ADELEKE AKINRINADE KAYODE

Senior Data Scientist | Machine Learning Engineer | Statistical Researcher

Python • R • SQL | Deep Learning • NLP • Bayesian Inference • Causal Inference | AWS • GCP • Azure | 4 Peer-Reviewed Publications | M.Sc. Statistics (Best Project Award)

+234 706 838 3770 • adelekeakinrinade1@gmail.com • kmexconsult.com • linkedin.com/in/Akinrinade • github.com/kmexa • Ibadan, Nigeria

PROFESSIONAL SUMMARY

Results-driven Senior Data Scientist, Machine Learning Engineer, and Statistical Researcher with 8+ years of end-to-end experience transforming complex data into high-impact decisions across healthcare, finance, and AI training domains. Dual fluency in both rigorous statistical theory and production ML engineering — a combination that enables robust model design grounded in mathematical first principles and delivered as reliable, scalable systems.

Expert in Python and R across the full data science stack: data collection, cleaning, and preprocessing; exploratory data analysis; statistical modelling (Bayesian inference, causal inference, survival analysis, discrete choice modelling); and production machine learning (XGBoost, LightGBM, TensorFlow, PyTorch, Hugging Face Transformers). Experienced with deep learning architectures (CNNs, RNNs, Transformers), NLP pipelines for structured and unstructured text, and reinforcement learning-adjacent AI evaluation workflows. Strong cloud computing background across AWS (SageMaker, S3, Lambda), GCP (BigQuery, Vertex AI), and Azure, with MLOps experience spanning Docker, GitHub Actions, Airflow, and CI/CD pipelines.

Published researcher with four peer-reviewed journal contributions in applied statistics and biomedical science. M.Sc. in Statistics (University of Ibadan, Best Project Award 2020) with undergraduate and postgraduate research spanning Bayesian experimental design, discrete choice modelling, and distribution theory. Former university lecturer in biostatistics and data science. Currently contributing to TreasuryIQ — a production-grade NBFC liquidity risk and AI-driven ALM intelligence platform at Technocolabs Softwares.

AREAS OF EXPERTISE

Statistical Modelling & Mathematics

- Bayesian inference (MCMC, WinBUGS/JAGS), causal inference, survival analysis, mixed-effects and regression models, discrete choice modelling (conditional logit, mixed logit), time series (ARIMA/SARIMA), experimental design, power analysis, probability theory — M.Sc.-level theoretical grounding applied to real-world analytical problems.

Machine Learning Engineering

- End-to-end ML pipeline design: data ingestion, feature engineering, model training (scikit-learn, XGBoost, LightGBM, Random Forest), hyperparameter optimisation, A/B testing, performance benchmarking, model deployment, and continuous monitoring. Production-grade ML delivery under code review and CI/CD standards.

Deep Learning & Neural Networks

- CNNs, RNNs, LSTMs, Transformers, attention mechanisms — TensorFlow, PyTorch, Keras. Fine-tuning pre-trained models (BERT, GPT, T5), neural architecture design, and inference optimisation for image, text, and tabular data applications.

Natural Language Processing (NLP)

- Transformer-based NLP (Hugging Face), text classification, named entity recognition, sentiment analysis, semantic search, word embeddings, vector databases, RAG architectures, tokenisation pipelines — applied to finance and healthcare unstructured data.

Data Collection, Cleaning & Engineering

- Multi-source data ingestion, preprocessing, schema validation, outlier detection, null imputation, deduplication, feature engineering — automated quality validation frameworks in Python and SQL. ETL/ELT pipeline design with Airflow, Spark, BigQuery, MongoDB, PostgreSQL, MySQL, and Amazon Redshift.

Data Analysis & Visualisation

- Exploratory data analysis, trend and pattern identification, statistical hypothesis testing, Monte Carlo simulation, econometric analysis. Data visualisation using Matplotlib, Seaborn, Plotly, ggplot2, Power BI, and Tableau — dashboards and reports for both technical and non-technical stakeholders.

