# Multidimensional Analysis of Reproductive Health Indicators in Africa: A Clustering Approach for Identifying Disparities and Formulating Targeted Policies

**Djahid Saidoun[1], Mohamed Amine Belaidi, Ahmed Derdiche**

**DOI:** **https://doi.org/10.29358/sceco.v0i42.599**
**How to cite:** **Saidoun, D., Belaidi, M. A. and Derdiche, A. (2025). Multidimensional Analysis of Reproductive Health Indicators in Africa: A Clustering Approach for Identifying Disparities and Formulating Targeted Policies.** ***Studies and Scientific Researches. Economics Edition, 42.*** **https://doi.org/10.29358/sceco.v0i42.599**
**Received:** **December 2024**
**Accepted:** **December 2025**

**Abstract**
*This research aims to interpret the key reproductive health indicators of African countries using cluster analysis. By classifying these countries based on their reproductive health indicators, this study highlights their strengths and weaknesses in achieving the Sustainable Development Goals, particularly those related to ensuring maternal and child health, access to healthcare, protection of reproductive rights, combating sexually transmitted diseases and harmful practices, and reducing maternal and child mortality rates. A combined descriptive-analytical approach was employed, enabling us to provide a comprehensive overview reflecting the reality of reproductive health in Africa. Additionally, advanced statistical methods were utilized through the implementation of two of the most significant clustering techniques. Furthermore, the TANAGRA 1.4.50 software, which encompasses a broad array of algorithms used in exploratory statistics, data analysis, and processing, was leveraged. The latest data on reproductive health indicators for African countries (42 countries), compiled from the United Nations Development Programme and the United Nations Population Fund in 2023, were also employed. The study revealed significant disparities in reproductive health indicators among African countries. These countries were classified into four main groups based on their priorities, the severity of their situation, and a set of demographics, social, and economic characteristics. Additionally, the study confirmed the existence of African countries categorized as being at risk, necessitating urgent interventions through the adoption of comprehensive strategic plans to strengthen families and support their reproductive health.*

**Keywords**
cluster analysis; reproductive health; infant mortality; maternal mortality; family planning

## Introduction

Reproductive health is considered very important in the life cycle and a lifestyle approach that includes the challenges faced by all individuals and communities at different stages of their lives. These challenges care about family planning, safe motherhood, pre- and postnatal care, treatment of sexually transmitted infections, infertility, and miscarriage.

Since the International Conference on Population and Development in Cairo (1994), ensuring access to reproductive mothers and adolescents health services has become a priority on the international agenda. As a result, studying reproductive health indicators has become indispensable as a crucial tool for policymaking and developing countries health problems (Velasco, Solsona, & Burgunder, 2011). Many UN agencies have prioritized providing data on the realities of families, with a focus on mothers and children. National household surveys, for example, offer statistics spanning the past 30 years to create health maps of countries.

[1] Corresponding author: Djahid Saidoun, email d.saidoun.etu@univ-blida2.dz

When we examine the statistics of reproductive health indicators in Africa, we find that some of these indicators remain relatively stable, sometimes undergoing minor changes. There is a variety between countries. Some African countries have achieved targets aligned with the Millennium Development Goals, while others remain at risk. Furthermore, these countries have witnessed a worrying decline in maternal and child health in recent years. Mortality rates have increased in some African countries facing severe humanitarian crises, or remained unchanged in others. The overall maternal mortality rate in Africa has reached 267 deaths per 100,000 live births, compared to 140 globally, a fact explained by the average age at childbirth for women, which stands at 65,9 (Organization World Health, 2023).

Among the primary causes of maternal mortality in Africa are severe bleeding, high blood pressure, infections related to pregnancy, complications from unsafe abortions, and pre-existing conditions that can be exacerbated by pregnancy (such as HIV/AIDS and malaria).

Regarding child mortality, according to the latest statistics, the infant mortality rate (under one year) has reached 44,1 deaths per 1,000 live births, while the under-five mortality rate has reached 63 deaths per 1,000 live births. These are very high rates that place Africa in a perilous position compared to the global average (27,3 and 36, respectively). Among the primary causes of child mortality, according to the World Health Organization, are preterm birth and complications related to birth (birth asphyxia and birth trauma), infections in newborns, and congenital malformations, in addition to a lack of quality care at birth or skilled care and treatment immediately after birth and in the early days of life (Organization World Health, 2023).

As for the total fertility rate in Africa, it remains high, reaching 4,07 live births per woman, with a crude birth rate of 31,1 births per 1,000 population. The total population of Africa has reached 1,481 billion, representing 18.17% of the world's population. Niger ranks first with a fertility rate of 7,15 children per woman, followed by Somalia with 6,12 children per woman (United Nations, 2023). Despite slight but noticeable improvements in some reproductive health indicators in African countries due to a series of strategies implemented by African governments in recent years as part of achieving the Millennium Development Goals set within the United Nations Development Programme, compared to global statistics, these countries still remain in a zone of risk and concern. This has led them to face future challenges that include a range of programs and policies aimed at reducing maternal and infant mortality, improving health conditions, and curbing the spread of diseases and epidemics such as HIV/AIDS, tuberculosis, Malaria, Ebola, and other deadly epidemics. Differences in reproductive health indicators among African nations are a direct reflection of the influence of social, economic, cultural, and environmental determinants that shape the living conditions of individuals and communities. The variations in maternal and child mortality rates, disease prevalence, and fertility levels are not random; rather, they are a result of interconnected factors such as poverty, education levels, women's empowerment, access to basic services, and even conflicts and security conditions. This theoretical framework allows us to move from merely describing indicators to understanding the deeper mechanisms that create and sustain these health disparities (Lobao & Saenz, 2002).

To deepen our understanding of these disparities, we can draw upon theoretical frameworks such as interactive structuralism theory, which explains how the physical and social environment in residential areas contributes to shaping health and well-being conditions and helps to interpret patterns of health inequality (Gieryn, 2000). Spatial and place theory further supports this understanding by emphasizing that places are not just backgrounds but active elements that influence health practices and individuals' perceptions of their health and well-being (Seamon & Sowers, 2008). Based on this, classifying countries into homogeneous groups according to reproductive health indicators not only highlights existing disparities but also points to the varying "reproductive capacities" of different communities. These capacities, in turn, stem from the differing abilities of countries to provide health infrastructure, education, and supportive social environments.

Building on this understanding, it becomes essential to adopt a developmental model for health equity when formulating targeted policies. This model aims not only to improve health indicators in themselves but also to address the structural, geographical, and social barriers that prevent the achievement of justice and fairness in the distribution of reproductive health resources and services (Braveman & Gruskin, 2003). Understanding these homogeneous groups ensures that designed health programs are comprehensive and geared toward achieving the maximum impact on the lives of individuals and communities in Africa by targeting specific gaps and promoting equitable access to care.

Through this discussion, the problem statement of our study crystalized, aiming to shed light on the reality of reproductive health indicators in African countries by highlighting the extent of the gap and disparities in strengthening health systems, through posing the following questions:

- What are the current levels and trends of reproductive health indicators in African countries?
- Are there significant spatial variations in reproductive health indicators among African countries?
- How can African countries be classified based on reproductive health indicators to identify key disparities, thereby supporting the formulation of targeted and effective health policies?

To encompass this problem, the following hypotheses were formulated:

- Reproductive health indicators in Africa are characterized by significant variation; despite slight improvements in some, other indicators may show stability or deterioration, especially in regions facing humanitarian and economic challenges;
- There are significant regional and international spatial disparities in reproductive health indicators among African countries, reflecting fundamental gaps in their health systems that negatively impact maternal and child health;
- African countries can be classified into homogeneous groups based on the levels of their reproductive health indicators, relying on clustering methods used in multivariate analysis, which allow for monitoring regional and international differences in reproductive health.

This study aims to encompass all aspects that contribute to human well-being and reproductive and sexual health, providing a safe and healthy environment and forming a sound family. Furthermore, it seeks to deduce the extent to which reproductive health indicators contribute to the organization and stability of societies and their utilization in preparing healthcare programs based on sustainable development principles.

This research also aims to:

- Identify the most important statistical methods used in cluster analysis that aim to process, analyze, and classify multivariate data;
- Understand the reality of the most important reproductive health indicators for African countries, monitor their trends and development, and determine the extent of spatial (regional and international) gaps and disparities;
- Provide a comprehensive overview of the classification of African countries according to reproductive health indicators to reveal similarities, weaknesses, and strengths among them. This is to enable decision-makers to monitor key challenges and interventions in promoting healthcare components and reducing maternal and infant mortality.

In this research, we relied on two main methodologies: the descriptive analytical approach, used for precise exploration of collected data, which serves as an initial step and a general overview of the reality of reproductive health in Africa. The advanced statistical approach was relied upon as a complementary method to describe cluster analysis techniques.

This study aimed to analyze the geographical variation in key indicators of maternal and child health across 27 countries in Sub-Saharan Africa, using nine (09) key indicators falling under the Sustainable Development Goals related to achieving health targets while increasing health inequality. Utilizing data from 27 Demographic and Health Surveys, mapping results showed significant differences (up to more than 20 percentage points) in the coverage of key interventions between contiguous sub-regions and within countries. The analysis also revealed the presence of spatial clusters with "hotspots" of very low coverage for all indicators, while on the other hand, some countries showed spatial clusters with high coverage.

## Review of Literature

Burgert-Brucke et al. (2015) focused on the geographical variation in key maternal and child health indicators across 27 Sub-Saharan African countries, using nine (09) indicators under the health-related Sustainable Development Goals and health inequality. Using survey data, maps revealed significant differences (up to >20%) in intervention coverage between regions and within countries, with "hotspots" of low coverage and "cold spots" of high coverage. This indicates the non-homogeneity of health challenges and the necessity for geographically targeted interventions.

Yaya et al. (2021) conducted a study on the spatial distribution of maternal mortality in Africa and the impact of social determinants of health, using data from 54 African countries for the period 2006-2018. Results showed that maternal mortality in Africa is very high, with an average maternal mortality ratio of 415 deaths per 100,000 live births in 2018, with a high concentration in Central and West Africa ("hotspots" in 7 countries). In contrast, South Africa and Namibia formed "cold spots" (groups with a low maternal mortality rate), while eight (08) North and West African countries formed "low-high" clusters, meaning they had a significantly low mortality rate yet were located within the range of high-rate countries. Regression and Bayesian network analyses also revealed that gender inequality and a shortage of skilled birth attendants are the strongest social determinants affecting maternal mortality disparities in Africa.

Udo and Doctor (2016) reviewed trends in facility-based deliveries in Sub-Saharan Africa and their association with neonatal mortality, using Demographic and Health Survey data for 27 African countries during 1990-2014. Results showed an increase in the average of facility-based deliveries from 44% to 57%. Rural areas experienced a larger increase (16%) compared to urban areas (6%), reducing the urban-rural gap by 2%. The study found an inverse association between facility-based deliveries and neonatal mortality in both surveys ($R2=0.20$, $p=0.019$ in the initial survey; $R2=0.26$, $p=0.007$ in the latter survey). It recommended equating rural areas with urban areas when formulating policies to reduce neonatal mortality and achieve Sustainable Development Goals.

Adams et al. (2023) analyzed the prevalence of HIV/AIDS in Sub-Saharan Africa during 1990-2018, classified countries by prevalence rates, and monitored the nature of the relationship between HIV prevalence and countries' socioeconomic status. Cluster analysis techniques (hierarchical and K-Means partitioning) were applied. Results revealed that HIV prevalence increased from 6.74% in 1990 to 9.13% in 1995, then decreased to 2.60% in 2018. Countries were divided into three groups, with Group 3 including only Lesotho and Eswatini. South Africa, Zambia, Zimbabwe, Namibia, Malawi, Mozambique, and Botswana recorded the highest prevalence rates, and the rest of the world was classified as part of Group 1.

Yaya et al. (2018) investigated the impact of women's empowerment on contraceptive use in 32 Sub-Saharan African countries, based on data from 474,622 women from Demographic and Health Surveys using multidimensional analysis and logistic regression models. Results revealed that 67.9% of women used contraception, ranging from 4.9% in Chad to 72% in Namibia, with significant variation among countries. Women's empowerment was positively and significantly associated with contraceptive use, with higher use among educated women, working women, and those from wealthy households. The study further asserts that promoting women's empowerment, including their decision-making autonomy, participation in paid work, reduction of domestic violence, and improved education, will lead to increased contraceptive use and improved maternal and child health services.

Habte et al. (2024) evaluated the level and determinants of maternal health service utilization (antenatal care, skilled birth attendance, and postnatal care) in West Africa, using data from 89,504 married women from 12 countries during 2013-2021. Multidimensional analysis was employed using STATATA software, Version 16. The study concluded that 66.4% of women partially utilized maternal health services, while 23.8% utilized them adequately. Togo recorded the highest adequate utilization rate (56.7%), while Nigeria recorded the lowest (11%). Results showed that maternal education, place of residence, wealth index, media access, health insurance, husband's attitude towards wife-beating, and women's decision-making autonomy are all important factors influencing maternal health service utilization.

