Uncovering What Is inside Data using Effective Visualizations

Vipin Arora[1], Xiang Zhang[1]

[1]Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN, 46214

Abstract

Generating information for decision making including new drug or treatment approval based on clinical trials involves several years of planning, execution and organization. Most of the tools that are currently used in preparing late stage submissions have not changed (multiple sets of static tables, listings and figures (TLFs)) that are generated based on pre-defined statistical analysis plans. Although many of these TLF's are the result of several hypotheses being tested for the effectiveness and safety of a new treatment, several of these TLFs are results of many sub groups (pre-defined and post hoc). With the availability of recent tools (Spotfire, Tableau, R etc.) there are significant opportunities that have become available to explore the data and provide quick responses to customer questions including how the treatment performed in subgroups etc. Statisticians and data analysts are using these tools to respond to customer questions in a more efficient manner. However, there is a need to ramp up the usage of these tools especially in the Pharma side. This talk will provide examples of data visualizations from complex data with portability without additional burden of QC.

Key Words: Visualizations; Spotfire; Tableau; Complex; R

1. Visualization Enhances Understanding of complex data

Visualization tools can enhance understanding the efficacy and safety data in a more flexible and user-friendly manner by converting the complicated data into simple yet powerful message. Below are examples where visualization tool could play a significant role:

- Subgroups (Preplanned and post-hoc) and performance on components of a specific endpoint
- Cross combination of subgroups of interests or due to recent updates in field and/or due to competitive landscape
- Investigate by displaying patient level data and follow the patients to a certain stage in the study (achieving efficacy or to occurrence of an event of special interest of safety)

1.1 Data Visualization considerations
We propose the following rules to consider when implementing the visualization tool for your data:
- Always keep the customer in mind;

- Data visualization is a journey and not the end goal;
- Build up the information rather than a big dump;
- Test the concept before roll out;
- Build flexibility and try to connect the past with current and future.

1.2 Selection of Visualization Tool(s)

Selecting visualization tool(s) is a critical step in the implementation of data visualization. With the growing ability of computer science, fancy and complex visuals are available nowadays. However, we would suggest thinking carefully on "which tool do I really need?" and the following thoughts may help answering this question:

- Historical aspects (knowledge base developed in the group/organization);
- Associated costs;
- Flexibility to interface;
- A single visualization tool may not work for various business needs within an organization;
    - A selection of customized visualization tools is generally a good approach (e.g., Spotfire, Tableau, R Shiny;
- Requires investment and motivation, however the payoffs are significant.
    - Training of key team members on selective visualization tools is the key.

1.3 A Few Extra Thoughts

- Big data and bite sizes of meaningful information;
- Step by step approach to understand complex data but stay focused;
- Adaptability and learnings;
- Additional use of available tools or customized versions.

2. Examples of Visuals

2.1 Complex data visualization

Figure 1 demonstrates treatment B provides faster and more reduction in pain (50% to 75% from Baseline);
Figure 2 demonstrates treatment B provides faster and more pain reduction (75% or more from Baseline).

