

Randi H Griffin

I'm a Data Scientist with a background in marketing and the social / biological sciences. I'm passionate about crafting creative yet practical solutions to problems and committed to fostering inclusive and collaborative team environments.

rgriff23@gmail.com | <http://www.randigriffin.com/> | <https://github.com/rgriff23> | www.linkedin.com/in/randigriffin

SKILLS

Programming: Python (pandas, scikit-learn, matplotlib); R (ggplot2, tidyverse, caret), SQL

Statistics: Generalized linear models, survival analysis, time series analysis, network analysis, meta-analysis

Machine learning: Classification, regression, clustering, forecasting, feature engineering, NLP

Other tools & techniques: git, bash, Airflow, R Shiny, dash, Periscope, Google Data Studio

EDUCATION

Ph.D. in Biological Anthropology, Duke University

May 2018

B.A. in Human Evolutionary Biology, Harvard University

May 2010

PROFESSIONAL EXPERIENCE

KAYAK Software Corporation, *Data Scientist (Performance Marketing)* Cambridge, MA, Nov 2018 – Present

- Developed predictive model of keyword revenue-per-click to support search ads algorithms, using NLP to provide accuracy gains for long tail keywords, resulting in a 300% increase in traffic and revenue for a given cost.
- Developed a counterfactual experimentation (XP) framework using causal inference models to 1. Select optimal treatment and control markets for future XPs, and 2. Measure lift and statistical significance of completed XPs.
- Developed a 90-day LTV forecasting system and dashboard used for quarterly budgeting and target setting.
- Developed and maintained ETLs to support programmatic marketing algorithms and reporting tasks.

Stand Up America, *Data Consultant (contract role)*

Remote, Oct 2020 – Present

- Conducted deep analysis of digital marketing campaigns which earned an Expy Award from the Analyst Institute.
- Built Periscope dashboards to track Get Out The Vote initiatives for the 2020 presidential election.

Insight Data Science, *Data Science Consultant*

Boston, MA, Sep 2018 – Nov 2018

- Built a dash app for a babysitting service that automatically geocodes user addresses, links them with census data, and uses logistic regression to estimate the probability that new users will subscribe to the babysitting app.

Duke University, *NSF Graduate Research Fellow*

Durham, NC, Sep 2013 – 2018

- Used multivariate GLMs to identify ecological predictors of primate skull morphology based on CT scans.
- Conducted simulation studies to evaluate statistical methods for estimating evolutionary ancestral states.
- Wrangled 10 years of data and used survival models to quantify parasite-mediated mortality in wild primates.
- Demonstrated fine-scale habitat segregation in mosquito communities using GLMMs and PCA, providing a recommendation of <20 meters for the minimum resolution of spatial data in mosquito-borne disease models.

Harvard University, *Research assistant*

Cambridge, MA, Sep 2011 – 2013

- Performed formal meta-analysis of 14 published studies and 164 effect sizes, with results contradicting the popular claim that elevated parasite loads in wild animals are 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.

INDEPENDENT PROJECTS

Scraping Olympic history: Scraped data on 135k Olympians and shared analysis on Kaggle. The dataset has >61K downloads as of Jan, 2021. <https://www.kaggle.com/heesoo37/olympic-history-data-a-thorough-analysis>

Twitterstorm analysis: Used social network and sentiment analysis to identify liberal and conservative clusters in a Twitterstorm (4.5k users, 5k tweets). https://github.com/rgriff23/Katie_Hinde_Twitter_storm_text_analysis

StackOverflow survey: Won \$1000 Kaggle Award for analysis of attitudes towards inclusion and ethics in the programming community. <https://www.kaggle.com/heesoo37/stack-overflow-2018-survey-age-gender-sexuality>

OPEN SOURCE CONTRIBUTIONS

Parsons: This project provides ETL connectors to integrate NGPVAN with other data sources commonly used by progressive political organizations: <https://github.com/move-coop/parsons/commits?author=rgriff23>

