Paper OS07

Choose your own adventure: a data story

Jiwan Heo, Atorus Research Inc., Toronto, Canada

Introduction

In the ever-evolving landscape of data visualization, transitioning from traditional statistical analysis system (SAS) approaches to a dynamic, multi-lingual environment presents both opportunities and challenges. As open-source technologies gain prominence in daily operations, the integration of visualizations becomes increasingly prevalent. A notable trend involves the replacement or supplementation of conventional Table, Listing, and Figures (TLF) with visual representations. While this marks a significant advancement, the full potential of visualizations often remains untapped, posing challenges for many organizations navigating this relatively new territory. In this paper, we explore a hierarchical approach to creating visualizations—a strategy that unveils a summary first, progressively adding layers of complexity upon drilling down. By addressing common obstacles and offering effective remedies, we aim to empower organizations to overcome barriers and unlock the true power of visualizations in their data analysis endeavors. This discussion serves as a practical guide, suggesting key strategies to enhance your visualization efforts and maximize the impact of your data representation.

Problem Statement:

Within the pharmaceutical industry, a prevailing paradigm involves submissions comprising an overwhelming number of files. Traditionally, the expectation is to meticulously sift through an extensive array of plots and tables, to extract insights on patterns and trends. This approach entails a mostly text-centric methodology where domain expertise is necessarily employed to discern patterns—often requiring a laborious effort. It's noteworthy that this methodology has been in practice for a considerable duration, and is often regarded as the standard, with limited exploration of alternative perspectives.

What's markedly underutilized in these text-heavy methods is the incredible perceptual abilities humans possess. People can swiftly scan, recognize what's important, and detect subtle visual changes, among other strengths. The reliance on text-centric approaches may inadvertently overlook the efficiency of human cognitive capabilities in processing complex information.

Solution:

Drawing inspiration from Ben Shneiderman's seminal work in his 1996 paper "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," we advocate adopting the "Visual Information-Seeking Mantra." The fundamental principle of this mantra is encapsulated in the sequence: "**overview first, zoom and filter, then details on demand.**"

Initiating with an overview allows for a holistic understanding of the data landscape. Subsequent steps involve zooming in on specific areas of interest and employing filters for targeted

exploration. Finally, details are revealed on demand, ensuring that human perceptual capabilities are optimally leveraged in the analysis process.

This visualization-centric methodology aims to address the limitations posed by traditional text-heavy methods, facilitating a more intuitive and effective means of extracting insights from voluminous datasets within the pharmaceutical domain.

Overview

We believe that offering an overview entails more than simply inundating the user with the entire dataset. Providing an overview should empower users to grasp the comprehensive picture, calling for a nuanced and concise summary. The first thing users encounter must be what's most important, ensuring a meaningful and efficient entry point into the dataset.

As designers, our task involves delving into the perspectives of stakeholders — understanding their desires, motivations, and the intricacies of their actions. By gaining insights into their tasks and the reasons behind their methods, we can discern the specific patterns or insights they seek. This comprehensive understanding guides us in crafting a thoughtful and effective overview that aligns with the users' needs and expectations.

We like to invite stakeholders to share their end-to-end process with us through a screenshare, enabling them to demonstrate their workflow without interruptions. Observing them perform their tasks while vocalizing their thought processes proves invaluable. To ensure the effectiveness of this approach, though, it's essential to enter the meeting with a basic understanding of their work. Without prior knowledge, the session may become challenging to follow, leading to a less productive use of time.

The next step involves reflecting on the session regarding its most important aspects. Once identified, your role is to contemplate the most effective way to present and communicate those key insights.

Zoom and Filter

The ability to control the content that users view through zoom and filter options enables them to concentrate on their specific areas of interest within a large dataset, eliminating irrelevant information. This functionality distinguishes visualizations from traditional text documents, freeing users from the need to sift through pages of data. Effectively, this functionality aids users in carrying out investigations, allowing them to formulate hypotheses and test them. It facilitates a more intimate exploration of the dataset, empowering users to delve into the details of their analysis. For example, users can focus solely on patients with recent adverse events or closely examine a specific section of a line graph.

