

Innovating Engineering Education Analysis through Creative Data Visualization

Aidan Kenny, Northeastern University

Dr. Andrew L. Gillen, Northeastern University

Andrew L. Gillen is an Assistant Teaching Professor at Northeastern University in the First Year Engineering Program and an affiliate faculty member to Civil and Environmental Engineering. He earned his Ph.D. in Engineering Education from Virginia Tech and B.S. in Civil Engineering from Northeastern University.

Innovating Engineering Education Analysis through Creative Data Visualization

Abstract

Data is the foundation of informed decision making and drives progress and innovation in virtually every aspect of modern society. It empowers us to uncover insights, patterns, and trends that inform strategy, policy, and advancement across diverse fields. However, without visualizations to make sense of the trends, data is often useless. Data sets of all sizes are often too much for us to understand without the proper visual tools to simplify them. This paper seeks to explore how creative data visualization can function as a form of data analysis, enhancing the viewers understanding and engagement with the data. For the purposes of this work, creative data visualization is defined as the process of utilizing design elements with data representations to improve understanding of trends and increase emotional response. Creative methods of visualization allow for more information to be conveyed in a simpler format resulting in a stronger response from the viewer. The goal of this work is to develop visualizations for three case studies that address relevant issues within the field of engineering education and demonstrate the effectiveness of creative data visualization strategies. The first case study addresses the sex imbalance within engineering institutions, highlighting a slowly changing future for the field. The second study delves into the devaluation of LGBTQ students in engineering education, emphasizing the importance of inclusivity. Lastly, the third case study creates a visual map of the holistic engineering education experience. Case studies 1 and 2 compare traditional and creative data visualization techniques, whereas case study 3 introduces new visualization for understanding the engineering education field. Bar plots, heatmaps, infographics, and systemograms are explored in this paper. This work not only enhances the understanding of the critical issues addressed in the case studies but also highlights the potential of creative data visualization in addressing multifaceted challenges.

Introduction

Academic research in all fields is driven and supported by data. It is fundamental to decision making as it provides evidence to support hypotheses, refute past claims, and give insight to patterns and trends. However, raw data alone is often difficult to interpret and may not be meaningful to the user. Visualization, on the other hand, offers a way to represent data in a simple and accessible format, thus making it easier to understand and analyze complex data sets [1]. Furthermore, traditional analysis techniques often pose barriers to interpretation by non-experts. There are a multitude of visualization methods ranging from simple tables and charts to elaborate interactive graphics and animations. Different techniques can highlight various aspects of the data. By using the appropriate visualization methods, it becomes easier to identify the underlying relationships [2]. Finally, visualizations evoke emotional responses and provide a compelling story that helps to communicate key findings [3]. By considering various data visualizations methods as well as the design principles used to present them, understanding and interpretation by the user can be improved.

In the 1890's, W.E.B. DuBois published artistic visualizations of African American civil rights violations that grabbed the public's attention and clearly displays the intended conclusions [4]. These pioneering data visualizations demonstrate the profound significance of creativity within

the field of data representation and analysis. Through his use of intricate hand-drawn charts, graphs, and maps, DuBois demonstrated that the presentation of data could go beyond statistics and become a powerful tool for conveying complex realities. For example, the visualizations not only illuminated racial disparities but also underscored the need for social change. Furthermore, his visualizations showcased the connection between visualization and the public's response. DuBois's works demonstrate that properly presented data can affect audiences' opinions, evoke emotions, raise awareness, and prompt action. His work shows that understanding the public's response to data allows for effective visualization techniques and demonstrates the necessity for more creative visualizations in modern fields to inspire change. A selection of DuBois's visualizations is shown below.



W.E.B DuBois Visualizations for the 1900 Paris World's Fair [4]

This project bridges fields of data science, engineering education, and visual design to demonstrate that visually appealing and novel data visualizations serve the author's purpose more effectively than traditional methods. Furthermore, creative data visualization methods can be a form of data analysis. In essence, data can be presented in unique ways such that it highlights important patterns, engages the viewer, and deepens understanding. Public data sets were utilized to investigate various kinds of visualizations. Given the increasing importance of large data sets to all aspects of society, the outcomes of this research will be widely applicable.

Methods

In order to innovate engineering education analysis through creative data visualization, the research was divided into several key phases.

Exploration

The initial stage of the research involved a literature review to gain more understanding of data visualization techniques, engineering education topics, and visual design principles. Academic databases, journals, and research papers were consulted to develop case studies that investigated relevant engineering education topics. While exploring data visualization methods, we made note of techniques that resulted in more creative final products. Simultaneously, publicly available engineering education data was procured from sources such as the National Center for Education Statistics (NCES). Engineering education is a colorful, dynamic, and deep field that addresses a wide range of topics. Thus, data was narrowed to focus on insights into student demographics, marginalized groups, educational outcomes, and engineering education resources. The

exploration phase allowed for uninhibited examination of data sets in order to identify patterns and develop research questions.

Abstraction

For each dataset that showed interesting patterns, researchers began by making a traditional visualization to assess the effectiveness of conventional data representations. It is difficult to define what makes a visualization technique creative versus traditional. For the purposes of this study, the traditional graphs were made with basic functions in Excel with enough detail to get the point across. Next, creative visualizations were made for the same data utilizing novel methods and creative flair to improve readability, impact, understanding, and emotional response. For this stage, many different visualization methods were created for a variety of research questions, valuing quantity of products over quality. Excel, Lucidchart, Tableau, and InDesign were all used to create new visualizations.