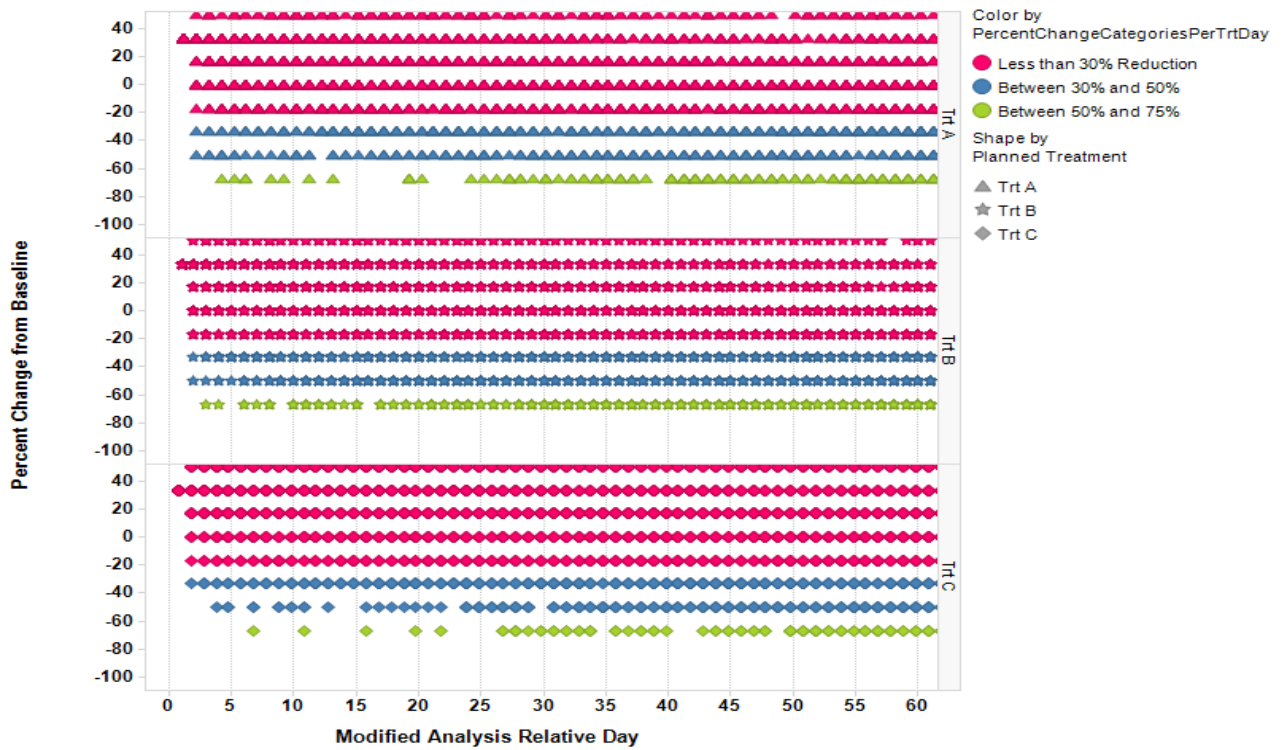**Percent Change from Baseline vs. Modified Analysis Relative Day for Pain**

Figure 1: Percent change from Baseline vs. Modified Analysis Relative Day for Pain

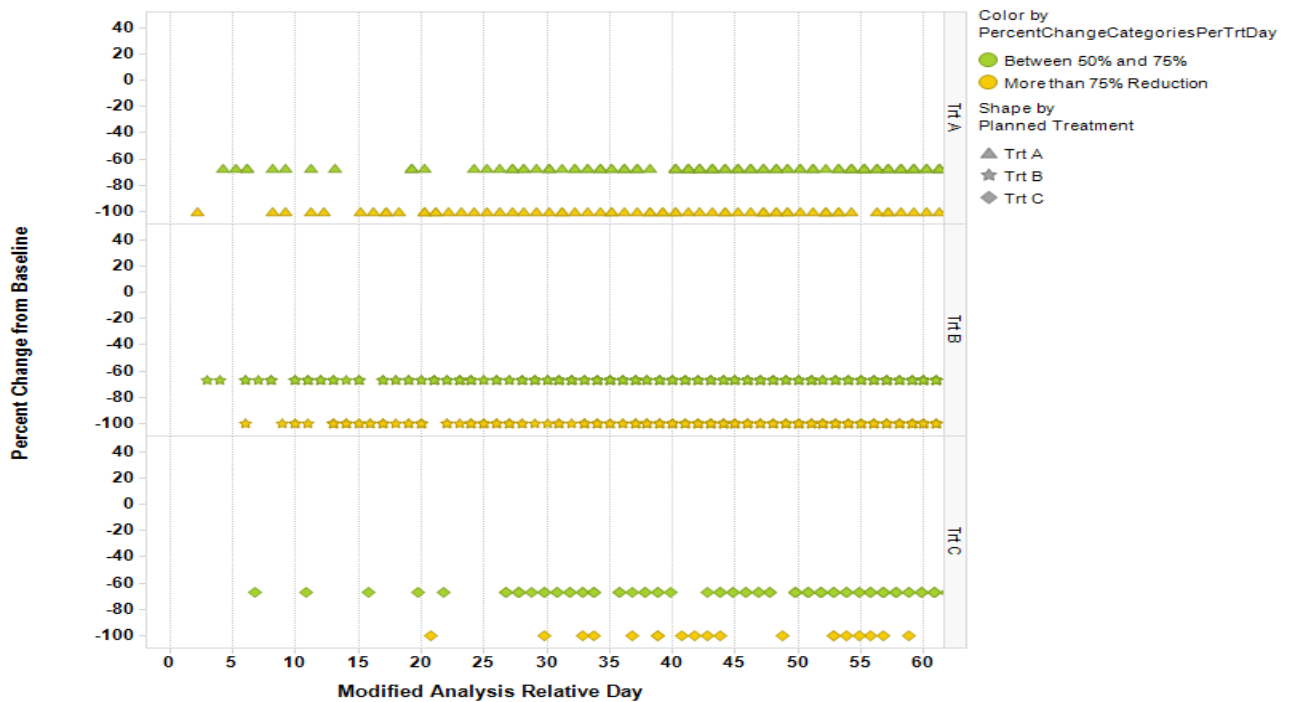**Percent Change from Baseline vs. Modified Analysis Relative Day for Pain**