Preparing for this design approach parallels the preparation for an overview. It involves closely observing how users navigate their tasks, noting when they pose specific questions and understanding the current processes that they employ to answer them. Subsequently, the goal is to enhance the intuitiveness and efficiency of these processes for the users.

Ensuring a productive collaboration with stakeholders involves setting the stage for them to articulate their workflow in a comprehensive manner. Encouraging them to provide a detailed,

non-simplified account is crucial. When stakeholders resort to a watered-down explanation, key details may inadvertently be omitted, leaving you perplexed about the coherence of the process. Therefore, it becomes your responsibility to acquire prior knowledge before your meeting to maximize the efficiency of your time together. This proactiveness is especially vital considering that opportunities to meet with stakeholders might be limited. For instance, interrupting them for clarification every time they use an acronym during the walkthrough of their investigative process is something you'd want to avoid.

Details on demand

This is the space where we encapsulate all the nuanced details found in traditional text documents. It's a valuable resource for those who opt to explore further. Ideally, this should be the concluding step for users, following their comprehension of the overview and the refinement of vast data to their specific interests.

Whether presented in the conventional style of a listing or as hovering information within an interactive visualization, the objective is to discreetly hold these detailed insights, unveiling them only upon request. This strategy ensures that these details don't overshadow the primary focus, allowing users to maintain their concentration on the aspects that matter most to them.

Sarepta example

Let's explore how our collaboration with the Medical Monitor team at Sarepta has translated our guiding principles into practical applications within the pharmaceutical industry.

At a broad level, the medical monitoring team at Sarepta meticulously reviews lab results to safeguard patient well-being during exposure to study drugs. These results are often categorized as “very low,” “low,” “normal,” “high,” or “very high,” each triggering distinct follow-up actions. For instance, a “low” blood pressure level may prompt a prescription for supplementary medications, while a “very low” reading may necessitate a dosage reduction or discontinuation of the study drug treatment. Beyond these apparent low or high values, they determine whether a specific reading constitutes an Adverse Event of Serious Interest (AESI), a Dose Limiting Toxicity (DLT), or a routine abnormality with minimal impact. The severity hierarchy follows a sequence of regular abnormality, AESI, and then DLT.

In their patient monitoring endeavors, the team delves into the full context behind each reading, especially when it signals potential issues. This involves scrutinizing adverse events and concomitant medications. This comprehensive approach enables them to decipher the reasons behind abnormal lab results. Once armed with a complete understanding, they make informed decisions on whether to intervene, prioritizing patient safety. It's worth noting that we've already introduced four datasets into the discussion, and we'll circle back to explore these in more detail.

Overview

As previously recommended, we initiated the process by engaging in meaningful discussions with the stakeholders, delving into a comprehensive understanding of their daily workflows. Prior to these meetings, we extensively collaborated with internal Sarepta teams to acquire the essential knowledge. Through this process, we keenly observed their priorities, identifying the paramount focus on detecting AESI or DLT events. This significance stems from the imperative

need to make prompt decisions in response to these events. Additionally, we recognized the importance of providing a holistic view of all events that transpired since the beginning of patient treatment.

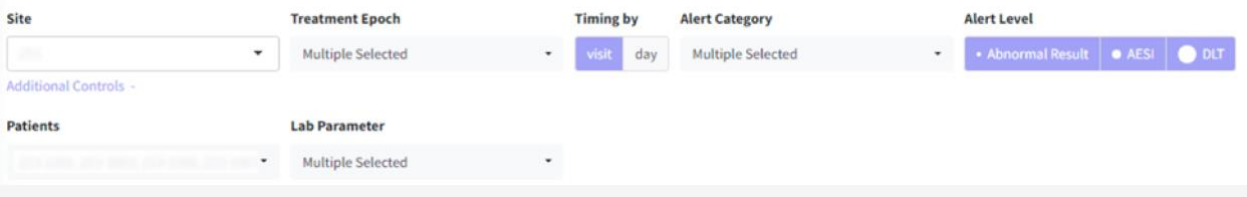
These insights provided us with two focal points for developing our visualizations. After thorough team discussions, here's what we've come up with.

Presented in the form of a beeswarm plot, each event is depicted as a circle, with size and darkness indicative of its severity. The circles align along their respective x-positions (visit week), forming clusters that avoid overlap. In instances of too many circles within one visit, they elegantly cluster together, even if it means they take up a lot of horizontal space. While this sacrifices a degree of x-axis precision, it ensures the display of each event distinctly, preventing circles being overlapped, and looking like a singular event.

