# Developing an interactive analytics dashboard for African Global country Indicators, with a particular focus on East African countries

Africano BYAMUGISHA
byamugisha.africano@students.mak.ac.ug
2023/HD05/04595U, Makerere University

Gideon MPUNGU
mpungu.gideon@students.mak.ac.ug
2023/HDO5/04617U, Makerere University

## ABSTRACT

Understanding development progress in any country requires accessible insights into global indicators. However, in the African context, there is limited accessibility to such insights with ease. Therefore as an attempt to provide a solution, this project develops an interactive analytics dashboard for visualizing and exploring global indicators' country profiles for African countries with a keen focus to East Africa. The dashboard facilitates comparative analysis across countries, regions, and other demographic characteristics for each of the variables. Utilizing reliable data source of the Gapminder data repository, the project analyzes diverse indicators across economic, social, environment and other key sectors using descriptive statistics, time series analysis, and correlation techniques. The methodology employed in this project synthesizes and advances previous work by incorporating established open-source tools, including R, Python, and Shiny, to facilitate an interactive analytical and visualization experience for a more engaging, exploratory interaction with the data for a non technical audience. Evaluating the effectiveness and impact of the dashboard, domain knowledge specialist reviewed and suggested feedback's through demos. The dashboard (can be accessed HERE) has mainly 3 sections of the sidebar which houses the menu options, the filters and slicers and the dashboard body that contains the visualisations and the descriptions. Future projects could improve this project by integrating predictive analytics, expanding the dashboard to cover all countries for better comparison, incorporating additional statistical analyses, and deploying it on a more robust server for improved performance.

## KEYWORDS

data analytics, interactive visualization, global indicators, African countries

## 1 MOTIVATION AND DOMAIN DESCRIPTION

### 1.1 Data Domain Definition

This project will focus on a wide range of global indicators that reflect the country profiles of African nations. These indicators cover various domains such as social, economic, health, trade, technology, and education, among others, that are relevant to the global development agenda. Global indicators are quantitative measures that capture the status, progress, and performance of countries on various aspects of development, such as poverty, health, education, gender equality, environment, trade, and innovation [1]. Global indicators can also provide a basis for cross-country and cross-regional comparisons and bench marking, as well as identifying best practices and areas for improvement [2][3]

### 1.2 Project Goal

The main goal of this project is to develop an interactive and comprehensive data analytics dashboard that visualizes and explores the global indicators for African country profiles. The dashboard will enable comparative analysis between countries, comparison among regions, and in-depth examination of Uganda's performance in relation to its East African neighbors. The dashboard is intended to provide policymakers, researchers, and investors with easy access to real-time and actionable insights about the global indicators that can inform their decision-making processes related to global indicators.

### 1.3 Motivation for this project

Data-driven decision-making is essential for guiding informed policy development, targeted interventions, and strategic resource allocation within a country's context. Rigorous analysis of the global indicators data is crucial for understanding complex development dynamics, uncovering disparities, identifying areas of progress, and highlighting potential challenges that require proactive intervention for a country's profile. Africa is a continent with immense diversity, potential, and opportunities, but also faces many challenges, such as poverty, inequality, conflict, disease, and environmental degradation [4]. Therefore, it is important to have easily accessible and understandable analytics on global indicators to inform evidence-based policies and actions that can foster inclusive and sustainable development in the continent [4].

### 1.4 Research Questions

- How do African countries compare to each other and global benchmarks across key global indicators?
- What dominant trends, both positive and negative, can be identified within African regions (East, West, North, Central) across various sectors?
- How has Uganda's performance on key global indicators evolved over time, and how does it compare to its East African neighbors?

### 1.5 Significance of Analysis

This analysis is valuable as it addresses the need for easily accessible, data-driven insights into African development as viewed on the global scale of the global indicators. The findings can shape development agendas, inform resource allocation, promote accountability, and ultimately contribute to improving the quality of life across the vast and diverse African continent and most especially Uganda. The analysis can also highlight the strengths and weaknesses of African countries, as well as the opportunities and threats they face, in relation to the global indicators. This can help identify the gaps and priorities for action, as well as the best practices and lessons learned, that can enhance the development outcomes and impacts of African countries. Furthermore, the analysis can also foster dialogue and collaboration among various stakeholders, such as governments, civil society, academia, private sector, and development partners, who are involved or interested in the development of Africa.

