)); /* supply the rank */
                                   %END;
  IF  OCEAN = 'N'  OR  FIPSTATE(STATE) = 'VT'
    THEN  POSITION = '8'; /* 1 cell below center */
  ELSE  POSITION = '9'; /* also justify left */
  OUTPUT;
END;
ELSE  DO; /* Draw arrow from label in Atlantic Ocean */
  FUNCTION = 'DRAW'; /* to center of coastal state. */
  SIZE = 1.00; /* (adjust thickness, if desired) */
  OUTPUT; /* Triggered by OCEAN=N obs */
```

```
  FLAG = 0; /* after OCEAN=Y same state. */
END;
IF  OCEAN = 'Y'  THEN  DO; /* Position tail of arrow. */
  FUNCTION = 'MOVE';
  IF  FIPSTATE(STATE) NE 'VT'
    THEN  X = X - .005;
  ELSE  Y = Y - .020;
  OUTPUT;
  FLAG = 1;
END;
RUN;

%IF &BLANK NE YES %THEN %GOTO LATER;

PROC  FREQ  DATA=MAPS.US;
TABLES  STATE*SEGMENT  /  OUT=COUNTS  NOPRINT;
RUN; /* count points on each segment of each state */

PROC  SORT  DATA=COUNTS;
BY  STATE  DESCENDING  COUNT;
RUN; /* put most complex segment in state first */

DATA  BIGSEG; /* keep biggest segment for each state */
SET  COUNTS; /* (there are only 5 multi-segment states) */
BY  STATE  DESCENDING  COUNT;
IF  (FIPSTATE(STATE) NE 'MI'  AND  FIRST.STATE)
OR  (FIPSTATE(STATE) EQ 'MI'  AND  NOT FIRST.STATE);
RUN; /* usually most complex segment is biggest */

DATA  DELOCEAN; /* drop states with anno in ocean */
SET  CENTERS;
IF  FIPSTATE(STATE) = 'HI'  THEN  DELETE;
IF  OCEAN = 'Y'  OR  LAG(OCEAN) = 'Y'  THEN  DELETE;
RUN; /* coastal states have two sets of coordinates */

DATA  INNER(KEEP=STATE  SEGMENT  X  Y);
MERGE  BIGSEG  DELOCEAN(IN=NOOCEAN); BY  STATE;
IF  NOOCEAN;
RUN; /* need coordinates only for states with blanking */

DATA  BOXES(KEEP=STATE  SEGMENT  X  Y);
SET  INNER; /* make boxes for inner states */
RX=X;  RY=Y; /* save coordinates of center */
X=.;  Y=.;  OUTPUT; /* counterintuitive, but necessary */
SELECT (FIPSTATE(STATE)); /* action depends on state */
WHEN ('FL') DO;
  X=RX-.0140;  Y=RY+.0125;  OUTPUT; /* upper left */
  X=X;  Y=Y -.0280;  OUTPUT; /* lower left */
  X=X +.0280;  Y=Y;  OUTPUT; /* lower right */
  X=X;  Y=Y +.0205;  OUTPUT; /* upper right point 1 */
  X=X -.0043;  Y=Y +.0075;  OUTPUT; /* upper right point 2 */
  X=X -.0237;  Y=Y;  OUTPUT; /* upper left, again */
END;
WHEN ('TN') DO;
  X=RX;  Y=RY+.0125;  OUTPUT; /* upper left point 1 */
  X=X -.0140;  Y=Y -.0010;  OUTPUT; /* upper left point 2 */
  X=X;  Y=Y -.0270;  OUTPUT; /* lower left */
  X=X +.0280;  Y=Y;  OUTPUT; /* lower right */
  X=X;  Y=Y +.0280;  OUTPUT; /* upper right */
  X=X -.0140;  Y=Y;  OUTPUT; /* upper left point 1, again */
END;
WHEN ('WV') DO;
  X=RX-.0073;  Y=RY+.0125;  OUTPUT; /* upper left point 1 */
  X=X -.0020;  Y=Y -.0015;  OUTPUT; /* upper left point 2 */
  X=X -.0002;  Y=Y -.0058;  OUTPUT; /* upper left point 3 */
  X=X -.0045;  Y=Y +.0005;  OUTPUT; /* upper left point 4 */
  X=X;  Y=Y -.0212;  OUTPUT; /* lower left */
  X=X +.0280;  Y=Y;  OUTPUT; /* lower right */
  X=X;  Y=Y +.0280;  OUTPUT; /* upper right */
  X=X -.0207;  Y=Y;  OUTPUT; /* upper left point 1, again */
END;
OTHERWISE DO;
  X=RX-.0140;  Y=RY+.0125;  OUTPUT; /* upper left */
  X=X;  Y=Y -.0280;  OUTPUT; /* lower left */
  X=X +.0280;  Y=Y;  OUTPUT; /* lower right */
  X=X;  Y=Y +.0280;  OUTPUT; /* upper right */
  X=X -.0280;  Y=Y;  OUTPUT; /* upper left, again */
END;
END;
RUN; /* all state fixes above specific to box size */