Upon first look, this chart immediately draws focus to the larger (thus important) circles. Additionally, they can traverse the entire patient journey, making overarching observations such as identifying a concentration of events in the Screening phase for most patients. This compromise in x-axis accuracy enhances the narrative aspect, as medical monitors can quickly identify where the important events are. Subsequently, during their exploration of significant events, they can delve deeper into the details, rendering the initial overview more focused and informative.

Zoom and Filter

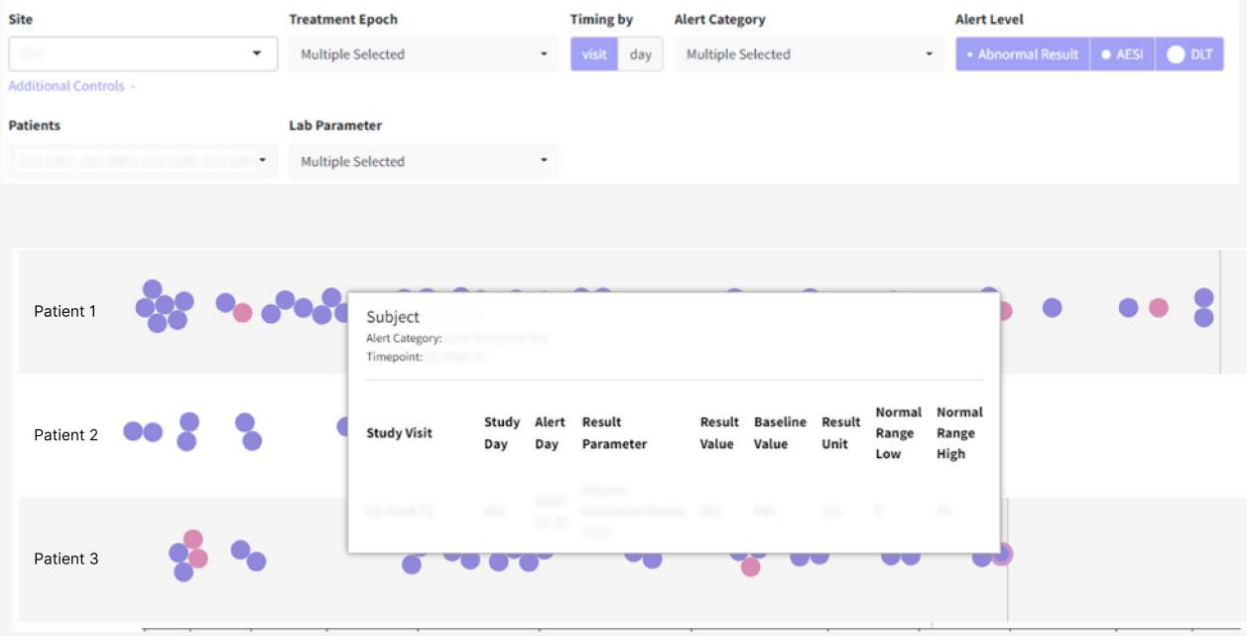
After initially exploring the overview, medical monitors now have a comprehensive understanding of the overall situation. Now, equipped with this insight, they are ready to delve deeper into specific patients, time periods, or event types.



In our design, we opted to not incorporate zoom functionalities, as the segmentation of data for each patient revealed the relatively short timeline rendered such features unnecessary. Instead, we provide users with a range of robust filter options, empowering them to tailor their analysis precisely. While the filters above are intuitively self-explanatory, it's noteworthy that users can filter specifically on AESI/DLT events, offering a powerful tool to significantly narrow down their search results.

Details on demand

Finally, in providing access to the raw data, we've designed two user-friendly approaches. Firstly, by hovering over a circle, users can effortlessly access key information presented in an aesthetically pleasing user interface. Additionally, for a more comprehensive view, users can navigate to a dedicated tab within the dashboard to explore the data in its original collection format. These options prove especially beneficial when stakeholders require a more traditional setting for sharing or recording their findings.