Evaluation

With a better idea of which visualizations revealed underlying patterns more effectively than others, the research questions were narrowed to three case studies. These case studies explored various niche topics in engineering education and were a guiding framework for the remainder of the research. The visualizations and case studies were chosen based on their effectiveness while also showing a diversity of techniques. For example, many of the visualizations from the earlier phase used color heatmaps on top of visualizations to add another dimension to the data. In the final visualizations, this was limited to case study 1 in order to fully explore other techniques. At this stage, the visualizations were remade with a higher degree of refinement, utilizing Excel Macros and Adobe InDesign to create more visually appealing products.

The visualizations were evaluated qualitatively by the researchers with an emphasis on how effectively the information was conveyed, the extent of viewer engagement with the visualization, and emotional response. First impression factors that the researchers looked for included clarity, coherence, and visual appeal. Visualizations were studied to check whether they effectively conveyed the trends and patterns within the data. Emotional response was more difficult to describe and is confounded by the subject of the data as well as viewer biases. This is still an important factor as emotional response to visualizations largely impacts how they are interpreted [13]. This methodology revealed differences between traditional and creative visualization techniques. However, the study is limited without the context of quantitative results and human subjects. This highlights the need for future work to fully understand the impact of creative data visualization on viewer understanding.

Case Study Design

This work demonstrates the effectiveness of creative data visualization through three case studies. The first case study addresses the sex imbalance within engineering institutions, highlighting a slowly changing future for the field. The second study delves into the devaluation of LGBTQ students in engineering education, emphasizing the importance of inclusivity. Lastly, the third case study creates a visual map of the holistic engineering education experience. For case studies 1 and 2, researchers compare the traditional visualizations with creative visualizations for the same data set. Case study 3 proposes a useful visualization for the field of engineering education as a whole. Throughout the research process, many visualization

techniques were explored. The current paper focuses on techniques such as bar plots, heatmaps, infographics, and systemograms.

Human Factors

It must be addressed that viewer interpretation of visualizations does not rely solely on the visualization itself. A variety of human factors and biases change how the viewer will be impacted by the data. First, familiarization with data visualizations would affect the viewer's level of understanding. For example, people in academia or data scientists who work with data visualization everyday may be predisposed to understand visual patterns. This may also increase the complexity of visualizations that are readily understandable. Visualization interpretation is heavily susceptible to biases. Confirmation bias may skew the viewers' understanding based on what they already believe. Anchoring bias is a common problem for design-heavy visualizations. The viewer may fixate on the elements that attract the most attention, causing them to miss the trend. Finally, cultural and socioeconomic differences could change the meaning of data. Personal values and experiences affect how people interpret colors, symbols, and time. Similarly, disabilities may inhibit proper communication of trends in data. Patterns communicated solely through color may lose their impact or not be noticed at all by people with color blindness. To combat this, researchers used redundancy. Important metrics are communicated in more than one way such as size, color, text, and patterning. This concept was utilized in making effective data visualizations for each case study.

Results and Discussion

Case Study 1: Sex Imbalance in Engineering Bachelor's Degree Completions

The field of engineering is often associated with innovation and advancement. However, a pervasive challenge within this discipline is the sex imbalance of its institutions and workforce. Despite recent societal efforts to promote gender equality, engineering continues to exhibit drastic underrepresentation of women. This carries issues related to equity as well as the diversity and innovation potential of engineering professions. Women comprise only 29% of the science and engineering workforce and the ratio of men to women varies widely based on specific fields. For example, in 2013, only 15% of engineers were women. This figure drops to 8% for mechanical engineers and 11% for electrical engineers [5, 6]. Science and engineering are necessary contributors for advancement in the global society, so it is crucial to understand the underrepresentation of women in these fields.

Data from the Integrated Postsecondary Education Data System (IPEDS) was used to construct Figure 1. The visualization shows the number of male and female students that complete bachelor's degrees in engineering for a given year between 2012 and 2021. The secondary axis on the right shows how the male to female ratio is changing over this time span, visualized in the gray line at the top of the graph. From this graph, certain trends can be deduced. In general, over the ten years, male and female degree completions have increased, thus the total number of engineering degrees is increasing as well. Near the end of the time span, this number levels out and decreases slightly. The second axis shows the male to female ratio in engineering for those years. As expected, it starts high (male-dominated) and decreases steadily. While this visualization conveys the important trends in the data, there are ways to improve the overall effect and avoid misleading representations.

