Emmanuel Onwubuya, Data Engineer

Hamburg, Germany, +4915739814757, emyraeleson@gmail.com

LINKS

LinkedIn, GitHub, Portfolio

PROFILE

Dynamic Data Engineer with over 4 years of experience in developing scalable data pipelines and managing cloud-based data platform migrations. Expertise in leveraging Python, SQL, and Spark to create robust data solutions that drive analytics and machine learning initiatives. Proven ability to enhance pipeline performance and ensure data integrity while translating complex business requirements into actionable insights. A strong advocate for innovation, currently exploring the integration of AI and large language models to optimize data workflows. Committed to delivering high-quality, analysis-ready datasets that empower data-driven decision-making.

EMPLOYMENT HISTORY

2024 — Present

Data Engineer, Wefra Life — Frankfurt,Germany

- Set up API data streams from Adverity and other third-party platforms, ingesting data into Azure Data Lake to handle diverse formats and sizes, and enabling seamless downstream processing.
- Migrated 200+ ETL models from dbt and Snowflake to Databricks, optimizing pipeline performance and ensuring scalability for evolving business needs; reduced query latency and improved data retrieval speeds by ~45–50%, significantly enhancing analytics efficiency.
- Engineered Databricks Bronze–Silver–Gold layers with reusable functions and CDC logic, improving automation, reducing redundancy, and ensuring robust data lineage; integrated Data Vault modeling principles to support scalable, auditable, and business-aligned data structures.
- Automated Databricks workflow deployment with Terraform, embedding monitoring and schema-change alerts to increase reliability and reduce downtime.
- Applied prompt-engineering and LLM-driven workflows to accelerate analytics and data manipulation tasks, improving productivity and enabling AI-assisted reporting pipelines.

2022 — 2024

Data Engineer, Accenture — Hamburg,Germany

- Led the development of a robust data warehouse solution by transforming large-scale, heterogeneous datasets into scalable PySpark and SQL pipelines, enabling unified data access across departments.
- Partnered with stakeholders across business, analytics, and executive teams to align data products with strategic KPIs, ensuring adoption and maximizing downstream business impact.
- Automated complex ETL workflows and scheduled data building processes using Skywise and Palantir data tools, improving data freshness, model readiness, and reducing manual processing efforts by over 30%.
- Optimized database performance and query execution through partitioning, indexing, and refactoring logic, reducing data load times by up to 40% and enhancing end-user responsiveness.

2021 — 2022

Junior Data Engineer, Domicil Real Estate Group — Munich,Germany

- Automated recurring ETL workflows using Python and shell scripts, improving data processing efficiency and enabling faster delivery of business insights.
- Collaborated with the finance team to implement real-estate acquisition and sales models, enhancing flexibility and contributing to an increase in successful acquisitions.
- Developed scalable data pipelines integrating APIs and third-party data sources into cloud data warehouses, ensuring reliable ingestion.

EDUCATION

Master in Data Analytics and Machine Learning, Universität Hildesheim, Germany

Bachelor of Engineering in Information Systems and Technologies, Voronezh State University, Russia

LANGUAGES

English
(Native)

Russian
(Conversational)

German
(Conversational)

SKILLS

Technical Skills

Python, SQL, PySpark, JavaScript, Bash, Linux, LLM Prompting, Airflow, Git, Terraform, MCP, Docker, CI/CD, REST APIs.

Data Platforms, Data Architecture & Cloud Infrastructure

Databricks, DBT, Snowflake, Supabase, Adverity, Data Vault,Star Schema, Medallion Architecture (Bronze–Silver–Gold), AWS, GCP, Azure, Dagster, Grafana, Tableau, Looker, Power BI.

Strengths

Analytical Thinking, Problem-Solving, Stakeholder Management, Agile Collaboration,Data Architecture & Modeling, Pipeline Optimization,Time Management, Ethical Judgment, Teamwork.

PROJECTS

Ufindar – An AI Powered University Search Web App

Created Ufindar, an AI-powered university search engine enabling students to discover universities through natural-language queries rather than rigid filters which integrated Lang-chain with Supabase and embedding-based semantic search to map user questions directly to relevant universities, programs, and metadata delivering contextual, accurate results. Deployed the working prototype here