Tips

Now that we've gone through the main principles to keep in mind when designing visualizations, we'd like to go over some common mistakes or gotcha's to aid your designing endeavours.

Sketches

A common challenge we often encounter is the tendency for individuals to delve into technical implementation prematurely. While tools like plotly and ggplot are exceptional for creating

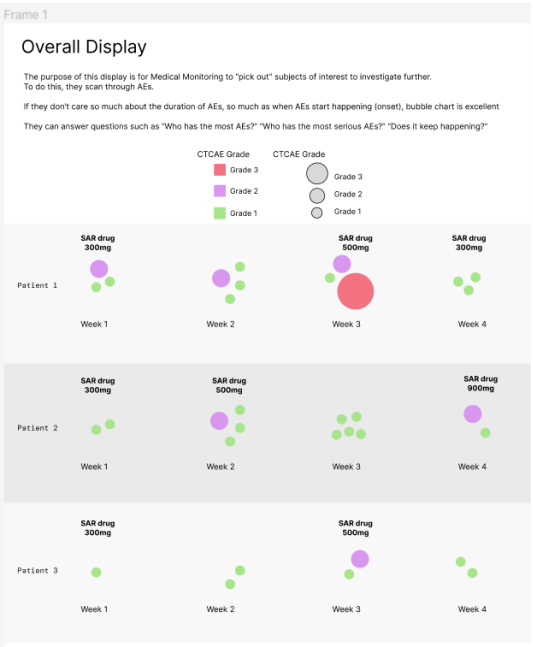
standard plots, they may fall short when it comes to more intricate or specific requirements. It's all too easy to dismiss innovative ideas with statements like, "How can I achieve that in ggplot?" or "That's a known bug in plotly," thereby prematurely eliminating potential options.

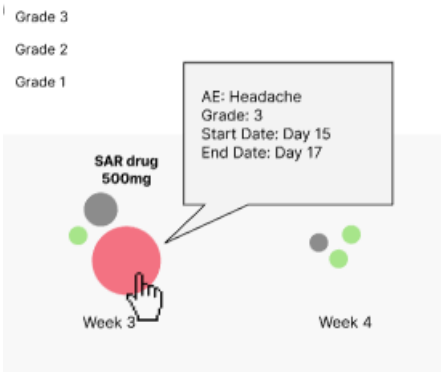
While scrutinizing technical aspects is undoubtedly a crucial step in any project, imposing limitations on oneself too early can hinder creative and unbiased brainstorming. Dedicate a significant portion of your time to brainstorming without immediately diving into coding. Consider taking a break from the screen and letting your ideas flow onto paper—messy notes, overlapping images, and all. The aim is to deliberately slow down the thought process, explore possibilities, and ultimately arrive at the best decision.

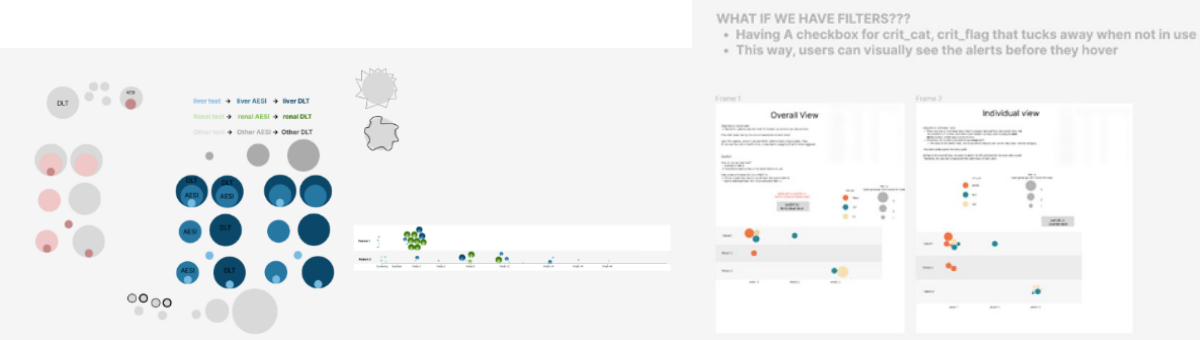
Repeated iterations of these visual brainstorming sessions are highly recommended. This approach allows you to test various ideas without becoming overly attached to any one of them. Even if you believe an idea is brilliant, others may not share the same perspective. Investing time in exploration before committing helps foster creativity and provides a solid foundation to revisit when facing challenges or confusion.