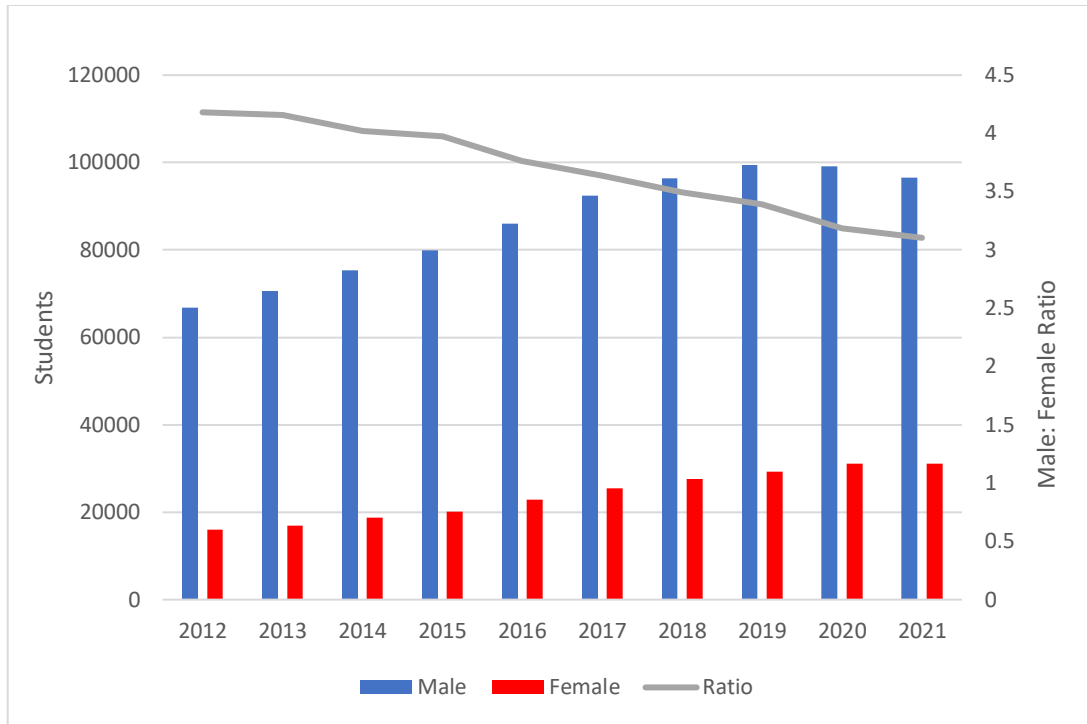


Figure 1: Number of male and female engineering students graduating from US universities per year 2012-2021, Secondary axis: Ratio of male to female engineering students per year 2012-2021 [7]

Figure 2 shows the same data as Figure 1 in a more unique way such that accurate trends can be understood while increasing the readability of the figure. Rather than showing a bar for male and female bachelor completions, one bar is shown representing the total bachelor completions for that year. The trend of increasing degree completions followed by a leveling out and slight decrease is still visible. Figure 2 represents the ratio of male to female degree completions with a color map. A high ratio of 5:1 is represented in red and the ideal ratio of 1:1 is represented in green. This method of visualization also shows that the male to female ratio is decreasing steadily, but it makes it more obvious that the ratio is changing slowly.

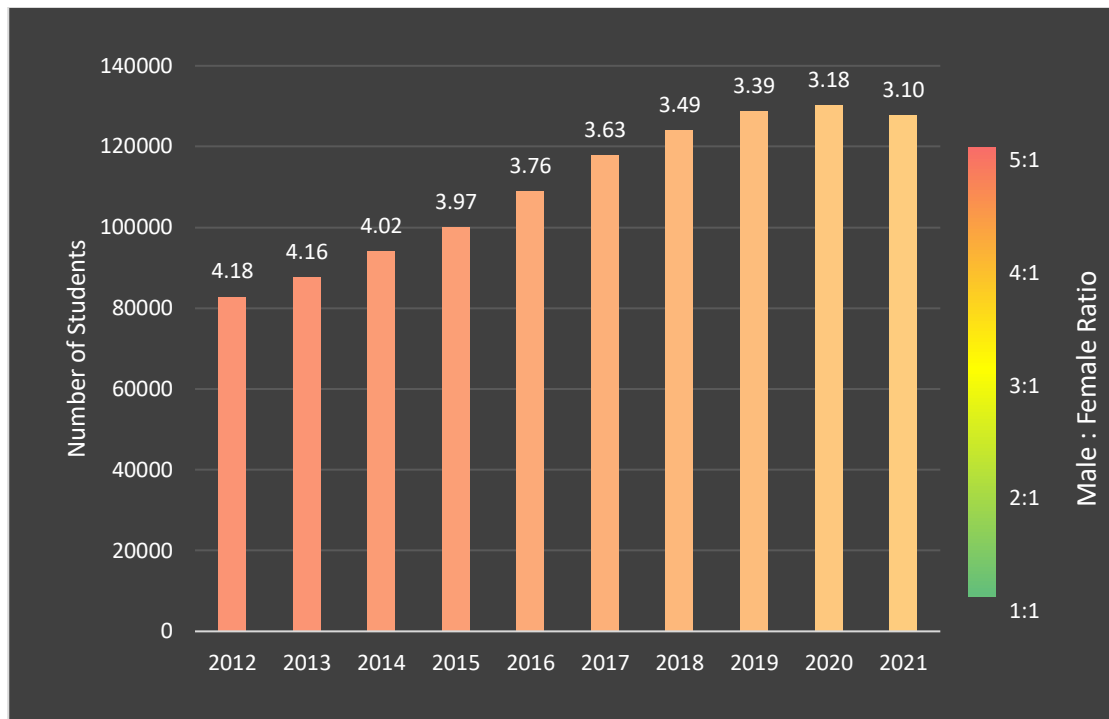


Figure 2: Total number of engineering bachelor completions in US universities per year 2012-2021, Heatmap: Ratio of male to female engineering bachelor completions per year 2012-2021, labels above bars describe ratio of male to female students

In terms of design elements, brighter and less basic colors both attract and keep the reader's attention compared to Figure 1. The blue, red, and gray dominant colors of Figure 1 do not stand out as much because the reader's eyes are more familiar with these hues [8]. Furthermore, with less objects to look at, the figure requires less mental effort from the viewer. Figure 1 shows three data sets as object: the number of degree completions by males, the number of degree completions by females, and the male to female ratio of degree completions. On the other hand, Figure 2 only shows one data set as an object, total degree completions, and one as a color gradient, male to female ratio. This method is less work for the viewer because both data sets occupy the same space so they can be more easily understood. Furthermore, redundancy is used to accommodate color blindness. The dataset describing the male to female ratio of engineering students is shown as both a color map and a text label.

