



PhUSE US Connect 2019

Paper PP12

Optimizing Over Time plots via Analytics

Kishore Pothuri, Vita Data Sciences, Waltham, MA, USA

ABSTRACT

The typical practice of generating time series and mean measurement over time plots is by using PROC SGPLOT and GPLOT in SAS®. Even though these are extensively used and known procedures, it is challenging and tedious for the end users to review these static outputs which are non-interactive in nature. With the help of analytic tools, SAS® and TIBCO Spotfire®, we can generate highly customizable and web-based plots which allow the end users to interact dynamically with real-time data and to efficiently analyze the results. Also, the corresponding data tables and listings can be generated within Spotfire and can be exported as csv, pdf or ppt for manual review, significantly reducing the review time and to make decisions effortlessly. In this paper, we will describe the methodology for the generation of over time plots in Spotfire using SAS® and also highlight the features of additional customizations to meet stakeholders' requirements.

INTRODUCTION

"A picture is worth a thousand words" is a commonly used phrase in every industry where there is data involved. In the clinical industry, the data is represented and analyzed in the form of tables, listings and figures/graphs (TLF or TLG) which are typically generated using SAS®. The conventional way of generating time series and mean measurement over time plots, is by using PROC SGPLOT and GPLOT in SAS®. Even though these are extensively used and known procedures, it is challenging and tedious for the end users to review these static outputs which are non-interactive in nature, without giving them an ability to explore and drill-down further. By combining the power of SAS® and TIBCO Spotfire®, we can generate highly customizable and web-based plots which allow the user to not only view data interactively but also provide the ability to identify trends, filter data, drill down to details and export the data as csv, pdf or ppt for more details. In this paper, let us discuss how one can make use of SAS® along with data visualization tool such as TIBCO Spotfire® and generate interactive time series plots.

OVERTIME PLOTS

A timeplot (sometimes called a time series graph) displays values against time. They are similar to x-y graphs, but while an x-y graph can plot a variety of "x" variables (for example, height, weight, age), timeplots can only display time on the x-axis. Unlike pie charts and bar charts, these plots do not have categories. Timeplots are good for showing how data changes over time and help in identifying and interpreting relationships in data. With the use of visualization tools, high resolution color graphic plots can be generated which allow the user to not only observe the trends and data patterns, but also analyze the data quickly and concisely.

In Clinical trials, patient data is continuously monitored and assessed at regular intervals for safety and efficacy. For this purpose, time series graphs are generated using SAS alone. However, the outputs are static and non-interactive in nature, without providing users the ability to explore and drill down the data further. With the help of visualization tools such as TIBCO Spotfire® in conjunction with SAS, all these challenges can be ruled out.

In this paper, let's see how one can make use of SAS statistical procedures along with data visualization tool such as TIBCO Spotfire® and generate interactive Overtime and Mean +/- Standard Deviation Plots.

COMPARISON OF OVERTIME PLOTS GENERATED USING SAS & TIBCO SPOTFIRE® IN CONJUNCTION WITH SAS

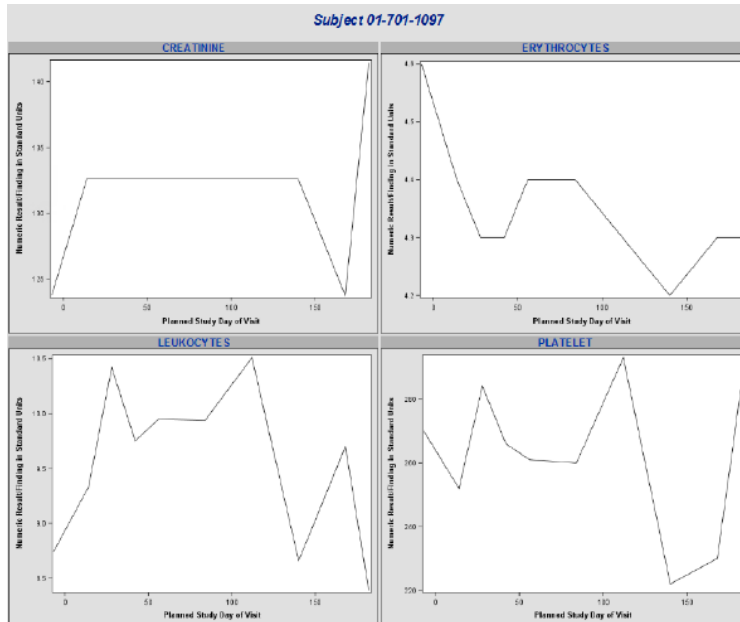


Figure 1: Lab Overtime Plots using PROC SGPLOT in SAS®

Fig. 1 Shows the screenshot of overtime plots for lab panels limited to four parameters (Creatinine, Erythrocytes, Leukocytes Platelet) for a particular subject. This plot is generated by using statistical procedure PROC SGPLOT in SAS, with 'Planned Study Day of Visit' on X-axis and the 'Result' of corresponding lab parameter on Y-axis.

