

# Randi H Griffin

Boston, MA

617-548-2608 | [rgriff23@gmail.com](mailto:rgriff23@gmail.com) | <https://rgriff23.github.io> | <https://github.com/rgriff23> | [www.linkedin.com/in/randigriffin](https://www.linkedin.com/in/randigriffin)

## SKILLS

**Languages:** Python (pandas, scikit-learn, matplotlib); R (ggplot2, tidyverse, caret); SQL; HTML/CSS

**Statistics:** Generalized linear models, survival analysis, multivariate statistics, meta-analysis, hypothesis testing

**Machine learning:** Classification, regression, clustering, feature engineering, NLP, dimension reduction

**Tools & techniques:** git, R Markdown, Jupyter, web scraping, data visualization, simulation, survey design

## EDUCATION

**Ph.D. in Evolutionary Anthropology,** Duke University

May 2018

**B.A. in Human Evolutionary Biology,** Harvard University

May 2010

## WORK EXPERIENCE

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

- Contributed production-grade Python and SQL code for ETLs and digital advertising bidding algorithms.
- Developed a rigorous counterfactual experimentation framework for evaluating ad bidding algorithms.
- Identified and corrected a flaw in our system for evaluating alternative forecasting methods, which were causing neutral changes to masquerade as up to 20% gains in accuracy.

**Northeastern University, Analytics Lecturer**

Boston, MA, Feb 2019 – Present

- Developed and taught a 6-week Masters-level course on data mining with R, receiving the highest possible teacher reviews (5/5 stars from 100% of students) in my first two sessions.

**Insight Data Science, Data Science Fellow**

Boston, MA, Sep 2018 – Nov 2018

- Developed a logistic regression pipeline in Python to predict which users will subscribe to a babysitting app.
- Generated novel features by geocoding 4K user addresses and linking them with geospatial census data.
- Built a dash app that allows the company to estimate the probability that new users will subscribe.

**Duke University, NSF Graduate Research Fellow**

Durham, NC, Sep 2013 – 2018

- Implemented multivariate GLMs in R to model the evolution of the primate skull in 3 dimensions, identifying diet, circadian rhythms, and mating systems as the major ecological drivers of skull shape.
- Designed computer simulations to evaluate statistical methods for estimating ancestral states in evolution, demonstrating that several newly developed methods are outperformed by existing simpler methods.
- Wrangled 10 years of longitudinal data and used survival models to provide the first demonstration that tapeworms reduce 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.
- Simulated pathogen transmission on social networks and identified network characteristics (clustering, centrality) that increase susceptibility to epidemic and endemic pathogens.
- Engaged in significant science outreach activities, including serving as a “Science Coach” in Duke’s BOOST program for encouraging K-12 students from underrepresented groups to pursue careers in STEM.

## SIDE PROJECTS

**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.

[github.com/rgriff23/Katie\\_Hinde\\_Twitter\\_storm\\_text\\_analysis](https://github.com/rgriff23/Katie_Hinde_Twitter_storm_text_analysis)

**Web scraping and EDA of Olympic history data:** Scraped and wrangled data on 135k Olympians from [www.sports-reference.com](http://www.sports-reference.com). This dataset has been downloaded >35k times (top 0.3%) on Kaggle as of Sep, 2019.

[github.com/rgriff23/Olympic\\_history](https://github.com/rgriff23/Olympic_history) [www.kaggle.com/heesoo37/olympic-history-data-a-thorough-analysis](https://www.kaggle.com/heesoo37/olympic-history-data-a-thorough-analysis)

**‘btw’ R package:** R wrapper for BayesTraits, an executable C++ program for fitting Bayesian phylogenetic models.

[github.com/rgriff23/btw](https://github.com/rgriff23/btw)

## PEER-REVIEWED PUBLICATIONS

- 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. Yapuncich. 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., Omotayo, 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

**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>

**USA Hockey Certified Coach.** Completed 4 years of training to attain the highest coaching certification provided by USA Hockey, and have coached youth teams ranging from the 12U to 19U age groups.