



Getting Started with D.C. Graph (alpha 2) Publication Graphics Software

<http://www.science.gmu.edu/csi779/drope/dcgraph.html>

By: Dan Rope
Rope_D@bls.gov

Poster Session for 1995 Joint Statistical Meetings
August 14, 1995

	<p><i>D.C. Graph</i> Alpha 2</p>	
<p>All Splus Row Label Plotting Code By: Prof. D.B. Carr, George Mason University (Splus Source Code Available via Anonymous FTP @ galaxy.gmu.edu)</p> <p>Graphical Front End By: Dan Rope, U.S. Bureau of Labor Statistics</p>		

Background

Mainstream statistical graphics packages often provide a convenient easy-to-use interface for creating traditional graphs but may have limited ability to customize graphs for special purposes. As computer graphic capabilities increase, statisticians are more able to address the concepts of data visualization and apply both statistical and cognitive science to developing quality graphics. As a result, ad hoc computer code is provided for other scientists to experiment with or produce these newer graphics. D.C. Graph uses functions developed in S-PLUS™ to produce row-labeled plots (“rowplot”) as described in "Converting Tables to Plots" by Dr. D.B. Carr (Technical Report No. 101 from the Center for Computational Statistics at George Mason University). Row-labeled plots provide a simple, effective way of looking at complex tables of data. A sophisticated graphical technique requires much flexibility and control over the resulting graph, which is provided in “rowplot”. D.C. Graph, created in Microsoft’s Visual Basic, is designed to be an intuitive graphical user interface that can create and run “rowplot” S-PLUS code in the Microsoft Windows environment. Emphasis was placed upon retaining power and flexibility while making the interface simple and convenient. A user is not required to write any S-PLUS code. Tools were included to perform common actions, and the design of D.C. Graph allows for implementation of future capabilities of “rowplot”. The software is currently in the Alpha stage which means that features may still be added and bugs still need to be found.

Disclaimer

This software is provided by the author free of charge for use at your own risk. It is NOT supported or distributed by the Bureau of Labor Statistics. The Bureau of Labor Statistics cannot be held responsible for any problems that may occur while using the software.

Hardware/Software Requirements

- At least 386 PC with a math co-processor
- At least 8MB RAM
- At least DOS 5.0 and Windows 3.1
- Splus 3.2 for Windows

Installation Instructions

To install D.C. Graph download and run the self extracting archive (dcg_a2.exe) available at <http://www.science.gmu.edu/csi779/drope/dcgraph.html>. Once the archive has extracted itself, you should run the setup.exe program. The D.C. Graph Setup program will install D.C. Graph in a directory of your choice (default of c:\dcgraph), and create a Program Manager group and item.

Reading Data

1. ASCII Data (Click: File/Read ASCII Data...)

At this point, the program is a little particular about what kind of data it will read. Basically, it should correctly read a dataset that is shaped like a table (a constant number of rows and columns) and:

- the first column contains the row labels--names for each row surrounded by double quotes (").
- the numeric data (NA for missing) are in columns 2-n
- all data are delimited by a consistent delimiter (space[s], comma, tab, etc.)
- the first row optionally can contain the panel titles--names for each column surrounded by double quotes.

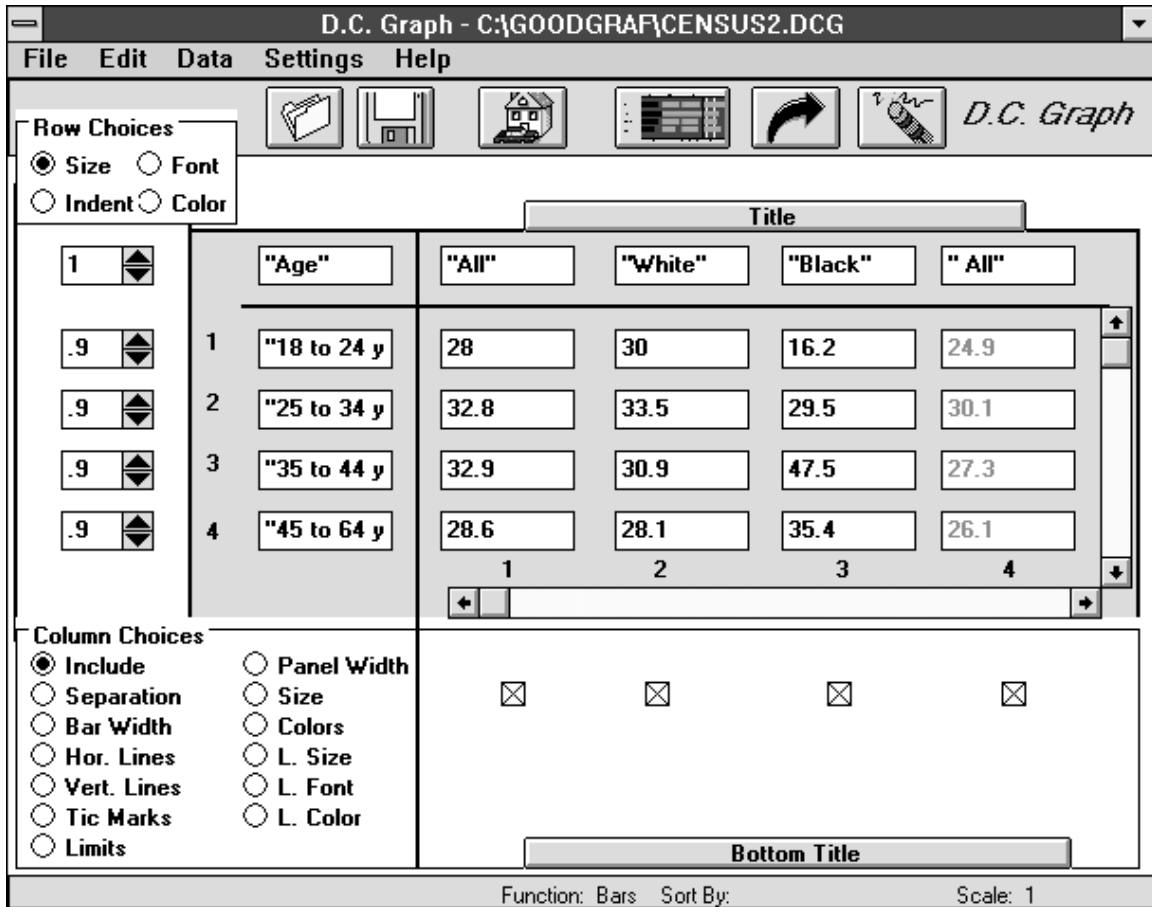
Most spreadsheet packages should have no problem writing out this kind of file. You will have a chance to tell D.C. Graph how many rows and columns are in your data set and which kind delimiter you are using (if D.C. Graph does not guess correctly).

