

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.

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

AWARDS

- Data Science for the 21st Century: National Science Foundation Research Traineeship
- 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 📖

WORK EXPERIENCE

RSTUDIO | SOFTWARE DEVELOPMENT INTERN • TIDYVERSE TEAM

June 2018 – Aug 2018 | Remote

- Full-time developer of widely-used open source R packages for data visualization, ggplot2 📄 (486K downloads/month) and scales 📄
- Worked within large existing codebases to resolve issues, submit bug fixes, add new features, write unit tests, and update documentation to provide increased functionality for user-defined scales, themes, and aesthetic manipulation for data visualization in R
- Prepared major version release: scales 1.0.0

GOOGLE | QUANTITATIVE ANALYST INTERN • GEODATA ANALYTICS TEAM

May 2016 – Aug 2016 | Mountain View, CA

- Queried, manipulated, and analysed mobile and web generated user-impressions data for tens of millions of features in the Google Maps database using SQL-like syntax (Dremel) & R
- Built and evaluated predictive regression models for average time to “maturity” of user-interactions with novel business features
- Fit spatially-implicit mixed effect regression models and evaluated model capacity for predicting user-interaction with novel business features

PROJECT EXPERIENCE

GRADUATE STUDENT RESEARCHER | UC BERKELEY

- Developed a R package to streamline best practices and automated report building for exploratory data analysis of telemetry data 📄
- Analyzed multi-banded MODIS and LandSat imagery to understand environmental context of animal movement data
- Conducted principal components analysis and k-means clustering on results from agent-based simulations
- Used high-capacity distributed computing cluster to execute long-running analyses in parallel

MODELING CHRONIC WASTING DISEASE | ALBERTA FISH & WILDLIFE

- Built an R Shiny application with infrastructure to load raw data, extract environmental covariates from rasters, model and map spatial risk 📄 📄
- Used Python (e.g. GeoPandas, Fiona, Shapely, ArcPy) & R (e.g. sf, raster, velox, mapview) libraries to manipulate spatial raster and vector data for analysis
- Built spatiotemporal models estimating disease prevalence and spread using a hierarchical Bayesian framework in WinBUGs and R

SOUTHWEST ALBERTA MONTANE RESEARCH PROJECT | U ALBERTA 📄

- Designed relational database and analysis flow for field-collected foraging & movement data from 182 radio-collared elk
- Implemented multivariate regression analyses in R, including mixed negative binomial regression and paired conditional logistic regression

TEACHING EXPERIENCE

VISITING INSTRUCTOR | HONG KONG UNIVERSITY | JANUARY 2018 📄 📄

DATA SCIENCE INSTRUCTOR | UC BERKELEY | 2017-2018 📄 📄

BIostatistics INSTRUCTOR | UC BERKELEY | FALL 2015