

Using SAS/GRAPH® to Create Visualizations That Also Support Tactile and Auditory Interaction

Ed Summers, Julianna Langston, Robert Allison, and Jennifer Cowley,
SAS Institute Inc., Cary, NC

ABSTRACT

Concepts, ideas, and analyses are commonly presented as graphics that are optimized for visual consumption. Does that mean blind students and professionals are out of luck? Not anymore.

This presentation demonstrates best practices for multimodal visualizations for blind users of the iPad and other touchscreen mobile devices. Multimodal visualizations allow blind users to interactively explore visualizations through touch, discover details through sound, and comprehend the essence of the visualizations without vision. The presentation explores how to use SAS/GRAPH® to create data visualizations that are optimized for access by blind users.

For brevity, this paper will focus on scatter plots. However, the live presentation will demonstrate best practices for a variety of common charts, plots, and maps. We will demonstrate how to create multimodal visualizations using SAS macros that encapsulate the best practices. Lastly, we will explore how an auditory channel can improve the usability of visualizations for sighted users.

INTRODUCTION

Data visualizations and infographics are commonly included in mainstream digital content such as Web pages, magazines, newspapers, books, textbooks, academic journals, and standardized tests. These graphics are optimized for visual consumption. Currently, visually impaired students and professionals do not have equal access to this digital content.

Scatter plots embody many of the challenges faced by users with visual disabilities. Scatter plots are used to determine the relationship between two continuous variables of a data set. For example, a scatter plot might be used to understand the relationship between the literacy rate of a nation (measured as a percentage of adults that are literate) and income (measured as the Gross National Income per person in U.S. dollars).

