UMAZAYI SUCCESS

Data Scientist / Data Analyst

(+234) 904 1124 474 • successumazayi@gmail.com • Kogi, Nigeria

LinkedIn • Portfolio • GitHub

PROFESSIONAL SUMMARY

Results-driven Data Scientist and Analyst with hands-on experience in machine learning, statistical analysis, and data wrangling across real-world datasets. Skilled in building and deploying end-to-end ML pipelines, feature engineering, and translating complex data into actionable business insights through Power BI dashboards and Python visualizations. Currently pursuing a B.Sc. in Computer Science at the University of Ilorin.

TECHNICAL SKILLS

Languages & Libraries: Python (Pandas, NumPy, SciPy, Scikit-learn, Matplotlib, Seaborn), SQL (MySQL)

Machine Learning: Supervised/Unsupervised Learning, Feature Engineering, Hyperparameter Optimization, Model Deployment (Streamlit)

Statistical Analysis: Hypothesis Testing, A/B Testing, Regression Analysis, Probability Distributions, Data Wrangling & Preprocessing

Tools & Platforms: Jupyter Notebook, Microsoft Power BI, Excel (VLookup, Pivot Tables, Conditional Formatting), ydata_profiling, Git/GitHub

PROJECTS

Malaria Diagnosis Detection | Personal Project — April 2026

GitHub: https://github.com/Success007T/Malaria-diagnosis-detection/ Live App: https://malaria-diagnosis-detection.vercel.app/

- Performed exploratory data analysis with ydata_profiling to surface data quality issues, patterns, and early diagnostic signals.
- Cleaned and prepared the dataset with Pandas to improve consistency, usability, and model readiness.
- Engineered predictive features that improved how the models captured relationships within the data.
- Trained and evaluated Random Forest and XGBoost classifiers, both achieving over 90% accuracy and recall.
- Optimized for recall to strengthen identification of likely malaria cases and reduce the risk of missed positives.

Data Science: Car Price Prediction | Personal Project — March 2026

GitHub: https://github.com/Success007T/Car-price-analsis-and-prediction-models Live App: https://car-price-prediction-app-by-success.streamlit.app

- Improved car price prediction accuracy by achieving $R^2 > 0.94$ across Linear Regression, Random Forest, XGBoost, and Decision Tree models by engineering 12+ features from raw data including depreciation curves and condition scoring after rigorous data wrangling and preprocessing.
- Reduced model error through advanced hyperparameter optimization using GridSearchCV, benchmarking 4 algorithms to identify the best-performing configuration.
- Accelerated decision-making for 3+ user personas by building an interactive Power BI dashboard visualizing price depreciation, make/model value retention, and condition analysis — deployed as a live Streamlit app.

Amazon Sales Analytics | Personal Project — February 2026

GitHub: https://github.com/Success007T/CodeAlpha_Tasks

- Uncovered a 0.21 correlation between discount depth and product rating across 1,400+ SKUs by conducting end-to-end statistical analysis and EDA — including discount vs. rating hypothesis testing, best-value scoring, and category-level breakdowns — using Python, Pandas, and NumPy.
- Increased analytical reusability by 100% by consolidating the full pipeline (EDA, sentiment analysis via TextBlob, and visualization) into a modular amazon_analysis.py script, eliminating repetitive notebook work.
- Surfaced actionable product intelligence through NLP-based sentiment analysis on customer reviews, identifying top-performing and underperforming categories using Matplotlib and Seaborn visualizations.

ST@40 Data Analytics Capstone | Capstone Project — February 2026

GitHub: https://github.com/Success007T/ST40_Capstone_Project_data_analysis

- Delivered sales performance insights to stakeholders by designing a professional Power BI dashboard — reducing time-to-insight by standardizing KPI reporting across datasets after thorough data wrangling and preprocessing in Python.
- Improved data quality for downstream analysis by performing systematic feature engineering and data cleaning, resolving missing values, outliers, and schema inconsistencies across the full sales dataset.

TechCrush Data Science Capstone — Agriculture Yield Prediction | Group Project (Team of 10) — February 2026

GitHub: https://github.com/Delkay-byte/Predictive-Agriculture-Analysis

- Contributed to a deployed machine learning model that predicted agriculture yield with measurable accuracy by collaborating on feature engineering, statistical analysis of key agricultural variables, and model evaluation across the full ML pipeline.

- Identified 5+ key agricultural intelligence factors driving yield performance by conducting exploratory data analysis and visualizations in Python, enabling targeted intervention strategies for the team's recommendations.
- Supported an A/B testing framework for model comparison across crop categories, validating the selected model's generalizability through cross-validation and performance benchmarking.

EDUCATION

Bachelor of Science, Computer Science — University of Ilorin, Kwara, Nigeria