Amos Ochieng Okutse

amos_okutse@brown.edu | 401-390-6037 | https://okutse.github.io/
20 Olive Street, Apt. 5, Providence, RI 02906

EDUCATION

Expected 2023	Brown University, Providence, RI ScM. Biostatistics Thesis: Machine Learning Methods for Missing data induced bias correction in randomized trials. Advisor: Prof. Joseph W. Hogan
2021	Jomo Kenyatta University of Agriculture & Technology, Nairobi, KE B. Sc. Biostatistics Thesis: An Ensemble-based Machine Learning Decision Framework for Patient Classification. Advisor: Dr. Jane A. Aduda

TECHNICAL SKILLS

Data Science:	Machine Learning, Statistical modeling, Exploratory data analysis, Feature engineering, Big data analytics, Data wrangling, and Database systems.
Programming languages:	R, Statistical Analysis System (SAS), Unix/Bash/PowerShell, MySQL, Stata, WINBUGS, GeoDa.

RESEARCH EXPERIENCE

2021- Present	Graduate Fellow Department of Biostatistics, Brown University, Providence, RI Advisor: Prof. Joseph Hogan
2019	Research & Data Analysis Intern Department of Biometry, Non-ruminant Research Institute, Kenya Agricultural & Livestock Research Organization, Kakamega, KE Supervisors: Fred Emongor and S. Mudekheri

PUBLICATIONS

- **Okutse, A. O.**, & Nyongesa, K. W. (2021). Differential Expression Analysis for the Identification of Survival Associated Genes in Primary Bladder Cancer using Micro-array Data. *International Journal of Undergraduate Research and Creative Activities, 13*(1). https://doi.org/10.7710/2168-0620.0306

- **Okutse, A. O.** (2020). Impact of Sleeping Duration on the Risk of Breast Cancer: A Systematic Review and Meta-Analysis of Population-Based Cohort Studies. *Reinvention: An International Journal of Undergraduate Research, 13*(1). https://doi.org/10.31273/reinvention.v13i1.530

PREPRINTS

- **Okutse, A. O.** (2019). Statistical Analysis of Non-Pair Matched Contingency Table Data: A Non-Technical Primer. A Preprint. https://10.13140/RG.2.2.13786.90561

- **Okutse, A. O.** (2019). A Naïve Bayes' Probabilistic Classifier for Modeling the Quality of Care in a Healthcare Setting. GRIN Verlag. ISBN: 978-3-346-14278-8. https://m.grin.com/document/535799

CONFERENCE PRESENTATIONS

2021   **Okutse, A. O.** and Athiany, H., "Trends of Socioeconomic Disparities in the
                Kenyan Child Malnutrition Statistics: Analysis of the Demographic and
                Health Survey, 2003-2014," in *Proceedings of the Sub-Saharan Africa Network
                (SUSAN) of the International Biometrics Society*, Nairobi, Kenya, Sep 2021

WORKSHOPS & CERTIFICATIONS

Fall   **The Sheridan Teaching Seminar (Certificate I)** – Critical reflection and inclusive
2022   classrooms, rhetorical practice & classroom communication, learning design, and
                student engagement. The Harriet W. Sheridan Center for Teaching & Learning,
                Brown University, Providence, RI.

Fall   **Brown Ethics & Responsible Conduct of Research (RCR).** Office of the Vice
2022   President for Research, Brown University, Providence, RI.

Summer   **Responsible Conduct of Research.** The Collaborative Institutional Training
2022   Initiative (CITI) under requirements set by Brown University.

Summer   **Principles of HIV/STI Research and Public Health Practice.** Department of
2022   Global Health, University of Washington, Seattle, Washington.

Fall   **An anti-racist Imperative for Public Health Data.** The New England Public
2021   Health Training Center.

Fall   **Getting Started with SAS Programming.** The SAS Institute through
2020   Coursera.

Spring   **Introduction to Statistical Genetics:** *Theory & Practice* held at Jomo Kenyatta
2021   University & Sino-Africa Joint Research Institute (JKUAT-SAJOREC), Department of
                Statistics & Actuarial Sciences, SAJOREC-JKUAT, KE.
                Facilitator: Prof. Dr. Bernard Omolo, University of South Carolina, Upstate,
                Spartanburg, South Carolina, USA.

RELEVANT PROJECTS

Fall 2021   **Modeling Infectious Disease Dynamics using Discrete Compartmental Models.**
                A collaborative class project with Kayla Finlayson and Yingjie Zhou in Statistical
                Programming with R under Dr. Alice J. Paul.

