

MICHAEL NSIAH-NIMO

DATA ANALYTICS PROFESSIONAL

Lincoln, Nebraska • +1 (915) 246-5125 • michael20129648@gmail.com • [Website](#)

PROFESSIONAL SUMMARY

- Applied Statistician and Data Analyst with 7+ years of experience analyzing real-world EHR and insurance claims data to uncover cost variability, billing inefficiencies, and disparities in care—particularly within underserved U.S.–Mexico border populations. Skilled in pricing and risk analysis using GLMs, stratified regression, and time-series modeling.
- Proficient in pricing and risk modeling using GLMs, stratified regressions, and time-series analysis to identify high-cost drivers and simulate the impact of clinical or policy changes
- Adept at delivering interactive dashboards (Power BI, R Shiny) that empower stakeholders to explore patterns, optimize reimbursement strategies, and inform grant or operational decisions
- Equipped with strong DevOps and cloud engineering capabilities, including building CI/CD pipelines (GitHub Actions, Jenkins), provisioning infrastructure via Terraform and Ansible, and deploying containerized analytics environments on AWS (ECS, Lambda, RDS, S3).
- Known for bridging statistical rigor with cloud-native infrastructure to drive insights, scalability, and cross-functional collaboration.

SKILLS

- **Database:** Oracle, SQL Server, MySQL, Netezza
- **Data & Programming:** Python (pandas, scikit-learn, XGBoost), R, SQL, Bash, PowerShell
- **Databases:** Oracle, SQL Server, MySQL
- **Visualization:** Power BI, Tableau, R Shiny
- **Cloud & Analytics Platforms:** AWS (EC2, S3, RDS, Lambda, SageMaker), Amazon Connect
- **Infrastructure & DevOps:** Terraform, CloudFormation, Ansible, Docker, Kubernetes, GitHub Actions, Jenkins
- **System Tools:** Linux/Windows Admin, Active Directory, DNS, DHCP, VMware
- **Data Engineering:** ETL scripting (YAML, JSON, Git Bash), data wrangling, automation workflows

CERTIFICATIONS

- **AWS Certified Solutions Architect – Associate**
- **HashiCorp Certified: Terraform Associate (003)**

WORK HISTORY

Research Statistician – Data Analytics & DevOps, 01/2020 to Current

Border Biomedical Research Center (BBRC) – El Paso, TX

Data Analytics Focus

- Led modeling and experimentation using real-world data (EHR, insurance cost, clinical procedures) to uncover cost drivers and inform reimbursement decisions in health-related financial analytics.
- Applied GLMs, mixed-effects models, and causal inference techniques to evaluate interventions, predict procedure costs, and guide grant-funded public health strategies.
- Conducted A/B testing and time-series analysis to assess operational improvements, pricing changes, and treatment interventions.
- Built dashboards (Tableau, Power BI, R Shiny) to communicate insights across analytics and leadership teams.
- Designed ML pipelines using Python and R to predict patient risk, cost overrun, and resource utilization.
- Created NLP pipelines (e.g., GPT, LangChain) to summarize clinical notes narratives from EpicPatient.

DevOps & Automation

- Built CI/CD pipelines with GitHub Actions and Jenkins for analytics model deployment and monitoring.
- Provisioned AWS infrastructure (SageMaker, Lambda, RDS, S3) with Terraform, enabling reproducible analytics environments and automated deployment and scaling of data processing environments using Docker and Kubernetes.
- Implemented observability tools (Datadog, CloudWatch) to monitor data pipelines and improve system uptime.

FEATURED PROJECT

Healthcare Cost Analysis Using Real-World EHR and Insurance Claims Data

NIH-funded Project | BBRC, University of Texas at El Paso

Tools: Python, SQL, Power BI, R Shiny, GLMs, Stratified Regression

- Extracted and prepared multi-source patient data from NIH All of Us and PHIX Health Information Exchange, integrating EHR and claims data covering visits, labs, imaging, diagnoses, and payments.

- Modeled cost variability of common procedures (e.g., labs, imaging) across patient subgroups using GLMs and stratified regression, adjusting for demographic and comorbidity risk factors.
- Evaluated pre- and post-policy intervention effects using causal inference and time-series methods, enabling assessment of new clinic workflows and billing protocols.
- Designed interactive dashboards for clinicians and administrators to visualize high-cost drivers and equity gaps in access and reimbursement.
- Findings directly influenced clinic pricing transparency, funding proposals, and local care pathway redesigns to improve affordability and reduce disparity.

