

Megan Stachura

Social-impact driven analytics engineer with 15+ years of experience building scalable data models and pipelines on cloud platforms, statistically analyzing and visualizing data, and managing collaborative projects using Databricks, Python, SQL, Apache Spark, R, and AWS

mstachur@gmail.com

[LinkedIn](#)

[Personal Website](#)

[Google Scholar Publications](#)

WORK EXPERIENCE

Analytics Engineer – Fred Hutch Cancer Center, 2025 to Present

- Design, implement, and improve SQL, Python, and R data pipelines in Databricks, transforming Epic EHR into OMOP Common Data Model to accelerate research collaboration
- Use asset bundles to source control scalable resources and CI/CD, and define permissions for HIPAA compliance
- Implement quality validation utilizing DQX library
- Author markdown templates and skills to document, standardize, and enforce best practices in development

Data Management Consultant – 2024 to 2025

- Built databases and software to streamline data collection, management, and analysis for growing teams, using Python, SQL, AWS RDS, and Shiny

Analytics Engineer, Analyst – Remitly, 2021 to 2023

- Architected, built, maintained, and documented marketing data pipelines using Python, Spark, SQL, Scala, AWS Redshift and S3, implementing unit and quality testing
- Designed new and improved existing dimensional data models to enable scalability, optimization, and self-service
- Collaborated with stakeholders to identify need, evaluate proposals, and carry out agile, version-controlled projects
- Identified product improvement opportunities through Tableau and Periscope dashboards and A/B experiments

Senior Scientist, Scientist II & III – Four Peaks Environmental Science & Data Solutions, 2018 to 2021

- Architected, built, and maintained data infrastructure using Python, MySQL, SQL Server, and AWS S3 and Lambda
- Led projects studying and mitigating impacts of hydroelectric dams on fish passage, serving as project manager for up to 4 projects at a time with budgets over \$600K annually, authoring 5 reports of project findings
- Built data dashboards/visualization and conducted statistical analyses using Python, R, and ArcGIS

Research Associate, Policy Fellow – U.S. National Marine Fisheries Service, 2014 to 2018

- Developed tools to automate collection and summary of data for regular reporting using R; trained users
- Developed complex statistical and simulation models in R

SKILLS & TOOLS

Data Engineering & Platforms

Databricks (Declarative Automation Bundles, Unity Catalog, Lakeflow Jobs, CLI), Apache Spark, SQL, Python (PySpark, pandas, numpy), R, AWS (Redshift, S3, Lambda, EC2, RDS, DynamoDB), Scala, Jenkins, Azkaban

Statistics & Machine Learning

Logistic regression, Bayesian hierarchical models, bootstrapping, cluster analysis

Data Dashboards & Visualization

Python (Django), R (Shiny), Tableau, Periscope, Power BI, Looker, ArcGIS

Communication & Collaboration

Git (GitHub, CodeCommit), Jira, Trello, Confluence, Notion, Slack, Teams, WordPress, HTML

VOLUNTEERING

- Native habitat restoration steward on Seattle's Duwamish River since 2016
- Founding member and leader of the Seattle Donut Economics Coalition since 2023
- Supporting foster families through Seattle Angels since 2023

EDUCATION

Recent Training

Regenerative Leadership (Regenerators Academy, 2024); Regenerative Economics and Finance (The Capital Institute, 2023-2024); Epic Cogito (2026)

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

Used Bayesian and multivariate statistics to identify patterns across fish stocks

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

Magna cum laude; minor in Mathematics