Figure 2: Percent change from Baseline vs. Modified Analysis Relative Day for Pain

2.2 Spider Diagram

Spider Diagram is able to provide an intuitive vision of the dynamic progression of key outcomes among different treatment arms and to compare different outcomes simultaneously among treatment arms.
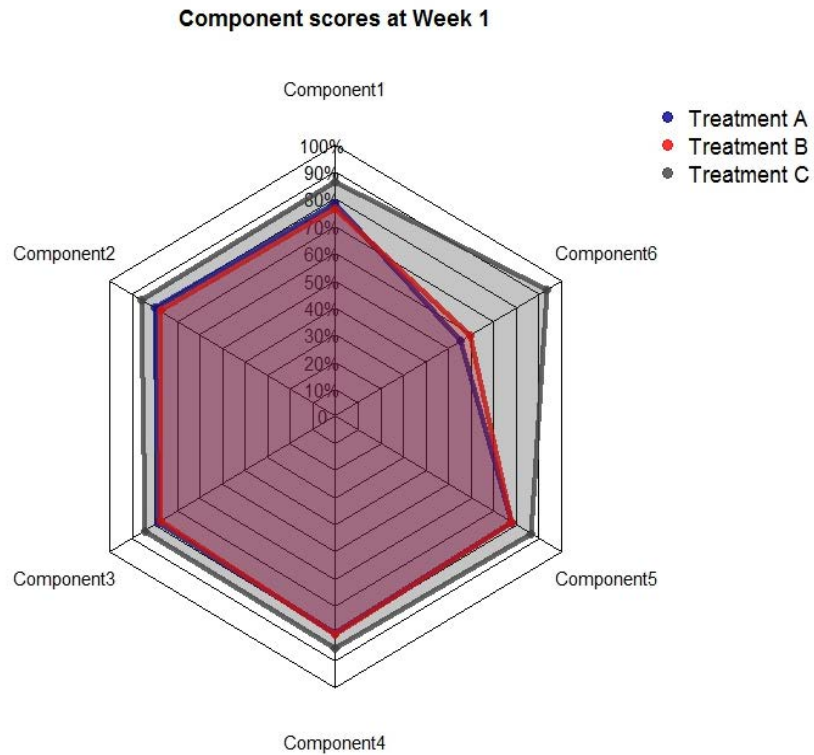

Figure 3: "%" means % of the mean value of corresponding component at baseline.

