Ava Hoffman

Fort Collins, CO

- 📱 (804) 687-7476
- 🌐 avahoffman.com
- ✉ avamariehoffman@gmail.com
- in /in/avahoffman
- ⌂ avahoffman

Skills

Languages

Python
SciKit-learn · NumPy · SciPy · Pandas · Statsmodels · Seaborn · Matplotlib · Bokeh · Gensim · Jupyter Notebook

Shell / command line
Git · SLURM · distributed computing · software compilation

R
RStan · RStudio · RMarkdown · leaps · lavaan · segmented · dplyr · reshape2 · bioconductor · sva · ggplot2 · ggtree · bayesplot · ggrepel · gridExtra · semPlot
bioinformatics: extensive, see GitHub

Databases

SQL (PostgreSQL)
Python interface · Amazon Web Services

Techniques

Statistics
hierarchical models · mixture models · multivariate statistics · dimensionality reduction · repeated measures · design matrices · hypothesis testing
primary tools: Rstan · JAGS · SAS

Supervised machine learning
Linear/nonlinear/logistic regression · linear discriminant analysis · random forest · time series analysis · structural equation modeling · validation, predictive check

Unsupervised machine learning
principal components analysis · k-means clustering · feature reduction

Other

LaTeX · **Markdown** · **Tableau**
some exposure to: **HTML / CSS** · **perl** · **Arduino**

Education

| 2018 | **PhD** · **Ecology** | Colorado State University, USA |
| 2012 | **BS** · **Biology** | University of Virginia, USA |

Experience

2018 — **Insight Data Science Fellow** — Remote Program

- Launched National Perks Project to improve visitor experience in crowded National Parks
- Leveraged large National Park Service, NOAA, & web-scraped datasets to forecast optimal time to visit parks using time series (FB Prophet and ARIMA models)
- Weighted user preferences for crowd level, plus minimum & maximum temperature with euclidean distance to personalize visitor experience
- Identified, defined, & structured data management with PostgreSQL & Amazon Web Services
- Built a customized web app using Git, Python, Flask, & Heroku providing visitors with an actionable recommendation for optimal visit time and access to further resources

2017 - 2018 — **USDA NIFA Predoctoral Fellow** — Colorado State University

- Designed Blue Grama Diversity Project to inform natural area stakeholders about genetic diversity in a foundational prairie grass, blue grama
- Pioneered genomic feature detection (sequencing) of a key grass species, discovering >9,000 genomic features to cluster populations & guide conservation from 15GB of data
- Modeled hierarchical linkages between genomics, populations, & plant appearance & developed analytical workflows using R and RStan
- Communicated genetic clusters in ggplot data visualizations
- Facilitated 6-member team collaboration for a large-scale project

2013 - 2017 — **PhD Researcher** — Colorado State University

- Discovered diversity within key prairie grasses in response to drought to guide management of grassland resources
- Optimized analytics pipeline for genes & contrasted fluctuation of >100,000 genes under different conditions using bash scripts
- Planned & implemented management of diverse data structures
- Quantified predictive relationships among plant, traits, & ecosystem by implementing trait feature reduction, Bayesian hierarchical models, path analysis, module clustering, literature data mining, linear regression, & meta-analysis in R, leading to domain innovation

Honors

| 2017 - 2018 | **Sustainability Leadership Fellow** | Colorado State University |
| 2016 - 2017 | **Vice President for Research Fellow** | Colorado State University |