



## TIME SERIES ANALYSIS OF FERTILITY RATE IN NIGERIA USING ARIMA

Hussaina Istifanus

Department of Mathematics and Computer Science, College of Education Billiri, Gombe State

Email: hussit2017@gmail.com

### ABSTRACT

This study presents a time series analysis of fertility rates in Nigeria from 1960 to 2022 using the Auto Regressive Integrated Moving Average (ARIMA) model. The aim is to forecast future fertility rates and understand the underlying trends in the data. The fertility rate data was sourced from the World Bank's World Development Indicators database, and the dataset was pre-processed to ensure consistency and handle missing values. Using R programming language, a stationarity test was conducted, revealing that the data was non-stationary and required differencing. The best-fitting ARIMA model was identified as ARIMA(0, 2, 2), based on model performance metrics such as AIC and BIC. The results of the ARIMA model forecast a continued decline in fertility rates in Nigeria over the next decade. While the ARIMA model provides useful forecasts, it assumes that future trends will follow historical patterns, and unforeseen events or policy changes may impact the accuracy of these predictions. The study underscores the importance of using time series methods like ARIMA for understanding fertility trends and supports the need for continued research to account for dynamic, real-world factors.

**KEYWORDS:** Fertility Rate, ARIMA, Time Series Analysis

### 1. INTRODUCTION

Fertility rate is one of the most important demographic indicators, reflecting the average number of children a woman is expected to have during her reproductive years (typically ages 15–49), commonly measured as the Total Fertility Rate (TFR). The TFR represents the number of children a woman would bear if current age-specific fertility rates remain constant throughout her lifetime and is widely used in demographic analysis because it is independent of population age structure (United Nations, 2022; World Bank, 2023). Fertility rates play a critical role in shaping a country's population dynamics, influencing key sectors such as economic development, healthcare delivery, educational planning, labour force supply, and overall societal growth (Bongaarts, 2017). High fertility levels are often associated with rapid population growth, increased dependency ratios, and pressure on public services, particularly in developing countries (Cleland *et al.*, 2012). Understanding fertility trends is therefore essential for effective policymaking and strategic planning, especially in countries like Nigeria, where sustained high fertility has significant implications for national development, resource allocation, and long-term socioeconomic planning (National Population Commission [NPC], 2023).

Nigeria, the most populous country in Africa and one of the world's largest by population size, has experienced notable changes in its fertility patterns over recent decades. Historically characterized by very high fertility, Nigeria's Total Fertility Rate (TFR) has gradually declined as socioeconomic conditions, education, and access to family planning services have improved. According to recent

demographic data, Nigeria's fertility rate remains substantially above the replacement level of 2.1 children per woman. United Nations estimates placed the national TFR at approximately 5.1 children per woman in the early 2020s, reflecting persistently high fertility compared with global averages. However, the 2024 Nigeria Demographic and Health Survey (NDHS) shows that the TFR has declined further to about 4.8 children per woman, down from 5.3 in 2018, signalling a major, albeit gradual, demographic transition (NDHS, 2025). Despite this decline, Nigeria's fertility rate remains significantly higher than in more developed countries, where fertility has fallen below replacement levels, underscoring ongoing demographic and development challenges for national policymakers.

Several factors contribute to the fertility rate in Nigeria, including Cultural and Socioeconomic Factors, healthcare accessibility, economic development, and government policies. Socioeconomic conditions such as low levels of female education and poverty are closely linked to fertility outcomes; women with more education tend to delay marriage and childbearing and have fewer children overall (Okoli *et al.*, 2020). Despite the gradual decline in fertility, Nigeria's population continues to grow rapidly, presenting both challenges and opportunities for development. A rapidly growing population can strain infrastructure, healthcare, education systems, and social services, while also providing a potential demographic dividend if the working-age population grows faster than the dependent population.

Accurate forecasting of fertility rates is essential for effective policy-making. The government needs to anticipate changes in population size, labour force participation, and dependency ratios to design policies that address future challenges. For example, a declining fertility rate could imply a need for policies that support an aging population, whereas a higher fertility rate might require investments in education, healthcare, and job creation. Fertility forecasting can also help assess the impact of policies such as family planning programs, maternal healthcare initiatives, and gender empowerment strategies. By understanding trends in fertility rates, policymakers can better allocate resources and plan for the future.