These studies collectively highlight the significant variability in reproductive and maternal and child health indicators across African countries. They jointly emphasize the urgent need to adopt a comprehensive and multidimensional approach to address reproductive health challenges in these nations. They consistently point to the existence of "hotspots" and "cold spots" for reproductive health indicators, indicating the non-homogeneity of health challenges across the continent. This underscores the importance of geographical and spatial analysis to understand the uneven distribution of health problems, in addition to the significance of examining other intervening factors such as socioeconomic status, maternal education level, women's empowerment, access to media, spousal attitude, and health insurance. Furthermore, these studies relied on large and comprehensive datasets from surveys, enhancing the reliability of the results and facilitating comparisons between countries and regions. The use of multidimensional analysis and advanced statistical techniques, such as cluster analysis, regression analysis, Bayesian networks, and logistic regression models, enabled the clear unveiling of complex relationships between variables and the identification of influential factors.

## Cluster analysis

Cluster analysis is a multivariate technique where the input data is presented in a matrix format, showing how similar each observation or case under study is to all other observations or cases (Wiley & Hartigan, 1975). As Estivill-Castro (2002) defined it, it is a set of procedures aimed at classifying and ordering a group of cases or variables in specific ways into groups called clusters, such that all cases classified within a cluster are homogeneous among themselves and share specific characteristics, differing from cases in other clusters.

According to Anil and Richard (2007), cluster analysis can be divided into two main groups, which are based on the structure of their output namely: hierarchical nonhierarchical (Partitioning) clustering methods. Hierarchical clustering also known as hierarchical cluster analysis is an algorithm that groups similar objects into groups called clusters. The clusters are merged (agglomerative methods) or split (divisive methods) step-by-step based on the similarity measure. In this section, we will highlight the most important of these methods based on the applied study.

### K-Means cluster analysis

His method is considered one of the most popular methods used in non-hierarchical cluster analysis. It is based on classifying several cases into homogeneous groups in terms of specific characteristics or features, relying on algorithms that can handle many observations. Estivill-Castro (2002) also termed this method "Quick Clustering", as it significantly reduces the time required for processing, analysis, and classification.

The k-means clustering methods have two phases of iteration namely: the assignment or initialization phase which involves an iterative process where each data point is assigned to its nearest centroid using Euclidean metric; the next is the centroid update phase, where clusters centroids are updated given the partition obtained by the previous phase. The iterative process stops when no data point change clusters or some maximum number of iterations is reached (Slonim et al., 2013).

### Hierarchical cluster analysis (HCA)

It is one of the hierarchical clustering methods that does not require prior knowledge of the number of clusters into which several cases are to be classified. This type of analysis is suitable for small samples, as it links similar groups in the data using a computationally efficient technique (Rencher, 2003).

The classification consists of a series of divisions starting from a single group containing all the elements to n groups, each containing one element (Everitt et al., 2011).

$$N(n, g) = \frac{1}{g}\sum_{k=1}^{g}\binom{g}{k}(-1)^{g-k}k^{n} \text{ formula (1).}$$

There are two types of hierarchical clustering: Single Linkage, the oldest and simplest method, also known as Nearest Neighbor. It considers distances or similarity coefficients between observations. The method starts by linking the two closest elements, forming a new cluster. This process continues based on the following formula (2) (Rencher, 2003):

$$D_{I,J} = \text{Min}\ (d_{i,j}) \text{ formula (2).}$$

Where:

- $D_{I,J}$ **:** Represents the cluster resulting from merging the two closest elements in the distance matrix;
- $d_{i,j}$ **:** Represents the distance between elements (i, j). Subsequently, we calculate the distance between the new cluster and all other clusters, denoted by K.

Complete Linkage Clustering, also known as Farthest Neighbor, is a clustering method that begins by grouping observations. It starts by merging the two closest observations to form an initial cluster. Unlike Single Linkage Clustering, Complete Linkage calculates the distance between the chosen cluster and all other clusters based on the farthest distance. This is typically calculated using the following formula (3) (Ayadda, 2012).

$$D_{(I,J)k} = \text{Max}\ (d_{ki}, d_{kj}) \text{ formula (3).}$$

### Expectation–Maximization (EM):

The Expectation-Maximization (EM) algorithm is a tool employed to find the parameter values that maximize the likelihood in a statistical model. This is particularly useful in scenarios where direct equation solving is intractable and the model involves latent variables in addition to unknown parameters. In other words, either there are missing data points or the model can be simplified by assuming the existence of additional, unobserved data points (Avinash, 2024).

This method relies on two alternating steps that are repeated until an optimal solution is reached (Bailey & Elkan, 1995).

- **Expectation step:** In this step, the expected value of the missing or incomplete data is calculated, assuming initial values for the parameters.

Estimate $\boldsymbol{Pr}\,[\mathbf{z} \mid \mathbf{x}, \boldsymbol{\theta}_\mathbf{n}]$, for all values of z. (construct the lower bound);

- **Maximization step:** In this step, the parameter values are updated in a way that increases the probability of obtaining the observed data, considering the expectations calculated in the previous step.

$$\boldsymbol{\theta_{n+1}} = \mathbf{arg\,max} \sum_{\mathbf{z}}^{\boldsymbol{\theta}} \mathbf{Pr}\ [\mathbf{z} \mid \mathbf{x}, \boldsymbol{\theta}_\mathbf{n}]\ \mathbf{log\,Pr}[\mathbf{x}, \mathbf{z}|\boldsymbol{\theta}]$$ **formula (4).**

This can be solved using traditional MLE methods with weighted samples. These two steps continue to alternate until the change in the parameter values stops, indicating the achievement of the optimal solution and thus linking all states until convergence occurs between them.

This is also a method for performing maximum likelihood estimation in the presence of latent variables. It does this by first estimating the values of the latent variables, then refining the model, and then iterating these two steps for all states until convergence occurs between them. It is an effective method that is commonly used for density estimation of missing data, such as clustering algorithms like the Gaussian mixture model.

Although the EM algorithm may sometimes be susceptible to local maxima when estimating parameters by maximizing the likelihood function of observed data, it possesses three characteristics that make it a user-friendly method (Hardie et al., 1997):

- Ability to handle a large number of variables simultaneously;
- Ability to find good estimates for all missing and incomplete data;
- Ability to model complex, multidimensional data clustering.

## Monitoring and analyzing key reproductive Health indicators in African Countries

Reproductive health indicators encompass various areas that collectively contribute to the formation of a cohesive family enjoying a healthy and safe life. These indicators are interconnected, varying across time, place, and the desired objective, such as access to healthcare and health education, ensuring the rights and needs of family planning, maternal and child health, as well as combating sexually transmitted diseases and harmful behaviors.

The African continent remains a focal point for researchers and experts due to the alarming trends in reproductive health indicators. Despite substantial financial allocations and efforts to restructure health systems and expand infrastructure across the continent's various regions, with the aim of reducing child and maternal mortality rates and improving overall health levels to achieve Sustainable Development Goal 5 by 2030, empirical evidence and recent statistics have revealed a significant global gap and identified numerous challenges.

### Child mortality rate

The child mortality rate is considered a crucial indicator reflecting a country's health level, as it directly mirrors the extent of its social, economic, and cultural development. The United Nations also considers it a significant measure for the Human Development Index (United Nations, 1994).

A report prepared by the United Nations Inter-Agency Group on levels and trends in child mortality by the Sustainable Development Goals region from 1990 to 2022 indicates a slight decrease in the under-five mortality rate in Africa in 2023 compared to 2022. It decreased from 64.49 deaths per 1,000 live births to 63.28 deaths per 1,000 live births. Infant mortality also witnessed a noticeable decrease during the same period, declining from 44,17‰ to 43,11‰.

Compared to the global rate, it is quite high. The under-five mortality rate worldwide reaches 36,15‰, while the infant mortality rate reaches 26,72‰ globally. To identify the region most affected by child mortality through monitoring the regional geographic distribution of these indicators, the data is represented as shown in Figure 1.

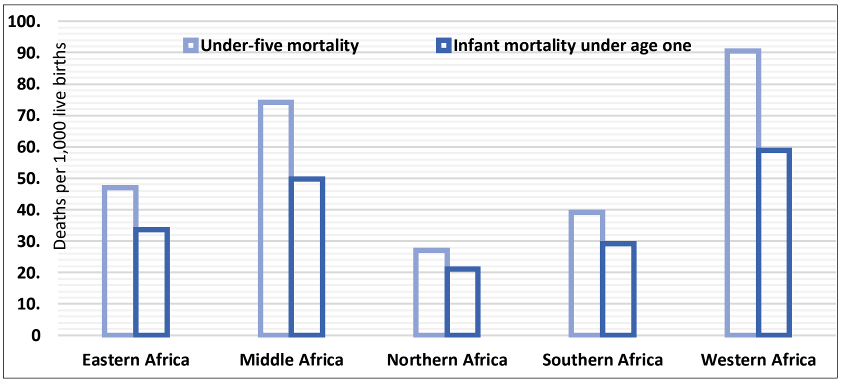



Figure 1 Distribution of child mortality in Africa by geographic regions in 2023.
Source: prepared by the authors based on United Nations data (2023)

As illustrated in the figure above, child mortality is significantly concentrated in West Africa, where deaths among children under five years reached their highest levels at a rate of 90,6 per 1,000 live births. Infant mortality rates in the same region were recorded at 58,9 per 1,000 live births. Central African countries follow, with respective mortality rates of 74,1 and 49,7 per 1,000 live births for children under five and infants. East African countries and Southern African countries exhibit similar rates, with mortality rates under five years of 47,02 and 39,2 per 1,000 live births, respectively. North African countries recorded comparatively lower child mortality rates of 27,17 and 21,07 per 1,000 live births for children under five and infants, respectively.
The distribution of child mortality by country is detailed in Figure 2, as follows:

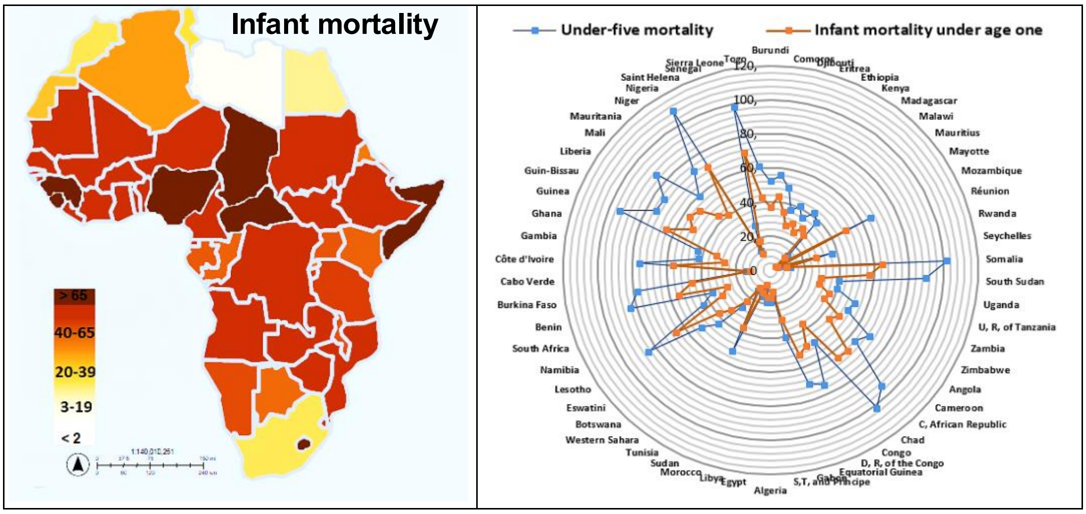



Figure 2 Distribution of Infant and Under-Five Mortality by Country for 2023.
Source: prepared by the authors based on United Nations data (2023)

Nigeria topped the list of African countries with the highest rates of infant and child mortality, recording a record infant mortality rate of 70,82‰ and an under-5 mortality rate of 109,28‰. Sierra Leone, Lesotho, the Central African Republic, Guinea, Chad, and Somalia followed, with infant mortality rates reaching 65‰ and under-5 mortality rates ranging between 85-102‰. These two rates were similar for South Africa, Côte d'Ivoire, Liberia, Benin, Mali, Equatorial Guinea, Mozambique, Congo, Angola, Burkina Faso, Cameroon, Mauritania, Togo, and Niger, with recorded infant mortality rates ranging between 40-57‰ and under-5 mortality rates between 58-94‰. Countries with the lowest recorded mortality rates were Libya, Seychelles, Mauritius, Cape Verde, Tunisia, Morocco, Algeria, and Egypt, with child mortality rates ranging between 8,9‰ and 18,9‰.

### Maternal mortality rate

In line with achieving Sustainable Development Goal 3.1 by 2030, the World Health Organization has issued a consensus statement and a strategic paper on ending maternal deaths, highlighting the need to increase coverage of quality maternal health care and improve women's ability to make their own decisions about their sexual and reproductive health (United Nations, 2022).