# 2 METHODOLOGY

## 2.1 Data Analytics Methods

This section describes the data sources, techniques, and tools that will be used:-

## 2.2 Data Sources

The primary data source for this project is Gapminder, a web-based platform that provides free access to hundreds of global indicators from various domains, such as health, education, economy, environment, and society. Gapminder collects and integrates data from multiple reputable sources, such as the World Bank, the United Nations, the World Health Organization, and the International Monetary Fund [5]. Gapminder dataset was complemented by other datasets from reliable and authoritative sources, such as the World Bank Open Data, the UNDP datasets, the UNStats datasets, and the OECD data library. These datasets provided additional and/or updated information on the global indicators that are relevant. The datasets will be selected based on their quality, availability, and compatibility with the Gapminder data.

## 2.3 Techniques

The following techniques were applied to perform the data analysis and visualization for the indicators:

- **Descriptive statistics:** To summarize the properties and characteristics of the indicators, descriptive statistics were employed and helped to understand the distribution and central tendencies of the data, as well as to identify potential outliers and anomalies.
- **Time series analysis:** was used to examine the changes and trends of the indicators over time. Time series analysis involved plotting the indicators overtime to identify the patterns.
- **Correlation and regression analysis:** This method was used to explore the potential relationships and associations between the indicators, and to quantify their strength and direction.
- **Advanced visualization techniques:** This used to present the data and the findings in a clear, attractive, and interactive way, using various types of charts, graphs, maps, and tables. Advanced visualization techniques involved using dynamic and responsive features, such as filters, sliders, tooltips, and animations, to enhance the user experience and engagement with the dashboard.

## 2.4 Software Tools

A combination of open-source tools were used to perform the data analysis, visualization, and dashboard development for the project. These tools include: R and Python - collaboratively used to perform the data cleaning, transformation, and statistical analysis; highcharter, and leaflet - for custom, interactive and reactive visualizations with features such as markers, popups, layers, and zooming; Shiny - Shiny was be used as the dashboard structure for the project; Shiny server - Shiny Apps server was used to run and manage the R/Python based applications, and handle the web requests to the dashboard.

# 3 STEP-BY-STEP PROCESS

The following steps were followed to conduct the data analysis and visualization and dasboard development:

## 3.1 Data Acquisition and Preprocessing

(1) Relevant datasets were Identified and collected from Gapminder, World Bank, UNDP, and other reputable sources, based on the research questions and the indicators of interest.
(2) Data cleaning was then conducted to address missing values, outliers, and inconsistencies in the dataset.
(3) Necessary data transformations were performed to ensure compatibility across datasets and facilitate comparative analysis.

## 3.2 Exploratory Data Analysis

(1) Descriptive statistics to understand the distribution and central tendencies of the key indicators, were employed.
(2) Visualizations and tables were utilised to identify patterns, potential outliers, and relationships among the indicators.

## 3.3 Comparative Analysis

(1) Interactive visualizations and tables were implemented to facilitate ranking countries based on selected indicators, across aspects of landlocked, UN state, .
(2) Designing visualizations (e.g., choropleth maps) to highlight regional trends within Africa, identifying areas with similar development trajectories.

## 3.4 Uganda – East Africa Case Study:

(1) Extracting Uganda-specific data and creating time-series visualizations to track its progress on key indicators over time.
(2) Designing comparative visualizations to benchmark Uganda against its East African neighbours, highlighting areas where it excels or lags.

## 3.5 Correlation Analysis

(1) Calculating correlation coefficients to measure the strength and direction of relationships between the key indicators.
(2) Developing interactive visualizations such as heat map, scatter plot.

## 3.6 Dashboard Development and Deployment

(1) Creating a dashboard layout and interface
(2) Integrating the data analysis and visualization components into the dashboard.
(3) Hosting and deploying the dashboard to the web, using the shiny server and the nginx webserver.

# 4 RELATED WORK

This project navigates the intersection of data analytics, interactive realtime visualization, and development studies, specifically targeting development indicators for East Africa and African Regional Comparison. The urgency to harness data for development is widely recognized, forming a substantial body of research and technical studies. Existing works, like those spearheaded by international organizations such as the World Bank [1] and United Nations Programmes [2] like UNStats, have established benchmarks in compiling, analyzing, and presenting global development indicators. Additionally, academic endeavors have explored various facets of visualization and analytics, seeking to improve data accessibility and decision-making [3,4]. This project extends these conversations by focusing on the visualization of East African country profiles, with a major focus on Uganda and also highlevel Comparisons