By using color rather than datapoints or a line to show another dimension of data, the slow trend of the changing sex ratio in engineering education is more clearly seen. This effect is amplified when compared to a projection of the ratio extended fifty years based on a Gaussian distribution. Gaussian distribution was chosen because it fit the S-curve requirements of asymptotic population growth that this model should follow and had the highest R^2 value of 0.9998. Figure 3 provides a useful perspective between the first two figures. Figure 1 shows the male: female sex ratio in engineering reducing steadily and significantly towards an even split. Figure 2 represents the same data but uses a color map to visualize the sex ratio. This method makes it more obvious that the sex ratio has decreased slightly over the ten-year span with much progress to go. The second method highlights that in 2021 the ratio did not even reach 3:1, or 25% female representation. This analysis reveals that different conclusions can be drawn from each graph,

but it does not deduce which is a more accurate representation of the truth. A prediction is needed to see how gender ratios in engineering may change in the future. Figure 3 shows that engineering bachelor completions will reach 1:1 around 2070. Fifty years from now is not a practical or acceptable goal for gender representation equality in engineering. Thus, Figure 2 is a more effective medium for conveying this persistent and slowly changing problem.

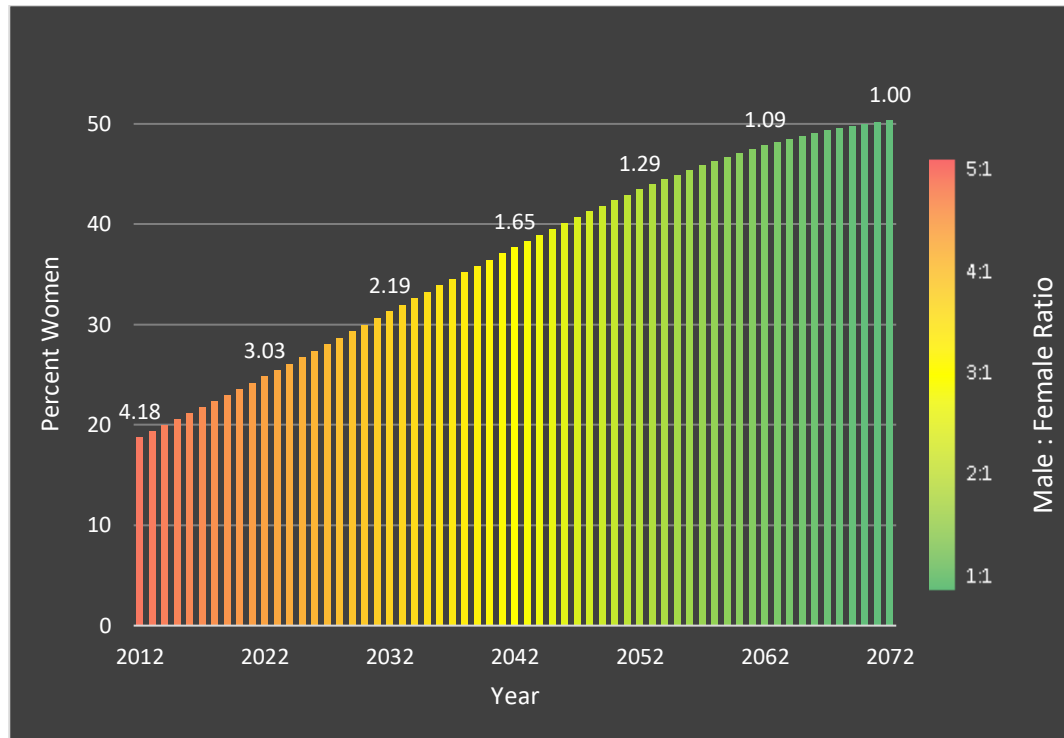


Figure 3: Fifty-year projection of male to female ratio of engineering students completing bachelor's degrees in US universities 2012-2072 according to Gaussian distribution, data labels above bars describe male to female ratio

Case Study 2: LGBTQ Devaluation in Engineering

While diversity in engineering is beneficial to creative innovation, there is still prevalent marginalization of lesbian, gay, bisexual, transgender, and queer (LGBTQ) students. In 2018, Cech et al. published a paper documenting the experience of LGBTQ students in engineering programs in comparison with non-LGBTQ students. The authors developed a large data set by surveying 1,729 students on the extent to which they agreed with various statements about their engineering experience (e.g., “accepted by students in department”, “treated as equally skilled student”) [9]. Of these students, 141 identified as LGBTQ, about 8.16% of the surveyed population. While the paper presented novel and impactful data, it did not include any graphs, charts, or other forms of visualization. This detracted from the potential impact that the data itself has on the reader. Visualizations help the reader process the information so that they can understand the issue being investigated. This is vital when it comes to LGBTQ belonging in engineering programs because students need to feel safe and accepted to take full advantage of their education [10]. Ignoring this problem could have long term effects on the mental health of LGBTQ students while also depriving the world of fully educated engineers. Figure 4 below shows a traditional visualization to help understand this data.