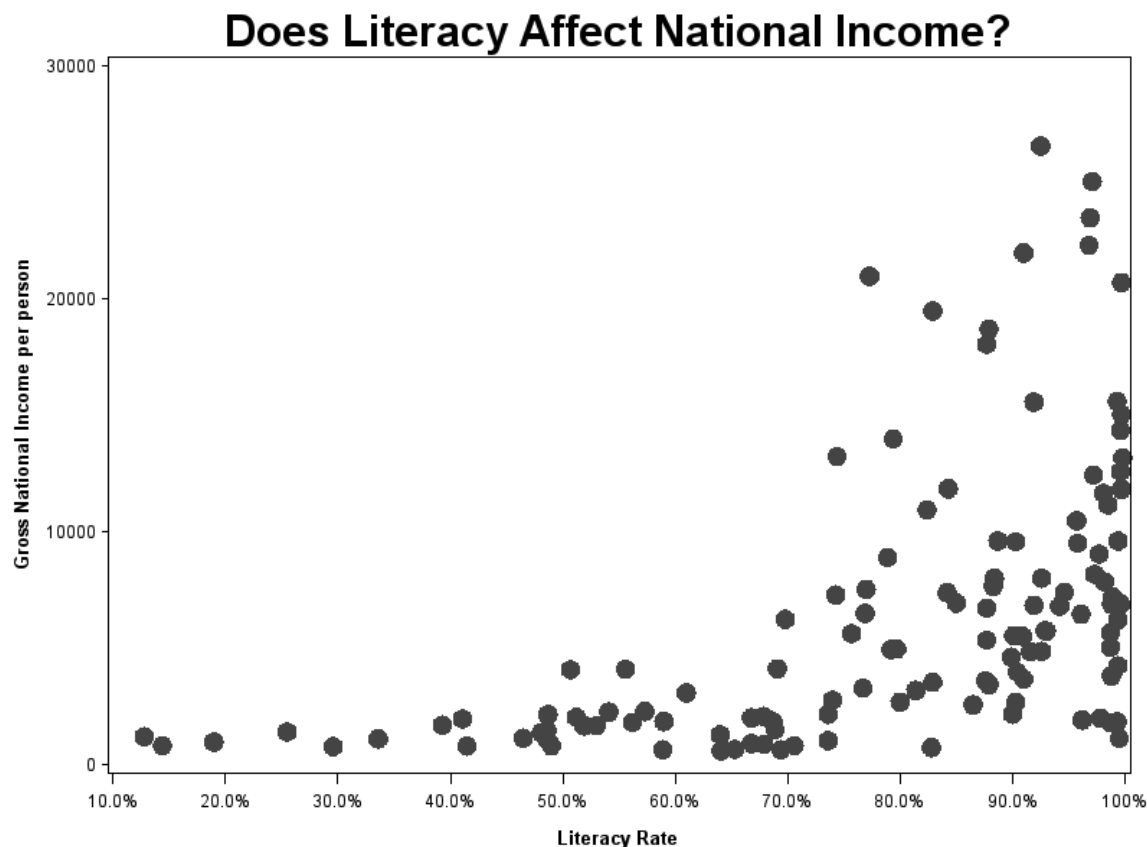


Figure 1: A Scatter Plot That Compares Literacy Rate against Gross National Income per Person

A scatter plot represents data spatially within a Cartesian plane. One variable is represented on the x-axis and another is represented on the y-axis. The location of each observation within the Cartesian plane is determined by the values of the variables represented on each axis.

Sighted users can read the text within a scatter plot to determine the value represented on each axis as well as the scale and unit of measurement. Sighted users can effortlessly identify patterns, dense areas, and outliers within a scatter plot. Within 500 milliseconds of viewing the scatter plot, a sighted user can obtain the “gist” of the relationship between the two variables.¹

We would like to enable equivalent experiences for users with visual impairments. SAS and its customers have a vested interest in the democratization of data. There is a shortage of people in the labor force with the required quantitative skills. For example, the McKinsey Global Institute estimates that the United States alone faces a shortage of 140,000 to 190,000 people with deep analytical skills, as well as 1.5 million managers and analysts to analyze big data and make decisions based on their findings.²

The primary objective of the work presented in this paper is to empower a broader set of users to succeed using SAS software. A secondary objective is to create more engaging experiences for users of all abilities.

REVIEW OF EXISTING INDUSTRY BEST PRACTICES

How would we use existing industry best practices to make the scatter plot in Figure 1 accessible for visually impaired students and professionals?

¹ Rayner, K.

² Manyika, J.

One industry best practice recommends that authors and developers provide a text equivalent for graphics within digital content.^{3 4} The best case outcome is that visually impaired users perceive the graphic second-hand through the experience of another person. As a result, blind users are deprived of the opportunity to form their own conclusions about the data. Furthermore, this approach might be sufficient for simple static graphics such as bar charts and pie charts; however, it is clearly insufficient for complex data visualizations such as scatter plots and interactive graphics that are common on the Web today.

Another best practice is to duplicate the data represented in a graphic in a table within the digital content so visually impaired users can explore the data in the table using a screen reader.⁵ Again, this approach might suffice for very simple data visualizations. However, it is clearly not sufficient for data visualizations, such as Figure 1, that use spatial relationships to convey information or visualizations of more than a few dozen data points. In addition, this approach does not allow blind users to perceive text annotations and other visual cues that are frequently included within graphics.

Another approach is to convert the digital graphic to physical form using an embosser.⁶ This approach provides excellent perception of the spatial relationships within the graphic. Text annotations and other visual cues can be accessed via audio solutions that work in conjunction with embossed tactile forms.^{7 8} However, this approach does not work for interactive graphics in mainstream content. It is also self-evident that the conversion of bits to atoms increases production cost, increases distribution costs, reduces portability, and reduces durability.

DESIGN GOALS

The following design goals guide the development of SAS best practices for mainstream data visualizations that can be accessed by users with visual impairments.

PERCEPTUAL PRECISION

First and foremost, an acceptable design must allow blind users to interactively explore graphics in mainstream digital content with enough perceptual precision to understand features within the graphic, spatial relationships among those features, and other context-sensitive details in a timely manner.

FIRST-HAND ACCESS

It is not acceptable for blind users to access data through the perception and judgment of another person, for example text descriptions in the alt or longdesc HTML attributes. All users need access to the original content so they can form their own conclusions based on their own perceptions. This is particularly important in educational settings.

