

Susan VanderPlas

802 17th St.

Auburn, NE 68305

☎ (515) 509-6613

✉ svanderplas@gmail.com

Eric Smith, Ph.D.

CMDA Program Leader

Statistics Department

Virginia Tech

Blacksburg, VA

November 5, 2014

Dear Dr. Smith,

I write to apply for the Computational Modeling and Data Analytics faculty position at Virginia Tech University. I am a PhD candidate in the statistics department at Iowa State University, and am defending my dissertation in February 2015. The CMDA program elegantly provides the skill set for modern data analytics, balancing mathematical and statistical training with programming and applied courses, mirroring my own academic journey from mathematics (and cognitive psychology) to bioinformatics to statistics.

My doctoral research focuses on human perception of statistical graphics, with the ultimate goal of designing graphics to effectively communicate statistical results. Research on graphical perception is conducted in many different fields, including psychology, statistics, human-computer interaction, and communication. In each of these fields, though, the research focuses either on extremely low-level effects, or very specific problems that are not applicable to most statistical graphics. Integrating an understanding of human perception and methodology from cognitive psychology and human-computer interaction in an interdisciplinary approach to this problem, I am exploring several facets of human perception that meaningfully impact the way we interact with statistical graphics. My dissertation covers perceptual distortions that impact variance estimation along nonlinear trend lines (the "sine illusion"), the influence of visual ability on graphical inference, and which features of a graph are most visually compelling. Two experiments examining the sine illusion were presented at JSM in 2013 and 2014, and one article has been accepted for publication in JCGS; another is nearly ready for submission. A paper detailing the relationship between visual ability and graphical inference is being prepared for publication as well, and I expect that the final study will be ready for publication review by March 2015.

Additionally, I collaborate with researchers in bioinformatics, genetics, and engineering, assisting with statistical analyses as well as data visualization. My masters' research, a collaboration between bioinformatics and materials engineering, presented a nonparametric algorithm designed to mimic human perception of peaks in spectroscopy data with higher resolution than manual peak identification methods. Currently, I am working with soybean geneticists at the USDA to explore, analyze, and visualize populations of soybean sequence data; the massive amount of data are particularly challenging to visualize, as even heavily summarized data can overload the human visual system easily.

In the future, I plan to investigate the perception of interactive graphics, creating guidelines for

interactive graphics optimized for the human visual system. Visualizations which respond to user attention dynamically and incorporate motion through animation or other transitions have the potential to more intuitively communicate experimental results. The additional complexity of such graphics may reduce the amount of information that can be encoded, due to the increased demands on working memory and attention. As web-based interactive visualizations become more common, it is important that statisticians design graphics that are not only visually attractive, but also clearly communicate the overarching message in perceptually appropriate ways.

In addition to my research, I have experience teaching both computational and introductory statistics classes. For several years, I have co-taught workshops on R programming as a resource for the Iowa State community (students, researchers, and local businesses), introducing the language and presenting advanced topics including data visualization, formatting and arranging data for analysis, linear models. Extending this series, I designed workshops using new software packages, such as `knitr` for reproducible research and Rstudio's `Shiny` package for creating interactive web applets. I also have experience teaching undergraduate and graduate introductory statistics lab courses for engineering, bioinformatics, social science, and business students. In both settings, I utilize frequent examples which allowed students to independently apply the course material in a guided setting, reinforcing students' understanding of the material and providing opportunities for self-assessment and feedback.

The interdisciplinary nature of the CMDA program is likely to attract a group of students who have technical skills but often need to communicate their results to audiences with non-technical or disparate backgrounds. As I specialize in communicating such results using statistical graphics, I would be interested in developing or teaching a class which focuses on strategies for communication of technical results in a non-technical manner, including the use of graphics and charts to effectively convey information.

Blacksburg is a beautiful area which I have been acquainted with from birth, as my father received his Ph.D. from the statistics department in 1989. Virginia Tech and the CMDA program provide a unique learning and research environment, and I would be honored to be considered for this position. If there are any additional materials I can provide, please feel free to contact me directly or look through my research and current projects on github. Thank you for your consideration, and I look forward to hearing from you soon.

Yours,

Susan VanderPlas