Sara A. Stoudt, PhD

Contact Information	Department of Mathematics, Bucknell University Lewisburg, PA http://sastoudt.github.io/ http://github.com/sastoudt	724-464-3179 sas072@bucknell.edu @sastoudt

Research Interests	applied and computational statistics in ecology, environmental science, and metrology, teaching the communication of statistics, data journalism

Education

University of California, Berkeley, Berkeley, CA

Ph.D., Statistics, Fall 2015 - Summer 2020
Advisors: Will Fithian (Department of Statistics) and Perry de Valpine (Department of Environmental Science, Policy and Management)
National Physical Science Consortium Fellow
National Science Foundation Research Traineeship: Data Sciences for the 21st Century, Environment and Society

Smith College, Northampton, MA

B.A., Mathematics and Statistics, 2015

Dissertation Research

- identifiability in species distribution and abundance models
- community metric estimation in latent factor joint species distribution models under model mis-specification
- statistics communication: *Communicating with Data: The Art of Writing for Data Science* (book written with Deborah Nolan, Oxford University Press, 2021)

Job Experience

Bucknell University Fall 2021 - currently
- tenure-track Assistant Professor in Mathematics Department

Berkeley Institute for Data Sciences Fellow Fall 2018 - Spring 2020
- Diversity and Inclusion Working Group
- Best Practices Working Group

Data Desk Intern - Los Angeles Times Summer 2019
- Census data aggregator
- Governor Newsom social media analysis

Data Science Intern - Farmers Business Network Summer 2018
- Crop yield prediction

Summer (Undergraduate/Graduate) Research Fellow Summers 2013-2017
- Statistical Engineering Division, National Institute of Standards and Technology
  - Measuring optical apertures for solar irradiance monitoring
  - Homogenization of surface temperature records
  - Errors-in-variables modeling for force calibrations
  - Interpolating atmospheric greenhouse gas fluxes
  - Evaluating the accuracy, consistency, and stability of measurements of the Planck constant

Teaching Experience

Assistant Professor, Mathematics Department, Bucknell University
- Statistics I Fall 2021
- Statistics II Spring 2022
- Statistical Inference Theory Fall 2021, Spring 2022

Lecturer, Statistical and Data Sciences Program, Smith College

- Communicating with Data (remote, almost 70 students) Fall 2020
- Introduction to Probability and Statistics (remote, about 45 students) Fall 2020, Spring 2021

Graduate Student Instructor (GSI), Statistics, UC Berkeley

- Co-developed and co-instructed new writing in statistics course (about 15 students) Fall 2017
- Co-developed and co-instructed new blogging for data science independent study (about 5 students) Spring 2018
- Outstanding GSI award

Miscellaneous Primary Instructor

- Data Storytelling Workshop (Correlation One's Women's Summit, 2020) (cancelled due to COVID-19)
- Introduction to Statistics in R Workshop (NICAR, 2020)
- Unleashing the power of biodiversity data with data science (Lewis and Clark workshop 2019, co-instructed with Ciera Martinez)

Miscellaneous Teaching Assistant

- Linear Regression Workshop (NICAR, 2020)
- Peer Data Review coach (Open News, 2019)
- R bootcamp (D-Lab/UC Berkeley Statistics Department, 2018)
- Software Carpentry (Data Science for the 21st Century Training Program, 2018)

Programming Languages

- Advanced: R (and Shiny)
- Experience With: Python, SQL, HTML, CSS, D3, JavaScript, bash, Matlab, Java, GIS, AMPL, Mathematica, NIMBLE, WinBUGS, LaTeX