AFFORDABLE MAINSTREAM HARDWARE

An acceptable design would allow blind users to access graphics using the same affordable mainstream hardware that is used by sighted users rather than expensive special equipment. Procurement departments, teachers, and corporate technical support staff are more likely to be familiar with mainstream devices. As a result, visually impaired students and professionals are more likely to get adequate hardware as well as technical assistance.

BORN-ACCESSIBLE

An acceptable design would allow authors of mainstream content to create graphics that are both compelling for sighted users and accessible for blind users when they are originally created. Ideally, authoring tools would produce accessible graphics by default with no extra effort required on the part of the author.

³ WCAG 2.0

⁴ Rothberg, M.

⁵ Rothberg, M.

⁶ Van Gerven, C.

⁷ Gardner, J. A.

⁸ Miele, J.

CRADLE-TO-GRAVE DIGITAL

An acceptable design would allow blind students and professionals to access graphics in digital form and avoid the cost of printing the graphics on a Braille embosser. However, the user should be able to print the graphics using a Braille embosser if additional perceptual precision is required.

VOICEOVER SCREEN READER

The information presented in this paper leverages the Voiceover screen reader. The Voiceover screen reader is included on all Apple iOS devices. Voiceover allows users with visual impairments to interact with iOS touchscreen devices in ways that are equivalent to the interactions used by sighted users. Voiceover achieves this goal in two ways. First, Voiceover exposes an alternative gesture set to the user and it acts as a gesture translator to the underlying application. Second, Voiceover vocalizes information about each element of an application using text-to-speech software.

To enable Voiceover on any iPad, iPhone, or iPod Touch that is running iOS 4 or later:

1. Open the Settings application.
2. Navigate to General > Accessibility > Voiceover.
3. Enable Voiceover using the toggle button at the top of the page.
4. Press the Voiceover Practice button to experiment with the Voiceover gesture set.

SAS SCATTER PLOT MACRO

We have defined a SAS macro that can be used to create scatter plots that are optimized for blind users of Apple iPad devices. The macro incorporates preliminary best practices for accessible scatter plots. The best practices were developed with the help of blind users and data visualization experts. They were refined in the SAS Usability Lab.

The macro produces graphics that are embedded within Web pages. They can be distributed across a university campus or around the world. Sighted and blind users can access the same graphics using a Web browser. They can both access the graphics first-hand and form their own conclusions about the data based on their own perceptions.

The macro produces an image file and one or more client-side image maps. An image map acts as an overlay that rests on top of the image file. Regions within the image map are annotated with appropriate descriptive text that is read by the Voiceover screen reader when a blind user touches the region.

SYNTAX

```
%scatterPlot(dataset, nameLoc, y_variable, x_variable, <y_denom>, <x_denom>, y_low, y_up, y_incr, x_low, x_up, x_incr, title, dot_size, radius, y_label, x_label);
```

Description of arguments:

- dataset – The name of the data set including the libref value that was specified with a LIBNAME statement.
- nameLoc – The name of the variable that contains the user-visible labels for each observation in the data set.
- y_variable – The name of the variable that will be represented on the Y axis.
- x_variable – The name of the variable that will be represented on the X axis.
- <y_denom> – A text string that contains the unit of measurement of the Y variable, for example, pounds, kilograms, and so on. A value is not required for this argument.
- <x_denom> – A text string that contains the unit of measurement of the X variable, for example, pounds, kilograms, and so on. A value is not required for this argument.
- y_low – A value that denotes the lower bound of the data that will be represented on the Y axis.
- y_up – A value that denotes the upper bound of the data that will be represented on the Y axis.
- y_incr – A value that denotes the increment of the labels on the Y axis.

- `x_low` – A value that denotes the lower bound of the data that will be represented on the X axis.
- `x_up` – A value that denotes the upper bound of the data that will be represented on the X axis.
- `x_incr` – A value that denotes the increment of the labels on the X axis.
- `title` – A text string that denotes the title of the scatter plot.
- `dot_size` – The size in pixels of the dots that are used to represent each observation in the scatter plot.
- `radius` – The size in pixels of the touch targets that are used to represent each observation in the scatter plot. It is usually beneficial for the touch targets to be slightly larger than the visual dots.
- `y_label` – A text string that denotes the label of the Y axis.
- `x_label` – A text string that denotes the label of the X axis.

