

AARON JONES

13930 Edgewater Drive, Norwalk, CA 90650
acjones1987@gmail.com • (562) 881-5206 • <https://www.linkedin.com/in/aaron-c-jones/>

WORK EXPERIENCE

- Blackwood Seven A/S,** Los Angeles, CA
Data Scientist, Marketing Sciences Jan 2018 – present
Analyst, Marketing Sciences May 2017 – Jan 2018
- Construct marketing mix models in the form of Bayesian hierarchical regression models to monitor media performance long-term by calculating costs per acquisition, KPI lift, and overall business decomposition, and guide the allocation of tens of millions in marketing spend quarterly
 - Build baseline models to quantify the impact of national media campaigns using machine learning, stacking, and bootstrapping, regularly achieving sub-5% holdout MAPEs
 - Research and develop an application, with R Shiny UI, that loops through numerous machine learning models to find the optimal counterfactual prediction and calculate lift
 - Teach instructional seminars for nontechnical team members on causal impact, using NLP on social media data, and the capabilities and best practices of modeling
- Southern California Coastal Water Research Project** Costa Mesa, CA
Data Analyst, Department of BiogeoChemistry May 2016 – May 2017
- Perform analyses for scientific articles, project proposals, and technical reports
 - Leverage numerous techniques, including visualization, regression (linear, logistic), principal component analysis, and classification (random forests, gradient boosting)
 - Construct algorithms that test ecological models via probabilistic resampling and that use integration to aggregate environmental production rates
- California State University, Fullerton**
Teaching Associate, Department of Mathematics Aug 2015 – Jan 2017
- Teach Introductory Probability and Statistics, and Intermediate Algebra

INDEPENDENT CONSULTING

- Aquatic Bioassay and Consulting Laboratories** Ventura, CA
Statistical Consultant May 2017 – present
- Facilitate the analysis of sparse environmental data by the nearest neighbor hot deck imputation of missing data
 - Determine the major drivers of biological condition using variable importance measurements from random forest models built to predict quality indices from Los Angeles River and San Gabriel Watershed Monitoring Program data
 - Visualize high dimensional data using principal component analysis and non-metric multidimensional scaling to uncover the underlying similarity between observations
- Black Swan Data** Lake Forest, CA
Data Science Consultant Jan 2017 – May 2017
- Research, program in Python, and deliver an image processing algorithm, utilizing support vector machines, that identifies company names in social media images
 - Algorithm trained on 112,800 image dataset built using the open-source datasets Chars74K and CIFAR-10 (Kaggle)

TOOLS AND LANGUAGES

Languages: Python, R, SQL, bash
Tools: AWS (Redshift, S3), MySQL, scikit-learn, caret, Git, R Shiny, Jupyter, Tableau, LaTeX

EDUCATION

- California State University, Fullerton**
M.S. Statistics May 2017
- Benson Scholarship for Graduate Mathematics Students (2016)
 - Relevant courses: Statistical Computing, Bayesian Statistics, Multivariate Analysis, and Time Series Analysis/Forecasting
- M.A. Music May 2015
- University of California, Santa Barbara**
B.M. Music, High Honors Jun 2009

PROFESSIONAL DEVELOPMENT

- Online Courses**
Applied Machine Learning in Python, University of Michigan (Coursera) Oct 2017