Sample ASCII Dataset:

```
" " "Private Industry" "State and Local Government"
"Wages and Salaries"      71.8 69.8
"Legally required benefits" 9.1 6.0
"Paid leave"              6.8 7.7
"Insurance"                6.9 7.8
"Retirement and savings"  2.9 7.8
"Supplemental pay"        2.4 .9
```

2. Splus Data Frames (Click: File/Open/Splus Data Frame...)

Because D.C. Graph directly uses Splus, it made sense to enable it to directly read Splus data frames. The object **MUST** be a data frame and must not contain characters (NAs are OK). The row labels and panel titles will be read from the `dimnames()` of the data frame. Just type the name of your Splus data frame into the dialog box and click OK.



Creating Graphics

The interface design is similar to a spreadsheet and displays the data that will be graphed. Next to each row and below each column is an input box that controls the appearance of that particular row or column (panel). The input boxes change depending upon which column or row choice is selected. Explanations for the choices appear at the bottom left of the screen when the mouse is moved over a column or row choice. Most modifications to the appearance of the graph are done by clicking on a column choice or a row choice and modifying the input boxes.

To create a graphic, simply click the "Generate" button on the toolbar. The graphic should appear in an Splus win.graph() window by default (see: **Menu Choices:** File/Send Output To.. for information on how to change this). Three different kinds of row-labeled plots are supported in the Alpha 2 version: bar plots, dot plots and arrow plots. The default is a bar plot which can be changed to a dot or arrow plot by clicking Settings/Function. An "arrow" plot displays the change in the values from one column to another. See "Converting Tables to Plots" in Technical Report No. 101 from the Center for Computational Statistics at George Mason University for more information about arrow plots.

The basic idea is to create a graph, examine it, go back to D.C. Graph, make improvements, create a new graph and repeat until you are satisfied with your graph. The quality of the resulting graph is primarily dependent upon the user's knowledge of the principles of data visualization and the display of data.

Colors

This is something that still needs work. For now, when you start D.C. Graph a new color scheme (Rowplot Greyscale) will be added to your Splus for Windows color schemes. The program is designed to work with the "Windows Standard" color scheme and "Rowplot Greyscale" color scheme. So far, each offers only 4 colors. The "Rowplot Greyscale" color scheme provides the colors shown in most of the graphs in Technical Report No. 101.

Fonts

Postscript fonts were taken from the ps.fonts object in Splus. These are available when sending output to a postscript file (see: **Menu Choices:** File/Send Output To..). All other output devices' font options are read directly from splus.ini, so you can add fonts using Splus's Options/Fonts.. dialog box and they will immediately be available in D.C. Graph.

The Toolbar:

(in order of appearance from left to right)



Open:

Click the Open button to retrieve previously saved D.C. Graph (.dcg) files from disk.



Save:

Click the Save button to save your work to disk. You will be prompted if you have not specified a filename; otherwise, the current file will be overwritten without prompting.



Home:

Click the Home button to return to the spreadsheet to the top, left corner of the data (row 1, column 1).



Generate:

Click the Generate button to create the graphic.



Repeat:

Click the Repeat button to avoid typing the same values into each column or row box. If a "column choice" is selected it will copy the value from the leftmost column choice box to all of the columns to the right (leaving the columns to the left alone). It works similarly with the "row choices" by copying the value from the topmost box to all boxes beneath, and has no effect on "arrow choices".



