MICHAEL NSIAH-NIMO, M.Sc.

RESEARCH SCIENTIST

My Website

CONTACT

Lincoln, Nebraska

michael20129648@gmail.com

Linkedin

EDUCATION

MSc Statistics
Focus: Machine Learning, Causal Inference, Causal Mediation Analysis Predicitive Analytics, Survival Analysis, Bayesian Statistics, Study Design

University of Texas at El Paso

2015 - 2017
BSc Actuarial Science

Kwame Nkrumah University of Science and Technology

2010 - 2014

SKILLS & TOOLS

- **Database Management and Standards:** Proficient in electronic data capture tools such as REDCAP and QUESTIONPRO, ensuring data quality, ethical considerations, integrity and compliance according to standards used in clinical trials and drug development including CDISC and ICH guidelines and regulatory authority guidance from the FDA, EMA.
- **Basic Science Research :** Expertise in planning, executing, reporting, and documenting experimental and observational studies, ensuring they are conducted efficiently, ethically, and in compliance with regulatory standards. This includes protocol development, SAP creation, CRO deliverables, and the production of Tables, Listings, and Figures
- **Statistical Methodologies and Analytical Techniques.** Proficient in Data Mining , Exploratory Data Analysis Visualization, Predictive Analytics, Experimental Design, Multivariate Analysis, Study Design and Sample Size Estimation, Probability Theory, Missing Data Imputation, Parametric and Non - Parametric Statistics , Multilevel Modeling, Survival Analysis, Longitudinal Data Analysis, Hypothesis Generation and Testing, AI, Machine Learning , Causal Inference
- **Programming Languages**: Skilled in Python (7+ years), R (6+ years) , SAS (5 + years), Stata, SQL , Excel, Skilled in developing and implementing reusable macros, Python classes, and SAS programs to create robust pipelines for data cleaning, variable selection, validation, analysis, and report generation
- **Libraries and Frameworks**: Proficient in R Tidyverse (ggplt2, dplyr, tibble, readr, tidyr, broom, modelr), R biostatistics, R shiny, glm, glme, lmer, RMarkdown, Git, GitHub, TensorFlow, NumPy, Pandas, Python statsmodels, Plotly, Scikit-Learn, Matplotlib, Seaborn), Tabelau and Power BI
- **Database Technologies**: Experienced with Google BigQuery, AWS, Snowflake, Azure,

SUMMARY

Over the past 7 years, I have built expertise in adapting research data and analytics, data visualization and prevalence informatics, analytical programming, statistical estimation and machine learning. I have consulted on **over 50 projects involving biomedical research, healthcare, educational, industrial, and marketable consumer services** with the goal of enhancing and simplifying complex business challenges to inform stakeholders decision and improve organizational profit and objectives.

WORK EXPERIENCE

Staff Research Statistician

JULY 2020 - PRESENT

Border Biomedical Research Center, University of Texas at El Paso

- Lead the Data Analytics team under the Associate Dean for Research driving a 20% increase in **extramural funding for cancer research** and managed successfully **analytics for faculty and student grant awards, research expenditures**, and other college data analytics for stakeholders, enhancing overall research productivity
- **Built a Faculty Retention Metric (FRM)**, to forecast potential faculty churn and identify the most significant contributors which enabled the institution retain top faculty talent, reduce disruptive churn, **and maintain consistent research output and reach a significant milestone in national research (Ro1) status**
- **Analyzed wearables (Fitbits) data** from the All of Us Research Program to track time series metrics, including physical activity, sleep quality, and heart rate variability, as indicators of physiological changes and stress which provided insights into cardiovascular and metabolic health, supporting assessment of biological age
- **Performed data mining and exploratory analyses** using Bayesian and Frequentist approaches across biological domains (physiological, immunological, metabolomic) and non-biological domains (social determinants of health from surveys, health records, healthcare claims and costs) **to identify the most informative features and categorical factor levels** to enhance statistical modeling, pattern recognition and prediction, and disease outcome estimation.
- **Developed and implemented health indices and risk scores applying gradient boosting and autoencoders with deep neural networks for immunological aging**, that predicts and tracks inflammation levels, aging-related processes and health disparities in Hispanic Origin Individuals (HOI) **to secure a national grant of over $1 million in research funding and improved patient outcomes by 25%**
- **Modeled the interaction of immune and stress responses using causal mediation analysis,** identifying a novel biomarker as key mediator of effects on metabolic health. Results suggest that stress pathways **may causally regulate inflammation-related metabolic dysfunction, informing potential therapeutic targets.**
- **Developed interactive dashboards in Jupyter Notebooks and R Shiny web apps to visualize key wearable metrics—**physical activity, sleep quality, and heart rate variability—**stratified by age groups,** enhancing ability to assess biological age across diverse populations
- Worked closely with the analytic unit to perform **comprehensive exploratory data analyses, study designs, hypothesis testing and statistical modeling for federal grant projects, informing model building** and **inferring causalities** significantly advancing research in socioeconomic and health disparities
- Developed a **ranked list of recommended matches** for targeting nonprofit organizations and **extramural grants, using text analysis and NLP tools t**o assess mission and vision statements as well as recently funded grants.
- Conducted **research and literature reviews on machine learning and statistical hybrid models**, focusing on **interpretability and prediction-first approaches** to optimize model effectiveness for research objectives, while staying current with state-of-the-art advancements.
- Built **data mining automation pipelines using advanced SQL, R tidyverse and python pandas queries to generate actionable insights from NIH research databases, such as All of US and NHANES**

