

# Henry Akwerigbe

## Software Engineer with Python and Data Analytical Skills

### SUMMARY

- 8+ years experience working with Python;
- 5 years of experience as a BI and 4 years of experience with Tableau;
- 8 years of experience with various data sets (ETL, Data Engineer, Data Quality Engineer);
- 3 years of experience with Amazon Web Services (AWS), Google Cloud Platform (GCP);
- Data Analytics/Engineering with Cloud Service Providers (AWS, GCP) - Experience working with MySQL, SQL, and PostgreSQL;
- Deep abilities working with Kubernetes (K8s);
- Hands-on scripting experience with Python; Microsoft Power BI, Tableau, Sisense, CI/CD principles, Data Validation, Data QA, SQL, Pipelines, ETL, and Automated web scraping.
- Pet web3 projects (solidity, wallet integration)
- Upper-intermediate English

### TECHNICAL STACK

<b>Main Technical Skills</b>	Python (8.5 yr.), Data Analysis (6 yr.), Google Cloud Platform (GCP) (4 yr.), Tableau (4 yr.), Microsoft Power BI (4 yr.)
<b>Programming Languages</b>	Python (8.5 yr.), JavaScript
<b>Python Frameworks and Libraries</b>	Flask, pandas, ScikitLearn, Python Pickle, Django Channels, PySpark
<b>Data Technologies/Analysis/Visualization</b>	Data Analysis (6 yr.), Data Testing (3 yr.), ETL, Apache Spark, Data Scraping, Data Modelling, Data Mining, Apache Airflow, Tableau (4 yr.), Microsoft Power BI (4 yr.)

<b>Databases &amp; Management Systems, ORM</b>	SQL, ETL, Apache Spark, MySQL, PostgreSQL, Microsoft SQL Server, AWS ElasticSearch, DynamoDB, RDBMS
<b>AI &amp; Machine Learning</b>	Machine Learning (SpaCy, ScikitLearn)
<b>Cloud Platforms, Services &amp; Computing</b>	Amazon Web Services (AWS) (3 yr.), Google Cloud Platform (GCP) (4 yr.), Heroku
<b>Amazon Web Services</b>	AWS ElasticSearch, AWS S3
<b>QA &amp; Test Automation</b>	Selenium Webdriver, Unit Testing
<b>Version Control</b>	Git
<b>Operating Systems</b>	Linux
<b>SDK, API, and Integrations</b>	RESTful API
<b>Deployment, CI/ CD, DevOps, Administration</b>	Pipeline, CI/CD, Kubernetes (K8s)
<b>Other Technical Skills</b>	Data Engineering (6 yr.), Team Collaboration (7 yr.), BIGData, Streamlit, Robotic Process Automation, Cronjob, Data pipelines / ETL, Parallelization, datasets, Machine, learning, web, RPA, Sisense

## WORK EXPERIENCE

### Python Engineer, Profile Import & Data Scraping (NDA)

Industry: HRTech

May 2023 - June 2023

**Overview:** Microservice for job resumes (profile) and Job description parser and scraper functionality that includes integration with LinkedIn, a popular workable, glassdoor-like platform, Google Docs, PDF & Word parsers.

#### Responsibilities:

- Web Scraper
- Data Parser for PDF, Word, Google Docs
- Machine Learning and text/content recognition
- API

- Automation testing for importing, high-load

**Technologies:** AWS, Restful API, Python, Pytest, Alluire, JavaScript, Docker, Kubernetes

## **Power BI Engineer (NDA)**

Industry: Real Estate

January 2023 - May 2023

**Overview:** Startup that revolutionizes the home equity market in the US. Our team is working on providing outstanding BI services with accessible data to decision-makers as well as streamlining the current services and their effectiveness.

### **Responsibilities**

- Design and develop Tableau dashboards;
- Produce well-designed, efficient code by using the best software development practices;
- Perform code reviews for compliance with the best engineering practices, coding standards, and quality criteria set for the projects;
- Provide suggestions to improve the architecture, and coding practices, build/verification toolset, and solve customer problems.

**Technologies:** Tableau, SQL, Snowflake

## **Data Engineer, Data management platform / Amazon E-Commerce Aggregator (NDA)**

Industry: E-Commerce

2020-August 2022

**Overview:** A next-generation consumer goods company reimagining how the world's most-loved products become accessible to everyone. We use a deep understanding of rankings, ratings, and reviews to identify and acquire quality brands and use world-class expertise and data science to make their products better or create new ones to meet changing customer demand.

### **Responsibilities:**

- Use Sisense to build dashboards for tracking updates to selected Amazon store brands for determined time periods. I used the interactive SQL palette

for querying the tables to filter the needed information (columns) to be displayed in the dashboard. This dashboard provides the data engineering manager with the necessary information to make decisions on store brands.

- Create and support ELT data pipelines built on Snowflake and DBT while ensuring high-quality data
- Develop and deploy data warehousing models, and support existing processes/ETLs (extract/transform/load), and functions (in Python/SQL/ DBT) in a cloud data warehouse environment using Snowflake, AWS services
- SQL statements and developing in Python
- Design and develop data pipelines (DAGs). Automation tests.

**Technologies:** Sisense, AirFlow, Snowflake, Python3, AWS S3, Hadoop, Spark, GitLab CI/CD, Kubernetes, LDAP.

## **Software Developer, AI Project (Valuesoft AI)**

Industry: Financial Services

September 2020 – July 2021

### **Responsibilities:**

- Develop automation workflows with RPA (UiPath).
- Set up and manage web-based cloud services on AWS EC2.

