

JOHN CHEGE MAINA

Junior Data Scientist

 0726 058 103

 jcmaina400@gmail.com

 [linkedin.com/in/john-chege-maina](https://www.linkedin.com/in/john-chege-maina)

PROFILE OBJECTIVE

Motivated junior data scientist aiming to leverage a strong understanding of data analytics techniques, statistical analysis, and data visualization tools to extract meaningful insights from large datasets. Proficient at interpreting complex data, collaborating with teams, and delivering solutions to align with business objectives. Eager to learn from seniors and contribute to the organization's data-driven decision-making process.

EDUCATION

28/08/ 2018 – 29/11/2022 **Bachelors in Mathematics and Computer Science**
Murang'a University of Technology

WORK EXPERIENCE

**13/09/2021 –
20/12/2021**

Kenya National Bureau of Statistics

- Executed field data collection and management processes, ensuring a 98% data accuracy rate and timely delivery of analysis reports.
- Efficiently accomplished and analyzed data, extracting valuable insights from trends, resulting in an 85% accuracy rate in identifying key patterns.
- Engineered research surveys and skillfully managed data collection process, achieving an impressive 95% response rate; provided a comprehensive dataset that served as a valuable resource for in-depth analysis.
- Effectively monitored and evaluated tasks, achieving a 95% completion rate and providing actionable feedback for process improvement.
- Demonstrated proficiency in data collection methods, achieving a 90% response rate and accurate representation of Kenya
- Fruitfully linked actual agricultural outputs to socioeconomic outcomes, providing essential insights for decision-making with an 85% correlation accuracy.
- Conducted survey experiments with a 95% response rate, ensuring a comprehensive dataset for analysis.

Attachee

PROJECTS

Data Analysis for Socio-Economic Decision Making

Collected and analyzed data on actual agricultural outputs and their impact on various socio-economic outcomes. Conducted regression analysis to identify significant correlations between agricultural performance and socioeconomic indicators. Prepared data visualizations and reports for policymakers. Outcome: Provided concrete data-backed recommendations for optimizing agricultural policies and improving socioeconomic outcomes.

Kenya Household Master Sampling Frame Analysis

Collected data through survey experiments, interviews, and questionnaires. Managed data entry and data cleaning processes for the sampled households. Conducted statistical analysis to generate demographic insights. Outcome: Provided accurate and comprehensive data on the household demographics, helping policymakers make informed decisions.

FIFA analysis

Conducted analysis of the 2022 FIFA World Cup data and produced an understanding of the players, teams, and matches. Managed to visualize and prepare a report.

Netflix movies and series analysis

Conducted statistical analysis of the data to produce insights about the movies and series. Outcome: provided comprehensive insight into the age, genre, year of release, movie directors, countries, and the popularity of the movies and series offered by Netflix.

United States Accidents Analysis

Performed data cleaning processes on the dataset. Analysed data to generate demographic insights. Prepared charts and provided concrete data-based commendations on strategies to curb road accidents.

Amazon food customer reviews analysis

Conducted sentiment analysis on the Amazon food customer reviews dataset. Generated insights on the opinions on the opinions of the customers. With this, I prepared visualizations and reports on the strategies to be put in place to improve sales.

Titanic data analysis and classification

Cleaned the Titanic dataset from Kaggle. Analysed to uncover some insights about the dataset. Through the analysis, managed to know which cabin had the highest number of occupants, which passenger class had the most people, between male and female, who survived and perished the most. Through the use of machine learning models, managed to classify the gender, deaths, cabins, and age of the Titanic travelers.

Programming languages popularity analysis Performed data cleaning for the dataset and conducted statistical analysis to generate insights. Provided precise outcomes about the popularity and usage of all the programming languages. Came up with a visualization for easy understanding.

SKILLS

- ✧ Microsoft Excel
- ✧ Python
- ✧ SQL(MySQL, MS Server, Postgre SQL)
- ✧ Power Bi
- ✧ Tableau
- ✧ Machine Learning
- ✧ Web Scrapping
- ✧ STATA
- ✧ SPSS
- ✧ Hypothesis Testing
- ✧ Sentiment Analysis
- ✧ Data Science
- ✧ Statistical Analysis

CERTIFICATION

Data Entry

Ajira Digital (2021)

Quantitative Data Analysis

Sonek Data School October – December (2022)

Business Intelligence Analyst

Sonek Data School January – May 2023.

Standard Bank Data Science Job Simulation

Forage – November 2023

British Airways Data Science Job Simulation

Forage – November 2023