

Data Entry & Automation Specialist

PROFESSIONAL SUMMARY

- Extensive expertise in data extraction, modeling, wrangling, statistical modeling, mining, machine learning, and data visualization, with 3+ years of experience as a Data and Automation Specialist.
- Proven domain knowledge and hands-on experience in the Health, Banking, and Manufacturing industries, enabling effective data-driven solutions tailored to specific business needs.
- Demonstrated ability to transform complex business requirements into manageable data formats and develop analytical models, algorithms, and data mining solutions that scale seamlessly across massive volumes of structured and unstructured data.
- Proficient in managing the entire data science project life cycle, from data acquisition and cleaning to engineering, feature scaling, statistical modeling, testing, validation, and data visualization.
- Strong proficiency in a wide range of machine learning algorithms and predictive modeling techniques, including Regression Models, Decision Trees, Random Forests, Sentiment Analysis, Naïve Bayes Classifier, Support Vector Machines (SVM), and Ensemble Models.
- Skilled in applying statistical methodologies such as Hypothesis Testing, ANOVA, Time Series Analysis, Principal Component Analysis, Factor Analysis, Cluster Analysis, and Discriminant Analysis. Experienced in time series analysis using models like AR, MA, ARIMA, GARCH, and ARCH.
- Strong programming skills in Python (2.x, 3.x) for developing robust analytic models and solutions. Proficient in utilizing the SciPy Stack packages, including NumPy, Pandas, SciPy, Matplotlib, and IPython, to implement data analysis and machine learning algorithms.

SKILLS

- **Programming and Scripting languages:** Python, HTML, R Programming
- **Automation & AI:** Familiarity with automation tools (Demeter AI, Carnegie, ChatGPT, Claude, Mid-journey)
- **Databases:** SQL, SAP, Microsoft Excel
- **Statistical Methods:** Hypothetical Testing, ANOVA, Time Series, Confidence Intervals, Bayes Law, Principal Component Analysis (PCA), Dimensionality Reduction, Cross-Validation, Auto-correlation
- **Machine Learning:** Scikit-Learn, NumPy, TensorFlow, PyTorch, Keras
- **Reporting Tools:** Tableau Suite of Tools 10.x, 9.x, 8.x which includes Desktop, Server and Online, Server Reporting Services (SSRS), Orange Data Mining, Exploratory.io
- **Tools:** Google Colab, Eclipse IDE, Jupyter Notebook, Visual Studio Code, RStudio, SPSS, Orange, Exploratory.io
- **Operating systems:** Windows, Linux, MacOS
- **Visualization Tools:** Power BI, Tableau, Google Charts, Plotly, Seaborn
- **Other skills:** Data Cleaning, Data Manipulation, Data Visualization, Data Wrangling, Database Administration EDA, Data Analysis, Data Mining, Critical Thinking, Communication & Presentation Skills, Problem-Solving

EDUCATION

- |   |             |
|---|-------------|
| • <b>Bachelor of Science   Mathematics</b>                          | <b>2023</b> |
| ○ Ondo State University of Science & Technology                     |             |
| • <b>Diploma   Scientific Computing and Python for Data Science</b> | <b>2020</b> |
| ○ World Quant University – USA                                      |             |
| • <b>Master of Science   Data Science &amp; Information System</b>  | <b>2024</b> |
| ○ Osiri University – Nebraska, USA                                  |             |
| • <b>Professional Diploma   Data Science</b>                        | <b>2024</b> |
| ○ African University of Science & Technology, Abuja                 |             |

## PROFESSIONAL EXPERIENCE

### Geeks for Growth

Marketing Operations Analyst -Data & AI (Fully Remote)

