

Creating Great Maps in ODS Graphics using PROC SGMAP

Kelly Mills, SAS Institute Inc.

ABSTRACT

With the introduction of the SGMAP procedure to SAS® ODS Graphics, geographic mapping has never been easier. This paper shows you how to unlock the geographic potential of your data in just a few statements to create rich, meaningful maps showing locations, regions, pathways, and labels. By mixing and combining statements, you can highlight important aspects of your data or reveal information hidden in the surrounding geography. With PROC SGMAP, the world is at your fingertips!

INTRODUCTION

PROC SGMAP is available in Base SAS® and has similar functionality of the GMAP procedure in addition to annotating items on your map.

Here is what can be accomplished with PROC SGMAP:

- Use the following to map backgrounds and regions:
 - Esri background maps
 - OpenStreetMap
 - Choropleth maps
- Use SGMAP statements to place data on maps:
 - SCATTER
 - BUBBLE
 - TEXT
 - SERIES
- The MAPSSAS data sets can be downloaded from SAS Maps Online with a Base SAS license for creating PROC SGMAP choropleth maps.
- PROC SGMAP and other mapping procedures are available in Base SAS:
 - the GEOCODE procedure
 - the GINSIDE procedure
 - the GPROJECT procedure
 - the GREDUCE procedure
 - the GREMOVE procedure
 - the MAPIMPORT procedure (for shape files)
 - the IMPORT procedure (for CSV files)
 - the %CENTROID macro

A SIMPLE EXAMPLE

Using your own data, you can easily plot data on a background map (Figure 1):

```
data large_cities;
  length city $20 population 8.;
  infile datalines dlm=',';
  input city population;
  state = 'TX';
datalines;
Houston, 23000000000
San Antonio, 15000000000
Dallas, 13000000000
Austin, 950000
Fort Worth, 854000
;

/* PROC GEOCODE to find latitude and longitude of cities */
proc geocode method=city
  data=large_cities
  out=geo_cities
  lookupcity=sashelp.zipcode;
run;

title 'Five Largest Cities in Texas';
proc sgmap plotdata=geo_cities;
  openstreetmap;
  bubble x=x y=y size=population;
run;
quit;
```



Figure 1. PROC SGMAP BUBBLE Plot on OpenStreetMap

OPENSTREETMAP, ESRI MAPS, AND CHOROPLETH MAPS

PROC SGMAP must use either the OpenStreetMap background, the Esri Map background, or a choropleth map.

OPENSTREETMAP

Specify the OpenStreetMap with the following code:

```
proc sgmap plotdata=geo_cities;
  openstreetmap;
  bubble x=x y=y size=population;
run;
```

ESRI MAPS

To use an Esri background map with PROC SGMAP, the "esrimap url=" parameter is needed:

```
proc sgmap plotdata=geo_cities;
  esrimap url='http://services.arcgisonline.com/arcgis/rest/services/
World_Street_Map';
  bubble x=x y=y size=population;
run;
```

Esri maps are used as the background for PROC SGMAP output by specifying a URL. Esri map services can be hosted on-premises to allow for custom background map creation. If an on-premises Esri server is not available, Esri provides a set of map services for public use. In either case, the format of the Esri map service URL is the same:

<http://<servername>/arcgis/rest/services/<mapServiceName>>

To switch the map type using links that Esri provides, go to <http://services.arcgisonline.com/arcgis/rest/services>, click on the different Services, and then remove "/MapServer" from the end of the URL.

For example, the link for ESRI Imagery World 2D is http://services.arcgisonline.com/arcgis/rest/services/ESRI_Imagery_World_2D/MapServer.

To use this link, set the above ESRIMAP URL in PROC SGMAP with the following statement:

```
esrimap url='http://services.arcgisonline.com/arcgis/rest/services/
ESRI_Imagery_World_2D';
```

CHOROPLETH MAPS

Choropleth maps can be created using shapefiles, CSV Files, the MAPSSAS data sets, or the MAPSGFK data sets that are shipped with SAS/GRAPH®.

The PROC SGMAP CHOROMAP statement is used to generate choropleth areas and can draw outlines of the areas. The choropleth areas can also have fill colors based on numeric or character response variables.

A choromap with no response variable draws only the outlines of areas:

```
proc sgmap mapdata=shapefiledata;
  choromap / mapid=mapid_var;
run;
```

The syntax for a CHOROMAP statement with a response variable is:

```
proc sgmap mapdata=shapefiledata maprespdata=respdata;
```

```

    choromap response_var / mapid=mapid_var id=id_var;
run;

```

The MAPID variable is from the MAPDATA data set. The response variable and the ID variable are from the MAPRESPDATA data set. MAPID and ID often are the same variable name.