EXAMPLES

This example plots the Gross National Income against the adult literacy rate for countries represented in the `sashelp.demographics` data set.

```
%scatterPlot(sashelp.demographics, name, gni, adultliteracypct, dollars, percent,
0, 30000, 10000, .1, 1, .1, Does Literacy Affect National Income?, 5, 2, Gross
National Income per Person, Literacy Rate);
```

GRAPH PARTS

There are several parts of each scatter plot that will always be present, so specific treatments of graph parts were developed to be inserted onto every graph.

PREPARATION OF THE ANNOTATE DATASET

The `annotate` facility will be invoked by `PROC Gplot`. Several of the macros that this paper will discuss produce only `DATA` steps that will be concatenated into one `annotate` data set. The following code does just that.

```
data anno_title;
  length html $ 1000 text $ 100 function $10;
  %annotitle;
run;
data anno_y;
  length html $ 1000 text $ 100 function $10;
  %annoy('white', &y_low, &y_up, &y_incr, &y_format, &y_denom, &y_label, 10, 5);
run;
data anno_x;
  length html $ 1000 text $ 100 function $10;
  %annox('white', &x_low, &x_up, &x_incr, &x_format, &x_denom, &x_label, 10);
run;
%ignore;
data anno;
  set ignore_error anno_title anno_y anno_x anno_markers;
run;
```

GRAPH TITLE

A sighted user can perceive, among other general facts, the layout of a plot in under a second. In order to provide this “glance exposure” to the visually impaired, a summary of the layout is provided under the title of the graph. The following code provides the `DATA` steps to produce these mark-ups for the `annotate` facility to process.

```
%macro annotitle;
  xsys='1'; ysys='3'; hsys='3'; when='a';
  function='label'; position='5'; size=5; x=50; y=98;
  color='black'; text="&title"; output;
  color='white';
```

```

x=0; y=94; function='move'; output;
x=100; y=100; function='box';
html="title='&title.. &descr, inactive' tabIndex=1 id='pagetitle' href";
output;
%mend;

```

The macro variable title is passed in by the developer, and the macro variable descr is based on best practices for the chart type and view being built.

Below is part of the chart that is produced by the %annotitle macro. In this particular case, the macro variable title, passed in by the developer, was “Does Literacy Affect National Income?” and the macro variable descr, which is generated, is “This is a standard view of a scatter plot that contains 125 points. Click on links below scatter plot to choose other view.”

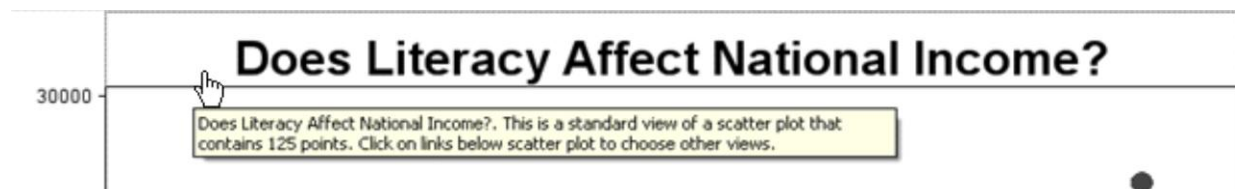


Figure 2: Example of Marked-Up Graph Title

PLOT AXES

In order to provide context for the content areas, we needed a way to mark up the axes. Taking inspiration from highways, where rumble strips alert drivers audibly to the edge of the road, we have marked up the axes with "rumble strips" to alert the user as to the edge of the content area. Furthermore, the rumble strips announce which axis is being touched, the value at that point on the axis, the units of measurement, and the axis label.

The following code produces the annotate data set to create the entire x-axis annotation.

```

%macro annoX(color, x_low, x_up, x_incr, x_format, x_denom, x_label, depth);
  xsys='2'; ysys='1'; hsys='3'; when='b'; size=2;
  color="color";
  do x_counter=x_low to %sysevalf(&x_up-&x_incr) by &x_incr;
    html="title='x = "||trim(left(put(x_counter, &x_format)))||
      "&x_denom &x_label, inactive' tabIndex=0 href";
    y=0; x=x_counter; function='move'; output;
    y=y-&depth; x=x_counter+&x_incr; function='box'; line=0; output;
  end;
%mend;

```

The macro variables x_low, x_incr, x_up, x_denom, and x_label are passed in by the developer when he calls the macro scatterPlot. When %scatterPlot calls %annoX, color will be “white,” and depth will be 10.

