REVISNG AUTOMAP USING SAS/AF® SOFTWARE

Carol A. Johnstone, Eastman Kodak Company

BACKGROUND

The objective of any mapping package is to display data geographically. Mapping is a powerful presentation format in that the results of analyses dependent in geography can be conveyed in a visual rather than tabular format. However, maps are often overlooked as a presentation tool because they are difficult to create.

The original AUTOMAP system was developed and presented at the SUGI-1984. It has U.S. and world mapping capabilities and was used by analysts at Eastman Kodak Company to display sales territory data and the results of consumer surveys.

The system required knowledge of SAS and specific answers to prompts. After training, analysts were reluctant to use the system as it was difficult to remember the system requirements.

Now converted to SAS/AF®, the system allows greater flexibility for users. A novice with no SAS experience can produce complex maps.

Automap is a user-friendly front-end to SAS/Graph Proc GMAP. It allows the user to create U.S. State/County and World/Regional maps. This paper discusses the components of the system:

• Create a Map Response Data Set
• Mapping Options: U.S. Maps/World Maps
• Viewing Already Created Maps

INTRODUCTION

Proc GMAP requires two SAS data sets to produce a map: a response data set and a map data set. The map data set contains the (x,y) coordinates of the geographical boundaries. The ID variable is the link between the map data set and the response data set. It creates the outline of the map.

<table>
<thead>
<tr>
<th>MAP Data Set: States in U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>-0.16911</td>
</tr>
<tr>
<td>-0.16884</td>
</tr>
<tr>
<td>-0.16894</td>
</tr>
<tr>
<td>-0.16904</td>
</tr>
<tr>
<td>etc.</td>
</tr>
</tbody>
</table>

The response data set contains the ID variable and the variable to be mapped (the response variable). For State/County maps the ID is the FIPS code; World maps, the ID is pre-assigned in the SAS World Map Data set.

<table>
<thead>
<tr>
<th>Response Data Set: Sales Tax by State</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>36</td>
</tr>
<tr>
<td>37</td>
</tr>
<tr>
<td>38</td>
</tr>
<tr>
<td>39</td>
</tr>
</tbody>
</table>

A response data set is a requirement of the system. It can be created within the system or the user can create it before entry.

Using both data sets, PROC GMAP allows one to overlay geographic boundaries and the response data to create the map picture.
USING THE AUTOMAP SYSTEM

The first option of the system is "Create a Map Data Set". It allows the user to select the desired map type and the states or countries for which they have data. (See Figures 1, 2, and 3) A SAS data set is created containing the ID variable and the response variable (variable to be shown on the map).

After the states or countries are selected, the SAS/AF® program screen creates a data set with a null value for the response variable (the value to be mapped). Next, SAS/AF® FSEDIT is automatically executed and the user fills in the value for the response variable. The variable name is preassigned the name RESPVAR and is numeric. In addition to RESPVAR, the FSEDIT screen displays the State or Country ID in both alpha and numeric format. These are protected fields which enable the user to fill in the data values accurately. (See Figure 4) For World Maps, the COUNTRY ID and Country Names are displayed in the same manner. The SAS/AF® screen also has PF Keys defined at the bottom to aid in its operation.

A state map with county boundaries is not an option in this section as it is very cumbersome and unlikely that a user would enter county data interactively. Most county maps are used in territory analysis and use sales data from existing files.

After the response data set is created, Option (2) U.S. Maps or Option (3) World Maps can be selected. (See Figure 1) This option consists of a 3-page SAS/AF® screen which allows the user to select the details to create their map. (See Figures 5, 6, and 7)

Components of this SAS/AF® screen are:

- Map Type
- Response Data Set Name
- SAS Program Filename
- Subsetting the Response Data Set
- Range Labels
- Legend Label
- Device Type
- GOUT Data Set Name
- Titles/Footnotes

MAP TYPE is different for the U.S. State/Counties and World Maps. (See Figures 5 and 6)

The ALL and SELECTED keywords relate to the size of the area that is represented. Example: If user had selected New York State and Map Type option (1) All States in the U.S. -- a U.S. map will display showing New York shaded in with boundaries of the other states. If option (2) Selected States in the U.S. -- New York State will display in larger format. (See Figures 14 and 15)

The map styles CHLOROPLETH, BLOCK, and PRISM are available. The OUTLINE style was added for users who just need an empty map. This was programmed in AUTOMAP by predefining the pattern statements to V=EMPTY and C=WHITE.

