Jonathan Barnes | Data Scientist

Minneapolis, Minnesota | Jonathanmbarnes@outlook.com | Jonathanmbarnes.github.io

RELEVANT SKILLS & KNOWLEDGE

Languages: R, Python, SQL, Bash.
Tools: Quarto, Shiny, APIs, Git, Linux, PowerShell, ArcGIS, Tableau, Azure, AWS.
Competencies: Data Visualization, Data Pipelines, Data Communication, Geospatial Visualization, Feature Engineering, Regression, Predictive Modeling, Machine Learning, Dimensionality Reduction, Clustering, Classification, Longitudinal Data Analysis, HIPAA, Informatics, Relational Databases, Electronic health Records (EHR), SNOMED, ICD-10.
Statistical and Machine Learning Methods: Random Forest, Gradient Boosting, Principal Component Analysis (PCA), Multiple Correspondence Analysis (MCA), Generalized Estimating Equations (GEE), Generalized Mixed Effects Models (GLMM), Lasso (L1), Ridge (L2), Elastic Net (L1/L2), Cox Proportional Hazard, Structural Equation Models (SEM).
Strengths: Problem-solving, Critical Thinking, Pattern Recognition, Integrity, Simplifying Concepts, Collaboration.

RELEVANT EXPERIENCE

Center for Public Health Systems, University of Minnesota School of Public Health, Minneapolis, MN
Graduate Research Assistant (May 2025 – Aug. 2025)

- Developing ETL pipeline that pulls federal spending from USAspending.gov API.
- Optimizing storage by using Parquet columnar data storage and big data techniques.
- Building Tableau dashboard communicating spending data in efficient and accessible manner.

Division of Biostatistics and Health Data Science, University of Minnesota School of Public Health, Minneapolis, MN
Graduate Teaching Assistant (Aug. 2024 – May 2025)
Spring 2025 - PUBH 6450 Online

- Graded assignments and quizzes, offering constructive feedback to guide student improvement.

Fall 2024 - PUBH 6450 In-Person

- Taught weekly lab sessions in R and SAS to develop student proficiency in statistical analysis and programming.
- Graded assignments and quizzes, offering constructive feedback to guide student improvement.

University of Minnesota Medical School, Minneapolis, MN
Graduate Research Assistant (Oct. 2023 – May 2024)

- Performed in-depth regression analyses on biomarkers for Dementia in NIA-funded Health and Retirement study.
- Developed proprietary classification groups based on plasma biomarker elevations.
- Designed Cox Hazard models to assess risk in study population to predict dementia risk.
- Programmed differential gene expression analysis using biomarker groups and recorded cognitive decline.

University of St. Thomas, St. Paul and Minneapolis, MN
ITS Support Services Intern & Senior Student Technician (Aug. 2021 – Aug. 2023)

- Recipient of the ITS Outstanding Student Award for contributions to ITS and Support Services.
- Identified and resolved process gaps, ensuring seamless workflows by proactively addressing vulnerabilities.
- Engineered and developed a Python application vital to ITS service desk function and operational efficiency.
- Mentored numerous junior students, equipping them to independently handle operations previously handled by me.
- Implemented digital signage across service desks and various campuses, providing easy information access.
- Authored key technical documentation and procedural guides for The Tech Desk, ITS staff, students, and faculty.
- Imaged, configured, and administered software on over 1,500 replacement computers, 2.5 times a normal Summer.
- Assembled in-house virtual lab to replace existing AWS EC2 instances and reduce costs.
- Resolved or consulted on at least 4,500 departmental tickets over two years.
- Served as the first point of contact for student workers at phone centers and walk-up desks.

EDUCATION

University of Minnesota School of Public Health, Minneapolis, MN
Master of Public Health – Public Health Data Science, GPA 3.73/4.00 (Expected Aug. 2025).

- President's Student Leadership & Service Award Nominee (2025).

University of St. Thomas, St. Paul, MN
Bachelor of Science – Data Analytics, Biology Domain (May 2023), Completed concurrently with full-time employment.