Peer-Reviewed Publications

- Nolan, D., and **Stoudt, S.** "The Promise of Portfolios: Training Modern Data Scientist" *Harvard Data Science Review*, Issue 3.3, July 2021.
- **Stoudt, S.**, and Possolo, A. "Statistical and Computational Tools for Metrologists" Chapter for Advanced Mathematical and Computational Tools in Metrology and Testing XII, Series on Advances in Mathematics for Applied Sciences, Vol. 90, pp. 109-125.
- **Stoudt, S.***, Vasquez, V.*, and Martinez, C*. "Principles for data analysis workflows" (*equal authorship) *PLOS Computational Biology* Volume 17, Number 3, 2021.
- **Stoudt, S.**, Pintar, A., and Possolo, A. "Coverage Intervals" *Journal of Research of National Institute of Standards and Technology* Volume 126, Number 126004, March 2021
- **Stoudt, S.**, Pintar, A., and Possolo, A. "Uncertainty Evaluations from Small Datasets" *Metrologia* Volume 58, Number 1, January 2021
- Possolo, A., Schlamminger, S., **Stoudt, S.**, Pratt, J. R., and Williams, C. J. "Evaluation of the accuracy, consistency, and stability of measurements of the Planck constant used in the redefinition of the International System of Units" *Metrologia* Volume 55, Number 1, December 2017
- **Stoudt, S.** "Geostatistical Models for the Spatial Distribution of Uranium in the Continental United States" *Advances in Geocomputation: Geocomputation 2015 - The 13th International Conference* Springer Advances in Geographic Information Science, 2017, pp. 325-334.
- **Stoudt, S.**, Badian-Pessot, P., Mahop, B. N., Earley, E., Menter, J., Flores, Y., Williams, D., Zhang, W., Maharajan, L., Bao, Y., Rosenbauer, L., Nguyen, V., Mendiratta, V., and Tania, N. "Modeling Internet Traffic Generations Based on Individual Users and Activities for Telecommunication Applications" *American Journal of Undergraduate Research* Volume 13, Issue 3, August 2016, pp. 53-65.
- Bartel, T., Possolo, A., and **Stoudt, S**. "Force Calibrations using Errors-in-Variables Regression and Monte Carlo Uncertainty Evaluations" *Metrologia* Volume 53, Number

3, June 2016, pp. 965-980(16).

- **Stoudt, S.**, Cao, Y., Udwin, D., and Horton, N. J. "What Percent of the Continental US is Within One Mile of a Road?" *Statistics Education Web,* 2014.

Publications In Progress

- **Stoudt, S.***, Goldstein, B. R.*, and de Valpine, P. (*equal authorship) "Identifying Engaging Bird Species and Traits with Community Science Observations" (in press, preprint on bioRxiv)
- **Stoudt, S.** "Collaborative Writing Workflows in the Data-Driven Classroom: A Conversation Starter" (under review)
- **Stoudt, S.**, Scotina, A, and Luebke, K "Supporting Statistics and Data Science Education with learnr" (under review)
- Hong, J., **Stoudt, S.**, and de Valpine, P. "Fast maximum likelihood estimation for general hierarchical models" (under review)
- **Stoudt, S.** "What would happen if...? Statistical Thinking as Speculative Fiction" (invited for special issue on Prediction, Innovation, and Futures for Vector, under review)
- **Stoudt, S.**, de Valpine, P., and Fithian, W. "Semi-parametric Methods for Assessing Identifiability of Species Distribution and Abundance Models" (in preparation for invited special issue)
- Nolan, D., and **Stoudt, S.** "Storyboarding as an Exploratory Data Analysis Companion" (in preparation)
- Horst, A., Kross, S. and **Stoudt, S.** "Technology Enabled Creativity: 10 Creative Activities for Data Science Classes That You Can Include Now" (in preparation)
- **Stoudt, S.**, de Valpine, P. and Fithian, W. "Stress Testing Latent Factor Approaches to Joint Species Distribution Models" (in preparation)