for African Regions through an interactive realtime dashboard, enabling stakeholders to engage with the data dynamically. This approach is informed by the gap observed in existing literature, where interactive data exploration tools tailored to Uganda global indicators are scarce. By integrating data from reputable sources like Gapminder [5], the project not only aligns with previous efforts to democratize access to development data but also introduces a novel interface designed for comprehensive, user-driven exploration. The methodology and tools chosen for this project reflect a synthesis and advancement of prior work. While the use of open-source tools such as R, Python, D3.js, and Shiny for data analysis and visualization is well-established in the field [5],[7]-[10], our project leverages these technologies to create a more integrated and interactive experience. This marks a deliberate shift from static reports and visualizations towards a more engaging, exploratory model, facilitating deeper insights into the data. In essence, this project complements existing research by providing a practical tool for exploring Uganda, East Africa and Regional Comparisons for the global in an interactive manner. It replicates the foundational goal of utilizing global indicators to illuminate development trajectories but distinguishes itself through the innovative application of interactive data visualization technologies. Such contributions are poised to enrich the discourse on sustainable development in Africa, offering new avenues for data engagement and policy formulation.

## 5 EVALUATION

- **Performance:** Measuring dashboard responsiveness and interactivity for ease of use and better user experience.
- **User feedback:** Conducting a random mini survey and usability inquiries to gather feedback on the dashboard's intuitiveness, clarity of visualizations, and overall value proposition.
- **Expert review:** Seeking evaluations from a data visualization expert and also domain experts such development specialists, policymakers to judge the relevance of insights and the dashboard's potential impact.
- **Bench marking:** Comparing visualizations and insights generated by the dashboard against findings from established reports such as the World Bank development indicators report, UNStats Word data dashboards to assess accuracy and consistency.

## 6 PRELIMINARY RESULTS

### 6.1 Dashboard Layout

This is illustrated in figure 1

### 6.2 Categorization of Indicators

The indicators were categorised followed the UNStats, World Bank and gapminder literature as follows :-

- Health: Indicators related to health, disease, life expectancy, etc. Economic: Economic indicators like GDP, trade, employment, etc.
- Environmental: Indicators related to the environment, land use, energy, etc.
- Social: Social indicators including education, poverty, demographics, etc.
- Technology and Infrastructure: Indicators related to technology use, infrastructure development, etc

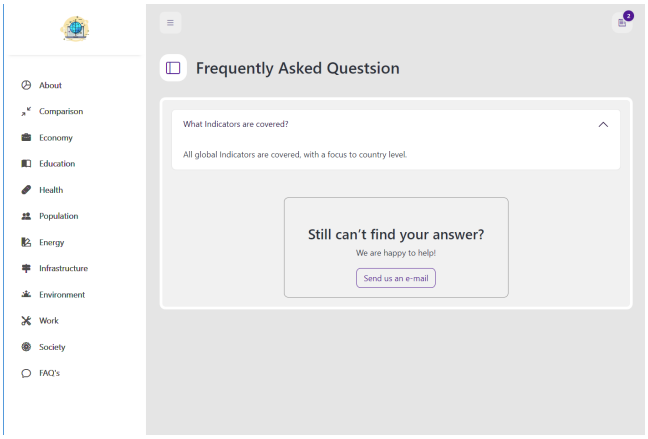


**Figure 1: A figure showing the dashboard layout and interface**

### 6.3 Strategy for Visualizing All Indicators

Given the large number of indicators, the approach used was to:

- Categorize Indicators: Grouping the indicators into broader categories (e.g., Health, Economy, Environment) for a more organized analysis.
- Modular Visualizations: Creating visualizations in a form of functions and modules that will be scaled on the dashboard for similar category.
- Automated Visualization: Developing a script that executed the visualisation modules / functions to automatically generate visualizations for each indicator with respect to user input from the dashboard.

### 6.4 Dataset Exploration

Alongside the preparation and consolidation of the datasets to a central repository, development of the charts and table summary for insights on the dashboard was performed. An exploration of some of the indicators were performed as shown in the proceeding subsections below, however this were scaled to all indicators in the dashboard with interactivity. (The exploration python notebook can be Accessed **HERE**)

### 6.5 Summary Statistics

The analysis examined the trends within the dataset, which contains a pool of historical, current, and projected values for a variety of indicators relevant to East African countries. The dataset's temporal range from 1600 to 2100 allows for extensive longitudinal studies. The indicators present a wide array, from negative values, which may indicate deficits or outflows, to positive figures reaching up to $1.62 \times 10^{11}$ , suggesting large-scale counts or inflows; as shown in figure 2. The data is comprised of a total of 480 unique indicators. These indicators cover a wide range of topics, including health, economy, environment, technology, and more. Given the extensive number of indicators, creating individual visualizations for each one would be quite extensive.