Statistical Data Analyst

College of Science , Research Enterprise
University of Texas at El Paso,

JULY 2018 - JULY 2020

- **Conducted time series analysis and forecasting on expenses, revenue, and grant funding data**, resulting in improved financial planning and resource allocation strategies
- **Formulated and applied mathematical models and other optimizing methods to develop and interpret information** that assists management with decision making, policy formulation, and other functions associated with the College of Science research enterprise.
- Worked on diverse datasets, leveraging both **parametric and non-parametric statistical methodologies** for federal and philanthropic foundation grants and increasing external funding
- Developed **comprehensive metrics to serve as financial determinants, analyzing key issues such as revenue, expenditure, and extramural funding performance, market trends, and economic disparities in the Borderplex region**
- Used **SQL, Tableau, and MapBox** to **extract, model, and visualize demographic data f**rom El Paso County, assessing socioeconomic divides within zip codes and their related financial implications **boosting targeted grant applications for health disparities research by 50%.**

MICHAEL NSIAH-NIMO, M.Sc.

RESEARCH SCIENTIST

Behavioral Skills

- Adept at using simple language without watering down the key ideas of the research and focus on the purpose and results of the study to capture audience comprehension and relatability
- Proven track record of seamless collaboration across interdisciplinary teams, contributing to project success.
- Adept in identifying challenges and implementing innovative solutions ensuring the integrity of the analyses.
- Experience in providing guidance and mentorship to junior team members fostering a collaborative work environment.
- Flexibility with changing priorities, strong attention to detail, ability to work well under pressure and take on unfamiliar tasks

PUBLICATIONS

- ***[In Review]*** *Exploring Socio-Behavioral Correlates of Metabolic and Inflammatory Risk in Hispanics Living Along the U.S./Mexico Border: A Pilot Study Concomitantly Collecting Survey Data, Blood and Hair Samples, and Physical Measures, Xu, Cai, Frietze, Gabriel, Leung, Ming-Ying,* ***Nsiah-Nimo, Michael,*** *Sanyal, Nilotpal, Begum, Khodeza Robles, Elisa Flores, Diana P , Mandal, Abhijit, Liang, Panfeng*

In Preparation

- *Biological Age Index of Chronic Inflammation as a Predictor of Cardiometabolic Phenotypes: Differential Associations with General, Central Obesity and Diabetes in Mexican Origin Hispanics, Robles, Elisa, Leung, Ming-Ying,* ***Nsiah-Nimo, Michael,*** *Sanyal, Nilotpal, Xu, Cai, Frietze, Gabriel, Begum, Khodeza , Flores, Diana P , Mandal, Abhijit, Liang, Panfeng*

- *Associations Between Social Support, Perceived Stress, and Metabolic Health Outcomes: A Cross-Sectional Analysis of General Obesity, Central Obesity, and Diabetes in the Hispanic Health Hub. Robles, Elisa, Leung, Ming-Ying,* ***Nsiah-Nimo, Michael,*** *Sanyal, Nilotpal, Xu, Cai, Frietze, Gabriel, Begum, Khodeza , Flores, Diana P , Mandal, Abhijit Liang, Panfeng*

CERTIFICATIONS

- **Probability Theory, Generalized Linear Models and Inferential Statistics (Duke University)** | May 2018 – September 2019
- **Supervised Machine Learning**: Regression and Classification | October 2022
- **Clinical Data Analysis with SAS** (Oreilly)