Other Publications

- **Stoudt, S.** Review of "Statistical Learning from a Regression Perspective" by Richard A. Berk *MAA Reviews*, December 2021.
- Nolan, D. and **Stoudt, S.** "Captions: The Unsung Heroes of Data Communication" *Proceedings of IASE 2021*, September 2021.
- **Stoudt, S.** Review of "Design of Observational Studies" by Paul R. Rosenbaum *MAA Reviews*, July 2021.
- Goeva, A., Jones, P., **Stoudt, S.**, and Trisovic, A. "Recipes for Connector Courses from the Early-Career Board Kitchen" *Harvard Data Science Review (Invited Commentary)*, June 2021.
- Frost, S*., Goeva, A*., Pombra, J.*, Seaton, W*. **Stoudt, S*.**, Trisovic, A*, Wang, C.*, and Zucker, C.*."Kaleidoscopic Perspectives on Practicum-based Data Science Education" (*equal authorship) *Harvard Data Science Review (Invited Commentary)*, February 2021.
- Nolan, D., and **Stoudt, S.** "Reading to Write" *Significance*, December 2020.
- Goeva, A*., **Stoudt, S*.**, and Trisovic, A*. "Toward Reproducible and Extensible Research: from Values to Action" (*equal authorship) *Harvard Data Science Review (Invited Commentary)*, December 2020.
- Frost, S*., Goeva, A*., Seaton, W*. **Stoudt, S*.**, and Trisovic, A*. "Early-Career View on Data Science Challenges: Responsibility, Rigor, and Accessibility" (*equal authorship) *Harvard Data Science Review (Invited Commentary)*, September 2020.
- Sholler, D., **Stoudt, S.**, Kennedy, C., Hoces de la Guardia, F. Lanusse, F., Ram, K. Ottoboni, K., Stuart, M., Vareth, M., Varoquaux, N. , Barter, R. , Geiger, R. S., Peterson, S., and van der Walt, S. "Resistance to Adoption of Best Practices: a report from the Berkeley Institute for Data Science's Best Practices in Data Science Series"
- Geiger, R. S., Sholler, D., Culich, A., Martinez, C., Hoces de la Guardia, F., Lanusse, R., Ottoboni, K., Stuart, M., Vareth, M., Varoquaux, N., **Stoudt, S.**, and van der Walt, S. "Challenges of Doing Data-Intensive Research in Teams, Labs, and Groups:

Report from the BIDS Best Practices in Data Science Series"

- Geiger, R. S., DeMasi, O., Culich, A. Zoglauer, A., Das, D. Hoces de la Guardia, F., Ottoboni, K., Fenner, M., Varoquaux, N., Barter, R., Barnes, R., **Stoudt, S.**, Dorton, S., and van der Walt, S. "Best Practices for Fostering Diversity and Inclusion in Data Science: Report from the BIDS Best Practices in Data Science Series"
- Sholler, D., Das, D., Hoces de la Guardia, F., Hoffmann, C., Lanusse, F., Varoquaux, N., Garcia, R., Geiger, R. S., McDevitt, S., Peterson, S., and **Stoudt, S.** "Best Practices for Managing Turnover in Data Science Groups, Teams, and Labs"
- **Stoudt, S.**, Santana, L., and Baumer, B. "In Pursuit of Perfection: An Ensemble Method for Predicting March Madness Match-Up Probabilities" *JSM 2014 Proceedings*

Non-Academic Writing and Multimedia

- "The Origins of Ordinary Least Squares Assumptions: Some Are More Breakable Than Others," *AMS Feature Column*, March 2022.
- Data Science Education from the Lens of Communication, *Data Science Education Program Podcast*, December 2021.
- Data Science by Design (DSxD): A Community of Creators, *ADSA Community Blog*, with Martinez, C and Vasquez, V., November 2021.
- "Why Do We Plot Data" (Explainer Zine) *Harvard Data Science Review*, with Blumenthal, K., Goeva, A., **Stoudt, S.**, Trisovic, A., and Trisovic, P., September 2021.
- Under Ice Ice Pressure Baby, CAUSEweb Fun Collection, June 2021
- COVID-19 Case Fatality Rate Bias Visual Explainer with Aleksandrina Goeva, Harvard Data Science Review, February 2021.
- "Data Visualization Beyond the Screen", Northeast Big Data Innovation Hub Blog, January 2021.
- scripts for 15 episodes of Study Hall: Data Literacy (produced by Arizona State University and the Crash Course team at Complexly) (released weekly throughout Fall 2020)
- Communicating with Data newsletter, jointly written with students in my Fall '20 course.
- Ecology for the Masses Stats Corner
- Interviewed for Music Journalism Insider newsletter
- Just how does Kidz Bop censor songs? (The Pudding)
- You Know Karen (The Pudding)
- Berkeley Science Review
- Logic Magazine
- Tidy Tuesday and #rstats blog
- What does probability mean anyway?
- Fixed, mixed, and random effects

