

BILLY KIPCHIRCHIR KOECH

[Portfolio Site](#)

✉ billykoech137@gmail.com

☎ +254 714 985 992

🌐 [Billy Koech](#)

🌐 [Billy Koech](#)

Professional Summary

A results-driven Data Scientist with over 2 years of experience and a strong academic foundation in data science and civil engineering, complemented by hands-on expertise in data analytics, machine learning, and web development. Proficient in Python, SQL, and tools like Power BI, Matplotlib, and Streamlit, I specialize in extracting insights from complex datasets and delivering actionable solutions. My recent work includes developing predictive models, leading youth employment data analysis across Africa, and contributing to interactive platforms like the Youth Employment Clock. Passionate about leveraging data to address societal challenges, I aim to create impactful solutions that drive innovation and informed decision-making.

Education

University of East London

Master of Science in Data Science

January 2024 to Present

University of Nairobi

Bachelor of Science in Civil Engineering.

2014 – 2020

Skills

- | | |
|---|--|
| • Programming Languages: | Python, SQL |
| • Data Visualization: | Power BI, Matplotlib, Seaborn, Plotly, Streamlit |
| • Data Analytics & Processing: | Pandas, NumPy, Excel, Google Sheets |
| • Machine Learning: | Sci-kit learn, XGBoost, Keras, TensorFlow |
| • Version Control: | Git, GitHub |
| • Development Tools: | VS Code, Jupyter Notebooks |
| • Presentations & Documentation: | Google Slides, PowerPoint, Google Docs, Word |
-

Work Experience

Associate Data Scientist | World Data Lab

October 2023 to October 2024

Key Responsibilities:

- Extracted and analyzed youth employment data from the ILO database and national statistics departments across Kenya, Uganda, Rwanda, Ghana, Nigeria, and Ethiopia.
- Conducted data cleaning, analysis, and forecasting to project youth employment trends by 2030, considering sector, population, and gender.
- Communicated findings through detailed presentations to senior management and clients, translating complex data into strategic insights.
- Contributed to the integration of employment data into an online interactive platform for visualizing youth employment trends in Africa.
- Facilitated a workshop with the Kenya National Bureau of Statistics (KNBS), presenting actionable insights on youth employment data in Kenya.
- Led a cross-functional internal team to collect and analyze youth employment data in Kenya, ensuring data accuracy and meaningful insights.

Impact:

- Contributed to the development of the Youth Employment Clock, an online interactive platform that monitors the employment status of young people in Africa | [LINK](#)

Data Science Intern | World Data Lab

April 2023 to September 2023

Key Responsibilities:

- Developed a predictive random forest model to impute missing data, enhancing data reliability and supporting analytics workflows.
- Led the collection of subnational population data for Kenya using web scraping techniques, streamlining data availability for stakeholders.
- Contributed to project documentation by detailing methodologies, findings, and actionable outcomes.
- Conducted in-depth analysis of MSME and Kenya Labour Market data, generating insights that supported strategic decision-making.

Impact:

- Increased company visibility by engaging in an online initiative, encouraging other interns to pursue roles within the organization | [LINK](#)
- Secured publication of the MSME analysis in Business Daily, demonstrating the value of data-driven insights to external audiences | [LINK](#)

Other Experiences - 2020 to March 2023

- Gained expertise in strategic planning, innovation, and project management by founding and leading a startup that developed SalesPoa, a platform aggregating deal information in Kenya.
- Enhanced analytical and problem-solving skills through design and quality assurance tasks during a road construction project as a Civil Engineering Intern.

Projects

1. YouTube Comments Sentiment Analyzer | [Try it out](#) | [Source code](#)

The YouTube Comments Sentiment Analyzer is an interactive web application that performs sentiment analysis on YouTube video comments. The application provides detailed video statistics and visualizes sentiment distribution using an interactive chart.

Streamlit | Pandas | Google YouTube Data API v3 | Hugging Face Transformers | Plotly | RoBERTa Model

2. PDFChat - Intelligent Document Conversation Assistant | [Try it out](#) | [Source code](#)

PDFChat is an advanced document interaction tool that transforms the way users engage with PDF documents. It leverages the power of LLMs and vector databases to enable natural conversation-style interactions with document content, making information retrieval more intuitive and efficient.

Python | Streamlit | python-dotenv | PyPDF2 | LangChain | LangChain-Community | OpenAI | FAISS (Facebook AI Similarity Search) | Git | HTML Templates | OpenAI API.

Certificates

- IBM Data Science Professional Certificate | [LINK](#)
- ALX Data Analytics Certificate | [LINK](#)

Awards

- MbeleNaBiz Business Plan Competition winner.