### 6.6 Correlation Analysis

The heatmap of the correlation matrix in figure 3, shows the relationships between different indicators. A strong negative correlation between child mortality and life expectancy, indicating that as child mortality decreases, life expectancy tends to increase; A positive correlation between life expectancy and primary school completion for girls, suggesting that improvements in education might be

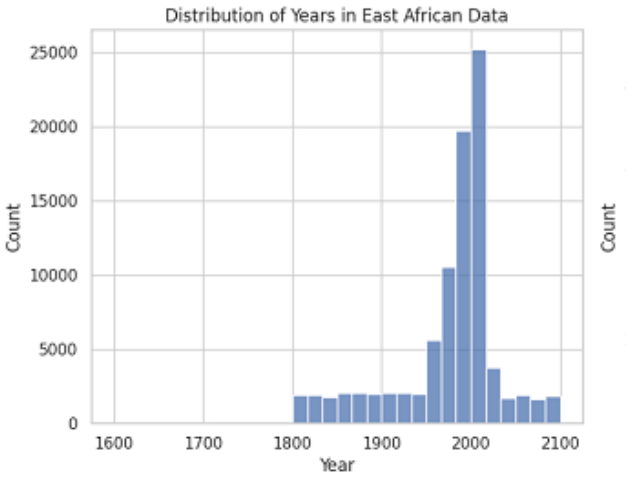


**Figure 2: A figure showing temporal coverage of the dataset**

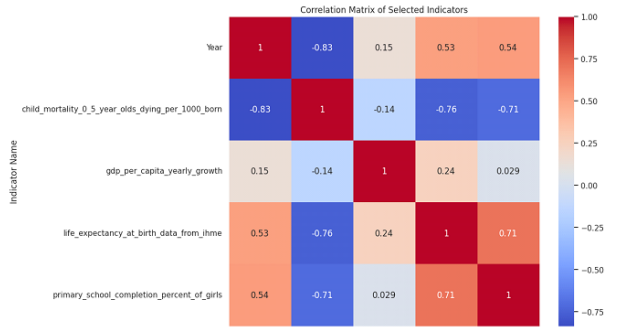


**Figure 3: Correlation heatmap for the few selected indicators**

associated with longer life expectancy. **Child Mortality vs. Life Expectancy**: The negative correlation suggests that as healthcare and living conditions improve, resulting in lower child mortality, life expectancy increases. **Education and Life Expectancy**: The positive correlation between education (primary school completion) and life expectancy highlights the importance of education in improving overall life quality.

*6.6.1 Country-Specific Trends and Comparisons .* Each country showed unique trends in these indicators, reflecting different developmental paths and policy impacts; as shown in figure 3.

## 6.7 Distribution of Key Indicators

Each of the selected indicators (life expectancy, GDP growth, child mortality, school completion for girls, and internet usage) shows distinct distribution patterns. This highlights the diverse nature of these indicators across East African countries. As shown in figure 4, Histograms illustrated the distribution of individual indicators, such as life expectancy and GDP growth, allowing us to observe the frequency and spread of data points. This aids in identifying typical ranges and any deviations from expected patterns.

## 6.8 Time Series Analysis

The time series analysis offers a narrative of progress and decline. For instance, the gradual upward trend in life expectancy at birth indicates improvements in healthcare and living standards. However, the sharp peaks and troughs in indicators such as GDP per

