

Smart Statistical Graphics – A Comparison Between SAS® and TIBCO Spotfire® In Data Visualization

Yi Gu, Roche Product Development in Asia Pacific, Shanghai, China

ABSTRACT

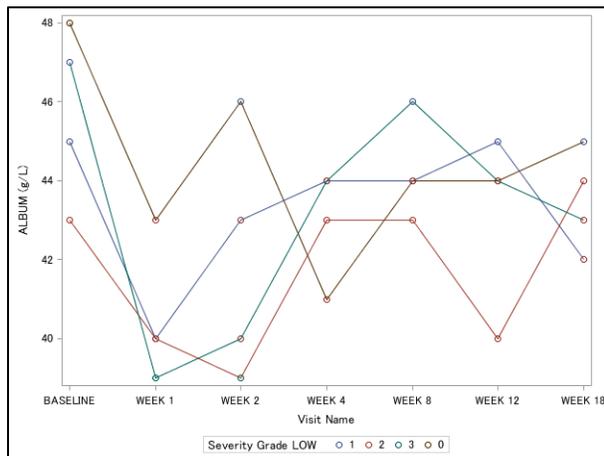
Known for the best analytic software, SAS has always been challenged by emerging analytic tools. TIBCO Spotfire 6.5, an analytics and business intelligence platform, which enables data visualization in an interactive mode, has been brought into the competition in pharmaceutical industry these years. With its increasing implementation in the field of safety monitoring and dose escalation, TIBCO Spotfire presents its superiority for exploratory analysis. On the other hand, the capability allowed in the efficiency, strong statistical analysis and data processing defends SAS software. This article is to demonstrate the two softwares in terms of data visualization and explore the possible integration between them.

INTRODUCTION

With the advent of new features in SAS 9.4, creating by-patient plot becomes more efficient. Procedures including SGPLOT and SGPANEL, enable plotting and combining plenty of information in one set of plot. This article focused on the commonly requested clinical graph, by-patient laboratory test plot, to discuss how SAS 9.4 and TIBCO Spotfire visualize laboratory results and introduce the interactive interface Spotfire has to supplement the basic visualization on the basis of SAS outputs.

1 – SCATTER PLOT COLORED BY GROUP

In Figure 1, the SGPLOT combined scatter plot and series plot together, using GROUP to visualize different subjects and grades. Each line represents a subject and displays the severity grades in the form of cycles in various colors. As is shown, the two green cycles at the very bottom has a low value of the lab test with a corresponding severity grade of three.



```
proc sgplot data=test_1;
  scatter x=visit y=aval/group=gradel;
  series x=visit y=aval/group=subjid ;
  yaxis label='ALBUMIN (g/L)';
run;
```

Figure 1. By Patient Plot of Laboratory Test Results, colored by Severity Grades

Similarly, TIBCO Spotfire allows the scatter plot display both the values of laboratory results and the corresponding severity grades. It enables coloring scatter marks by changing COLORS in PROPERTIES of a scatter plot. Figure 2 shows the window of the PROPERTIES of a scatter plot, and how COLORS can be adjusted.

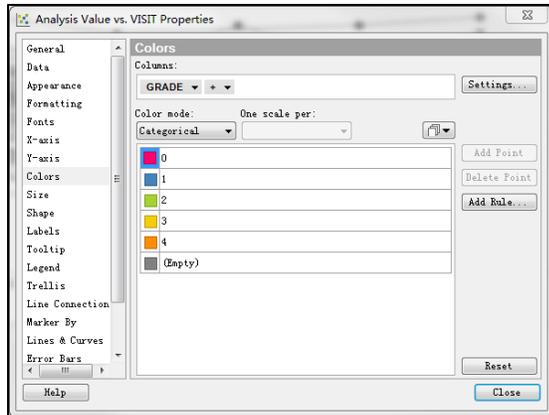


Figure 2. Properties Window of the scatter plot in Spotfire

In some cases, the classification levels have not yet been identified, especially in exploratory analyses. For example, the classification of the percentage of change from baseline, which is important for clinical trial results interpretation.

By utilizing TIBCO Spotfire, users are able to input of the percentage of interest for classification and visualization.

