ABSTRACT

TIBCO Spotfire is an analytics and business intelligence platform which enables data visualization in an interactive mode. Users can create heat maps and map charts using inbuilt functions in Spotfire. The easiest way to understand a heat map is to think of a cross table or spreadsheet which contains colors instead of numbers. The default color gradient sets the lowest value in the heat map to dark blue, the highest value to a bright red, and mid-range values to light gray, with a corresponding transition (or gradient) between these extremes. Heat maps are well-suited for visualizing large amounts of multi-dimensional data and can be used to identify clusters of rows with similar values, as these are displayed as areas of similar color. Patterns in heat maps are clear, because colors are used to display the frequency of observations in each cell of the graph. Also, Map chart can be useful to show the population density in the world map. This paper will demonstrate some basic heat maps and map chart created using spotfire.

INTRODUCTION

TIBCO Spotfire is an analytics and business intelligence platform which enables data visualization in an interactive mode that has grown in popularity within the healthcare and various industries over the last few years. With its increasing implementation in the field of safety monitoring and data review, TIBCO Spotfire presents its capabilities for exploratory analysis.

Heat Maps are extremely versatile and efficient in drawing attention to trends, and it’s for these reasons they’ve become increasingly popular within the analytics community, but that’s just the tip of the iceberg as to why. While other data visualizations must be interpreted – either by analysts or business users – Heat Maps are innately self-explanatory. The darker the shade, the greater the quantity (the higher the value, the tighter the dispersion, etc.). When existing data visualizations are paired with Heat Maps, their ability to rapidly communicate key data insights to the viewer is greatly enhanced.

Heat maps use colors to communicate numeric data by varying the underlying values that represent red, green, and blue (RGB) as a linear function of the data. You can use heat maps to display spatial data, plot big data sets, and enhance tables. You can use colors on the spectrum from blue to red to show population density in a US map. In fields such as epidemiology and sociology, colors and maps are used to show spatial data, such as how rates of disease or crime vary with location. With big data sets, patterns that you would hope to see in scatter plots are hidden in dense clouds of points. In contrast, patterns in heat maps are clear, because colors are used to display the frequency of observations in each cell of the graph. Heat maps also make tables easier to interpret. For example, when displaying a correlation matrix, you can vary the background color from white to red to correspond to the absolute correlation range from 0 to 1. You can shade the cell behind a value, or you can replace the table with a shaded grid. Also, Map chart can be useful to show the population density e.g. site, subject, different parameters in the world map. This paper focuses on creating heat maps and map charts using inbuilt function and properties in Spotfire on a fly.

TECHNIQUE AND MECHANISM

The general process of creating data visualizations in TIBCO Spotfire is as follows:

1. Import Data in Tibco Spotfire containing the required variables.
2. Create visualizations e.g. Heatmap and Map chart as per user specifications.

TIBCO SPOTFIRE OVERALL VIEW

Below display will provide you a brief overview of the TIBCO Spotfire Development area. The
development area consists of the following four main windows which can be resized as per need:

1. **Data:** This window will provide the list of all data sets and variables available for the visualization. This window can be closed from the view tab.

2. **Filters:** This window will provide the list of variables available for sub setting the data. This list includes the common variables and modified flags added by the data prep macro. This window can be closed from the view tab.

3. **Details-on-Demand:** This window will provide a data set view of selected data in the visualization. It will provide information about the data set used for the visualization and the list of variables available in the data set. This data can be exported into Excel or .CSV files for further evaluation. This window can be closed from the view tab.

4. **Visualization Area:** This area contains all the visualizations. Multiple visualizations (e.g. graphs, bar chart, tree map, pie chart, box plot, heat map, map chart) can be added in one tab.

**HEAT MAP USING TIBCO SPOTFIRE**

The easiest way to understand a heat map is to think of a cross table or spreadsheet which contains colors instead of numbers. The default color gradient sets the lowest value in the heat map to dark blue, the highest value to a bright red, and mid-range values to light gray, with a corresponding transition (or gradient) between these extremes.

The example below shows how the values in the table are displayed as color gradients in the heat map cells. Here you can see that each value in the cross table is represented by specific color.
EXAMPLE 1: RACE DISTRIBUTION

Below interactive visualization heatmap is created using the Demographics (DM) from the SDTM database. This heat map will give an interactive view using RGB color gradient where user can easily identify the population density based on the intensity of color e.g. red is high and blue is low.

**Step 1:** Import the demographic data in Spotfire. See display 3 below.

**Step 2:** Right click on the table and switch to heat map visualization (see display 4 below).

**Step 3:** Right click and select properties. Select variable sex on X axis, variable race on Y axis, select cell values as unique subject count (see display 5 and 6 below). User can adjust the properties as per user specifications.
Display 4. Switching from Table view to Heat map view

Display 5. Select Properties
Display 6. Select variables for X and Y axis

Step 4: To understand the heat map user can add a cross table view and heat map on same visualization this will provide one-one matching between the value and color in heat map through tool tip information (see display 7 below).

Display 7. Race distribution heat map and table view

EXAMPLE 2: RACE DISTRIBUTION AND SCREENING OUTCOME

Below interactive visualization heatmap is created using the Demographics (DM) from the SDTM database. This heat map will give an interactive view using RGB color gradient were user can easily identify the population density based on the intensity of color e.g. red is high and blue is low. Here screening outcome heatmap and cross-table are added with the race distribution heatmap to give a good
overview of study. Since, the x-axis variable sex is common in both heat maps the color gradient will be combination of data values in both heat maps (see display 8 below).

