

VALENTINE NTHIGAH

DATA SCIENTIST & ANALYST

✉ wawiranthigah20@gmail.com ☎ +254790024211 📍 47791-00100 [in](#) linkedIn [G](#) Github

PROFESSIONAL SUMMARY

- knack for crafting insightful visualizations, reports, and data storage tools for project needs
- Depth of knowledge in collecting and organizing large datasets, performing statistical analysis, and developing data models.
- Skilled in SQL, Python, machine learning, and data visualization and committed to staying abreast of emerging trends in data analytics.

PROFESSIONAL EXPERIENCE

Commercial Analyst

Burn Manufacturing USA LLC

02/2024 – present
Nairobi, Kenya

- **Sales:** Track the monthly distribution of approximately 10,000 stoves across four countries.
- **Distribution:** Develop scalable ETL pipelines and assess performance across different regions within each country.
- **Business Registration:** Use KOBO to track and analyze B2B registrations for approximately 5,000 new businesses per end-user quarter.
- **Expense Tracking:** Review and optimize the allocation of \$10,000–\$100,000 in operational expenses at the country level.
- **Stock Balance Model:** Implement an XGBoost Regressor model that predicts stock levels for 20 main distributors and evaluates stock performance based on inventory aging aiming for less than 10%.
- **AD-Hoc Tasks:** Complete 2 predictive marketing projects per quarter, with impact measured by successful execution achieving a 10% increase in marketing ROI.

Data Analyst Consultant

Ponea Health Ltd

2021 – 2024
NAIROBI, Kenya

- Analyzing \$ 500K+ of monthly marketing spend data to optimize customers' interest.
- Development of data models through regression analysis by optimizing \$ 1.5M customer's advertising spend by media type through multivariable regression analysis.
- Maintain 500+ database backup and performance reports and file systems for the program's data
- Designed comprehensive reports, dashboards, and visualizations to form the Excel data effectively.
- Utilizing SQL to generate insights on clients' churn and renewal rates from data tables with 1000 rows in SQL.

QUANTITATIVE FIELD INTERVIEWER

KENYA NATIONAL BUREAU OF STATISTICS

01/2022 – 07/2022
NAIROBI, Kenya

- Conducted quantitative and analytical field data collection tasks during the Kenya Demographic and Health Survey (KDHS) 2022.
- Created a good rapport with the respondents and ensured the confidentiality of the data collected from the respondents.
- Gathered data and analyzed the distribution trend from the sampled household cases
- Compiled timely daily field reports and submitted them to the county statistic officer (CSO)

EDUCATION

BACHELOR OF SCIENCE IN PURE AND APPLIED MATHEMATICS,,
KENYATTA UNIVERSITY

2016 – 2021
NAIROBI, Kenya

SKILLS

Statistical& analytical skills:

SPSS, STATA, advanced excel, python, machine learning, sklearn, PyTorch, XGBoost, NLP,VMYSQL, MS Word, PowerPoint, Outlook.

ETL (Extract, Transform, Load) Processes, Identify patterns, correlations, and outliers, and Provide meaningful insights

Data Visualization Tools

Tableau, Seaborn, PowerBi, Excel, matplotlib

Other skills

Time management, Verbal and written communication, Interpersonal and diplomacy, organizational and coordination skills

CERTIFICATES

SQL - MySQL for Data Analytics and Business Intelligence

2023

UDEMY

QUANTITATIVE ANALYSIS AND WEB ANALYTICS

2021 – 2022

DALWORTH

NAIROBI, Kenya

- SPSS, STATA, ADVANCED EXCEL

DATA ANALYTICS & SCIENCE

05/2023 – 05/2024

AZUBIAFRICA

Machine Learning

hamoye HDSC Spring

PROJECTS

ALX Real-Time Stock Market Analysis Pipeline

2024

Built a real-time stock market analytics system

- using Kafka, Spark, and NoSQL databases.
- Integrated machine learning algorithms to predict stock prices with 80% accuracy.

Indian startup ecosystem

10/2023

an investigation of the ecosystem and a propose of the best course of action
skills

- Python- libraries employed include; (nSQLAlchemy, Numpy, pandas, dotenv, pyodbc, seaborn, matplotlib.pyplot and SciPy.Stats), MySQL and Tableau

HAMOYE TEAM PROJECT

04/2023 – 06/2023

International football prediction

data pipeline

- Predicting the outcome of football matches
- skills: Google COLAB, Python Libraries, Numpy, Pandas, Matplotlib.pyplot, Seaborn, Plotly. express