2.3 Rainfall Plots

Rainfall plots are able to illustrate the progression of the outcome of interest at patient-level at each time point of interest. Simultaneously, this visual are also able to provide the comparison at population-level across treatment arms. The stacked bar chart (upper left corner) to provide comparison (% of subjects on each status) across 3 treatment arms at each study visit; the line graph (upper right corner) to provide comparison (% of subjects on status of interest) across 3 treatment arms longitudinally.
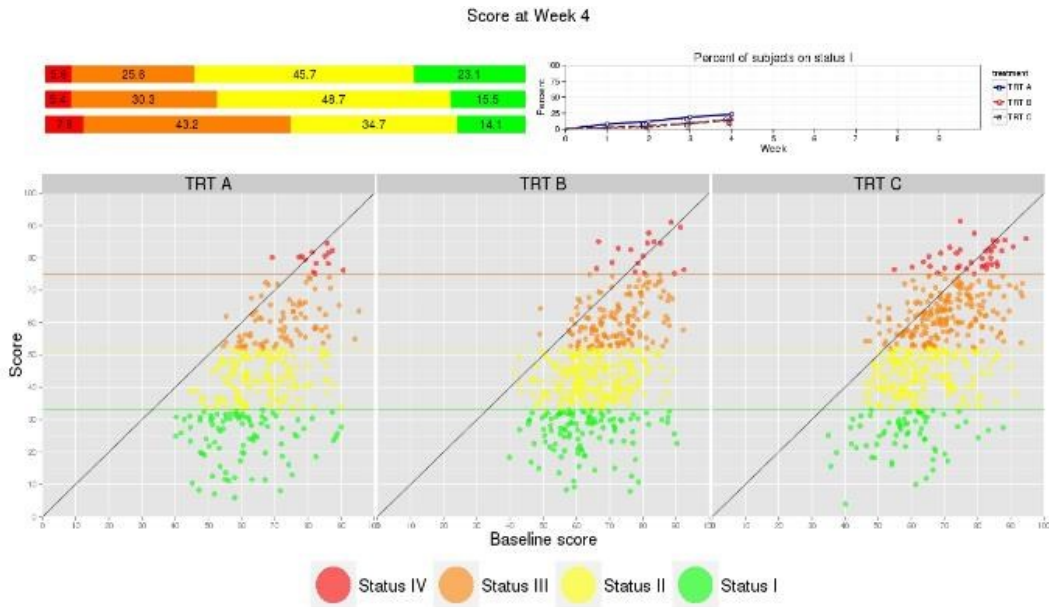
Figure 4: Rainfall Plots.

