

Nathalie Jones, M.S.

(770) 508-4908

github.com/njones738

nathalie.jones711@gmail.com

Summary

Data scientist with 2+ years of experience in data visualization, ML/AI, survey design, and predictive/descriptive analytics. Passionate about turning messy data into actionable insights and building tools that advance equity, fairness, and human-centered design.

Education

Kennesaw State University – Kennesaw, Ga	Graduation
<i>Master of Science in Data Science and Analytics</i>	Dec 2023
<i>Bachelor of Science in Computational and Applied Mathematics</i>	Dec 2021
<i>Minor in Applied Statistics and Data Analytics</i>	

Skills

Languages & Tools:

Python, R, SQL, SAS, Git, Tableau, MS Office, LaTeX, HTML

Soft Skills & Collaboration:

Human-Centered Design, Equity-Driven Analysis, Technical Writing, Presentation Delivery, Collaboration, Agile/Iterative Workflow

Recent Work Experience

1. Outlier & Data Annotations – AI model grader	Jan 2024
• ‘Graded’ AI responses to math (all levels) and science prompts for accuracy issues, hallucinations, and readability	– present
2. CARES Research Lab – Graduate Research Assistant	Aug 2022
• Conducted research on foster and homeless youth in higher education	–
• Developed and deployed large survey to 60K post-secondary faculty	Dec 2023
3. IHG Hotels & Resorts: Summer Internship – Data Science Internship	Jul 2023
• Analyzed the ROI of a summer promotional campaign	–
• Presented findings to head of advanced analytics	Aug 2023

Corporate Projects

1. “Detecting Data Oddities” – Team Project built in R, Python. Sponsored by IHG Hotels & Resorts	Dec 2023
• Wrangled Analyzed 460M+ hotel reservations for IHG using R and Python, developing anomaly detection tools with statistical methods	
2. “Predicting Customer Churn” – Team Project built in R, Python. Sponsored by Southern Company	Dec 2023
• Built and evaluated churn prediction models for Southern Company using R and Python, identifying key risk drivers and boosting classification accuracy	
3. “Predicting Email Click” – Team Project built in R, Python. Sponsored by IHG Hotels & Resorts	May 2023
• Built a high-performing XGBoost model (AUC: 95%) to predict email clicks across 18M+ records for IHG, driving actionable marketing segmentation	
4. “Modelling Chatbot Experience” – Team project built in R, Python. Sponsored by Southern Company	Dec 2022
• Improve chatbot intent classification and reduce query repetition for Southern Company’s internal AI assistant using BERTopic and SpaCy	
5. “Using Logistic Regression to Build Credit Scores” – Coded in R, SAS. Sponsored by Equifax	May 2022
• Built a logistic regression model for Equifax that classified credit risk with 85% concordance and (in theory) generated \$114K profit per 1K customers	

Poster Projects Papers & Presentations	1. “An Efficacy Rating for March Madness Tournament Seeding” – Coded in Python, SAS	
	<ul style="list-style-type: none"> Developed an 'Efficacy Rating' for the NCAA March Madness tournament, enhancing tournament seeding fairness and predictability 	Dec 2023
	2. “Charting the Course: A Data Driven Approach” – Coded in R, Python	
	<ul style="list-style-type: none"> Scraped PhD program information and faculty data from 220 programs across 60 U.S. universities to analyze and compare programs 	Nov 2023
	3. “Mapping Post-Secondary Transfer Rates” – Coded in R, Python	
	<ul style="list-style-type: none"> Studied the association between institutional transfer rates and the number of schools within 100 miles of each other 	May 2023
	4. “Why Withdraw?” – Coded in R, Python, SAS	
	<ul style="list-style-type: none"> Continued research into the CollegeScorecard with a longitudinal analysis on post-secondary withdrawal rate 	Dec 2022
Publications	5. “Classification of Pell Institutions” – Coded in R, Python	
	<ul style="list-style-type: none"> Used R and Python to clean, structure, and create/compare several models including XGBoost, PCA, Random Forest, and Logistic Regression 	Dec 2021
	6. “Does the Pell Grant Come with a Price?” – Coded in R, Python	
	<ul style="list-style-type: none"> Spatially joined the U.S. Census with the CollegeScorecard to analyze associations between debts accumulated by either an independent or dependent student 	Dec 2021
	7. “Two-Layer Neural Network” – Coded in Python	
	<ul style="list-style-type: none"> Created a 0, 1, or 2-layer neural net in Python as a class object to predict whether an individual in 1994 earned \$50K or more from Census data 	May 2021
	8. “Access to Higher Education” – Coded in R	
	<ul style="list-style-type: none"> Conducted research using parametric and nonparametric methods to study student outcomes and institutional attributes associated with schools where a majority proportion of students receive a Pell Grant 	May 2021
Publications	1. “Faculty Awareness Project” – First Author	
	<ul style="list-style-type: none"> Developed and deployed a survey to explore faculty knowledge of campus services that aid students experiencing homelessness 	Jun 2024
Publications	2. “Breaking the Cycle” – Co-Authored Article	
	<ul style="list-style-type: none"> Co-authored an article summary of work conducted by my lab director where she evaluated college support programs for former foster youth 	Jan 2024
Conferences	1. Southeastern SAS User Group 2023	Oct 2023
	2. Harvard’s National Collegiate Research Conference 2022	Jan 2022
	3. Posters on the Hill 2022	Apr 2022
	4. Posters at the GA Capitol 2022	Apr 2022
	5. Kennesaw State University Bi-Annual Analytics Day	Apr/Nov 2021 – 2023