SHAPEFILES, CSV FILES, AND PROC GPROJECT

The MAPIMPORT procedure is used to import polygon shapefiles:

```

proc mapimport out=campuses
  datafile='u:\dallas\campuses\DISD_Campuses.shp' ;
run;

```

CSV files are imported using the IMPORT procedure:

```

proc import datafile="u:\hurricane2005\Year.2005.ibtracks_wmo.v03r10.csv"
  out=tracks
  dbms=csv
  replace;
  datarow=4;
  getnames=no;
  guessingrows=1000;
run;

```

Unprojected latitude and longitude values must be specified when plotting on OpenStreetMap and Esri Maps. The GPROJECT procedure might be needed to change the plot coordinates.

Here's an example of how this procedure looks. The "id objectid" statement is needed because many shapefiles use the OBJECTID variable to organize data:

```

proc gproject data=data_in out=data_out
  from="<parameters>"
  to="<parameters>";
  id objectid;
run;

```

Shapefiles include a PRJ file. This file includes a line that looks something like the following text:

```
NAD_1983_StatePlane_Texas_North_Central_FIPS_4202_Feet.
```

This is an indication of what parameters should be used for the "gproject from=" setting above.

A good place to find these projections is <http://spatialreference.org>. For the above example, search for NAD_1983_StatePlane_Texas_North_Central_FIPS_4202_Feet on the Spatial Reference website and you'll find ESRI:102738. Click on that and then click the link for the Proj4 to reveal the following text string:

```
+proj=lcc +lat_1=32.13333333333333 +lat_2=33.96666666666667
+lat_0=31.66666666666667 +lon_0=-98.5 +x_0=600000.0000000001 +y_0=2000000
+ellps=GRS80 +datum=NAD83 +to_meter=0.3048006096012192 +no_defs
```

This is the text string that should be used for the "from=" setting in PROC GPROJECT. The "to=" setting should always be EPSG:4326 for OpenStreetMap and ESRI maps in PROC SGMAP.

The PROC GPROJECT for this example is shown in the following code:

```
proc proc gproject data=data_in out=data_out
  from="+proj=lcc +lat_1=32.13333333333333 +lat_2=33.96666666666667
+lat_0=31.66666666666667 +lon_0=-98.5 +x_0=600000.0000000001 +y_0=2000000
+ellps=GRS80 +datum=NAD83 +to_meter=0.3048006096012192 +no_defs "
  to="EPSG:4326";
  id objectid;
run;
```

Another place to find PROJ.4 definitions is in the SASHELP.PROJ4DEF data set. For the **above example, search the description column for 'StatePlane Texas North' and find that ESRI:102738 is a match.**

Use the SASHELP.PROJ4DEF data set in PROC GPROJECT to make the code easier to read:

```
proc gproject data=datain out=dataout
  from="ESRI:102738"
  to="EPSG:4326";
  id objectid;
run;
```

AN EXAMPLE USING SHAPEFILES

This example uses shapefiles downloaded from the Dallas Independent School District website to outline school districts and plot the locations of the high schools (Figure 2):

```
/* Download Dallas, TX Campus Locations shapefiles from here: */
/* https://www.dallasisd.org/Page/29132 */
/* Run the program below to import the shapefile and use SGMMap */

ods _all_ close;
ods html file="campuses.html";

/* Import the scatter data */
proc mapimport out=campuses
  datafile="u:\dallas\campuses\DISD_Campuses.shp";
run;
/* Shapefile is in Texas State Plane Coordinates, */
/* Texas North Central Zone. Reproject this into */
/* WGS 84 for the background map. */
/* PRJ information used from SASHELP.PROJ4DEF. */
proc gproject data=campuses out=campuses_wgs84
  from="ESRI:102738" /* Existing projection */
  to="EPSG:4326"; /* Desired projection */
  id;
run;
/* Only include high schools */
data campuses_wgs84; set campuses_wgs84;
  where level="High";
run;

/* Import the choropleth map data */
proc mapimport out=hs_boundaries
  datafile="u:\dallas\hs_boundaries\DISD_High.shp";
run;
/* Shapefile is in Texas State Plane Coordinates, */
```

```

/* Texas North Central Zone. Reproject this into */
/* WGS 84 for the background map. */
proc gproject data=hs_boundaries out=hs_boundaries_wgs84
      from="ESRI:102738" /* Existing projection */
      to="EPSG:4326"; /* Desired projection */

  id;
run;
/* Create one entry per region */
proc sort data=hs_boundaries_wgs84 out=regions nodupkey;
  by name;
run;

/* Plot the imported shapefiles */
title1 'SGMap with Choromap Outlines and Transparencies';
title2 'Dallas High School Districts and School Locations';
proc sgmap mapdata=hs_boundaries_wgs84
      maprespdata=regions
      plotdata=campuses_wgs84;

  esrimap

  url='http://services.arcgisonline.com/arcgis/rest/services/World_Topo_Map';
  choromap name / mapid=name id=name
      lineattrs=(thickness=2px color=green) transparency=.3
      name="Districts";
  scatter x=x y=y / markerattrs=(symbol=trianglefilled size=10px
      color=red);

  keylegend "Districts" / title="High School";
run;

quit;
ods _all_ close;

