

Brilliant Jepkogei Kiptoo

Data Scientist

Sydney, NSW · brilliantkip@gmail.com · 0424 964 808 · [LinkedIn](#) · [GitHub](#)

Home

[Machine Learning](#)[Python](#)[Statistics](#)[Data Modelling](#)[Streamlit](#)

Summary

Data Scientist with strong academic training in Statistics and Data Science and hands-on experience building predictive models, performing advanced analytics, and deploying data-driven solutions. Experienced in transforming complex datasets into actionable insights using machine learning, statistical modelling, and data visualisation techniques.

Core Skills

- Data preprocessing, feature engineering, and model evaluation
- Machine learning (regression, classification, clustering, time series)
- Statistical modelling and hypothesis testing
- Exploratory data analysis and insight generation
- Model interpretation and communication of results

Tools & Technologies

- Python (Pandas, NumPy, Scikit-learn)
- R, SQL
- Matplotlib, Seaborn, Plotly
- Streamlit, Jupyter Notebook

- Git, Docker, Cloud platforms (GCP / Azure)

Professional Experience

Data Analyst / Data Scientist — Capstone Construction Limited

July 2022 – Present · Remote

- Applied statistical and machine learning techniques to analyse operational and project datasets.
- Developed predictive and analytical models to assess trends, risks, and performance patterns.
- Performed feature engineering and data transformation to improve model accuracy.
- Built analytical dashboards and reports to communicate insights to stakeholders.
- Ensured data quality through validation, anomaly detection, and documentation.

Economist / Statistician Intern — Kenya National Bureau of Statistics (KNBS)

June 2021 – June 2022

- Performed statistical analysis on large datasets using Python, R, and SPSS.
- Supported modelling and trend analysis for socio-economic indicators.
- Generated visualisations and analytical outputs for national statistical reports.

Junior Statistician — Kenya Forestry Research Institute (KEFRI)

June 2019 – May 2021

- Conducted exploratory data analysis and statistical modelling for research projects.
- Applied regression and inferential techniques to forestry and socio-economic datasets.
- Supported research publications through data-driven insights.

Education

- **Master of Data Science and Innovation** — University of Technology Sydney (Feb 2023 – Dec 2024)
- **Bachelor of Science in Statistics** — Meru University of Science & Technology (Sep 2013 – Mar 2019)

Professional Development

- Machine Learning & Applied Statistics
- Data Science Capstone Project
- English (Proficient), Swahili (Fluent)

Selected Projects

- **Eat Well, Live Well:** Built a machine learning model to predict obesity and diabetes risk and deployed a Streamlit web application.
- **Sentiment Analysis Tool:** Developed an NLP model using Doc2Vec and LSTM to analyse financial news sentiment.
- **CSV Explorer:** Created an interactive Streamlit application for exploratory data analysis on uploaded datasets.

Download

[Download CV \(PDF\)](#)

Upload the PDF later into `/assets/cv/`.