

Nichodemus Werre Amollo, MSc

Lead Data Scientist & Research Evaluation Specialist

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

SUMMARY

Data science and social impact analytics expert with 8+ years of experience translating complex research into actionable evidence for youth development and education programs across East Africa. Proven expertise in causal inference, impact evaluation design, and building scalable data systems that drive evidence-based decision-making. Strong command of R, Python, and SQL for statistical modeling, econometric analysis, and longitudinal data exploration. Hands-on experience designing RCTs, implementing field data collection systems, and communicating sophisticated findings to non-technical stakeholders. Passionate about leveraging rigorous analytics to unlock potential and measurably improve youth outcomes.

CORE COMPETENCIES

Research & Analytics: Causal Inference (RCTs, PSM, DiD) | Longitudinal Data Analysis | Econometric Modeling | Impact Evaluation Frameworks | Theory of Change Development | Hypothesis-Driven Feature Engineering | Survey Design & Implementation | Mixed-Methods Research

Data Engineering & Tools: Python (pandas, scikit-learn, statsmodels) | R (tidyverse, ggplot2, lme4, causalmle) | SQL (BigQuery, PostgreSQL) | Git/GitHub | Data Pipelines & ETL | Field Data Systems (ODK, SurveyCTO, KoboToolbox, DHIS2)

Data Visualization & Communication: Tableau | Looker | ggplot2 | Data Storytelling | Policy Briefs | Stakeholder Reporting | Executive Dashboards | Non-Technical Translation

Specialized Methods: Time Series & Forecasting | Survival Analysis | Heterogeneous Treatment Effects | Propensity Score Matching | Difference-in-Differences | Machine Learning (clustering, random forests, classification) | Thematic Coding & Qualitative Integration | Text Analysis & NLP

PROFESSIONAL EXPERIENCE

Lead Data Scientist & Program Evaluation Director

*Georgetown University - Regional Evaluation Hub / Apr 2023 – Present
East Africa (Kenya, Uganda, Tanzania, Rwanda)*

- Architected Theory of Change frameworks for 3+ major youth development and livelihood programs, translating behavioral science theory into measurable causal hypotheses to identify underlying mechanisms driving youth success and employment outcomes.
- Designed and executed advanced impact evaluations including RCTs and quasi-experimental studies (PSM, DiD) across agriculture, health, and livelihoods sectors; drove policy recommendations adopted by national governments and major funders.
- Built scalable data infrastructure from ground zero—implemented digital data collection systems (ODK, SurveyCTO) across 15+ field sites with 5,000+ participants; established automated dashboards and real-time monitoring systems reducing reporting time by 60%.
- Conducted heterogeneous treatment analysis to examine program impacts across demographics and rural/urban contexts; surfaced actionable insights that pivoted product design for youth targeting and resource allocation.
- Synthesized complex findings into compelling narratives for boards, donors (USAID, Global Fund, Gates Foundation), and government partners; translated statistical results into "Go/No-Go" decision frameworks.

Senior Research Data Scientist

Kenya Medical Research Institute (KEMRI) | Apr 2022 – Mar 2025

Nairobi / Central & Eastern Uganda (remote deployments)

- Designed longitudinal cohort analysis tracking 2,000+ youth across rural coffee-farming zones; built survival and competing-risk models to forecast health and livelihood outcomes; identified 4 key predictive factors informing early intervention targeting.
- Developed forecasting models (ARIMA, exponential smoothing) for health surveillance and early warning systems; reduced alert lag time from 2 weeks to 3 days, enabling faster policy response.
- Engineered data pipelines integrating health surveillance (DHIS2), agricultural data, and socioeconomic variables; cleaned and harmonized messy real-world datasets across multiple collection methods.
- Collaborated with community health networks to ground analytical frameworks in local context; co-designed data tools that balanced rigor with field feasibility, achieving 95%+ data quality standards.

