

MERCY KUBANIA

Nairobi, Kenya | +254 703725893 | mercykubz@gmail.com | [LinkedIn Profile](#)

DATA ANALYST | ANALYTICS ENGINEER

Data Analyst turned analytics engineer with 5 years in health and pharmacy data. Skilled in building scalable data pipelines from source to reporting, dbt models (3 years) with tests, CI/CD from dev to prod and scheduled runs, data migration, and patient-data de-identification, and in cutting manual data cleaning by about 90% with AI-assisted workflows.

TECHNICAL SKILLS

- **Languages:** Python (Advanced), SQL (Advanced)
- **Warehousing and ingestion:** Snowflake (cloud data warehouse, equivalent to BigQuery), Rivery
- **Transformation and version control:** dbt Cloud (3 years): models, tests, documentation; GitHub
- **Modelling and loading:** star schema, incremental loading, data migration
- **Data protection:** patient-data de-identification and anonymization; role-based access
- **BI and apps:** Streamlit (self-hosted), Looker, Metabase, Excel, Google Sheets
- **Automation and AI:** n8n, AI-generated insights

WORK EXPERIENCE

Data Analyst / Data Lead | Afyanalytics (Nairobi, Kenya)

Jan 2026 - Present

Early-stage healthcare analytics startup. Own the data stack from raw data to reporting in a small team serving hospital, pharmacy and retail clients.

- Architected the data platform cost-first: **Snowflake** warehouse consolidating **5+ sources and 1.5M records**, with dashboards in **self-hosted Streamlit**, so no per-seat BI licence cost compared with Power BI.
- Built and owned the **staged pipeline from raw data to analytics-ready tables** in Snowflake: data modelling, cleaning, and patient data **de-identification and anonymization**. Access runs on **role-based tiers**: roles define what each group can see, and users are assigned to roles.
- Built an **AI-assisted n8n workflow** that standardises HMIS clinical data and free-text doctor notes into **ICD-11 codes**. **Manual work dropped about 90%**, with human validation for the ~5% that need review.
- Designing the move to **self-hosted dbt Core** on Snowflake, replacing dbt Cloud, to keep costs low: rebuilding the tested **dev-to-prod CI/CD** flow from Maisha Meds, with environments, Git workflow, governance and team onboarding.
- Delivered **client analytics**: metrics for **3 hospitals** on what happened, why and what to do next, presented to facility stakeholders. Also a supermarket dashboard on churn, segmentation and revenue, and **AI-generated insights** that keep client dashboards current without manual rewrites.
- Solved a **new-branch stocking problem** for a pharmacy chain without pharmacy data: integrated distance to hospitals, gyms and salons, a regional wealth index and other chains' product data with branch data, then used **KNN** to recommend first-month stock (client demo).
- Introduced **Notion** as the team's hub for completed projects, data dictionaries and metric definitions. Supervised an intern and support 2 teammates.

Junior Data Analyst | Maisha Meds (Nairobi, Kenya)

Dec 2021 - Apr 2025

Analytics engineering scope: Snowflake, Rivery, dbt Cloud (3 years), GitHub.

- **Migrated** data from **3 source systems** (covering pharmacy, CRM, supply chain and inventory) into **one unified Snowflake ecosystem** (Rivery, dbt), **cutting fragmentation by 80%**. Validated with stakeholders to establish the source of truth and monitored existing dashboards for changes.
 - Designed and implemented **20+ dbt models** in dbt Cloud and GitHub, with every change **PR-reviewed** and promoted from dev to prod through **CI/CD**, dbt tests and alerts guarding data quality, and YAML documentation. **Scheduled dbt Cloud jobs** kept BI tools on current data, with failure alerts.
 - Maintained **data dictionaries** through data quality audits and managed dashboard access controls for **50+** technical and non-technical users.
 - Automated price and volume estimates for **30+ medical products** by modelling **8 years of trade data** in Snowflake and dbt, used by Business Development to target private-sector opportunities.
 - Defined country-specific **fraud and compliance indexes** (normalisation, weighting, composite scoring) implemented across **4 countries**, automating audits for **500+ pharmacies**, plus a performance analysis of **4,000+ pharmacies**.
 - Built SQL metrics and **10+ Looker and Metabase dashboards** for KPI tracking and external stakeholder reporting.
- Began as an intern, Sep 2021 - Dec 2021:** malaria dashboards and POS data cleaning in SQL and Python (pandas).

Data Science Intern | SANCOM Ltd (Nairobi, Kenya)

Jan 2021 - May 2021

- Applied clustering and hotspot detection to **875,000+** open San Francisco crime records.

EDUCATION

B.Sc. in Computer Science

Jomo Kenyatta University of Science and Technology, Juja

Sept 2015 - Nov 2019

Certificate in Data Science

Moringa School, Nairobi

Mar 2020 - Aug 2020