CS573 Prospectus

Xioshuai (Maksim) Li
Data Science
Worcester Polytechnic Institute
Worcester, MA, USA
xli3@wpi.edu

Akim Ndlovu
Computer Science
Worcester Polytechnic Institute
Worcester, MA USA
andlovu@wpi.edu

Noëlle Rakotondravony
Computer Science
Worcester Polytechnic Institute
Worcester, MA USA
ntrakotondravony@wpi.edu

1. Introduction

The current surge covid related visualizations due to the viruses novelty has brought about many questions related to visual design principles. Many design principles are overlooked as designers rush to chart the data at hand. This however has long lasting consequences as it not only creates anxiety when faced with a pandemic like covid 19, but also reduces confidence from the audience, in using visual communication channels to effectively and accurately communicate information [1]. When visualizations are developed without considering design principles they often tend to be ambiguous and in most cases misrepresent or overlook essential information.

Although technical aptitude is imperative in developing appealing visual designs, a critical analysis of the audience, information, analysis requirements and decisions to be made is quintessential. A chart designed for scientists wouldn't be suitable for the general public, the opposite being true, as their interpretation and the desired information from the visualization would probably vary [2]. Based on this philosophy it is therefore important to ensure that design principles are adhered to, whilst taking into account the relevance of the information projected by the visualization to its audience.

For this project we critically analyse a visualization of the availability of hospital beds in the state of Massachusetts shown in Figure 1. The visualization encodes a line mark with a vertical spatial position channel for the number of beds, and a horizontal spatial position channel for the different categories that the beds fall into. The visualization implements the use of a color channel to further split the categories into ICU, Non ICU and Alternate Beds. The project will highlight how the data visualization succeeds or fails to communicate the desired message effectively. In addition to the critical analysis we propose an alternative interactive visualization for conveying the same information.

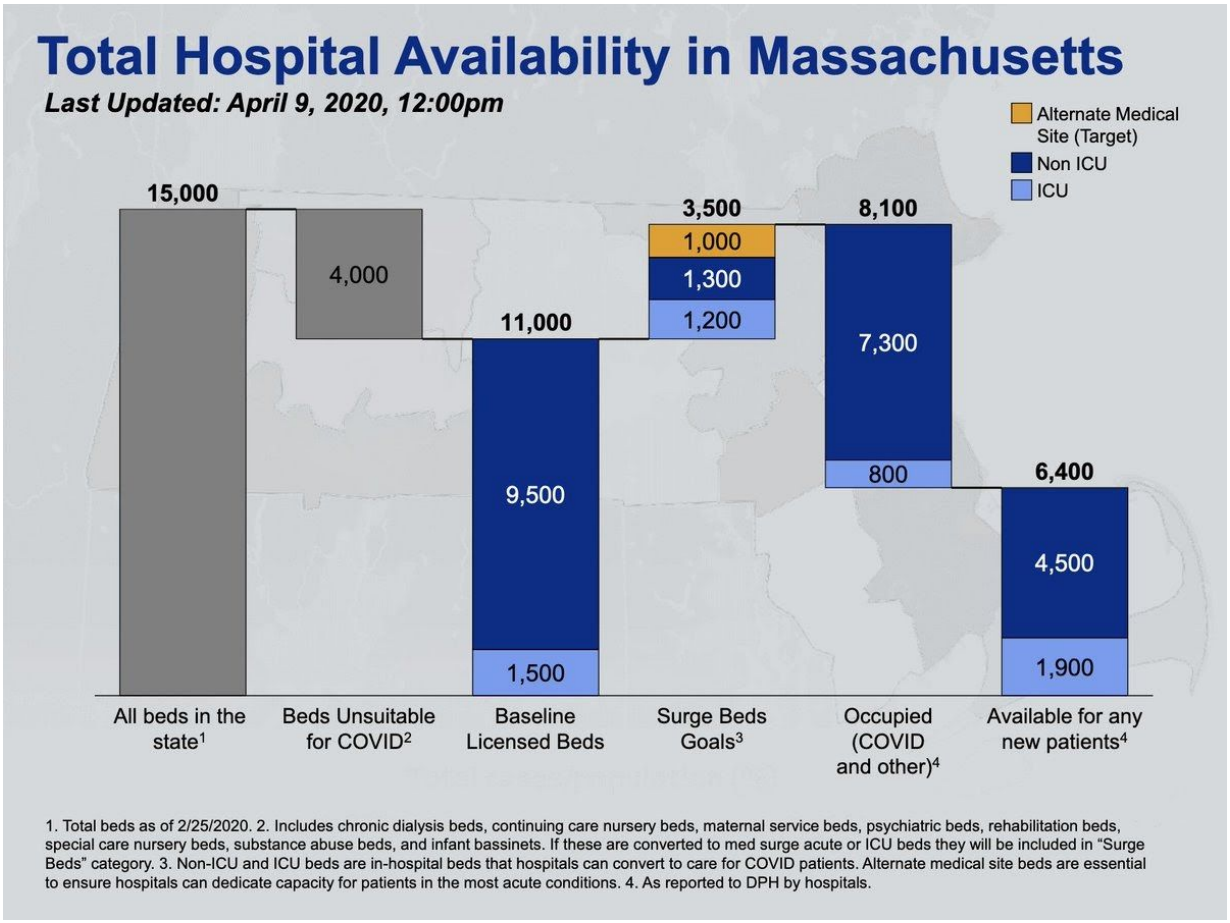

Figure 1: Total Hospital Bed Availability in Massachusetts [3].

