Andrew J. Bibian

Richmond, CA • (805) 204-8149 • ajbibian@gmail.com • bibsian.github.io

SPECIALIZED SKILLS

• **Languages:** Python (4 years), R (7 years), SQL (4 years)
• **Analytics:** Bayesian Modeling, Regression, Numerical Simulations, Random Forest, NLP
• **Programming Tools:** PostgreSQL (4 years), MSSQL (1 year), AWS (4 years), Cluster Computing (2 years), Git (4 years), Bash (3 years)
• **Project Management Tools:** Agile (1 year), Jira (2 year)

EXPERIENCE

Data Scientist/Engineer   September 2018 – Present
CredibleMind - Sausalito, CA - Seed funded startup, name pending, 7 FT employees, 90% remote
• Develop, deploy, and monitor a deep learning text classifier for filtering incoming news content.
• Manage, monitor, and deploy python data pipelines with AWS and the Node.js Serverless framework.
• Develop python libraries for gathering data through external API's and web scraping (headless chrome).
• Develop ETL processes with AWS, python, and SQL for reporting on platform content, generating gap analyses, deriving aggregated metrics, and proposing new content to platform managers.

Data Science Fellow   June 2018 – September 2018
Insight Data Science - San Francisco, CA – 7-week post-doctoral training program (6% hold no PhD)
• Created SnapDogg, a web application for predicting Instagram 'Likes' from user's potential post.
• Engineered metrics from images, text, and timestamps to predict user 'Likes' and increase model accuracy by 40% from baseline.
• Deployed a random forest regression for predicting 'Likes' using AWS, Flask, and Python.

Data Analyst   March 2017 – April 2018
Wood Mackenzie - Houston, TX
• Collaborated with domain experts to develop analyses quantifying risk for drilling exploration.
• Developed a data pipeline to aggregate spatio-temporal data for integration with existing data products.
• Led weekly meetings on project timelines, priorities, and participated in Agile development.
• Led internal workshops for upskilling with Python and collaborated with colleagues to lean workflows (time savings of 64 hours per year).

Data Analyst   August 2015 – March 2017
Rice University - Houston, TX
• Managed a team of three implementing in house ETL tools for aggregating data.
• Developed protocols to sort, identify, standardize, and compile hundreds of independent time series datasets on population dynamics into one PostgreSQL repository.
• Designed, maintained, and updated database records, schemas, and handled database migrations on local and remote servers (Linux & AWS).

Graduate Researcher (National Science Foundation Fellow)   August 2012 – December 2015
Rice University - Houston, TX
• Numerically analyzed and parallel processed 10 GB of data on university clusters (R & bash scripts).
• Used power analyses to determine experimental sample size requirements.
• Developed novel hierarchical Bayesian models (linear and logistic regression) for analyses.
• Used stochastic simulations to assess the accuracy and precision of statistical methods.

EDUCATION

M.A. Ecology & Evolutionary Biology - Rice University   2015
B.S. Wildlife, Fish, and Conservation Biology - University of California, Davis   2010