Senior Research Data Manager & Evaluation Lead

JOOUST & VLIR-UOS Regional Program | Nov 2020 – Apr 2023

Kenya, Uganda, Rwanda

- Managed multi-country datasets across sustainable agriculture and health integration projects (3 countries, 8,000+ observations, 150+ variables); conducted longitudinal analysis to assess long-term livelihood gains and program sustainability.
- Implemented rigorous RCT evaluation (randomization, blinding, fidelity monitoring) for youth skills training program; applied causal inference methods (ANCOVA, DID) to estimate treatment effects; findings shaped national curriculum adoption.
- Built Theory of Change measurement frameworks collaboratively with program teams; designed surveys and instruments that operationalized social science concepts into measurable indicators.
- Developed interactive Tableau dashboards for M&E tracking visible to field teams, program managers, and government partners; established automated alerts for underperforming indicators.

Research Data Analyst – Rural Development & Youth Livelihoods

LERIS Hub | Sep 2017 – Nov 2020

Kenya / Uganda Border Regions

- Designed and conducted mixed-methods evaluations for youth employment and farmer accountability initiatives; integrated quantitative impact analysis with qualitative case studies to explain "what works" and "why."
- Analyzed multi-source datasets (surveys, administrative records, qualitative interviews) and generated 8+ evidence briefs for NGO partners and government ministries that directly informed policy and program design decisions.
- Applied causal inference techniques (propensity score matching, regression discontinuity) to estimate treatment effects in non-randomized settings; publications in peer-reviewed outlets.
- Supported survey design and questionnaire development drawing on economic theory and behavioral science; piloted instruments and iterated based on real-world data quality.
- Conducted longitudinal follow-up studies tracking youth outcomes over 18+ months; analyzed attrition patterns and applied multiple imputation to handle missing data in econometric models.
- Supervised data collection teams across two countries; implemented data quality protocols and trained 30+ enumerators on survey administration and field protocols.

Data Systems Analyst & Evaluation Specialist

Lake Region CDI | May 2016 – Aug 2017

Western Kenya

- Designed and rolled out digital data collection infrastructure (ODK) across climate- and agriculture-vulnerable regions; trained 40+ community data collectors and supervisors on systems, data quality standards, and field protocols.
- Conducted impact analysis on climate adaptation programs; merged survey data with administrative records to assess livelihood outcomes and identify key drivers of resilience in smallholder farming communities.
- Produced regional evidence briefs synthesizing climate vulnerability and agricultural productivity data; findings shaped county-level climate adaptation strategy and influenced resource allocation decisions.
- Built baseline data systems for 3+ development projects; established monitoring frameworks and dashboards enabling real-time reporting to government and donor partners.

EDUCATION

Master of Science, Statistics & Data Science • JOOUST (*In Progress, Expected Jun 2026*)

Bachelor of Science, Statistics (Honours) • University of Nairobi (2016)

PROFESSIONAL CERTIFICATIONS & ADVANCED TRAINING

Graduate Certificate, Social Impact Analytics (Distinction) • University of San Francisco | Jan 2025 – Dec 2025

Certificate, Research Methodology & Quantitative Methods • Johns Hopkins Bloomberg School of Public Health | Apr 2018 – Sep 2019

Certificate, Ethics in Human Subjects Research • Johns Hopkins Bloomberg School of Public Health | Jun 2019 – Jul 2019

Advanced R for Data Analysis • DataCamp | 2023

Data Science Professional Certificate • DataCamp | 2023

Google Data Analytics Specialization • Google Career Certificates | 2021

AWS Certified Cloud Practitioner • Amazon Web Services | 2022

Biomedical Research Ethics • CITI Program | 2024

Economic Evaluation • University of Washington | 2024

TECHNICAL SKILLS

Languages: R (Advanced), Python (Advanced), SQL (Advanced), Stata (Advanced)

Statistical Methods: Causal Inference | RCTs | PSM | DiD | IVs | Time Series | Survival Analysis | HLM | SEM

ML & Modeling: Classification | Clustering | Regression | Random Forests | Feature Engineering | Anomaly Detection

Data Tools: Python (pandas, scikit-learn, statsmodels) | R (tidyverse, ggplot2, lme4) | SQL (BigQuery, PostgreSQL)

Visualization: Tableau | Looker | ggplot2 | Power BI | Jupyter

Languages: English (Native/Fluent), Swahili (Fluent)