

Megan Stachura

Mission-driven data scientist with over 10 years of experience managing, statistically analyzing, and visualizing data for decision making using Python, R, SQL, and AWS

MeganStachura@gmail.com
www.MeganStachura.com
www.github.com/mstachur

EXPERIENCE

Data Scientist - Four Peaks Environmental Science & Data Solutions, 2018 - Present

- Develop Python scripts and AWS pipelines to automate the transformation of raw data files into a standard format, correct data issues, extract useful data features, and produce summary metrics and statistical analyses to support client decision making
- Validate, manage, and query data in 4 MySQL and SQL Server databases
- Design and build interactive data dashboards and complex data visualizations using R and Python
- Project manager for up to 4 projects at a time, coordinating client and team communications and work to meet deliverable timelines and budgets
- Lead internal diversity committee

Research Associate, Policy Fellow, & Intern - U.S. National Marine Fisheries Service, 2009 & 2014 - 2018

- Developed a simulation model of recreational fisheries in R, integrating data from 36 sources and utilizing zero-inflated negative binomial regression
- Wrote R and R Markdown scripts to repeatably extract, clean, summarize, and analyze data and generate text, tables, and figures for regular reporting. Proactively identified projects that could benefit from these scripts and processes
- Implemented a logistic regression model to estimate the impacts of fishing methods on discarded fish survival rates
- Authored 16 web articles to communicate complex science and policy information to a general audience

Research Scientist & Graduate Fellow - University of Washington, 2010 - 2014

- Developed novel hypotheses, compiled and cleaned data, and applied numerous statistical modeling techniques in R (e.g., Bayesian hierarchical models, cluster analysis) to evaluate environmental influences on fish species, utilizing commonalities across species to gain statistical power
- Published 5 peer-reviewed papers and gave 8 formal presentations

SKILLS

Data Processing, Preparation & Management

Python (pandas, numpy), R (dplyr, R Markdown), SQL (MySQL, SQL Server), AWS (S3, Lambda, EC2, Elastic Beanstalk), Excel

Statistics & Machine Learning

Logistic regression, Bayesian hierarchical models (JAGS), zero-inflated negative binomial regression, bootstrapping, linear mixed models, cluster analysis, principal component analysis, dynamic factor analysis, deep learning (fastai, PyTorch)

Data Visualization

R (Shiny, ggplot2, base graphics), Python (matplotlib, seaborn), Tableau

Communication & Collaboration

Authored 15 scientific publications and 16 general audience articles; gave 18 formal presentations; Git (CodeCommit, GitHub)

EDUCATION

Master of Science in Aquatic & Fishery Sciences - University of Washington, 2013

Courses included R programming, data visualization, linear and non-linear regression, multivariate statistics, and Bayesian statistics

Bachelor of Science in Marine Science & Biology - University of Miami, 2010

Magna cum laude; minor in Mathematics