Cloud Computing & MLOps

- AWS (SageMaker, S3, Lambda, EC2, Redshift), GCP (BigQuery, Vertex AI, Cloud Storage), Microsoft Azure ML — cloud-based model training, deployment, and data processing. Docker, GitHub Actions, Jenkins, Airflow CI/CD — reproducible, production-grade MLOps pipelines.

Applied Research & Technical Communication

- Four peer-reviewed publications in indexed international journals. Manuscript preparation, peer review support, and statistical methodology consulting for academic and industry clients. Precise technical documentation and stakeholder communication in both written and verbal formats.

TECHNICAL SKILLS

Python: pandas, NumPy, SciPy, scikit-learn, XGBoost, LightGBM, TensorFlow, PyTorch, Keras, statsmodels, NLTK, SpaCy, Hugging Face Transformers, LangChain, Matplotlib, Seaborn, Plotly, SQLAlchemy, pytest, subprocess — expert-level

R: tidyverse, caret, brms, survival, ggplot2, data.table, lme4, nnet, mlogit — expert-level statistical modelling and data analysis

SQL & Databases: PostgreSQL, MySQL, BigQuery, Amazon Redshift, MongoDB, SQLite — complex queries, window functions, CTEs, stored procedures, query optimisation, large-scale data extraction

ML & DL Frameworks: scikit-learn, XGBoost, LightGBM, Random Forest, TensorFlow, PyTorch, Keras, Hugging Face Transformers — supervised, unsupervised, deep learning, NLP, and ensemble methods

Cloud Platforms: AWS (SageMaker, S3, Lambda, EC2, Redshift), GCP (BigQuery, Vertex AI, Cloud Storage), Microsoft Azure ML

Data Engineering & Orchestration: Apache Spark, Hadoop, Airflow, ETL/ELT design, Docker, Jenkins, GitHub Actions, CI/CD, Linux, Bash/CLI

Statistical Software: SPSS, SAS, Stata, MATLAB, WinBUGS/JAGS (Bayesian MCMC), Mathematica, Jupyter Notebooks

Visualisation & BI: Matplotlib, Seaborn, Plotly, ggplot2, Power BI, Tableau

Version Control: Git, GitHub (PRs, branching, code review, Actions)

Additional Languages: Java, MATLAB, LaTeX, Bash

PROFESSIONAL EXPERIENCE

Senior Data Scientist & ML Engineer — Contract | **Technocolabs Softwares (TreasuryIQ)** *2025 – Present | Remote*

- Architecting and deploying an end-to-end AI-driven liquidity risk and ALM intelligence platform for NBFCs — leading the full data science pipeline from multi-source data ingestion and preprocessing through feature engineering, ML model development, validation, and production deployment.
- Engineering scalable Python and SQL data pipelines with automated quality validation (schema checks, outlier detection, integrity auditing) for cash-flow forecasting and liquidity stress testing model inputs.
- Developing and benchmarking predictive ML models for liquidity risk scoring, applying rigorous performance testing and iterative improvement cycles based on stakeholder-defined evaluation criteria.
- Conducting exploratory data analysis on complex financial datasets — identifying ALM structural mismatches, liquidity patterns, and stress indicators that directly shaped platform architecture decisions.
- Producing clear technical documentation, model architecture write-ups, and analytical reports for cross-functional remote stakeholders; contributing to sprint delivery via structured GitHub PR code reviews under production engineering standards.

Senior Data Scientist & ML Researcher — Freelance | **Upwork / Turing** *2023 – Present | Remote*

- Delivered 20+ end-to-end data science and ML projects for clients across finance, healthcare, and environmental science — spanning statistical modelling, deep learning, NLP pipelines, Monte Carlo simulation, survival analysis, A/B testing, and econometric analysis.