%IF &NOBOXSTA EQ %THEN %GOTO LATER;

PROC  SORT  DATA=&NOBOXSTA  OUT=NOBOXES; BY  STATE;
RUN;

DATA  BOXES;
MERGE  BOXES  NOBOXES(IN=NOBOX); BY  STATE;
IF  NOBOX  THEN  DELETE;
RUN;

%LATER:

DATA  &OUTMAP;
SET  MAPS.US
%IF  &BLANK EQ YES  %THEN  %DO;
  BOXES
%END;
          ;
BY  STATE  SEGMENT;
RUN; /* interleave map & box coordinates */

%MEND USANNO3;
```

Appendix 8: USANNO3 Macro

```
DATA INDATA;
INFILE I;
INPUT STATEABB $ ATTEND;
STATE = STFIPS(STATEABB);
RUN;
%MAPCOLNC(DATA=INDATA,CHOROVAR=ATTEND,NCLUS=7)
RUN;
 /* goptions & title statements here */
LEGEND1 LABEL=NONE VALUE=(F=CENTX H=0.8)
        SHAPE=BAR(3.4,0.64) DOWN=2;
PROC GMAP DATA=INDATA MAP=MAPS.US ALL;
ID STATE;
CHORO ATTEND / LEGEND=LEGEND1 DISCRETE COUTLINE=BLACK;
FORMAT ATTEND FMTCOLNC.;
RUN;
```

Appendix 9: Program for Seven Color Map (Figure 4)