June 2023 – Present

#### Responsibility:

- Processed and managed large marketing data for seamless integration into the company's proprietary database using ScaleSERP and Carnegie platform.
- Automated data workflows by implementing Demeter AI Automation, improving operational efficiency and accuracy in marketing data processing.
- Optimized marketing tracking by ensuring Hermes URLs were correctly applied to structured data for enhanced attribution and performance analysis.
- Utilized Excel (Pivot Tables, VLOOKUP/XLOOKUP, and advanced formulas) to clean, consolidate, and analyze marketing data before integration.
- Conducted routine data audits and quality checks to maintain marketing data integrity, troubleshoot discrepancies, and improve database accuracy.
- Developed and maintained Excel-based dashboards and reports to track key marketing performance metrics, enabling data-driven decision-making.
- Collaborated with designers, developers, analysts, front-end engineers, and marketing teams to ensure smooth data flow and optimized workflows across departments.
- Created and maintained Standard Operating Procedures (SOPs) for data imports, updates, and troubleshooting steps to improve marketing data consistency.
- Communicated effectively via Slack, Loom video and project management tools to coordinate data-related tasks, resolve issues, and enhance workflow automation in marketing operations.

### Strictly4data Group

Remote Data Scientist

June 2020 – 2023

#### Responsibility:

- Collaborated with data engineers and operation team to implement ETL process, wrote and optimized SQLqueries to perform data extraction to fit the analytical requirements.
- Performed data analysis by using Hive to retrieve the data from Hadoop cluster, SQL to retrieve data fromRedShift.
- Explored and analyzed the customer specific features by using Spark SQL.
- Performed univariate and multivariate analysis on the data to identify any underlying pattern in the data andassociations between the variables. Performed data imputation using Scikit-learn package in Python.
- Participated in features engineering such as feature intersection generating, feature normalize and label encodingwith Scikit-learn preprocessing.
- Used Python 3.X (numpy, scipy, pandas, scikit-learn, seaborn) and Spark 2.0 (PySpark, MLlib) to develop variety of models and algorithms for analytic purposes.
- Developed and implemented predictive models using machine learning algorithms such as linear regression,classification, multivariate regression, Naive Bayes, Random Forests, K-means clustering, KNN, PCA and regularization for data analysis
- Conducted analysis on assessing customer consuming behaviors and discover value of customers with RMF analysis; applied customer segmentation with clustering algorithms such as K-Means Clustering and HierarchicalClustering.
- Built regression models include: Lasso, Ridge, SVR, XGboost to predict Customer Life Time Value.
- Built classification models include: Logistic Regression, SVM, Decision Tree, Random Forest to predict CustomerChurn Rate.

**Responsibility:**

- Proficient in handling highly imbalanced Fraud datasets, employing techniques such as under sampling, oversampling with SMOTE, and cost-sensitive algorithms using Python Scikit-learn to improve fraud prediction accuracy.
- Developed complex Spark SQL queries to analyze data and meet specific business requirements, showcasing a strong command over data manipulation and extraction in distributed computing environments.
- Utilized MapReduce/Spark Python modules in Hadoop on AWS to build predictive analytics and machine learning models, demonstrating proficiency in leveraging big data frameworks for advanced data analysis.
- Ensured data quality and integrity by performing data cleaning and validation using Pandas and Numpy, contributing to the overall reliability of the analytical process.
- Engaged in feature engineering tasks such as generating feature intersections, normalizing features, and label encoding using Scikit-learn preprocessing, enhancing the effectiveness of predictive models.
- Improved fraud prediction performance by implementing random forest and gradient boosting algorithms for feature selection, leveraging Python Scikit-learn's capabilities.
- Applied a variety of machine learning algorithms including Naïve Bayes, KNN, Logistic Regression, Random Forest, SVM, and XGboost to identify loan default risks, showcasing expertise in credit risk assessment.
- Developed an ensemble model combining Ridge, Lasso Regression, and XGboost to predict potential loan default losses, demonstrating an ability to utilize diverse algorithms for comprehensive risk analysis.
- Contributed to the standardization of documentation and data modeling practices, gathering and translating application requirements into data models, and supporting the adoption of best practices related to data and applications.