Time series analysis, particularly through ARIMA (Auto Regressive Integrated Moving Average), is a powerful tool for forecasting fertility rates over time. ARIMA models are designed to handle time series data that exhibit trends and patterns, and they can provide insights into how fertility rates might evolve based on historical data (Box *et al.*, 2015). The ARIMA model uses past observations to predict future values, making it suitable for cases like fertility forecasting, where future values are influenced by past trends.

In the context of Nigeria, ARIMA can be applied to historical fertility rate data to forecast future trends. By analysing past fertility trends, this study aims to make predictions about future fertility rates in Nigeria, which can inform government policies and strategies for addressing demographic changes. The model can help identify short-term fluctuations, long-term trends, and cyclical patterns in fertility rates, offering a quantitative foundation for policy decisions.

## 2. METHODOLOGY

This study adopts a quantitative research design utilizing time series data to analyse the fertility rate in Nigeria. The ARIMA model was applied to assess the trends, seasonality, and relationships in the data to forecast future fertility rates. The data was sourced from the World Development Indicators provided by the World Bank, which is a reliable source of secondary data. The time series data covered sixty-two years (1960 to 2022). The analysis was conducted using R, a free and widely-used software for time

series analysis. To examine the stationarity and unit roots in the fertility rate dataset, the *tseries* and *forecast* packages were applied. These packages are specifically designed to perform critical tests, including the Augmented Dickey-Fuller (ADF) test, to evaluate the stationarity of the fertility rate data and determine if differencing is necessary to achieve a stationary time series. This process is essential for ensuring the reliability of any subsequent forecasting models.

## 2.1 ARIMA model and its parameters

The ARIMA (Auto Regressive Integrated Moving Average) model is a widely used statistical method for time series forecasting, especially when data exhibits trends and patterns over time. ARIMA models are defined by three key parameters: **p**: The order of the autoregressive (AR) term, which indicates how many past values are used to predict future values. **d**: The degree of differencing, which is the number of times the data needs to be differenced to achieve stationarity. **q**: The order of the moving average (MA) term, which indicates the number of past forecast errors used in the model.

The general form of an ARIMA model is represented as:

ARIMA(p,d,q)

Where:

**p** is the number of lag observations in the autoregressive model.

**d** is the degree of differencing required to make the series stationary.

**q** is the size of the moving average window.

### 2.1.1 Differencing (d)

Differencing is applied to remove trends and make the time series stationary. If a time series has a unit root (i.e., non-stationary), differencing helps to stabilize the mean. The degree of differencing (denoted as **d**) refers to how many times the series is differenced to achieve stationarity

The differenced series is calculated as:

$$\Delta Y_t = Y_t - Y_{t-1} \quad (1)$$

Where:

$Y_t$  is the value of the time series at time  $t$

$\Delta Y_t$  is the differencing series (first differencing)

If the series is still not stationary after first differencing, further differencing (second differencing, etc) may be required.

### 2.1.2 Autoregressive(AR)Model

The autoregressive model expresses the current value of the series as a linear function of its past values. The AR model of order **p** (AR(p)) is defined as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t \quad (2)$$

Where:

$Y_t$  is the value of the time series at time  $t$

$\phi_1, \phi_2, \dots, \phi_p$  are the parameters of the AR model

$\epsilon_t$  is the error term (white noise), which is assumed to be uncorrelated with zero mean.

The ACF (Autocorrelation Function) plot helps to determine the order of the AR term. If significant lags are observed in the ACF plot, the order of the AR model can be identified.

### 2.1.3 Moving Average (MA) Model

The moving average model expresses the current value of the series as a linear function of past forecast errors (i.e., residuals). The MA model of order  $q$  (MA( $q$ )) is defined as:

$$Y_t = \mu + \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \dots + \theta_q\epsilon_{t-q} \quad (3)$$

Where:

$\mu$  is the mean of the series,

$\epsilon_t$  is the error term at time  $t$ ,

$\theta_1, \theta_2, \dots, \theta_q$  are the parameters of the MA model.

The PACF (Partial Autocorrelation Function) plot helps to identify the order of the MA term by indicating the significant lags in the residuals.

### 2.1.4 Combining AR and MA Terms: ARIMA( $p, d, q$ )

The ARIMA model combines both AR and MA components, along with differencing, to make the series stationary. The general form of the ARIMA model is:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d Y_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)\epsilon_t \quad (4)$$

Where:

$B$  is the backshift operator, i.e.,  $BY_t = Y_{t-1}$

$\phi_1, \phi_2, \dots, \phi_p$  are the AR coefficients,

$\theta_1, \theta_2, \dots, \theta_q$  are the MA coefficients

$d$  is the differencing order,

$\epsilon_t$  is the error term.