Limitations: The output is static in nature; the SAS code needs to be updated every single time if the user chooses to view different lab parameters of interest for a different subject; data drill down is not possible; Also, the underlying source data cannot be viewed simultaneously.

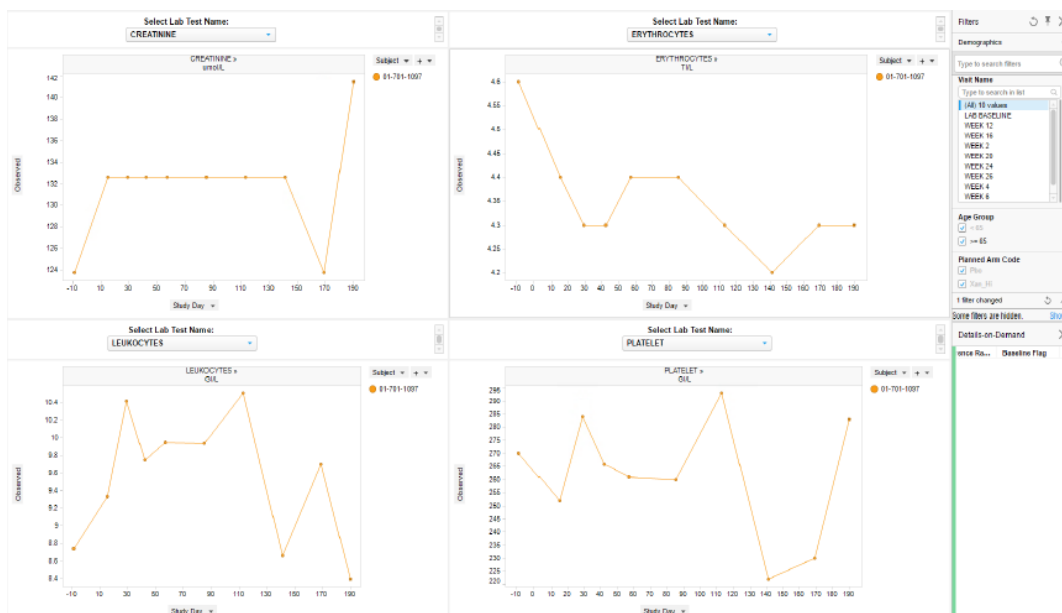


Figure 2: Lab Overtime Plots using TIBCO Spotfire® from our VDS Tag® Data Analytics Instance

Fig. 2 Shows the screenshot of overtime plots for Lab panels limited to four parameters (Creatinine, Erythrocytes, Leukocytes Platelet) for a particular subject. This plot is generated by using TIBCO SPOTFIRE® with 'Planned Study Day of Visit' on X-axis and the 'Result' of corresponding lab parameter on Y-axis.

Advantages: The output is interactive in nature; the user can select the lab parameter of interest by using the 'Select Lab Test Name' drop down, allowing the user to efficiently analyze the results across multiple parameters; the plot is dynamically updated based on the selection of filters available on the right side of the page and also the corresponding underlying source data can be viewed on 'Details On Demand' and the data listing can be exported as csv, pdf or ppt for further data exploration if needed.