Data Analyst – Data Analytics and Experimentation | CoS Research, 07/2018 to 07/2020

University of Texas – El Paso, TX

- Performed data extraction, cleaning, and transformation of multi-source datasets (survey, transactional, and clinical records) to support exploratory analysis, predictive modeling, and machine learning pipelines.
- Conducted statistical and ML modeling in SAS, R, and Python, applying GLMs and tree-based methods (Random Forests, Gradient Boosting) to identify key drivers of health and socioeconomic outcomes.
- Developed and validated predictive models for outcome forecasting and risk stratification, leveraging cross-validation, feature selection, and dimensionality reduction techniques (PCA, LASSO).
- Led stratified and large-scale analyses to uncover disparities and trends, generating actionable insights for strategic planning and performance improvement initiatives.
- Contributed to experimental study design, hypothesis testing, and intervention evaluation using mixed-effects modeling and time-series forecasting approaches.
- Mentored junior analysts, providing training on SAS macros, Python/R machine learning libraries (scikit-learn, caret), and data visualization/BI tools to expand team analytics and modeling capabilities.

Cloud DevOps Engineer, 03/2022 to 07/2025

Technology Excellence Services – Remote (Freelance)

- Architected and administered secure, scalable AWS environments using EC2, S3, VPC, Route53, and IAM, optimized infrastructure for cost efficiency, high availability, and compliance with cloud security best practices
- Designed and automated CI/CD pipelines with GitHub Actions and CloudFormation, integrating Flyway for seamless database migrations and delivering faster, more reliable deployments across development, staging, and production
- Utilized CloudWatch and CloudTrail for comprehensive monitoring, setting alarms, and tracking user activities and API usage to maintain security and operational integrity
- Built cloud infrastructure for advanced workloads including machine learning and AI model pipelines, leveraging AWS services for large-scale data processing, model training, and inference while reducing compute and storage costs

System Administrator, 07/2014 to 07/2015

MTN Ghana – Winneba, Ghana

- Resolved complex system, application, and network issues through root-cause analysis, minimizing downtime and improving service reliability
- Automated and optimized system processes, boot sequences, and deployments, accelerating operational readiness.
- Provisioned, configured, and secured Linux servers with core infrastructure services (Apache, file servers, firewalls, directory services) to support scalable, high-availability environments
- Enhanced performance and stability by diagnosing and fixing DNS, NFS, and Apache issues, and optimizing TCP/IP

SELECT PORTFOLIO PROJECTS

Quantifying Sales Uplift with Causal Impact Analysis - [View project](#)

- **Context:** To measure the real-world impact of a marketing intervention on sales, I needed to isolate its effect from external trends and seasonal noise.
- **Action:** Applied CausalImpact (Bayesian structural time-series modeling) on customer sales data, estimating what sales would have been without the intervention.
- **Result:** Determined a statistically significant uplift, accurately attributing change to the campaign rather than external variation.
- **Growth / Next Steps:** Refined my ability to evaluate intervention effectiveness and ROI—directly applicable to modeling pricing or offer impacts in product analytics scenarios.

Assessing Campaign Performance Using Chi-Square(A/B) Test for Independence - [View project](#)

- **Context:** A marketing team needed evidence on whether two different mailer designs drove different customer responses.
- **Action:** Applied Chi-Square hypothesis testing to compare mailer performance across customer segments, assessing whether observed differences were statistically significant.
- **Result:** Identified which mailer significantly outperformed the other in select demographics—enabling informed targeting decisions.

- **Growth / Next Steps:** Reinforced my foundational experiment analysis skills, especially segmentation — a key capability when evaluating pricing experiments or user A/B testing in fintech settings.

Predicting Customer Loyalty Using ML (Regression) - [View project](#)

- **Context:** A grocery retailer wanted to forecast which customers are likely to stay loyal based on purchase patterns and demographics.
- **Action:** Built regression models using purchase behavior data to predict loyalty scores. Conducted feature selection and model tuning to improve prediction accuracy.
- **Result:** Produced a reliable model for forecasting customer retention—providing actionable insights to improve loyalty programs.
- **Growth / Next Steps:** Strengthened my ability to model behavior and retention—analogueous to predicting borrower loyalty or product acceptance in lending analytics.

Education

MSc: Statistics, 2017

University of Texas - El Paso

BSc: Actuarial Science, 2014

Kwame Nkrumah University of Science and Technology