

EMEKA CHIDOKA

Data Engineer

Wolverhampton, United Kingdom | chidokaemeka@gmail.com | +44 7448 786 966 | www.chidokaemeka.com |
linkedin.com/in/emeka-chidoka-a4a633158 | github.com/Emzzy123

PROFESSIONAL PROFILE

Data Engineer with a background in software development, data analytics and regulated healthcare operations. Hands on experience building cloud based data pipelines, ETL and ELT workflows, analytics warehouses and reporting ready datasets using Python, SQL, AWS, Azure, Snowflake, PySpark, dbt, Airflow, Docker and GitHub. Strong understanding of data modelling, data quality, API integration, web scraping, data lineage, orchestration and stakeholder reporting. Able to combine software engineering discipline, analytical thinking and clear documentation to deliver reliable data solutions for business users and technical teams.

CORE TECHNICAL SKILLS

Programming and Querying: Python, SQL, Pandas, NumPy, PySpark, API calls, web scraping, CSV, JSON

Cloud Data Engineering: Azure Data Factory, ADLS Gen2, Azure Databricks, Azure Blob Storage, Azure Synapse concepts, Azure SQL concepts, Azure Key Vault, AWS S3, AWS Glue, Snowflake, cloud storage, cloud pipeline patterns

Warehousing and Transformation: Snowflake, dbt, ETL, ELT, staging models, intermediate models, marts, snapshots, data lineage

Data Modelling and Quality: dimensional modelling, fact tables, dimension tables, star schema, data validation, dbt testing, row count checks, duplicate checks

Orchestration and DevOps: Apache Airflow, Docker, Git, GitHub, GitLab, README documentation, environment configuration

Analytics and Databases: Power BI, Tableau, Excel, PostgreSQL, MySQL, MongoDB, dashboard reporting, KPI analysis

PROFESSIONAL EXPERIENCE

Data Engineer | Amdari | Contract | Remote | Apr 2026 to Present

- Support cloud based data engineering projects using Azure, AWS, Python, SQL, PySpark, Snowflake, dbt, Docker and API integration.
- Build and support ETL and ELT workflows that move raw data into cleaned, validated and analytics ready datasets.
- Use PySpark and SQL for scalable data transformation, data cleaning and preparation of structured datasets.
- Develop dbt models across staging, intermediate and mart layers to improve reusability, testing and maintainability.
- Assist with API calling, data validation, pipeline documentation and data lineage to improve reliability and traceability.
- Collaborate with team members to troubleshoot pipeline issues, communicate blockers and support delivery of reporting ready outputs.

Data Engineer | 10Alytics | Contract | Remote | Sept 2025 to Apr 2026

- Completed hands on data engineering training focused on Python, SQL, AWS, PySpark, dbt, Snowflake, GitHub, web scraping, API integration and capstone projects.
- Built end to end data engineering projects covering data ingestion, transformation, modelling, orchestration and reporting.
- Used Python for data extraction, automation, file handling, API calls, web scraping, data cleaning and workflow scripting.
- Applied SQL to transform datasets, write reusable analytical queries and prepare reporting ready tables.
- Worked with AWS concepts including S3, Glue and cloud based storage patterns for ingestion and processing workflows.
- Used PySpark to process larger datasets, apply transformations and understand distributed data processing concepts.
- Built dbt and Snowflake models including staging, intermediate and mart layers, using modular SQL to create clean analytics outputs.
- Created fact tables, dimension tables and aggregate reporting models to support business intelligence and dashboard reporting.
- Implemented data quality checks, duplicate checks, row count validation and transformation testing to improve trust in pipeline outputs.
- Used Git and GitHub for version control, README documentation and professional portfolio presentation.

Data Specialist | 10Alytics | Full time | Remote | Jan 2024 to Sept 2025

- Analysed large datasets using Python, SQL and Excel to identify trends, support reporting and produce actionable insights.
- Designed Tableau and Power BI dashboards to visualise KPIs and support data driven decision making.
- Applied predictive analytics and statistical analysis techniques to support forecasting and business performance analysis.
- Strengthened practical knowledge of data cleaning, exploratory data analysis, data visualisation and stakeholder communication.

Admin and Support Staff | Cygnet Healthcare Ltd | Wolverhampton, UK | May 2023 to Present

- Maintain accurate digital and physical records in a regulated healthcare environment, supporting compliance, audits and operational reporting.
- Coordinate data entry, documentation and information flow across clinical teams while ensuring accuracy, confidentiality and consistency.
- Support reporting, record keeping and administrative workflows requiring strong attention to detail and reliable data handling.

- Contribute to improved workflow efficiency by organising information clearly and supporting communication between teams.

PHP Developer | Retortal Ltd | Remote | Aug 2022 to Apr 2023

- Developed and maintained Symfony based web applications, improving reliability through continuous integration and delivery practices.
- Supported production troubleshooting, root cause analysis and incident remediation for business critical web applications.
- Re architected web application setup on AWS Cloud, improving performance and scalability for ecommerce environments.
- Worked with Git, databases, APIs and structured development workflows to deliver stable technical solutions.

Database Specialist | Noaj Empire | Contract | Nigeria | Sept 2019 to Oct 2020

- Cleaned, validated and processed raw datasets to improve data integrity and reporting quality.
- Supported database administration tasks including maintenance, optimisation, troubleshooting and access to reliable data.
- Developed and optimised SQL queries to extract meaningful insights from operational databases.
- Produced reports and analysis to support business decisions and database improvement activities.

SELECTED DATA ENGINEERING PROJECTS

Azure Data Engineering Pipeline: Built a cloud data pipeline using landing, bronze, silver and gold layers, external locations and transformation logic for analytics ready outputs. Tools: Azure Data Factory, ADLS Gen2, Databricks, dbt, Synapse concepts, GitHub.

Snowflake and dbt Analytics Warehouse: Created raw, staging, analytics and reporting schemas with dbt models for cleaned staging views, dimensions, facts and aggregate reporting tables. Tools: Snowflake, dbt, SQL, Power BI.

Self Service Data Pipeline Platform: Collaborated on a data engineering platform and contributed to features in my section, including a configurable pipeline framework for CSV and API ingestion, metadata tracking, run logging, quality checks and generated Airflow DAGs. Tools: Python, DuckDB, Airflow, YAML, Docker, GitHub.

Airflow Wikipedia Pipeline: Orchestrated extraction, transformation and loading of Wikipedia stadium data using Airflow tasks and Python pipeline functions. Tools: Apache Airflow, Python, Docker, CSV, Azure storage concepts.

EDUCATION

MSc Computer Science, University of Wolverhampton, UK | Merit | 2020 to 2021

Postgraduate Diploma in Innovative Technology Engineering, Waterford Institute of Technology, Ireland | Second Class Upper | 2018 to 2019

BSc Information System Science, Regent University College of Science and Technology, Ghana | Second Class Upper | 2011 to 2015

INTERESTS AND HOBBIES

Cloud data engineering, open source learning, technical documentation, data communities, competitive gaming, esports strategy and continuous professional development.