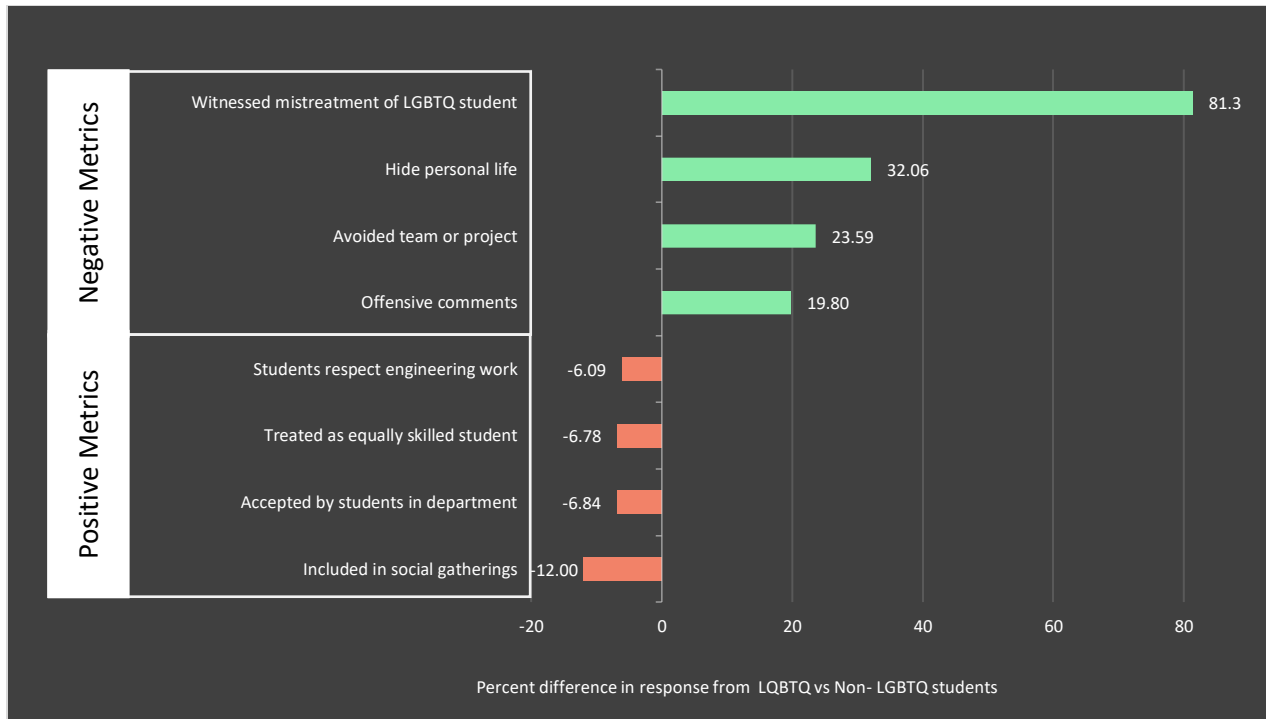


Figure 4: Difference in response between LGBTQ and non-LGBTQ students for eight statements about their experience in engineering programs (n = 1729)

The figure shows the difference in response by LGBTQ students and non-LGBTQ students to a survey about acceptance and belonging in engineering programs. It displays the response to four positive and four negative phrases. The obvious trend is that in all cases LGBTQ students are more likely to agree with negative metrics, such as feeling as though they needed to hide their personal life, more than non-LGBTQ students. Conversely, LGBTQ students are more likely to disagree with positive metrics, such as feeling other students respect their engineering work. Another important conclusion from the figure is that LGBTQ students are much more likely to notice mistreatments of another LGBTQ student than a non-LGBTQ student is. Visually, the figure does not draw attention to support the trends, The metrics on the left blend together so there is little differentiation between positive and negative metrics. The original data such as the average response for each group is missing which detracts from the impact of difference between the groups. Figure 5 below shows an alternative way to represent this data.