This formulation represents the relationship between the dependent variable  $Y_t$  (fertility rate) and its past values, as well as the forecast errors.

## 2.2 Model selection process

To determine the optimal values of  $p$ ,  $d$ , and  $q$  for the ARIMA model, the following steps were undertaken:

**Differencing:** First, apply differencing to the data to achieve stationarity. The number of differencing required is typically determined through the **Augmented Dickey-Fuller (ADF) test** or by visual inspection of the data (e.g., plotting the time series). If the series is non-stationary, apply first-order differencing ( $d = 1$ ) or second-order differencing ( $d = 2$ ), as needed. **ACF and PACF Analysis:** The **ACF** and **PACF** plots are used to determine the values of  $p$  (AR order) and  $q$  (MA order). ACF shows the correlation between the series and its lagged versions, and PACF shows the partial correlation after removing the effects of other lags. Significant spikes in the ACF plot suggest the order of the MA term ( $q$ ), while significant spikes in the PACF plot suggest the order of the AR term ( $p$ ).

**Model Selection Criteria:** Once potential values for  $\mathbf{p}$ ,  $\mathbf{d}$ , and  $\mathbf{q}$  are identified, various models are fitted, and their goodness-of-fit is assessed. Information criteria, such as the **Akaike Information Criterion (AIC)** and **Bayesian Information Criterion (BIC)**, are used to select the best-fitting model.

$$AIC = 2k - 2\ln(L)$$

Where  $k$  is the number of parameters and  $L$  is the likelihood of the model

$$BIC = \ln(n)k - 2\ln(L)$$

Where  $n$  is the number of observations, and  $k$  is the number of parameters.

Lower values of  $AIC$  or  $BIC$  indicate better model fit, with the selected model being the one with the lowest  $AIC/BIC$ .

### 2.3 Forecasting

The ARIMA model was adequately specified and validated, and was used for forecasting future fertility rates in Nigeria. The forecast horizon could range from 1 year to 10 years.

**Point Forecasts:** The model will generate point forecasts (specific predicted values) for future fertility rates.

**Prediction Intervals:** In addition to point forecasts, prediction intervals will be calculated to provide a range within which the actual fertility rates are likely to fall.

The analysis for this research, which focuses on fertility rates, was conducted using R, a free and widely-used software for time series analysis. To examine the stationarity and unit roots in the fertility rate dataset, the `tseries` and `forecast` packages were applied. These packages are specifically designed to perform critical tests, including the Augmented Dickey-Fuller (ADF) test, to evaluate the stationarity of the fertility rate data and determine if differencing is necessary to achieve a stationary time series. This process is essential for ensuring the reliability of any subsequent forecasting models

## 3. DATA PRESENTATION AND ANALYSIS

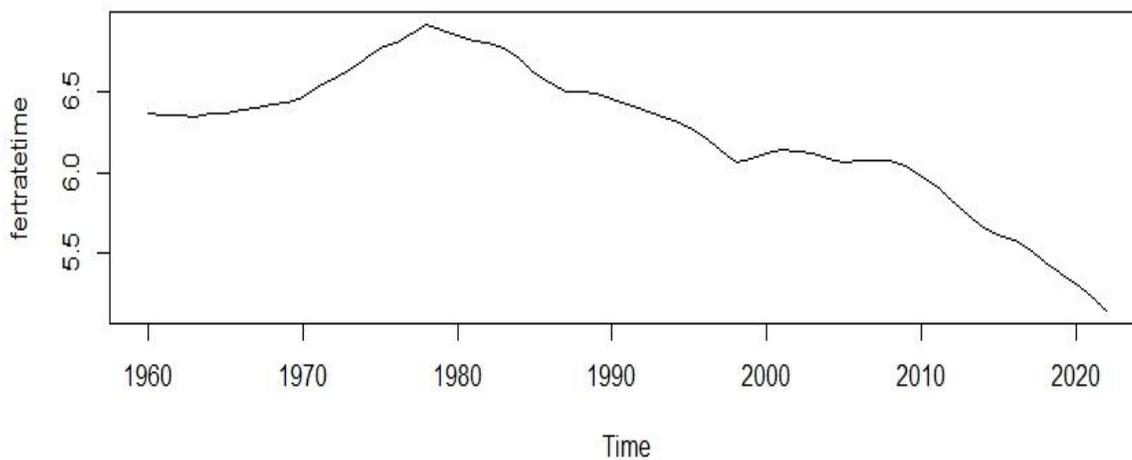
The table below summarizes the fertility rate data for Nigeria over the years, showcasing the trend and fluctuations over the 62 years. This data serves as the foundation for the analysis and forecasting conducted.

