Randi H Griffin, Ph.D.

Boston, MA

617-548-2608 | rgriff23@gmail.com | https://rgriff23.github.io | https://github.com/rgriff23 | www.linkedin.com/in/randigriffin

SKILLS

Languages – R (ggplot2, tidyverse, caret), Python (pandas, scikit-learn, matplotlib, seaborn, dash); prior experience with MySQL, HTML/CSS
Statistics - Generalized linear models, survival analysis, multivariate statistics, meta-analysis, hypothesis testing
Machine learning - Classification, regression, clustering, feature engineering, NLP, validation
Tools & techniques - git, Bash, R Markdown, Jupyter, web scraping, data visualization

WORK EXPERIENCE

Insight Data Science – *Fellow & Consultant for Datenight Babysitting App,* Boston, MA, Sep 2018 – Pres

- Developed predictive models for user subscriptions and bookings for the Datenight Babysitting App.
- Worked closely with CEO to compile and process data on users from a MySQL database.
- Generated features by geocoding 4K user addresses and linking them with socioeconomic and demographic data from the Canadian Census.

Duke University – *NSF Graduate Research Fellow,* Durham, NC, Sep 2013 – 2018

- Implemented multivariate GLMs to model 3D point clouds derived from 122 microCT scans of primate skulls to identify ecological predictors of skull morphology.
- Designed computer simulation studies to evaluate two new statistical methods for estimating ancestral states in evolutionary biology, demonstrating that the new methods perform far less effectively than several alternatives.
- Performed survival analysis on 10 years of longitudinal data to demonstrate substantial health costs of tapeworm infections in wild monkeys.
- Demonstrated fine-scale habitat segregation (< 20m) in mosquito communities using GLMMs and PCA.
- Instructed graduate students and faculty on data analysis at an annual workshop (2013-2015); lectured and led discussion groups for three undergraduate courses; and supervised one undergraduate thesis project.

Harvard University – *Research assistant,* Cambridge, MA, Sep 2011 – 2013

- Maintained and queried a MySQL database with ~20K records of parasites reported in wild mammals.
- Performed meta-analysis of 14 studies and 164 effect sizes, with results indicating that parasite infections in wild primates are not driven by human-caused habitat disturbance.
- Simulated pathogen transmission on social networks and identified network characteristics (clustering, centrality) that increase susceptibility to epidemic and endemic pathogens.

PROJECTS

'btw' R package – R wrapper for BayesTraits, an executable C++ program for fitting Bayesian phylogenetic models. https://github.com/rgriff23/btw

Web scraping Olympic history data - Scraped and wrangled data on 135k Olympians from www.sports-reference.com. This dataset has been downloaded >10k times (top 0.3%) on Kaggle as of Sep, 2018. https://github.com/rgriff23/Olympic_history

Twitterstorm analysis - Compiled data on 4.5k users and 5k tweets in a politically-charged Twitterstorm, then used social network and sentiment analysis to identify liberal and conservative clusters. https://github.com/rgriff23/Katie_Hinde_Twitter_storm_text_analysis

EDUCATION

Duke University, *PhD in Evolutionary Anthropology* — May 2018
Harvard University, *BA in Human Evolutionary Biology* — May 2010