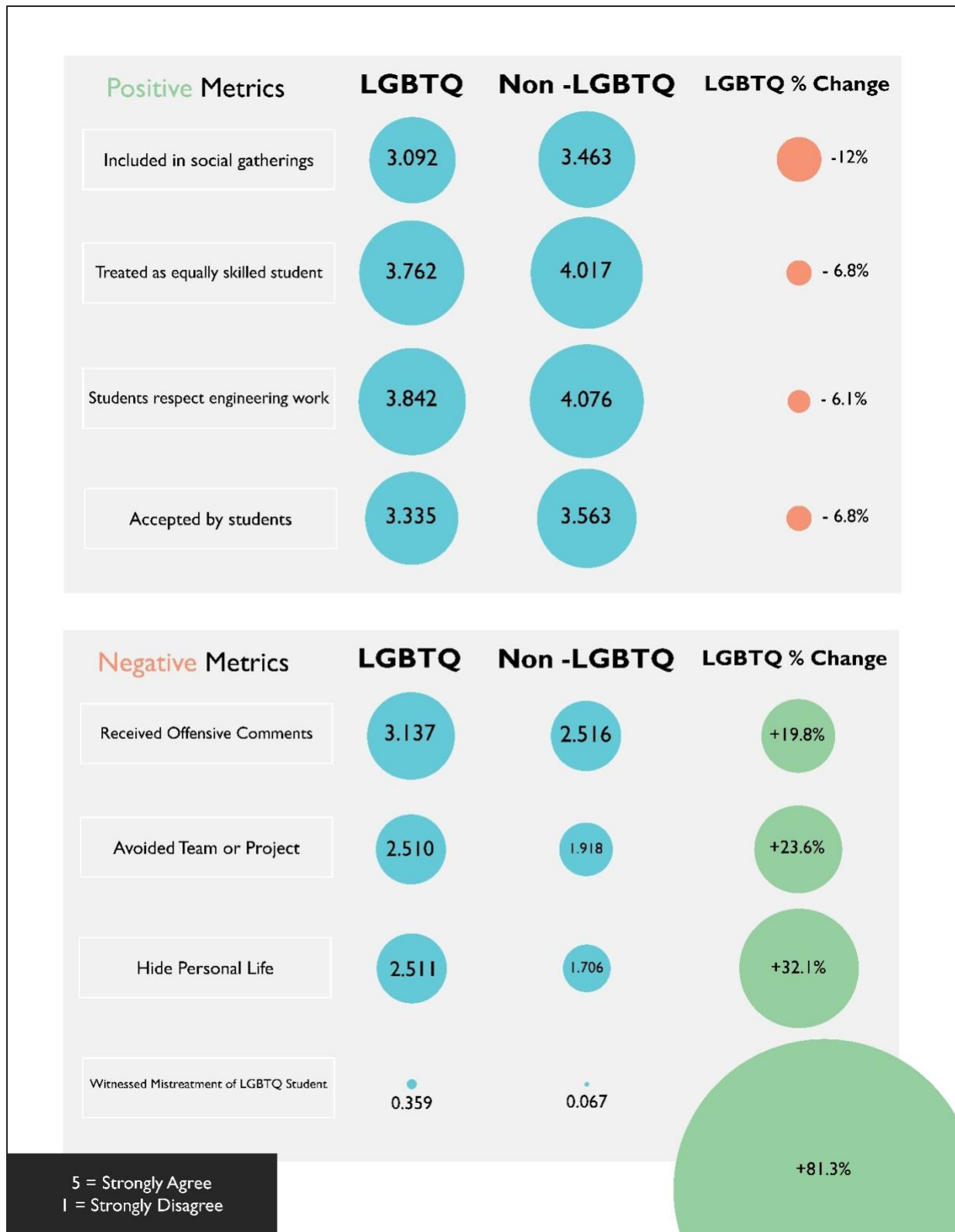


Figure 5: Infographic showing devaluation of LGBTQ students in engineering programs (N = 1729, 141 of whom identify as LGBTQ)

Figure 5 presents the same data as Figure 4 but uses an infographic as the medium. This creates a more visually appealing and understandable visualization. The number shown in the bubble is the average response of the surveyed students in which a 1 means strongly disagree and 5 means strongly agree. The general trend of the data is preserved. LGBTQ students are more likely to agree with the negative metrics about their experience and less likely to agree with positive metrics about their experience. Furthermore, the infographic adds the average response for LGBTQ and non-LGBTQ students, providing necessary context to the data. For example, the size of the top bar in Figure 4 denoted percent difference response to the phrase “witnessed mistreatment of LGBTQ student” is much larger than all the other bars. This large size difference may suggest that LGBTQ harassment is rampant in engineering programs. With the added context in Figure 5, it is clear that very few students both LGBTQ and non-LGBTQ have witnessed the mistreatment of an LGBTQ student. Rather, the data shows that LGBTQ students are more likely to notice this mistreatment than non-LGBTQ students. The added context in Figure 5 is important to interpreting the data in an accurate way. The infographic is a good medium for this because it is not restricted by the limits of a graph or chart. Things can be rearranged in a freeform layout so that it creates a cohesive story.

The infographic uses the size of objects to show comparisons between data points. This is the same as a bar chart, but the infographic utilizes bubbles for novelty. The bubbles fit the freeform layout of an infographic, having a more impactful effect than bars on a chart. Figure 5 also excels at visually separating the positive and negative phrases in a way that Figure 4 does not. They are physically separated in groupings and denoted with color labels for positive and negative metrics. This use of color complements the color that shows the positive and negative percent difference in response from LGBTQ and non-LGBTQ students. Overall, Figure 5 is much more appealing to look at. Neat elements, aesthetic colors, and clear patterns draw and hold the viewers’ attention. This is a key factor in improving understandability and the impact that the data has on the viewer.

Case Study 3: Engineering Education Systemogram

Visualizations are useful tools even when they are not displaying numerical data. They can still convert complicated topics into easily digestible formats. Systemograms are a type of infographic used to map out complex systems [11]. They display an array of key elements within the system, connected by lines that represent a variety of relationships. Given its usefulness in understanding intricate systems, it should be helpful in mapping the engineering education process. A huge number of factors affect the education of new engineers. From elementary school to graduate school, students are exposed to STEM curriculum, experiential learning, career development, and other external factors that contribute to them becoming an engineer. Having a systemogram that compiles this information could be used by students, teachers, professors, and administrators to refine the system for everyone’s benefit. The systemogram of the engineering education system is shown below in Figure 6.