```
%MACRO MAPCOLNC(DATA=,CHOROVAR=,NCLUS=,METHOD=CENTROID,
USERCOLS=NO,USERCOL1=WHITE,USERCOL2=CYAN,USERCOL3=MAGENTA,
USERCOL4=YELLOW,USERCOL5=BLUE,USERCOL6=GREEN,USERCOL7=VIOLET);
PROC CLUSTER DATA=&DATA OUTTREE=C NOPRINT METHOD=&METHOD;
VAR &CHOROVAR;
RUN;
PROC TREE DATA=C OUT=T N=&NCLUS NOPRINT;
COPY &CHOROVAR;
RUN;
PROC SORT DATA=T OUT=S; BY CLUSTER &CHOROVAR;
RUN;
PROC UNIVARIATE DATA=S NOPRINT;
BY CLUSTER;
VAR &CHOROVAR;
OUTPUT OUT=U MIN=MIN MAX=MAX;
RUN;
PROC SORT DATA=U OUT=TOFORMAT; BY MIN;
RUN;
%MACRO VARLIST(VARCOUNT=1);
%DO I = 1 %TO &VARCOUNT;
 B&I
%END;
%MEND  VARLIST;
DATA _NULL_;
SET TOFORMAT;
%GLOBAL %VARLIST(VARCOUNT=2*&NCLUS) ;
CALL SYMPUT('B'||LEFT((2 * _N_) - 1),TRIM(LEFT(MIN)));
CALL SYMPUT('B'||LEFT((2 * _N_)    ),TRIM(LEFT(MAX)));
RUN;
PROC FORMAT;
VALUE FMTCOLNC
%DO I = 1 %TO &NCLUS;
%LET M = %EVAL((2 * &I) - 1);
%LET N = %EVAL((2 * &I)    );
  %IF "&&B&M" ¬= "&&B&N" %THEN %DO;
   &&B&M - &&B&N = "&&B&M"'-'"&&B&N"
  %END;
  %ELSE                         %DO;
   &&B&M - &&B&N = "&&B&M"
  %END;
%END;
 ;
RUN;
PATTERN1 V=MSOLID C=&USERCOL1;
%IF &USERCOLS=YES %THEN %DO;