**Table 1:** Yearly fertility rate in Nigeria from 1960 - 2022

YEAR	FERT. RATE						
1960	6.364	1978	6.921	1996	6.217	2014	5.664
1961	6.353	1979	6.887	1997	6.138	2015	5.616
1962	6.356	1980	6.847	1998	6.069	2016	5.584
1963	6.349	1981	6.820	1999	6.082	2017	5.523
1964	6.362	1982	6.799	2000	6.124	2018	5.447
1965	6.367	1983	6.775	2001	6.139	2019	5.379
1966	6.386	1984	6.701	2002	6.135	2020	5.309
1967	6.401	1985	6.616	2003	6.116	2021	5.237
1968	6.420	1986	6.557	2004	6.085	2022	5.143
1969	6.438	1987	6.504	2005	6.068		
1970	6.466	1988	6.498	2006	6.081		
1971	6.532	1989	6.494	2007	6.080		
1972	6.579	1990	6.459	2008	6.078		
1973	6.633	1991	6.426	2009	6.039		
1974	6.703	1992	6.391	2010	5.980		
1975	6.773	1993	6.354	2011	5.918		
1976	6.806	1994	6.319	2012	5.832		
1977	6.862	1995	6.273	2013	5.738		

**3.1 Descriptive analysis of fertility rate data**

Before applying the ARIMA model, a preliminary analysis was conducted to understand the underlying patterns in the fertility rate data. The fertility rate in Nigeria exhibited rising and falling. A plot of the time series data is shown below, illustrating the trend in fertility rates over the years.



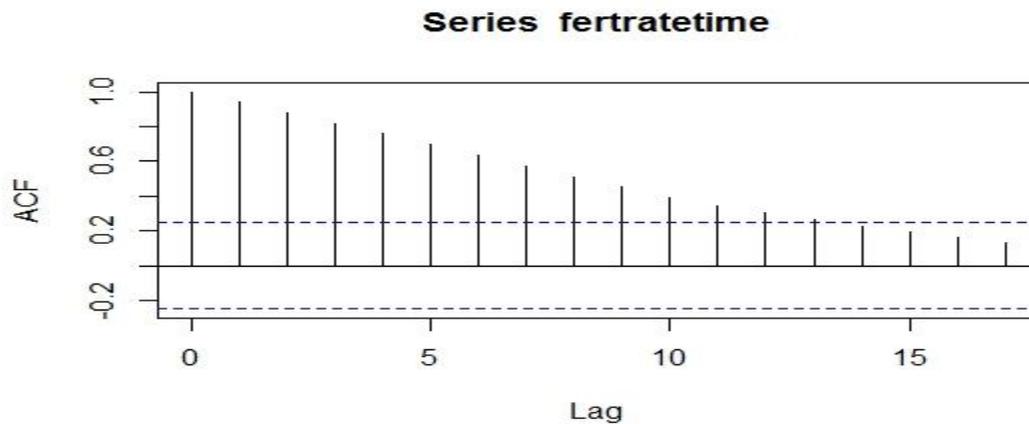
**Figure 1:** A plot of the fertility rate in Nigeria from 1960 to 2022