As creators of new tools, it's paramount to comprehend stakeholders' perspectives, divorcing ourselves from our own technical mindset. Instead of stating technical limitations like "plotly doesn't support that," engage stakeholders in conversations about their aspirations. Discussing technical challenges may inadvertently force them into a developer's mindset, leading to withheld information or simplified explanations. It's more beneficial to encourage stakeholders to share their dream scenarios entirely before making decisions about project scope.

Here are some screenshots of brainstorm we did on Figma.





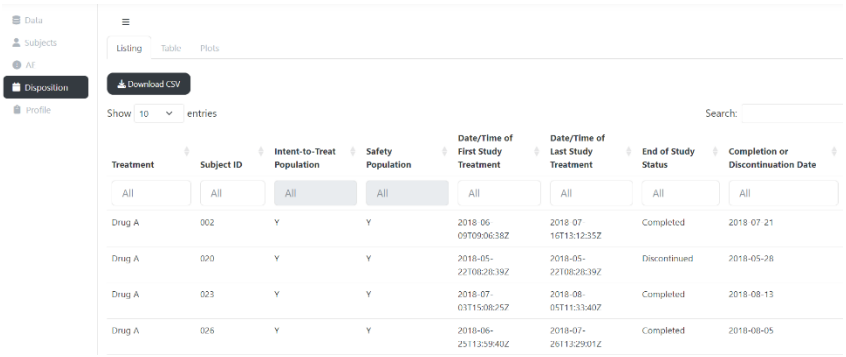


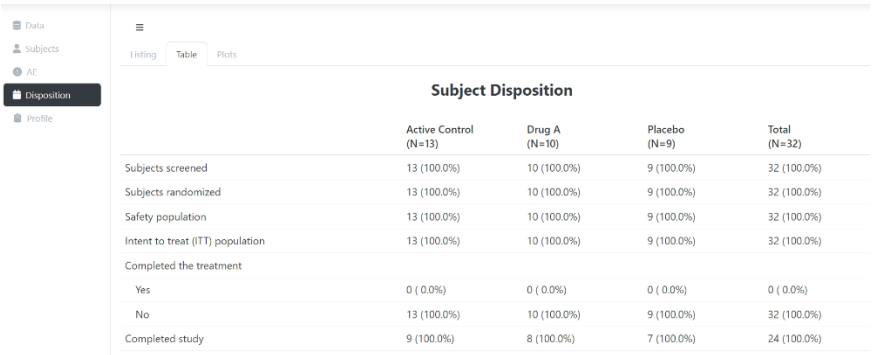
Keep things together!

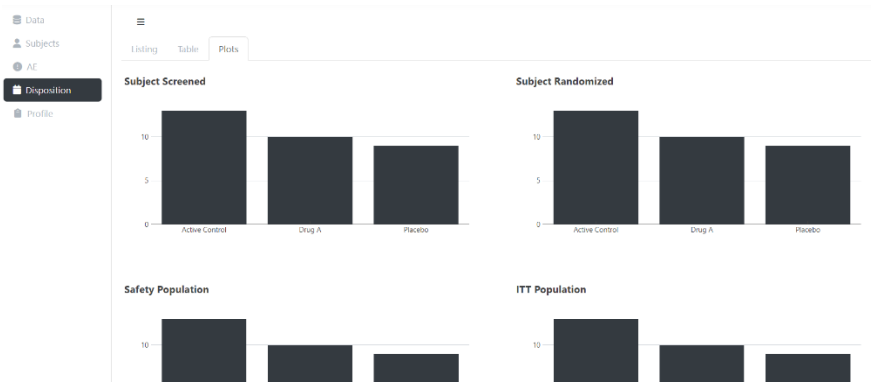
A crucial consideration in designing user interfaces is the strategic arrangement of elements, particularly grouping those that should be viewed together. While this might seem straightforward, it often proves challenging. Frequently, users find themselves toggling between dashboard tabs, scrolling extensively, or resorting to opening separate files to access the diverse information required for their tasks.

