Dana Paige Seidel

National Science Foundation Data Science Fellow & UC Berkeley PhD Candidate with 10 years experience in data wrangling and statistical modeling in R specifically for spatial and time series data. Contributor to the popular ggplot2 package & co-author of scales package. Available full-time as of summer 2019.

CONTACT

✉ dana.seidel@gmail.com
🌐 https://www.danaseidel.com
📞 612.232.6708   🐦 @dpseidel
🐙 dpseidel   📄 dpseidel   🔗 dana-seidel

EDUCATION

UNIVERSITY OF CALIFORNIA, BERKELEY
PHD IN ENVIRONMENTAL SCIENCE, POLICY, & MANAGEMENT
Expected June 2019 | Berkeley, CA

UNIVERSITY OF ALBERTA
MSc IN BIOLOGICAL SCIENCES
Aug 2014 | Edmonton, AB, Canada
Concentration in Ecology

CORNELL UNIVERSITY
BSc IN NATURAL RESOURCES
May 2011 | Ithaca, NY
Concentration in applied ecology
College of Agriculture and Life Sciences
Graduated *cum laude*

SKILLS

• Data Analysis • R • SQL
• Data Visualization • Biostatistics
• Regression Analysis • Open Science
• Geospatial Analysis (R, Python, ArcGIS)
• Experimental Design
• Open Source Development in R
• Version Control, Git/Github
• Testing & Continuous Integration
• Teaching • R Markdown • LaTeX
• Shiny • Tidyverse

AWARDS

• Data Science for the 21st Century: National Science Foundation Research Traineeship ($32K, 2015-2017)
• Outstanding Graduate Student Instructor Award, UC Berkeley 2018
• Letter of Commendation for Excellence in Teaching, Univ. of Alberta 2012

PUBLICATIONS

For full list of my publications, please see Google Scholar 🎓

EXPERIENCE

RSTUDIO | SOFTWARE DEVELOPMENT INTERN • TIDYVERSE TEAM
June 2018 – Aug 2018 | Remote
- Full-time developer of widely-used open source R packages ggplot2 🐙 (486K downloads/month) and scales 🐙
- Worked within large codebases to resolve issues, submit bug fixes, add new features, write tests, and update documentation to provide increased functionality for user-defined scales, themes, and aesthetic manipulation for data visualization in R.

GOOGLE | QUANTITATIVE ANALYST INTERN • GEODATA ANALYTICS TEAM
May 2016 – Aug 2016 | Mountain View, CA
- Explored impressions data for millions of features in the Google Maps database using Dremel, R
- Assessed average time to impressions maturity for novel business features using hazard models.
- Fit mixed effect regression models and evaluated model capacity for predicting impressions of novel business features based upon spatial relations with mature features.

RESEARCH

GRADUATE STUDENT RESEARCHER | UC BERKELEY
- Developed a R package to streamline best practices and automated report building for exploratory data analysis of animal movement data.
- Built a R Shiny application for the Government of Alberta to update spatial risk maps of infectious disease using rare-events logistic regression & spatial libraries in Python & R. 🌐🐙
- Built spatiotemporal models estimating disease prevalence and spread using a hierarchical Bayesian framework in WinBUGs and R.

SOUTHWEST ALBERTA MONTANE RESEARCH PROJECT 🌐
- Designed a database and analysis flow for foraging & movement data from 182 radio-collared elk - totaling 856K cleaned GPS samples.
- Implemented multivariate regression analyses in R, including mixed negative binomial regression and paired conditional logistic regression

TEACHING

VISITING INSTRUCTOR | HONG KONG UNIVERSITY
January 2018 | Movement Ecology in R Workshop
Co-developed and led an 8-day workshop teaching ecological methods of spatial and movement analysis in R & RStudio. 🌐🐙

DATA SCIENCE INSTRUCTOR | UC BERKELEY
Fall 2017 ESPM 157 • Spring 2018 ESPM 288
Taught upper division course teaching collaborative & reproducible data science including programming in R & RStudio, relational databases, data management, version control, remote data and APIs, cloud computing, & R Markdown. 🌐🐙

BIOSTATISTICS INSTRUCTOR | UC BERKELEY
Fall 2015 ESPM 173
Taught upper division biostatistics course including basic hypothesis testing, probability, experimental design, linear models, and programming in R & RStudio.