

MIKE NJERU MURANGO

IBM Certified Data Scientist | Business Intelligence Analyst

Mikemurango00@gmail.com

+254791303899, Nairobi, Kenya

[LinkedIn](#) | [Github](#)

<https://murangomike.github.io/New-Portfolio/>

SUMMARY

Strategic and detail-oriented Analytics and Machine Learning Engineer with a strong background in data modeling, machine learning model development, and algorithm implementation. Experienced in building and optimizing data pipelines, performing SQL-based transformations, and monitoring data quality. Skilled in leveraging machine learning techniques to solve complex problems and drive data-driven decision-making. Adept at collaborating with both analytics and technology teams to design scalable, efficient data infrastructure that enhances impact measurement. Passionate about using advanced analytics and machine learning to drive social impact and optimize processes for better outcomes.

EXPERIENCE

Data Science Consultant

Mar 2024-Current

- Conducted data collection via mobile data tools, web scraping using Python libraries (BeautifulSoup, Requests), and **API integrations** for scalable and automated data ingestion.
- Performed comprehensive data wrangling, cleaning, and preprocessing on complex structured and semi-structured datasets using Python (Pandas, NumPy), SQL, and BigQuery to support analytics, modeling, and reporting workflows.
- Built and orchestrated robust **ETL/ELT pipelines** using **dbt**, **Apache Airflow**, and custom Python scripts; extracted data from CRM, ERP, and HR systems (e.g., Oracle, Microsoft RDS, REST APIs) and loaded into AWS S3, BigQuery, and data lakes for scalable processing.
- Leveraged **Google BigQuery** for large-scale SQL analysis—joining disparate datasets, optimizing query performance using partitioning and clustering, and developing materialized views to power dashboards and BI tools.
- Collaborated using **Git for version control** and team-based development, managing branches, code reviews, and merge requests in tools like GitHub and GitLab.
- Implemented **CI/CD pipelines** using **GitHub Actions** and **Jenkins**, automating testing, builds, and deployment processes.
- Containerized data workflows and Python services with **Docker**, building reproducible Docker images for scalable deployment across development, staging, and production environments.
- Developed and deployed **RESTful Flask APIs** for ML models and data services, enabling integration with web and mobile frontends; hosted applications in production environments using Docker, Gunicorn, and NGINX on AWS EC2 and Heroku.
- Created and delivered insightful **dashboards** and reports using Power BI, Looker Studio, and Tableau, enabling both technical and non-technical stakeholders to act on data-driven

insights.

- Developed a **market basket analysis** model using the Apriori algorithm, which optimized in-store product placements and bundling strategies for a retail client.
- Built and deployed **machine learning models** for churn prediction and retention forecasting, enabling proactive customer retention strategies.
- Conducted A/B testing experiments on pricing strategies for a retail client, identifying optimal pricing tiers that boosted revenue and customer satisfaction.
- Built and trained a **Computer Vision model** using Convolutional Neural Networks (CNNs) with PyTorch, allowing customers to identify retail products from photos for an improved shopping experience.
- Applied reinforcement learning techniques (Multi-Armed Bandits: Epsilon-Greedy, Thompson Sampling) in a marketing context to dynamically optimize ad creatives and media spend, significantly improving CTR and conversion rates.
- **Collaborated cross-functionally** with data engineers, developers, and business units to deliver production-grade data products and decision-support systems in cloud-native environments.

Data Analyst - Supp Merchant

Mar 2023 - Apr 2024

- Conducted data cleaning and pre-processing using **Excel, Power BI, and Python**, ensuring data quality for analysis.
- Designed survey instruments and conducted quantitative and qualitative research projects.
- Utilized Programming languages such as Python to create **ETL pipelines** as well as build and maintain data dictionaries.
- Spearheaded **python automations** for data wrangling by creating field level standards like date formats, currency handling, null values and NLP cases etc.
- Performed advanced **data mining, statistical analysis**, and forecasting for various departments in the organization.
- Led **technical documentation** of data processes, business requirements and testing requirements.
- Developed and tracked Key performance indicators to measure effectiveness of data initiatives.
- Maintained excellent communication with government counterparts and other stakeholders to ensure consistent updates and smooth coordination of data-driven initiatives.
- Established **GDPR-compliant** data access controls for the entire organization.

