

Group 5 work

Charles Nyakumbo

Dotun Famoriyo

Visa QiTing

Overview of the data

The Africa Millennium Development Goals (MDGs) data refers to a collection of statistical indicators and data related to the progress of African countries towards achieving the Millennium Development Goals, which were a set of eight global development goals established by the United Nations in 2000. These goals aimed to address various social, economic, and environmental challenges in developing countries, including poverty reduction, education, gender equality, healthcare, and environmental sustainability.

Data types

- 📌 Country
- 📌 Name
- 📌 Goal
- 📌 Indicator
- 📌 Social Group
- 📌 Name Social Group
- 📌 Units
- 📌 Scale
- 📌 Frequency
- 📌 Date
- 📌 Value

<https://data.humdata.org/dataset/africa-millennium-development-goals/resource/16fab191-570c-4c59-b6d2-f0507e9b30ff>

Data Collection and cleaning Process

1. **Identify Reliable Data Sources:** Begin by identifying reputable and reliable sources of data related to the Africa Millennium Development Goals (MDGs). These sources can include international organizations, national statistical agencies, research institutions, and databases specific to the MDGs, such as the United Nations Development Programme (UNDP) or World Bank databases.
2. **Define Specific Data Requirements:** Determine the specific data requirements for your analysis or business case. This includes identifying the relevant indicators, time, geographical coverage, and any specific criteria or filters needed for your analysis.

3. **Search and Access Data Sources:** Once the data requirements are defined, search the selected data sources for the desired data. This can be done through online platforms, data portals, or by directly contacting the data providers. Many organizations provide free access to their data, while others may require registration or subscription.
4. **Understand Data Structure and Format:** Familiarize yourself with the structure and format of the data. Data may be available in various formats such as spreadsheets, CSV files, APIs, or online databases. Understand the variables, units of measurement, and any data limitations or caveats mentioned by the data source.
5. **Retrieve and Extract Data:** Retrieve the relevant data from the selected sources based on your defined data requirements. This may involve downloading datasets, using APIs to access data programmatically, or extracting specific data subsets from larger datasets.
6. **Clean and Preprocess Data:** Data retrieved from sources may require cleaning and preprocessing before analysis. This involves removing duplicates, handling missing values, standardizing variables, and ensuring consistency in data formats. Data cleaning and preprocessing steps are crucial for accurate and reliable analysis.
7. **Validate and Verify Data:** Validate the accuracy and consistency of the retrieved data by cross-checking with other trusted sources or conducting internal data verification processes. This helps ensure the reliability and integrity of the data.
8. **Organize and Store Data:** Once the data is cleaned and validated, organize it in a structured manner for easy access and analysis. Store the data in a secure and well-documented repository, ensuring proper data governance and data protection practices.

Challenges Encountered

- a. **Data Availability:** Not all data may be readily available or accessible. Some sources may have limited data coverage, gaps in certain indicators, or restricted access due to licensing or subscription requirements. To overcome this, alternative sources can be explored, or efforts can be made to request data directly from the relevant organizations or authorities.
- b. **Data Quality and Reliability:** Data quality issues, such as missing values, inconsistencies, or errors, can impact the reliability of the analysis. To address this, data validation techniques and cross-referencing with multiple sources can be employed to ensure accuracy and reliability. Data cleaning and preprocessing

techniques, such as imputation or outlier detection, can also be applied to address data quality issues.

- c. **Data Consistency and Standardization:** Different data sources may use different units of measurement, definitions, or methodologies, making it challenging to compare or combine data. Standardization efforts, such as converting units, harmonizing definitions, or applying common frameworks, can help ensure consistency across the data and facilitate meaningful analysis.
- d. **Time Period and Granularity:** Data availability may be limited to specific time periods or may lack the desired granularity for the analysis. In such cases, interpolation or extrapolation techniques can be used to estimate missing data points or aggregate data to the desired level of granularity. However, caution should be exercised in using extrapolated or estimated data, and the limitations should be acknowledged.
- e. **Data Bias and Representation:** Data may exhibit biases or limitations in terms of geographical coverage, population representation, or sector-specific focus. It is important to be mindful of these biases and limitations when interpreting and generalizing the findings. Supplementing the data with qualitative research or local insights can help provide a more comprehensive understanding of the context.
- f. **Data Security and Privacy:** Some data sources may have privacy concerns or restrictions on data usage due to confidentiality or legal requirements. Adhering to data protection and privacy regulations, obtaining necessary permissions, and anonymizing sensitive data can help address these concerns while ensuring compliance.

Overcoming these challenges requires a combination of careful data selection, thorough validation and cleaning procedures, and appropriate data handling techniques. It is crucial to document any assumptions, limitations, and steps taken to address data challenges to maintain transparency and integrity in the analysis. Additionally, seeking expert advice or consulting with domain specialists can provide valuable insights and guidance in navigating data-related challenges.

Interpretation and insights

a) Key findings

- ✓ Data gaps and quality: EDA can help identify data gaps and issues related to data quality, such as missing or inconsistent data. This insight is crucial for understanding the limitations of the dataset and interpreting the findings accurately.
- ✓ Regional disparities: EDA can reveal regional disparities in MDG achievement within Africa. It can highlight areas where certain regions or countries have performed better or worse compared to others, providing insights into the factors contributing to these disparities.
- ✓ Success stories and best practices: EDA can highlight success stories and best practices from countries that have achieved notable progress in specific MDG areas. These insights can provide valuable lessons for other countries and inform strategies for future development goals.

b) Data relationships (Relating the data to the supply chain and local production)

Our survey was based on the gaps in the domestic producers and local market in Africa, Kenya as the area of focus.

