**UBIOWORO RUKEVWE**

Hungary|+36304889889|Ubiorukevwe@gmail.com|[linkedin.com](http://www.linkedin.com/in/rukevweubioworo) |[Github.com](http://www.github.com/rukevweubio)

**PROFESSIONAL SUMMARY**

Dynamic and results-oriented Data Engineer with a proven track record in architecting and deploying high-performance data solutions. Leveraging expertise in Python, SQL, and Airflow, I specialize in creating scalable data pipelines that streamline operations and drive business intelligence. My proficiency in Azure Cloud and Azure Databricks empowers me to design innovative data models and analytics frameworks that transform complex datasets into strategic insights. With a full-stack approach to data engineering, I excel in every stage of the data lifecycle from extraction and transformation to modeling and deployment. Passionate about harnessing emerging technologies, I am committed to optimizing data processes that enhance decision-making and propel organizational growth.

**CORE COMPETENCIES**

* **Data Analysis**: SQL, Python, data profiling, data visualization
* **Data Integration**: ETL pipelines, data modeling, cloud-based solutions (Azure, GCP).
* **Technical Skills**: SQL, Python, Apache Airflow, Azure Data Factory, Snowflake, Mageai, database administrator
* **Soft Skills**: Communication, problem-solving, teamwork, attention to detail

**TECHNICAL SKILLS**

* Programming Languages Python, Sql
* ETL Tools: Apache Airflow, Azure Datafactory, Azure Dataflow ,Mageai
* ER Design:Draw.io,Lucid chart,Quickdb
* Warehouse: Snowflake, motherduck
* Big Data; pyspark ,snowpark
* Containerization: Docker
* Data modeling; Data build tools (DBT)
* Cloud: Google Cloud , Azure
* Database: MySQL, PostgreSQL, NoSQL
* Other Software such, Microsoft Power BI, DAX,
* Collaborative tools Jupyter Notebook, VSCode, Google Colab
* Version Control: GitHub, Git
* Linux :linux system administration and bash scripting

**PROFESSIONAL EXPERIENCE**

**Data Engineer** – Esque kft, Hungary **Sep. 2023 – Present**

* Utilized Azure Data Factory and google cloud (GCP) to optimize cloud-based ETL data pipelines workflows, enhancing efficiency by 30% and streamlining data flow from diverse sources.
* Designed and implemented robust and scalable data pipelines to automate the ingestion, processing, and delivery of structured and unstructured data, improving pipeline reliability by 30%.
* Improved data accuracy by 25% through the implementation of transformation and cleaning processes, ensuring high-quality data for advanced analytics and reporting.
* Designed and deployed large-scale distributed applications, automating the ingestion and processing of structured and unstructured data, reducing manual workload by 25%.
* Collaborating with a cross-functional team of senior data scientists and analysts in working on a variety of complex and challenging problems
* Implemented Linux operations for data workflows, including shell scripting and command-line automation, streamlining processes by 15%.
* Optimized cloud-based ETL processes using Azure Data Factory and Data Flow, increasing workflow efficiency by 30% and streamlining data flow from various sources.
* Converted and loaded 500,000+ records daily from JSON, CSV, and Parquet formats into a snowflake data warehouse, achieving 99.9% data accuracy and reducing ETL runtime by 20% using optimized schema designs.

**Data Engineer – Delta steel Company ,Nigeria** April. 2023 – August 2023.

* **Boosted data fetching efficiency by 30%** through query optimization and indexing, while maintaining ETL pipelines that ingested data from multiple sources using Python, Azure Data Storage Gen 2, and Azure Storage Blob
* **Built and maintained scalable ETL pipelines** with Apache Spark, Airflow, and Azure Data Factory, processing large datasets efficiently and cutting data processing time by 20%.
* Supported large-scale distributed applications, enabling the seamless integration of multiple data sources with PostgreSQL ,MySQL and MongoDB databases .
* **Automated data ingestion and transformation**, enhancing real-time analytics by 40% and reducing manual tasks by 80%
* Conducted SQL-based data profiling, analysis, and extraction, improving data accuracy by 25% and enabling better decision-making.
* Automated the provisioning of 50+ Google Cloud Storage buckets and BigQuery datasets using Terraform, reducing infrastructure deployment time by 40% and ensuring 100% compliance with predefined security and naming standards.

**EDUCATION**

* **B Eng. Tech Mechanical Engineering**– Ahmadu bello University Kaduna State, Nigeria **2013**

****KEY ACHIEVEMENTS****

* **Improved data accuracy by 25% through the implementation of data transformation and cleaning processes.**
* **Automated data workflows using Linux shell scripting, reducing processing time by 15%.**
* **Designed and deployed scalable data pipelines, improving data accessibility and reliability by 30%.**