In short, this macro starts from the lower bound of the x-axis and works its way rightward by the assigned increment, creating a strip of invisible boxes, each marked up with title text.



Figure 3: Example of Marked-Up X-Axis

ACCOUNTING FOR ERRORS

The annotations generated for the axes are technically built in a negative field. The annotate facility will return an error, but process the data set until the maximum number of errors are encountered. The following data set will set

that maximum number of errors to a sufficiently large number so as to keep the annotate facility running.

```
%macro ignore;
  data ignore_error;
    function='seterror'; size=5000; output;
    function=''; size=.; text=''; output;
  run;
%mend;
```

SCATTER PLOTS

When interacting with a scatter plot, users might need to understand data distribution, correlation, density, and outliers. Since different tasks require different visual approaches, said approaches were simulated with five different views for visually impaired users. The scatter plot macro generates a separate HTML file for each of the five views.

- The standard view annotates each data point represented in the scatter plot with a label that includes the x and y value of that data point. When the user touches a data point, Voiceover vocalizes the label.
- The 3 by 3 Grid View divides the Cartesian plane into a 3 by 3 grid. Each sector of the grid is annotated with a label that specifies the number of data points within that sector. When a user touches the sector, Voiceover reads the label. This interaction allows a blind user to explore the Cartesian plane with their finger and quickly build a mental model of the spatial distribution of the data.
- The 9 by 9 Grid View divides the Cartesian plane into a 9 by 9 grid. Each sector of the grid is annotated with a label that specifies the number of data points within that sector. The user interaction with this view is similar to the 3 by 3 Grid View, but it allows the user to build a more detailed mental model of the spatial distribution of the data.
- The Horizontal Slice View divides the Cartesian plane into nine horizontal slices of exactly the same size. Each slice is annotated with a label that specifies the number of data points within that slice. This view allows the user to quickly find outliers on the y-axis.
- The Vertical Slice View divides the Cartesian plane into nine vertical slices of exactly the same size. Each slice is annotated with a label that specifies the number of data points within that slice. This view allows the user to quickly find outliers on the x-axis.

STANDARD VIEW

In order to produce the general annotations, the description must first be generated. For the standard view, and only for the standard view, the numbers of points that will be plotted are part of the description. The following code counts the number of points that will be plotted.

```
PROC SQL noprint;
  SELECT count(&y_variable)
    INTO :numvars
    FROM &dataset
    WHERE &x_variable ne . and &y_variable ne .;

%let descr=This is a standard view of a scatter plot that contains %trim(&numvars)
points. Click on links below scatter plot to choose other views.;
```

See Figure 2 for the description that the above produces.

Once the general annotations have been generated, the specific annotations for the plot are next. For the standard view, each individual data point is marked-up with three pieces of information: the data point's label, x-value, and y-value. The following produces the annotate data set with these mark-ups.

```
data anno_markers;
  set &dataset;
  x=&x_variable; y=&y_variable;
  xsys='2'; ysys='2'; when='b';
  function='pie'; style='pempty'; color="white";
  rotate=360; width=0.01; size=&radius;
```

```

html="href alt='||trim(left(&nameloc))|| ", " ||
trim(left(put(&x_varibale, &x_format))) || " &x_denom &x_label, "||
trim(left(put(&y_variable, &y_format))) || "y_denom &y_label"';
run;

```

The macro variables `nameloc`, `dataset`, `radius`, `x_variable`, `x_denom`, `x_label`, `y_variable`, `y_denom`, and `y_label` are passed in by the developer when he calls `%scatterPlot`. See Figure 4 for an example of the mark up.

