Susan VanderPlas

Curriculum Vitae

802 17th St.
Auburn, NE 68305
515-509-6613
srvanderplas@gmail.com
srvanderplas.com
srvanderplas

Education

2011–15 **Doctor of Philosophy in Statistics**, *Iowa State University*.

2009–2011 **Master of Science in Statistics**, *Iowa State University*.

2005–2009 **Bachelor of Science**, *Texas A&M University*.
Major: Psychology and Applied Mathematical Sciences (Statistics), Minor: Neuroscience

Professional Experience

2015 **Statistical Analyst**, *Nebraska Public Power District*.
Conducted analyses to improve NPPD's data-driven decision making, including analysis of safety, profitability, and reliability data.

- Worked with IT and Strategic Management to develop a plan for analytics/data science maturity at NPPD.
- Designed a mentoring program to train individuals as embedded data scientists to increase the ability of NPPD to effectively utilize data.
- Modeled employee turnover to identify individuals likely to retire or resign.
- Established automated statistical monitoring of plant conditions, department turnover, and human performance errors.
- Predicted likely direction of tornadoes based on location and topological factors to establish the risk of tornado guided missle debris damaging critical equipment.
- Evaluated the risk of climate fluctuations on operational readiness.
- Identified site conditions statistically associated with water accumulation in radiation detectors at a nuclear plant.
- Improved engineering margin in thermal limits management in a nuclear reactor core.

2015 **Postdoc**, *Iowa State University*, Ames, IA.
Office of the Vice President for Research

- Evaluated faculty funding start-up packages to explore links between start-up funding and future productivity.
- Explored natural variation and underlying trends in grant receipts across Iowa State over a 20 year period.

2014 **Consultant**.
Developed web applications, interactive data displays, and statistical analyses for clients including the Iowa Soybean Association and Iowa State USDA Extension office.
Example 1: Nitrogen Deficiency in Corn, Example 2: Crop Yield Forecast

Research Interests

STATISTICAL GRAPHICS AND VISUALIZATION

- Visual Inference
- Human Perception
- Interactive Graphics
- Graphics for "Big" Data

DATA SCIENCE

- Data Mining
- Pattern Recognition

Publications

Submitted (2017) Heike Hofmann and **Susan VanderPlas**. All of this has happened before. all of this will happen again: Data science. *Journal of Computational and Graphical Statistics*, Submitted (2017).

2017 **Susan VanderPlas** and Heike Hofmann. Clusters beat trend!? testing feature hierarchy in statistical graphics. *Journal of Computational and Graphical Statistics*, 26(2):231–242, 2017.

Lindsay Rutter, **Susan VanderPlas**, Dianne Cook, and Michelle Graham. ggeanealogy: An R package for visualizing genealogical data. *Journal of Statistical Software*, 2017.

2016 **Susan VanderPlas** and Heike Hofmann. Spatial reasoning and data displays. *IEEE Transactions on Visualization and Computer Graphics*, 2016.

2015 **Susan VanderPlas** and Heike Hofmann. Signs of the sine illusion - why we need to care. *Journal of Computational and Graphical Statistics*, 24(4):1170–1190, 2015.

2010 Fadi Towfic, **Susan VanderPlas**, Casey A Oliver, Oliver Couture, Christopher K Tuggle, M Heather West Greenlee, and Vasant Honavar. Detection of gene orthology from gene co-expression and protein interaction networks. *BMC bioinformatics*, 11(Suppl 3):S7, 2010.

2009 Rachel Hull, Heather Bortfeld, and **Susan Koons**. Near-infrared spectroscopy and cortical responses to speech production. *The open neuroimaging journal*, 3:26, 2009.

In Progress **Framed Plots** User Study of the estimation accuracy of pie charts and mosaic charts with and without frames indicating unaligned segments of the population.

Context Mediated Graph Perception User Study of the influence of contextual information and expertise on the perception of polar plots used to display information in a compass-like setting.

Teaching Experience

Teaching Assistant

Spring 2013 — **Statistical Methods for Research**, *Iowa State University Stat Dept.*, Ames, IA.
Held office hours and graded labs and tests for Stat 401, a class composed primarily of graduate engineering students.

