

SARAH INIOBONG UKO

Data Engineer • Pipeline Architect • Analytics Specialist

Nigeria | ukosarahiniobong@gmail.com | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

PROFESSIONAL SUMMARY

I turn data chaos into automated, reliable infrastructure. A Data Engineer with nearly 2 years of hands-on experience and AltSchool Africa training, I've built 9+ production pipelines and logged 4,950+ engineering hours designing ETL/ELT systems using Airflow, dbt, BigQuery, Snowflake, Python, and SQL that replace manual workflows with scalable, self-running pipelines. When data is broken, slow, or untrusted, I'm the engineer organisations call to fix it for good.

CORE COMPETENCIES

DATA ENGINEERING	DATA ANALYTICS	CLOUD & INFRASTRUCTURE
Apache Airflow (Orchestration)	Python: Pandas / NumPy	Snowflake Data Warehouse
dbt (Data Build Tool)	SQL: PostgreSQL / MySQL	Google BigQuery
ETL / ELT Pipeline Development	Power BI & DAX Measures	MongoDB & NoSQL Design
Data Modeling (Star / 3NF)	Advanced Excel / Power Query	Linux / Bash Scripting
Pipeline Orchestration & DAGs	Statistical Analysis & KPIs	Git & Version Control

KEY IMPACT METRICS

9+ Production ETL Pipelines <i>40% reduction in manual reporting time</i>	550+ Complex SQL Queries Authored <i>Stored procedures + optimised indexes</i>	4,950+ Verified Engineering Hours <i>18+ end-to-end data products delivered</i>
--	---	--

FEATURED PROJECTS: END-TO-END DATA ENGINEERING

1. CommercePulse Event Pipeline: Retail & Financial Transaction Analytics

Stack: MongoDB · Google BigQuery · Apache Airflow · Python · dbt

 **GitHub:** [Link](#)

- Engineered a hybrid batch/streaming event pipeline ingesting high-volume retail transaction data from MongoDB into BigQuery, architected to handle the scale of millions of daily POS and digital payment events directly mirroring the transaction infrastructure at Moniepoint.
- Applied creative engineering to resolve a critical reporting bottleneck: replaced a fragile manual export workflow with a fully automated Airflow DAG that eliminated 100% of ad-hoc data pull requests, cutting reporting latency by over 60%.

- Built end-to-end ELT models in dbt to transform raw commerce event streams into clean analytical layers, enabling stakeholders to track sales velocity, inventory turnover, and transaction anomalies in near real-time.
- Designed micro-batch streaming logic alongside scheduled batch loads so that pipeline latency could be tuned per use-case, a pattern applicable to Moniepoint's mixed real-time and aggregate financial reporting needs.

Outcome: Raw events → dbt-transformed warehouse → dashboard-ready data. Zero manual intervention post-deployment.

2. Code to Care ER Insights: Healthcare ETL & Operational Analytics

Stack: Python (Pandas, NumPy) · PostgreSQL · Power BI · DAX

 **GitHub:** [Link](#)

- Designed and deployed a Python-based ETL pipeline extracting fragmented operational data from multiple ER source systems, standardising schemas, and loading clean records into a PostgreSQL analytical warehouse.
- Automated 6 recurring data transformation jobs that previously required 3+ hours of manual preparation per reporting cycle, reducing operational reporting time by approximately 40% and freeing analysts for higher-value tasks.
- Engineered KPI models using DAX inside Power BI to surface ER triage time, patient throughput, and resource utilisation metrics enabling administrators to make scheduling decisions 25% faster.
- The pipeline architecture directly mirrors fintech operations: multi-source ingestion → standardisation → warehouse load → BI layer, the same pattern applicable to Moniepoint's transaction data workflows.

Outcome: Automated ETL reduced manual reporting effort by ~40%; operational insights available in near real-time.

