Seth Hall

| Mebane, NC 27302 | hall.seth.b@gmail.com | (919) 428-4921 |

EDUCATION

University of North Carolina at Chapel Hill | Chapel Hill, NC

B.S. in Statistics and Analytics, Minor in Data Science — Expected May 2027

PROFESSIONAL SUMMARY

Statistics and Analytics major seeking a data-focused internship. Experience with predictive modeling, machine learning, and optimization, with a focus on turning data into clear, actionable insights.

TECHNICAL SKILLS

Programming: Python, R, GAMS; PyTorch
Libraries/Tools: PyTorch Geometric, Jupyter, SciPy, PuLP, Excel, RStudio, QGIS
Methods: Regression, predictive modeling, hypothesis testing, optimization/decision analysis, data cleaning, data visualization
Math: Probability, linear algebra, calculus for statistics
Business Ops: Regulatory compliance, EIN registration, website architecture
Certification: Microsoft Office Specialist

EXPERIENCE

Surefyr Software Solutions — Founder & Operator | Mebane, NC — Jul 2025–Present

- Founded and legally registered an LLC; handled EIN registration, compliance, and operational setup
- Designed early-stage website architecture while learning foundational web development concepts

STEAMHEALS — Nonprofit Co-Founder & VP of Operations — Aug 2024–May 2025

- Co-founded an international nonprofit delivering STEM + health/hygiene education to underserved students in the U.S., India, and Thailand
- Partnered with schools and a university to reach hundreds of students through scalable, teacher-led workshops
- Built curriculum and recruited volunteers to support long-term program growth

PROJECTS

NBA Game Outcome Prediction — Course Project — Spring 2026

- Built predictive models in Python to forecast NBA game offensive rebounds using team game data, rolling pregame features, and schedule-based rest indicators
- Engineered matchup variables from recent offensive rebound rates, missed-shot volume, possession proxies, and back-to-back context while avoiding same-game data leakage
- Compared linear regression, ridge, lasso, elastic net, Poisson, and random forest models using chronological out-of-sample testing and selected the best model by MAE

Heartbeat Classification with Sequence Models — Course Project — Spring 2026

- Trained RNN, LSTM, and GRU classifiers in PyTorch on MIT-BIH Arrhythmia heartbeat sequences (187 time steps) to predict 5 rhythm classes
- Compared models using accuracy/weighted F1, learning curves, and confusion matrices to interpret errors

Graph Classification with Graph Convolutional Networks — Course Project — Spring 2026

- Implemented GCN graph classifiers on the ENZYMES dataset using PyTorch Geometric, including a manual graph convolution and a PyG GCNConv version with global mean pooling
- Evaluated with accuracy, precision, recall, weighted F1, and confusion matrix using an 80/10/10 train/val/test split and hyperparameter tuning

Bias in Healthcare Algorithms — Independent Research Project — Fall 2024

- Replicated and expanded Obermeyer et al.'s work on bias in healthcare algorithms in R/RStudio
- Simulated synthetic datasets and applied statistical parity, equalized odds, and disparate impact metrics to quantify bias
- Proposed methodological improvements and ethical policy recommendations to reduce inequities