Among the factors that have hindered real progress towards achieving the SDGs in Africa, including the goals related to reducing maternal mortality, are fragile humanitarian situations, conflict and post-conflict situations, and disasters. According to the Fragile States Index 2023, ten (10) countries were in a very high alert situation: Burkina Faso, Yemen, Somalia, South Sudan, Syria, the Democratic Republic of Congo, the Central African Republic, Chad, Sudan, and Afghanistan (United Nations, 2022). These countries have experienced high maternal mortality rates, with an average of 551 deaths per 100,000 live births in highly fragile states, more than double the global average.

According to a report released by the World Health Organization (WHO) and the United Nations Children's Fund (UNICEF) on maternal mortality trends between 2000 and 2020, Nigeria experienced a staggering increase in maternal deaths, topping the list of African countries. The country accounted for over a quarter of all global maternal deaths, representing 28.5%, or approximately 82,000 maternal deaths in 2020 (Bradley, 2012).

As for the United Nations' statistics in 2023, the global maternal mortality ratio was estimated at 232 deaths per 100,000 live births, a 35,1% decrease compared to 2000 over 23 years. Monitoring this indicator at the African continental level revealed significant disparities in maternal survival rates among countries. Compared to other regions of the world, sub-Saharan Africa was the only region with an extremely high maternal mortality ratio, estimated at 532 deaths per 100,000 live births. This represents 68% of global deaths and is distributed regionally as follows: 744 in West Africa, 529 in Central Africa, 338 in East Africa, and 147 in Southern Africa.

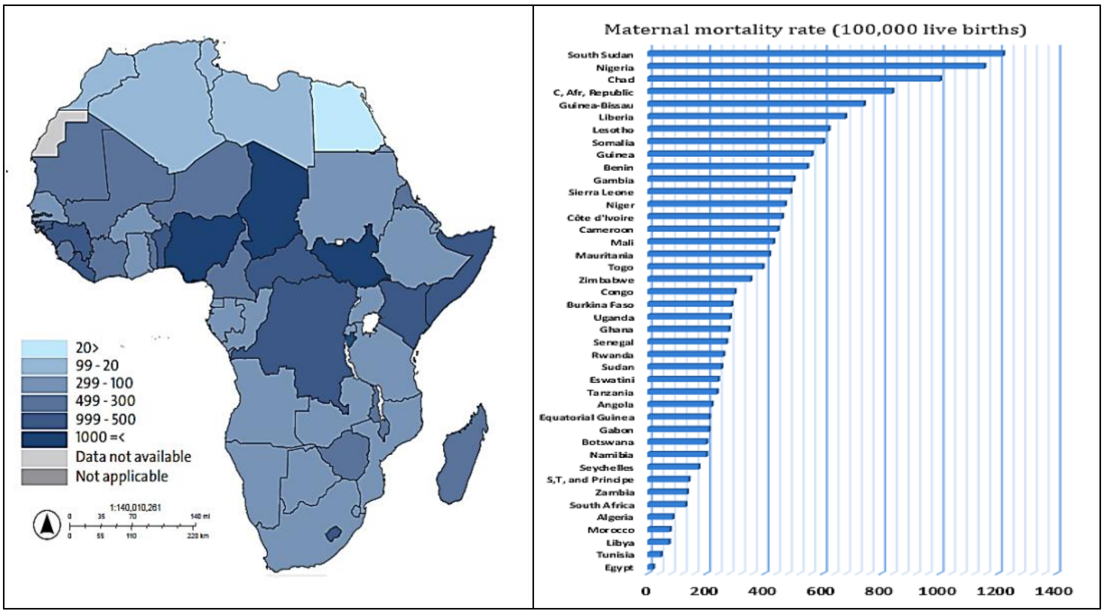



Figure 3 Geographical Distribution of Maternal Mortality Rates in Africa.
Source: prepared by the authors based on UNICEF and WHO data (2020)

When examining maternal mortality rates by African countries, it becomes evident that South Sudan and Nigeria have experienced exceptionally high rates. These nations recorded 1,118 and 1,153 maternal deaths per 100,000 live births, respectively. Additionally, several other countries have also shown elevated rates ranging between 500 and 1,000 deaths, including Chad, Central African Republic, Guinea-Bissau, Liberia, Lesotho, Somalia, Guinea, Benin, and Gambia.

Furthermore, 22 African countries reported maternal mortality rates between 200 and 300 per 100,000 live births, with most situated in Central and Southern Africa. Examples include Sierra Leone, Niger, Mali, Togo, Zimbabwe, Congo, Burkina Faso, Ghana, Senegal, Rwanda, Sudan, Tanzania, and Angola.

In contrast, countries that have achieved relatively low maternal mortality rates, as classified by the global standard of 20-99 deaths per 100,000 live births (Wang, 2012), are Algeria, Morocco, Tunisia, Egypt, and Libya.

## Levels of family planning use

The use of safe family planning is a fundamental human right. Family planning is crucial for gender equality and is a significant factor in achieving demographic dividend and economic growth. This was affirmed in the International Conference on Population and Development (ICPD), adopted by 179 countries in Cairo in 1994, and reaffirmed in the 2030 Agenda for Sustainable Development, emphasizing universal access to sexual and reproductive health care, including family planning. This enables all couples and individuals to exercise their right to decide freely and responsibly the number and spacing of their children, and to have access to information and means to do so (Warren & John, 2007).

Progress towards this goal is monitored globally through indicators (3.7.1: Percentage of women of reproductive age 15-49 years using modern contraceptives) (UNICEF, 2018). Studies have shown a generally inverse relationship between fertility and contraceptive use within countries, although this relationship can vary depending on the mix of contraceptives used by women, the occurrence of abortion, marriage and sexual activity patterns, as well as various other socioeconomic factors (United Nations, 2015).

The unmet need for family planning also reflects the percentage of fertile women who wish to space or limit their births but are not using any contraceptive method. A woman is also considered to have an unmet need if she is pregnant at the time of data collection but reported that the pregnancy was unintended (Bassey, 2023). This indicator measures the gap between women's reproductive intentions and their contraceptive behavior at the population level.
According to the latest statistics in 2023, globally, 77% of women who want to avoid pregnancy use modern contraception. Sub-Saharan Africa has the lowest proportion at 56%, compared to other regions. Additionally, 37% of women in North Africa use modern contraception.

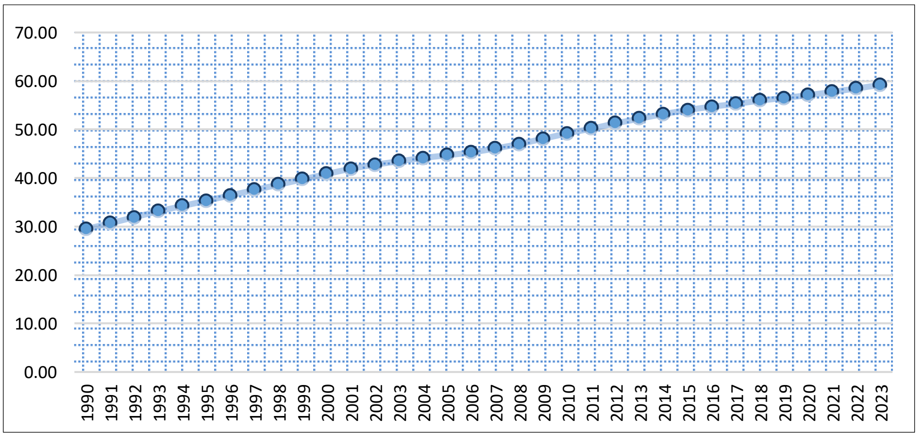



Figure 4 Proportion of demand for modern contraceptive use among women (aged 15-49) seeking to avoid pregnancy.
Source: prepared by the authors based on United Nations data (2023)

By tracking the evolution of this indicator, it is evident that all African countries experienced a significant increase in the modern contraceptive prevalence rate between 1990 and 2023. In fact, Sub-Saharan and North Africa recorded the most substantial global rise in meeting the demand for family planning through modern methods. This rate surged from 29,5% in 1990 to 59,2% in 2023. The distribution of this indicator across African countries is illustrated in Figure (5).

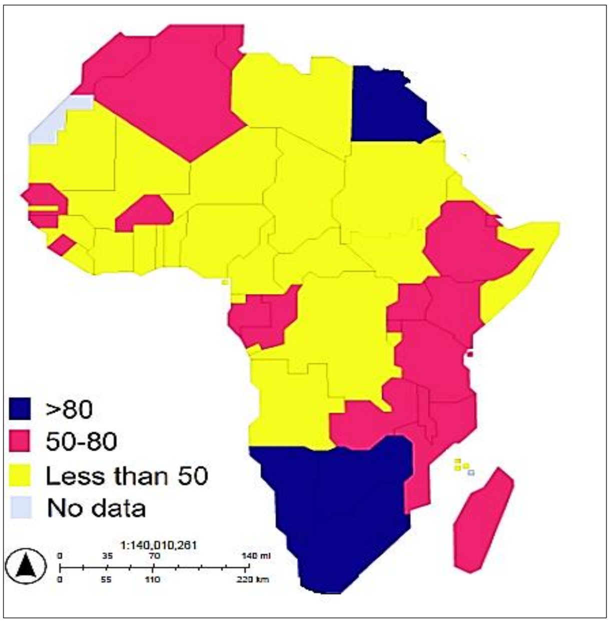



Figure 5 Percentage of women of reproductive age (15-49 years) whose need for family planning is met through modern contraceptive methods.
Source: United Nations, Population Division, Family Planning Indicators (2023)

The map reveals that over 80% of women of reproductive age (15-49) in South Africa, Namibia, Botswana, Egypt, Zimbabwe, Lesotho, and Eswatini meet their modern contraceptive needs. Conversely, in 22 African countries, less than half of women with a need for family planning are using modern methods. Among these, seven North African countries Algeria, Tunisia, Morocco, Senegal, Sierra Leone, Burkina Faso, Gabon, Congo, Ethiopia, Uganda, Zambia, Mozambique, Tanzania, Kenya, Rwanda, the Democratic Republic of Congo, Burundi, Malawi, and Guinea – have rates between 60-80%. The remaining African countries report rates below 50%.

Regarding the use of other methods such as male sterilization, intrauterine devices (IUDs), and traditional methods, there has been a slight global decrease of 2% in usage, with no significant changes over time. However, these methods have been significantly impacted by early efforts to introduce modern contraceptives through family planning programs, as seen in North Africa where intrauterine devices have become a preferred reverse method (WHO, 2021). Despite this slight decline in traditional methods, their usage remains widespread, particularly through the national family planning program in Thailand, which has seen widespread acceptance of IUDs, condoms, oral contraceptives, and hormonal injections due to the early involvement of auxiliary midwives in their distribution, leading to a rapid increase in their use (WHO, 2023).

Oral contraceptives represent the largest share of contraceptive use and are most widely used in North Africa, exceeding 70% of total use in Algeria and Morocco. Male condoms account for over 30% of use in Gabon, reaching 61%. Injectable contraceptives have the highest share of contraceptive use at 50% or more of total use in Ethiopia, Liberia, Malawi, and Zambia. Meanwhile, contraceptive vaccines are the most commonly used method of contraception in five countries, all located in Sub-Saharan Africa.

## Interpretation of the most important reproductive health indicators using cluster analysis

Having previously conducted a descriptive analysis of the state of reproductive health in African countries, we now turn to analyzing these indicators using cluster analysis. We will employ three main methods: K-means clustering, hierarchical clustering analysis (HCA), and expectation maximization (EM) clustering. These methods will help us to understand the strength and direction of the association between these indicators, as well as the importance of each indicator and its impact on the reproductive health of African countries.

### Method and procedures

In this study, we will employ cluster analysis, which is a statistical method that includes a set of mathematical techniques that enable us to form multivariate data into groups (clusters). The aim is to group observations with originally heterogeneous characteristics into homogeneous and similar groups (Romesburg, 2004).

Likewise, this analysis is considered an appropriate statistical technique for exploring the underlying structure of the dataset, as it works to reduce the number of cases and observations and classify them into similar observations according to some appropriate criteria, without resorting to prior knowledge of group membership or number, based on data merging algorithms (Ramdeen & Yim, 2015).

Through this approach, we will employ the following clustering methods:

- K-Means clustering ;
- Hierarchical Cluster Analysis (HCA) ;
- Expectation Maximization (EM) clustering ;

The use of K-Means, HCA, and EM algorithms is supported by a strong statistical and methodological rationale, enabling us to uncover complex disparities and classify African countries in a manner that supports the formulation of more effective and sustainable health policies. We chose to employ them in this study for several reasons:

- These algorithms are powerful tools for classifying 42 African countries into homogeneous groups (clusters) based on their levels of reproductive health indicators. Each algorithm uses different similarity criteria, thereby providing multiple insights into the underlying structure of the data;
- These methods enable us to identify regional and international disparities in reproductive health, pinpointing "hotspots" with significant challenges and "coldspots" with good performance;
- By integrating cluster analysis with Principal Component Analysis (PCA), we can extract principal components. Based on the clusters formed by the algorithms, this provides a deep understanding of the critical factors that shape the trajectory of reproductive health, particularly the studied aspects such as geographic location, socioeconomic status, women's empowerment, and health equity;
- Verifying the robustness of the results: Since each algorithm has its own assumptions, obtaining similar clusters across these methods confirms that the discovered structure is genuine to the data and not merely an artifact of a specific algorithm's choice. Furthermore, these methods allow for a deeper exploration of the data's structure: K-Means identifies distinct clusters, HCA reveals hierarchical relationships, and EM excels at handling complex and irregularly shaped clusters.