The Response Data Set Name and Response Variable Name will be automatically filled in if option (1) Create a Map Data Set was executed.

Users are given the option of saving the SAS program code by specifying their own SAS program filename. The default program name is AUTOMAP. This option gives them the ability to save multiple SAS mapping code in individual files. The SAS code can then be re-used at a later date.

By entering a logical expression, users are able to subset their response data set. This feature allows the user to use an existing response data set, such as all states in the U.S., and create a map of a different subset of the data. For example, a map could be created showing only those states that have a Sales Tax less than 5% instead of all possible Sales Tax ranges in the response data set.

To add a customized legend, the user can put range labels on the response variable. There are error checks built in so that numeric and character variables are formatted accurately. After selecting the appropriate range label type, the user will fill in range labels on a subsequent menu. The values of the variables will be automatically filled in by the system. (See Figures 9, 10, and 11)
A legend label can also be added to the response variable.

The device type is limited to IBM3179G and IBM3279 type terminals, but can be expanded upon request.

A GOUT graphics catalog name is a feature which allows users to save the output produced by GMAP. If additional copies are needed, they can be generated without executing the SAS program again.

A maximum of three titles and footnotes are optional to annotate the map.

See Figures 1 through 16 for an example of the AUTOMAP System Screens and Sample Maps.

INTERNAL SYSTEM

The AUTOMAP system is written in SAS/AF® and uses SAS mapping data set libraries residing on a CMS mini disk. To invoke AUTOMAP a user can just type the command AUTOMAP in CMS.

SAS map data libraries, MAPS and MAPW, were customized for the system. The COUNTIES data set, purchased from SAS Institute Inc., contains unprojected and unreduced coordinates for all states. When using unprojected latitude and longitude coordinates in a two-dimensional plane, the map will be reversed. The U.S. county map was projected, reduced, and then broken down by state. The whole world was broken down by region, then projected and reduced. The data library MAPS contains U.S. state and county boundaries and MAPW contains boundaries for the other countries.

There are four SAS program generators at the core of the system. A user response data set generator, a SAS mapping program generator, SAS format program generator, and a map data set generator.

The format program generator is only executed if the option RANGE LABELS is selected. An error checking routine validates the response variable type to ensure correct format syntax. If the variable is NUMERIC-DISCRETE or CHARACTER, an output data set from Proc Freq is used to create the distribution of values. To put these values on the SAS/AF® screen, the SYMPUT function is used to create SAS macro variables which are displayed. The NUMERIC-CONTINUOUS option formats this type of numeric variable to the range increments the user specifies.

After all prompts are answered, a SAS mapping program is built, default name AUTO$MAP SAS A, on the user's A disk. After creating this program the system executes it and displays the map on the terminal screen.

AUTOMAP SAS/AF® ENHANCEMENTS

The system was redesigned in SAS/AF® to allow users to operate with on-line help and little training. Menu screens and program screens were kept at a minimum to allow the system flow to be easily understood.

It is designed for the novice SAS user, but it can also help the experienced SAS programmer generate skeleton SAS code for mapping applications.

The flexibility of creating the response data set during the AUTOMAP system eliminates the need to remember how to create a map data set containing proper ID variables.

FUTURE ENHANCEMENTS

An option which would allow users to create a U.S. Map with customized region boundaries.

The ability to annotate the created maps interactively with the cursor would be a desirable feature for users. This may be a reality in SAS/Graph version 6.0.

An option to include Alaska and Hawaii on State Maps would also be a desirable feature for users.

Welcome to AUTOMAP

Select an OPTION and press Enter
1 Create a Mapping Data Set
2 U.S. Maps: States/Counties
3 World Maps
4 Hardcopy Output
5 View Created or Example Maps
6 Brief System Definition
7 Exit
PF1=HELP PF3=END

Figure 1. Primary MENU
Create a Map Data Set

Select an OPTION and press Enter
1. US Map - States
2. Whole World
3. Asia, Africa, Australia
4. Africa, Middle East
5. Asia, South Pacific
6. Europe
7. Latin America
8. North America
9. Return to Primary Menu

PF1=HELP PF3=END

Figure 2. Sub-Menu: Create Map Data Set

Select States for Map
Press PF3 to Submit or type Cancel to Return
Response Data Set Name:  
Type an 'X' by the state(s) desired in map

_ALABAMA  _COLORADO
_ARKANSAS  _CONNECTICUT
_ARIZONA  _DELAWARE
_CALIFORNIA  _(etc.)