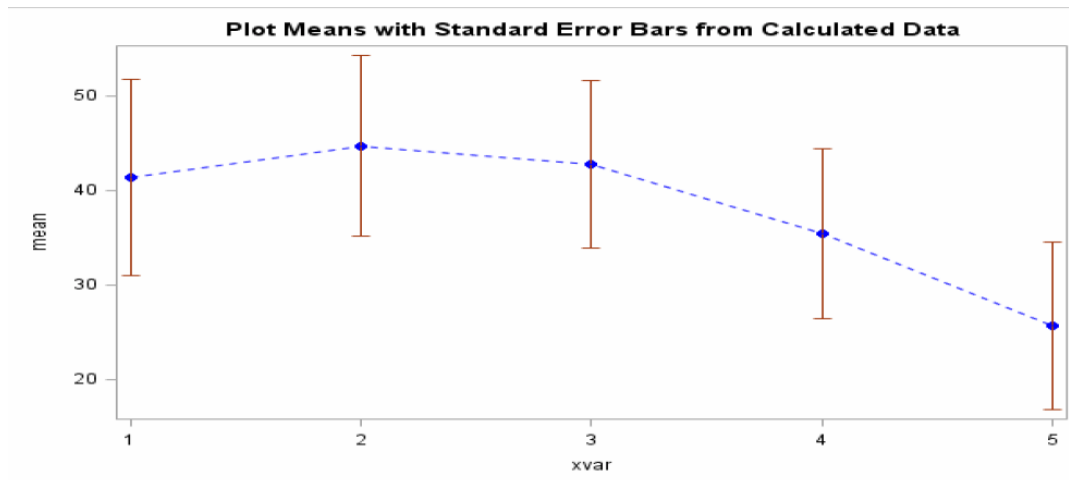


Figure 3: Means Plot with Standard Deviation using PROC GPLOT in SAS®

Fig. 3 Shows the screenshot of a means plot with standard deviation generated by using PROC GPLOT in SAS.

Limitations: The output is static in nature; the SAS code needs to be updated every single time if the user chooses to view the mean values for Change from Baseline/Percent Change from Baseline for parameters of interest; data drill down is not possible; Also, the underlying source data cannot be viewed simultaneously.

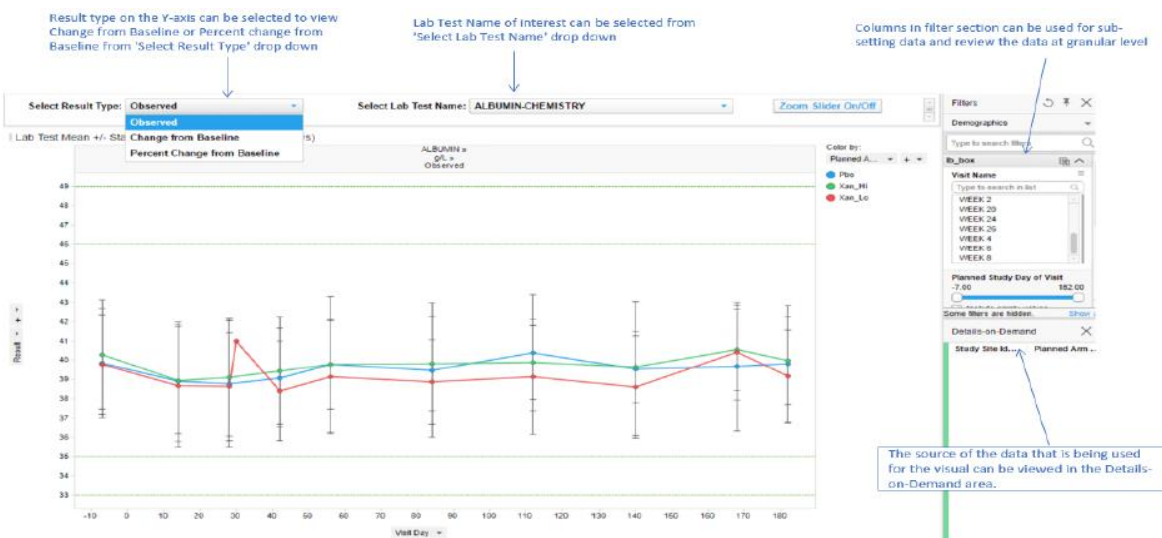


Figure 4a: Screen shot of the Lab Test Mean +/-Standard Deviation plot from our VDS Tag® Data Analytic instance - annotated to show how one can select pre-defined data filters to only display value of interest

Fig 4a. Shows the screen shot of lab test mean $\pm$ standard deviation plot for lab test 'Chemistry' generated by using TIBCO Spotfire® with 'Planned Study Day of Visit' on X-axis and 'Mean of the Results' of corresponding lab parameter on Y-axis. The 'Select Result Type' drop down can be used to select either 'Change from Baseline' or 'Percentage Change from Baseline' apart from the standard result.



Figure 4b: Screen shot of the Lab Test Mean $\pm$ Standard Deviation plot from our VDS Tag® Data Analytic instance- restricted to display Lab Test Mean $\pm$ Standard Deviation Values based on the selection of a data filter Visit Name='WEEK 6'

Fig 4b. Shows the screen shot of lab test mean $\pm$ standard deviation plot for lab test 'Chemistry' generated by using TIBCO Spotfire® with 'Planned Study Day of Visit' on X-axis and 'Mean of the Results' of corresponding lab parameter on Y-axis. In addition, the data filter on right is selected to display the data belonging to only 'WEEK 6'.

