Megan Stachura

Data scientist with over ten years of experience preparing, managing, statistically analyzing, and visualizing data for decision making using Python, R, SQL, and AWS; passionate about contributing to a world where everyone can thrive

MeganStachura@gmail.com
www.MeganStachura.com

EXPERIENCE

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

- Develop Python scripts and Amazon Web Services pipelines to automate the transformation of raw data files into a standard format, identify and correct data issues, extract useful data features, and produce summary metrics and statistical analyses to support client decision making
- Design and implement both interactive and static data visualizations in Python and R (Shiny)
- Validate, manage, and query data in MySQL and SQL Server databases
- Project manager for up to 4 projects at a time, coordinating client and team communications and work to meet deliverable timelines

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

- Developed a simulation model of recreational fisheries in R, integrating economic and ecological data from 36 different sources and predicting fish catch using 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
- Wrote web content to communicate complex science and policy information to a general audience

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

- Applied numerous statistical modeling techniques in R to evaluate and predict environmental influences on fish species, utilizing commonalities across species to gain statistical power
- Reviewed literature, formulated novel hypotheses, researched and compiled relevant data, and identified and implemented appropriate statistical techniques

SKILLS

Data Processing, Preparation & Management

Python (pandas, numpy), R (dplyr, R Markdown), SQL (MySQL, SQL Server), Amazon Web Services (S3, Lambda, EC2, Elastic Beanstalk, CodeCommit), 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

Data Visualization

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

Communication/Collaboration

Authored 15 scientific publications and 16 general audience articles; gave 18 formal presentations; Git

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