PATTERN2 V=MSOLID C=&USERCOL2;
PATTERN3 V=MSOLID C=&USERCOL3;
PATTERN4 V=MSOLID C=&USERCOL4;
PATTERN5 V=MSOLID C=&USERCOL5;
PATTERN6 V=MSOLID C=&USERCOL6;
PATTERN7 V=MSOLID C=&USERCOL7;
                          %END;
%IF &NCLUS=7 %THEN %DO;
PATTERN7 V=MSOLID C=GRAY33; /* 80% black fill */
PATTERN6 V=MSOLID C=GRAY59; /* 65% black fill */
PATTERN5 V=MSOLID C=GRAY80; /* 50% black fill */
PATTERN4 V=MSOLID C=GRAYA6; /* 35% black fill */
PATTERN3 V=MSOLID C=GRAYCC; /* 20% black fill */
PATTERN2 V=MSOLID C=GRAYF2; /*  5% black fill */
                     %END;
%ELSE %IF &NCLUS=6 %THEN %DO;
PATTERN6 V=MSOLID C=GRAY33; /* 80% black fill */
PATTERN5 V=MSOLID C=GRAY66; /* 60% black fill */
PATTERN4 V=MSOLID C=GRAY99; /* 40% black fill */
PATTERN3 V=MSOLID C=GRAYCC; /* 20% black fill */
PATTERN2 V=MSOLID C=GRAYF2; /*  5% black fill */
                            %END;
%ELSE %IF &NCLUS=5 %THEN %DO;
PATTERN5 V=MSOLID C=GRAY33; /* 80% black fill */
PATTERN4 V=MSOLID C=GRAY66; /* 60% black fill */
PATTERN3 V=MSOLID C=GRAYB3; /* 30% black fill */
PATTERN2 V=MSOLID C=GRAYE6; /* 10% black fill */