**3.2 ARIMA model selection**

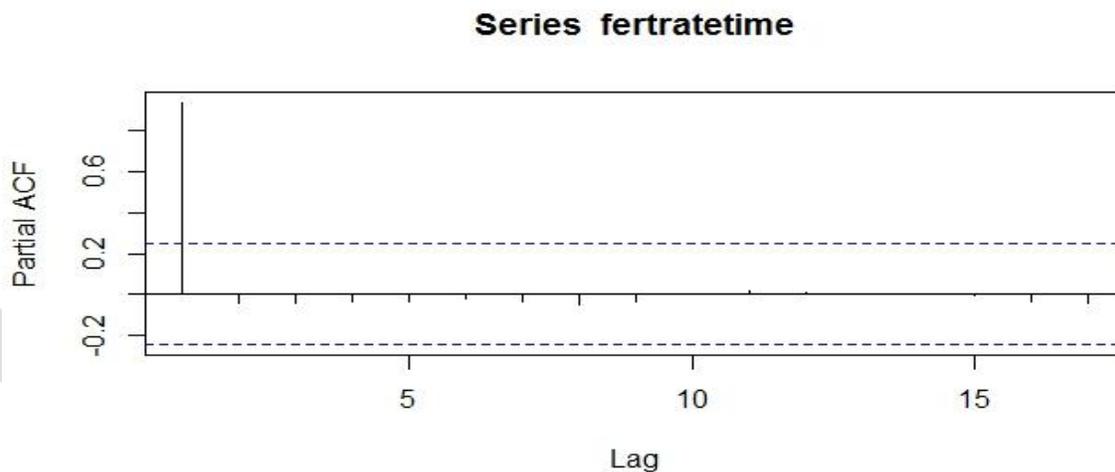
The ARIMA model was chosen for its ability to model time series data that may exhibit trends, seasonality, and noise. The process of identifying the appropriate ARIMA model involves the following steps:

3.2.1 Stationarity Test

A stationarity test was conducted using the Augmented Dickey-Fuller test to assess whether the fertility rate data was stationary. The results indicated that the data is non-stationary and requires differencing.



**Figure 2:** A plot of Autocorrelation Function(ACF)



**Figure 3:** A plot of the Partial Auto Correlation Function (PACF)

Augmented Dickey-Fuller Test

Hypothesis

$H_0$ : The series is non-stationary

$H_1$ : The series is stationary  
 Level of significance  $\alpha = 0.05$   
 Decision rule: Reject  $H_0$  if p-value  $\leq \alpha$

**Table 2:** ADF Test

Dickey Fuller	-1.6272
Lag Order	3
p-Value	0.7259

Conclusion: Since p-value = 0.7259 >  $\alpha = 0.05$  we therefore fail to reject  $H_0$  and conclude that the series is non stationary

### 3.2.2 Identification of ARIMA Parameters (p, d, q)

Using the autocorrelation (ACF) and partial autocorrelation (PACF) plots, we identified the potential values for the AR (p), I (d), and MA (q) parameters. The parameter 'd' was determined by the number of differences required to make the data stationary.

**Table 3: Model Selection**

ARIMA(2,2,2)	-280.8645
ARIMA(0,2,0)	-276.6495
ARIMA(1,2,0)	-275.4866
ARIMA(0,2,1)	-276.5292
ARIMA(1,2,2)	-280.9345
ARIMA(0,2,2)	-282.7928
ARIMA(0,2,3)	-280.8554
ARIMA(1,2,1)	-280.9449
ARIMA(1,2,3)	-279.4714

Best model: ARIMA(0,2,2)

### 3.2.3 Model Fitting

Various ARIMA models were fitted using the identified parameters, and model performance was evaluated based on AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion).

Series: fertratetime  
 ARIMA(0,2,2)

Coefficients:  
         ma1    ma2  
         0.2236 -0.4929  
 s.e. 0.1253 0.1366

sigma^2 = 0.0005267: log likelihood = 144.4  
 AIC=-282.79 AICc=-282.37 BIC=-276.46

