

Paper RIV-27

How to Make an Impressive Map of the United States with SAS/Graph® for Beginners

Sharon Avrunin-Becker, Westat, Rockville, MD

ABSTRACT

Have you ever been given a map downloaded from the internet and asked to reproduce the same thing using SAS complete with labels and annotations? As you stare at all the previously written, brilliant SAS/Graph conference papers, you start feeling completely overwhelmed. The papers assume you already know how to get started and you feel like a clueless clone not understanding what you are missing. This paper will walk you through the steps to getting started with your map and how to add colors and annotations that will not only impress your manager, but most importantly yourself that you could do it too!

INTRODUCTION

In great anticipation of the 7th Star Wars movie release, a fictional researcher has decided he wants to track characteristics of loyal Star Wars fans. He has developed a survey and has posted it on various fan based sites. He wants to track how many fans have completed the survey across the United States and has divided the states into four regions in hopes of identifying where the most loyal fans live.



Figure 1. Example of a simple map not done with SAS

HOW TO GET STARTED

The first time I tried to create a map using SAS I read some incredible papers on the topic, but I had a hard time getting past the first step. I knew I had a SAS/GRAPH license, but I didn't know how to access any of the maps. These maps are important because they contain the information for SAS to know where to draw the boundary lines. I finally noticed over in the Explorer window under Libraries a directory I had never clicked on before called Maps. The SAS datasets located here are what you need to be able to tell SAS what kind of map you are trying to draw.

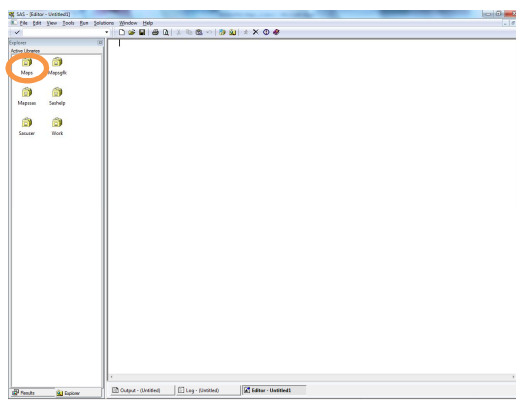


Figure 2. Location of Maps directory

GET COUNTS BY STATE

Now that we know where to find the information for the maps, let's get the data ready to be used. First, we are going to get counts by states and attach the region number to each dataset. Of course, experienced programmers will be able to do all this in one easy PROC SQL but for the beginners I broke it down into two easy steps to better understand what is happening.

```
/* Step 1 – getting counts by region and adding region number */
%macro doit (reg= );
  proc sql;
    create table x&reg as
    select state , count(*) as stcnt
    from soutregion
    where newreg = &reg
    group by state ;
  quit;

  data xx&reg;
    set x&reg;
    region = &reg;
  run;
%mend doit;

%doit (reg=1);
%doit (reg=2);
%doit (reg=3);
%doit (reg=4);
```

/* Step 2 - set your datasets back together */

```
data allx;
  set xx1
      xx2
      xx3
      xx4
  ;
  run;
```

GET COUNTS BY REGION

Let's also grab counts by region which we will use later.

```
proc sql;
  create table z as
  select region, sum(stcnt) as stcntx
  from allx
  group by region ;
quit;
```

FINISH GETTING DATA READY

The next step is to get your data looking a bit better for displaying on a map. First, we will add the variable statecode so we can merge this data in with the maps datasets. Then we will add a zero (0) to any states where a survey may not have been completed. Finally, we will add a new variable called text which contains the state abbreviation and the number of surveys completed for that state.

```

data allx2 ;
  length STATECODE $2.;
  set allx;
  STATECODE = state;
  if stcnt = . then stcnt = 0;
  text=cats(STATE)||' ('||cats(stcnt)||')';

RUN;

```

BEFORE ADDING TEXT AND LEGENDS

Before we start adding in text and labels let's take a quick look at what the code and map looks like at this point if we plan to shade our map by regions. The map dataset US will be used to define the longitude and latitude lines for each state. The CHORO option below specifies the variable we want to use to determine the shadings in the choropleth map.

```

proc gmap data=allx2 map=maps.us;
  id STATECODE;
  choro region ;
run;
quit;