The hypothesis is: ***There is a significant gap between the quality suppliers and resellers in the supply chain.***

The survey objectives.

- a. To determine the relevance of technology used if any in the supply chain.
- b. To ascertain the major problems in the supply chain
- c. To establish the impact of middlemen on cost of products.
- d. To determine the extent of data usage in decision making within the supply chain.

Major issues in the supply chain as per the surveys conducted.

- ❖ Wastage since it is not easy to predict retail selling cycle.
- ❖ Increased cost of products to end user – Wholesale and distributor margin, cost of transportation.
- ❖ Poor technology
- ❖ Delays

Major relationships between gaps in the supply chain in Kenya and the Africa development goals data.

In the MDG data analysis the major goals based on the issues affecting Africa are;

Goal 1: Eradicate extreme poverty and hunger.

Goal 2: Achieve universal primary education.

Goal 7: Ensure environmental sustainability.

Goal 8: Develop a global partnership for development.

Goal 1: Eradicate extreme poverty and hunger.

Wastages in the supply chain can be controlled when extreme poverty and hunger is eradicated. Reducing wastages in the supply chain can contribute to addressing poverty and hunger by ensuring efficient utilization of resources.

Goal 2: Achieve universal primary education.

Achieving universal primary education can play a crucial role in addressing poor access to technology, particularly in Africa. By combining efforts in education and technology, there is a greater likelihood of addressing the issue of poor technology access and ensuring equitable opportunities for all.

Goal 7: Ensure environmental sustainability.

Ensuring environmental sustainability can contribute to reducing delays in the supply chain by addressing various factors that can disrupt the flow of goods and services.

- ✚ Efficient Resource Management
- ✚ Climate Change Mitigation and Adaptation
- ✚ Sustainable Transportation and Logistics
- ✚ Waste Reduction and Recycling

Goal 8: Develop a global partnership for development.

Developing a global partnership for development can play a significant role in addressing supply chain gaps by fostering collaboration, knowledge sharing, and resource mobilization.

By establishing a global partnership for development, stakeholders can work together to identify and address supply chain gaps. Collaboration, knowledge sharing, technology transfer, investment, and policy coherence can help build resilient, efficient, and sustainable supply chains that contribute to global development goals.

c) Limitations

When analyzing the Africa Millennium Development Goals (MDGs) data, it is essential to consider potential biases that may exist in the dataset. These biases can impact the analysis and findings, potentially leading to skewed or incomplete conclusions.

- i. **Data Availability Bias:** The availability of data may vary across countries and indicators. Some countries may have more comprehensive data collection systems than others, leading to uneven representation and potential underreporting of certain indicators. This can introduce bias in the analysis, as countries with limited data may be excluded or have less reliable data.
- ii. **Reporting Bias:** The accuracy and consistency of data reporting can vary between countries and over time. Inaccurate or inconsistent reporting can introduce bias and

affect the comparability of data across countries and regions. Differences in reporting practices or data collection methodologies can impact the reliability and validity of the analysis.

- iii. **Sample Selection Bias:** Data collection processes may involve sampling methods that could introduce biases. For example, surveys or studies might have excluded certain population groups or regions, leading to underrepresentation and potential bias in the findings. It is important to understand the sampling methods employed and the limitations they may introduce.
- iv. **Measurement Bias:** The measurement of MDG indicators may vary across countries and time periods. Differences in measurement techniques, data collection protocols, or data quality can introduce bias and affect the comparability and accuracy of the analysis. It is important to consider measurement inconsistencies when interpreting the findings.
- v. **Contextual Bias:** The MDG data may not fully capture the contextual factors that influence development outcomes. Socio-cultural, political, and economic factors unique to each country can impact progress towards the goals but may not be fully reflected in the data. Failing to account for these contextual factors can introduce bias and limit the understanding of the underlying drivers of development.
- vi. **Response Bias:** Surveys or questionnaires used to collect data on MDG indicators may be subject to response bias. Individuals or households may provide inaccurate or biased responses due to social desirability bias, recall bias, or other factors. This can affect the reliability and validity of the data and impact the analysis and findings.

Recommendations

1. It is important to acknowledge these potential biases and their impact when analyzing the Africa MDGs data. Researchers and analysts should exercise caution, critically evaluate the data sources and methodologies, and consider the limitations and potential biases when interpreting the findings. Sensitivity analyses and validation techniques can also be employed to assess the robustness of the results and mitigate potential biases.
2. While addressing poverty and hunger can contribute to reducing wastages in the supply chain, it is not a standalone solution. A comprehensive approach involving policy interventions, investment in infrastructure, education, and sustainable agricultural practices is necessary to achieve significant improvements in supply chain efficiency and reduction of wastages.
3. While achieving universal primary education is an important step in addressing poor technology access, it is necessary to recognize that further efforts are needed to bridge the digital divide fully. This includes targeted policies and investments in technology infrastructure, teacher training, digital content development, and internet connectivity. By combining efforts in education and technology, there is a greater likelihood of addressing the issue of poor technology access and ensuring equitable opportunities for all.

4. By integrating environmental sustainability principles into supply chain management, businesses can create more resilient, efficient, and agile supply chains. This, in turn, can help reduce delays, enhance operational efficiency, and improve overall sustainability performance.

References

Grunert, K. G., Verbeke, W., & Wills, J. (2014). Sustainability labels on food products: Consumer motivation, understanding and use. *Food Policy*, 44, 177–189. <https://doi.org/10.1016/j.foodpol.2013.12.001>

Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, 64(5), 416–424. <https://doi.org/10.1086/257839>

Powell, B., & Lal, D. (2008). Making poor nations rich : entrepreneurship and the process of economic development. <https://www.marketsandmorality.com/index.php/mandm/article/download/193/pdf>