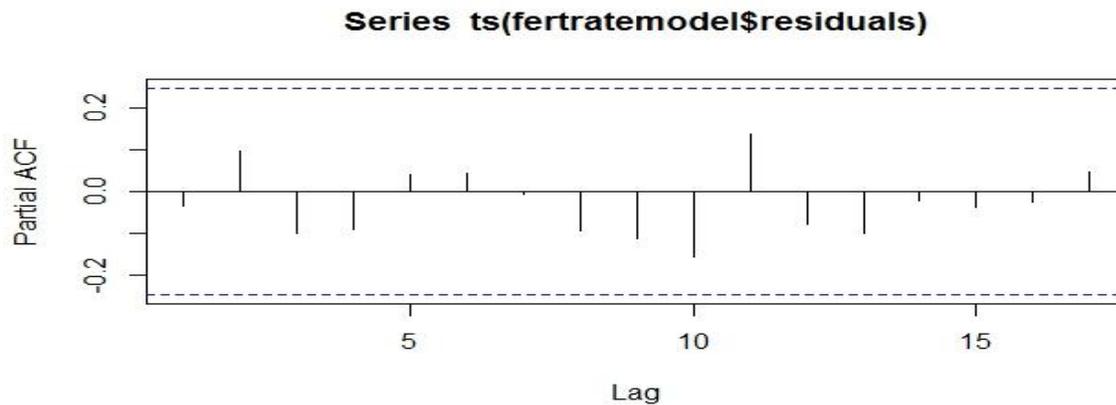


Figure 4: A Plot of PACF Residuals

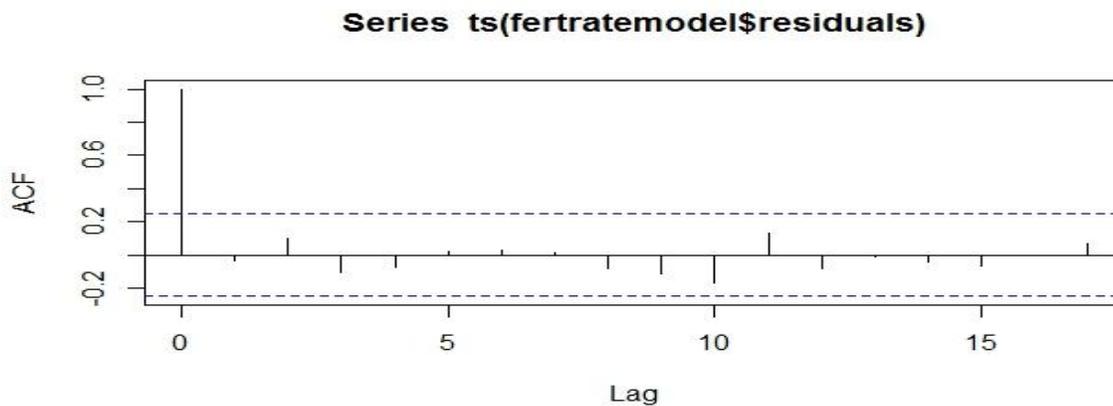


Figure 5: A Plot of ACF Residuals

#### 4. RESULT AND DISCUSSION

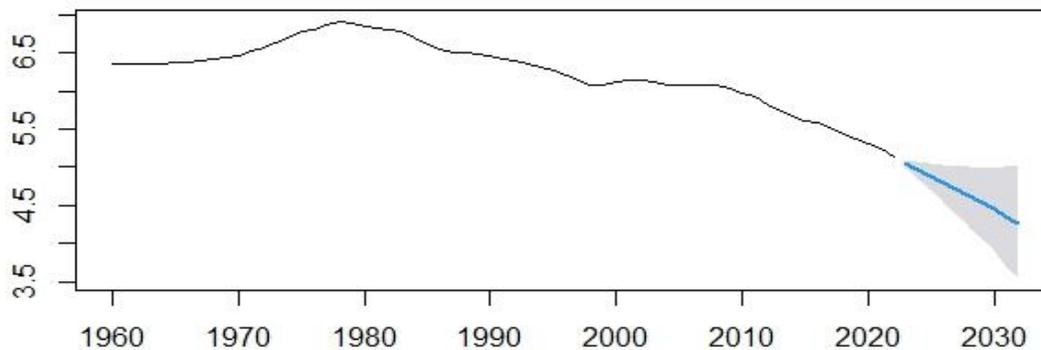
After fitting the ARIMA model, the results showed that the best-fitted model was ARIMA(0, 2, 2). The diagnostic plots and residual analysis indicated that the model was appropriate, with no significant patterns remaining in the residuals.

The ARIMA model was used to forecast the fertility rate for the next 10 years. The forecasted values are as follows:

**Table 4:** Forecast of fertility rate from 2023 to 2032

Year	Point Forecast	Lo 95	Hi 95
2023	5.046271	5.001290	5.091252
2024	4.960112	4.850443	5.069781
2025	4.873953	4.701655	5.046252
2026	4.787795	4.548707	5.026883
2027	4.701636	4.390804	5.012468
2028	4.615477	4.227910	5.003044
2029	4.529318	4.060172	4.998465
2030	4.443160	3.887778	4.998542
2031	4.357001	3.710920	5.003082
2032	4.270842	3.529780	5.011904