```

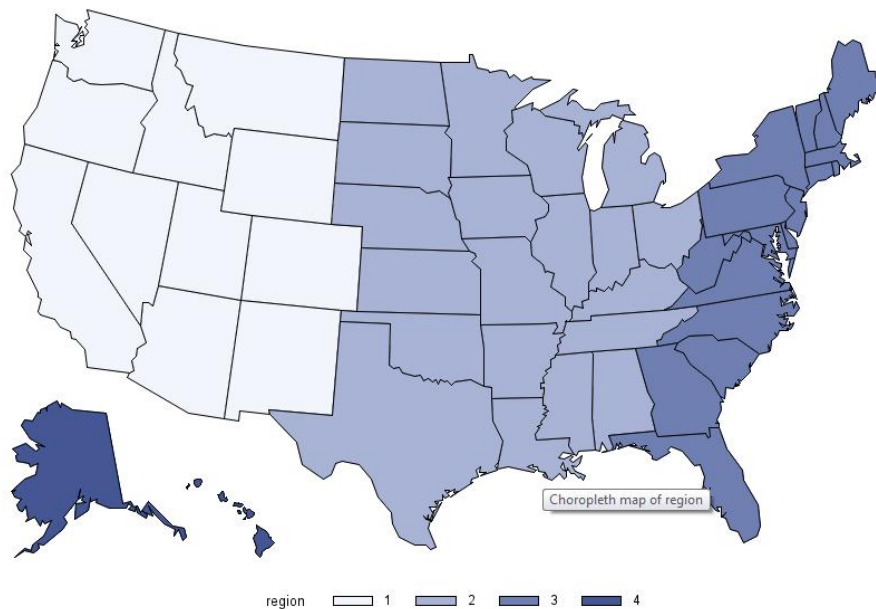


Figure 3. Output from PROC GMAP

READY TO USE THE MAP DATA

We are finally ready to start using the maps datasets. First we will sort the datasets so they can be merged together. We will need to first merge in US2 which contains the FIPS codes for the states (the variable for the FIPS codes is called STATE).

```

proc sort data=allx2 out=sallx2(drop=state); by STATECODE; run;
proc sort data=maps.us2 out=sus2(keep=STATE STATECODE); by STATECODE; run;

data mapfips;
  merge sallx2
        sus2
        ;
  by STATECODE;

```

```
run;
```

THEN WE WILL NEED TO MERGE IN WITH USCENTER TO GET THE PROJECTED LONGITUDES AND latitudes for the radians (these variable are named X and Y), so SAS knows where to place the text.

```
proc sort data=mapfips out=smapfips; by STATE; run;
proc sort data=maps.uscenter out=suscenter(keep=STATE X Y) nodupkey;
      by STATE; run;

data mapfips2;
  merge smapfips (in=a)
        suscenter (in=b)
  ;
  by STATE;
  if a;
run;
```

This data step defines where and when the labels will be placed (in this case the labels will be centered on each states X–Y axis after the map boundaries are drawn).

```
data stlabel;
  length function $ 8 position $ 1
        text $ 20 style $ 30;
  set mapfips2;
  retain
xsys ysys '2'
hsys '3'
position '5'
function 'label'
size 1.40
color 'black'
style "'Tahoma/bo'"
when 'a'
;
run;
```

Finally we run the same code as before with the annotation specified

```
proc gmap data=allx2 map=maps.us;
  id STATECODE;
  choro region / annotate=stlabel ;
run;
quit;
```

HERE IS THE OUTPUT...

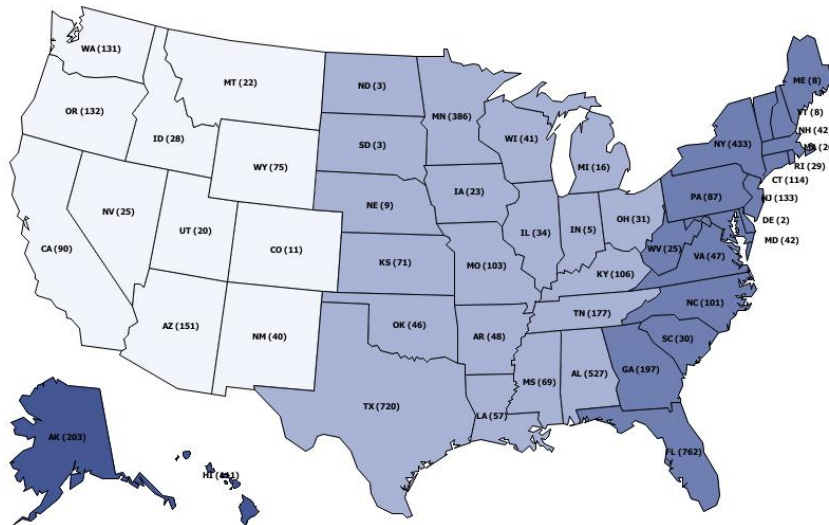


Figure 4. Output from PROC GPLOT with annotation specified

SPECIFYING COLORS

If you would like to switch your colors you can do this by adding in a pattern statement. The v in the code below is the value statement specifying you want a solid pattern.