Data Analyst Volunteer - UN Development Programme Oct

2023 - Dec 2023

- Analyzed data from multiple sources, including **Excel** and cloud databases, to assess progress on AU2069 goals.
- Created comprehensive **Power BI dashboards** to visualize AU2069 themes and objectives, delivering clear, actionable insights to senior stakeholders.
- Ensured data integrity by implementing thorough validation checks and conducting quality assessments.
- Collaborated with cross-functional teams to implement data governance practices,

enhancing the reliability of data used in decision-making.

- Recommended automation opportunities to streamline repetitive reporting tasks.

Junior Data Analyst - Radiant Minds

Feb 2019 - March 2023

- Analyzed and interpreted complex datasets using **Power BI, Excel, and Python** to identify business trends and opportunities.
- Cleaned, transformed, and modeled data to ensure accuracy and reliability for decisionmaking.
- Developed interactive dashboards in Power BI and Excel, providing real-time insights for business stakeholders.
- Built predictive models to forecast trends and support strategic planning.
- Provided actionable insights through data storytelling, presentations, and strategic reports to drive business growth.
- Collaborated with internal and external stakeholders to align data analysis with business objectives and decision-making.
- Automated reporting workflows using **Power Query, DAX, and Python**, reducing manual effort and improving efficiency.
- Optimized business processes by leveraging data-driven strategies and insights.

Public Management Officer Intern - Nairobi City County Jan 2018 - April 2018

- Conducted data analysis using **Excel**, generating actionable insights for public management projects.
- Presented findings to supervisors and directors (**non-technical**), simplifying complex data for decision-making.
- Led data collection efforts from various slums in Nairobi County using mobile data collection techniques.
- Collaborated with community outreach teams to ensure accurate and reliable data gathering.
- Assisted in creating reports that tracked program effectiveness and community impact.
- Coordinated with stakeholders to address data-related challenges

EDUCATION

- **Data Science**
World Quant University Mar 2023 - Aug 2024
- **Business Analytics Specialization**
(People analytics, Operations analytics, Customer Analytics Accounting Analytics)
The Wharton School of the University of Pennsylvania Dec 2023- Mar 2024
- **Business and Marketing Strategies Specialization**
University of London Dec 2023- Mar 2024
- **Bachelor of Science in Management and Development**
Jomo Kenyatta University of Agriculture and Technology (JKUAT) Sep 2015 - Nov 2019

- **Certificate in Advanced Computer Application (CSK)**
(Data Analysis, Accounting & Database Management)
Annex Computer College

April 2015

CERTIFICATIONS

- **Professional Certificate, Data Science**
2024 International Business Machines (IBM)
- **Project Management**
2024 Microsoft Corp.

Sep 2023 - Jun

Oct

TECHNICAL SKILLS

Data Modeling & Transformations: dbt, SQL, data warehousing, incremental models, snapshotting, partitioning, clustering

• **Data Quality & Governance:** dbt tests, data freshness, anomaly detection, data validation, logging, naming conventions, data lineage

• **Data Engineering & ETL:** ETL/ELT, Airflow, dbt, Python, schema design, API ingestion, REST APIs, batch processing, streaming, orchestration

• **Version Control & Collaboration:** Git, GitHub, GitLab, Bitbucket, pull requests, branching, documentation, Confluence, Jira, cross-functional teams

• **DevOps & Deployment:** CI/CD, GitHub Actions, Jenkins, Docker, Docker Compose, containers, virtual environments, artifact versioning

• **Tools & Technologies:** dbt, SQL, Git, SurveyCTO, ODK, TextIt, Power BI, Python (Pandas, NumPy), Google BigQuery, PostgreSQL, MySQL, AWS (S3, RDS), GCP, VS Code, Jupyter, Linux, Bash

SOFT SKILLS

Stakeholder Engagement, Clear and Concise Communication, Problem Solving, Ability to Collaborate, Prioritization, Continuous learning, adaptable, Project management
Risk Management, Analytical Skills, Interpersonal Skills, Team player, Pragmatic Thinking, Personal Management, Self Management, Ethics, Innovation, Confidence, Report Writing, Interpersonal Communication