

# CHUKWUDI OGBUTA

✉ Chukwudiogbuta@gmail.com    ☎ (234)-90-6605-8806    📍 330102, Warri, Nigeria    🔗 Portfolio website

🌐 LinkedIn Profile    🐙 Github Profile

## 👤 PROFILE

Highly skilled Analytics Engineer with 5 years of experience in data management, analysis, and engineering. Proven track record in building robust datasets, developing ETL pipelines, and performing database administration and monitoring. Strong ability to create efficient macros, templates, and utilities for data cleaning and reporting, with a focus on ensuring data integrity through meticulous validation and quality control. Experienced in data analysis, reporting, and cross-functional collaboration to support decision-making and optimize data workflows.

## 🧠 SKILLS

- Python Programming
- SAS Programming
- Apache Airflow
- Apache Spark
- Machine Learning for Data Engineering
- SQL and Relational Databases (MySQL, PostgreSQL, SQL Server, etc.)
- Microsoft Excel
- Microsoft Power BI
- SDTM & ADaM
- R Programming
- ETL/ELT Data Pipeline Orchestration
- Apache Kafka
- Bash Shell Scripting
- NoSQL Database (Cassandra DB, Mongo DB, CouchDB)
- AWS Data Engineering Services (QuickSight, Glue, RDS, S3 Bucket, RDS, etc.)
- Microsoft PowerPoint
- Tableau
- Jreview

## 📁 PROFESSIONAL EXPERIENCE

### Clinical Data Programmer, *Clinfocus*

2021 – present  
Nigeria

- Build SAS datasets from clinical database.
- Build ETL pipelines using AWS services.
- Database Administration and Monitoring.
- Develop SAS macros, templates and utilities for data cleaning and reporting.
- Utilize SDTM guidelines to build datasets.
- Communicate with an internal team to create deliverables for pharmaceutical and biotechnology clients.
- Implement analyses specified in the protocol or the Statistical Analysis Plan (SAP) while working with the project statistician.
- Ensure CRF meets the guidelines of the protocol and check for consistency and adequacy.
- Write SAS programs to generate tables, listings, figures and analysis datasets.
- Review CRF annotations and data specifications.
- Identify, write and program edit checks per the data validation plan or data management plan.
- Study management reports using SAS.
- Validate the programmed analysis datasets, tables, listing and figures.
- Perform analyses defined in the statistical analysis.
- Prepare clinical and statistical summary reports.
- Communicate with programming and statistics leads.
- Utilize SAS programming skills within protocol team and perform all programming required for clinical trial analysis and reporting.
- Perform quality control on final reports.
- Develop SAS coding and table templates for preparing, processing and analyzing clinical data.

- Developed and maintained data pipelines using AWS services such as S3 buckets, RDS (Relational Database Service), Glue, and Lambda functions to ingest, process, and analyze petroleum exploration and production data.
- Designed and implemented ETL processes to extract data from diverse sources, including drilling logs, seismic surveys, and production databases, and load it into centralized data repositories for analysis.
- Collaborated with geologists, engineers, and data scientists to understand data requirements and optimize data workflows for exploration, drilling, reservoir management, and production optimization projects.
- Built and managed data lakes and data warehouses to store and organize vast amounts of structured and unstructured data from oil wells, pipelines, refineries, and other upstream and downstream operations.
- Implemented data governance policies and procedures to ensure data quality, security, and compliance with industry regulations such as GDPR and SOX within the petroleum industry.
- Developed custom data processing scripts and workflows using Python, SQL, and AWS Glue to clean, transform, and aggregate data for downstream analytics and reporting.
- Created dashboards and visualizations using AWS QuickSight to monitor key performance indicators (KPIs), track drilling operations, and analyze production trends.
- Conducted performance tuning and optimization of database queries, ETL jobs, and data processing workflows to improve efficiency and reduce costs associated with data storage and processing.

## PROJECTS

---

### **Horizon Oil & Gas Batch ETL Pipeline, Streamlining Data Processing for Oil Exploration**

Led the development of a batch ETL pipeline to extract, transform, and load data from diverse oil wells. Enhanced data analysis capabilities for oil exploration and production teams, facilitating informed decision-making.

### **Generator Sensor Data Streaming Pipeline, Real-time Data Analysis for Predictive Maintenance**

Developed a real-time data streaming pipeline using Apache Kafka and Apache Spark to ingest and analyze sensor data from industrial generators. Implemented machine learning models for predictive maintenance, resulting in proactive equipment upkeep and reduced downtime.

### **Machine Learning for Predictive Maintenance, Utilizing Apache Spark for Industrial Asset Management**

Utilized Apache Spark and Python to analyze sensor data from generators, enabling predictive maintenance strategies. Implemented machine learning algorithms to forecast maintenance requirements and optimize equipment performance. Contributed to cost savings and improved operational efficiency.

## CERTIFICATIONS

---

### **Bachelor of Electrical Engineering, University of Benin**

CGPA: 4.2 out of 5.0 overall (Second Class upper division)

### **IBM Data Analytics with Excel and R Professional Certificate**

### **Google Data Analytics Professional Certificate**

### **Communication and Interpersonal Skills at Work Certificate, University of Leeds (Online)**

### **Business Management, King's London College (Online)**

### **Project Management Professional Certificate**

### **Customer Service and Relationship Management Certificate, Exford Global**