Step 1: Calculate *Percentage of change from baseline* [PCHG] in SAS 9.4:

```
data test;
  PCHG = abs ( (AVAL-BASELINE) /BASELINE ) ;
run;
```

Step 2: Load data into Spotfire and insert Scatter Plot as the visualization

Step 3: Create a text box for inputting user-defined value of abnormal percentage of change from baseline
 Insert Property Control >> Control type: Input field >> Create a new property *\$PctSelect*

Step 4: Add a calculated column named as "ABNORMALITY"

```
case
  when [Percentage Change from Baseline] > ${PctSelect} then "Abnormal"
  when [Percentage Change from Baseline] <= ${PctSelect} then "Normal"
end
```

Step 5: Edit COLORS in the scatter plot PROPERTIES and choose the new added column ABNORMALITY for coloring

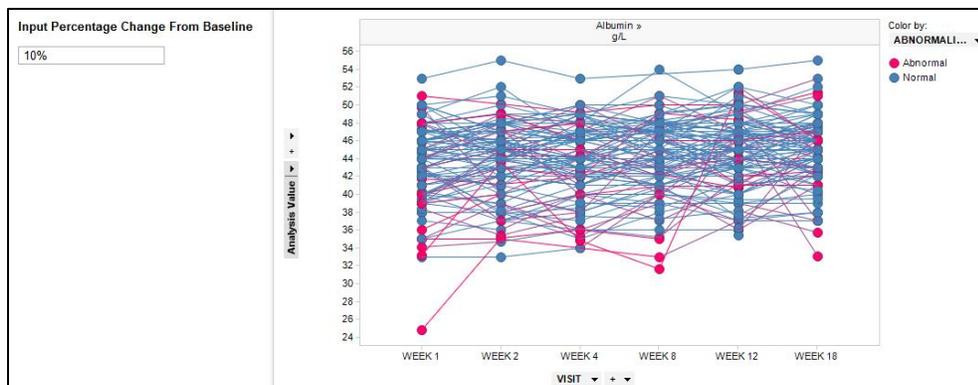
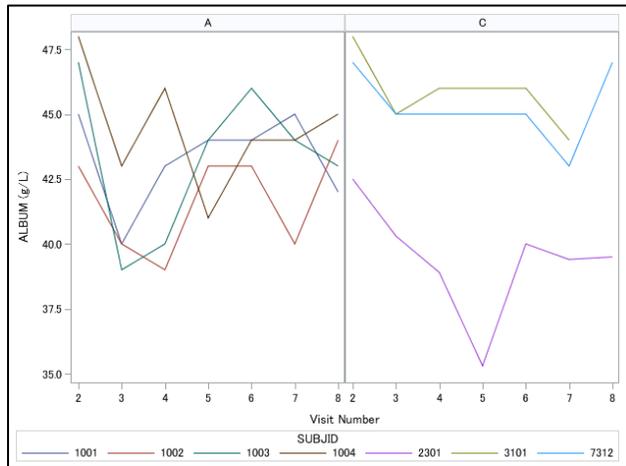


Figure 3. Input values to classify percentage of change from baseline and display in the by-patient plot using TIBCO Spotfire

As is shown in Figure 3, the user-defined value for differentiate is 10%, which means in the plot, Abnormal marks in red are those test results with an increase or decrease change from baseline more than 10%.

2 – PANEL PLOTS

Panel plots are designed to produce graphs based on classification levels, which effectively display the data among the categorized groups. The by-patient plot in Figure 4 is classified by treatment groups, showing the changes in lab test results over time among different treatment groups.



```
proc sgpanel data=test_2;
  panelby armcd/novaname;
  series x=visitnum
  y=aval/group=subjid ;
  rowaxis label='ALBUM (g/L)';
run;
```

Figure 4. By Patient Plot of Laboratory Results paneled by Treatment Groups

In TIBCO Spotfire, the classification can be altered based on user's request by choosing variables from an inserted drop down list. As is show in Figure 5, SEX was chosen as the variable for panel. Following are the steps for the visualization:

Insert Property Control >> Control type: Drop Down List >> Create a new property *\$ChoosePanel* >> Set property value through Column Selection of data table.

Edit TRELIS in the scatter plot PROPERTIES to have panels split by *\$ChoosePanel*

