Cornelius **Kiplimo Tanui**

✉ kiplimocornelius@gmail.com | ⚙ corneliustanui | 🔗 Cornelius Tanui | 📍 Nairobi, Kenya

Statistician, Data Manager, Data Analyst, Data Scientist, Statistical Programmer, Data Configurations Engineer, Statistical Trainer.

Education

Master of Science, Research Methods	**Jomo Kenyatta University of Agriculture and Technology**, 2025 - 2026
Certificate, Lean Six Sigma Black Belt	**The Council for Six Sigma Certification (CSSC)**, 2025
Certificate, Monitoring and Evaluation in Global Health	**University of Washington**, 2025
Certificate, Data Protection	**Strathmore University, CIPIT Law**, 2024
Certificate, Medidata Rave EDC Certified Principal Investigator(link)	**Medidata**, 2023
Bachelor of Science, Applied Statistics with Computing, (1st Class Honours)	**Rongo University, Kenya**, 2014 - 2018

Professional Experience

Lead Clinical Data Science Programmer   **ICON Plc, Nairobi, Kenya**
13 Jan 2025 - Present

- Program data review listings in Databricks using Python (Pyspark, Pandas, etc) and SQL and visualizations using Spotfire.
- Review and provide input into all trial set-up documentation.
- Translate protocol and any applicable de-risking into the creation and maintenance of detailed technical specifications, mapping documentation including aCRF, data transfer agreements, etc.
- Creation of Data Review Model (DRM) based on client's standards and SDTM datasets based on current SDTM standards.
- Creation of back-end edit checks, listings, and visualizations to support risk-based central monitoring, e.g., Quality Tolerance Limits (QTLs), Critical to Quality (CtQ), etc., and ongoing clinical and safety data review.
- QC of all clinical programming deliverables.
- Perform unblinding, data-cut activities for planned reporting milestones and database locks.
- Development and validation of electronic submission packages that comply with submission requirements from FDA and other regulatory agencies as needed.
- Support the development of standards and drive implementation within the organization

Clinical Data Coordinator (Internal   **PHASTAR, Nairobi, Kenya**
01 Feb 2021 – 10 Jan 2025

- Clinical Database configurations in elluminate for ETL/ELT, Qlik Visualizations (CDA and CTOA), Graphical Patient Profiles, Patient Workbooks.
- Clinical Database configurations in Veeva CDB for ETL/ELT.
- Third party data validation against Data Transfer Agreement (DTA).
- eCRF and edit check programming in Medidata Rave, and Medrio.
- Generating and reviewing Work Instructions, Standard Operating Procedures, and Templates.
- Data standardizations (CDASH/SDTM and internal standards) using SQL, SAS, and R.
- Dashboard programming in R/Shiny, Tableau, Qlik Sense, and JReview to support data reviews and risk-based quality monitoring (RBQM) of trial sites.
- Web application programming using R/Shiny and Python/Dash with CI/CD deployment using Git and Docker.
- Training and onboarding junior staff, developing and reviewing training materials.

Statistician   **EED Advisory, Nairobi,**
                                                                                      Kenya
02 Jun 2020 – 29 Jan 2021
- Writing and reviewing proposals to bid for projects in line with ToRs, computing sample sizes, and advising on statistical methodological aspects of surveys.
- Programming data collection tools in ODK, KoBoToolbox, ONA, Survey Solutions
- Training enumerators and supervising regional, national, and local surveys.
- Developing statistical analysis plans, writing data cleaning and analysis codes in Stata and R to generate tables and charts.
- Leading data analysis teams and writing and reviewing reports for client submissions.

Data Analyst   **Digi-Plus,**
                                                                                      Thika, Kenya
06 Jan 2020 – 28 Apr 2020
- Data cleaning and analysis using R, Stata, and SPSS.
- Report writing.