Awards

- RStudio Diversity Scholar 2020
- Berkeley Institute for Data Science Fellow 2018-2020
- Outstanding Graduate Student Instructor award 2018
- National Physical Science Consortium Fellow 2015-2018
- Data Sciences for the 21st Century: Environment and Society Graduate Training Program 2015-2017
- Gertrude M. Cox Scholarship 2015
- 2nd Place USPROC Undergraduate Research Project Competition 2015
- Best Poster Award: Geocomputation Conference 2015
- Elected to Mu Sigma Rho 2015
- Ann Kirsten Pokora Prize for excellence in mathematics at Smith College 2015
- Goldwater Scholar 2014
- Elected to Phi Beta Kappa Society 2014

- First Place: Statistics in Sports Undergraduate Research Competition at Joint Statistical Meetings 2014
- Honorable Mention Undergraduate CLASS Project Competition 2014
- Best in Show: Five College Data Fest 2014, 2015
- Elected to Sigma Xi (student) 2013
- Suzan Rose Benedict Prize for excellence in mathematics at Smith College 2013
- Smith College STRIDE (Student Research in Departments) Scholarship 2011-15

Grants

- awarded (co-PI) Code for Science and Society Virtual Event Grant, 2021.
- awarded (co-PI) Academic Data Science Alliance Career Development Network Seed Grant, 2021.
- awarded Curriculum Enhancement Grant from Smith College's Design Thinking Initiative to fund mailing maker kits to Communicating with Data students, Fall 2020.
- awarded grants from UC Berkeley Wellness Fund, UC Berkeley Chancellor's Advisory Committee on Student Services and Fees, and UC Berkeley Student Technology Fund Committee to fund the "Fostering diverse and inclusive data science at Berkeley" series at BIDS (2019)

Service

- MathBites Deputy Editor, Spring 2022-ongoing
- AMS Feature Column Contributor, Spring 2022-ongoing
- Undergraduate Statistics Project Competition judging, February 2022
- New York Times What's Going On In This Graph Moderation, January 2022
- Code review for "How do public officials make Land Bank decisions? Artificial Intelligence may seek patterns" October-December, 2021
- Bucknell Mathematics Department Curriculum Committee, Fall 2021-ongoing
- Bucknell Mathematics Department Seminar Committee, Fall 2021-ongoing
- TWS Ask a Biometrician Workshop (early-career consultant), November 1st, 2021
- Bucknell Mathematics Alumni Panel Moderator, September 2021
- co-organizer of Data Science by Design (DSxD) initiative, including Creator Conf in May 2021 and the Future of Data Science Anthology (early 2022)
- organizer of graduate school panel for Smith College Statistical and Data Sciences Program and Five College Statistics students, November 2020
- member of Harvard Data Science Review Early Career Board, Fall 2020-ongoing
- Women in Statistics and Data Science Rotating Twitter Curator, August 2020
- Co-organizer of Code and Coffee East Bay for the Bay Area Women in Machine Learning & Data Science Meetup group, Spring 2019 - Spring 2020
- Member of BIDS Diversity and Inclusion Working Group, Fall 2018 - Spring 2020
- We Are RLadies Rotating Twitter Curator, October 2018
- Co-president of the Statistics Graduate Student Association (SGSA), Fall 2017 - Spring 2018
- Co-organizer of UC Berkeley DataFest, Springs 2016 - 2018
- Co-organizer of SGSA Gender Issues Roundtable Discussion, Fall 2016
- Co-organizer of Statistics Graduate Student Association Diversity Discussion, Spring 2017
- Statistics Graduate Student Association Diversity Affairs Member, Fall 2016 - Spring 2017
- Graduate Student Volunteer: "Roadless America" Interactive Activity for Cal Day 2016 and 2017

Relevant Activities

- Blogdown website production for the Murmuration project
- Author of base R to dplyr vignette 2020
- Author of base R to stringr vignette 2019

- Developer of interactive visualization for the Wealth Tax Simulator 2019
- NCAR Graduate Workshop on Environmental Data Analytics (by application) 2016
- San Francisco Estuary Institute consulting 2016
- SAMSI International Temperature Initiative (by application) 2014

Presentations

- Revision Beyond the Copyedit: Encouraging Students to Reorganize, Restructure, and Refocus in Writing Assignments
  Talk, ICOTS, September 2022, with Deborah Nolan