- ***Relevant Graduate Project Work***

Efficacy of Treatment Sessions in a Medical Emergency Department - Utilized Interaction Plots, Hasse Diagrams and Linear Mixed Models to analyze and conduct hypothesis tests to reveal patterns of illicit drug use, such as cocaine and cannabis, among individuals, gender and age

PROFESSIONAL CONFERENCES

- Microsoft AI Research Forum, 2024
- University of Hawaii Bioinformatics Data Science Workshops, 2023
- NIH NIGMS Cloud Based Biomedical Research, 2023
- Better Data for More Equitable Research : Research America Alliance, 2023
- RCMI Seminars at Meharry Medical College, 2023
- College of Science Stakeholders' on Health Disparities and Cancer Research, 2016 ,2017, 2018, 2019, 2020, 2021, 2022, 2023
- JSM Conference 2017, Diversity Scholar
- Cardwell Foundation Seminar, 2016

Statistical Analyst Intern | Project Planning and Study Design Analysis

University of Texas at El Paso, Center for Institutional, Research and Planning — SUMMER 2018

- Collaborated with a team of researchers on multiple **project planning, design, and analysis** to address institutional and educational policy issues at UTEP
- Facilitated academic publications presentations utilizing Big Data for institutional research and university stakeholders

Lecturer/Research Statistician

AUGUST 2015 - JULY 2018

Alcohol Research Group, University of Texas at El Paso

- **Analyzed treatment efficacy in emergency department settings** by employing linear mixed models and interaction plots, uncovering significant patterns of illicit drug use, such as cocaine and cannabis, across different age groups and genders, to inform tailored intervention strategies.
- **Applied Hasse Diagrams and advanced statistical techniques to** conduct hypothesis testing, providing actionable insights on substance use behaviors, supporting clinical teams in optimizing treatment sessions for Hispanic health initiatives.
- Collaborated with Dr. Amy Wagler on multiple statistical research projects, **leading to advancements in statistical methodologies that improved data interpretation** and **contributed to manuscripts reviewed for publications.**
- **Applied Design of Experiments (DOE) techniques to clinical data analysis,** increasing experimental accuracy and reliability by 20%, resulting in more consistent and valid research outcomes.
- Instructed over 150 students in advanced mathematics, achieving a 90% pass rate while **mentoring students on end-of- semester projects**, enhancing their understanding and application of statistical techniques in real-world scenarios

BUSINESS AND HEALTH PROJECTS

Compressing Feature Space For Classification Using PCA

- Applied Principal Component Analysis (PCA) ***to reduce the dimensionality of 50 cytokines to capture variation in estimated biological ages among Hispanic Origin Individuals leading to an increase in model performance***

Enhancing Targeting Accuracy Using ML

- Built a model that would accurately predict the customers that would sign up for a delivery club. ***This allowed for a much more targeted approach when running the next iteration of the campaign***

Predicting Customer Loyalty Using ML

- Built a predictive model using Random Forest in Python to estimate customer loyalty scores for a company's data agency ***leading to a 30% increase in the ability to contact customers with promotional material***

"You Are What You Eat" Customer Segmentation

- Used k-means clustering on grocery transaction data to split out customers into distinct "shopper types" ***to accurately target customers with relevant content & promotions***

Fruit Classification Using A Convolutional Neural Network

- Built & optimized a Convolutional Neural Network to classify images of fruits, with ***the goal of helping a grocery retailer enhance & scale their sorting & delivery processes.***

Quantifying Sales Uplift With Causal Impact Analysis

- Analyzed customer retail practices to quantify the increase in sales attributed to customers joining the company's delivery club compared to what they would have spent without the club***. This analysis revealed a significant 41.1% uplift in sales among club members, indicating the club's positive impact on customer spending***

Professional Contributions to Federal and Non - Federal Grants Projects

- ***NIH Project No : 5U54MD007592 -27, SUB IDS : 8320, 8321, 8317***
- ***NIH Project No : 5U54MD007592 -29, SUB IDS : 8320, 8321, 8317***
- ***NIH Project No: 1C06OD032074-01 & 3U54CA272167-02S1***

- National Institute of Health (NIH) U54's for the College's Research Core :
- NIH All of Us Data for Preliminary Data for Health Disparities Research
- NIH U54 Clinical Research
- NIH U54 Socio behavioral Research
- NIH Covid 19 U54 Supplements NIH C06 Grants for the Borderplex Biomedical Research Core NIMHD Grants for Health Disparities, RCMI Grants for Hispanic Health Disparities Paso Del Norte Health Research Grants