Data visualization story line

There is a total of 15,000 bed in the state, of which

4,000 are unusable for covid and

11,000 are baseline licensed beds

Of the 11,000 usable beds for covid,

9,500 are non ICU beds

1,500 are ICU beds

As of the time the chart was published, of the baseline beds,

8,100 are occupied for both covid and other → leaving a total of 2900 available beds (all categories)

Of the occupied beds,

7,300 are non ICU → leaving 9,500 - 7,300 = 2,200 available non ICU beds

800 are ICU beds → leaving 1,500 - 800 = 700 available ICU beds

Moreover, the state targets a total of 3,500 surge beds that can be made available as ICU or non ICU beds if needed:

1,000 are from alternate medical site → making a total of 2,200 + 1000 = 3,200 non ICU beds

1,300 non ICU beds → making a total of 3,200 + 1300 = 4,500 non ICU beds

1,200 ICU beds → making a total of 1,900 ICU beds

Therefore, the total number of available beds is 4,500 + 1,900 = 6,400 beds

2. One-sentence description

A critical analysis of the use of color and magnitude scales for the visualization of hospital bed availability in the state of MA.

3. Project Type

The project will provide a critical analysis or evaluation of the current Massachusetts hospital availability visualization. Our ultimate goal is to enhance the infographic by adding interactivity to cater for what-if scenarios.

4. Audience

The project will be used as an academic reference for the data visualization community as well as the general public and for communicating information related to the distribution of hospital beds for the state of Massachusetts.

A significant number of visualizations have been published on covid-19 [4, 5, 6], however the most pertinent question is whether these visualizations communicate information in a simplistic and effective manner. Covid data has been made publicly available [7], fostering the increased number of visualizations being developed. Although these visualizations prove to be much better than trying to decode a table of numeric values, it is essential to ensure that due process is adhered to [1].

This project will attempt to implement data visualization best-practice techniques in order to enhance the visualization. In the event that the project falls short of its expectations, a section highlighting possible areas of improvement and further research, will be clearly documented.

5. Approach

5.1 Details

We intend to redesign the visualisation in question thereby making it easier to comprehend. This will be done using a Parallel Set chart, a tree map, and a line chart. The visualizations will augment each other to communicate the breakdown of hospital bed availability in Massachusetts. We will take advantage of parallel sets to clearly illustrate hierarchical levels of bed categories and capacity in Massachusetts. For each level, there will be a fastidious set of tree maps to show the detailed distribution of hospital bed capacity in terms of their functionality. We will also employ the use of a line chart to present the trend of the COVID-19 pandemic in Massachusetts to allow the audience to quickly assess COVID-19 cases in relation to hospital bed availability in Massachusetts.

5.2 Evidence for Success

Our alternative visualization(s) will leverage on the results a critical analysis of the current visualization, along with the implementation of established data visualization techniques best suited for visualizing hierarchical data [8]. The importance of interactive data visualizations is well recognised in the data visualization community and is often encouraged in the design of successful visualizations [9]. For this project we also propose an interactive chart, where the users can walk back and forth through the data.

6. Best-case Impact Statement

For this study, we critically evaluated the visualization of total hospital bed availability in Massachusetts and identified shortcomings in the interpretation of the visualisation. Furthermore we proposed the adoption of well established charts (parallel sets, tree maps and line charts), to tell a story about Hospital Bed Availability in Massachusetts. Our proposed visualization provides the audience with the ability to carry out what-if analysis given the added interactivity. As further research, we propose a feedback analysis study of the proposed visualization by the target audience. The results of the proposed analysis can thereafter be used as a measure of success for the proposed visualization. In conclusion, we believe that the use of best practises in developing visualizations should not be frowned upon, irregardless of the urgency associated with the development of these visualizations.

7. Major Milestones

- Identify the data that the original author wanted to communicate
- Identify the target audience
- Critique the data visualization
- Brainstorm alternative visualizations
- Evaluate the proposed ideas and select the most suitables charts based on their relevance
- Implement the proposed visualizations

8. Obstacles

8.1 Major Obstacles

- The true interpretation of the current visualization and the audience it was designed for remains a hypothesis.

8.2 Minor Obstacles

- Implementing the ideal visualization(s) for the appropriate audience and data presented.
- What kind of visualization(s) channels should be selected for presenting the COVID-19 cases information.
- Augmenting the visualization(s) with interactivity.

9. Resources Needed

No additional resources will be required to complete the project.