2012–2013 — **Introduction to Business Statistics II**, *Iowa State University Stat Dept.*, Ames, IA.
Taught undergraduate business students statistical methods and use of JMP statistical software. Responsibilities included holding office hours and evening help sessions, developing lab materials, managing the course website on Blackboard, and grading labs, homework, and tests.

2011 — **Statistical Methods for Research**, *Iowa State University Stat Dept.*, Ames, IA.
Taught graduate social science students statistical methods and use of SAS statistical software. Responsibilities included teaching lab sessions, creating lab materials, holding office hours and grading homework and lab materials.

2011 — **Empirical Methods for Comp. Sci.**, *Iowa State University Stat Dept.*, Ames, IA.
Held office hours and graded homework for Stat 430, a class composed of graduate bioinformatics and computer science students.

Other

2017 — **Business Intelligence Embedded Agent Program**, *Nebraska Public Power District*, Columbus, NE.
Designed a program to mentor NPPD employees, providing instruction in data science and opportunities to apply new skills within the company. One-on-one and group mentoring sessions were used to create a sense of community and to reinforce skills learned through online courses.

2017 — **R Course Instructor**, *Nebraska Public Power District*, Columbus, NE.
Taught a 3-day internal course on using R for data analysis, including use of ggplot2, dplyr, tidyr, knitr, and stringr.

2013–2014 — **R Course Instructor**, *Iowa State University Stat Dept.*, Ames, IA.
Designed and conducted workshops to teach R skills to the members of the university and local business community. Workshop topics included an introduction to R, ggplot2, data management with plyr, reshape2, and stringr, package development, document creation with knitr, linear models, and creating web applets with Shiny.

Presentations

2017 — **A Bayesian Approach to Visual Inference**, *JSM Contributed Session*.

2016 — **Clusters Beat Trend!? Testing Feature Hierarchy in Statistical Graphics**, *JSM Contributed Session*.

2015 — **Visual Aptitude and Statistical Graphics**, *InfoVis*.

2015 — **Animint: Interactive Web-Based Animations Using Ggplot2's Grammar of Graphics**, *JSM Invited Session*.

2015 — **Animint: Interactive, Web-Ready Graphics with R**, *Great Plains R User Group*.

2014 **The curse of three dimensions: Why your brain is lying to you**, *JSM Student Paper Award Session*.

2014 **Do You See What I See? Using Shiny for User Testing**, *JSM Panel*.

2013 **Signs of the Sine Illusion – why we need to care**, *JSM Contributed Session*.

Service

2017 **Graphics Section Representative to the Council of Sections**, *ASA*.

Theses

Dissertation

Title *The Perception of Statistical Graphics*

Committee Dr. Heike Hofmann (Chair), Dr. Dianne Cook, Dr. Sarah Nusser, Dr. Max Morris, Dr. Erin McDonald, Dr. Stephen Gilbert

Abstract Research on statistical graphics and visualization generally focuses on new types of graphics, new software to create graphics, interactivity, and usability studies. Our ability to interpret and use statistical graphics hinges on the interface between the graph itself and the brain that perceives and interprets it, and there is substantially less research on the interplay between graph, eye, brain, and mind than is sufficient to understand the nature of these relationships. This dissertation further explores the interplay between a static graph, the translation of that graph from paper to mental representation (the journey from eye to brain), and the mental processes that operate on that graph once it is transferred into memory (mind). Understanding the perception of statistical graphics will allow researchers to create more effective graphs which produce fewer distortions and viewer errors while reducing the cognitive load necessary to understand the information presented in the graph.

MS Creative Component

Title *Nonparametric statistical analysis of Atom Probe Tomography spectra*

Committee Dr. Alyson Wilson (Chair), Dr. Alicia Carriquiry, Dr. Krishna Rajan

Research Projects

Perception of Statistical Graphics

2015 **Independent Research**, Auburn, NE.

Designed and analyzed experiments to understand human perception of statistical graphics and optimized graphics to clearly communicate statistical results.

- Hierarchy of Graphical Features: Which features of statistical graphics dominate the perceptual experience? Do colored points matter more than trend lines? (Paper)
- Reproducibility of Plots in the 1870 Statistical Atlas (Paper)
- Bayesian Analysis of Statistical Lineups (Working Project)

2012–2015

PhD Research, *Iowa State University*, Ames, IA.