## **Python Instructor, (NDA)**

Industry: EduTech

March 2020 – March 2021

### **Responsibilities:**

- Teach Python programming to students.
- Develop a curriculum to be used for teaching Python programming and data analysis.

## **Python Developer, IoT-leveraged agricultural tech company (Sensegrass)**

Industry: AgriTech

A project on monitoring and reporting sample data from agricultural plants on a field of land.

May 2020 – August 2020

### **Responsibilities:**

- Hands-on setting up, maintaining, and deploying services to AWS EC2.
- Automated web scraping of data from webpages using Selenium.
- Carried out multi-processing and parallelizing of code with PySpark.
- Used Spark for 2 cases of data processing in an ELT phase:

1. Data was collected from drones and other specialized bots were used to physically survey the land area and take samples from the soil and air for properties such as soil pH, moisture content, specific gravity, etc for different types of crops planted on the field. This data was received gotten in real-time, and placed on a queue to be loaded into AWS DynamoDB. The transformation involves converting some data properties from the queue such as temperature from degree Celsius to the Kelvin scale, moisture content from cubic centimeters to cubic meters, etc. The transformed data is then loaded into AWS s3.

2. Process large batch data averaging 10 million rows with spark: There were cases where I had to transform data on a different database containing historical data to consolidate the currently maintained tables in another database. The historical data contains millions of rows of IoT-generated values. To optimize speed and memory usage for transformation, I used Python's implementation of Spark (Pyspark) to carry out the same transformation technique on the batch data to backfill the current table in the database.

## **IT Analyst (FieldworkAfrica)**

Industry: Market Research

July 2019 – 2021

### **Responsibilities:**

- Developed data visualizations on PowerBI and Tableau to track areas of high and low drink consumption to establish which areas are potentially viable to push a new drink to.
- Provided daily and historical data reports and visualizations to the technical director. Daily and historical reports included tracking the

coverage of data collection in geographical areas, and providing updates on data quality checks and target data samples.

- Developed and maintained cloud services on the Google cloud platform.
- Developed questionnaire scripts on ODK for market research.
- Led a data collection team of 10 people.
- Performed data analysis using data tools, visualizations, and dashboards.
- Used PowerBI and Tableau to provide daily and historical data reports and visualizations to the technical director. Daily and historical reports included tracking the coverage of data collection in geographical areas, providing updates on data quality checks and target data samples.

## **Python Developer, (NDA)**

Industry: IT Services & IT Consulting

Jan 2019 – April 2019

### **Responsibilities:**

- Working on websites back-end with Flask and Django.
- Maintaining SQL databases for proper scaling.
- Ensuring proper test units are integrated to promote clean codes.

## **Data Science Trainee, (DSN)**

Industry: Data Science Academic Research

2017

### **Responsibilities:**

- Implemented optimization algorithms.
- Carried out analytics with Microsoft Azure for prediction models.
- Generated various visualization models for data analytics with Power BI and Seaborn.

## **Campus Ambassador, (ScholarX)**

Industry: EduTech

July 2016 - December 2016

### **Responsibilities:**

- Promoted ScholarX mobile App on designated campuses and social

platforms for the company achieving 1000 downloads on the Google Play Store.

## **Engineering Intern, (Nivafer Steel Co.)**

Industry: Oil & Gas / Construction

April 2015 - July 2015

### **Responsibilities:**

- Assisted in a supervisory management role and design engineering in various structural steel processes.

### **Tableau Experience Highlights:**

#### 1. Real Estate Project:

- Developed and maintained predictive algorithms for US house prices using machine learning techniques such as regression and classification
- Created interactive data visualizations for real estate agents and investors using Tableau
- Analyzed a variety of data points on comparables for single-family homes and condos, including location, property age, and amenities
- Assessed factors like ARV (After Repair Value), square footage, year built, number of beds and baths, garages, and local market conditions
- Developed user-friendly dashboards to display real-time market trends and property values, enabling investors to make informed decisions quickly
- Collaborated with a team of data scientists and engineers to continuously improve algorithms and visualizations

#### 2. Tableau Specialist in Market Research Project:

- Utilized Tableau for comprehensive daily and historical data reporting and visualization to support decision-making processes.
- Provided data insights and visualizations to the technical director, enabling a better understanding of market dynamics and trends.
- Created a range of custom dashboards for daily and historical reports that

covered various aspects such as sales, customer demographics, and product performance.

- Monitored and analyzed data collection coverage in target geographical areas to ensure accurate representation of the market.
- Conducted regular data quality checks, including data validation and cleaning, to maintain high data accuracy and reliability.
- Collaborated with data engineers and analysts to optimize data collection methods and improve overall data quality

## **EDUCATION**

- **Yaba College of Technology**  
2018 - 2019  
Higher National Diploma (HND), Mechanical Engineering
- **Yaba College of Technology**  
2013 - 2016  
National Diploma (ND), Mechanical Engineering

## **CERTIFICATIONS**

- **Python Developer Certificate** (Sensegrass) 2020
- **Google Scholarship Android Basics** - 2018
- **Certificate of Completion** (DSN 2nd Data Science Boot Camp) 2017
- **Certificate of Proficiency in Human Resources and Skill Acquisition**  
2014
- **Certificate of Participation ACM** (Association for Computing Machinery)  
2017
- **Big Data Foundations** (Level 1) 2017
- **Data Science Foundations** (Level 1) 2016