Jesse Ngugi Gichuki

Data Analyst | Data Scientist

Nairobi, Kenya | +254 799 117179 | ngugigichuki21@gmail.com

linkedin.com/in/ngugiauraa | github.com/ngugiauraa | ngugi.pages.dev | dev.to/ngugiauraa

Professional Summary

Data Analyst and Data Scientist with hands-on experience building ETL pipelines, machine learning models, and business intelligence dashboards. Trained through the LuxDevHQ Data Science Bootcamp and applying those skills at CoffeeKenya.Coop, turning messy transactional and member data into clean PostgreSQL databases, Power BI dashboards, and monthly board reports. Skilled in Python, SQL, Power BI, Excel, and Apache Airflow, with a focus on data cleaning, KPI tracking, and predictive modeling.

Skills

- **Programming:** Python, SQL, PostgreSQL, Java, JavaScript
- **Data Analysis & BI:** Power BI, DAX, Power Query, Microsoft Excel, dashboards, KPI tracking, data modeling, reporting, variance analysis, unit economics
- **Machine Learning:** Scikit-Learn, predictive modeling, classification, regression, feature engineering, model evaluation, class imbalance handling
- **Data Engineering:** ETL pipelines, Apache Airflow, data cleaning, data ingestion, database management, DBeaver
- **Libraries:** Pandas, NumPy, SciPy, Matplotlib, Seaborn, Anaconda
- **Tools & Cloud:** Git, GitHub, AWS, Azure, Google Cloud, DigitalOcean

Professional Experience

Data Analyst, CoffeeKenya.Coop, Nairobi, Kenya | 2024 – Present

- Built automated Python scripts to clean transactional and member datasets, handle missing values, and standardize variables, reducing manual data preparation by 35%.
- Developed Power BI dashboards and multi-entity Excel dashboards using Power Query and DAX to track producer deliveries, payments, and unit economics.
- Designed and managed relational PostgreSQL tables in DBeaver to store cleaned agricultural data and improve query performance.

- Produced monthly board reporting packs with variance analysis and trend insights for cooperative leadership.
- Supported financial planning and analysis (FP&A) through recurring reporting and KPI tracking.

Projects

Customer Churn Prediction (Machine Learning) | github.com/ngugiauraa/churn-predictor

- Built an end-to-end machine learning workflow in Python and Scikit-Learn to predict customer churn.
- Performed exploratory data analysis with Seaborn and Matplotlib to identify key churn drivers such as contract type and service type.
- Cleaned and converted raw text fields into model-ready numeric features, with missing-value imputation.
- Used a stratified train-test split to preserve class balance, and packaged training, feature, and prediction scripts with an API.
- Tuned decision threshold to hit 80%+ precision, reducing false positives in targeting at-risk users.

Sales Dashboard ETL Pipeline (Data Engineering) | github.com/ngugiauraa/sales-dashboard-etl

- Designed a data pipeline to ingest, clean, and load daily sales revenue data into an analytical database.
- Scheduled and monitored pipeline tasks with Apache Airflow DAGs to prevent upstream failures.

Education & Certifications

- **Professional Certificate in Data Science & Analytics**, LuxDevHQ | 2025
- **Data Modeling & Visual Analytics (Power BI and Excel)** | 2024