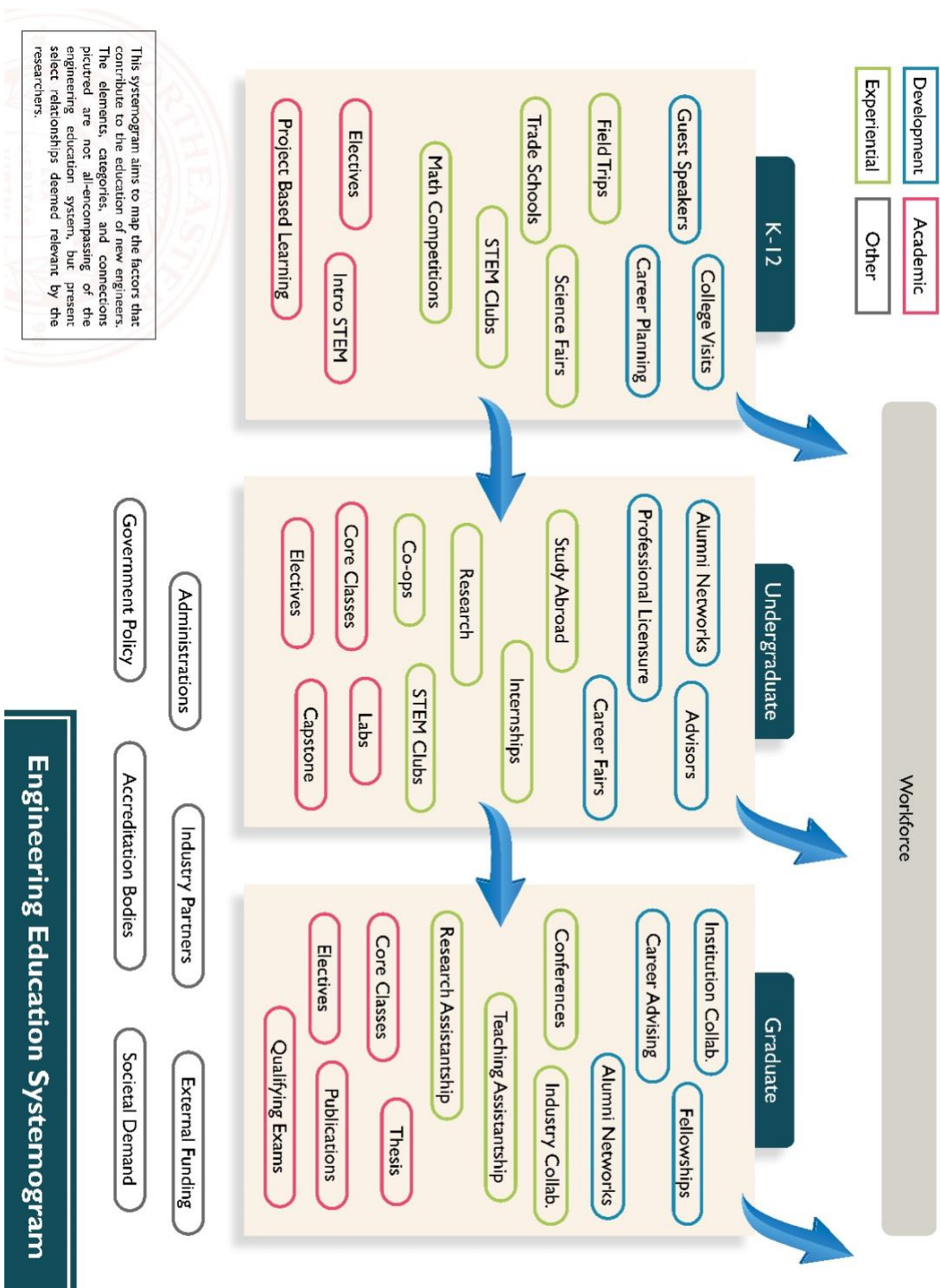


Figure 6: Systemogram of student flow through engineering education system

The figure above uses multiple visual elements to create an understandable and creative representation of data. First, the design clearly demonstrates flow to mimic the path of students through the engineering education system. Arrows show these possible paths. For example, after each stage of engineering education (K-12, undergraduate, graduate), there is a path to join the workforce which is shown via the blue arrows. The figure uses color to group the factors. This increases the understandability of the visualization. At first glance, there is an obvious grouping of education level, but the colors associated with the individual factors highlight trends across the groupings. The education levels stand out due to their location and size on the page, as well as the use of shadows to give the impression that it is off the page. Finally, the use of repetition improves the visual appeal and impact of the data. The shape of the elements serves as another style of grouping which allows the user to scan the figure much quicker. Furthermore, the repetition shows which elements are important. The individual factors are the smallest thing on the page, but the repetitive pattern draws the reader's attention. The use of visual design elements makes this systemogram easy to read and understand, making it a useful tool for the field of engineering education. While traditional systemograms have lines showing relationships between elements, these were omitted to further enhance the readability of the figure.

Limitations

There are several limitations to this project and enhancing data visualizations in general. To start, the labels used in the case studies, such as “sex” and “LGBTQ” mirrored what was used in the databases from which the information was drawn. While this ensures we are not misrepresenting the data that was collected, it limits us from using other terminology that might be more inclusive or representative of people's lived experience. We recognize that gender, like many other identities, is a continuum and, while categorizing can be helpful for representing data, it can also be exclusionary.

In addition, it would be irresponsible to talk about data visualizations without considering accessibility. While developing creative data visualizations, it is important to keep in mind best practices in accessible design such as using high contrast colors and alt text for digital works. This paper is just a starting point for exploring more compelling data visualizations. More work needs to be done to develop these for a variety of potential audiences.