THE COLORS ARE SPECIFIED...

```
pattern1 v=ms c=yellow;
pattern2 v=ms c=lightGreen;
pattern3 v=ms c=lightPurple;
pattern4 v=ms c=lightblue;
```

HERE IS THE OUTPUT...

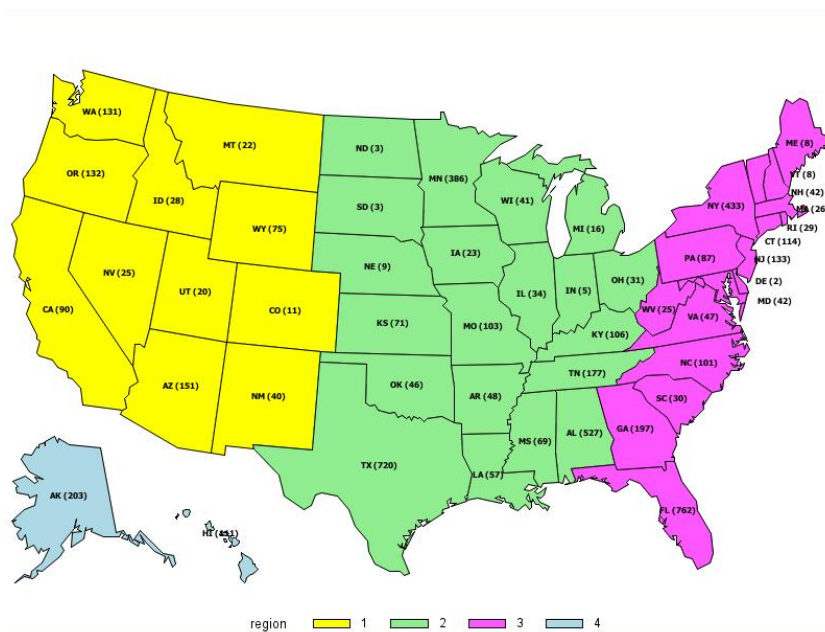


Figure 5. Output from PROC GMAP with color specified

ADDING IN A LEGEND

As nice as this table looks, you have no idea what these colors represent. Now is the time to add in a legend. The legend has lots of options associated with it. You can assign where on the graph you want to display your legend, you can specify the labels on the values, and you can also have a frame around your legend. You can even add in macro variables to display totals. Don't forget that the legend statement needs to be added to your PROC GMAP code for it to be displayed.

THE SIMPLE LEGEND SPECIFICS...

```
LEGEND1 LABEL=(HEIGHT=1 POSITION=TOP JUSTIFY=CENTER
"Total Star Wars Fans Across All States")
      VALUE=(HEIGHT=1 "R2D2 Region" "Ewok Region"
              "Chewbacca Region" "Yoda Region")
ACROSS=1 DOWN=4 POSITION = (bottom outside right) FRAME
MODE=PROTECT offset=(-25 pct);

proc gmap data=allx2 map=maps.us;
  id STATECODE;
  choro region / annotate=stlabel legend=legend1;
run;
quit;
```

HERE IS THE OUTPUT...

