Onyeka Muoka
Data Scientist

Email: muokaonyeka@yahoo.com | Phone: +447309120764 | LinkedIn: linkedin.com/in/onyeka-muoka |
Portfolio - GitHub: github.com/Onyeka280

Professional Summary: Analytical and resourceful data scientist with 8 years of experience in healthcare business operations. Proven track record in managing day-to-day activities, aligning KPIs with strategic goals, and driving performance. Skilled in data pre-processing, predictive modelling, machine learning, and data visualization. Strong background in handling high-volume datasets and building automated data pipelines. Brings a results-driven mindset and deep understanding of operational challenges to deliver actionable insights. Passionate about using data science to improve outcomes and support evidence based decision making.

Key Skills

- Programming Languages: Python (Keras, Tensorflow, NLTK, Scikit-learn, Hugging Face Transformers (pipeline), LDA, BRETopic, Statsmodels), SQL
- Data Visualization: Matplotlib, Seaborn, Tableau, Power BI
- Machine Learning: Regression, Classification, Clustering, Time-Series Analysis, Anomaly Detection, Neural Networks
- Databases: MySQL, PostgreSQL
- Statistical Analysis: Hypothesis Testing, Regression Analysis
- Soft Skills: Strong communication, Problem-solving, Collaboration, Project Management

---

Projects

Topic Modelling and Sentiment Analysis with BRETopic and FinBERT
- Designed a user-friendly interface for supervisors at the Bank of England's Prudential Regulation Authority to detect sentiment, extract actionable insights and risk indicators, and perform comparative analysis across financial institutions using quarterly results announcements, helping them quickly spot early warning signs in banks' financial health and reduce manual review time.
- Tools: Python (Phi4, BERTopic, FinBERT), Tableau

Customer Segmentation with K-Means Clustering
- Segmented 68,000+ e-commerce customers using K-Means in Python, applying RFM and CLV metrics to identify five key clusters and deliver insights that improved customer engagement.
- Tools: Python (Scipy, Scikit-learn, PCA & t-SNE), Power BI

Anomaly Detection for Ship Engine Monitoring
- Developed an anomaly detection system for ship engine data in Python by cleaning sensor inputs and applying IQR, One-Class SVM, and Isolation Forest, then visualized performance deviations using PCA plots in Seaborn, enabling early detection of potential engine failures, reducing downtime risk and supporting predictive maintenance planning.
- Tools: Python (Scikit-learn, One-Class SVM, Isolation Forest, PCA)

Experience

Data Science Career Accelerator
University of Cambridge | February 2025 – Present | United Kingdom
- Developed machine learning models (Auto-ARIMA, XGBoost, LSTM, and hybrid approaches) to identify sales patterns and forecast weekly sales for Nielsen BookScan, improving forecast accuracy by 15% and enabling more data-driven sales planning.
- Built a Python-based diabetes prediction model achieving 92% precision, enabling automated risk screening for early intervention and lowering processing time by 40%.
- Analyzed 40,000+ customer reviews from Google and Trustpilot using NLP techniques (topic modelling, sentiment and emotion analysis, LLMs) to identify key drivers of negative feedback, uncover location-specific issues, and generate actionable recommendations that improved customer experience for PureGym's 2M+ members.
- Designed and deployed a sentiment analysis pipeline using NLTK and Scikit-learn, automating classification of 1,000+ monthly reviews, cutting processing time by 50% and speeding up product decisions by 20%.
- Developed a student dropout prediction model using deep neural networks and XGBoost in Python, enabling early identification of at-risk students and contributing to a 15% reduction in dropout rates.

Clinical Support Specialist (Phlebotomy)
Royal Free London NHS Trust and Lewisham and Greenwich NHS Trust | February 2024 – Present | United Kingdom
- Collected, recorded, and maintained accurate patient data, ensuring high standards of data quality, integrity, and compliance with the Trust policies and the Data Protection Act (2018).
- Monitored and analysed patient activity records, identifying patterns in service use and contributing to audit processes that informed operational improvements.
- Used digital systems (iCare, SwiftQueue and other electronic health record platforms) to document and manage data, supporting real-time decision-making for patient care.

- Ensured compliance with infection control and health and safety standards by systematically documenting and maintaining records, enabling data analysis to identify incident patterns, evaluate risks, and generate insights to inform policy, audits, and service improvements.
- Investigated, documented, and reported incidents, errors, and complaints, ensuring findings were translated into actionable lessons and process improvements.
- Supported training and mentoring of new staff, including documenting workflows and providing guidance on accurate data entry and reporting.
- Collaborated with multidisciplinary teams to improve patient services, aligning data insights with clinical needs and service delivery priorities.
- Managed stock levels and resource tracking, highlighting demand patterns and supporting operational planning.
- Contributed to surveys, audits, and service evaluations by collecting, collating, and analysing data to support decision-making.

Operations Manager
Citiscan Diagnostics Centre | April 2016 – Sept 2023| Awka, Nigeria
- Led the transformation of a basic ultrasound unit into a full-service diagnostics centre by analysing patient demand data, introducing X-ray, lab testing, and EMR systems—boosting diagnostic volume by over 70%.
- Deployed a custom POS and billing system after reviewing financial inconsistencies and usage logs, enabling automated reporting, reducing errors, and cutting billing time by 60%.
- Led the building and implementing of data driven scheduling and inventory tools using Excel, reducing patient wait times by 40% and cutting stock wastage by 15%.
- Collaborated with QA teams to assess operational metrics and audit findings, leading to measurable improvements in service standards and compliance.
- Tracked and visualized operational KPIs using dashboards and reports to advise leadership on cost-efficiency, staff productivity, and service trends, driving informed strategic decisions.

Education
Data Science Career Accelerator
University of Cambridge | February 2025 – October 2025

Master of Science in International Business
University of Chester | 2023 – 2024
Thesis: "The impact of influencer marketing on customers intention to purchase in the gaming industry"

Bachelor of Science in Marketing
University of Calabar, Nigeria | 2011 – 2015

Volunteering
DataKind UK
Data Scientist, August 2025 – Present

Royal Free London NHS Trust
PACS/RIS Administrator, November 2025 - Present

Certifications
Data Science Career Accelerator – University of Cambridge (2025)
Full Stack Data Science Training – 10alytics (2025)
Python - Kaggle Certificate of Completion (2025)
Pandas - Kaggle Certificate of Completion (2025)
Learning Microsoft SQL Server 2022 – Linkedin Learning (2025)
SQL Server Integration Services - Linkedin Learning (2025)

Additional Information

- Languages: English (Fluent), Igbo (Conversational)
- Interests: Machine Learning, Artificial Intelligence, Data Ethics