The methods mentioned above were used based on the TANAGRA program (TANAGRA 1.4.50), which is a program that includes a wide range of algorithms used in the field of exploratory statistics and data analysis and processing, in addition to multidimensional graphing that facilitates the visualization of the largest possible amount of data in an interactive manner. It also helps decision-makers in making immediate decisions.

### Study Sample and Statistical Treatment

This study was based on the latest data issued by the United Nations Population Division (UN Population Division) for 2023, which included the most important reproductive health indicators for 42 out of 54 African countries. This sample is sufficient for conducting a multidimensional analysis of reproductive health indicators in Africa.

A set of variables related to reproductive health were identified and are summarized as follows:

Table 1 list of variables considered in the study.

| N° | Variables | N° | Variables |
|---|---|---|---|
| 01 | Population | 12 | Anemia_prevalence_women |
| 02 | Births | 13 | Life_expectancy |
| 03 | Total_deaths | 14 | HIV_prevalence_all_ages |
| 04 | Maternal_deaths | 15 | Total_fertility_rate |
| 05 | Under_five_mortality_rate | 16 | HIV_prevalence_15_49 |
| 06 | Infant_mortality_rate | 17 | Crude_birth_rate |
| 07 | Life_expectancy_women | 18 | Comprehensive_health_coverage_index |
| 08 | Infant_mortality | 19 | Cause_of_death |
| 09 | AIDS_prevalence_total | 20 | Health_expenditure_per_capita |

| | | | |
|---|---|---|---|
| 10 | Death_rate | 21 | Birth_rate |
| 11 | Fertility_rate | 22 | Physicians |

Source: elaborated by the authors

**Data Limitations**

Despite the analytical value of this study in identifying the differences among African countries regarding reproductive health indicators, it is important to acknowledge certain limitations related to the dataset. These limitations are summarized as follows:

- **National-level aggregation:** The study relies on data aggregated at the national level. While this is useful for comparing countries, it overlooks significant internal variations within a single country, such as differences between urban and rural areas or among different socioeconomic groups;
- **Potential for underreporting:** The data is often based on national reporting systems and population surveys. As a result, this study faces the potential issue of underreporting in reproductive health data. This raises concerns about the accuracy and completeness of the data and, consequently, the reliability of the conclusions;
- **Incomplete country coverage:** The study is limited to only 42 out of 54 African countries due to the unavailability of data for the remaining countries. This means that the findings may not fully represent the entire continent. The excluded countries may have unique characteristics that could influence the overall patterns or be explained by other national policies.

To ensure the accuracy and reliability of the conclusions, the statistical analysis was preceded by a crucial step: data processing. This included testing the randomness of missing values and evaluating the suitability of applying the results of the three clustering algorithms (K-Means, HCA, EM), which led to the identification of four (04) clusters for each algorithm.

- **Testing randomness of missing Data and imputation:** This analysis aims to examine the problem of missing data within the used dataset to avoid obtaining biased results, carried out in two stages. First, Little's MCAR test was conducted to determine if data loss was random. Second, missing values were imputed using the Expectation-Maximization (EM) algorithm. The results are presented in Table 2 below.

Table 2 Missing Data Test and Imputation

| Variables | N | Missing | | Mean (1) | Means EM.[a] (2) | Deviation (2)-(1) | % |
|---|---|---|---|---|---|---|---|
| | | Eff.[b] | % | | | | |
| Population | 42 | 0 | 0,0 | 27733495,88 | 27733495,88 | 0,00 | 0,0 |
| Births | 42 | 0 | 0,0 | 856012,85 | 856012,86 | 0,00 | 0,0 |
| Total_deaths | 40 | 2 | 4,8 | 30,43 | 30,69 | 0,26 | 0,9 |
| Maternal_deaths | 37 | 5 | 11,9 | 3,72 | 3,67 | -0,05 | -1,4 |
| Under_five_mortality_rate | 42 | 0 | 0,0 | 192954,71 | 192954,71 | 0,00 | 0,0 |
| Infant_mortality_rate | 39 | 3 | 7,1 | 7,79 | 7,89 | 0,09 | 1,2 |
| Life_expectancy_women | 42 | 0 | 0,0 | 63,53 | 63,53 | 0,00 | 0,0 |
| Infant_mortality | 42 | 0 | 0,0 | 37897,02 | 37897,02 | 0,00 | 0,0 |
| AIDS_prevalence_total | 42 | 0 | 0,0 | 30,06 | 30,06 | 0,00 | 0,0 |
| Death_rate | 42 | 0 | 0,0 | 39,77 | 39,77 | 0,00 | 0,0 |
| Fertility_rate | 41 | 1 | 2,4 | 42,08 | 43,04 | 0,95 | 2,2 |
| Anemia_prevalence_women | 37 | 5 | 11,9 | 3,98 | 4,00 | 0,01 | 0,4 |
| Life_expectancy | 42 | 0 | 0,0 | 63,12 | 63,13 | 0,00 | 0,0 |
| HIV_prevalence_all_ages | 42 | 0 | 0,0 | 45,78 | 45,79 | 0,00 | 0,0 |
| Total_fertility_rate | 42 | 0 | 0,0 | 125,91 | 125,91 | 0,00 | 0,0 |
| HIV_prevalence_15_49 | 38 | 4 | 9,5 | 1,91 | 2,03 | 0,12 | 5,8 |
| Crude_birth_rate | 38 | 4 | 9,5 | 1,31 | 1,44 | 0,13 | 8,8 |
| Comprehensive_health_coverage | 39 | 3 | 7,1 | 39,71 | 40,87 | 1,16 | 2,8 |
| Cause_of_death | 35 | 7 | 16,7 | 48,85 | 48,67 | -0,19 | -0,4 |
| Health_expenditure_per_capita | 42 | 0 | 0,0 | 4287,92 | 4287,93 | 0,00 | 0,0 |
| Birth_rate | 38 | 4 | 9,5 | 4,77 | 4,70 | -0,08 | -1,6 |
| Physicians | 39 | 3 | 7,1 | 0,42 | 0,42 | 0,00 | -0,8 |

| |
|---|
| a. Test MCAR de Little : Khi-deux = 290,111, DF = 237, **Sig. = ,010** |
| b. Total number of missing data is **41, representing 4.4% < 30%.** |

Source: elaborated by the authors using SPSS version 23

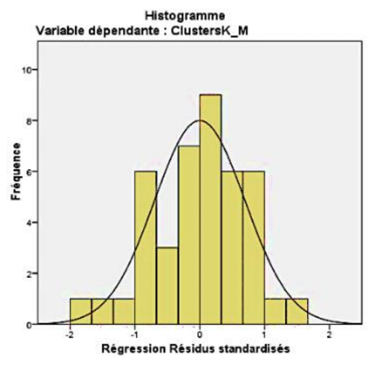
The results indicate that the proportion of missing data in the sample amounted to 41 missing values in total, representing 4.4%, which is relatively low. More importantly, Little's Test showed statistical significance (Sig. = 0.010), negating the assumption that data were Missing Completely At Random (MCAR). On the other hand, we observe slight differences between the means before and after imputation. Therefore, we conclude that the treatment was necessary and sound.

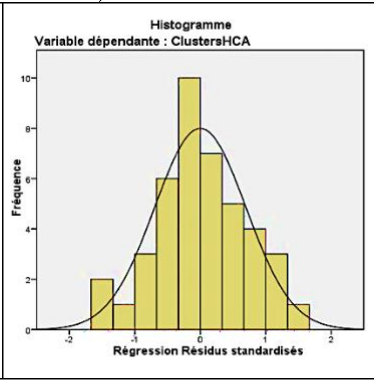
- **Goodness-of-Fit Test:** To verify the quality of the clustering method outputs, Regression Analysis was applied in this analysis. The extracted clusters from each algorithm were taken as the dependent variable, while reproductive health indicators were used as independent variables. Table 3 below illustrates the results of the Analysis of Variance (ANOVA) and the coefficient of determination ($R^2$).

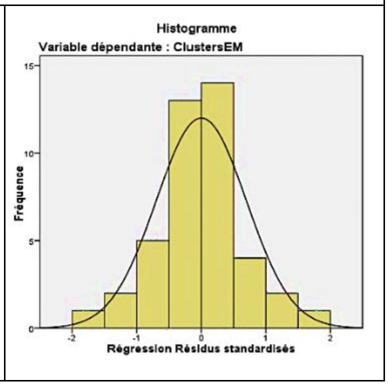
Table 3 Results of Clustering Methods Suitability Test

| Clustering methods | N° Clusters | Degree of Correlation | | | ANOVA[a] | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | R | R-squared | Ajst R-squared | Model | Sum of Squares | df | Mean Square | F | Sig. |
| K-Means | 04 | ,912[a] | ,831 | ,654 | Regression<br>Residual<br>Total | 25,331<br>5,145<br>30,476 | 21<br>20<br>41 | 1,206<br>,257 | 4,689 | ,001[b] |
| HCA | 04 | ,968[a] | ,936 | ,870 | Regression<br>Residual<br>Total | 24,792<br>1,684<br>26,476 | 21<br>20<br>41 | 1,181<br>,084 | 14,019 | ,000[b] |
| EM | 04 | ,918[a] | ,843 | ,678 | Regression<br>Residual<br>Total | 38,687<br>7,218<br>45,905 | 21<br>20<br>41 | 1,842<br>,361 | 5,105 | ,000[b] |

a. Dependent Variable: Clusters (K_m, HCA, EM).
b. Predictors: Independent Variables (21 variables).




Source: elaborated by the authors using SPSS version 23

From Table 3, we observe that all clustering methods (K-Means, HCA, EM) resulted in regression models with very high statistical significance, with correlation coefficients estimated at (0.91, 0.96, 0.91) respectively. The coefficient of determination explains that the three models account for a variance ranging from 83% to 93% of the total variance. However, the HCA method clearly demonstrates the best performance, explaining the highest percentage of variance (R-squared 0.936) and having the highest correlation value, which indicates its greater effectiveness in this scenario. Furthermore, from the three scatter plots, we observe that the distribution of standardized residuals for all methods is closer to a normal distribution, which enhances the reliability of the regression results for this model. Since all three methods produced regression models with very high statistical significance, this suggests that each can be a valid means for creating "Clusters" that can be explained by the independent variables.

- **Clustering justification:** to ensure the optimal number of groups extracted in each clustering method, we used the methods shown in Figure 6.

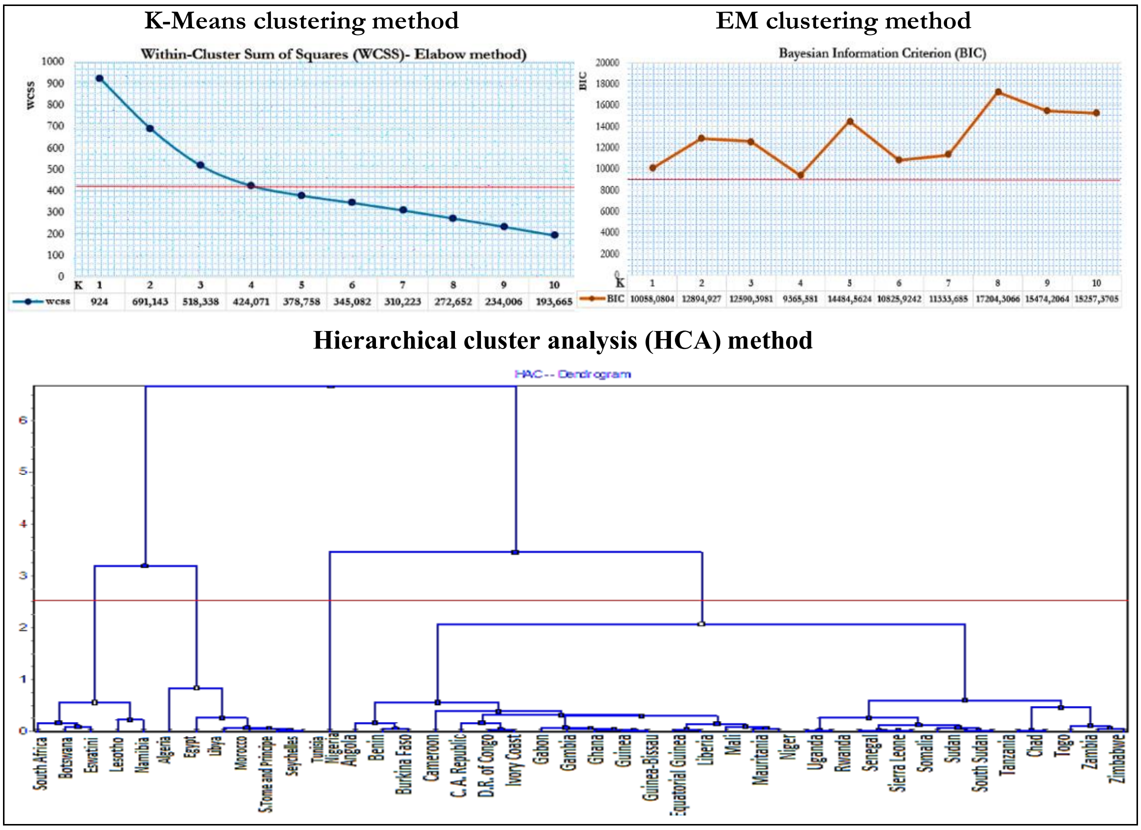



Figure 6 Justification for the optimal number of clusters in each clustering method.