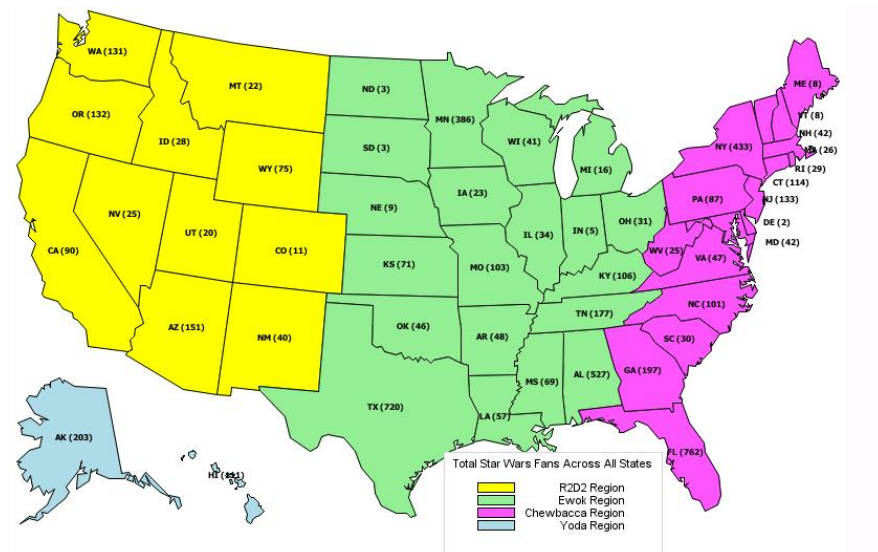


Figure 6. Output from PROC GMAP with legend

ADDING IN MACRO VARIABLES...

```
DATA _null_;
  SET z;
  if region = 1 then CALL SYMPUTX('reg1tot',stcntx);
  if region = 2 then CALL SYMPUTX('reg2tot',stcntx);
  if region = 3 then CALL SYMPUTX('reg3tot',stcntx);
  if region = 4 then CALL SYMPUTX('reg4tot',stcntx);
RUN;
```

```

LEGEND1 LABEL=(HEIGHT=1 POSITION=TOP JUSTIFY=CENTER "Total Star Wars Fans
Across All States")
VALUE=(HEIGHT=1 "R2D2 Region (&regltot.)" "Ewok Region (&reg2tot.)"
          "Chewbacca Region (&reg3tot.)" "Yoda Region (&reg4tot.)")
ACROSS=1 DOWN=6 POSITION = (Bottom outside right) FRAME
MODE=PROTECT offset=(-22 pct);

```

HERE IS THE OUTPUT...

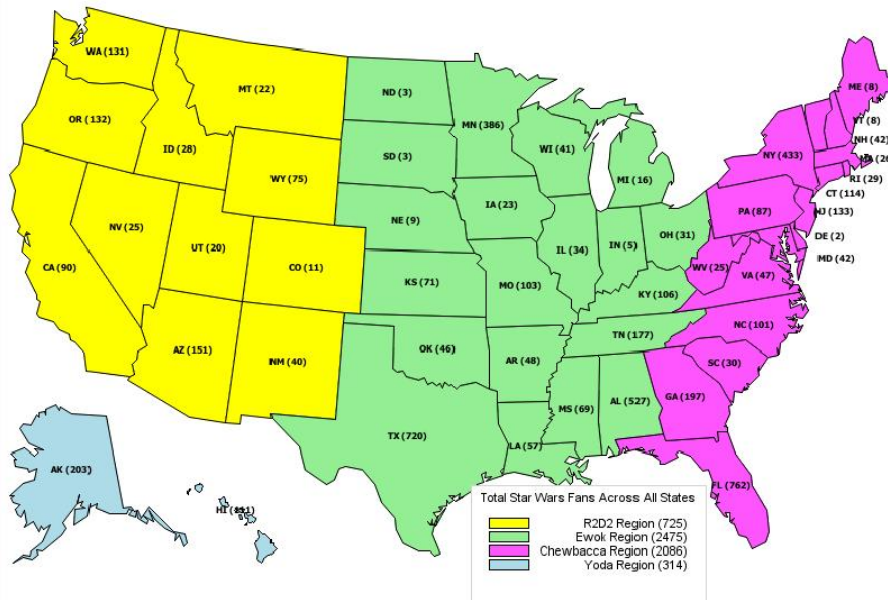


Figure 7. Output from PROC GPLOT with totals specified

ADDING IN MAIN POINTS OF INTEREST

Next we want to see if survey returns were higher in states where a Star Wars experience or convention had occurred within the past 20 years. To do this we select the cities we want from the USCity dataset in the Maps directory. Since some city names may be located in more than one state, you can also specify the state you are interested in. Then you can specify the color, size, position, and shape of your overlay city markers. Then merge it back in to the dataset with your previous state definitions.

```

data citystar;
  length function color $ 8 position $ 1
         text $ 20 style $ 30;
  retain xsys ysys '2' hsys '3' when 'a';
  set maps.uscity(keep=x y city statecode state);
  if (STATECODE='FL' and city='Orlando') or
     (STATECODE='CA' and city='Anaheim') or
     (STATECODE='CO' and city='Denver') or
     (STATECODE='IN' and city='Indianapolis') or
     (STATECODE='CA' and city='Los Angeles') or
     (STATECODE='TX' and city='San Antonio');

  /* Create the observation for a star */
  function='symbol'; style='marker'; text='Z';
  color='red'; size=1.25; output;

  /* Create the observation for a city */
  function='label'; style="'Albany AMT/bold'"; text=city;
  color='red'; size=1.5; position='2'; output;
run;

data stlabel2;

```



```
set citystar
  stlabel (drop=stcnt region)
;
run;
```

HERE IS THE OUTPUT...