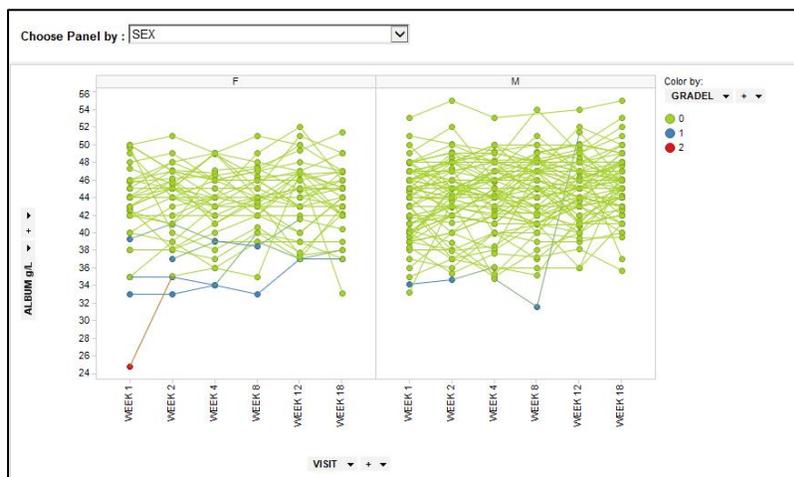


Figure 5. Choose Panel by variables for the by patient plot in TIBCO Spotfire

3 – INDIVIDUAL OR MULTIPLE SUBJECTS OF INTEREST

A general by-patient plot provides an overall picture of how laboratory test results distribute and makes outlier visualized. For details by individual, SAS Macro can be utilized to select unique subjects of interest and draw the corresponding by-patient plot.

```
%macro sbj(subjid);
```

```

data test_2;
  set test;
  if SUBJID in ("&subjid");
run;

proc sgpanel data=test_2;
  panelby armcd/novarname;
  series x=visitnum y=aval/group=subjid ;
  rowaxis label='ALBUM (g/L)';
run;
%mend sbj;

```

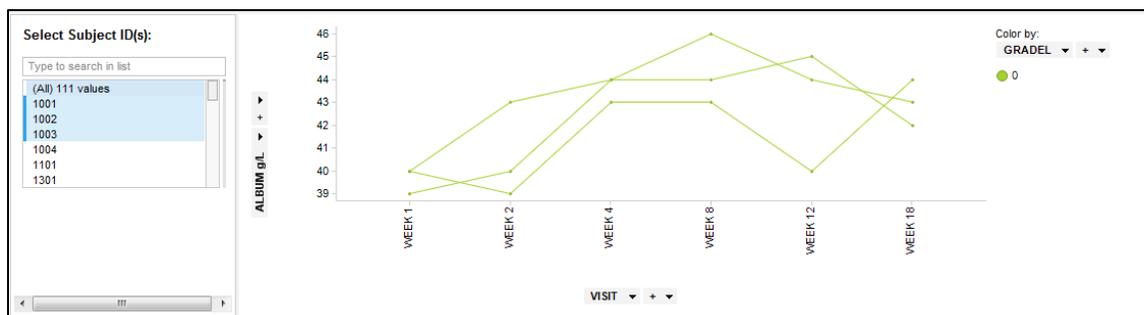


Figure 6. By-patient lab test plot with selected subjects

TIBCO Spotfire utilizes an embedded filter to visualize lines for unique or multiple subjects of interest. Moreover, further information, e.g. adverse events, can be tracked by marking. The red dot at the bottom of the plot is highlighted in Figure 7 and correspondingly, the adverse events information for this subject was listed in the table above. With the linkage of the two data sets, user can easily seek deeper and detailed clinical information of the subjects of interest.



Figure 7. Marked scatter point in by-patient lab test plot with related Adverse Events.

CONCLUSION

In terms of reporting capability of graphics, both SAS 9.4 and TIBCO Spotfire present a high level capability of data visualization. And regarding to the users' experience, TIBCO Spotfire provides an interactive platform for exploratory analysis. With its simplicity to adjust axes symbols and text, and its ability of user-input query to export visualization, TIBCO Spotfire enables faster data review, quality assessment and process improvement. As is called "the de factor end user interface of choice for the strategic assessment of clinical data", TIBCO Spotfire saves time consumed through a traditional ad-hoc process of creating statistical graphics. While still following a standard process including on-demand development, quality review and final production, TIBCO Spotfire leaves the feasibility for customers' modification.

On the other hand, SAS offers unparalleled analytic power and data processing capability. In the examples of this article, data were all pre-processed via SAS 9.4 and then imported to TIBCO Spotfire for plotting. Deployment and integration would definitely leverage the efficiency and effectiveness of the visualization.

REFERENCES

SAS® 9.4 Graph Template Language: Reference, Fourth Edition. Available at:

<https://support.sas.com/documentation/cdl/en/grstatgraph/67882/HTML/default/viewer.htm#n1dlakx61v72in1k3ebm8rz18qd.htm>

Chen, Mina (2016). *Meaningful Presentation of Clinical Trial Data with Multiple Y-Axes Graph*.

Sanjay Matange (2016). *Clinical Graphs Using SAS*.

Phillips, M. (2015). *Tibco spotfire : a comprehensive primer*.

TIBCO Spotfire. *Spotfire Analytics – Transforming Clinical Development*. Available at:

<http://spotfire.tibco.com/assets/blt9ebddf0e420042ec/spotfire-clinical-development.pdf>

CONTACT INFORMATION

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

Name: Ivy Gu

Enterprise: Roche (China) Holding Ltd.

Address: 4F, Building 11, 1100 Longdong Avenue, Pudong New District

City, State ZIP: Shanghai, China, 201203

Work Phone: +86-021-2892 3949

E-mail: ivy.gu@roche.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.