**K-Means Clustering:** the optimal number of clusters (K) for K-Means was determined using the Elbow Method. The K-Means algorithm was run ten times, with K values ranging from 1 to 10. For each run, the Within-Cluster Sum of Squares (WCSS) was calculated, which measures how cohesive the points within each cluster are.

The plot of WCSS values against K shows a sharp decrease in WCSS for K=1, K=2, K=3, and K=4. However, the decrease becomes much less pronounced when moving to K=5. This significant change in the rate of decrease forms the "elbow" on the plot, indicating that K=4 is the optimal number of homogeneous clusters.

**EM Clustering:** to confirm the optimal number of clusters for EM (Expectation-Maximization) clustering, the Bayesian Information Criterion (BIC) was used. Similar to K-Means, the clustering method was run ten times (with K from 1 to 10). For each run, the BIC value was calculated using the formula:

$$BIC = -2 * log(L) + K * log(n)$$

Where: ***L*:** The maximum likelihood of the model;

***K*:** The number of parameters in the model;

***n*:** The number of data points.

The lowest BIC value indicates the best model, which balances model complexity with its fit to the data. The plot of BIC values shows that the lowest value is 9365.5, which corresponds to K=4. This confirms the existence of four homogeneous clusters using the EM method.

**Herarchical Cluster Analysis (HCA):** the dendrogram visually represents the hierarchical clustering process, with each country starting as an independent cluster. Moving up from the bottom, we can observe 5 groups at a height of 1. As we move to heights 2 and 3, some of these groups merge. The horizontal red line at approximately 2.5 was chosen as the criterion to cut the dendrogram. This criterion clearly divides the countries into 4 distinct clusters.

## Study Results

### K-Means clustering method

This method was applied as it requires prior knowledge of the number of clusters through which the number of cases is grouped, and considering the number of countries proposed for the study (42 African countries), the number of clusters was determined to be four K = 4, the results were as follows:

Table 4 Distribution of countries by clusters using the K-Means averaging method.

| - **Clusters** | | **04** |
|---|---|---|
| - **Trials** | | **05** |
| - **R-Square** | | **0,9402** |
| **Clusters** | **Size** | **Pays** |
| Cluster n°1 | 04 | South Africa, Democratic Republic of Congo, Egypt and Tanzania. |
| Cluster n°2 | 01 | Nigeria. |
| Cluster n°3 | 26 | Benin, Botswana, Central African Republic, Eswatini and Gabon. |
| Cluster n°4 | 11 | Gambia, Guinea, Guinea-Bissau, Equatorial Guinea, Lesotho, Liberia, Libya, Mauritania, Namibia, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Sudan, Chad, Togo, Tunisia, Zambia and Zimbabwe. |

Source: elaborated by the authors based on the outputs of the TANAGRA program

By applying this method as noted, it was confirmed that four homogeneous clusters were obtained in terms of the variables used, and this was done over five (05) clustering stages. It is also noted that the correlation coefficient R-Square indicates the existence of a strong correlation within the clusters, which explains the extent of similarity between the countries within each cluster. The results of the analysis of variance and intra-cluster distance according to each variable are distributed as shown in Table 5.

Table 5 Analysis of variance and intra-cluster distance by each variable.

| **Attribute_Y** | **Statistical test** | | **Cluster centroids** | | | |
|---|---|---|---|---|---|---|
| | **Fisher's F Value** | **Proba** | **Cluster n1** | **Cluster n2** | **Cluster n3** | **Cluster n4** |
| Population | 199,21 | **0,000** | 85 708 002,00 | 223 804 640,00 | 8 306 068,65 | 34 746 582,18 |
| Births | 126,36 | **0,000** | 2 553 009,50 | 8 109 136,00 | 215 887,65 | 1 092 571,55 |
| Total_deaths | 221,36 | **0,000** | 451 077,25 | 2 747 861,00 | 60 210,35 | 180 587,18 |
| Maternal_deaths | 198,28 | **0,000** | 7 330,00 | 82 000,00 | 1 313,19 | 3 148,18 |
| Under_five_mortality_rate | 2,22 | **0,010** | 25,72 | 87,36 | 44,47 | 41,83 |
| Infant_mortality_rate | 1,83 | **0,016** | 27,18 | 70,82 | 41,31 | 37,91 |
| Life_expectancy_women | 1,60 | **0,021** | 65,92 | 50,85 | 62,69 | 64,26 |
| Infant_mortality | 1,48 | **0,023** | 30,32 | 70,82 | 41,35 | 37,91 |
| AIDS_prevalence_total | 1,20 | **0,032** | 5,90 | 1,30 | 5,58 | 1,38 |
| Death_rate | 2,50 | 0,074 | 7,84 | 12,28 | 8,16 | 7,03 |
| Fertility_rate | 2,49 | 0,075 | 3,95 | 5,06 | 2,98 | 4,36 |
| Anemia_prevalence_women | 1,34 | 0,277 | 35,03 | 55,10 | 39,48 | 42,26 |
| Life_expectancy | 1,33 | 0,280 | 65,48 | 53,87 | 62,95 | 65,07 |
| HIV_prevalence_all_ages | 1,19 | 0,325 | 1,36 | 0,34 | 1,63 | 0,33 |
| Total_fertility_rate | 1,10 | 0,360 | 3,45 | 5,06 | 3,88 | 4,36 |
| HIV_prevalence_15_49 | 1,06 | 0,378 | 2,20 | 0,44 | 2,32 | 0,55 |
| Crude_birth_rate | 0,92 | 0,441 | 26,45 | 36,22 | 29,41 | 32,35 |
| Comprehensive_health_coverage_index | 0,81 | 0,495 | 56,50 | 38,00 | 48,00 | 47,64 |
| Cause_of_death | 0,69 | 0,564 | 38,96 | 65,21 | 46,50 | 44,83 |
| Health_expenditure_per_capita | 0,65 | 0,585 | 175,28 | 69,76 | 140,55 | 78,46 |
| Birth_rate | 0,56 | 0,642 | 29,35 | 36,22 | 29,41 | 32,36 |
| Physicians | 0,02 | 0,996 | 0,48 | 0,38 | 0,41 | 0,40 |

Source: elaborated by the authors based on the outputs of the TANAGRA program

Table 5 shows the analysis of variance for each of the studied variables using the cluster groups. The statistical test ANOVA indicates the significance level for each of the variables: population, births, deaths, comprehensive health coverage index, under-five mortality rate, infant mortality rate (rate), life expectancy (women), infant deaths, and maternal deaths. We find that these variables are significant at a significance level of (0.05), which indicates that they are important variables in classifying countries into clusters. However, the remaining variables were not statistically significant, which indicates that these variables did not contribute significantly to the ranking of countries in the clusters using this method.

### Hierarchical cluster analysis (HCA) method

We used data covering 42 countries. Therefore, the hierarchical clustering algorithm (HCA) is suitable for estimating the number of clusters. We used hierarchical agglomerative clustering, which starts by considering each case as a separate cluster. Then, similar clusters are merged together, and we can choose the number of clusters that achieves the highest height in the hierarchical clustering tree. After applying this method, four groups were extracted, which are distributed according to Table (6).

Table 6 Distribution of cases by clusters using the HCA method.

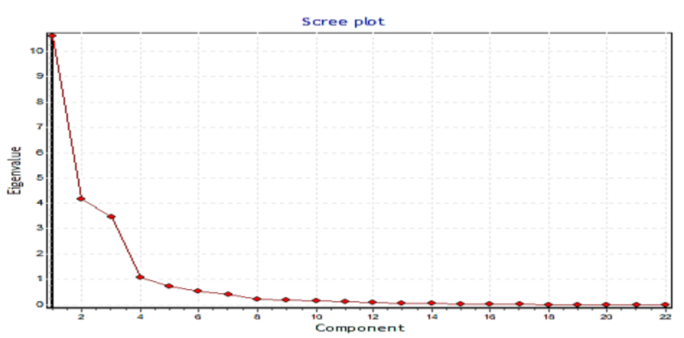
| Clusters | Size | Pays |
|---|---|---|
| Cluster n°1 | 01 | Nigeria. |
| Cluster n°2 | 29 | Angola, Benin, Burkina Faso, Cameroon, Central African Republic, Democratic Republic of Congo, Ivory Coast, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Equatorial Guinea, Liberia, Mali, Mauritania, Niger, Uganda, Rwanda, Senegal, Sierra Leone, Somalia, Sudan, South Sudan, Tanzania, Chad, Togo, Zambia and Zimbabwe. |
| Cluster n°3 | 05 | South Africa, Botswana, Eswatini, Lesotho and Namibia. |
| Cluster n°4 | 07 | Algeria, Egypt, Libya, Morocco, Sao Tome and Principe and Seychelles Tunisia. |

Source: elaborated by the authors based on the outputs of the TANAGRA program

To understand the nature of the relationship between the variables and the number of cases studied, we used Principal Component Analysis (PCA). This method allowed us to identify the most influential indicators on the level of reproductive health in African countries by extracting the latent roots of the variables and their level of saturation in the clusters.
Extracting the eigenvalues of the principal factors matrix, which are the latent roots of the correlation matrix, and the proportion of the total variance of the studied variables explained by each factor.

Table 7 Latent roots and total explained variance of the variables.

| Axis | Eigen value | Difference | Proportion (%) |
|---|---|---|---|
| 1 | 10,602 | 6,457 | 48,19% |
| 2 | 4,144 | 0,684 | 18,84% |
| 3 | 3,460 | 2,369 | 15,73% |




Source: elaborated by the authors based on the outputs of the TANAGRA program.

The latent roots represent the total contribution of each variable to a factor in the matrix, each individually. The first factors are those with the largest latent roots, followed by the next.
To accurately extract the primary factors that explain these variables, we employed the following criteria:

- Kaiser Index: According to this index, factors whose latent roots are greater than one (2) are retained. Accordingly, from Table 05, we can see that the first three latent roots are greater than (2). Therefore, these factors are retained and the remaining factors are discarded;
- Cattle Index: This index relies on a graphical method to identify primary factors. From the figure showing the latent roots of all variables, we can see that the latent roots of the first three factors lie in the steep slope region and above the stable part. This indicates that these factors should be retained and the remaining factors should be discarded.

Based on these indicators, the following primary factors are extracted:

- Factor 1 has the largest latent root, equal to (10,60), and explains 48,19% of the total variance in reproductive health variables;
- Factor 2 has a latent root of (4,14) and explains 18,84% of the total variance in reproductive health variables;
- Factor 3 has a latent root of (3.46) and explains 15,73% of the total variance in these variables.

Therefore, we can explain most of the variables in the first three factors with a total variance of (82,76%) of the total variance of the studied variables, which is a high and sufficient percentage to summarize all the variables in specific factors.

Based on the graph illustrating the explained variance index, which is determined by the largest eigenvalue and the highest explained total variance ratio for all the investigated variables, the maximum value that the sum of the three eigenvalues can attain represents the proportion that can be elucidated by the total variance of reproductive health variables. This value is estimated at (11,95), with a total variance ratio of factors estimated at (82,76%).

The saturation values of the previously represented axes are considered correlation coefficients between them. In other words, the higher the saturation value of the primary axis, the clearer and more interpretable the variable or cluster is. From the table below, we can observe the saturation of the variables on the axes as follows.

Table 8 Level of saturation of variables by latent roots.

| **Axis 2** | |
|---|---|
| **Variables** | **Saturation value** |
| Comprehensive_health_coverage_index | 0,921 |
| Cause_of_death | -0,903 |
| Crude_birth_rate | -0,898 |
| Birth_rate | -0,898 |
| Under_five_mortality_rate | -0,896 |
| Life_expectancy_women | 0,889 |
| Infant_mortality | -0,883 |
| Infant_mortality_rate | -0,875 |
| Total_fertility_rate | -0,866 |
| Life_expectancy | 0,826 |
| Physicians | 0,789 |
| Health_expenditure_per_capita | 0,714 |

| **Axis 2** | |
|---|---|
| **Variables** | **Saturation value** |
| HIV_prevalence_15_49 | -0,855 |
| AIDS_prevalence_total | -0,843 |
| HIV_prevalence_all_ages | -0,775 |

| **Axis 3** | |
|---|---|
| **Variables** | **Saturation value** |
| Total_deaths | 0,864 |
| Population | 0,828 |
| Maternal_deaths | 0,814 |
| Births | 0,797 |

Source: elaborated by the authors based on the outputs of the TANAGRA program

The first principal component (PC1) accounts for the majority of the variance at 48.19%. It shows a strong positive correlation with the Universal Health Coverage Index (0.921). Conversely, it is negatively correlated with several indicators: causes of death (-0.903), birth rate (-0.898), births (-0.898), under-5 mortality rate (-0.896), infant mortality (-0.883), and total fertility rate (-0.866). This component largely reflects the level of good health and reproductive development. High positive values indicate improved healthcare and a reduction in both mortality and uncontrolled births, while high negative values point to significant challenges in these areas.