- Audiences and Arguments: Teaching the Strategies of Effective Data Communicators
  Invited Talk, JSM, August 2022
- Wrong but Useful? Identifiability Regimes of Species Distribution and Abundance Models Under Model Mis-specification
  Talk, ISEC, June 2022, joint work with Perry de Valpine and Will Fithian

- Draw me a map v. paint me a picture: helping students know when to write about process and when to write about findings
  Talk, eCOTS, May 2022, with Deborah Nolan

- A Logical Lark? The Promise and Pitfalls of Community Science Data
  Invited Talk, University of Maryland Behavior, Ecology, Evolution, and Systematics Seminar, April 2022, joint work with Ben Goldstein, Perry de Valpine, and Will Fithian

- Going Analog: Emerging from the Weeds and Broadening our Audience
  Invited Talk, Iowa State University Graphics Group, March 2022

- Data Science by Design - Experimenting With a New Community of Practice For Data Creatives
  Talk, ADSA Annual Meeting, November 2021 (postponed - COVID, January 2022), joint work with Ciera Martinez and Valeri Vasquez

- All Species Distribution and Abundance Models Are Wrong, Which Are Useful?
  Talk, TWS Annual Conference, November 2021, joint work with Perry de Valpine and Will Fithian

- Building Responsible Data Science Workflows
  Panel, PyData Global, October 2021, with Valentin Danchev, Ben Marwick, Dr. Brandeis Marshall, Kirstie Whitaker, Thibault Lestang, and Yacine Jernite

- HDSR's First Zine
  Lightning Talk, Cut+Paste: Zines for Science Communication, October 2021, joint work with Kelly Blumenthal, Aleks Goeva, Ana Trisovic, and Pavle Trisovic

- Captions: The Unsung Heroes of Data Communication
  Talk, IASE, September 2021, joint work with Deborah Nolan

- Storyboarding as Part of the Process of Statistical Investigation
  Lightning Talk and Live Discussion, USCOTS Beyond Session, June 2021, with Deborah Nolan

- Communicating with Data: Practicing and Teaching
  Talk, Data Science for Social and Environmental Justice Virtual Writing and Research Development Group, June 2021, joint work with Deborah Nolan

- Data Communication
  Panelist, National Workshop on Data Science Education, UC Berkeley, June 2021, with Jeff Leek, Amelia McNamara, Trity Pourbahrami, and Deborah Nolan

- Creativity to Learn, Creativity to Teach
  Session Co-lead, Data Science by Design Creator Conf, May 2021, with Allison Horst and Sean Kross

- Generalized Additive Models: Allowing for some wiggle room in your models
  Invited Talk, Social Science Data Lab, Mannheim Centre for European Social Research at the University of Mannheim, Germany, March 2021

- From No Prereqs to Data Storytellers: Borrowing from the Journalist's Toolbox to Teach Students to Communicate with Data
  Talk, Computation+Journalism Conference, February 2021

- The Art of Writing for Data Science
  Guest Lecture, Communicating Data and Statistics course, Columbia Journalism School, February 2021

- Ideas Swap on Using Kits for Hands-on Making
  Panel, Smith College, January 2021

- Species Distribution and Abundance Models: The Good, The Bad, and The Not Identifiable
  Talk, Smith College Sigma Xi Lectures, November 2020

- Joint Species Distribution Models: Are they walking the walk or just talking the talk?
  Talk (Invited), Boston Women in Machine Learning and Data Science, October 2020

- Econometrics Meets Ecology
  Talk (Invited), USFCA students at Women and Diversity in Economics Club, October 2020

- Communicating with Data: How and where does it fit in the data science curriculum?
  Breakout Session Co-lead, Academic Data Science Alliance Annual Meeting, October 2020, with Deborah Nolan

- Principles for data-intensive research workflows: Guidance for the classroom and the computational laboratory
  Breakout Session Co-lead, Academic Data Science Alliance Annual Meeting, October 2020, with Valeri Vasquez and Ciera Martinez

- An overview of and lessons learned from hosting a data science workshop series for undergraduate students from under-represented backgrounds
  Co-presenter, Academic Data Science Alliance Annual Meeting, October 2020, with Orianna Demasi and Stacey Dorton