Conclusion

The presented case studies explore the critical role of creative data visualization in enhancing the understanding and impact of various aspects of engineering education. It is important and beneficial to look beyond traditional data representation methods and towards more innovative, visually appealing, and creative approaches. The first case study addressed the issue of gender disparity in engineering. Use of color as a representation of data grabs attention and increases readability. The second case study investigates the marginalization of LGBTQ students in engineering education via an infographic. Finally, the third case study introduces the idea of a systemogram for visualizing the complex engineering education system. In each case, a creative approach offers a more compelling narrative and facilitates a better understanding of these challenges in engineering education programs.

The findings from this work are highly relevant to engineering education. With data being a driving factor in engineering decision making, it is crucial that future engineers understand how to properly present data and how data visualization can affect its interpretation. The researchers

envision the results of this work being incorporated into educational materials. Curricula could focus on how design elements such as color, form, value, contrast, and balance can be used in data representation. Interdisciplinary projects could combine creative data visualization with engineering concepts to bridge this gap. Finally, future work could compile a repository of visualization recommendations and their effects on viewer understanding. While creative data visualization is widely important, it is particularly relevant in engineering education. This field often deals with simplifying complex topics, enhancing engagement with curricula, and building communication skills. Thus, these creative data visualization case studies could be used in engineering education curricula and research.

A repository of creative visualization techniques is one key avenue of future research. This work shows promising results and highlights the importance of understanding visualization techniques. First, the testing of these visualizations needs to be expanded to a larger audience. Each visualization could be evaluated by human subjects on a quantitative scale in categories such as emotional response, impact, and clarity. In this kind of study, researchers could better understand the data interpretation of diverse audiences such as those with visual impairments. With a grasp on how understanding changes for people with disabilities, researchers could fully explore other forms of communication. This work was limited to print media visualizations. An extensive line of research could be done on videos, animations, and physical representations of data. Finally, work could be done to understand student learning outcomes. After incorporating this work into curricula, researchers could assess retention of information and application of visualization concepts on students' broader learning goals.

While creative data visualization can alter perception of data, it is important to distinguish the difference between this and manipulating data to show intended outcomes. Creative visualization adds design elements that amplify patterns that are already present. This is evident in case studies 1 and 2 in which the traditional visualization shows the same trends as the unique ones. The creative methods only increase readability to convey the trends of the data. Creative visualization is not akin to misrepresenting data. In conclusion, this paper underscores the significance of creative data visualization as a means to enhance the effectiveness of conveying complex information and critical issues in engineering education. Visual media is vital for properly understanding numerical and non-numerical data sets. The research demonstrates the potential for creative data visualization to drive change, evoke emotions, and inspire action in the field of engineering education and beyond.

References

- [1] K. Healy, *Data visualization: A practical introduction*. Princeton, NJ: Princeton University Press, 2018.
- [2] P. Fox and J. Hendler, "Changing the Equation on Scientific Data Visualization," *Science*, vol. 331, no. 6018. American Association for the Advancement of Science (AAAS), pp. 705–708, Feb. 11, 2011. doi: 10.1126/science.1197654.
- [3] Q. Li, "Data visualization as creative art practice," *Visual Communication*, vol. 17, no. 3. SAGE Publications, pp. 299–312, Apr. 17, 2018. doi: 10.1177/1470357218768202.
- [4] W. E. B. (William E. B. Du Bois), "A series of statistical charts illustrating the condition of the descendants of former African slaves now in residence in the United States of America," Library of Congress, <https://www.loc.gov/pictures/item/2013650370/> Accessed Jul. 24, 2024.
- [5] National Science Board. Science and Engineering Indicators 2016. Arlington, VA: National Science Foundation (NSB-2016-1).
- [6] L. Landivar, "Disparities in STEM Employment by Sex, Race, and Hispanic Origin", United States Census Bureau. Report Number ACS-24, 2013. Accessed Jul. 12, 2023. Available: <https://www2.census.gov/library/publications/2013/acs/acs-24.pdf>
- [7] U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 2022, Accessed from Data USA: Engineering. <https://datausa.io/profile/cip/engineering>
- [8] C. G. Healey, "Choosing effective colours for data visualization," *Proceedings of Seventh Annual IEEE Visualization 1996*, San Francisco, CA, USA, 1996, pp. 263–270, doi: 10.1109/VISUAL.1996.568118.
- [9] E. A. Cech and W. R. Rothwell, "LGBTQ Inequality in Engineering Education," *Journal of Engineering Education*, vol. 107, no. 4. Wiley, pp. 583–610, Oct. 2018. doi: 10.1002/jee.20239.
- [10] B. Kutsyruba, D. A. Klinger, and A. Hussain, "Relationships among school climate, school safety, and student achievement and well-being: a review of the literature," *Review of Education*, vol. 3, no. 2. Wiley, pp. 103–135, Jun. 2015. doi: 10.1002/rev3.3043.
- [11] D. Blair, J. T. Boardman, and B. J. Sauser, "Communicating strategic intent with systemigrams: Application to the network-enabled challenge," *Systems Engineering*, vol. 10, no. 4. Wiley, pp. 309–322, Sep. 19, 2007. doi: 10.1002/sys.20079.
- [12] S. Haroz, R. Kosara, and S. L. Franconeri, "ISOTYPE Visualization," *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Apr. 18, 2015. doi: 10.1145/2702123.2702275.
- [13] Y. Wang *et al.*, "An Emotional Response to the Value of Visualization," in *IEEE Computer Graphics and Applications*, vol. 39, no. 5, pp. 8–17, 1 Sept.-Oct. 2019, doi: 10.1109/MCG.2019.2923483.