The second principal component (PC2) explains a significant portion of the variance at 18.84%. It is strongly and negatively correlated with three indicators: HIV prevalence (ages 15-49), total HIV prevalence, and HIV prevalence for all ages (-0.855, -0.843, -0.775, respectively). Consequently, this component primarily represents the burden of HIV and its impact on reproductive health.

The third principal component (PC3) accounts for a substantial 15.73% of the variance. It is positively correlated with life expectancy for women (0.889), mortality (0.864), population (0.828), crude birth rate (0.797), physicians per 1,000 people (0.789), and per capita health expenditure (0.714). It also has a negative correlation with maternal deaths (-0.814). Therefore, this component reflects general demographic indicators, health resources, and life expectancy, with an inverse relationship to maternal deaths. This suggests that an increase in resources and improved demographic conditions lead to a reduction in maternal mortality.

After identifying the clusters and principal components, we will now discuss the relationship between these clusters and the extracted components based on their loadings as shown in Figure 7.

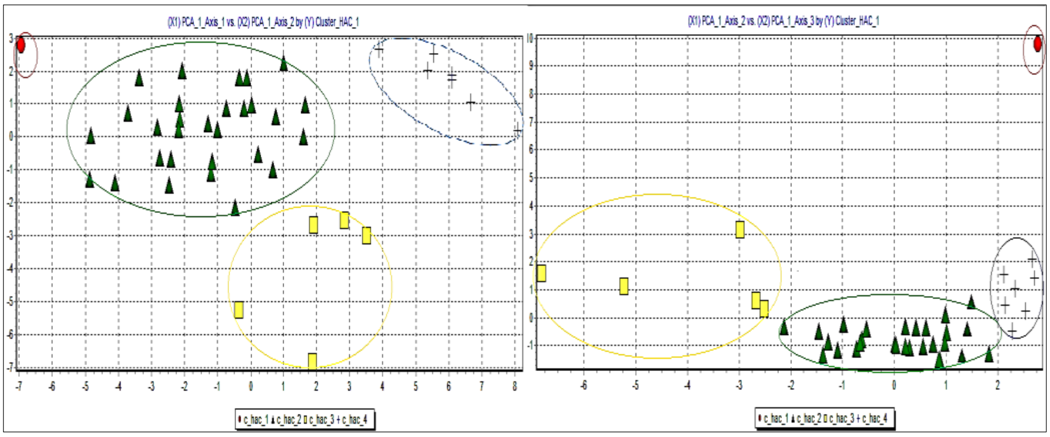


Figure 7 Representation of cluster saturation on axes.

Source: elaborated by the authors based on the outputs of the TANAGRA program.

Integrating factor analysis with hierarchical clustering provided a deeper interpretation of the clusters based on their axis saturation values.

The first cluster, which contains Nigeria, shows a strong negative saturation with the first axis. This indicates significant healthcare challenges, high infant and birth mortality rates, and low levels of health and reproductive development. A strong positive saturation with the third axis, however, points to a large population and relatively lower maternal mortality rates compared to other countries. This combination of factors suggests poor reproductive and health characteristics.

The second cluster (29 countries) also shows a strong negative saturation with the first axis. This indicates that this group faces major challenges in reproductive health indicators, such as high mortality rates (infant, child) and birth rates, along with a weak universal health coverage index. This cluster includes the countries with the poorest performance in reproductive health.

The third cluster (5 countries) shows a strong negative saturation with the second axis, which points to high levels of HIV prevalence. While these countries may perform better in other aspects of reproductive health (not strongly correlated with the first and third axes), the high prevalence of HIV is a major health challenge for them.

The fourth cluster (7 countries) has a significant positive saturation with the first axis. This indicates good performance in reproductive health indicators, such as a high universal health coverage index and low mortality (infant, child) and birth rates. This cluster includes the countries with the best performance in reproductive health, often found in North Africa and some small islands, which tend to have higher levels of development.

### Expectation maximization (EM clustering method)

Unlike the K-Means clustering method, which uses the mean values of the clusters as their centers, the Expectation-Maximization (EM) clustering method is flexible in determining the centers of the clusters by following the Gaussian distribution. Accordingly, the description of the shape of the clusters depends on two inputs: the mean and the standard deviation. As a result, the imposed constraints are reduced.

We applied the same steps followed in the K-Means method with setting the number of clusters to four K=4, and the results were as follows:

Table 9 Country distribution by clusters using the EM-Clustering algorithm.

| - Clusters | | 04 |
|---|---|---|
| - Trials | | 05 |
| - R-Square | | 0,893 |
| **Clusters** | **Size** | **Pays** |
| Cluster n°1 | 06 | Algeria, Egypt, Libya, Morocco, Seychelles and Tunisia. |
| Cluster n°2 | 14 | Angola, Benin, Burkina Faso, Central African Republic, Democratic Republic of Congo, Ivory Coast, Guinea, Mali, Niger, Nigeria, Sierra Leone, Somalia, South Sudan and Chad. |
| Cluster n°3 | 06 | South Africa, Botswana, Eswatini, Guinea-Bissau, Lesotho and Namibia. |
| Cluster n°4 | 16 | Cameroon, Gabon, Gambia, Ghana, Equatorial Guinea, Liberia, Mauritania, Uganda, Rwanda, Sao Tome and Principe, Senegal, Sudan, Tanzania, Togo, Zambia and Zimbabwe. |

Source: elaborated by the authors based on the outputs of the TANAGRA program

By applying this method as shown in Table 9, it was confirmed that four homogeneous clusters were obtained in terms of the variables used, and this was done through 5 clustering stages. It is also noted that the correlation coefficient (R) indicates the existence of a strong correlation within the groups, which explains the extent of similarity between the countries within each group. As for the results of the analysis of variance and distance within the clusters according to each variable, they are distributed as shown in the table below.

Table 10 Analysis of variance and intra-cluster distance by each variable.

| Attribute_Y | Statistical test | | Cluster centroids | | | |
|---|---|---|---|---|---|---|
| | Fisher's F Value | Proba | Cluster n1 | Cluster n2 | Cluster n3 | Cluster n4 |
| Birth_rate | 44,139 | 0,000 | 17,5 | 37,5 | 24,4 | 31,2 |
| Death_rate | 18,700 | 0,000 | 5,8 | 9,4 | 10,0 | 6,7 |
| Life_expectancy | 39,570 | 0,000 | 74,8 | 58,9 | 60,0 | 64,7 |
| Infant_mortality | 27,538 | 0,000 | 12,7 | 57,6 | 41,5 | 34,6 |
| Crude_birth_rate | 38,262 | 0,000 | 17,5 | 36,6 | 24,4 | 31,2 |
| Infant_mortality_rate | 24,161 | 0,000 | 12,7 | 56,7 | 41,3 | 34,6 |
| Under_five_mortality_rate | 32,473 | 0,000 | 10,4 | 67,0 | 42,7 | 34,4 |
| Total_fertility_rate | 29,480 | 0,000 | 2,4 | 5,0 | 3,0 | 4,0 |
| Life_expectancy_women | 39,751 | 0,000 | 75,5 | 57,8 | 60,9 | 64,0 |
| Cause_of_death | 84,935 | 0,000 | 10,4 | 58,5 | 44,9 | 48,2 |
| Health_expenditure_per_capita | 17,977 | 0,000 | 301,9 | 44,2 | 299,1 | 66,5 |
| HIV_prevalence_15_49 | 36,325 | 0,000 | 0,4 | 0,5 | 7,8 | 1,2 |
| Comprehensive_health_coverage_index | 48,525 | 0,000 | 69,5 | 36,9 | 57,3 | 47,4 |
| AIDS_prevalence_total | 26,613 | 0,000 | 2,2 | 1,2 | 17,4 | 3,2 |
| Physicians | 42,885 | 0,000 | 1,5 | 0,1 | 0,5 | 0,2 |
| HIV_prevalence_all_ages | 12,172 | 0,000 | 1,5 | 0,4 | 4,5 | 0,7 |
| Anemia_prevalence_women | 12,071 | 0,000 | 29,8 | 48,3 | 31,9 | 40,0 |
| Fertility_rate | 3,472 | 0,025 | 2,4 | 4,4 | 2,6 | 3,4 |
| Maternal_deaths | 1,747 | 0,034 | 302,5 | 10 457,1 | 379,8 | 1 849,9 |
| Births | 1,395 | 0,259 | 714 755,3 | 1 447 205,8 | 225666,2 | 628070,6 |
| Population | 0,996 | 0,405 | 35 936 232,5 | 39 665 735,6 | 11 824 971,8 | 20 182 457,1 |
| Total_deaths | 0,610 | 0,613 | 172 397,7 | 319 801,6 | 125 021,7 | 115 147,4 |

Source: elaborated by the authors based on the outputs of the TANAGRA program.

The Table 10 shows the ANOVA analysis of variance for each of the studied variables using the cluster groups. The statistical test ANOVA indicates a significance level for the majority of the variables (18 variables) of (0.05), indicating that these are important variables that contribute significantly to the classification of countries into clusters. However, the remaining variables, namely crude birth rate, population size, and total deaths, were not statistically significant, indicating that these variables did not contribute significantly to the ordering of countries in the clusters using this method.

To identify the nature of the relationship between the variables and the number of cases studied, we employed factor analysis in the same manner as hierarchical cluster analysis (HCA). We restricted our analysis to examining the correlation between the primary axes and the extracted clusters using the graphical representation of cluster saturations by the axes, as shown in Figure 8.

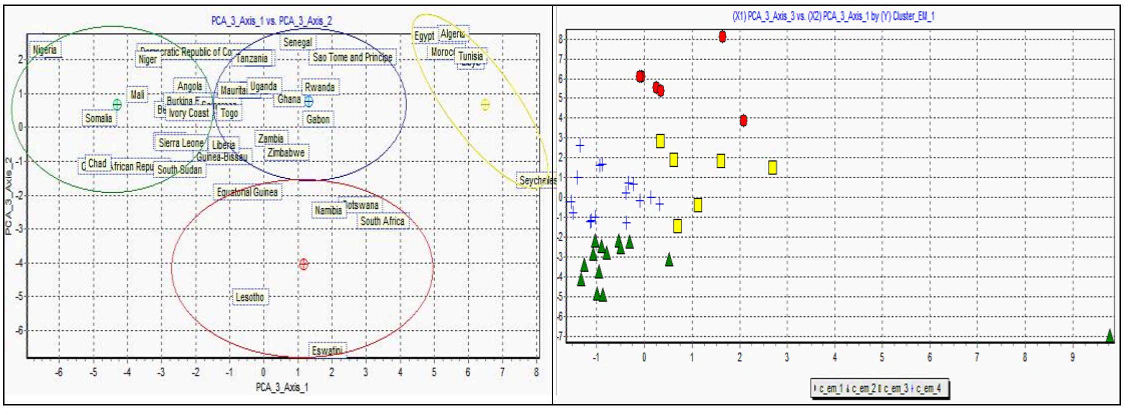



Figure 8 Representation of cluster saturation on the axes using the EM-Clustering approach.
Source: elaborated by the authors based on the outputs of the TANAGRA program.

Our analysis of the cluster saturation representation with the axes extracted using factor analysis of the studied variables reveals the first cluster, which includes six (06) countries, is strongly positively correlated with the first axis. The second cluster is strongly positively correlated with the second axis and strongly negatively correlated with the third axis and includes (14) countries. The third cluster, which includes six (06) countries, is negatively correlated with the second axis and positively correlated with the third axis. The fourth cluster includes (16) countries and is negatively correlated with the first axis.

By studying the distances between the clusters, which reflect the significant differences in the economic, social, and health status among African country groups, where each cluster represents a relatively homogeneous group of countries. Table 11 shows the differences between the average indicators of the groups and the average of the cluster centers.