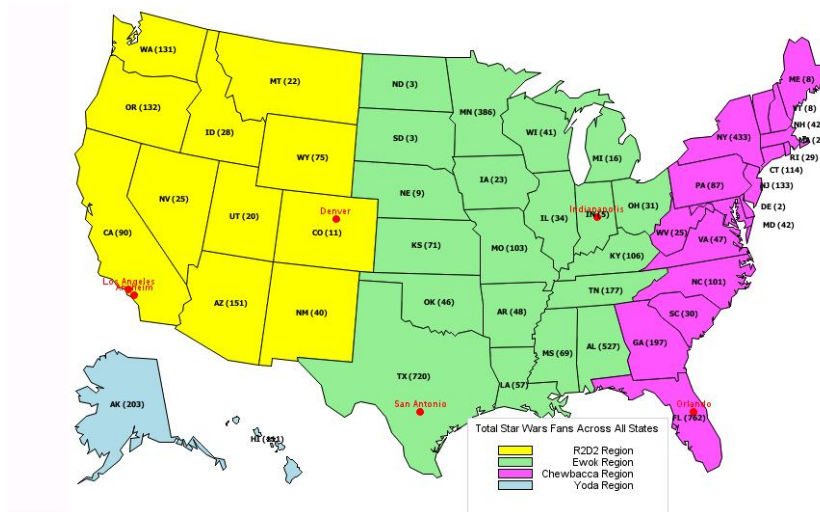


Figure 8. Output from PROC GMAP with target cities defined

ADDING IN TITLES

Titles and footnotes are added the same way in SAS as they always are by using an appropriate title or footnote statement

```
title j=c h=2 "Star Wars Is More Than A Movie, It's An Obsession";
footnote j=l h=1 "Where will you be December 18, 2015?";
```

HERE IS THE OUTPUT...

Star Wars Is More Than A Movie, It's An Obsession

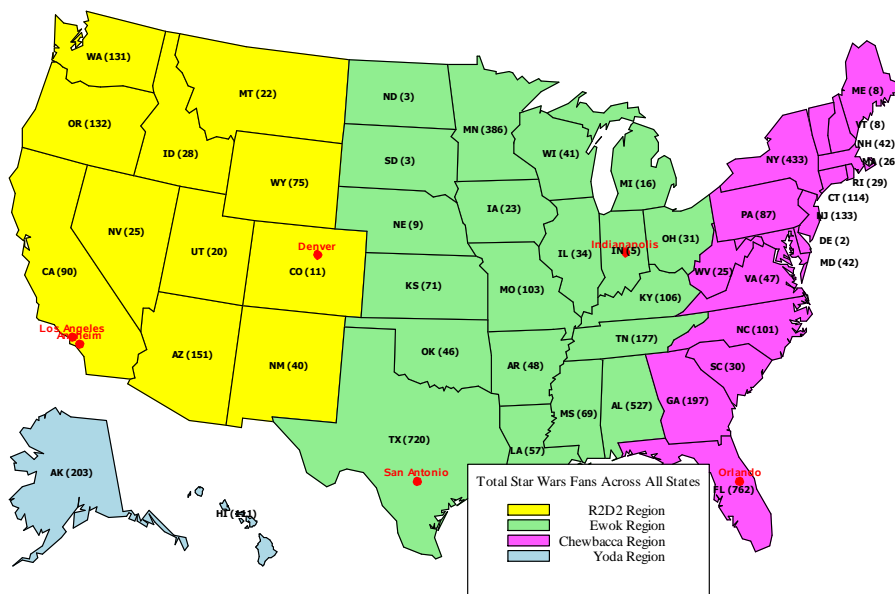


Figure 9. Output from PROC GMAP with a title and footnote

CONCLUSION

This paper touched on some basic code to get the beginner SAS programmer started on creating maps. Sometimes the information out there can be so thorough but complex that it is hard to break down the first few steps to even know how to begin programming. Once you have the basic understanding of how the code comes together you can turn to other papers and start creating even more complex maps.

REFERENCES

Zdeb, M. "The Basics of Map Creation with SAS/GRAPH", SUGI 29

Eberhart, M. "Make the Map You Want with PROC GMAP and the Annotate Facility", NESUG 2008

ACKNOWLEDGMENTS

Many thanks to Michael Raithel, Rick Mitchell, and Sarah Woodruff.

CONTACT INFORMATION

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

Sharon Avrunin-Becker

Westat

1650 Research Blvd. – WB432

Rockville, MD 20850

Work Phone: (301) 738-8384

Email: SharonAvruninBecker@westat.com

DISCLAIMER

The contents of this paper are the work of the author and do not necessarily represent the opinions, recommendations, or practices of Westat.

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.