International Journal of Basic Science and Technology
March, Volume 10, Number 1, Pages 1 – 6 http://www.ijbst.fuotuoke.edu.ng/ 1
https://doi.org/ 10.5555/RGGF8084 ISSN 2488-8648

Modelling Birth Rate in Otuoke, Bayelsa State, Nigeria: A Time Series Approach

[1]Eli, I.C. and [2]Barinaadaa, J.N.
[1]Department of Mathematics and Statistics Federal University Otuoke, Bayelsa State, Nigeria
[2]Department of Mathematics and Statistics Ignatius Ajuru University of Education, Port Harcourt, Nigeria.

Article Information

Article # 10001
Received: 9th June 2023
1st Revision: 1st Dec. 2023
2nd Revision:8th Jan. 2024
Acceptance:4th Feb. 2024
Available online:
3rd March 2024

Key Words
Birth rate, ARIMA,
Differencing,
Estimation

Abstract

This study seeks to determine a suitable Time series model for predicting birthrate in Bayelsa State, Nigeria, using Time series Box- Jenkins method. Monthly birth data from April, 2015 to April 2022 collected from Federal Medical Centre Otueke Bayelsa State was used. The Augmented Dick Fuller test was used to test for unit root. The unit root test revealed that the data was not stationary. The series was difference and stationary at the first differencing. Based on of various selection criteria Autoregressive Integrated Moving Average, ARIMA (1, 1, 1) was identified as an appropriate model for the forecast of birthrate. The Ljung-Box *Q* test used to test the adequacy of the model. The residual of the model was checked using Ljung-Box residual test and it shows no sign of autocorrelation in the residual. Thus, Autoregressive Integrated Moving Average ARIMA (1,1,1) model was recommended for forecast of birth-rate in Bayelsa state.

***Corresponding Author**: Eli, I.C..; eliic@fuotuoke.edu.ng

Introduction

Forecasts of monthly births can help predict the fluctuating trends of a country's population distribution, which is also important in estimating the increase in population. The importance of population forecasting in the context of national, state, and local government budgeting and legislation in medical services, security, and infrastructure cannot be emphasized (Bravo and Magalhães, 2010). To appropriately design programs to reduce mortality among kids and fertility, it is also necessary to understand the concept of birth rate, especially among developing nations experiencing population change (Soest & Saha, 2018). Several independent studies have found a link between particular periodic birth rates, reproductive customs, and the risk of a child dying (Asefa *et al.,* 2000).

Research findings indicate the occurrence of seasonal and trend patterns in birth rate data. As a result, utilizing models based on time series to forecast birth and mortality rates has grown in popularity in the past few decades. A time series is any random variable that is observed sequentially throughout time. Analyzing time series data requires analyzing the characteristics of the response variable with respect to time as the independent variable.

The time variable is used as a point of reference for estimating the target variable for predicting or forecasting purposes. Hussein (1993), for example, analyzed Egypt's crude birth rate and crude death rate from 1992 to 2010 employing the autoregressive integrated moving average (ARIMA) model; the study used a time series approach to anticipate Egypt's rates of births and deaths beyond 2010. Furthermore, Adsera (2006), Abel *et al.* (2013), Lee (1992), Ahlo (2012), Lee and Tuljapurkar (1994), Graham *et al.* (2008), Keilman *et al.* (2002), Tayman (2007), Kirk (1996), Rau (2007), Palloni, and Rafalimanana (1999) amongst others have also worked on the application of time series models for birth rate and demographic forecasting.

The autoregressive integrated moving average (ARIMA) is one of the time series models commonly used. The model is also known as the ARIMA (p, d, q) model, where p is the autoregressive component, q is the moving average component and d represents differencing, p,d and q are all integers larger than or equal to 1. The ARIMA model is appropriate when the data is stationary. For a stationary series, the mean, variance and autocorrelation function remain constant throughout time.

The evaluation of time series models for birth and death in Otuoke has received little attention thus, the main focus of the present study is to determine a suitable ARIMA model for the forecast monthly Birth rate in Otuoke, Bayelsa State, Nigeria using the time series method developed by Box- Jenkins.