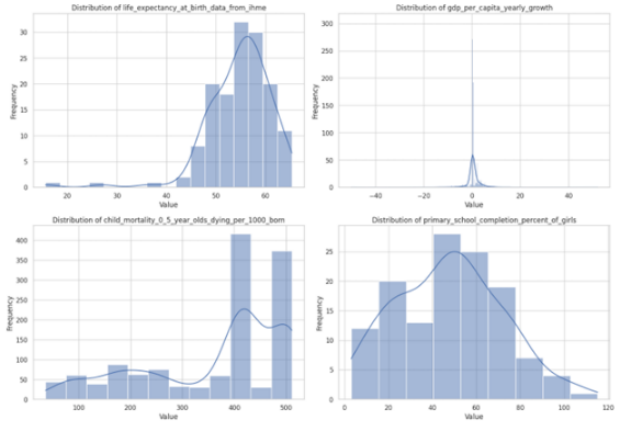


**Figure 4: A figure Showing histograms for the few selected variables**

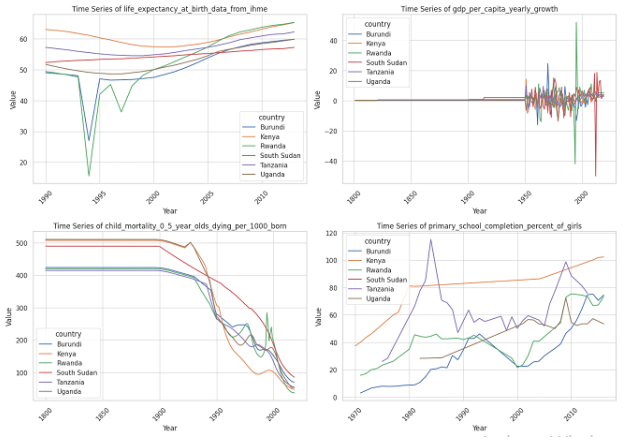


**Figure 5: A figure showing time series line charts for a few selected indicators**

capita yearly growth highlight the volatility and susceptibility of economies to external shocks. Each country presents unique trends, reflecting their individual development journeys and the influence of policies and global events. For example, the time series for primary school completion rates among girls shows varied trajectories, hinting at the differing impacts of educational policies across these nations.; as illustrated in figure 5.

# 7 IMPLICATIONS FOR APPLICATIONS

The insights gained from this analysis hold profound implications for applications in policy-making, investment decisions, and research directions. For instance, the correlation between child mortality and life expectancy could inform healthcare policy prioritization and resource allocation. Similarly, the educational attainment data can guide interventions aimed at bolstering educational infrastructure and opportunities for girls. The analysis underscores the need for nuanced policies that consider the intricate web of socioeconomic factors influencing development indicators. Furthermore, these findings can stimulate further research into the causal relationships between different indicators, potentially uncovering new avenues for enhancing development outcomes.

# 8 DASHBOARD OVERVIEW

## 8.1 Deployment

The dashboard was successfully deployed on the free hosting service provided by shinyapps.io, ensuring seamless access and user interaction. It is also integrated with GitHub, enabling continuous deployment and iterative development in response to user feedback. The dashboard is publicly accessible and can be visited at the following link: **HERE**.

## 8.2 Dashboard Sections

The dashboard's design is intuitive, comprising three primary sections: the sidebar, filters and slicers, and the main body content.

*8.2.1 Sidebar.* The sidebar serves as the navigation pane, featuring essential options such as 'About' and 'FAQs' which provide information about the dashboard's purpose and usage as seen in figure 6.

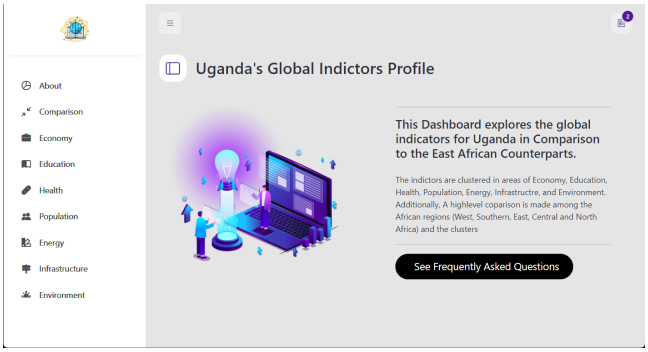


**Figure 6: The dashboard sidebar and about page**

As illustrated in figure 7, the 'Comparison' feature allows users to conduct a detailed comparative analysis of indicators by country, region, UN state membership, and landlocked status. Additional menu options are tailored to specific indicator categories including economy, health, society, environment, infrastructure, population, environment, among others, as illustrated in Figure 6.

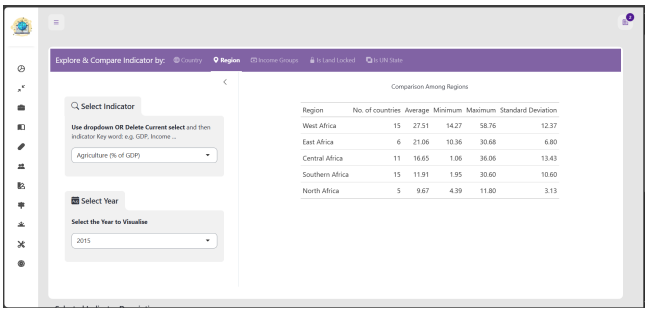


**Figure 7: The Comparison dashboard page**

*8.2.2 Filters and Slicers.* Each category within the sidebar is equipped with filters and slicers. These tools empower users to select indicators of interest and visualize data for particular years, enhancing the dashboard's interactivity and analytical capabilities.