'btw' R package: R wrapper for the BayesTrait phylogenetic modeling software. <https://github.com/rgriff23/btw>

UNIVERSITY TEACHING

Northeastern University, Lecturer (Masters in Analytics)

Boston, MA, Feb 2019 – Present

- **Capstone Course (2 semesters):** In each semester, I managed 6 teams of 5 students as they complete an analytics project for a sponsoring company.
- **Developed Surveys and Guidelines** to be used by all Analytics Capstone Courses to aid the formation of balanced project teams at the start of the semester and the collection of useful peer-feedback at the end of the semester.
- **Data Mining in R (3 quarters):** Developed original materials and received excellent teacher ratings.

PEER-REVIEWED PUBLICATIONS

Fox, S.D., **Griffin, R.H.**, Pachankis, J.E. 2020. Minority Stress, Social Integration, and the Mental Health Needs of LGBTQ Asylum Seekers in North America. *Social Science & Medicine*, 246, 112727.

Schneider-Crease, I.A., **Griffin, R.H.**, Gomery, M.A., Bergman, T.J., and J.C. Beehner. 2017. High mortality associated with parasitism in geladas (*Theropithecus gelada*) in the Simien Mountains National Park, Ethiopia. *American Journal of Primatology*, 79(9).

Schneider-Crease, I.A., **Griffin, R.H.**, Dorny, P., Noh, J.C., Handali, S., Chastain, H.M., Wilkins, P.P., Nunn, C.L., Snyder-Mackler, N., Beehner, J.C., and T.J. Bergman. 2017. Identifying wildlife reservoirs of neglected taeniid tapeworms: non-invasive diagnosis of endemic *Taenia serialis* infection in wild primates. *PLOS Neglected Tropical Diseases*, 11(7): p.e0005709.

Griffin, R.H., and G.S. Yapunich. 2017. A critical comment on the 'multiple variance Brownian motion' model of Smaers et al. (2016). *Biological Journal of the Linnean Society*, 121(1): 223-228.

Reiskind, M., **Griffin, R.H.**, Janairo, M.S., and K.A. Hopperstad. 2016. Mosquitoes of Field and Forest: The Scale of Habitat Segregation in a Diverse Mosquito Assemblage. *Medical & Veterinary Entomology*, 31(1): 44-54.

Griffin, R.H., and G.S. Yapuncich. 2015. The Independent Evolution method is not a viable phylogenetic comparative method. *PLoS ONE* 10(12): e0144147.

Coburn, R.A., **Griffin, R.H.**, & S.D. Smith. 2015. Genetic basis for a rare floral mutant in an Andean species of Solanaceae. *American Journal of Botany* 102(2): 264-272.

Young, H., **Griffin, R.**, Wood, C.L., and Nunn, C.L. 2013. Does habitat disturbance increase infectious disease risk for primates? *Ecology Letters*, 16(5): 656-663.

Cooper, N., **Griffin, R.**, Franz, M., Omatayo, M., and Nunn, C.L. 2012. Phylogenetic host specificity and understanding parasite sharing in primates. *Ecology Letters* 15(12): 1370-77. [Science Daily press release](#)

Griffin, R.H., Matthews, L.J., and Nunn, C.L. 2012. Evolutionary Disequilibrium and Activity Period in Primates: A Bayesian Phylogenetic Approach. *American Journal of Physical Anthropology* 147:409-416.

Griffin, R.H. and Nunn, C.L. 2011. Community structure and the spread of infectious disease in primate social networks. *Evolutionary Ecology* 26(4): 779-800.

ATHLETIC ACHIEVEMENT & COACHING

South Korean 2018 Olympic Team and Women's Ice Hockey National Team Player, 2015-2018.

<https://today.duke.edu/2018/03/duke-olympian-will-soon-defend-her-phd>

Harvard Women' Ice Hockey, 2006-2010. ECAC Student-Athlete of the Year Finalist, 2010.

USA Hockey Certified Coach. Completed 4 years of training and have coached youth teams from ages 12 to 19.