KEZIAH GICHEHA

Gmail | LinkedIn | Portfolio | GitHub

+254717254185

Nairobi, Kenya.

Professional Summary

Data Scientist with a proven ability to extract meaningful insights from data and drive business impact. Experienced in machine learning, predictive analytics, and data visualization, with a strong track record in projects such as Customer Churn Prediction and NLP-based Depression Detection. Proficient in Python, SQL, and Power BI, leveraging data-driven strategies to optimize decision-making. With a Computer Technology and Project Management background, I bring a unique blend of technical expertise and leadership to solve complex problems. Passionate about using data to unlock opportunities and fuel business growth.

Education

MBA in Project Management – Daystar University, 2025.
- Relevant Coursework: Quantitative Techniques, Stakeholder Management, Project Management, Strategic Management

B. Tech in Computer Technology – The Technical University of Kenya, Dec 2018.
- Relevant Coursework: Database Administration, Object Oriented Programming

Certifications:
- Data Science Certificate, Moringa School, 2024
- Microsoft Certified: Power BI Data Analyst, 2025

Key Skills

- Version Control: Git, GitHub
- Programming Languages: Python, SQL
- Data Cleaning & Preparation: Data Wrangling, Feature Engineering, Data Transformation
- Data Visualization: Tableau, Matplotlib, Seaborn, Power BI
- Machine Learning: Supervised Algorithms (Regression & Classification), Unsupervised Learning Algorithms: Clustering
- Database Management: MySQL, PostgreSQL
- Reporting Techniques
- Business Process Improvement
- Stakeholder Management
- Storytelling with Data

Projects

NLP Depression Detection
Description: Developed a natural language processing (NLP) model that classifies Reddit posts as either depressive or non-depressive.
Tools & Technologies: Python, EDA, ETL, NLTK, Scikit-learn, StreamLit
Key Contributions:
- Achieved 90% accuracy, showcasing linguistic markers of depressive tendencies.
- Deployed a live app for real-time use. (**Moodlens**)

Predicting Customer Churn
Description: Developed models to predict customer churn, enabling proactive retention strategies.
Tools & Technologies: Python, Scikit-learn, Logistic Regression, Decision Trees
Key Contributions:
- Achieved 82% model accuracy and identified key churn predictors, improving retention by 25%.

Sentiment Analysis of Social Media Posts
Description: Performed sentiment analysis on tweets related to Apple and Google products
Tools & Technologies: Python, NLTK, Sentiment Analysis, Logistic Regression, Support Vector Machine (SVM)
Key Contributions:
- Based on the analysis, the SVM model is recommended for production use due to its superior overall accuracy and handling of neutral sentiment.

Career History

Data Analyst, Freelancer
December 2024 – Present
- Developed Actionable Dashboards – Created interactive Power BI dashboards for a retail client, enabling executives to track KPIs and improve operational efficiency.
- Improved Business Decision-Making – Analyzed customer purchasing behaviour for an e-commerce client, leading to a 15% increase in sales by optimizing product recommendations.
- Enhanced Data Accuracy & Efficiency – Cleaned and structured messy datasets, reduced reporting errors by 30%, and cut data processing time by 40%.
- Optimized Marketing Strategies – Conducted sentiment analysis on social media data, helping a brand increase engagement by 20% through data-driven content adjustments.

System Administrator, ECW Ltd, Nairobi, Kenya
Jan 2022 – Aug 2024
- Supported stakeholder decisions by providing insights into project performance, helping achieve 100% compliance and revenue growth.
- Ensured data accuracy in reports that influenced key decision-making for stakeholders.
- Generated and analyzed over 30 reports per month using Excel to track project performance.
- Optimized SQL queries, improving query speed by 10%.
- Acted as manager during the absence of the duty manager, ensuring smooth daily operations.

Technical Support Engineer – XC360 EA LTD Nairobi, Kenya
Sept 2022 – Mar 2021
- Identified and resolved potential data issues, increasing operational efficiency by 15% and preventing delays in reporting.
- Streamlined IT operations through remote support and troubleshooting, ensuring accurate and efficient data handling.

Data Assistant, FHI360, Nairobi, Kenya
Feb 2019 – Oct 2019
- Managed and organized large datasets, ensuring data integrity and accessibility for analysis, leading to a 20% reduction in processing time.
- Conducted data entry, cleaning, and validation tasks, improving the accuracy and reliability of reports by 15%.

References

Available upon request