Designed and analyzed experiments to understand human perception of statistical graphics and optimized graphics to clearly communicate statistical results.

- The Sine Illusion in Statistical Graphics: How does this common illusion effect the information we take in from graphs?
  - Won the ASA Student Paper Award (2014) for the Graphics Section (Paper)
  - Created Shiny applets to demonstrate the illusion and test it's effect.
- Statistical Graphics and Visual Aptitude: How are spatial reasoning abilities related to the ability to read statistical graphics? (Paper)
- Hierarchy of Graphical Features: Which features of statistical graphics dominate the perceptual experience? Do colored points matter more than trend lines? (Paper)

Visualization of Genetics Data

2013
2015

Research Assistant, USDA Soybean Genome Project, *Iowa State University*, Ames, IA.

- Analyzed large quantities of soybean genetics data to identify inheritance, important genes, single nucleotide polymorphisms, and copy number variation.
- Created interactive applets presenting the data and appropriate graphics designed to encourage exploration of the results by biologists.
- Assembled a database of known soybean parantage to facilitate further research, and created an interactive applet to display the lineage of any variety in the database.

Advisor: Dr. Dianne Cook, Dr. Michelle Graham

Software Development

2013
2014

Statistics Teaching Applets, *Iowa State University*, Ames, IA.

Created and redesigned web-based applets to teach statistical techniques interactively. Applets covered topics such as the method of least squares, ANOVA, k-means, regression diagnostics, and t-tests.

2013
2015

Animint Developer, *R Project*, Google Summer of Code.

Worked to develop the `animint` package for R to translate `ggplot2` into d3 interactive JavaScript graphics. Participated in the project in 2013, adding support for all ggplot2 geoms as well as most scales and axes. Returned to serve as a mentor for the project in 2014 and 2015.

Materials Informatics

2010
2011

M.S. Research, *Iowa State University*, Ames, IA.

Worked with materials scientists and engineers to develop and implement nonparametric methods for automatic peak detection in mass spectroscopy data. Fit systems of differential equations to spectroscopy data based on theoretical concepts from quantum physics to facilitate inference about the atomic structure of a material.

Other

2012

Research Assistant, Iowa Dept. of Transportation, *Iowa State University*, Ames, IA.

Developed a hierarchical Bayesian model to determine the effectiveness of road interventions on traffic accidents and fatalities. Discovered a previously unknown error in the data used in prior analyses using exploratory techniques, and developed a method to compensate for the missing data.

2009

Research Rotations in Bioinformatics, *Iowa State University*, Ames, IA.

Explored applications of the EM algorithm to next-generation sequencing data error detection and modeled the relationship between age and fertility in reptiles. Each project lasted about 6 weeks; rotations were structured to allow new students to explore several facets of bioinformatics, and included wet-lab and computational experiences.

2009

NSF Research Experience for Undergraduates, *Iowa State University*, Ames, IA.

Worked with biologists and bioinformaticians to compare homologous gene expression in humans, pigs, and mice.

2008

NSF Research Experience for Undergraduates, *University of Nebraska*, Lincoln, NE.

Created a mathematical model describing electrical impulse transmission and decay along neurons with varying states of myelination.

References

Heike Hofmann

Professor
Department of Statistics
Iowa State University
2413 Snedecor Hall
Ames, IA 50011
✉ hofmann@iastate.edu
☎ 515-294-8948

Dianne Cook

Professor
Department of Econometrics and Business Statistics
Monash University
E762A Menzies Building
20 Chancellors Walk
Clayton, VIC 3800
✉ dicook@monash.edu
☎ +61 (0)3990 52608

Michelle Graham

Scientist
United States Dept of Agriculture
Iowa State University
1575 Agronomy Bldg
Ames, IA 50011
✉ michelle.graham@ars.usda.gov
☎ 515-294-3626

Alyson Wilson

Professor
Department of Statistics
North Carolina State University
2311 Stinson Drive
Campus Box 8203
Raleigh, NC 27695-8203
✉ alyson_wilson@ncsu.edu
☎ 919-515-1901