Nichodemus Werre Amollo, MSc

Lead Data Scientist & Research Evaluation Specialist

Nairobi, Kenya | nichodemuswerre@gmail.com | +254 725 737 867 | gondamol.github.io

Professional Summary

Data Scientist and Research Lead with over 8 years of experience building scalable, evidence-based engines for social impact. Deep expertise in **Causal Inference**, **Theory of Change design**, and **Machine Learning** for development. Proven ability to translate complex behavioral data into actionable product strategies, having architected data systems for large-scale RCTs and longitudinal studies across East Africa. Hands-on expert in **R, Python, and SQL**, dedicated to solving the "causal puzzles" of youth employment and livelihood improvement.

Technical Skills & Tools

Domain	Tools & Techniques
Languages	Python (Pandas, Scikit-learn), R (Tidyverse, lme4, Shiny), SQL (BigQuery), Stata
Causal Inference	RCT Design, Difference-in-Differences (DiD), Propensity Score Matching (PSM), DAGs
Machine Learning	Random Forests, Clustering (K-Means), Predictive Modeling, NLP (Text Analysis)
Data Engineering	ETL Pipelines, BigQuery, ODK/SurveyCTO (Field Data), Git/GitHub
Visualization	Tableau, Power BI, Looker, ggplot2, R-Shiny Dashboards

Professional Experience

Lead Data Scientist & Program Evaluation Director

Georgetown University - Regional Evaluation Hub | *East Africa (Kenya, Uganda, Rwanda)*

Apr 2023 – Present

- **Strategic Architecture:** Architected the **Theory of Change** frameworks for 3+ major youth livelihood programs, translating abstract social science theories into measurable causal hypotheses and predictive features.
- **Impact Evaluation:** Directed the end-to-end execution of varied **RCTs and quasi-experimental evaluations**, utilizing Difference-in-Differences and PSM to isolate program impact from economic noise.
- **Operational Optimization:** Built a scalable data infrastructure from the ground up, implementing **ODK/SurveyCTO** systems across 15+ field sites (5,000+ participants) and automating reporting pipelines, reducing data latency by **60%**.

- **Heterogeneity Analysis:** Conducted advanced heterogeneity analysis to identify distinct impacts on marginalized rural girls. **Pivot 1:** These insights led to a fundamental redesign of the enrollment strategy, increasing uptake by **25%** in target demographics.
- **Stakeholder Influence:** Synthesized complex econometric findings into "Go/No-Go" policy briefs for funders (USAID, Gates Foundation), directly securing **$2M+** in follow-on funding based on evidence.

Senior Research Data Scientist
Kenya Medical Research Institute (KEMRI) | *Nairobi*

Apr 2020 – Mar 2023

- **Predictive Modeling:** Developed and deployed **ARIMA and exponential smoothing models** for health surveillance, reducing outbreak alert lag times from 2 weeks to 3 days.
- **Longitudinal Analysis:** Designed a multi-year cohort analysis tracking **2,000+ youth**, using survival models to identify early warning signs of livelihood dropout.
- **Data Pipeline Engineering:** Engineered robust ETL pipelines integrating unstructured qualitative data (field notes) with structured health metrics (DHIS2), creating a unified "360-degree" view of beneficiary health.
- **Capacity Building:** Mentored a team of 30+ junior analysts and data collectors, establishing a culture of code review, version control (Git), and reproducible research.

Research Data Manager & Evaluation Lead
JOOUST & VLIR-UOS Regional Program | *Kenya, Uganda, Rwanda*

Nov 2017 – Mar 2020

- **Causal Discovery:** Implemented rigorous **RCT evaluations** (randomization, fidelity monitoring) for skills training programs. Applied ANCOVA and Fixed Effects models to estimate treatment effects with high precision.
- **Dashboard Development:** Created interactive **Tableau/R-Shiny dashboards** enabling non-technical program managers to monitor KPIs in real-time, replacing static monthly PDF reports.
- **Agile Data Operations:** Managed the cleaning and harmonization of "messy" real-world data across 3 countries and 8,000+ observations, ensuring **99% data completeness** for final analysis.

Education

University of San Francisco
Graduate Certificate, Social Impact Analytics | *Jan 2025 – Dec 2025*

- **Grade:** Distinction
- **Focus:** Inferential statistics, RCT design, Machine Learning for program targeting, Causal Inference frameworks.

Jaramogi Oginga Odinga University of Science and Technology (JOOUST)
MSc in Statistics and Data Science | *Expected 2026*

- **Focus:** Advanced Statistical Methods, Data Mining, Computational Statistics.

Maseno University
BSc in Statistics | *2016*

Certifications

- **Research Methodology & Quantitative Methods** | Johns Hopkins Bloomberg School of Public Health (2019)
- **Ethics in Human Subjects Research** | Johns Hopkins Bloomberg School of Public Health (2019)
- **Advanced R for Data Analysis** | DataCamp
- **AWS Certified Cloud Practitioner** | Amazon Web Services