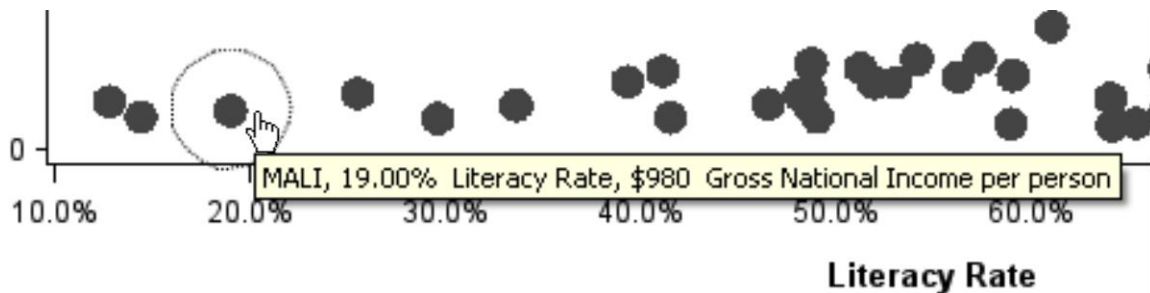


Figure 4: Example of Marked-Up Data Point

GRID VIEWS

Producing the description for the grid views requires no calculations. The description is simply a summary of the layout and a statement of how the view can be navigated. The following builds the 3 by 3 Grid View.

```

%let descr=This is a view of a scatter plot that divides the Cartesian plane into a
3 by 3 grid. Click on a sector of the grid to switch to the 9 by 9 Grid View.
Click on links below the plot to choose other views.

```

The following builds the 9 by 9 Grid View.

```

%let descr=This is a view of a scatter plot that divides the Cartesian plane into a
9 by 9 grid. Click on a sector of the grid to switch to the Standard View.
Click on links below the plot to choose other views.

```

The calculations aren't required until the plot area annotations, where the number of data points in each section needs to be counted. Each count will be stored in array, which will need to be globalized, done here:

```

%do i=1 %to 81;
  %global var&i;
%end;

```

Once the array is initialized, the calculations can begin. First, the bounds of each section are calculated. The following is calculated for the 3 by 3 Grid View.

```

/* span y = the range of the y values, divided by 3 */
%let spany=%sysevalf(%sysevalf(&y_up-&y_low)/3);

/* span x = the range of the x values, divided by 3 */
%let spanx=%sysevalf(%sysevalf(&x_up-&x_low)/3);

%do i=0 %to 2;
  %do j=0 %to 2;
    %let leftSide=%sysevalf(&x_low+%sysevalf(&spanx*&j));
    %let rightSide=%sysevalf(&leftSide+%spanx);
    %let bottom=%sysevalf(&y_low+%sysevalf(&spany*&i));
    %let top=%sysevalf(&bottom+%spany);
    %let num=%sysevalf(%sysevalf(3*&i)+&j+1); /* Section # */

    %count(&num, &bottom, &top, &leftSide, &rightSide);
  %end;
%end;

```


Secondly, the points in each section are counted and stored in an array.

```
%macro count(num, bottom, top, leftSide, rightSide);
  PROC SQL noprint;
    SELECT count(&y_variable);
    INTO :var&num
    FROM &dataset
    WHERE &y_variable>=&bottom
      AND &y_variable<=&top
      AND &x_variable>=&leftSide
      AND &x_variable<=&rightSide;
%mend;
```

Again, the macro variables data set, y_variable, and x_variable have been passed in by the developer. Now that the array has been populated with values, the annotate data set can be built for the plot area.

```
%let sector1=south western; %let sector2=southern; %let sector3=south eastern; %let
sector4=western; %let sector5=central; %let sector6=eastern; %let sector7=north
western; %let sector8=northern; %let sector9=north eastern;

data anno_markers;
  xsys='1'; ysys='1'; hsys='3'; when='b';
  style='empty'; color='white';
  %do i=0 %to 2;
    %do j=0 %to 2;
      %let num=%sysevalf(%sysevalf(3*&i)+&j+1);
      function='move';
      x=%sysevalf(&j*33); y=%sysevalf(&i*33); output;
      html="title='||trim(left(&&var&num))|| ' data points in &&sector&num
        sector, click here to switch to 9 by 9 view' href='fine_view.htm'";
      function='bar';
      x=%sysevalf(%sysevalf(&j+1)*33);
      y=%sysevalf(%sysevalf(&i+1)*33); output;
    %end;
  %end;
run;
```

See Figure 5 for an example of the mark up the above would create.

