

BRIAN KIPKEMBOI

📞 0448 748 242

✉ briankipkemboi52@gmail.com

[in](#) [Brian Kipkemboi](#)

🌐 briankipkemboi.com

📍 Cairns City, QLD 4870

ABOUT MYSELF

Detail-oriented and analytically driven Data Science professional with professional experience spanning healthcare data systems, statistical analysis, program evaluation, and community services. I am currently completing a Master of Data Science (Professional) at James Cook University with expertise in machine learning, natural language processing, computer vision, and applied analytics. Experienced in building end-to-end data pipelines, deploying ML models on edge devices, and applying NLP techniques to real-world datasets. Demonstrates strong capability in designing data collection frameworks, analysing complex datasets, and translating statistical outputs into actionable insights. I am skilled in Python, R, SQL, PyTorch, and modern data science tools. Background in frontline homelessness case management and disability support brings a uniquely human-centered lens to data-driven problem solving.

CORE COMPETENCIES

- Statistical Analysis & Modelling
- Healthcare Data Systems (HMIS, DHIS, EHR)
- Data Collection & Survey Design
- Data Cleaning, Validation & Transformation
- Quantitative & Qualitative Analysis
- Reporting & Data Visualizations
- Monitoring & Evaluation (M&E)
- Research Methodology
- Database Management
- Evidence-Based Decision Support
- Machine Learning & Predictive Modelling
- Natural Language Processing & Sentiment Analysis
- Computer Vision & IoT Systems
- Web Scraping & Data Pipeline Engineering
- Deep Learning & Transformer Architectures
- Homelessness Case Management & Crisis Intervention

PROFESSIONAL EXPERIENCES

AI Research Intern

JCU Innovation & Industry Lab - Mind Scrapers Team, Cairns | February 2026 - Present

Roles:

- Completed WIL placement (EG5300) under Dr. Samantha J. Horseman, delivering 200 hours of applied AI research across five concurrent industry projects.
- Built and deployed a real-time computer vision system using YOLOv8 and OpenCV for agricultural quality detection on Raspberry Pi edge devices with LoRa/Meshtastic IoT telemetry.
- Developed a live Flask dashboard displaying quality percentages, stock levels, and sales inference from edge-transmitted sensor data.
- Conducted full NLP sentiment analysis pipeline (VADER, TextBlob, LDA topic modelling) on 1,200+ social media comments with web scraping via Apify.
- Built three progressive GPT-style transformer language models from scratch in PyTorch, including RAG retrieval and web-scraping capabilities.
- Collaborated with industry partners including Wärtsilä (maritime AI) and presented at investment workshops and the CAVE 3.0 immersive lab launch.

Homelessness Case Manager & Residential Youth Worker

Anglicare Far North Queensland, Cairns | May 2024 - Present

Roles:

- Delivered holistic, end-to-end case management to individuals and young people experiencing or at risk of homelessness across community and residential settings.
- Conducted comprehensive intake assessments, strengths-based assessments, risk assessments and safety planning in accordance with legislation and service agreements.
- Advocated on behalf of clients with housing providers, specialist homelessness services, mental health, AOD, legal and domestic and family violence services.
- Managed crisis situations and behavioral escalations using de-escalation techniques and sound professional judgement within established guidelines.
- Maintained accurate, confidential and timely documentation within approved systems in line with organizational policies, quality standards, and KPI-driven frameworks.

Disability Support Worker

Enabling Pathways, Cairns | October 2023 - June 2024

Roles:

- Collected and recorded client health and behavioral data systematically, ensuring accuracy and consistency for monitoring progress and supporting evidence-based healthcare planning decisions.
- Analysed daily client activity patterns and behavioral trends to assist healthcare professionals in evaluating care effectiveness and improving personalized care strategies.
- Maintained structured documentation systems, ensuring compliance with healthcare regulations, data integrity standards, and organizational reporting requirements for audit and review purposes.
- Collaborated with multidisciplinary healthcare teams to interpret client data, contributing to development and adjustment of individualized care and support plans.
- Monitored nutritional intake records, and health indicators, identifying patterns and reporting anomalies to support clinical decision-making processes.
- Applied data recording and reporting techniques to track client outcomes, supporting evaluation of service quality and continuous improvement initiatives.
- Assisted in compiling client progress summaries and reports, translating observational and recorded data into meaningful insights for healthcare providers.

- Ensured accurate data capture during daily care activities, supporting statistical tracking of client wellbeing indicators and service delivery outcomes.