- Minors – Biology (Biochemistry & Computational Biology courses), Applied Statistics.
- Capstone - Development of Obesity and Nutritional Trends Over Time in the United States Using NHANES.

LEADERSHIP AND VOLUNTEER EXPERIENCE

Biostatistics Community Outreach & Engagement (BCOE) | *Member (Sep. 2023 – Present)*
School of Public Health Student Senate (SPHSS) | *BSHD Representative (Oct. 2023 – May 2025)*

- *Internal Committees – Advocacy, Constitution Revision, SPHere Lounge Improvement.*

Center for Health Interprofessional Programs (CHIP) Executive Council | *SPH Representative (Oct. 2024 – May 2025)*
Health Sciences Student Consultative Committee (HS SCC) | *SPH Representative (Nov. 2023 – May 2024)*

HIGHLIGHTED PROJECTS

Mapping Misery – Predicting Poor Physical and Mental Health Days in the U.S.

- Developed predictive models for poor physical and mental health days in 3,000+ U.S. counties using R.
- Engineered clusters based on 10 years of trends across 15 metrics using Gaussian Mixture Models (GMM) by transforming variables into 8 level categories with NAs for robust Multiple Correspondence Analysis (MCA).
- Achieved an R^2 of 0.844 (RMSE: 0.263) for physical health days with Elastic Net Regression.

Skills & Tools: Data Pipelines, PCA, MCA, GMM Clustering, Feature Engineering, Longitudinal Data Analysis, Elastic Net Regression, Predictive Modeling, R, Geospatial Visualization, GIS, Data Visualization, Data Communication.

Tech Desk Tools Python *(ITS Outstanding Student Award)*

- Built a GUI-based desktop application for the University of St. Thomas Tech Desk to streamline common support tasks such as ticket creation, password resets, system diagnostics, and resource access.
- Redesigned and modernized legacy AutoHotKey application in Python, improving functionality and future scalability.
- Integrated role-specific admin features for senior student technicians based on assigned Azure roles.
- Deployed the application to all staff and student workers at the University of St. Thomas's Tech Desk.

Skills & Tools: Python, PowerShell, Bash, Dashboards, Data Communication, GUI Development, Active Directory, Azure, Technical Writing, Packaging and Deployment.

Predicting Child Mortality and Cesarean Delivery Outcomes with Generalized Linear Mixed Models.

- Built Generalized Linear Mixed Models (GLMM) to investigate factors affecting child mortality and cesarean delivery.
- Determined a 1,124.5% increase in the odds of cesarean delivery for hospital births (95% CI: 618.7%–1,986.3%, $p < 0.001$) — a result consistent with logical expectations.

Skills & Tools: GLMM, R, Data Cleaning, Data Pipelines, Data Visualization, Feature Engineering, Collaboration.

London Tube Transit Dashboard

- Created an interactive map dashboard in R to visualize live data from the London Tube system.
- Incorporated route geometries, station locations, and enabling real time updates from a live API.

Skills & Tools: R, Shiny, GIS, APIs, Data Pipelines, Dashboards, Data Visualization, Data Communication.

Data Inspection Dashboard | https://jonathanbarnes.shinyapps.io/DescriptiveApp/

- Developed an interactive tool that allows users to conduct baseline Exploratory Data Analyses (EDAs) without requiring R knowledge and accepts common dataset formats in public health.
- Automates and facilitates data cleaning, NA handling, imputation, and produces a data table with inline distribution plots and summary statistics.

Skills & Tools: R, Shiny, Data Pipelines, Dashboards, Data Visualization, Data Communication, JavaScript.

Biostatistics Reference Guide

- Provides step-by-step instructions and R code for linear regression, logistic regression, Poisson regression, Cox proportional hazards models, survival analysis, ordinal regression, and log-binomial regression.
- Includes practical tips for R-based modeling and coding.

Skills & Tools: R, Regression Analysis, Survival Analysis, Logistic Regression, Poisson Regression, Ordinal Regression, Log-binomial Regression, Tidyverse, ggplot2, Technical Writing.