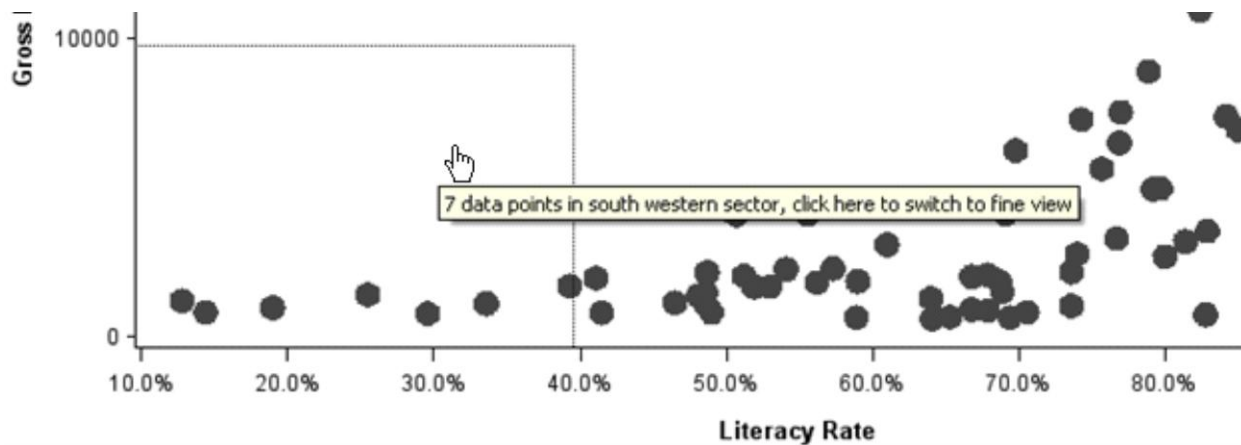


Figure 5: Example of a 3 by 3 Grid View Mark-Up

SLICE VIEW

Like the grid views, producing the description for the slice views requires no calculations; it is simply a statement.

```
%let descr=This is a view of a scatter plot that divides the Cartesian plane into a
3 by 3 grid. Click on a sector of the grid to switch to the 9 by 9 Grid View.
Click on links below the plot to choose other views.
```

As with the grid view, the bounds of each slice must be calculated. The following code calculates the dimensions of a horizontal slice.

```
/* span = the range of the y values, divided by 9 */
%let span=%sysevalf(%sysevalf(&y_up-&y_low)/9);
%do i=1 %to 9;
  %let min=%sysevalf(&y_low+%sysevalf(&span*&sysevalf(&i-1)));
  %if &i=9; %then
    %let max = &y_up;
  %else
    %let max = %sysevalf(&min+%span);
  %count(&i, &min, &max, &x_low, &x_up);
%end;
```

The macro count is the same as the one that was used for the grid view. With the freshly populated array, the annotate data set for the plot area can be produced.

```
data anno_markers;
  xsys='1'; ysys='1'; hsys='3'; when='b';
  style='empty'; color='white';
  %do i=1 %to 9;
    function='move'; x=0; y=%sysevalf(%sysevalf(&i-1)*11); output;
    html="title=''||trim(left(&&var&num))|| " data points in slice &i, click
      here to switch to 9 by 9 view' href='fine_view.htm'";
    function='bar'; x=100; y=%sysevalf(&i*11); output;
  %end;
run;
%let descr=This is a view of a scatter plot that divides the Cartesian plane into a
3 by 3 grid. Click on a sector of the grid to switch to the 9 by 9 Grid View.
Click on links below the plot to choose other views.
```

See Figure 6 for an example of the mark-up the above would create.