For instance, imagine examining a basic line chart displaying lab values over time in a Shiny app. Should the need arise to review adverse events alongside these findings, users might be compelled to open a separate Excel file, locate the pertinent patient within the specified timeframe, mentally note details, and then return to the app.

This back-and-forth navigation can strain the eyes and diminish focus. A common workaround is integrating adverse events listings within the app, organized in a separate tab. A prevalent dashboard structure includes a sidebar with different domains, each featuring various tabs such as listing, TFL table, or a visual.

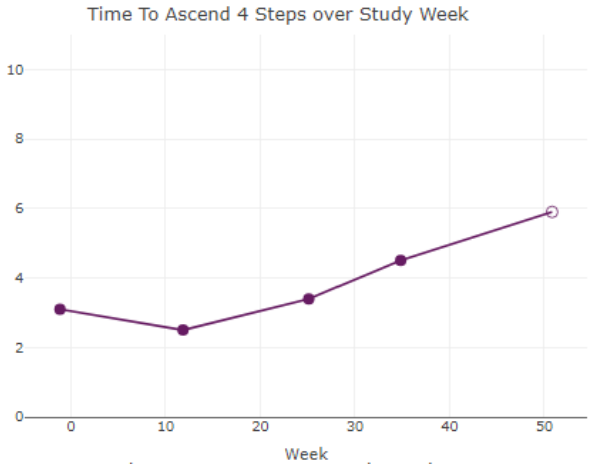




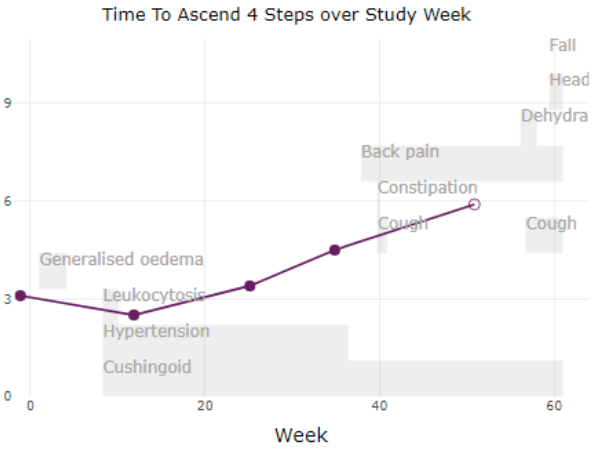


While this mitigates the need for opening a separate file, switching pages introduces a mental reset that disrupts continuity. Although there's no one-size-fits-all solution, we have found success in applying the "overview first, zoom and filter, then details on demand" principle to alleviate some of the challenges in this process.

Consider the following line chart depicting lab values:



Now, suppose we wish to investigate events around the second measurement point where it begins to increase. Typically, one would examine adverse events to identify potential factors affecting the patient's progress. Instead of navigating elsewhere for this information, we can overlay the adverse events on the line graph with adjusted opacity.



While this may introduce some clutter, it unmistakably indicates adverse events around the second measurement, providing insight into potential explanations for the lab value increase. Notably, users can explore this possibility without navigating to a separate location. From here, users can commit these events to memory and delve into more details.

We offer various filtering options at the top of the page, including an adverse event SOC/PT multiple selection that updates choices for each patient. Additionally, a listing at the bottom of the page presents raw data for user review.

By strategically placing UI elements, we propose an enhanced approach that minimizes the need for users to navigate between different sources for data retrieval. This allows them to concentrate on their core tasks without the cognitive load associated with information retrieval from disparate locations.

Conclusion

Implementing this design philosophy within the pharmaceutical industry presents a profound opportunity to elevate the collective efficiency of its operations. Venturing into the realm of visualization initiatives may initially appear daunting, particularly for individuals accustomed to generating RTFs/TFLs using predefined table of contents and well-established macros. The transition from familiar routines to the creation of something entirely novel can indeed be challenging. However, within this challenge lies the potential for transformative change. This paper serves as a comprehensive guide, offering practitioners a tangible checklist to navigate the uncharted waters of visual design. It aims to empower those accustomed to predefined structures by providing reassurance and guidance on their journey toward embracing a more creative and efficient approach.

Contact Information

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

Author Name: Jiwan Heo

Company: Atorus Research Inc.

Email: jiwan.heo@atorusresearch.com