

## NIXON NG

EMAIL: [nixonngzh@gmail.com](mailto:nixonngzh@gmail.com) | MOBILE NUMBER: +447842763752 | GITHUB: <https://github.com/nixonngzh>

### PROFILE

Results-driven Data Engineer with 2 years of consultant experience transforming commercial data ecosystems. Slashed migration timelines by 67% while ensuring 99.99% data accuracy for FTSE 100 clients. Architected and implemented scalable ETL/ELT pipelines processing 20GB+ weekly across AWS and Azure. Currently advancing expertise through MSc Data Science with focus on Big Data, Machine Learning, and AI to deliver next-generation data solutions.

### PROFESSIONAL EXPERIENCE

#### Data Engineering Consultant, Kubrick Group, United Kingdom

Jan 2023- Sept 2024

##### Overview

- Built and managed data pipelines and airflow jobs for client company, resulting in 300% increase in data loading/checks efficiency for client product suite.
- Developed advanced skills in Data Modelling, SQL, Python, and Data Wrangling techniques.
- Utilized an agile test-driven development approach to develop robust scripts using automated Git CI/CD pipelines.
- Produced robust end-to-end data pipelines using cloud technology such as Azure Data Factory and Spark.

#### Data Engineer, IAG Loyalty (Client)

- Performed data quality analysis and generated client report suites, ensuring data discrepancy rate below 0.01% through automated validation using Tableau and SQL.
- Orchestrated Airflow DAGs on data refreshes of 20GB/week, using Matillion and Python scripts with AWS Lambda, Glue, EC2 & Redshift.
- Productionized Matillion ETL data flow pipeline jobs from AWS into Snowflake using Data Vault design model.
- Managed Github DDL, Proc, UDF SQL scripts for Unit-testing on Airline customer data.

#### Data Project Lead (Kubrick Internal)

- Developed an AI-powered chatbot using LangChain, ML clustering, and LLMs to enhance customer interactions.
- Utilized SQL in Dbt for data transformation, data cleaning, data quality checks and data lineage.
- Established comprehensive data schemas, data dictionaries, stakeholder lists and critical data attributes (CDAs).
- Generated synthetic data using Python to represent Salesforce incoming feed into Snowflake through Azure Cloud.

#### Virus Research Assistant, University of Bath, United Kingdom

Mar 2021- Aug 2021

- Conducted epidemiological modeling to analyze COVID-19 transmission across UK cities using WHO data.
- Applied time-series forecasting and geospatial analysis to assess virus spread.
- Created interactive dashboards in Tableau to visualize insights for academic research.

### EDUCATION

#### City University of London, MSc in Data Science

Sept 2024- Oct 2025

- Modules:** Principles of Data Science, Machine Learning, Big Data, Natural Language Processing, Computer Vision, Neural Computing (AI). Projects available on <https://github.com/Nixonngzh/Portfolio-Guide>

#### University of Bath, MEng in Chemical Engineering

Sept 2017- Jul 2022

- Overall 2:1.** Engineering Product Design & Real-Time Python Simulation, Process Modelling and Control.

### VOLUNTEERING

#### Covid-19 Relief Campaign Co-organizer, HappyBankCrew Food Bank

Aug 2021- Oct 2021

- Led fundraising efforts, securing RM15,000 (£3,000) within 2 weeks for emergency food supplies.
- Managed stakeholder relations, engaging donors, food banks, and community media.

### SKILLS AND CERTIFICATIONS

#### Skills

- Languages & Analytics:** Python (Pandas, Numpy, Scikit-Learn, PyTorch), Advanced SQL, JavaScript, Cypher
- Data Engineering:** ETL/ELT Pipelines, Data Vault Modeling, Matillion, dbt, Apache Spark, Kafka
- Cloud Platforms:** AWS (S3, Lambda, Glue, Redshift), Azure Cloud, GCP, Snowflake, Databricks
- Databases & Visualization:** PostgreSQL, MSSQL, Teradata, Neo4j, Tableau, PowerBI, Plotly Dash, Grafana
- DevOps:** Git, CI/CD, Docker, Kubernetes, Apache Airflow, Automated Testing

#### Certifications

- Databricks Lakehouse Fundamentals, Snowflake Hands-on Essentials, Neo4J Certified Professional, AWS Certified Cloud Practitioner (In Progress)