```

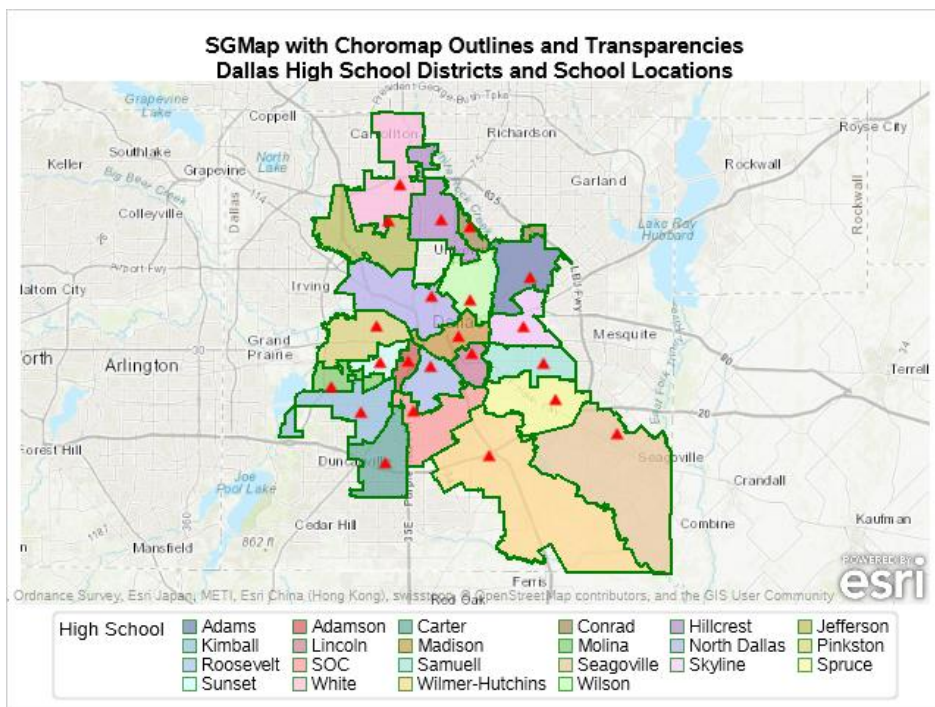


Figure 2: PROC SGMAP CHOROMAP and SCATTER Plot on an Esri Map

PROC SGMAP SERIES PLOTS

Using the PROC SGMAP SERIES statement is a little different from plotting SCATTER, BUBBLE, or TEXT. If there is more than one line segment, then GROUP should be specified or a missing Y-value should be inserted to keep the lines separate.

Here is an example using the GROUP statement (Figure 3):

```
/* Download the 2005 NOAA Hurricane CSV file from here: */
/* https://www.ncdc.noaa.gov/ibtracs/index.php?name=wmo-data */
/* Run the program below to import the shapefile and use SGMMap */
ods _all_ close;
ods html file="hurricanes.html";

/* Import the map data */
proc import
datafile="u:\dallas\hurricane2005\Year.2005.ibtracs_wmo.v03r10.csv"
  out=tracks
  dbms=csv
  replace;
  datarow=4;
  getnames=no;
  guessingrows=1000;
run;
/* Keep only some variables */
data tracks(keep=var3 var4 var5 var6 var7 var9-var11);
  set tracks;
run;
/* Rename variables */
data tracks; set tracks;
  rename var3=num var4=basin var5=sub_basin var6=name var7=time var9=y
  var10=x var11=wind;
run;

/* Limit data to North America */
data tracks; set tracks;
  where basin='NA';
run;

/* Plot the 2005 hurricanes in North America */
title1 'SGMap with Series Plot';
title2 '2005 North America Hurricanes';
proc sgmap plotdata=tracks;
  openstreetmap;
  series x=x y=y / group=name;
run;

quit;
ods _all_ close;
```