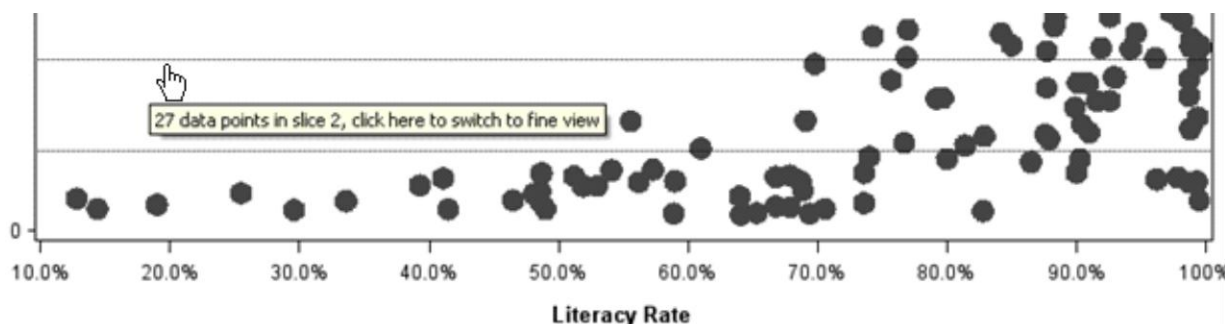


Figure 6: An Example of a Horizontal Slice View Mark-Up

OUTPUT

There are a couple of best practices to be aware of in the construction of the output.

1. The title tag of the output page should have the same title as the graph. This is accomplished with (title="&title"), as shown below.
2. The viewport should be set so that the image takes up as much space as it reasonably can on the iOS device. To accomplish this, the following meta tag should be part of the produced page:

```
<meta name="viewport" content="width=device-width">
```

 This is printed with the ODS statement metatext as shown below.

```
/* The value of &page will be one of the following: standard, horizontal, vertical,
   coarse (3 by 3 Grid View), or fine (9 by 9 Grid View). */

ODS HTML file="&page._view.htm" (title="&title")
  metatext="name='viewport' content='width=device-width'";
symbol1 value=dot height=&dot_size.pct color=gray44;
axis1 label=(angle=90 "&y_label") order=(&y_low to &y_up by &y_incr) minor=none;
axis2 label=("&x_label") order=(&x_low to &x_up by &x_incr) minor=none;

PROC GPLOT data=&dataset anno=anno;
  Plot &y_variable*&x_variable / vaxis=axis1
                                haxis=axis2;
```

CONCLUSION

Currently, blind students and professionals do not have equal access to graphical representations of data. This paper defines how to use SAS/GRAPH to create data visualizations that are optimized for access by those users on Apple iOS devices using the Voiceover screen reader. The paper defines preliminary best practices for scatter plots. The paper describes a SAS macro that embodies the best practices.

REFERENCES

- Rayner, Keith. 1998. "Eye Movements in Reading and Information Processing: 20 Years of Research." *Psychological Bulletin*. 372-422. Washington, DC: American Psychological Association
- Manyika, James, et al. 2011. "Big data: The next frontier for innovation, competition, and productivity." McKinsey Global Institute.
<http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation>
- "Web Content Accessibility Guidelines (WCAG) 2.0." <<http://www.w3.org/TR/WCAG20/#text-equiv>>
- Rothberg, M. and B. Gould. 2010. "Effective Practices for Description of Science Content within Digital Talking Books." *Proceedings of The 25th Annual International Technology & Persons with Disabilities Conference*. San Diego, CA: CSUN.
- Van Gerven, C. 2011. "Making Tactile Graphics." *Proceedings of The 26th Annual International Technology & Persons with Disabilities Conference*. San Diego, CA: CSUN.
- Gardner, J. A. and V. Bulatov. 2008. "Making Scientific Graphics Accessible with ViewPlus IVEO." *Proceedings of The 23rd Annual International Technology & Persons with Disabilities Conference*. Los Angeles, CA: CSUN.
- Miele, J. and S. Landau. 2011. "Audio-Tactile Interactive Computing with the Livescribe Pulse Pen." *Proceedings of The 26th Annual International Technology & Persons with Disabilities Conference*. San Diego, CA: CSUN

CONTACT INFORMATION

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

Ed Summers
SAS Campus Drive
SAS Institute Inc.
ed.summers@sas.com
<http://www.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.