PF1=Help PF7=Backward PF8=Forward

Figure 3. Creating Response Data Set Screen

Edit SAS Data Set:  
Press PF3 to Submit:
STATE NAME:  NEW YORK
STATE ID:  36
RESPVAR:  
Note: For an Outline Map -> Enter 0

PF3=End PF7=Backward PF8=Forward

Figure 4. Entering Respvar Values

U.S. Map Options

Page 1 of 3

Press PF3 to Submit or type Cancel to Return
Map Type:  
Map Style:  
1. All States in US
2. Selected States in US
3. All Counties in State(s)
4. Selected Counties in State(s)

Response Data Set Name:  
Response Variable Name:  
SAS Program Filename:  
Subset Response Data Set? (Y/N)
If Yes, Enter Logical Expression (Var1 > 10)

PF1=Help PF7=Backward PF8=Forward

Figure 5. Page 1 of US Map Options

U.S. Map Options

Page 2 of 3

Press PF3 to Submit or type Cancel to Return
Range Labels for Response Variable?  
1. No Legend
2. YES, Response Variable NUM (continuous)
3. YES, Response Variable NUM (discrete)
4. YES, Response Variable CHARACTER
5. Use Default

Legend Label for Response Variable? (Y/N)
If YES, Enter Text: (ex. Expenditures 1988)

DEVICE?  
1. IBM3179G  2. IBM3279
GOUT Data Set Name:  

PF1=Help PF7=Backward PF8=Forward

Figure 6. Page 2 of US Map Options

620
U.S. Map Options

Press PF3 to Submit or type Cancel to Return

Titles and Footnotes

1.
2.
3.

Footnotes

1.
2.
3.

PF1=Help PF7=Backward PF8=Forward

Figure 7. Page 3 of US Map Options

World Map Options

Press PF3 to Submit or type Cancel to Return

Map Type: _
Whole World 1 2
Am Region 3 4
Africa, Mid.East 5 6
Asia, S.Pacific 7 8
Europe 9 10
Latin America 11 12
North America 13 14

The only difference between the US and World Map Menus is the Map Type Selection.

PF1=Help PF7=Backward PF8=Forward

Figure 8. World Map Options

Range Labels - Numeric (DISCRETE)

Press PF3 to Submit or type Cancel to Return

Enter the Range Labels for: RESPVAR

Value = Label
1 = NORTHEAST
2 = SOUTHEAST
3 = MIDWEST
4 = CENTRAL
5 = SOUTHWEST
6 = NORTHWEST

PF1=Help PF7=Backward PF8=Forward

Figure 9. Example of Numeric Discrete Range Labels

Range Labels - Numeric (CONTINUOUS)

Press PF3 to Submit or type Cancel to Return

Enter the BREAK POINTS of range:

(minimum)
50
150
360

(maximum)

PF1=Help PF7=Backward PF8=Forward

Figure 10. Example of Numeric Continuous Range Labels

Range Labels-Character Response Variable

Press PF3 to Submit or type Cancel to Return

Enter the Range Labels for: RESPVAR

Value = Label
NE = NORTHEAST
SE = SOUTHEAST
MW = MIDWEST
CN = CENTRAL
SW = SOUTHWEST
NW = NORTHWEST

PF1=Help PF7=Backward PF8=Forward

Figure 11. Example of Character Range Labels

Hard Copy Output

Press PF3 to Submit or type Cancel to Return

GOUT Data Set Name: ___.GMAPS

Hard copy of all Subset of Maps Stored in the above GOUT data set will be sent to the selected plotter below.

Select Maps:
1. ALL
2. SUBSET (must be familiar with GREPLAY commands)

PF1=Help PF7=Backward PF8=Forward

Figure 12. Example of Hardcopy Output Screen
View Created or Example Maps
Select an Option and press Enter
1 Your Created Maps
2 U.S. Maps
3 World Maps
4 U.S. Region Maps
5 Network Maps
6 Return to Primary Menu

Figure 13. Viewing Created or Example Maps Menu

Figure 14. All States in U.S.

Figure 15. Selected States in U.S.

Figure 16. World Maps - Europe All Countries

REFERENCES


SAS and SAS/AF are registered trademarks of SAS Institute Inc., Cary, NC, USA. No endorsement by Eastman Kodak is implied.