*8.2.3 Main Body Content.* The main body of the dashboard is methodically divided into two sections. The first section showcases dynamic visualizations and tables, as depicted in Figure 7, facilitating an engaging data exploration experience. The second section provides a comprehensive description of each indicator, complete with details, sources, and tags, as demonstrated in Figure 8.

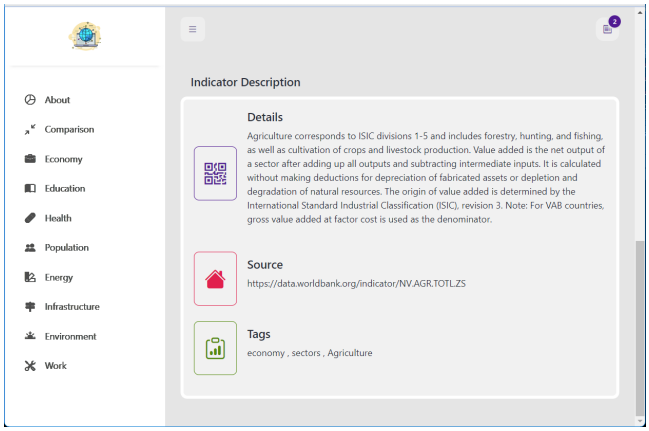


**Figure 8: The Indicator description dashboard section**

# 9 SAMPLE OUTPUTS

In order to experiment with the functionalities of the dashboard, let us assume the role of a research officer intending to explore the trends in Agriculture's contribution to the GDP, aiming to inform decision-making processes.

## 9.1 Comparative Analysis

*9.1.1 By Country.* The comparative analysis by country, as depicted in the chloropleth map (Figure 9), underscores disparities in the allocation of GDP towards agriculture across nations in the year 2020. Notably, countries such as Niger, Mali, Ivory Coast, and Ethiopia exhibit higher dedications to agricultural contributions compared to counterparts like Libya, South Africa, and Somalia. This observation alludes to the potential influence of geographical positioning, governmental policies, and prioritization of agricultural endeavors as key economic activities. Further in-depth analysis would help to unveil nuanced insights.

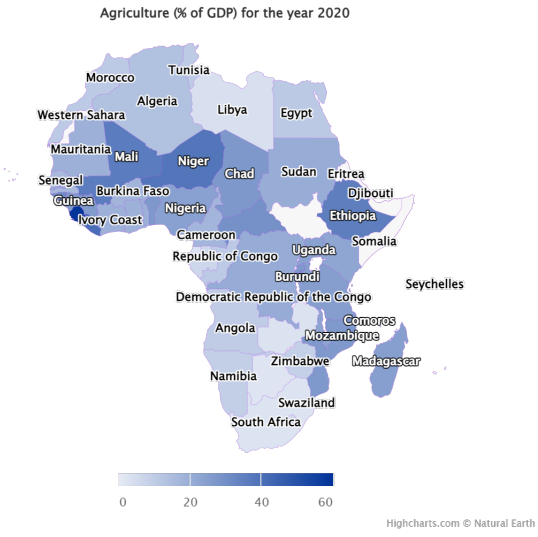


**Figure 9: A map of Africa showing the proportion of Agriculture to GDP by countries in 2020**

*9.1.2 By Region.* An examination of agriculture's GDP allocation by region, presented through summary statistics in Table 1, reveals interesting patterns. On average, West African nations lead with the highest proportion (27.66%) of GDP derived from agriculture, followed by East and Central African regions. Additionally, Southern Africa reports the lowest allocation (2.21%), with West Africa showcasing the widest range (54.9

| Region | No. of countries | Average | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|---|
| West Africa | 15 | 27.66 | 16.15 | 59.49 | 11.86 |
| East Africa | 5 | 24.56 | 20.88 | 28.62 | 3.09 |
| Central Africa | 11 | 16.02 | 1.68 | 35.56 | 11.11 |
| Southern Africa | 15 | 12.44 | 2.21 | 35.8 | 11.29 |
| North Africa | 5 | 9.92 | 3.59 | 13.78 | 3.78 |

**Table 1: Summary comparison among regions for Agriculture as proportion of GDP**

*9.1.3 By Landlocked Status.* As illustrated in Figure 10, countries with coastal access tend to exhibit a higher proportion of agricultural contribution to GDP, surpassing the average of 19%. Notably, a substantial majority (70.6%) of nations exceeding this average are coastal, while only 29.4% lack coastal access. This observation underscores the potential impact of geographical features on economic specialization and development trajectories.