Advantages: The outputs are interactive in nature; the user can select the lab parameter of interest by using the 'Select Lab Test Name' and also can change the result type from 'Observed' to either 'Change from Baseline' or 'Percent Change from Baseline' by using the 'Select Result Type' drop down; the plot is dynamically updated based on the selection of filters available on the right side of the page and also the corresponding underlying source data can be viewed on 'Details On Demand' and the data listing can be exported as csv, pdf or ppt for further data exploration as needed.

The source data that is being used for the visual can be viewed in the Details-on-Demand area. Figure 5 shows the Details-on-Demand panel presenting the underlying source data of the visual in Figure 4b. When a particular data point is marked in the visual then all the rows of the data of that part will be presented in the Details-on-Demand panel. This functionality can also be used to drill down into specific details and the data rows can be exported to excel sheet. In this example, a data listing of source data for Visit Name='WEEK 6' is presented which can be exported into csv for internal review and communication.

Details-on-Demand											
Study Site Id...	Planned Arm ...	Country	Age Group	Unique Subje...	Lab Test or E...	Lab Test or E...	Reference Ra...	Category for ...	Reference Ra...	Numeric Res...	Standard Units
701	Pbo	USA	< 65	01-701-1015	ALB	ALBUMIN	3.3	CHEMISTRY	4.9	37.00	g/L
701	Pbo	USA	< 65	01-701-1015	ALP	ALKALINE PH...	35	CHEMISTRY	115	43.00	U/L
701	Pbo	USA	< 65	01-701-1015	ALT	ALANINE AMIL...	6	CHEMISTRY	34	26.00	U/L
701	Pbo	USA	< 65	01-701-1015	AST	ASPARTATE A...	9	CHEMISTRY	34	26.00	U/L
701	Pbo	USA	< 65	01-701-1015	BASO	BASOPHILS	0	HEMATOLOGY	0.2	0.04	G/L
701	Pbo	USA	< 65	01-701-1015	BLU	BILIRUBIN	0.2	CHEMISTRY	1.2	5.13	umol/L
701	Pbo	USA	< 65	01-701-1015	BUN	UREA NITRO...	4	CHEMISTRY	24	4.28	mmol/L
701	Pbo	USA	< 65	01-701-1015	CA	CALCIUM	8.4	CHEMISTRY	10.3	2.30	mmol/L
701	Pbo	USA	< 65	01-701-1015	CHOL	CHOLESTEROL	156	CHEMISTRY	300	5.64	mmol/L
701	Pbo	USA	< 65	01-701-1015	CK	CREATINE KI...	21	CHEMISTRY	169	64.00	U/L
701	Pbo	USA	< 65	01-701-1015	CL	CHLORIDE	94	CHEMISTRY	112	105.00	mmol/L
701	Pbo	USA	< 65	01-701-1015	CREAT	CREATININE	0.7	CHEMISTRY	1.4	88.40	umol/L
701	Pbo	USA	< 65	01-701-1015	EOS	EOSINOPHILS	0	HEMATOLOGY	0.57	0.08	G/L

Figure 5: Screen shot of the Details-on-Demand window displaying underlying source data based on the marking on the Lab Test Mean $\pm$ Standard Deviation page - restricted to display source data corresponding to 'WEEK 6'.



CONCLUSION

The over time plots developed with use of analytical tools such as TIBCO Spotfire® are highly customizable, allow the end user to not only view data interactively but also provide the ability to identify trends, filter data, drill down to details and export data as csv, pdf or ppt for more details. With very minimal data transformation the plots can be generated very quickly, which also help in reducing the review time significantly and to make decisions effortlessly.

REFERENCES

[1] Busa, B., & Pothuri, K. (2018). Building Dashboards for Real-Time Insights into Clinical Data.
<https://www.lexjansen.com/phuse-us/2018/dv/DV17.pdf>

[2] Plotting Time Series Data
<https://support.sas.com/rnd/app/ets/examples/gplot/index.html>

[3] Means Plot with Standard Deviation using calculated Data
<https://support.sas.com/kb/42/515.html>

CONTACT INFORMATION

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

Author Name: Kishore Pothuri
Company: Vita Data Sciences
Address: 281 Winter St, suite# 100
City / Postcode: Waltham, MA, 02451
Work Phone: 781-833-0259
Email: kpothuri@vitadatasciences.com
Web: www.vitadatasciences.com

Brand and product names are trademarks of their respective companies.