Susan Vanderplas
Statistics Department
susan.vanderplas@unl.edu

What Do You See? Perception, Algorithms, and Statistical Graphics

Statistical graphics and models are powerful tools to summarize data and support human decision making; however, studies of chart perception are a patchwork of different methods and conflicting results. Chart design guidelines are often based on opinion, not empirical study, rendering many scientific communications sub-optimal or ineffective. This is alarming: effective science communication is critical for cultivating public trust in the scientific process and ensuring that decision makers accurately interpret supporting information. Addressing these challenges, the PI's long-term career goal is to examine statistical graphics with the goal of *helping people use data more effectively*, and to apply this research to educate and inspire a new generation of scientists while building public trust in science.

Toward this vision, the overall **research goal** of this CAREER program is to advance understanding of the use, design, and perception of data visualizations. Three research objectives (ROs) support this goal. RO1: Create a framework for comprehensive graphical testing across multiple levels of user engagement. RO2: Empirically validate common chart design guidelines, measuring the impact of design on task performance. RO3: Assess the impact of measurement methods on experiments evaluating statistical graphics.

Integrated with these research efforts, the overall **education goal** is to leverage visualization research to motivate statistical learning and improve data-driven decision making in society. Three education objectives (EOs) address this goal. EO1: Develop and implement experiential learning activities in statistical graphics for undergraduate introductory statistics courses. EO2: Create graduate course modules for K-12 educators that connect ongoing research to engaging, hands-on classroom activities for teaching statistics, math, and science. EO3: Improve the penetration of visualization research beyond academia by incorporating summaries of empirical studies in resources used by data scientists, industry analysts, and researchers in STEM disciplines.

Intellectual Merit This work will expand our understanding of graphical communication by empirically and systematically examining chart design with a focus on comprehensive task-based testing, producing nuanced guidelines for chart design based on intended use and accessibility considerations. It is often difficult or even misleading to generalize the results of graphics studies because small experimental design and implementation decisions have a large effect on the results, yet vary across otherwise similar studies. This project will assess the role of these decisions on empirical user studies, with the dual goals of providing contextual data to interpret results from historical studies and informing the design of future studies. Incorporating experiential learning and graphics research into statistics courses and STEM education in K-12 will advance knowledge of pedagogy and motivate student interest and engagement in STEM classes. Finally, collaborating with online visualization resources will provide opportunities to identify obstacles to adoption of best practice guidelines among practitioners.

Broader Impacts The proposed research will provide a solid foundation for future visualization research by developing, validating, and implementing new methods in a common framework that can be easily reused. The general public will benefit from better visual science communication resulting from nuanced, empirically driven data visualization research. In addition, collaboration with a popular data visualization website will bridge the gap between academic research and implementation, using summaries of empirical research curated by undergraduates to support and augment chart selection tools. Another collaboration with Nebraska Math and Science Summer Institutes will engage K-12 educators with graduate coursework that connects current research in statistics and data visualization with fun activities that reinforce critical concepts such as variability, experimentation, and sampling. A similar undergraduate experiential learning initiative will incorporate graphics research and reflective writing to introductory statistics courses, reinforcing material and connecting it to ongoing research. If successful, these activities will introduce and reinforce STEM concepts through fun, nonthreatening, and even artistic activities that minimize math anxiety while supporting student learning and increasing general scientific literacy. Data visualizations are broadly useful across the population, providing insight, supporting decision making, and encouraging exploration; as a result, investments in graphics research and education have broad, wide, and deep potential impacts for aspiring and practicing scientists as well as the general public.