Statistician

Nuru International, Migori| November 2022 - September 2023

Roles:

- Designed and developed survey instruments, questionnaires, and structured data collection frameworks to ensure statistically valid, reliable, and representative datasets for analysis and reporting.
- Collected, processed, validated, and managed large datasets from multiple sources, ensuring high standards of data quality, accuracy, consistency, and integrity across all statistical operations.
- Performed advanced statistical analyses including regression, hypothesis testing, confidence intervals, and multivariate modelling to generate meaningful insights supporting evidence-based decision-making.
- Utilised statistical software such as SPSS and STATA to conduct complex data manipulation, predictive modelling, and quantitative analysis for program evaluation and research purposes.
- Prepared analytical reports, dashboards, charts, and data visualisations using Excel and other tools to effectively communicate statistical findings to technical and non-technical stakeholders.
- Monitored and improved data collection processes by performing data cleaning, coding, validation, and error-checking to ensure reliable datasets suitable for accurate statistical interpretation.
- Supported project planning, monitoring, and evaluation activities by providing data-driven insights and recommendations to enhance program performance and strategic outcomes.
- Maintained comprehensive documentation of statistical methodologies, procedures, and outputs while ensuring compliance with ethical standards, data confidentiality, and organisational reporting requirements.

Statistician

Reale Hospital, Eldoret| November 2021 - November 2022

Roles:

- Designed and implemented statistical studies, surveys, and observational research to support hospital decision-making, operational planning, and evaluation of healthcare programs and services.
- Collected, validated, and cleaned healthcare datasets, ensuring data accuracy, completeness, and reliability for statistical analysis and reporting within hospital information systems.
- Developed and managed health databases, organizing large volumes of patient and operational data to support efficient analysis, reporting, and evidence-based decision-making processes.
- Applied advanced statistical techniques including regression analysis, hypothesis testing, and trend analysis to interpret complex healthcare data and identify key performance indicators.
- Prepared comprehensive statistical reports, visualizations, and summaries to effectively communicate findings and insights to hospital management and multidisciplinary stakeholders.
- Provided data-driven recommendations to support policy development, operational improvements, and strategic planning aimed at enhancing patient care and service delivery.
- Utilised statistical software such as R, Python, and SPSS to perform data analysis, predictive modelling, and ensure reproducibility and accuracy of statistical results.
- Ensured compliance with data governance, confidentiality, and ethical standards while handling sensitive health information, and supported staff through training in statistical methods and data management practices.

EDUCATION & CERTIFICATIONS

Master of Data Science (Professional) | James Cook University | September 2023 - Present

Bachelor of Science in Applied Statistics | Rongo University | September 2017 - December 2021

- Second Class Honours (Lower Division)

Certificate III in Individual Support | Community Services | Completed

Certificate in Computer Software and Maintenance | Royalnet Institute, Kitale | February 2017 - June 2017

- Credit I

Kenya Certificate of Secondary Education (KCSE) | Lelmokwo High School | January 2013 - December 2016

- Score: 52 Points

Kenya Certificate of Primary Education (KCPE) | Sunrise Academy | January 2002 - December 2012

- Score: 353 / 500

TECHNICAL SKILLS

- **Statistical Tools:** R, SPSS, STATA, Python (NumPy, pandas, SciPy)
- **Data Tools:** MS Excel (Advanced), Access, Pandas, Jupyter Notebook, Google Colab, AWS SageMaker
- **Platforms & Deployment:** Git, GitHub, Netlify, Vercel, Docker, Flask, React, Streamlit
- **Systems:** DHIS2, HMIS, EHR Systems, Raspberry Pi, LoRa/Meshtastic IoT, Linux
- **Techniques:** Regression Analysis, Sampling Methods, Data Modelling, Machine Learning (Scikit-learn, XGBoost), Deep Learning (PyTorch, TensorFlow), NLP (NLTK, VADER, TextBlob, LDA), Computer Vision (OpenCV, YOLOv8), Web Scraping (BeautifulSoup, Selenium, Apify), Transfer Learning (MobileNetV2), RAG Systems, Transformer Architectures
- **Visualization:** Tableau, Power BI, Matplotlib, Seaborn, Plotly, Streamlit Dashboards
- **Languages:** Python, R, SQL, JavaScript, HTML/CSS, Bash

PERSONAL SKILLS

- Analytical thinking and problem-solving ability
- Strong written and verbal communication skills
- Ability to interpret complex datasets accurately
- Attention to detail and data accuracy focus
- Team collaboration and stakeholder engagement
- Time management and multitasking ability
- Adaptability in dynamic healthcare environments
- Data storytelling and technical presentation
- Cross-cultural communication and stakeholder engagement

- Crisis management and de-escalation
- Rapid prototyping and end-to-end project delivery

AREAS OF INTEREST

- Biostatistics and Public Health Analytics
- Predictive Modelling and Machine Learning
- Health Data Systems and Digital Health
- Epidemiological Research
- Data-driven Policy and Decision Making
- Statistical Consulting and Research Design
- Natural Language Processing & Low-Resource Language AI
- Computer Vision for Agriculture & IoT Systems
- Civic Technology & AI for Social Impact
- Deep Learning & Transformer Model Development

REFERENCES

Available on Request