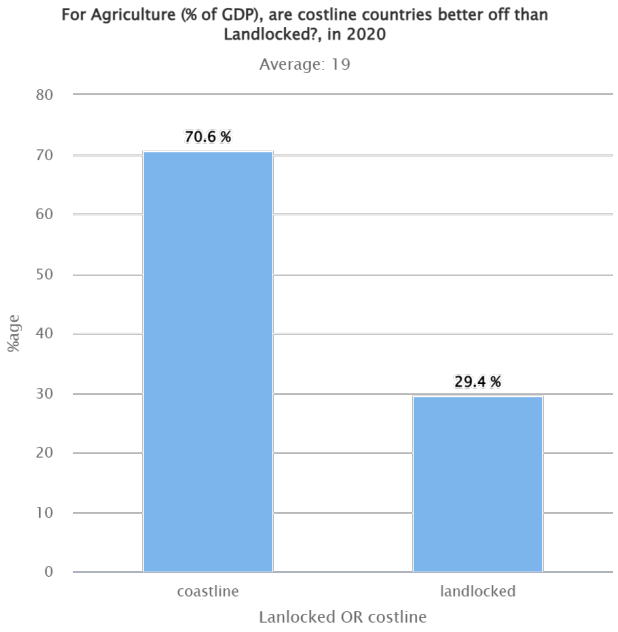


**Figure 10: Average proportion of agriculture contribution to GDP by landlocked status in 2020**

*9.1.4 By UN State Status.* Figure 11 highlights that a majority (52.9%) of UN states surpass the overall average (18.63%) of GDP contribution from agriculture in 2020. This suggests varying degrees of reliance on agriculture as a fundamental economic pillar among UN member states, with implications for policy formulation and development strategies.

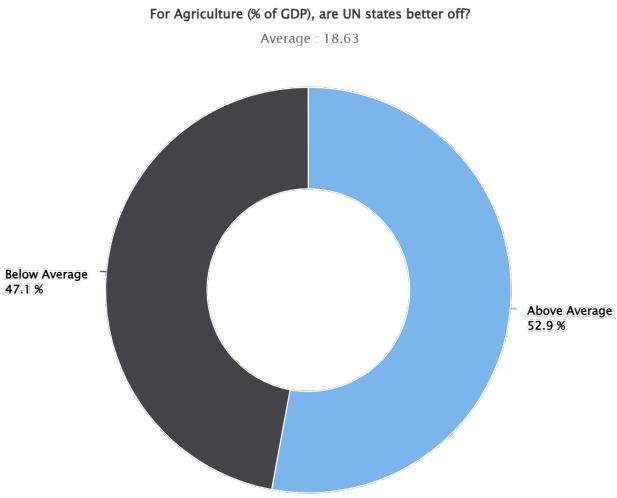


**Figure 11: Average proportion of agriculture contribution to GDP by UN State status in 2020**

## 9.2 Uganda Over Time

An examination of Uganda's agricultural contribution to GDP over time, as depicted in Figure 12, reveals a consistent downward trend. This trend signals a potential shift in economic priorities away from agriculture towards other sectors, potentially impacting the country's overall economic performance and development trajectory. Possible factors contributing to this trend could include urbanization, industrialization, and changes in agricultural productivity.

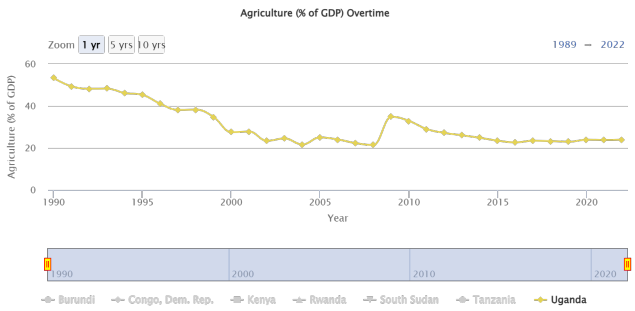


**Figure 12: Agriculture (% of GDP) for Uganda Overtime**

## 9.3 East African Comparison

Comparative analysis among East African countries reveals a general decline in the percentage contribution of agriculture to GDP over time (Figure 13). Notably, Burundi exhibits the highest contribution, while Kenya reports the lowest among the surveyed nations. This downward trend suggests evolving economic policies and activities aimed at enhancing productivity and diversifying the economy, potentially at the expense of agricultural productivity. It may be indicative of a broader regional shift towards industrialization and services, with implications for food security, rural livelihoods, and economic resilience.