- Designed and implemented NLP solutions using transformer-based deep learning models (Hugging Face, PyTorch) for finance domain text classification and entity extraction — fine-tuned pre-trained models and benchmarked output quality with evidence-based evaluation reports.
- Audited and corrected a Monte Carlo financial projection model in Python/Excel for a US client — traced structural data transformation errors, applied targeted fixes, and improved forecast accuracy across multiple model versions with comprehensive documentation.
- Debugged a national-scale BigQuery/Python ETL pipeline for Worth Rises (US non-profit) — read unfamiliar codebase, ran diagnostic scripts, diagnosed multi-state data gaps, and validated corrected outputs through systematic testing.
- Produced panel ARDL econometric analyses and ML modelling pipelines supporting peer-reviewed submissions (Virtus Interpress, Finance Research Letters, IEEE Access, Journal of African Business).
- Passed Turing SQL and ML Engineer Bench assessments; authored an original competitive programming problem ('Chained Factorizations') accepted through Turing's Problemsetter Evaluation — demonstrating algorithmic depth and task design capability.

Principal Data Scientist & Statistical Consultant | **Lufemos Consult***May 2023 – Present | Ibadan, Nigeria*

- Lead all data science and statistical consulting engagements — developing predictive models, ML pipelines, NLP solutions, and visualisation dashboards for clients in healthcare, finance, and education.
- Designed and deployed end-to-end ML pipelines (scikit-learn, XGBoost, TensorFlow) with automated data quality validation, performance testing, and cloud-based deployment (AWS, GCP).
- Applied advanced statistical methodologies — Bayesian inference, causal inference, survival analysis, mixed logit modelling — to solve complex real-world client problems across regulated industries.
- Communicated analytical findings and model outputs to non-technical stakeholders through Power BI dashboards, Python visualisations, and precise written reports; directly informing strategic business decisions.

Lecturer — Data Science, ML & Statistics | **MOA Professional Institute (City & Guilds)***Jan 2022 – Apr 2023 | Ibadan, Nigeria*

- Designed and delivered data science and ML curriculum for City & Guilds of London Institute diploma programmes — covering Python, R, statistical analysis, machine learning, and data visualisation.
- Supervised 15+ capstone projects through complete data science workflows from data collection through modelling, validation, and stakeholder presentation; conducted code reviews and produced written technical feedback.
- Received consistently above-average teaching evaluations; conducted hands-on Python, R, and SPSS practical workshops.

Lecturer II (Biostatistics & Research Methods) | **University of Medical Sciences***Aug 2018 – Dec 2021 | Ondo, Nigeria*

- Taught applied biostatistics, epidemiology, and research methodology to undergraduate and postgraduate medical and public health students; promoted from Assistant Lecturer to Lecturer II based on research output and teaching excellence.
- Published peer-reviewed statistical research in indexed international journals — distribution theory, clinical data analysis, and epidemiological modelling.
- Provided statistical consulting to medical researchers and supervised 10+ thesis projects spanning regression, survival analysis, and epidemiological modelling in R and SPSS.

Statistical Data Analyst — Volunteer | **Centre for Citizens with Disabilities (CCD)***Jan 2020 – Present | Nigeria*

- Analysed survey and demographic data to support evidence-based disability advocacy; produced statistical reports and data visualisations informing policy recommendations and grant applications.

KEY PROJECTS

TreasuryIQ — AI-Driven NBFC Financial Intelligence Platform — *Technocolabs Softwares 2025*

- Stack: Python, SQL, scikit-learn, XGBoost, pandas, GCP BigQuery, Docker, GitHub Actions | Production-grade ML platform for NBFC liquidity risk scoring and ALM classification — full pipeline from multi-source data engineering through ML model training, validation, performance testing, and stakeholder reporting.

AML / Financial Crime Detection — Five-Model ML Pipeline — *Lufemos Consult Research 2024*

- Stack: Python, scikit-learn, XGBoost, LightGBM, NetworkX, ARIMA, NLTK, SpaCy, statsmodels | Designed a five-model ensemble ML classification system for AML transaction anomaly detection integrating network graph analysis, ARIMA time-series features, and NLP extraction from unstructured transaction narratives.