Defaults:

Click the Defaults button to return the selected row or column choice to the default value. The values for all of the columns (or rows) will be reset.

Menu Choices

File

Open

D.C. Graph File: Same as the Open button

Splus Data Frame...: See **Reading Data**, 2. Splus Data Frames

Read ASCII Data...: Allows you to choose an ASCII file to import. See **Reading Data**, 1. ASCII Data

Save: Same as the Save button

Save As...: Allows you to specify new filename for current file

Send Output To...: Allows you to chose where the graphics will be sent.

To Printer: Graphics will be sent to the printer using the Windows Printer Driver specified in Splus's File/Print Setup dialog box.

To PS file: Graphics will be sent to a postscript file specified by clicking Options... and filling in the name of the postscript file.

To win.graph(): Graphics will be sent to the Splus Graphics "win.graph()" window.

To Clipboard: Graphics will be "copied" into the clipboard and displayed in a dialog box when graph is complete.

Generate, View Code...: Generates all Splus code to create the current graphics and writes it to the specified file (which, due to a bug, must already exist).

Exit: Prompts to save changes and exits D.C. Graph (not Splus for Windows)

Edit, Cut, Copy, Paste: The usual cutting, copying, pasting in Windows.

Data

Sort By...: Sorts the data by the selected panel title *when the graph is generated*

Scale By...: Scales the data by the inputted value *when the graph is generated*.

Transpose: Transposes the rows and columns in the current file. Note: all column and row settings are lost after this is done. You should save your work before doing this.

Settings:

Rows...: Provides options for row settings that are global to the graph as opposed to specific for any particular row. Includes options are: width of area for row labels, number of rows in a group, spacing above first line and below last line, spacing for the row title line, right justified row labels.

Margins...: Provides dialog box for adjusting the margins of the graphic. The outer dark box represents the paper, the red cross-hashed box is the plotting region, and the dashed box is the outer margin (for titles).

Function: Changes the type of plot to a bar plot, dot plot, or arrow plot. See: **Creating Graphics**

Panel Titles...: Provides options for panel (column) titles that are global as opposed to specific for any particular column. Included options are: size, font, color, and spacing.

X-Axis Positions...: Provides options for coordinates for the X-axes. The three numbers directly correspond to the Splus “mgp” settings. The “Alternate” check box automatically enters numbers to make the x-axes alternate positions between panels for easier reading. *Note: This dialog box needs work!*

In Color: Toggles color schemes between “Windows Standard” color scheme (black, red, green, blue) and the “Rowplot Greyscale” color scheme. See: **Colors**

Help: Not available yet!

Titles

Proper and informative titles are essential companions to quality graphics. D.C. Graph provides space for 3 (or more--see: **Tips and Tricks**) title lines and 1 (or more) bottom title lines. The bottom title line often indicates units if the data have been scaled, etc. To create a title, click on the Title (or Bottom Title) button. The dialog box provides options for the text of the title, size, font, spacing, indentation, and an option to choose the panel to center the title over.

Tips and Tricks

Using “\n”:

Splush uses ‘\n’ to indicate a line break. Therefore, any text that contains a ‘\n’ will break and go to the next line. This is useful for splitting a title line into two lines or a panel title that is unusually long.

Splitting Long Row Labels:

For row labels, the “\n” trick may not work so well if the next line overwrites the next row label. In these cases, it is usually better to split the troubled row into two rows. One with half of the text and the data, and the other with the other half of the text and NAs or zeros for the data. Unfortunately, currently the only way to do this is to modify the original ASCII file and re-read it which is not very satisfying since all D.C. Graph settings done up to this point would be lost.

Bottom Title Margin Problem

Most often the bottom title will appear in the middle of the X-axis labels by default. To get around this, first put in some “Top Space” in the bottom title dialog box (perhaps about 3 or so). Then, you may want to modify the margins (See: **Menu Choices**, Settings/Margins...) by separating the bottom margin from the graphics area. *Note: Once this is done, you will probably see an Splush “Graphics Warning” the next time you generate your graph. This warning message usually has no effect on the graphic.*

Creating Graphs on Different Platforms (Unix, etc.)

This is not the most convenient use of D.C. Graph, but it is possible to generate the code, move it up to another platform and run it assuming Splush is installed on that platform. File/Generate, View Code... creates a file that contains all of the code necessary to create the graph. The code assumes that the "rowplot", "bars", "arrow", "dot", and "title.text" Splush functions are installed and available in Splush's search() path on the platform that will be used. These functions are available via anonymous ftp at galaxy.gmu.edu under /pub/submissions/eda/rowplot. Read the first few lines of the generated code to determine if anything needs to be commented out or changed [i.e. win.graph() does not work on Unix].

Pasting Graphics to other Applications

If File/Send Output To... is set to "Clipboard", then the graphics will automatically be copied into the Windows clipboard and displayed in a window when the graph is complete. Clicking OK makes the window disappear, but the graph remains in the clipboard. To insert the graphic within another application, simply start the application and paste it (usually done by clicking Edit/Paste). Note that the object (i.e. the graph) that is pasted is a Window's metafile, so it should be fairly easy to scale or edit.