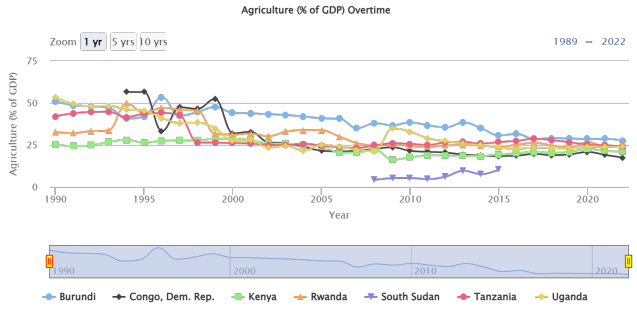


**Figure 13: Agriculture (% of GDP) for East African countries Over time**

## 9.4 Interpretation of the Experimental Results

The findings from the comparative analysis provide valuable insights into the dynamics of agricultural contribution to GDP across countries and regions in Africa. The disparities observed underscore the multifaceted nature of agricultural development, influenced by geographical, political, and economic factors.

The higher proportion of GDP derived from agriculture in certain West African nations suggests a significant reliance on agriculture as a primary economic activity in those regions. Conversely, the declining trend in agricultural contribution observed in Uganda and other East African countries reflects broader shifts in economic structures and priorities.

These trends have implications for policy formulation and development strategies. While diversification of economies away from agriculture may offer opportunities for industrialization and urbanization, it also poses challenges in terms of food security, rural livelihoods, and environmental sustainability.

In conclusion, the dashboard facilitates a comprehensive understanding of a given indicator, enabling informed decision-making processes and interventions aimed at promoting sustainable development and economic growth.

## 10 EVALUATION RESULTS

A group of 7 individuals comprising of 2 development officers specialist and 5 peers were approached to provide feedback and improvement suggestion of the dashboard on interactivity, usability and responsiveness. The responses were collected by one on one demos. The feedback collected included provision of download option for the tables and visualisations, and providing details about selected indicators in real time, which were incorporated in Version 2 of the dashboard.

Additionally, a frequently asked question section was added inte menu options with contact information to enable continued collection of feedback for iterative improvements to the dashboard.

## 11 FUTURE WORK

Based on the findings and learnings from the project, there are several improvements that can be made to this project, including the following:

- Integrating Predictive Analytics: Building models that can forecast future trends based on historical data could be immensely valuable for planning and policy-making. This would involve using machine learning techniques to predict the trajectory of key indicators on the dashboard.
- Expanding the dashboard to capture all countries in the world, which would facilitate comparison among continents, trading blocks and other demographic groupings.
- Including more statistical analysis such as Bivariate and Multivariate analysis including indicator interactions, country influence, time series decomposition, and many others.
- Deploying the dashboard on a more resourced server to enable faster loading time and handling multiple users at the same time as compared to the free current server which limits performance of the dashboard.

## 12 REFERENCES

[1] The World, B. World Development Indicators. *The World Bank* (2023).

[2] *Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development.* UNSD.

[3] Measuring Distance to the SDG Targets - OECD. *www.oecd.org* (2022).

[4] African Development, B. African Economic Outlook. *African Development Bank - Building today, a better Africa tomorrow* (2019).

[5] Gapminder Gapminder: Unveiling the beauty of statistics for a fact based world view. *Gapminder.org* (2023).

[6] United Nations Development Programme (UNDP). "Human Development Reports," UNDP.

[7] M. N. K. Boulos, et al. "Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation," International Journal of Health Geographics, vol. 10, no. 1, pp. 45.

[8] Gapminder. "Unveiling the beauty of statistics for a fact-based worldview.

[9] T. Y. Lin, J. J. Vandenberghe. "Interactive Data Visualization: Foundations, Techniques, and Applications," 2nd Edition.

[10] R. Heer, B. Shneiderman. "Interactive Dynamics for Visual Analysis," ACM Queue, vol. 10, no. 2.

Final Report Submitted On 20th April 2024