                            %END;
%ELSE %IF &NCLUS=4 %THEN %DO;
PATTERN4 V=MSOLID C=GRAY33; /* 80% black fill */
PATTERN3 V=MSOLID C=GRAY80; /* 50% black fill */
PATTERN2 V=MSOLID C=GRAYCC; /* 20% black fill */
                            %END;
%ELSE %IF &NCLUS=3 %THEN %DO;
PATTERN3 V=MSOLID C=GRAY66; /* 60% black fill */
PATTERN2 V=MSOLID C=GRAYB3; /* 30% black fill */
                            %END;
%ELSE %IF &NCLUS=2 %THEN %DO;
PATTERN2 V=MSOLID C=GRAY80; /* 50% black fill */
                            %END;
%MEND  MAPCOLNC;
```

Appendix 10: MAPCOLNC Macro

Figure 1. Map of San Francisco SUGI 14 Attendance, By State PROC GMAP Default Ranges Unacceptable

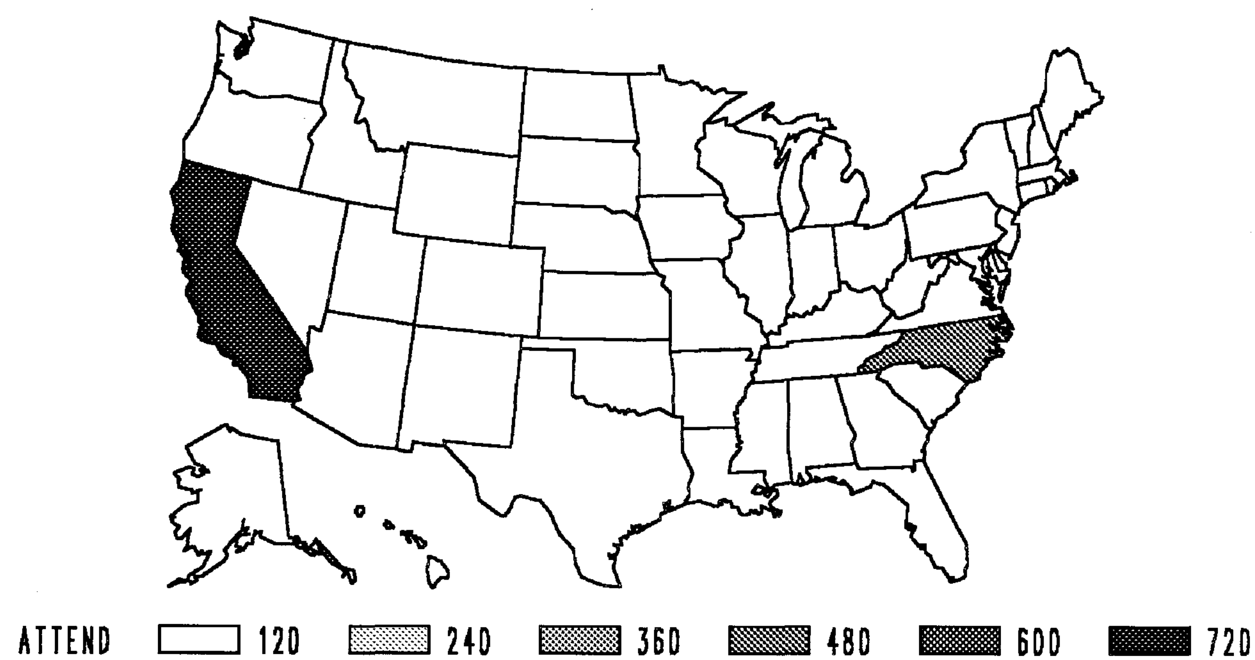

Figure 2. Map of San Francisco SUGI 14 Attendance, With State, Count, Rank, & City Star

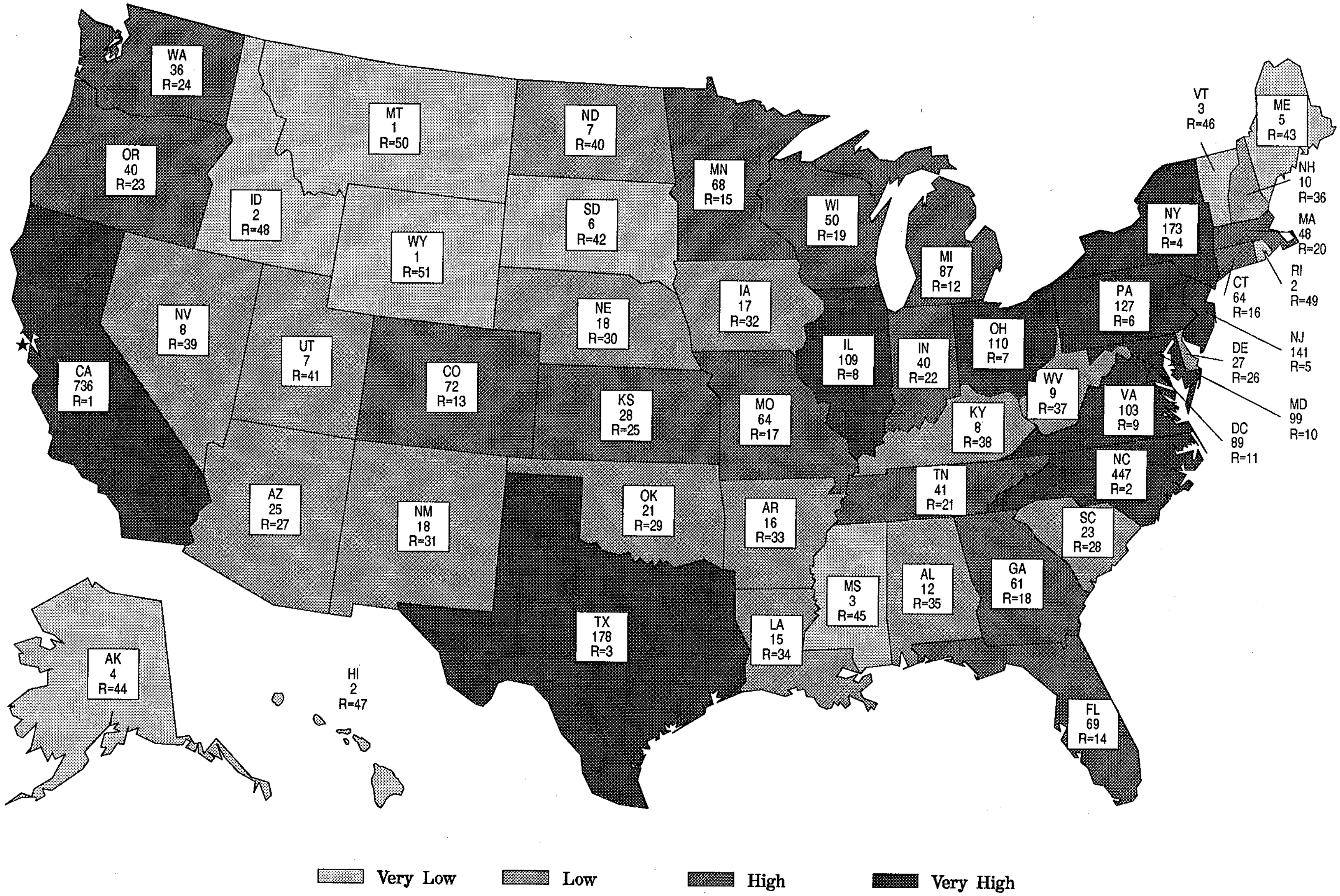

Figure 3. Vespucci Color Map Company – 1994 Sales (in millions of dollars) and Change in Sales vs. 1993

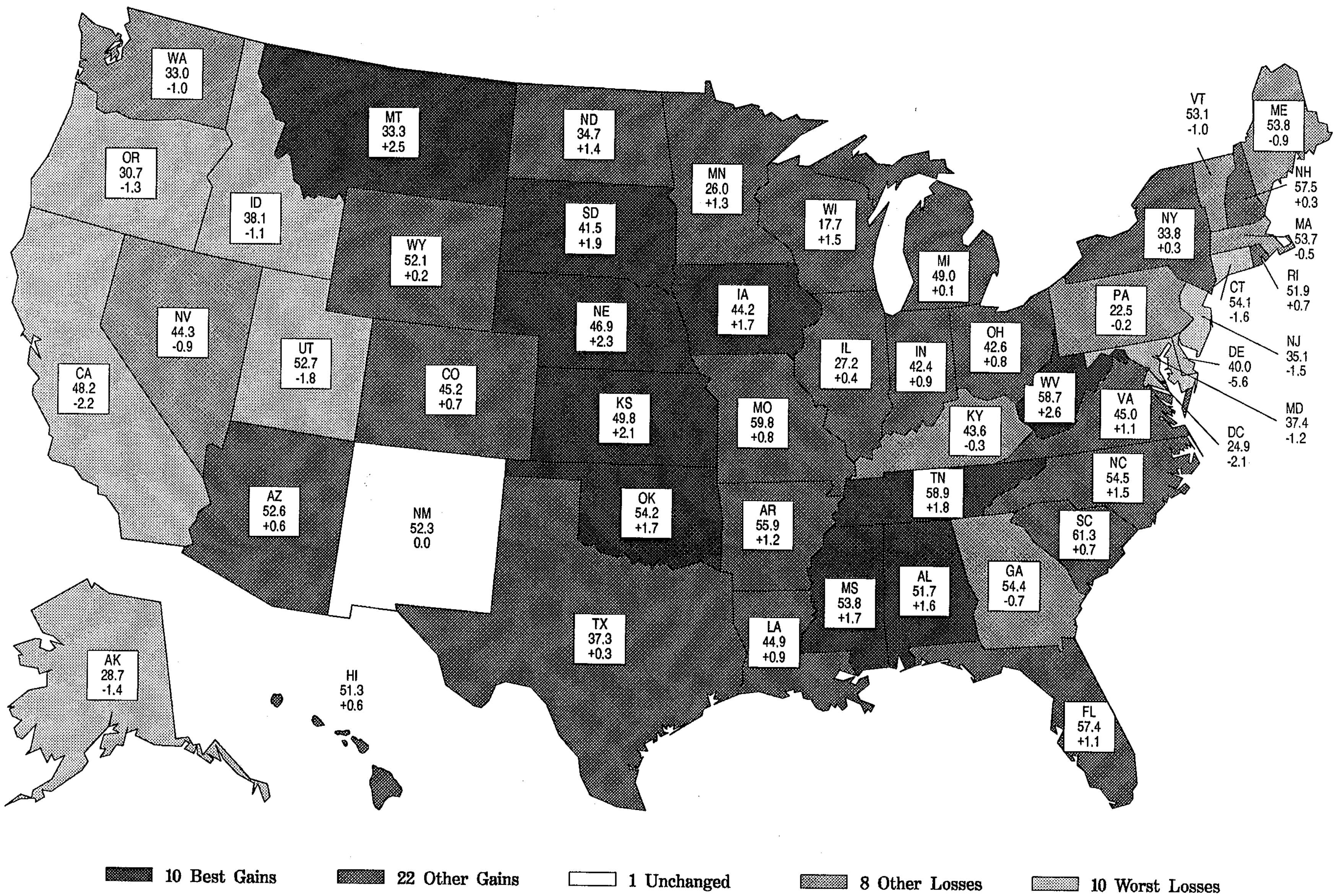

Figure 4. Maps of San Francisco SUGI 14 Attendance, By State – Ranges from PROC CLUSTER METHOD=CENTROID