Materials and Methods

Autoregressive (AR) Model

The autoregressive process of order k denoted by AR (k) is a time series model given by

$$X_t = \sum_{i=1}^{k} \vartheta_i X_{t-i} + \varepsilon_t$$

$$= \vartheta_1 X_{t-1} + \vartheta_2 X_{t-2} + \vartheta_3 X_{t-3} + \cdots + \vartheta_k X_{t-k} + \varepsilon_t \qquad (1)$$

Where the $\vartheta_i s\ i = 1,2,3, \ldots, k$ are the model parameters and ε_t is a sequence of independently and identically distributed (iid) random variables with $E(\varepsilon_t) = 0$ and $var(\varepsilon_t) = \sigma_e^2$

Autoregressive models are based on the idea that the current value of the series, can be explained as a linear combination of past values, together with a random error in the same series.

Moving Average (MA) Model

Given that ε_t is a sequence of independently and identically distributed (iid) random variables with $E(\varepsilon_t) = 0$ and $var(\varepsilon_t) = \sigma_e^2$, then

$$X_t = \sum_{i=1}^{r} \theta_i \theta_{t-i} = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \cdots + \theta_r \varepsilon_{t-r} \qquad (2)$$

is the rth order moving average MA(r) model.

Autoregressive Moving Average (ARMA) Model

Autoregressive Moving Average (ARMA) Model is a mixture of the AR and MA models. Thus, a time series is said to follow an autoregressive moving-average process of order k and r, i.e ARMA (k, r), process if:

$$X_t = \sum_{i=1}^{k} \vartheta_i X_{t-i} + \varepsilon_t + \sum_{i=1}^{r} \theta_i \theta_{t-i}$$

$$X_t = \vartheta_1 X_{t-1} + \vartheta_2 X_{t-2} + \vartheta_3 X_{t-3} + \cdots + \vartheta_k X_{t-k} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \cdots + \theta_r \varepsilon_{t-r} \qquad (3)$$

$$X_t - \vartheta_1 X_{t-1} - \vartheta_2 X_{t-2} - \vartheta_3 X_{t-3} - \cdots - \vartheta_k X_{t-k} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \cdots + \theta_r \varepsilon_{t-r}$$

$$\left(1 - \vartheta_1 B - \vartheta_2 B^2 - \vartheta_3 B^3 - \cdots - \vartheta_k B^k\right) X_t = \left(1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \cdots - \theta_r B^r\right) \varepsilon_t$$

$$\vartheta(B) X_t = \theta(B) \varepsilon_t$$

Where $\vartheta(B) = 1 - \vartheta_1 B - \vartheta_2 B^2 - \vartheta_3 B^3 - \cdots - \vartheta_k B^k$ and $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \cdots - \theta_r B^r$, the backward shift; $B^k X_t = X_{t-k}$

Frequently, after achieving stationarity, time series models contain AR(k) and MA(r) of certain orders which in combination could be used for forecast. The model is usually referred to as the ARMA (k, r) process where k is the order of the autoregressive part and r is the order of the moving average part.

International Journal of Basic Science and Technology
March, Volume 10, Number 1, Pages 1 – 6 http://www.ijbst.fuotuoke.edu.ng/ 2
https://doi.org/ 10.5555/RGGF8084 ISSN 2488-8648

ARIMA Models

The ARIMA models can further be extended to non-stationary series by allowing the differencing of the data series resulting in ARIMA models. The general non-seasonal model is known as ARIMA (k, d, r): where with three parameters; k is the order of autoregressive, d is the degree of differencing, and r is the order of moving average. The ARIMA(k,d,r) model is denoted as follows:

$$\vartheta(B)\nabla^d X_t = \theta(B)\varepsilon_t$$

Where $\nabla(B) = 1 - B$ is the differencing parameter applied d times, and d is a non-negative integer.

$$\vartheta(B) = 1 - \sum_{i=1}^{r} \vartheta_i B^i \text{ and } \vartheta(B) \text{ is the AR operator}$$

$$\vartheta(B) = 1 - \sum_{j=1}^{r} \theta_j B^j \text{ and } \theta(B) \text{ is the MA operator}$$

Data

The data used in this work are of secondary sources. The comprised of monthly birth rate in federal medical center Otueke, Bayelsa State, Nigeria from April 2015 to April 2022.