Data Manager/Data Analyst   **AMREC Consultants,**
                                                                                      Kisumu, Kenya
13 Oct 2018 – 27 Dec 2019
- Designing data collection tools in ODK, KoBoToolbox, ONA, SurveyCTO, and CommCare.
- Training enumerators and supervising data collection exercises.
- Data cleaning and analysis using R, Stata, and SPSS.
- As a training coordinator, I developed training curricula for R and Stata, and trained and mentored over 50 University students on data analysis and data management.

Personal Projects_____

1. [2019 KPHC Analytics](#)   An R Shiny web dashboard application to demonstrate various findings of the 2019 Kenya Population and Housing Census. The app is auto-deployed using GitHub workflows and Docker CI/CD.
2. [miscellaneousR](#)   This is an R package hosted on [GitHub](#). The package contains convenient functions for performing simple to complex tasks appertaining to data wrangling.
3. [Quick Stats Reporter](#)   This R Shiny web application is an extremely useful tool for quick data analysis and reporting. It determines on its own what kind of table and graph to generate depending on the type of variable(s) selected.
4. [Fuzu Students Banding](#)   A Machine Learning approach to banding of university students in Kenya for award of scholarships. A more scalable alternative to Means Testing Instrument.

Memberships and Associations_____

1. **Member**   Kenya National Statistical Society (KNSS)
2. **Member**   International Biometrics Society – Kenya (IBS-Ke)
3. **Member**   Association for Clinical Data Management (ACDM)
4. **Part-time Statistical Trainer**   DTE Consultancy

Certifications_____

1. 2021   SAS Programming 1: Essentials, **SAS**
2. 2020   Microsoft Excel - Advanced Excel Formulas & Functions, **Udemy**
3. 2020   Microsoft Power BI - A Complete Introduction [2020 EDITION], **Udemy**
4. 2020   Science and Machine Learning with R, **Udemy**
5. 2020   Power BI for Data Science, **Udemy**
6. 2018   Certificate of Good Clinical Practice, **NIDA**

Selected Publications

1. **Tanui, C.** (2018, August 11). *Modelling consumer price index (CPI) in Kenya for the period 2000 to 2017 using autoregressive integrated moving averages—ARIMA(P, D, Q)*. Afribary. https://afribary.com/works/modelling-consumer-price-index-in-kenya
2. Ronoh, D., Mentis, D., Odera, S., Sego, T., & **Tanui, C.** (2025). *Application of GIS in sub-national energy planning in Kenya: Integrating primary data into a least-cost electrification model using OnSSET (case study of Narok County, Kenya)*. Practice Note. World Resources Institute. https://doi.org/10.46830/wripn.23.00040

Skills

1. **Languages**   English, Swahili, Kalenjin.
2. **Programming**   R, Python, Julia, SAS, Stata, SQL, SPSS, Docker, Git, HTML, CSS.
3. **GIS**   QGIS, Google Earth.
4. **Project Management**   MS Project, Git (GitHub, GitLab, Bitbucket), Jira, Confluence.
5. **Fields**   Clinical Trials, Public Health, WASH, Climate Action, Renewable Energy.
6. **Statistical Knowledge**   Time series analysis (ARIMA, GAM), survival analysis, panel data, predictive modeling (random forests, decision trees, SVM), non-parametric analysis, Bayesian inference, dimensionality reduction (SEM, PCA, MCA), generalized linear models (OLS, Poisson), design and analysis of experiments.
7. **Soft skills**   Excellent communication skills, good time management skills.

Referees

Dr. Vincent Were	**Mr. Edwin Nyagesoa**	**Ms. Pamela Adede**
Designation: Statistician	**Designation**: Statistician	**Designation**: Data Manager
Company: AMREC Limited	**Institution**: KEMRI/CDC	**Company**: PHASTAR
Mobile: To be provided upon request	**Mobile**: To be provided upon request	**Mobile**: To be provided upon request
Email: To be provided upon request	**Email**: To be provided upon request	**Email**: To be provided upon request