3. Coffee-Insights Framework: SQL Optimisation & Inventory Intelligence

Stack: PostgreSQL · SQL (CTEs, Window Functions, Stored Procedures) · Power BI

 **GitHub:** [Link](#)

- Designed a fully normalised (3NF) relational database for multi-location coffee shop operations covering product inventory, supplier procurement, sales transactions, and automated restocking event triggers.
- Authored 100+ optimised SQL queries, complex CTEs, window functions, and correlated subqueries to track inventory depletion rates, flag low-stock events, and auto-generate daily sales summaries.
- Built stored procedures to schedule vendor procurement workflows and generate business health reports automatically, eliminating manual stock checks and reducing over-ordering incidents to near zero.
- The event-driven alerting and threshold monitoring logic is directly transferable to payment infrastructure: the same pattern underpins transaction anomaly detection and merchant balance monitoring in fintech platforms.

Outcome: 35% reduction in average query runtime; inventory discrepancy rate dropped to near zero post-deployment.

ADDITIONAL DATA ENGINEERING EXPERIENCE

EduHub MongoDB Platform: Advanced NoSQL Data Management

Stack: MongoDB · Python · Aggregation Pipelines · Schema Design

 **GitHub:** [Link](#)

- Architected a MongoDB document-model platform for a digital education service, designing flexible schemas for high-cardinality students and course datasets with dynamic attribute requirements.
- Engineered advanced aggregation pipelines using \$lookup, \$unwind, \$group, and \$facet operators to produce complex multi-dimensional reports on student engagement and course completion rates.
- Optimised collection index strategies that reduced analytical query execution times by 50%+ while improving write throughput for concurrent ingestion workloads.

MiniStore Analytics: E-commerce Database Design & Optimisation: Retail Transaction Infrastructure

Stack: PostgreSQL · Python · Advanced Indexing · Partitioning

 **GitHub:** [Link](#)

- Engineered a full-scale relational database for a Nigerian e-commerce platform handling high-concurrency retail transactions, the same class of merchant payment data infrastructure that underpins Moniepoint's POS network.
- Implemented table partitioning and composite index strategies that reduced query response times by 45% under peak transaction load.
- Designed automated database triggers for low-stock alerts, duplicate transaction detection, and fraud pattern flags at the DB layer, reducing application-layer processing overhead significantly.

Retail ETL Pipeline Implementation: Legacy Data Migration & Automation

Stack: Python · PostgreSQL · Pandas · CSV/Excel Sources

 **GitHub:** [Link](#)

- Automated 8+ ETL processes extracting transactional data from legacy CSV/Excel sources and loading clean, deduplicated records into a centralised PostgreSQL repository, reducing manual data handling by 40%.
- Engineered validation and error-handling logic catching schema mismatches, null violations, and duplicate keys at ingestion, achieving a 99%+ data quality pass rate across all production loads.

DATA ANALYTICS EXPERIENCE

YouTube Metrics & Trending Analysis: Executive BI Dashboard

Stack: Power BI · DAX · Python · SQL

- Led cross-functional analysis of global content trending data, authoring sophisticated DAX measures to compute rolling averages, YOY growth, and virality scores for C-suite stakeholder reporting.
- Replaced a weekly manual reporting cycle with an always-on Power BI dashboard, giving stakeholders real-time visibility into content performance metrics.

Supply Chain Operational Analysis: Logistics Intelligence

Stack: SQL · Power BI · Advanced Excel

 **GitHub:** [Link](#)

- Performed deep-dive SQL analysis on supply chain logistics data, uncovering critical lead-time bottlenecks and vendor inefficiencies that informed procurement strategy.
- Produced a BI report surfacing cost-saving opportunities that supported route optimisation decisions across regional distribution networks.

EDUCATION

AltSchool Africa - School of Data → *Primary Engineering Foundation*

Diploma in Data Engineering • One-Year Full Programme

March 2025 – March 2026

Federal University of Technology, Owerri (FUTO)

Bachelor of Science in Microbiology

2021 – Present

CERTIFICATIONS & TRAINING

- AltSchool Africa - Data Engineering Diploma • March 2025 – March 2026
- Axia Africa - School of Data (Data Analysis) • March 2025 – August 2025
- The Female Media Network (Tefem) - Data Analysis Bootcamp & Mentorship Certification • September 2024 – March 2025

References available upon request