Test for stationarity:

The ARIMA model is suitable for stationary time series data, thus, a test for stationarity (unit root test) of the data is necessary before further analysis. The unit root test for the data was carried out using the Augmented Dick Fuller (ADF) Test.

Results

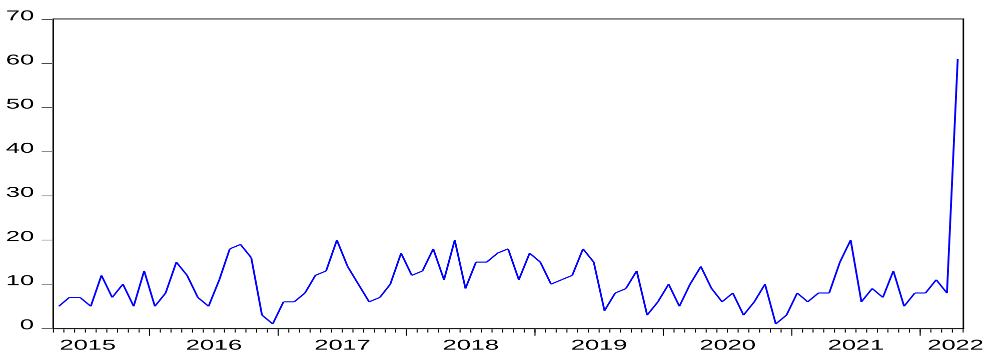


Fig.1 Time Plot of Birth Rate

Table 1: Unit Root Test of Monthly Birth Series

Null Hypothesis: MONTHLYBIRTH has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.139385	0.0614
Test critical values:	1% level	-3.510259	
	5% level	-2.896346	
	10% level	-2.585396	

*MacKinnon (1996) one-sided p-values.

March, Volume 10, Number 1, Pages 1 – 6 http://www.ijbst.fuotuoke.edu.ng/
https://doi.org/ 10.5555/RGGF8084 ISSN 2488-8648

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(MONTHLYBIRTH)
Method: Least Squares
Date: 11/06/23 Time: 20:42
Sample (adjusted): 2015M05 2022M04
Included observations: 84 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MONTHLYBIRTH (-1)	-0.710488	0.171641	-4.139385	0.0001
C	7.822300	1.892134	4.134115	0.0001

R-squared	0.172841	Mean dependent var	0.666667
Adjusted R-squared	0.162754	S.D. dependent var	7.705941
S.E. of regression	7.051026	Akaike info criterion	6.767745
Sum squared resid	4076.791	Schwarz criterion	6.825621
Log-likelihood	-282.2453	Hannan-Quinn criter.	6.791011
F-statistic	17.13451	Durbin-Watson stat	1.384478
Prob(F-statistic)	0.000084		

Table 1 shows the ADF result for the monthly birthrate data, demonstrating the nonstationary nature of the data. The data was differenced and the plot of the first difference is given in figure2 below

