

Steve Broll

✉ sb2643@cornell.edu 🏠 stevebroll.github.io 📄 [stevebroll](#) 📺 [stevebroll](#)

Education

Cornell University

PHD IN STATISTICS

Ithaca, New York

Aug 2020 - May 2025

Texas A&M University

B.S. IN STATISTICS, MINOR IN MATHEMATICS

- GPA: 3.85

College Station, Texas

Aug 2016 - May 2020

Graduate Research Experience (Funded)

Department of Statistics

RESEARCH ASSISTANT

Cornell University

January 2021 - May 2022

- Studied effects of two drugs on the viral load and metabolite abundances in patients with Tuberculosis
- Used a method incorporating batch effect correction with latent factors and variable selection via empirical Bayes to study which metabolites showed significantly different behavior under different drugs
- Developed a method to model viral load in patients taking the more successful drug, inducing sparsity with a group penalty for the metabolite abundances

Outcomes: Paper in progress, expected to submit by end of Summer 2022.

Department of Statistics

RESEARCH ASSISTANT

Cornell University

May 2022 -

- Using variations on standard methods for modeling over-dispersed count data
- Incorporating batch effect correction via latent factors to detect relatively small signal among batch effect noise in ATAC data

Outcomes: Current work in progress.

Undergraduate Research Experience (Funded)

Undergraduate Research, Dr. Irina Gaynanova

STUDENT RESEARCHER, FUNDED BY DEPARTMENT OF STATISTICS

Texas A&M University

May 2019 - Present

- Developed R package for computing metrics and producing visualizations for Continuous Glucose Monitor data
- Emphasized ease of use and clear documentation
- Created Shiny Application to provide point-and-click interface with full functionality of the package

Outcomes: Built an R package 'iglu' associated [Shiny App](#) with [demo](#), and [poster presentation](#). Developed skills in R, R package development, version control, and Shiny.

Summer Institute in Biostatistics (SIBS)

STUDENT RESEARCHER

North Carolina State University

Jun 2018 - Jul 2018

- Attended classes covering a wide scope of biostatistics applications and implemented methods in labs using R and SAS
- Analyzed comorbidities of subjects who suffered a myocardial infarction with a binary response of 90 day survival
- Performed hypothesis testing for individual comorbidities and fit sparse logistic regression model to predict odds of 90 day survival

Outcomes: Completed group project with [poster presentation](#). Learned about a wide scope of biostatistics applications as well as motivations for a PhD in Statistics or Biostatistics.

Undergraduate Internship, Dr. Alan Dabney

INTERN

Texas A&M University

Jun 2017 - Aug 2017

- Studied standardized test performance across cohorts of College of Nursing
- Conducted exploratory data analysis and clustering methods to identify covariates useful for predicting test scores
- Provided visualizations and results to faculty advisor and College of Nursing collaborator

Outcomes: Produced report in a Jupyter notebook using Python. Gained experience with formal presentations and working with a collaborator outside statistics.

Undergraduate Research Experience (Unfunded)

Capstone Project

Texas A&M University

RESEARCH GROUP MEMBER

Sep 2019 - PRESENT

- Worked in group project modeling monthly temperature anomaly data from Berkeley Earth
 - Fit seasonal ARIMA time series model to global anomaly data and pixelwise regression models to North America land surface grid
- Outcomes:** Produced three short reports. Final report and poster presentation in progress.

Texas A&M Institute of Data Science (TAMIDS) Competition

Texas A&M University

FIRST PRIZE COMPETITION TEAM MEMBER

Apr 2018 - Apr 2018

- Investigated trends in over 110 million Chicago taxi rides from 2013-2017
 - Trained ARIMA time series model on four years taxi data and tested predictive performance with the final year
- Outcomes:** Contributed to [report](#) chosen as a finalist, and [presented](#) with team to earn first prize.

Undergraduate Research, Dr. Boris Hanin

Texas A&M University

STUDENT RESEARCHER

Oct 2018 - Dec 2019

- Met biweekly with research group from Computer Science department discussing potential techniques for training of neural networks
 - Simulated test data to be used for classification with neural networks
- Outcomes:** Produced simulated data and found interest in deep learning for classification.