Table 11 Discrepancies between the cluster center and the average of the groups

| Indicators | Cluster n°1 | | Cluster n°2 | | Cluster n°3 | | Cluster n°4 | |
|---|---|---|---|---|---|---|---|---|
| | **Average between the Groups*** | **Cluster Center**** | **Average between the Groups*** | **Cluster Center**** | **Average between the Groups*** | **Cluster Center**** | **Average between the Groups*** | **Cluster Center**** |
| Number of Countries | 5 | | 1 | | 29 | | 7 | |
| Population | 13847031,6 | 52526822,3 | 223804632,0 | 153144750,6 | 48780959,6 | 17266829,4 | 30835607,1 | 29221653,1 |
| Births | 260736,0 | 1327506,7 | 8109136,0 | 5481312,7 | 1622875,1 | 556363,9 | 613568,4 | 924422,1 |
| Birth_rate | 23,4 | 20,6 | 36,2 | 32,3 | 29,3 | 31,1 | 19,0 | 32,1 |
| Fertility_rate | 2,3 | 2,9 | 5,1 | 4,2 | 3,5 | 3,4 | 2,6 | 3,9 |
| Total_deaths | 147188,8 | 265290,9 | 2747861,0 | 1873581,2 | 470756,4 | 99462,4 | 147962,4 | 208392,1 |
| Death_rate | 10,3 | 6,5 | 12,3 | 11,5 | 8,2 | 7,4 | 5,8 | 8,1 |
| Life_expectancy | 59,6 | 71,7 | 53,9 | 55,9 | 63,1 | 63,4 | 73,9 | 61,4 |
| Infant_mortality | 39,3 | 3663,2 | 70,8 | 191475,1 | 43,1 | 11355,7 | 12,7 | 2547,0 |
| Crude_birth_rate | 23,4 | 20,5 | 36,2 | 32,3 | 29,3 | 31,5 | 19,0 | 31,8 |
| Infant_mortality_rate | 39,0 | 17,5 | 70,8 | 61,0 | 42,9 | 40,0 | 12,7 | 45,5 |
| Under_five_mortality_rate | 40,3 | 15,5 | 87,4 | 72,5 | 47,5 | 41,3 | 10,2 | 51,2 |
| Total_fertility_rate | 2,8 | 2,8 | 5,1 | 4,4 | 3,9 | 4,1 | 2,6 | 4,2 |
| Life_expectancy_women | 60,8 | 72,3 | 50,8 | 54,2 | 63,0 | 62,8 | 74,6 | 60,8 |
| Cause_of_death | 42,0 | 19,9 | 65,2 | 58,4 | 42,2 | 49,0 | 13,4 | 50,6 |
| Health_expenditure_per_capita | 311,6 | 259,7 | 69,8 | 146,2 | 101,3 | 87,9 | 274,1 | 112,4 |
| HIV_prevalence_15_49 | 8,2 | 1,0 | 0,4 | 2,9 | 2,3 | 1,4 | 0,4 | 2,3 |
| HIV_prevalence_all_ages | 4,6 | 1,4 | 0,3 | 1,7 | 1,3 | 0,9 | 1,3 | 1,4 |
| anemia_prevalence_women | 29,4 | 31,5 | 55,1 | 47,4 | 43,3 | 40,1 | 31,8 | 39,5 |
| Comprehensive_health_coverage | 59,6 | 61,0 | 38,0 | 44,4 | 50,0 | 44,4 | 68,0 | 40,9 |
| Maternal_deaths | 433,8 | 2645,0 | 82000,0 | 54793,3 | 12948,3 | 2177,6 | 260,6 | 4856,2 |
| AIDS_prevalence_total | 19,5 | 3,4 | 1,3 | 6,7 | 4,6 | 3,4 | 1,9 | 5,6 |
| Physicians | 0,5 | 1,2 | 0,4 | 0,4 | 0,5 | 0,3 | 1,4 | 0,3 |

(**) Cluster center by taking the average of the index for the three clustering methods (K-means, EM and HCA).
(*) Mean Index Value Across Countries in Each Group.

Source: elaborated by the authors based on the outputs of the TANAGRA program

## Interpretation and discussion of results

The analysis of reproductive health indicators in Africa reveals significant variations among countries, highlighting a complexity that requires precise understanding and tailored policy responses. The interpretation of the clusters obtained using the EM-clustering algorithm and other statistical techniques provides an overview of the main challenges and successes encountered across the continent. From this research paper, the following conclusions were drawn:

### Interpreting the results of applying clustering methods

The results from the K-Means and Hierarchical Cluster Analysis methods show similarity in the ordering of countries within the clusters, with their distribution as follows: (01, 26, 04, 11) for the first method, and (01, 29, 05, 07) for the second method. In contrast, the EM-Clustering method yielded a different distribution of countries across clusters, with a distribution of (06, 14, 06, 16).

### Discussion of the results of the axes

Principal Component Analysis (PCA) was employed to understand the nature of the relationship between variables and the number of cases (countries) by extracting the latent roots of the variables. The results indicated the existence of three (03) primary factors explaining the majority of the variables, with a cumulative variance of 82.76% of the total variance of the studied variables. These axes are:

- First Axis: This axis includes variables related to child mortality and healthcare. Data from this axis show that increased access to maternal health services is associated with a significant reduction in infant mortality (Okonofua, 2018). This axis also includes all Sub-Saharan African countries with low birth rates in health facilities and a high proportion of births without qualified medical personnel, as confirmed by Udo and Doctor (2016). Other studies on infant mortality trends (Benn & Kurt, 2014; Moens et al., 2015) emphasized that these countries represent a "dilemma" regarding women's and children's health. They also did not experience significant reductions in infant mortality between 1990-2011 and highlighted prominent high-risk factors, primarily environmental, socio-economic factors, and maternal mortality (survival), lack of water and sanitation, and female education. In this regard, Wallis (2018) confirmed that investing in primary healthcare and maternal health is essential for improving health outcomes, as it has become a priority to address disparities and follow the approach of countries that have made significant positive progress towards achieving the Sustainable Development Goals;
- Second Axis: This axis carries variables explaining the prevalence of Human Immunodeficiency Virus (HIV), highlighting the need for comprehensive and integrated strategies to combat this epidemic. These results include countries with high HIV prevalence, which is consistent with Adams et al. (2023), which identified African countries suffering from a high prevalence of this disease during 1990-2018, particularly Lesotho, Eswatini, South Africa, Zambia, Zimbabwe, Namibia, Malawi, Mozambique, and Botswana. Other survey studies on new evidence for adolescent sexual and reproductive health and rights in the West and Central African region (Okonofua, 2024). Proved that research groups found that the West and Central African region as a whole experiences the worst global reproductive health indicators concerning rates of unwanted pregnancies, unsafe abortions, HIV, and sexually transmitted infections. Therefore, prevention and treatment programs must be strengthened and integrated into reproductive health policies (UNAIDS, 2020);
- Third Axis: This axis largely explains variables including population size, mortality, crude births, and maternal mortality. It comprises countries characterized by high fertility rates due to demographic and socio-economic factors associated with contraceptive use. This is evident in the results of a Study by Yaya et al. (2018), which observed a significant decline in contraceptive use in many African and Sub-Saharan African countries, such as Chad, Gambia, Guinea, Angola, Mali, Burundi, Comoros, Burkina Faso, and Nigeria. Regarding the hotspots for maternal mortality rates, these are linked to countries suffering from low utilization of adequate maternal health services with a high need for health interventions (hotspots), including the entire West African region and Nigeria, as confirmed by the results of Studies by Habte et al. (2024) and Burgert-Brucker et al. (2015). Therefore, these countries must adopt family planning and reproductive health policies to control these dynamics and improve living conditions (Council, 2019).

### Discussion of cluster results

Based on the analysis of how the clusters correlate with the extracted axes, the results of the groups were interpreted as follows:

- Nigeria, as the sole country in the first group, tops the list of African nations suffering the most from reproductive health indicators. It records the highest rates of infant and under-five child mortality, in addition to the highest maternal mortality rates linked to infectious diseases and complications from pregnancy and childbirth. This reveals a severe deficiency in basic healthcare. This situation is exacerbated by a complex set of demographic factors, including high birth and fertility rates that place immense pressure on the fragile health system. Institutional challenges further compound the issue, such as restrictive abortion laws, a shortage of medical staff, administrative bureaucracy, and weak capacity of health facilities. This aligns with Study by Oredegbemi et al. (2025) which demonstrated that Nigeria continues to face critical gaps regarding these indicators. Moreover, socioeconomic factors play a pivotal role. Extreme poverty, political instability, and internal conflicts contribute significantly to the deepening of internal geographical disparities and limit women's and girls' access to essential health services, which in turn widens the crisis of inequality in healthcare access (Barreix et al., 2015);
- Southern African countries-South Africa, Botswana, Eswatini, Lesotho, and Namibia are grouped together due to their high HIV prevalence rates. East and Southern Africa is the region most affected globally, with approximately 20.8 million people living with the virus in 2023. There were 260,000 AIDS-related deaths, and women and girls accounted for 63% of all new infections. In South Africa, 12.8% of the population lives with the virus, while the rate in Eswatini reaches 19.58%. This complex reality is due to several interlocking factors: demographic factors like a large youth population (the most sexually active group), high fertility rates, and migration; institutional factors such as weak health systems, a lack of awareness and education programs, and poor coordination between government agencies; economic factors, most notably increasing poverty, unemployment, and inequality, which can lead some to engage in transactional sex, in addition to the high cost of treatment; and finally, social and cultural factors, such as the prevalence of multiple sexual partners and casual relationships, as well as the stigma, discrimination, and misconceptions surrounding the disease (UNAIDS, 2019);
- Countries like Algeria, Egypt, Libya, Morocco, Sao Tome and Principe, Seychelles, and Tunisia are clustered together due to their similarities in reproductive health indicators, such as low maternal and child mortality rates and better availability of reproductive healthcare services. This relative success in health is attributed to a combination of factors. Socio-culturally, increasing health awareness, along with women's education and empowerment, has improved health practices. Economically, relative stability and financial investment in healthcare have allowed resources to be directed toward improving key indicators. Institutionally, commitment to national programs, training of specialized medical staff, and regional cooperation have enhanced the effectiveness of health systems. Demographically, high levels of urbanization, smaller populations in island nations, and progress in demographic transition have facilitated greater universal health coverage and improved access to family planning methods, reflecting an integrated effort to achieve positive health outcomes. This was proven in Study by Yaya et al. (2021), where these countries were classified among those with low maternal mortality rates due to their moderately efficient health systems that allow them to control reproductive health indicators.
- Twenty-nine (29) African countries, predominantly located in sub-Saharan Africa, form a cohesive group characterized by negative reproductive health indicators and significant challenges. This region is among the most affected globally, struggling with extreme poverty, weak economic development, and escalating conflicts and humanitarian crises. This reality is reflected in weak health systems and infrastructure, poor maternal and prenatal care, and a widespread prevalence of anemia among women of reproductive age. These countries also record very high birth and fertility rates (ranging from 4 to 6 children per woman). The situation is further exacerbated by gender inequality, low levels of education, the spread of diseases and epidemics, and cultural factors such as traditional beliefs and practices that hinder the adoption of family planning.

Regarding the evaluation of some important principles included in achieving reproductive health requirements, such as the principles of health equity, the right to health, and the principle of women's empowerment, the majority of previous studies in this field indicate that all African countries in general suffer from the widespread phenomenon of violence against women (including intimate partner violence and sexual violence) across all socioeconomic classes (Edouard, 2024). Furthermore, most African countries focus their reproductive health programs on traditional services and have a pro-rich distribution in the use of modern contraceptives without incorporating the principles of the International Conference on Population and Development regarding equality, social justice, and human rights. This deficiency makes them vulnerable to the effects of fluctuating external policies, This was proven by the studies of Okonofua (2025), Adeyinka et al. (2019) and Fentie et al. (2023). Furthermore, (14) Sub-Saharan African countries were classified among the poorest regions in the world according to the Multidimensional Poverty Index. (Batana, 2008).

### Policy Implications and Research Perspectives

The results of this study show that it is essential to adopt a differentiated approach based on national and regional specificities. Strengthening health systems, especially in the most vulnerable countries, is a priority. Interventions must be based on strong evidence and adapted to local contexts to be effective. Future research prospects should include longitudinal analyses to understand long-term trends. Accordingly, some solutions can be formulated, according to each group, as follows:

- Nigeria, as a unique group, faces distinct challenges characterized by the highest maternal and child mortality rates, weak health systems, and ongoing conflicts. Addressing this situation necessitates significant investment in integrated primary healthcare, including mobile health teams and comprehensive maternal care programs. It is crucial to intensify campaigns to increase the availability of modern family planning methods and raise awareness about their importance, while also tackling socio-economic factors such as girls' education and malnutrition. The World Health Organization (2020) called on Nigeria to take immediate action to increase resource mobilization and redesign health policies to address the existing social determinants of health.
- The second group's primary challenge is the high prevalence of HIV. Solutions for this group revolve around integrated preventative and treatment programs for HIV, with a focus on expanding testing and treatment, as well as mother-to-child transmission prevention programs. Promoting comprehensive sexual health education and combating stigma and discrimination are also critical. The Joint United Nations Programme on HIV/AIDS (UNAIDS, 2019) also called for integrating HIV care into reproductive health services in "one-stop shop" clinics, deeming this a crucial step to improving the situation in these countries.
- North African countries and small island states show relatively better reproductive health indicators. The focus for this group should be on sustaining these gains, enhancing the quality of healthcare, and making continuous investments in universal health coverage. They must also prepare for future challenges, such as focusing on adolescent sexual and reproductive health, and addressing non-communicable diseases. Furthermore, these countries can play a leading role in sharing their experiences and best practices with other African nations facing greater challenges. The World Bank (2020) emphasized that despite the general progress in the health sector, the problem of access to healthcare, especially in rural areas of these countries, persists. It called for ensuring the equitable distribution of health resources to achieve universal health coverage that reaches all segments of society.
- The group of 29 countries in sub-Saharan Africa faces multidimensional and deep-rooted challenges concerning reproductive health indicators. This requires building resilient and comprehensive health systems that strengthen primary healthcare, intensify targeted maternal and child health programs, and focus on universal vaccination programs and the management of common diseases. They must also invest in human capital by increasing the number and training of health workers and empowering women. Additionally, addressing underlying socio-economic factors through the formulation of multidimensional poverty reduction programs, crisis preparedness, and increased awareness of family planning requirements is vital. UNICEF (2021) affirmed the urgent need for comprehensive and multifaceted interventions to combat malnutrition, enhance health infrastructure, and invest in maternal education.