Display 8. Multiple heat maps and cross table on one visualization

EXAMPLE 3: ADVERSE EVENTS

Adverse events (AEs) are a key focus of any safety review, for example the ability to quickly locate and isolate relevant AEs especially serious one.

Below interactive visualization heatmap is created using the Adverse Events (AE) from the SDTM. For, AE summary heatmap, serious adverse events were selected on the x-axis and y-axis is group by system organ class and derived dictionary term. Here user can easily identify a particular adverse event by severity and color index and later drill down the adverse events on patient level. The tooltip displays some summary information about the respective area (see display 9 and 10 below).

Display 9. Frequency of AE heat map properties
EXAMPLE 4: VITAL SIGNS SCATTER PLOT HEAT MAP TYPE VIEW

Scatter plots display a marker at the intersection of the values of an X variable and a Y variable. In contrast, heatmaps divide the graph into rectangular (or hexagonal) bins and use colors to show how many observations fall in each bin. If you have a large number of data points, then ordinary scatter plots, fit plots, residual plots, and so on become hard to interpret. If you have enough data, then points merge into large blobs that do not always reveal the underlying structure of the data. Heat maps differentiate more clearly between the denser and less dense portions of the data.

Below interactive visualization is created using the Vital Signs (VS) data set from SDTM database. It consists of a scatter plot of vital signs numeric results/findings over study timeline (see display 11 below). It is not straight forward to give scatter plot a density heat map type view. But, user can follow steps below to create the density heat map from scatter plot.
Display 11. Vital signs scatter plot

**Step 1**: Go to the properties and check the auto-bin column for both X and Y axis variables. This will give grid type view on the plot. Later, right click on the slider above the variable name and add the number of bin as per the view requirements. See display below 12 and 13 below.

Display 12. Select Auto-bin column from scatter plot
Display 13. Add number of bin to the scatter plot

**Step 2:** Select the shape in properties as “Tiles” marker and select marker by to “None”. Later, select the color columns as row count and pick the available color gradient scheme or create your own (see display 14). This will give the scatter plot a heat map type density view shown in side by side display 15 below. This map is useful to quickly identify the outliers based on colors.

Display 14. Select the color gradient for heat map type view
Display 15. Scatter plot and density type heat map view side by side

MAP CHART USING TIBCO SPOTFIRE

Map charts allow you to position your data in a context, often geographical, using different layers. The layers can be either data layers, such as marker layers or feature layers, or reference layers such as map layers. Here we will use the map layers inbuilt in Spotfire to visualize our geographical data as marker layers.

EXAMPLE 5: MAP CHART REPRESENTING STUDY SITES ACROSS GLOBE.

Below interactive visualization map chart is created using the Demographics (DM) from the SDTM database. This can be quickly created by switching the demographics data to map chart and later selecting the require variable and properties (these steps are already cover in example 1). See display 16 and 17 below.
Display 16. Demographics data
Below map chart is visualizing the location of sites across the world. This map chart can be helpful to determine the study site density across the globe. User can adjust the color and size through properties.

Display 17. Map chart representing various site across the globe

EXAMPLE 5: MAP CHART REPRESENTING MEDICARE PROVIDER ACROSS USA.
Below interactive visualization map chart is created using the Medicare service provider data from medicare.gov. This map chart will represent different hospital site (Medicare providers) across the USA. See display 18 for table view of data.
Display 18. Medicare providers across USA

Below map chart will visualize Medicare providers in USA by state, county, and city. User can update the color and size using the properties dialogue. See display 19 and 20 below.

Display 19. Medicare providers across the USA
Display 20. State-wise view of Medicare providers in USA

If user want to give a heat map type view to the Medicare providers data to observe the density, then it is possible to adjust the size by unique provider ID count through properties. Later, user can adjust the color gradient (Min and Max) to give a heat map type view (see display 21 below). This map is very useful to quickly identify densely populated area across USA.

Display 21. Heat map type view of Medical providers across USA
CONCLUSION

TIBCO Spotfire provides an interactive platform for exploratory analysis using data from various fields. Using inbuilt functions and adjusting the properties in Spotfire user can create various heatmap and map chart quickly. With its simplicity to adjust axes symbols and text, and its ability to export data for further user analysis/query, TIBCO Spotfire enables faster data review, quality assessment and process improvement.

REFERENCES


http://spotfire.tibco.com/
https://data.medicare.gov/Hospital-Compare/Heatmap/k7is-d4t4

ACKNOWLEDGMENTS

Thanks to Ryan Wilkins, Lindsay Dean, Philip Bartle, Ken Borowiak, Richard DAmato, Lynn Clipstone, and PPD Management for their reviews and comments. Thanks to my family for their support.

CONTACT INFORMATION

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

Ajay Gupta, M.S.
PPD
3900 Paramount Parkway
Morrisville, NC 27560
Work Phone: (919)-456-6461
Fax: (919) 654-9990
E-mail: Ajay.Gupta@ppdi.com, Ajaykailasgupta@aol.com

DISCLAIMER

The content of this paper are the works of the authors and do not necessarily represent the opinions, recommendations, or practices of PPD.

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.