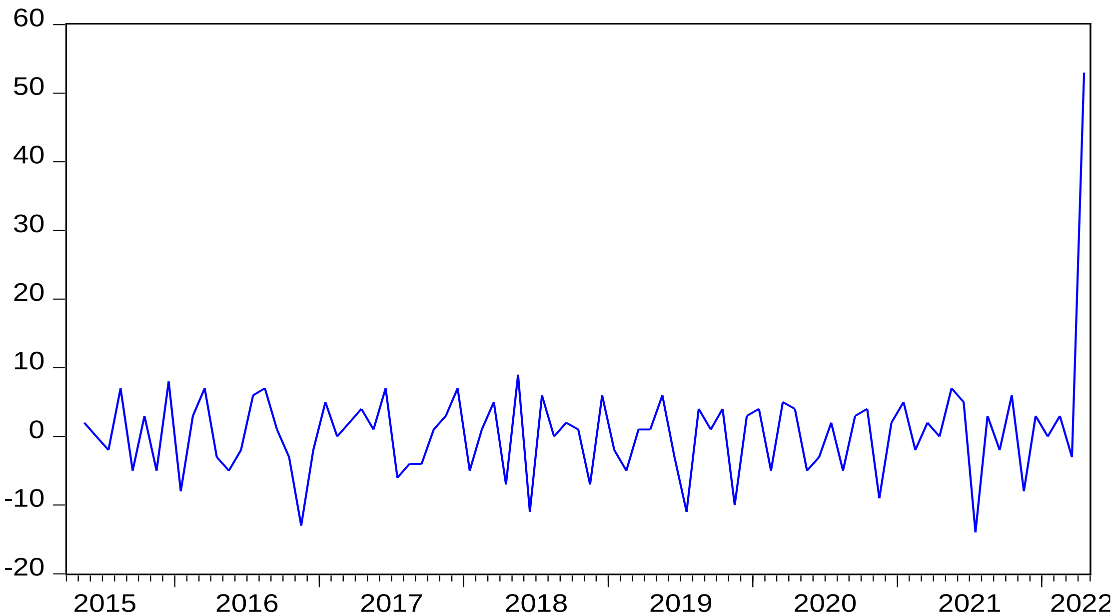


Fig.2:Time Plot of the first difference

International Journal of Basic Science and Technology
March, Volume 10, Number 1, Pages 1 – 6 http://www.ijbst.fuotuoke.edu.ng/ 4
https://doi.org/ 10.5555/RGGF8084 ISSN 2488-8648

Table 2: Unit root test fort the differenced series.

Null Hypothesis: D(DMONTHLYBIRTH) has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.025997	0.0000
Test critical values:	1% level	-3.514426	
	5% level	-2.898145	
	10% level	-2.586351	

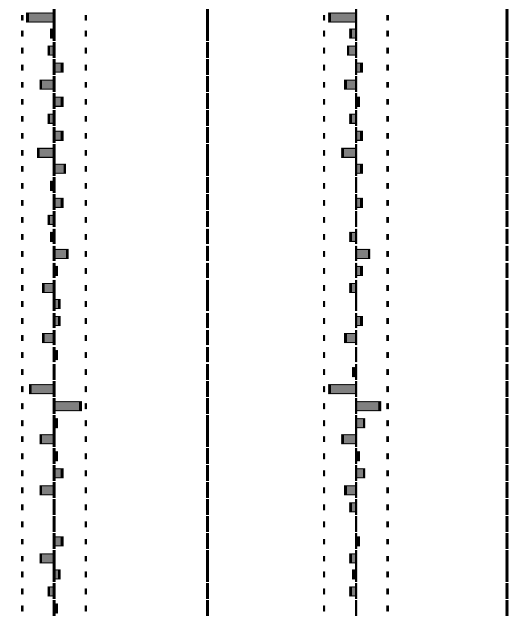
*MacKinnon (1996) one-sided p-values.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DMONTHLYBIRTH,2)
Method: Least Squares
Date: 11/06/23 Time: 20:50
Sample (adjusted): 2015M09 2022M04
Included observations: 80 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(DMONTHLYBIRTH(-1))	-3.228095	0.402205	-8.025997	0.0000
D(DMONTHLYBIRTH(-1),2)	1.191971	0.297880	4.001510	0.0001
D(DMONTHLYBIRTH(-2),2)	0.443130	0.146787	3.018868	0.0035
C	0.579052	0.926766	0.624809	0.5340

R-squared	0.759483	Mean dependent var	0.587500
Adjusted R-squared	0.749989	S.D. dependent var	16.57793
S.E. of regression	8.289154	Akaike info criterion	7.116480
Sum squared resid	5221.966	Schwarz criterion	7.235581
Log likelihood	-280.6592	Hannan-Quinn criter.	7.164231
F-statistic	79.99517	Durbin-Watson stat	1.575921
Prob(F-statistic)	0.000000		

Table 2 shows that the Augmented Dickey-Fuller test, at the first difference, is 0.00. Thus, the series is considered to be stationary at the first differencing. We can progress to determine a suitable model for the series.