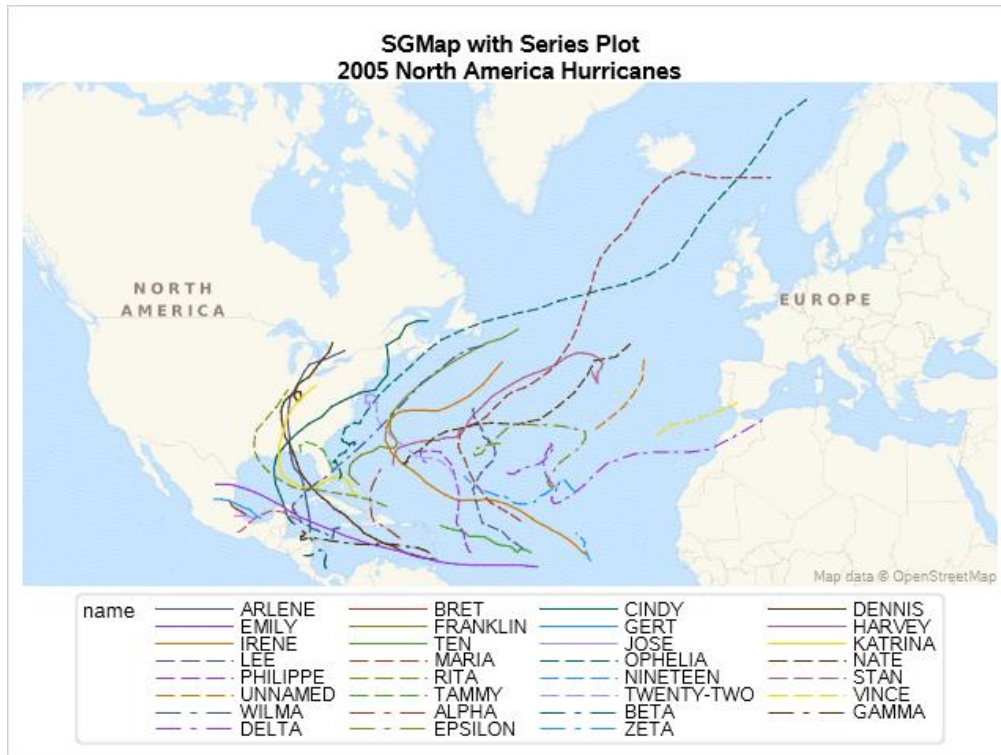


Figure 3: PROC SGMAP SERIES Plot with Groups

If the GROUP statement is not used, then a missing Y-value must be inserted into the data (Figure 4):

```

/* Add missing y value for each series segment */
proc sort data=tracks out=tracks; by name; run;
data tracks; set tracks;
  by name;
  if last.name then do;
    output;
    y=.; output;
  end;
  else output;
run;
run;

/* Plot the 2005 hurricanes in North America */
title1 'SGMap with Series Plot';
title2 '2005 North America Hurricanes';
proc sgmap plotdata=tracks noautolegend;
  openstreetmap;
  series x=x y=y;
run;

```




Figure 4: PROC SGMMap SERIES Plot with No Groups

MAPSSAS DATA SETS FOR CHOROPLETH MAPS

The SAS Maps can be downloaded from the [SAS Maps Online > Downloads > Archived Maps](#) page. Use the [V92 Maps download link](#) to get data sets with a Base SAS license. To use these maps, save the data set to a folder and then use the CIMPORT procedure. Here is an example of importing a SAS Map:

```
libname maps 'C:\maps\v92maps_all';
proc cimport infile='C:\maps\v92maps_all\v92maps_all.cpt'
  library=maps;
run;
```

The GfK GeoMarketing maps are also available on the [Archived Maps page](#), but a SAS/GRAPH license is required.