2.4 Sankey Plots

Sankey Plots is helping in showing the flow of categories over time or over disease severity.
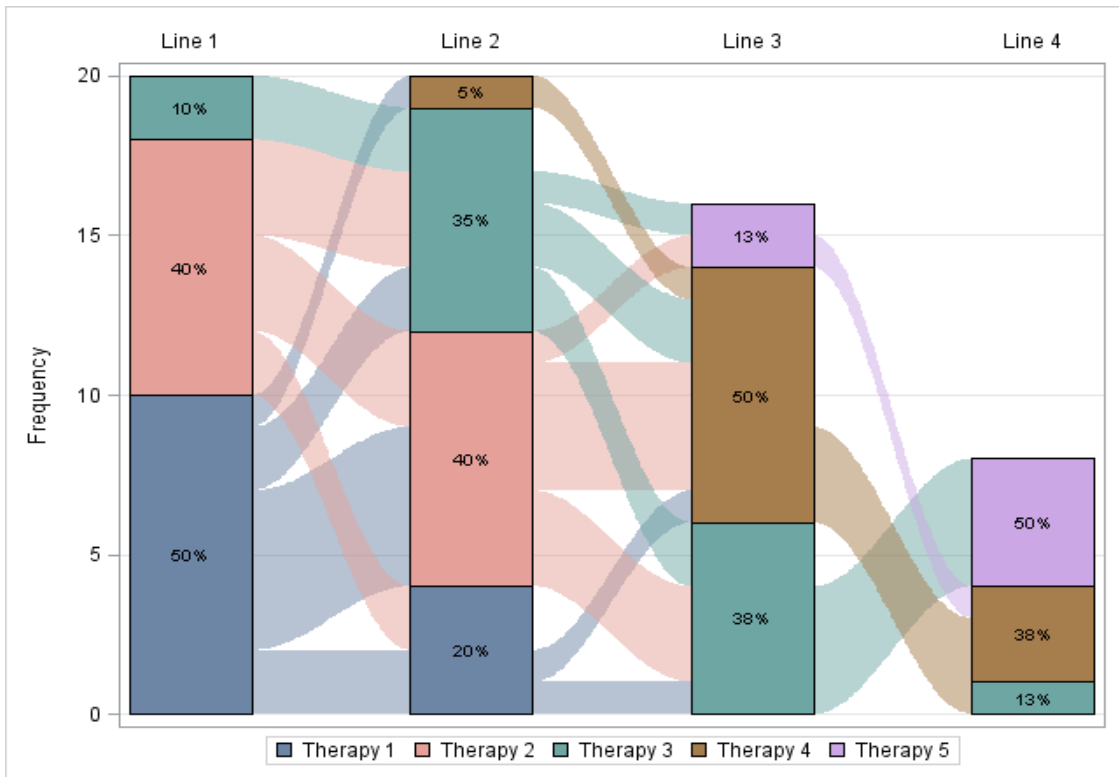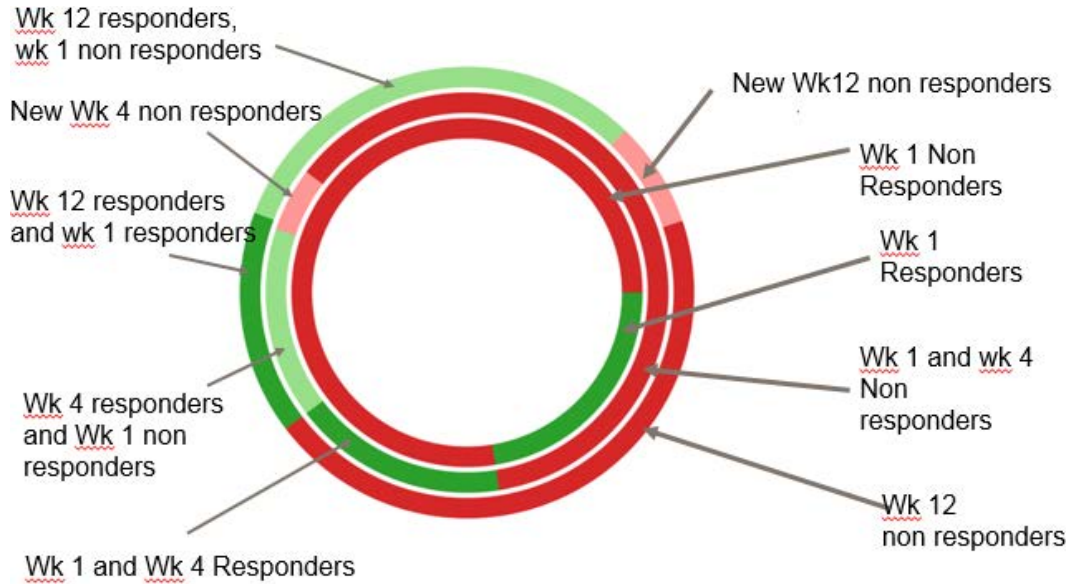

Figure 5: Sankey Plots.

2.5 Donut Plots

Donut plots are able to show the responders over time.

3. Summary

- Data Visualization along with Table and Listings provides a great combo for understanding complex data including trends and relationships;
- Data visualization provide a platform to confirm or generate new hypotheses;
- Data Visualization is beneficial to respond to questions from the audience without wait on the generation of new Tables and Listings.