Undergraduate Research, Dr. Huiyan Sang

Texas A&M University

STUDENT RESEARCHER

Jan 2018 - Feb 2019

- Explored spatio-temporal functional MRI scan data and techniques to connect motor functions to brain activity
 - Focused on modeling activity across brain surface and extending 2-dimensional sparse clustering techniques to 3-dimensional scan data
- Outcomes:** Presented [poster](#), studied techniques that sparked interest in spatial and spatio-temporal statistics.

Undergraduate Research, Dr. Alan Dabney

Texas A&M University

RESEARCH GROUP MEMBER

Oct 2016 - Mar 2017

- Implemented models from *Introduction to Statistical Learning: with Applications in R* on baseball data
 - Used logistic regression and quadratic discriminant analysis (QDA) to model player's chances of making the Hall of Fame
- Outcomes:** Fit a range of predictive models in R, collaborated with peers in statistics, and presented poster

Leadership

Department of Statistics; College of Science

Texas A&M University

PEER MENTOR

Sep 2017 - Present

- Developed and sustained relationships with and provided assistance and support for first-year students
- Maintained regular contact with six assigned students, including weekly one-hour meetings

Teaching

BTRY 6010

Cornell University

TEACHING ASSISTANT

Fall 2020

- Ran lab section for introductory statistics course for graduate students from a variety of fields
- Demonstrated applications of course material in R, and reviewed key results in combinatorics and statistics

cRamp: Boot camp in R

Texas A&M University

INSTRUCTOR

Sep 2018 - Nov 2018

- Designed beginner R course designed for students taking the introductory calculus-based statistics course
- Created [lecture notes](#), taught weekly in-class lectures, and recorded online lectures for students' convenience

Technical Skills

Advanced, R, R package development, R Markdown, Shiny, ggplot, LaTeX

Intermediate, Python, Git, Stan

Basic, C++, Rcpp/RcppArmadillo

Honors & Awards

TEXAS A&M UNIVERSITY

Recipient , Dr. Newton Service Award	2020
First Prize Team , Texas A&M Institute of Data Science Undergraduate Competition	2018
Recipient , Statistics Department Scholarship	2017, 2018
Recipient , President's Endowed Scholarship	2016 - Present

Presentations

TEXAS A&M UNIVERSITY

Statistics Undergraduate Project Showcase , <i>Spatiotemporal Analysis of Berkeley Earth Temperature Anomaly Data</i>	Dec 2019
Statistics Undergraduate Research Poster Session , <i>Interpreting Blood Glucose Data in R with iglu</i>	Sep 2019
Texas A&M Institute of Neuroscience Annual Symposium , <i>Spatiotemporal Mixed Modeling of task-fMRI Data</i>	Apr 2018
Texas A&M Student Research Week , <i>Spatiotemporal Mixed Modeling of task-fMRI Data</i>	Apr 2018
Texas A&M Institute of Data Science Competition Finalist Presentation , <i>Forecasting Median Daily Fare with ARIMA Models</i>	Apr 2018
Texas A&M Student Research Week , <i>Predictive Models using Discriminant Analysis on the MLB Baseball Hall of Fame</i>	Apr 2017

NORTH CAROLINA STATE UNIVERSITY

Final Group Presentation , <i>SYMPHONY Data</i>	Jul 2018
--	----------

Service

Directed Reading Program

Cornell University

MENTOR

Spring 2021-

- Directed undergraduates each semester in reading papers and textbooks related to topic of their choice, culminating in a final project and presentation

Texas A&M Math and Stats Fair

Texas A&M University

VOLUNTEER

Feb 2018, Feb 2019

- Set up statistics room and ran demonstrations of basic statistical concepts for K-12 students

Statistics Undergraduate Student Association

Texas A&M University

OFFICER

Aug 2018 - April 2019

- Helped organize and run events and socials for statistics undergraduates