Here is an example of PROC SGMMap using the MAPSSAS data sets with a choropleth map (Figure 5):

```
/* ***** */
/* This example demonstrates using MAPSSAS datasets with PROC SGMMap */
/* to outline the states and territories of Australia. */
/* */
/* PROC GREDUCE is also used to add the DENSITY variable to the map */
/* dataset so that the jagged coastline can be smoothed. */
/* ***** */
ods _all_ close;
ods html file="australia.html";
```

```

/* Get the MAPSSAS library from your download location */
libname maps "c:\maps\v92maps_all";
proc cimport infile='c:\maps\v92maps_all\v92maps_all.cpt' library=maps;
run;

/* Convert from radians to degrees */
data australia; set maps.austral;
  long = -long * 45/atan(1);
  lat = lat * 45/atan(1);
run;

/* PROC GREduce to add the Density variable with values 1 through 6 */
proc greduce data=australia out=australia_reduced;
  id;
run;

title1 'PROC SGMAP Choromap using MAPSSAS Datasets';
title2 'PROC GREduce used to Smooth Coastline';
proc sgmap mapdata=australia_reduced;
  esrimap url='http://services.arcgisonline.com/arcgis/rest/services/
Canvas/World_Light_Gray_Base';
  choromap / mapid=id
  density=low /* Use lower density for smoothing */
  lineattrs=(color=green thickness=2);
run;

quit;
ods _all_ close;

```

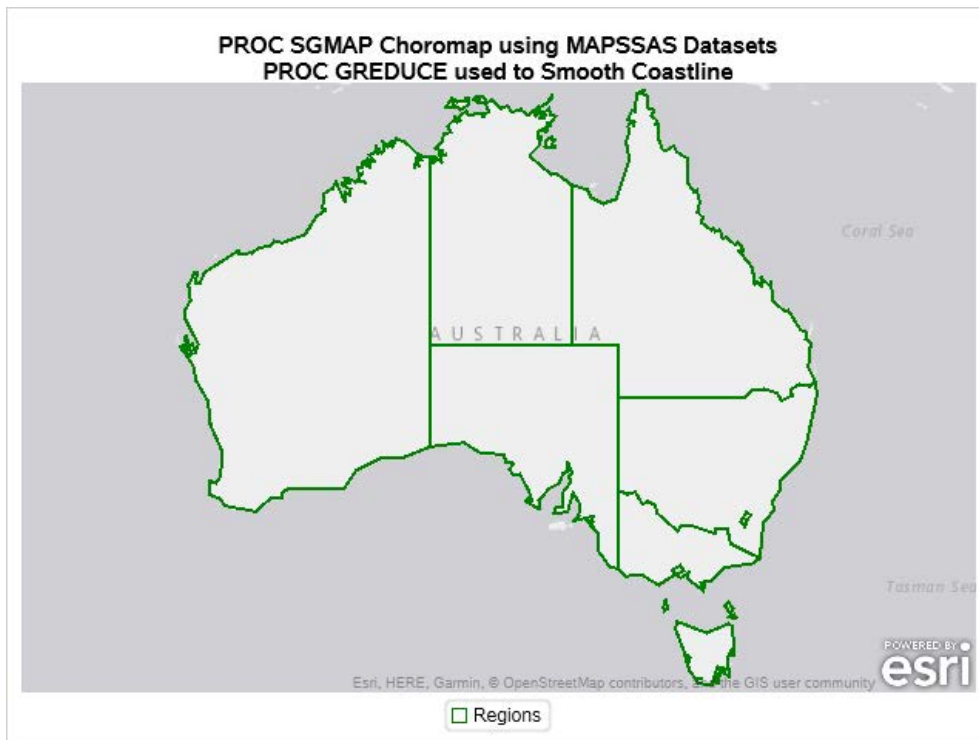


Figure 5: PROC SGMAP Using MAPSSAS Data Sets

CONCLUSION

PROC SGMAP and other mapping procedures are now included with Base SAS and SAS® University Edition.

Using OpenStreetMap and Esri maps allow you to easily produce informative and attractive maps without using SAS/GRAPH annotation.

For more examples of PROC SGMAP and other SAS ODS Graphics, see the latest [SAS/GRAPH and Base SAS 9.4: Mapping Reference](#) and the [Graphically Speaking Blog](#).

ACKNOWLEDGMENTS

The development team at SAS that made PROC SGMAP possible:

- Ed Odom, Sr Software Developer, SAS BI R&D Mapping and Imaging Technology
- Dan Heath, Principal Software Developer, SAS BI R&D Scientific Visualization
- Jeff Phillips, Principal Software Developer, SAS BI R&D Mapping and Imaging Technology
- Prashant Hebbar, Principal Software Developer, SAS BI R&D Scientific Visualization

RECOMMENDED READING

- [SAS/GRAPH and Base SAS 9.4: Mapping Reference](#)
- [SAS Graphically Speaking Blogs](#)

CONTACT INFORMATION

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

Kelly Mills
Kelly.Mills@sas.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.