Randi H Griffin

I am a data scientist with strong people and project management skills. I'm passionate about crafting creative yet practical solutions to problems, communicating complex ideas clearly, and 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, SQL, R
Statistics: Generalized linear models, survival analysis, time series, network analysis, meta-analysis
Machine learning: Classification, regression, clustering, forecasting, feature engineering, NLP
Generative AI: Prompt engineering, commercial APIs, open-source models, automated testing and evaluation, red teaming
Other tools & techniques: git, bash, Airflow, streamlit, dash, flask, shiny, FastAPI, dashboard software (e.g., Tableau, PowerBI, Google Data Studio, Sisense, and more)

EDUCATION

Ph.D. in Evolutionary Anthropology, Duke University   May 2018
B.A. in Human Evolutionary Biology, Harvard University   May 2010

PROFESSIONAL EXPERIENCE

Boston Consulting Group, *Lead Data Scientist (BCG X)*   Boston, MA, Oct 2023 – Present
- *Client:* Led a team of 3 data scientists in delivering an ETL and web application to support Reliability Centered Maintenance optimization of mission-critical equipment for a U.S. Federal Agency.
- *Client and internal:* As part of the GenAI Red Team, led testing and evaluation projects that produced actionable insights for multiple GenAI use cases, including several internal HR chatbots, customer-facing sales chatbots, and a chatbot to support call-center workers at a U.S. state agency.
- *Internal:* Spearheaded ideation and development of ARTKIT, an open-source Python library for automated testing and evaluation of GenAI applications; managed a team of up to 12 software engineers and data scientists over a 5-month period; consistently received strong upward feedback for creating a healthy team environment, leading with vision and integrity, and prioritizing individual coaching, mentorship, and constructive feedback: https://github.com/BCG-X-Official/artkit

Boston Consulting Group, *Senior Data Scientist (BCG X)*   Boston, MA, Apr 2021 – Oct 2023
- *Internal:* Served as a Delivery Excellence coach, providing technical advice and risk mgmt support for ~20 client project teams, including code reviews and red teaming of high-risk applications
- *Client:* Anchored an Analytics Center of Excellence for a leading biopharmaceutical company for 8 months, supporting ~60 clinical trial contract proposals with potential revenues of $2.2B
- *Client:* Developed and productized ML models and an experimentation framework for a revenue management system at a major international airline, meeting project deadlines and requirements
- *Client:* Developed interpretable ML models and reporting dashboards to flag likely delays in material arrival times as part of a Supply Chain Control Tower for a major U.S. defense industry contractor

KAYAK Software Corporation, *Data Scientist (Marketing Analytics)*   Cambridge, MA, Nov 2018 – 2021
- Developed predictive model of keyword revenue-per-click to support search ads algorithms, with accuracy gains yielding in a 300% increase in traffic and revenue for a given cost.
- Developed a counterfactual experimentation (XP) framework using Causal Impact to select optimal treatment and control markets for future XPs, and to measure lift and significance of completed XPs.
- Developed a 90-day LTV forecasting system and dashboard used for budgeting and target setting.

- Developed and maintained ETLs to support marketing algorithms and reporting tasks.

Stand Up America, *Data Consultant (contract)*   Remote, Oct 2020 – Dec 2020
- Conducted analysis of 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 estimates the probability that new users will subscribe to the app.

Duke University, *NSF Graduate Research Fellow*   Durham, NC, Sep 2013 – 2018
- Used multivariate GLMs to identify ecological predictors of primate skull shape based on CT scans.
- Conducted simulation studies to evaluate statistical methods for reconstructing ancestral states.
- Used survival models to quantify parasite-mediated mortality in wild primates.
- Demonstrated fine-scale habitat segregation in mosquito communities using GLMMs and PCA, recommending <20 meters minimum resolution of spatial data in mosquito-borne disease models.

Harvard University, *Research assistant*   Cambridge, MA, Sep 2011 – 2013
- Performed meta-analysis of 14 published studies and 164 effect sizes to test and reject the 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

Olympic history analytics: Scraped data on 135k Olympians and shared analysis on Kaggle (>66K downloads as of April, 2021). https://www.kaggle.com/heesoo37/olympic-history-data-a-thorough-analysis

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

Stack Overflow survey analytics: Won $1000 Kaggle Award for analysis of Stack Overflow inclusion and ethics survey. https://www.kaggle.com/heesoo37/stack-overflow-2018-survey-age-gender-sexuality

OPEN-SOURCE CONTRIBUTIONS

ARTKIT (maintainer): Python library for developing efficient, GenAI-assisted testing and evaluation pipelines for GenAI applications: https://github.com/BCG-X-Official/artkit/commits?author=rgriff23

Parsons (contributor): 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 (author): R wrapper for the BayesTrait 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 completed 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

Zimmerman, A., Fox, S., **Griffin, R.**, Nelp, T., Thomaz, E.B.A.F., Mvungi, M., Mmbaga, B.T., Sakita, F., Gerardo, C.J., Vissoci, J.R.N., Staton, C.A. 2020. An analysis of emergency care delays experienced by traumatic brain injury patients presenting to a regional referral hospital in a low-income country. *PLoS ONE* 15(10): e0240528.

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., Omotayo, M., and Nunn, C.L. 2012. Phylogenetic host specificity and understanding parasite sharing in primates. *Ecology Letters* 15(12): 1370-77.

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. Four years of training and coached youth teams aged 12-19.