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.

CONTACT

dpseidel@berkeley.edu
https://www.danaseidel.com
612.232.6708 @dpseidel
dpseidel dpseidel dana-seidel

EDUCATION

UNIVERSITY OF CALIFORNIA, BERKELEY
PhD in Environmental Science, Policy, & Management
Expected May 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

• R • Data Analysis • Tidyverse
• geospatial analysis (R, Python, ArcGIS)
• Version Control, Git/Github
• Open Science • SQL
• Data Visualization • Biostatistics
• Regression Analysis
• Experimental Design • Teaching
• RMarkdown • Shiny
• Open Source Development in R
• Webpage development

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

- Participated in open source R development for ggplot2 (486K downloads/month) and scales packages
- Learned the internals of a large existing codebase to resolve issues, submit bug fixes, add new features, write tests, and update documentation to provide better functionality for user-defined scales, themes, and aesthetics manipulation during data visualization in R.

GOOGLE | Quantitative Analyst Intern • GeoDataAnalytics
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 in 4 cities across US and India based upon spatial relations with mature features

RESEARCH

Graduate Student Researcher | UC Berkeley

- Combined, cleaned, & analyzed time series data from 39 collared elephants, 2M cleaned GPS samples.
- Developed an R shiny application for updating spatial risk maps of infectious disease using rare events logistic regression & spatial libraries in Python and R.
- Built spatiotemporal models estimating disease prevalence and spread using a hierarchical Bayesian framework in WinBUGs and R.

Southwest Alberta Montane Research Project

- Designed 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.