March, Volume 10, Number 1, Pages 1 – 6 http://www.ijbst.fuotuoke.edu.ng/
https://doi.org/ 10.5555/RGGF8084 ISSN 2488-8648



	AC	PAC	Q-Stat	Prob
1	-0.176	-0.176	2.6940	0.101
2	-0.016	-0.049	2.7175	0.257
3	-0.048	-0.062	2.9225	0.404
4	0.057	0.037	3.2117	0.523
5	-0.083	-0.072	3.8380	0.573
6	0.042	0.016	4.0019	0.676
7	-0.040	-0.033	4.1524	0.762
8	0.055	0.037	4.4422	0.815
9	-0.101	-0.083	5.4282	0.796
10	0.074	0.037	5.9593	0.819
11	-0.013	0.008	5.9770	0.875
12	0.046	0.036	6.1867	0.906
13	-0.029	0.003	6.2729	0.936
14	-0.010	-0.031	6.2827	0.959
15	0.083	0.097	6.9978	0.958
16	0.019	0.037	7.0380	0.973
17	-0.079	-0.049	7.7081	0.972
18	0.027	0.002	7.7890	0.982
19	0.028	0.035	7.8731	0.988
20	-0.074	-0.067	8.4940	0.988
21	0.017	0.007	8.5276	0.992
22	-0.004	-0.015	8.5298	0.995
23	-0.161	-0.187	11.595	0.976
24	0.187	0.165	15.790	0.896
25	0.020	0.052	15.840	0.920
26	-0.085	-0.098	16.734	0.917
27	0.018	0.023	16.776	0.937
28	0.058	0.047	17.214	0.944
29	-0.083	-0.074	18.114	0.942
30	-0.006	-0.025	18.118	0.957
31	0.005	0.001	18.122	0.968
32	0.043	0.010	18.383	0.974
33	-0.090	-0.030	19.539	0.969
34	0.029	-0.013	19.662	0.976
35	-0.044	-0.046	19.941	0.981
36	0.020	-0.003	20.004	0.986

Fig.3 Correlogram of birth rate data after first difference

Table 3 Estimate of ARIMA Model

Dependent Variable: DMONTHLYBIRTH
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 11/14/23 Time: 08:55
Sample: 2015M05 2022M04
Included observations: 84
Convergence achieved after 23 iterations

Coefficient covariance computed using an outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.142059	0.272340	0.521622	0.6034
AR(1)	0.206107	0.382689	0.538575	0.0007
MA(1)	-0.857705	0.193780	-4.426189	0.0000
SIGMASQ	50.05519	3.705391	13.50875	0.0000

R-squared	0.146902	Mean dependent var	0.666667
Adjusted R-squared	0.114911	S.D. dependent var	7.705941
S.E. of regression	7.249686	Akaike info criterion	6.857966
Sum squared resid	4204.636	Schwarz criterion	6.973719
Log likelihood	-284.0346	Hannan-Quinn criter.	6.904498
F-statistic	4.591952	Durbin-Watson stat	1.383635
Prob(F-statistic)	0.005117		

Inverted AR Roots	.21
Inverted MA Roots	.86

March, Volume 10, Number 1, Pages 1 – 6 http://www.ijbst.fuotuoke.edu.ng/
https://doi.org/ 10.5555/RGGF8084 ISSN 2488-8648