Embodying this approach requires intensifying in-depth research and studies based on reliable databases. The scope of indicators should be expanded to include the quality of care, beneficiary satisfaction, gender-based violence, and achieving health equity. Modern data analysis methodologies, such as Bayesian models, should be applied to identify causal relationships and deduce the influential factors (social, economic, cultural, and security...). This will help decision-makers evaluate policies and redirect interventions effectively.

## Conclusions

Reproductive health is a vital component of quality of life and community development, encompassing crucial issues such as family planning, maternal and child care, and the control of sexually transmitted diseases. Since the establishment of the International Conference on Population and Development in Cairo in 1994, these issues have gained international prominence and become a priority across all countries, especially in African nations, which continue to face immense challenges in this field. Achieving comprehensive reproductive health in this continent is not merely a health objective; it is a fundamental pillar for sustainable human and social development. By adopting a multi-sectoral, evidence-based approach that considers local specificities, Africa can overcome its current challenges and achieve a future where mothers, children, and adolescents enjoy health and well-being.

This study clearly revealed the extent of stark disparities and pressing challenges facing reproductive health in Africa. Through the use of a descriptive analytical method and an advanced statistical approach based on multi-dimensional analysis of vital indicators using clustering methods to classify African countries based on reproductive health indicators, exploring relationships between variables, and identifying underlying factors explaining disparities, it became clear that the continent stands at a crossroads requiring innovative policy responses that account for the limited successes achieved in some regions and the alarming deterioration in others.

This study, based on the literature, demonstrated the importance of cluster analysis and factor analysis, which enabled us to group homogeneous countries, revealing common patterns in strengths and the magnitude of health challenges for regions suffering from larger gaps in health systems. This procedure can help policymakers design more effective health interventions and programs that contribute to achieving sustainable development goals, tailored to the needs of each group. The study also concluded the necessity of focusing on critical factors shaping the trajectory of reproductive health, particularly geographical location, socioeconomic status, women's empowerment, and the principle of health equity.

In light of these findings, researchers in this field can interpret the reality of reproductive health in African countries by summarizing the set of influencing variables and representing them as primary, well-interpretable factors. However, the generalizability of these results depends on several other factors, whether they have a direct or indirect impact on reproductive health. These include, but are not limited to, environmental factors, genetic factors, nutritional patterns, type of breastfeeding, and other operational factors such as the type and volume of data used, the place and time of its collection, and its accuracy and reliability.

## References

Adams, S. O et al. (2023). Cluster analysis of HIV/AIDS incidence in Sub-Saharan Africa (1990–2018). *Int J Epidemiol Health Sci, 4*(4), 1-11. doi:10.51757/IJEHS.4.2023.701311.

Adeyinka, A. E et al. (2019). Systematic review and meta-analysis of the association between dimensions of inequality and a selection of indicators of Reproductive, Maternal, Newborn and Child Health (RMNCH). *Journal of Global Health, 9*(1), 1-13. doi:10.7189/jogh.09.010429.

Anil K, J., & Richard C, D. (2007). *Algorithms for Clustering Data.* Englewood Cliffs, New Jersey, Michigan State University, United States of America: Prentice-Hall.

Avinash, K. (2024). *Expectation-Maximization Algorithm for Clustering Multidimensional Numerical Data.* Purdue University, United States of America: RVL Tutorial Presentation.

Ayadda, M. (2012). *Quantitative Models and Methods in Planning and Their Computer Applications.* Amman, Jordan: Dar Al-Hammed.

Bailey, L., & Elkan, C. (1995). Unsupervised learning of multiple motifs in biopolymers using expectation maximization. *Machine Learning, 21*(1), 51-80. https://link.springer.com/article/10.1007/BF00993379

Barreix, M et al. (2015). Experience from a multi-country initiative to improve the monitoring of selected reproductive health indicators in Africa. *BJOG: An International Journal of Obstetrics & Gynaecology, 124*(S1), 205-212. doi:10.1002/ijgo.12105.

Bassey, E. (2023). *Fragile States Index_annual report.* Fund For Peace. https://fragilestatesindex.org/wp-content/uploads/2023/06/FSI-2023-Report_final.pdf,

Batana, Y. (2008, August). Multidimensional measrement of poverty in Sub-Saharan Africa. *Oxford Poverty & Human Development Initiative(Working Paper No.13)*, 1-35. OPHI, Oxford.

Benn KD, S., & Kurt, S. (2014). Global infant mortality trends and attributable determinants – an ecological study using data from 192 countries for the period 1990–2011. *Population Health Metrics, 12*(1), 29. doi:10.1186/s12963-014-0029-6.

Bradley, S. (2012). *Revising Unmet Need for Family Planning. DHS Analytical Studies N°25.* ICF International, Calverton, Maryland- United States of America. https://www.dhsprogram.com/pubs/pdf/AS25/AS25%5B12June2012%5D.pdf,

Burgert-Brucker, C et al. (2015). *Geographic Variation in Key Indicators of Maternal and Child Health Across 27 Countries in Sub-Saharan Africa.* DHS Spatial Analysis Reports , ICF International, Rockville, Maryland, USA.

Council, P. (2019). *Fertility Trends and Maternal Health: Implications for African Demographics.* New York, United States of America: Population Council.

E. Udo, I., & V. Doctor, H. (2016). Trends in health facility births in sub-Saharan Africa: An analysis of lessons learned under the Millennium Development Goal framework. *African Journal of Reproductive Health, 20*(3).

Edouard, L. (2024). Violence against women: Diverse facets. *African Journal of Reproductive Health, 28*(10), 12-16. doi:10.29063/ajrh2024/v28i10.1.

Estivill-Castro, V. (2002). Why so many clustering algorithms: a position paper. *ACM SIGKDD Explorations Newsletter, 4*(1), 65-75.

Everitt, B., Landau, S., Leese, M., & Stahi, D. (2011). *Cluster Analysis* (5 ed.). Londen: Wily Series.

Fentie, E. A et al. (2023). Modern contraceptive utilization among reproductive-age women in sub-Saharan African countries: a decomposition analysis. *BMC Health Services Research, 23*(185), 4-10. doi:10.1186/s12913-023-09172-6.

Gieryn, T. (2000). A Space for place in sociology. *Annual Review of Sociology*, 463-496.

Habte, A et al. (2024). Predictors of maternal health services uptake in West African region: a multilevel multinomial regression analysis of demographic health survey reports. *Reproductive Health, 21*(45), 1-17. doi:10.1186/s12978-024-01782-5.

Hardie, R., Barnard, K., & Armstrong, E. (1997). Joint MAP registration and high-resolution image estimation using a sequence of undersampled im ages. *IEEE Transactions on Image Processing, 6*(12), 1621-1633. doi:10.1109/83.650116. https://pubmed.ncbi.nlm.nih.gov/18285233/

Lobao, L., & Saenz, R. (2002). Spatial Inequality and Diversity as an Emerging Research Area. *Rural Sociology, 64*(04), 497-511.

Moens, K et al. (2015). Symptom clusters in people living with HIV attending five palliative care facilities in two Sub-Saharan African countries: A hierarchical cluster analysis. *PLOS ONE, 10*(5). doi:10.1371/journal.pone.0126554.

Okonofua , F. (2024). New evidence on adolescent sexual and reproductive health and rights from the West and Central African region. *African Journal of Reproductive Health, 28*(8s), 11-14. doi:10.29063/ajrh2024/v28i8s.

Okonofua, F. (2018). Correlation between maternal healthcare access and infant mortality rates in Africa. *The Lancet Global Health, 6*(7), 750-760.

Okonofua, F. (2025). Where are the "rights" in SRHR? *African Journal of Reproductive Health, 29*(1), 11-14. doi:10.29063/ajrh2025/v29i1.1

Oredegbemi, A et al. (2025). Two decades of women's sexual and reproductive health and rights in Nigeria: Successes, challenges, and opportunities. *African Journal of Reproductive Health, 29*(1), 25-37. doi:10.29063/ajrh2025/v29i1.3.

Organization World Health. (2023). *world Health statistics 2023: monitoring Health for the SDGS, sustainable development goals, 2023,.* World Health Organization.

Ramdeen, K., & Yim, O. (2015). Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. *The Quantitative Methods for Psychology, 11*(1), 08-21. doi:10.20982/tqmp.11.1.p008. https://www.tqmp.org/RegularArticles/vol11-1/p008/p008.pdf

Rencher, A. (2003). *Methods of Multivariate Analysis* (2 ed.). New York, United States of America.

Romesburg, H. (2004). *Cluster Analysis for Researchers.* North Carolina, United States of America: Lulu Press.

Seamon, D., & Sowers, J. (2008). Place and Placelessness, . *Human Geography*, 43-51.

Slonim, N., Aharoni, E., & Crammer, K. (August 2013). Hartigan's K-Means Versus Lloyd's K-Means-Is It Time for a Change?. *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, (pp. 1677-1684).

https://www.ijcai.org/Proceedings/13/Papers/249.pdf
UNAIDS. (2019). *HIV and AIDS in Southern Africa: Regional Report.* Geneva: Joint United Nations Programme on HIV/AIDS.
UNAIDS. (2020). *Integrating HIV Prevention into Reproductive Health Services: Best Practices.* Geneva: Joint United Nations Programme on HIV/AIDS.
UNICEF. (2018). *The latest Levels and Trends in Child Mortality: Report 2018 from UNICEF and partners in the UN Inter-Agency Group for Child Mortality Estimation.*
UNICEF. (2021). *Maternal and Child Health in Sub-Saharan Africa: A Comprehensive Review. United Nations Children's Fund.* New York, United States of America.
United Nations . (2022). *World Population Prospects.* Department of Economic and Social Affairs, Population Division (2022a).
United Nations. (1994). *Programme of Action of the International Conference on Population and Development.* Cairo: Department for Economic and Social Information and Policy Analysis. https://digitallibrary.un.org/record/209240?v=pdf,
United Nations. (2015). *Resolution adopted by the General Assembly. Transforming our world: the 2030 Agenda for Sustainable Development.* Department of Economic and Social Affairs, Population Division. https://digitallibrary.un.org/record/3923923,
United Nations. (2022). *World Contraceptive.* Department of Economic and Social Affairs, Population Division (2022b).
United Nations. (2023). *Sustainable Development Goals (SDGs) region.* Department of Economic and Social Affairs, Population Division.
Velasco, C., Solsona, M., & Burgunder, V. (2011). Structural Violence and Maternal Mortality. *African Population Studies Union and Princeton University. In the 6th African population conference.* Ouagadougou.
Wallis, A. (2018). Addressing gaps in maternal, neonatal, and child health for achieving SDG 2030 in West Africa. *African Journal of Reproductive Health, 22*(4), 9-15.
Wang, C. (2012). History of the Chinese Family Planning program: 1970–2010 Contraception. *international reproductive health journal, 85*(6), 563-569.
Warren, C., & John, A. (2007). *The Global Family Planning Revolution: The Emergence of Thailand's National Family Planning Program.* World Bank. The International Bank for Reconstruction and Development.
WHO, U. U. (2021). *Trends in maternal mortality 2000 to 2020.*
Wiley, J., & Hartigan, A. (1975). *Clustering Algorithms.* New York, United States of America: INC.
World Bank. (2020). *Rural Healthcare Access in North Africa: Challenges and Opportunities.* World Bank Group.
World Health Organization. (2020). *Global Health Observatory.: Reproductive Health Indicators in Africa.* Geneva: World Health Organization (WHO).
World Health Organization. (2023). *World Health Statistics 2023: monitoring Health for the SDGS, sustainable development goals.* World Health Organization (WHO).
Yaya, S et al. (2018). Women empowerment as an enabling factor of contraceptive use in sub-Saharan Africa: a multilevel analysis of cross-sectional surveys of 32 countries. *Reproductive Health, 15*(214), 1-12. doi:10.1186/s12978-018-0658-5.
Yaya, S et al. (2021). Disparities in pregnancy-related deaths: spatial and Bayesian network analyses of maternal mortality ratio in 54 African countries. *BMJ Global Health, 6*(2), 1-10. doi:10.1136/bmjgh-2020-004233.

**Djahid Saidoun**
University of Blida 2 –Lounici Ali
d.saidoun.etu@univ-blida2.dz

**Mohamed Amine Belaidi**
National Higher School of Statistics and Applied Economics. Kolea, Algeria
Belaidi.amine1@gmil.com

**Ahmed Derdiche**
University of Blida 2 –Lounici Ali
demderdiche@yahoo.fr