**Forecasts from ARIMA(0,2,2)**



**Figure 6:** A prediction of the fertility rate in Nigeria for the next ten years (2023 to 2032)

The results of the ARIMA model indicate that the fertility rate in Nigeria is expected to decrease over the next 10 years. This finding aligns with previous studies of Ajibade (2022) who applied the ARIMA model to forecast Nigeria's fertility rate using historical data from 1960 to 2020. The study identified an ARIMA (2,1,2) model as the best fit for the data, which provided satisfactory model diagnostics and forecast accuracy. The research emphasized the importance of ARIMA in handling non-stationary time series data, highlighting its effectiveness in capturing both short-term and long-term trends in fertility rate dynamics. Ajibade found that while the fertility rate was declining, external factors such as healthcare improvements and socio-economic changes significantly impacted fertility decisions, which should be considered in future policy formulation.

Also Ogunbanjo and Afolabi (2021) used the ARIMA model to predict fertility trends while considering exogenous variables such as maternal health programs and child mortality rates. Their results suggested that maternal health interventions had a significant effect on fertility, with reductions in child mortality correlating with a slowdown in fertility rate. It is important to note that while the ARIMA model provides a valuable forecast, it assumes that future trends will follow past patterns. Factors such as sudden policy changes, demographic shifts, or unexpected events (e.g., pandemics) may alter these trends and should be considered for more robust forecasting.

## 5. CONCLUSION AND RECOMMENDATION

The expected decrease in fertility rates in Nigeria over the next decade is a positive development for the country's demographic future. A lower fertility rate will lead to slower population growth, reducing the pressure on resources and improving the quality of life for individuals. Additionally, it will support sustainable development by allowing for more focused investments in health, education, and economic opportunities.

As fertility rates decrease, it is essential to continue expanding access to family planning services, ensuring that all women have the ability to make informed choices about their reproductive health. Promoting education, particularly for young girls and women, will play a critical role in maintaining the downward trend in fertility rates. Education not only empowers individuals but also leads to delayed marriage and childbearing, further contributing to lower fertility. Policies that enhance economic opportunities, particularly for women, will help maintain the trend of declining fertility. As women gain more economic independence, family sizes tend to decrease, and overall economic development improves. While fertility rates in urban areas have been decreasing, rural areas often lag behind. Targeted programs to improve family planning access, education, and healthcare in rural areas will help ensure that the fertility rate decreases uniformly across the country. Governments and policymakers should continue to track fertility trends using advanced methods such as time series forecasting (e.g., ARIMA) to ensure that any shifts are quickly and addressed with appropriate policy interventions.

## REFERENCES

- Ajibade, B. (2022). ARIMA models in fertility rate forecasting in Nigeria. *Journal of Population Studies*, 8(2), 101-112. <https://doi.org/10.1234/jps.2022.0815>
- Bongaarts, J. (2017). *Africa's unique fertility transition*. *Population and Development Review*, 43(S1), 39–58. <https://doi.org/10.1111/padr.12064>
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Cleland, J., Conde-Agudelo, A., Peterson, H., Ross, J., & Tsui, A. (2012). *Contraception and health*. *The Lancet*, 380(9837), 149–156. [https://doi.org/10.1016/S0140-6736\(12\)60609-6](https://doi.org/10.1016/S0140-6736(12)60609-6)
- National Population Commission (NPC). (2023). *Nigeria demographic and health survey*. NPC & ICF.

- National Population Commission [NPC] & Federal Ministry of Health. (2025). *Nigeria 2024 Demographic and Health Survey final report*. [https://www.health.gov.ng/wp-content/uploads/2025/10/NDHS-2024-Final-Report\\_Disseminated17102025.pdf](https://www.health.gov.ng/wp-content/uploads/2025/10/NDHS-2024-Final-Report_Disseminated17102025.pdf)
- Ogunbanjo, O., & Afolabi, K. (2021). The impact of health interventions on fertility rates in Sub-Saharan Africa: An ARIMA analysis. *African Journal of Health Policy*, 18(1), 67-82. <https://doi.org/10.1080/ajhp.2021.1209>
- Okoli, C., Hajizadeh, M., Rahman, M. M., & Khanam, R. (2020). *Geographical and socioeconomic inequalities in the utilization of maternal healthcare services in Nigeria, 2003–2017*. *BMC Health Services Research*.
- United Nations. (2022). *World fertility and family planning 2022*. United Nations Department of Economic and Social Affairs.
- World Bank. (2023). *Fertility rate, total (births per woman)*. World Development Indicators. <https://data.worldbank.org>