Guppy L. Stott

gs69042@uga.edu

EDUCATION

University of Georgia — **Athens, GA**
Bioinformatics, PhD — *2021 - Present*
NC State University — **Raleigh, NC**
Statistics, Masters — *2018 - 2020*
Clemson University — **Clemson, SC**
Mathematical Sciences, BS — Genetics Minor — *2013 - 2017*

RESEARCH

University of Georgia: Institute of Bioinformatics — **Athens, GA**
PhD Candidate — *2021 – Present*
- Investigating subsampling methods for phylogeographic analyses
- Applying Bayesian phylogeography to high-pathogenicity avian influenza
- Developed a Nextflow pipeline for SARS-CoV-2 pandemic response activities
- Developed graph database for integration of phylogenies and patient metadata

ORISE at the Centers for Disease Control and Prevention — **Atlanta, GA**
Influenza Data Science Fellow — *2020 – 2021*
- Developed a high-throughput reassortment detection algorithm for influenza viruses

Clemson University: Department of Mathematical Sciences — **Macauley Lab**
Undergraduate Research — *August 2016 – 2017*
- Worked on a boolean model of the glutamine operon using Sage, a Python-based mathematical programming language

Clemson University: Department of Genetics and Biochemistry — **Liang Lab**
Undergraduate Research — *August 2013 – 2016*
- Ran experiments on tulip poplar with a focus on lignin production (possible benefits for biofuel research)

PROFESSIONAL EXPERIENCE

TIAA — **Charlotte, NC**
Sr. Data Scientist — *2020*
- Led team (3) on advisor supervisory modeling, promoting more efficient audit methodologies
- Consulted with teams throughout the organization to solve business problems around data quality, enrollments, and workflow optimization, translating business needs to statements of work and analysis plans
- Began an enterprise-wide training program (over 1,500 enrollees) to bring the existing workforce up to speed on data science, statistics, and analytics processes which we are expanding due to the success of the program

Data Scientist — *2018 – 2019*
- Built a graph algorithm for the prioritization of ETL consolidation efforts and the identification of PII risk
- Modelled prospect identification using a random forest algorithm (among other iterations)
- Built survivorship models for participant post-enrollment activities (e.g. time to obtain web access or meet with an advisor)
- Designed and built a graph database solution for data lineage and client retention modeling

Data Science Technical Associate — *2017 – 2018*
- Built outlier detection models for institutional metrics in Hadoop Hive and Spark
- Automated the matching of hires to managers using game theoretic models (100 hours saved/year)
- Analyzed cyber-fraud data through the development of a JavaScript (D3) application

IT Summer Intern — *Summer 2016*
- Developed data visualizations around internal social media to find technology influencers for software discovery in JavaScript and CytoScape

---

PROFESSIONAL SERVICE

Athens Science Observer
Editor-in-Chief — *2023-Present*
Associate Editor — *2022 - 2023*
University of Georgia Department of Bioinformatics Graduate Student Association
Chair, Graduate Student Satisfaction Survey Committee — *2023 - Present*
Member, Bioinformatics Departmental Seminar Speaker Committee — *2023 - Present*

| *Chair, Bioinformatics Departmental Seminar Speaker Committee* | 2022 - 2023 |
| *Treasurer* | 2022 - 2023 |

University of Georgia Graduate Student Association

| *Representative, Department of Bioinformatics* | 2023 - Present |
| *Diversity, Equity, and Inclusion Committee* | 2023 - Present |

University of Georgia Graduate Pride

| *Secretary* | 2023 - Present |

TIAA Technical Associate Council

| *Vice President, Charlotte Chair* | 2017 - 2018 |

PROFESSIONAL AFFILIATIONS

American Society of Virologists (ASV)

| *Member* | 2023 - Present |

Society for Epidemiologic Research (SER)

| *Member* | 2023 - Present |

International Society of Non-Binary Scientists (ISNBS)

| *Member* | 2023 - Present |

TEACHING EXPERIENCE

| **TIAA** | **Charlotte, NC** |

Instructor: Technical Associate Python Training — 2018-2020
- Introduced new Technical Associates (new college graduate rotational program) to data analysis process
- Taught new employees basic Python, EDA, and basic statistics to help in their data analysis roles

Digital Course Designer: Fundamentals of Data Science - Enterprise training — 2019-2020
- Headed up our data science training series which teaches members of the company how to identify problems that are a good fit for data science solutions and how to make better use of our data
- Designed course selections for 3 audiences, business leaders, engineers, and analysts, to help them understand some basics of data utilization and data science in particular

Instructor: ML Models in Production Workshop (University of North Carolina-Charlotte Data Science Initiative) — 2019
- APIs for data consumption and building topic extraction models
- Model pickling, building an API in Flask for model consumption

Instructor: Spark Workshop (University of North Carolina-Charlotte Data Science Initiative) — 2018
- Detailed tips and tricks in Python and using the Jupyter development environment
- EDA in Spark and working with Hadoop, basics of map reduce

| **University of North Carolina - Charlotte** | **Charlotte, NC** |

Teaching Assistant: Data Analytics — 2018 – 2019
- Materials covered: statistics, Python, web visualization (D3 and dashboarding), databases, Tableau, Hadoop, and machine learning
- Helped develop lesson plans, host office hours, and provide guidance on project development

| **Clemson University: Academic Success Center** | **Clemson, SC** |

Community of Practice Leader for the Supplemental Instruction Program — 2015 – 2017
- Created an internal team to automate administrative tasks using JavaScript (90 hours saved/year)
- Advised (10-15) SI Leaders by leading weekly practice meetings and conducted observations of their sessions
- Awarded CoP Leader of the Year Award

Supplemental Instruction Leader — 2014 – 2015
- Led (3-8) undergraduate students in instructional sessions for Calculus II

SKILLS

Computer Languages: SQL, Python, R, SAS, SageMath, Cypher, C++, JavaScript, D3, HTML, Matlab, Java, C

Software: Tableau, Hadoop (Hive and Impala), Neo4J, Cytoscape, Spark

Certifications: Neo4j Graph Database Developer, SAFe 4 Agilist

PRESENTATIONS

| **Phylogenetic Analysis Subsampling Tool (PAST)** | **2023** |
| *CEIRR Network Annual Meeting* | *Poster* |

Phylogeny and Metadata Network Database Pipeline for Epidemiologic Surveillance 2022
CEIRR Network Annual Meeting *Poster*
Characterization of redbay (Perseaborbonia L.) gSSR markers 2014
American Society of Plant Biology Annual Meeting *Poster*
Unraveling the Basal Angiosperm CAD Genes Involved in Lignin Biosynthesis 2014
Clemson Creative Inquiry Poster Forum *Poster*

PUBLICATIONS

Phylogeny and Metadata Network Database for Epidemiologic Surveillance 2022
Preprint-Biorxiv *https://doi.org/10.1101/2022.04.19.488067*
Diversity level of genomic microsatellites in redbay (Perseaborbonia L.) generated by Illumina sequencing 2015
Journal of Plant Science and Molecular Breeding 4:2 *http://dx.doi.org/10.7243/2050-2389-4-2*

Science Communication

A Tale of Two Shelterbelts 2022
Athens Science Observer *https://athensscienceobserver.com/2022/11/30/a-tale-of-two-shelterbelts/*
Finding Comfort in Uncertainty 2021
Athens Science Observer *https://athensscienceobserver.com/2021/11/16/finding-comfort-in-uncertainty/*