- Saying "Yes": A Data Memoir
  Talk (Invited), RLadies Amherst, April 2020

- Groove is in the Heart and the Data

Lightning Talk, RLadies San Francisco, December 2019

- Talking With the Public About Data Science
  Panelist, Moore-Sloan Data Science Environment Annual Summit, November 2019, with Joshua Tucker, Andrea Jones-Rooy, and moderator Meredith Broussard

- Diversity and Inclusion in Data Science
  Talk and Discussion Lead, Moore-Sloan Data Science Environment Annual Summit, November 2019, on behalf of BIDS Diversity and Inclusion Working Group

- Goodness-of-Fit Checks and Diagnostic Plots for Hierarchical Joint Species Distribution Models
  Poster, American Fisheries Society and The Wildlife Society Joint Annual Conference, September 2019, with Will Fithian and Perry de Valpine

- Species Distribution and Abundance Models: The Good, The Bad, and The Not Identifiable
  Talk and Poster, Berkeley Statistics Annual Research Symposium, UC Berkeley March 2019, with Will Fithian and Perry de Valpine

- Identifiability in the Wild: Econometrics Meets Ecology
  Talk, Third Annual Berkeley-Stanford Econometrics Jamboree, UC Berkeley, November 2018, with Will Fithian and Perry de Valpine

- Clarifying the Identifiability Controversy in Species Distribution Modeling
  Poster, Berkeley Statistics Annual Research Symposium, UC Berkeley March 2018, with Will Fithian and Perry de Valpine

- Sampling-Based Approaches to Maximum Likelihood Estimation for Latent Variable Models
  Poster, Berkeley Statistics Annual Research Symposium, UC Berkeley March 2017, with Johnny Hong and Perry de Valpine

- Interdisciplinary Graduate Education in Data Science: DS421 NRT
  Poster, Berkeley Institute for Data Science Data Science Faire, UC Berkeley May 2017, with DS421 Cohort 1

- Streamlining Climate Model Accessibility for Integration into Site-Specific Life Science Research
  Talk, Data Science for the 21st Century Annual Symposium, UC Berkeley May 2017, with Jenna Baughman

- Sampling-Based Approaches to Maximum Likelihood Estimation for Latent Variable Models
  Poster, BSTARS, UC Berkeley March 2017, with Johnny Hong and Perry de Valpine

- Uncertainty Quantification and Statistics
  Talk (Invited), NIST Presentation to SPIRAL students, July 2016 and July 2017

- Geostatistical Models for the Spatial Distribution of Uranium in the Continental United States
  Plenary Talk (Invited), First Electronic Undergraduate Statistics Research Conference, October 2015

- Internet Traffic Generation
  Talk (Invited), MAA Mathfest, August 2015, with Erika Earley, Yadira Flores, and Jordan Menter

- "Big Force" Calibrations: An Errors in Variables Approach
  Talk, Summer Undergraduate Research Fellow Colloquia, National Institute of Standards and Technology, August 2015, with Antonio Possolo and Tom Bartel

- Geostatistical Models for the Spatial Distribution of Uranium in the Continental United States
  Poster, Geocomputation, May 2015

- Correcting Temperature Records for Biases Unrelated to the Climate
  Talk, Summer Undergraduate Research Fellow Colloquia, National Institute of Standards and Technology, August 2014 (also given at WIMIN Conference in September 2014), with Antonio Possolo

- The Perfect Bracket: Machine Learning in NCAA Basketball
  SPEED poster and presentation, Joint Statistical Meetings, August 2014, with Loren Santana and Ben Baumer

- Taking a Closer Look at Learning: Factors Associated with Changes in Academic Performance During the Transition from Elementary to Middle School
  Poster, Women in Statistics Conference, May 2014, with Dana Hsu, Anna Rockower, and Katherine Halvorsen

- Measuring Optical Apertures for Solar Irradiance Monitoring
  Plenary Talk (Invited), Summer Undergraduate Research Fellow Colloquia, National Institute of Standards and Technology, August 2013, (also given at SMATH conference in September 2013) with Maritoni Litorja, and Antonio Possolo