

# **REUBEN NYAMWAYA OMWENGA**

## **Statistician & Data Analyst**

+254 799 613 635 | [reubenomwenga5@gmail.com](mailto:reubenomwenga5@gmail.com) | Nairobi, Kenya

---

## **Technical Summary: Data Management, Quality Assurance, and Insights**

### **UN-Habitat – Strengthening Urban-Rural Linkages in Intermediary Cities Project, Kenya**

*Consultant: Data Collection Assistant | May – June 2024*

#### **Project Overview**

This project, aligned with UN-Habitat's Urban-Rural Linkages Guiding Principles and Framework for Action, aimed to generate evidence on integrated territorial development by examining migration patterns, access to services (e.g., water, markets, health), and economic flows between urban intermediary cities and surrounding rural areas. Data collection targeted five Kenyan intermediary cities Nyeri, Machakos, Kisii, Eldoret, and Nakuru focusing on household-level indicators to inform policies for reducing spatial inequalities and enhancing food systems, housing, and basic services. The initiative supports Kenya's Vision 2030 and the New Urban Agenda, emphasizing participatory, evidence-based interventions for sustainable urbanization.

Over 1,200 household surveys were collected to map urban-rural convergences, with a focus on vulnerable populations (e.g., rural migrants in urban peri-areas). My role involved full-cycle data processes, from tool design to validation, ensuring analytical robustness and comparability across diverse contexts skills directly transferable to multi-country programs like PASGR's AYPReS.

#### **My Role & Key Responsibilities**

- **Tool Design & Deployment:** Programmed multilingual (English & Swahili) digital questionnaires in SurveyCTO and ODK Collect, incorporating skip logic, geolocation, and real-time validation to align with research questions on service access and economic linkages.
- **Field Implementation:** Supervised 25 enumerators during tablet-based data collection in urban-rural transects; conducted 15% of interviews personally to ensure cultural sensitivity in Ekegusii-speaking areas.
- **Quality Assurance:** Established a Google Sheets-based real-time monitoring dashboard for daily checks on completeness, outliers, and consistency; resolved 95% of issues on-site.
- **Data Processing & Synthesis:** Cleaned and merged datasets using Python (Pandas for deduplication, NumPy for aggregation) and Excel (Power Query for transformations); generated summary statistics and visualizations for policy briefs.
- **Stakeholder Collaboration:** Worked with UN-Habitat's Policy, Legislation, and Governance Section (PLGS) team to translate findings into actionable insights, including intersections with food systems and rural revitalization.

## Key Achievements & Metrics

- **Efficiency Gains:** Completed data collection 4 days ahead of schedule, achieving a 98.7% completeness rate (exceeding project target of 95%).
- **Error Reduction:** Identified and corrected 312 skip-logic and entry errors in real-time, reducing post-collection cleaning time by 40%.
- **Impactful Deliverables:** Delivered a cleaned master dataset (n=1,200+ records, 150+ variables) with codebook and metadata; contributed to a preliminary policy brief highlighting urban-rural investment opportunities.
- **Scalability:** Tools designed were adaptable for French contexts, supporting UN-Habitat's pan-African work in countries like Senegal and Côte d'Ivoire.

## Sample Data Quality Dashboard (Excerpt – June 5, 2024)

*Real-time monitoring table showing survey progress by city (generated via Google Sheets linked to SurveyCTO).*

<b>City</b>	<b>Surveys Collected</b>	<b>Completeness Rate</b>	<b>Avg. Interview Duration</b>	<b>Key Issues Resolved</b>
Nyeri	252	99.1%	24 minutes	45 geolocation tags
Machakos	238	98.3%	26 minutes	62 skip-logic errors
Kisii	241	98.9%	23 minutes	78 language toggles
Eldoret	229	97.8%	28 minutes	51 outlier values
Nakuru	240	98.7%	25 minutes	76 duplicates
<b>Total</b>	<b>1,200</b>	<b>98.7%</b>	<b>25 minutes</b>	<b>312</b>

## Key Insights Delivered to UN-Habitat Team

Analysis revealed that households with cross-boundary members (urban workers residing rurally) are 2.4x more likely to invest in agricultural improvements, underscoring the need for integrated territorial planning to boost rural economies and reduce urban migration pressures. This finding, derived from regression modeling in Python (e.g., logistic regression on investment variables), directly informed recommendations for URL-GP localization in Kenya's intermediary cities.

## Tools & Methodologies Applied

- Survey Platforms:** SurveyCTO (primary), ODK Collect (backup).
- Data Processing:** Python (Pandas for cleaning/aggregation, Matplotlib/Seaborn for visualizations), SQL queries for subset analysis, Excel (PivotTables, VBA for automation).
- Quality Protocols:** Followed UN-Habitat's data standards for robustness, including cross-validation against secondary sources (e.g., Kenya National Bureau of Statistics).

This experience reinforced my expertise in ensuring data tools align with policy-relevant outputs, much like the analytical support needed for AYPReS. Full datasets, cleaning scripts, and Jupyter notebooks available upon request.

**Contact for Questions**

Reuben Nyamwaya Omwenga  
+254 799 613 635 | [reubrenomwenga5@gmail.com](mailto:reubrenomwenga5@gmail.com)