NLP & Transformer Deep Learning — Finance Text Analytics — *Freelance — Upwork / Lufemos2023–2024*

- Stack: Python, Hugging Face Transformers, PyTorch, TensorFlow, SpaCy, NLTK | Implemented and fine-tuned transformer-based NLP models for finance domain text classification and entity extraction on large-scale unstructured data; rigorous performance benchmarking and model architecture documentation throughout.

Worth Rises Prison Telecom Data Pipeline — *Freelance — Upwork 2023*

- Stack: Python, SQL, BigQuery, MongoDB, ETL | Debugged and optimised a national-scale ETL pipeline ingesting prison telecom pricing data; resolved multi-state data integrity gaps through systematic codebase analysis and validated corrected outputs, restoring 100% national data coverage.

Healthcare Discrete Choice Experiment — *MSc Thesis — University of Ibadan 2019*

- Stack: R, Python, SPSS, ggplot2, conditional logit, mixed logit | Designed a 500+ respondent DCE survey, applied advanced mixed logit modelling to quantify healthcare willingness-to-pay and preference heterogeneity across Oyo State — awarded Best Project, Department of Statistics, 2020.

EDUCATION

M.Sc. Statistics (Statistical Design of Investigation) | *University of Ibadan 2018 – 2021*

CGPA: 5.00/7.00 | WES Equivalent: 3.22/4.00 | Best Project Award, Dept. of Statistics, 2020 | PhD-eligible (confirmed Nov 2021)

- Thesis: Age Preference for Life-saving Programs in Oyo State — Discrete Choice Experiment; conditional & mixed logit modelling across 500+ respondents; full data collection, cleaning, EDA, and advanced statistical modelling pipeline in R and SPSS.

B.Sc. Statistics — Second Class Honours (Upper Division) | *University of IlorinOct 2012 – Sep 2016*

CGPA: 4.14/5.00 | WES Equivalent: 3.49/4.00

- Undergraduate Project: Bayesian Analysis of Clinical Trial Data — MCMC sampling (WinBUGS/JAGS), D-optimal and A-optimal experimental design criteria; rigorous probabilistic modelling and statistical documentation throughout.

PROFESSIONAL CERTIFICATIONS

Advanced Statistics for Data Science Specialization — *Coursera / Johns Hopkins University* *2025*
Advanced Linear Models for Data Science 1 & 2 — *Johns Hopkins University* *2025*
Mathematical Biostatistics Boot Camp 1 & 2 — *Johns Hopkins University* *2025*
Machine Learning with Python — *IBM / Coursera* *2025*
Introduction to Data Engineering — *IBM / Coursera* *2025*

PEER-REVIEWED PUBLICATIONS

[J.1] Obubu et al. (2019). Useful Generalization of the Inverse Lomax Distribution: Statistical Properties and Application to Lifetime Data. American Journal of Biomedical Science & Research.

[J.2] Obubu et al. (2019). On Making an Informed Decision Between Four Exponential-Based Continuous Compound Distributions. Journal of Advances in Applied Mathematics, 4(2).

[J.3] Obubu et al. (2019). New Generalization of Length Biased Exponential Distribution with Applications. Journal of Advances in Applied Mathematics, 4(2).
[J.4] Obubu et al. (2019). Lung Cancer: A Chronic Disease Epidemiology; Prevalence Study. Asian Journal of Advanced Research and Reports, 3(4), 1–7.

RESEARCH INTERESTS

Bayesian Hierarchical Modelling • Causal Inference • Discrete Choice Experiments • Financial Crime Analytics • NLP for Finance & Healthcare • AI Model Evaluation & Alignment • Probabilistic Machine Learning

Languages: English (Native), Yoruba (Native), Nigerian Pidgin (Fluent) **|** **Availability:** Open to remote contract, consulting, and research collaboration engagements. **|** **References:** Available on request.