                - Developed a dynamic and interactive web-based application using Shiny
                  Dashboards and the R statistical programming environment to simulate
                  infectious disease dynamics using discrete compartmental models including the
                  SI, SIR, and SEIR. The application provides an interactive interface for
                  modeling and visualizing trends of infectious diseases, and allows data
                  download given parameters including the type of model to simulate, the
                  infection rate, the total susceptible population, the total number infected, rate
                  of recovery, total number recovered, total population exposed, the simulation
                  time (in days), and also explores the effect of an intervention e.g., vaccination
                  on the subsequent trends in the disease dynamics.
                  https://idiseases.shinyapps.io/indiseases/

Spring 2022	**Bayesian Learning Artificial Neural Networks for Modeling Survival Data.**

A class project for Statistical Learning and Big Data with Dr. Alice J. Paul, collaborated with Naomi Lee.

- Developed an instructive, introductory extension post using the R statistical programming environment, Python, HTML, and gene expression profiles data from patients with primary bladder cancer from the Gene Expression Omnibus (GEO) to show how Bayesian inference can be applied in modeling survival in patients with primary bladder cancer using single-layer artificial neural networks with Bayesian Learning.
  https://okutse.github.io/bayesian-networks/

Fall 2022 Statistical Machine Learning Methods for *Listeria Monocytogene* foodborne disease source attribution.

A class project in Practical Data Analysis with Dr. Alice J. Paul, collaborated with Zexhuan Yu and Rophence Ojiambo.

- This is a collaborative project with Dr. Ernest Julian (Co-chair with the Centers for Disease Control and Prevention [CDC] and the Food and Drugs Administration [FDA] of the Healthy People 2030 Foodborne Illness Reduction Committee). In the project, we use statistical machine methods for multiclass classification including Random forests, Naïve Bayes Classification, and Bayesian Additive Regression Trees (BART) to predict the food source of a *Listeria monocytogene* strain and isolate to propose a robust and throughput predictive framework for driving action in the fight against foodborne illnesses using data from the National Center for Biotechnology Information (NCBI) Pathogen Detection Database.

Spring 2019 A critical evaluation of "The effects of body weight loss and gain on arterial hypertension control: a prospective observational study."

A class project in Research Methodology with Dr. Peter O. Juma.

- Carried out an in-depth review and critique of the paper based on its methodological soundness and suitability of the implemented analysis.

SCHOLARSHIPS & AWARDS

2021	**NAMBARI: The Moi-Brown Partnership for Biostatistics Training.** The Global Health Initiative [GHI], Brown University, Providence, RI.
2021	**Department of Biostatistics Tuition Scholarship:** Department of Biostatistics, School of Public Health, Brown University, Providence, Providence, RI.

PROFESSIONAL AFFILIATIONS

2021 - Present	**International Biometric Society [IBS].** Kenya Chapter, Nairobi, Kenya.
2021 - Present	**Kenya National Statistical Society [KNSS],** Nairobi Region, Nairobi, Kenya.
2017 - Present	**Jomo Kenyatta University Biostatistics Student Association [JKUBSA],** Jomo Kenyatta University, Nairobi, KE.

LEADERSHIP

2019	**Project Team Lead:** The First Jomo Kenyatta University and the Kenya Authority for Parking Services (KAPS) Hatch-hack hackathon. Led a team that emerged third-runners-up for developing a dynamic and interactive web-based application portal for patient classification using ensemble-based machine learning methodologies.
2022	**Peer Mentor:** Department of Biostatistics & the Graduate School First Years Peer Mentoring Program.

RELEVANT COURSEWORK

Jomo Kenyatta University of Agriculture & Technology	Introduction to time series analysis, Calculus, Ordinary & Partial Differential Equations, Categorical Data Analysis, Multivariate Methods, Nonparametric methods, Introduction to Spatial Statistics & Epidemiology, Introduction to Design & Analysis of Clinical Trials, Introduction to Bayesian Inference, Design and Analysis of Sample Surveys, Design and Analysis of Experiments, Regression Analysis, Survival Analysis, Infectious Disease Modeling, Statistical Genetics.
Brown University	Statistical programming with R, Applied Generalized Linear Models, Probability & Statistical Inference, Introduction to Epidemiologic Methods, Applied Multilevel Models, Applied Longitudinal Models, Statistical Learning and Big Data, Practical Data Analysis, Causal Inference & Missing Data.