10. Related Publications

The addition of interactivity on static info visualizations affords the potency to walk back and forth through the data, presenting multiple views and scenarios which were previously impossible [9]. The study assesses the impact of embedding text into infographics, signifying that most people lack interest in the aforementioned visualisations and suggest a concept of suggested interactivity to engage the audience. Given that interactivity is a key component in our proposed solution, this study works as the underpinning of this project.

Affording the audience the ability to walk back and forth through the data whilst telling a story also forms an integral part of this project [10]. The golden standard for this project is to be able to provide a visualization which will not only engage the audience but allow them to get as much information out of it, pursuing questions and uncovering patterns [11]. The study provides best practise guidelines to that effect.

Study [12], gives an in depth analysis of hybrid sankey diagrams which are similar in nature to parallel sets. As we intend to implement parallel sets as a proposed solution, an in-depth understanding of the when, how and why of these charts is imperative for this project.

Furthermore the proposed solution seeks to provide the ability of carrying out what-if scenarios, in order to provide an in-depth analysis of future requirements based on data variations. Study [13], gives an overview of how the variations in covid-19 cases can be used to anticipate an increase in the demand of hospital bed availability, minimizing the risk of an overwhelmed health care system.

11. Success Factors

In order for this work to be publishable, we need:

- The designers perspective of the visualization
- A survey of the usability of our proposed visualization, in order to assess its effectiveness
- An evaluation process to measure and quantify the extent to which our proposed interactive visualization is better than the initial visualization proposed by the designer
- To incorporate “what-if” analysis into the visualization
- To design a tool that can be usable across all states to support an unambiguous decision-making process

References

[1] Makulec, A., 2020. 10 Considerations Before You Create Another Chart About COVID-19. [online] Tableau. Available at: <https://www.tableau.com/about/blog/2020/3/ten-considerations-you-create-another-chart-about-covid-19> [Accessed 17 April 2020].

[2] Emery, A., 2015. Why Your Audience Matters In Data Visualization. [online] TechChange. Available at: <https://www.techchange.org/2015/05/21/audience-matters-in-data-visualization/> [Accessed 17 April 2020].

[3] Massachusetts Department of Health, 2020. Total Hospital Availability In Massachusetts. [image] Available at: <https://twitter.com/TommyVitolo/status/1248691712471584769/photo/1> [Accessed 17 April 2020].

[4] Sccgov.org. 2020. Coronavirus (COVID-19) Data Dashboard - Novel Coronavirus (COVID-19) - County Of Santa Clara. [online] Available at: <https://www.sccgov.org/sites/covid19/Pages/dashboard.aspx> [Accessed 17 April 2020].

[5] Brittain, A., Mellnick, T., Keating, D. and Fox, J., 2020. How A Surge Of Coronavirus Patients Could Stretch Hospital Resources In Your Area. [online] Washington Post. Available at: <https://www.washingtonpost.com/graphics/2020/investigations/coronavirus-hospitals-data/> [Accessed 17 April 2020].

[6] Schoen, J. and Mangan, D., 2020. *Coronavirus Cases: These States Face Biggest Potential Shortfalls In Hospital ICU Beds*. [online] CNBC. Available at: <https://www.cnbc.com/2020/04/06/coronavirus-cases-states-with-biggest-hospital-bed-shortfalls.html> [Accessed 17 April 2020].

[7] GitHub. 2020. Cssegisanddata/COVID-19. [online] Available at: <https://github.com/CSSEGISandData/COVID-19> [Accessed 17 April 2020].

[8] Munzner, T. and Maguire, E., 2014. Visualization Analysis & Design. 1st ed. CRC Press, pp.94 - 114, 200 - 216.

[9] Boy, J. et al. (2016) Suggested Interactivity: Seeking Perceived Affordances for Information Visualization. IEEE Transactions on Visualization and Computer Graphics. [Online] 22 (1), 639–648.

[10] Segel, E. & Heer, J. (2010) Narrative Visualization: Telling Stories with Data. IEEE Transactions on Visualization and Computer Graphics. [Online] 16 (6), 1139–1148.

[11] Heer, J. & Shneiderman, B. (2012) Interactive Dynamics for Visual Analysis. Queue. [Online] 10 (2), 30–55.

[12] Lupton, R. . & Allwood, J. . (2017) Hybrid Sankey diagrams: Visual analysis of multidimensional data for understanding resource use. Resources, Conservation & Recycling. [Online] 124 (C), 141–151.

[13] Cavallo JJ , Donaho DA , Forman HP . Hospital capacity and operations in the coronavirus disease 2019 (COVID-19) pandemic: planning for the Nth patient.

[14] W. A. Pike, J. Stasko, R. Chang, and T. A. O'Connell. The science of interaction. Information Visualization , 8(4):263–274, 2009.

[15] J. van Wijk. The value of visualization. IEEE Visualization , pages 79–86, 2005

[16] Correll, M., 2020. Visualization Design Principles For The Pandemic. [online] Medium. Available at: <https://medium.com/@mcorrell/visualization-design-principles-for-the-pandemic-e65388280d16> [Accessed 17 April 2020].