## OTHER PROFESSIONAL EXPERIENCE

**Facilitator** *Tableau for Data Visualization Training*

May 2017

Lagos State Planning Commission & UNFPA

- Led a 3-day training program, enhancing participants' skills in statistical analysis and data management.

**Programming Instructor** *Digital Skills Literacy Workshop*

JTech Foundation, Ghana

- Delivered a 2-month training course on fundamental programming concepts, enabling participants to explore advanced topics.
- Conducted a workshop on Effective Online Communication and Microsoft Office 365, achieving an 87% score.

**Facilitator** *Data Science & AI Training*

*Data Science Network (DSN)*

- Conducted classes on data analysis, predictive modeling, and automation, equipping professionals with practical AI skills.
- Trained participants in using data science techniques for health data analytics, focusing on predictive models for disease detection and patient outcome analysis, contributing to the improvement of healthcare decision-making processes.

**Facilitator** *Data analysis, Visualization, Excel, PowerPoint, and Tableau Training*

Nasarawa Investment and Development Agency (NASIDA) | Strictly4data

- **Upcoming:** Leading a week-long training program for 8 participants, focused on extensive data analysis, building Excel analytical reports, PowerPoint, and Tableau skills.
- This program will include hands-on projects and cater to both beginner and intermediate learners, enhancing their capabilities in data analysis, presentation, and visualization.

# UNIVERSITY OF NIGERIA



## DIPLOMA CERTIFICATE

This is to Certify that

**BILAU DAUDA**

having satisfied all the requirements laid down by the

University, is awarded this Diploma in

**ARTIFICIAL INTELLIGENCE**

on this day **Tuesday 31st of December, 2024**

VICE-CHANCELLOR

REGISTRAR



## CERTIFICATE

of Completion

This certificate is awarded to:

***Dauda Bilau***

DREG006276

for successfully completing the requirements for:

**Certificate Programme in Applied Data Science Program with AWS Cloud  
Computing, Big Data Analytics with PySpark, ML Ops**

**November 30th, 2024**



Dean

Certificate No. : th66ewYNCm

# Scientific Computing and Python for Data Science

Applied Data Science Module: Unit I

THIS CERTIFICATE OF COMPLETION IS AWARDED TO

**Dauda Bilau**

WITH HONORS

BY WORLDQUANT UNIVERSITY

IN PARTNERSHIP WITH THE DATA INCUBATOR

this, the twelfth day of September, two thousand and twenty



The Data Incubator



# Certificate of Service

Has been awarded to

**Dauda Bilau**

**for**

**Facilitating**

**AI Invasion 2021;  
Introduction to Machine Learning**

(30 hours Python Programming, Introduction to Artificial Intelligence and Machine Learning)

From 21 – 25 June, 2021



*Bayo Dekambi*  
Conveners' Signature

**Authorization ID: DSNAI0005807**



FEDERAL MINISTRY OF  
COMMUNICATIONS AND  
DIGITAL ECONOMY  
Leveraging ICT for National Development

**NATIONAL INFORMATION TECHNOLOGY  
DEVELOPMENT AGENCY**

**NITDA**

# Certificate of Participation

Awarded to

*Dauda Bilau*

For attending the course on

*Basic digital skills and productivity tools*

Awarded on: *4th June, 2021*

**Kashifu Inuwa Abdullahi, CCIE**  
Director General/CEO





NSESA FOUNDATION



# Certificate of Participation

*Nsesa Foundation presents this Certificate to*

*Danda Bilan*

*Has successfully completed SuaCode's "Introduction to Programming" course offered by Nsesa Foundation and developed in the following: Variables, Conditionals, Functions and Basic Concepts in the Processing programming language.*

PRINCE STEVEN ANNOR  
*Project Lead*



GEORGE BOATENG  
*President*