Date: 11/14/23 Time: 08:50
Sample: 2015M04 2022M04
Included observations: 84
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.008	-0.008	0.0051	
		2	0.009	0.009	0.0117	
		3	-0.048	-0.048	0.2208	0.638
		4	0.000	-0.000	0.2208	0.895
		5	-0.074	-0.073	0.7179	0.869
		6	0.039	0.035	0.8559	0.931
		7	-0.028	-0.027	0.9306	0.968
		8	0.047	0.040	1.1435	0.980
		9	-0.020	-0.016	1.1804	0.991
		10	0.148	0.142	3.3241	0.912
		11	0.041	0.052	3.4925	0.942
		12	0.020	0.015	3.5328	0.966
		13	-0.014	0.007	3.5521	0.981
		14	-0.001	-0.002	3.5522	0.990
		15	0.062	0.090	3.9527	0.992
		16	-0.017	-0.025	3.9840	0.996
		17	-0.080	-0.076	4.6820	0.994
		18	0.015	0.007	4.7051	0.997
		19	-0.035	-0.036	4.8421	0.998
		20	-0.108	-0.133	6.1649	0.995
		21	-0.029	-0.058	6.2617	0.997
		22	-0.049	-0.072	6.5372	0.998
		23	-0.080	-0.104	7.3043	0.997
		24	0.146	0.142	9.8821	0.987
		25	0.015	-0.018	9.9115	0.992
		26	-0.090	-0.109	10.910	0.990
		27	-0.010	0.025	10.923	0.993
		28	-0.020	-0.009	10.974	0.996
		29	-0.102	-0.085	12.346	0.993
		30	-0.009	0.019	12.357	0.995
		31	-0.032	-0.007	12.492	0.997
		32	-0.018	-0.004	12.539	0.998
		33	-0.092	-0.066	13.740	0.997
		34	0.038	-0.003	13.954	0.998
		35	-0.008	-0.005	13.964	0.999
		36	-0.021	-0.018	14.033	0.999

Fig 4. Ljung-Box Q test of the residual

Conclusion

In this study we empirically evaluated the ARIMA model that will be appropriate for forecasting the monthly birth rate in Bayelsa State, Nigeria. The time plot of the series shown in Figure 1 follows an irregular sig-sag pattern showing both upward and downward movement. The Augmented Dickey Fuller test revealed that the monthly time series of birth rates contains unit root. The data was seen to be stationary in the first difference. An appropriate forecasting model was obtained using the Box-Jenkins method. Based on the behavior of ACF and PACF plots, the ARIMA (1,1,1) model was identified as the most adequate model. The model adequacy was confirmed through Ljung-Box *Q* test. The residual of the model was checked using Ljung-Box and it shows no sign of autocorrelation in the residual. We recommend that further studies should be carried out on forecasting Bayelsa State's monthly birth rate using the model.

References

Abel, G., Bijak, J., Forster, J., J., Raymer, J., Smith, P., W.F. and Wong, J., S. (2013) Integrating uncertainty in time series population forecasts: An illustration using a simple projection model. *Demographic Research*, 29 (43), 1187-1226.

Adsera, A. (2006). An economic analysis of the gap between desired and actual fertility: The
case of Spain. Review of Economics of the Household, 4(1):75–95.

Alho, J., J. Bravo and Palmer, E. (2012). Annuities and Life Expectancy in NDC. Gender, *Politics, and Financial Stability*, 2(2), 95 - 436.

Asefa, M., Drewett, R., &Tessema, F. (2000). A birth cohort study in South-West Ethiopia to identify factors associated with infant mortality that are amenable for intervention. Ethiopian Journal of Health Development, 14(2); 161-168

https://doi.org/ **10.5555/RGGF8084** ISSN 2488-8648

Arvind, S. and Gupta, P.C. (2012). Predicting the Number of Blood Donors through their Age and Blood Group by using Data Mining Tool‖ *International Journal of Communication and Computer Technologies* 1(6), 26-34

Bravo, J. M., Coelho, E., and Magalhães, M. G. (2010). Mortality projections in Portugal. Work session on demographic projections, EUROSTAT-EC Collection: Methodologies and working papers, Theme: Population and Social Conditions, 241–252.

Box G., Jenkins, M., G. (1976). Time series analysis: Forecasting and control, revised edition, Oakland, Ca: Holden-Day.

Graham, W., Ahmed, S., Stanton, C., Abou-Zahr, C., and Campbell, O. (2008). Measuring maternal mortality: An overview of opportunities and options for developing countries. *BMC Medicine*, 6(1):12.

Han, J. & Kamber, M. (2006). *Data Mining: Concepts and Techniques*, (2nd ed.). Amsterdam: Morgan Kaufmann

Hand, D., Heikki M., H., and Smyth P. (2011). *Principles of Data Mining*. Cambridge: MIT Press

Hill, K, El Arifeen S, Koenig M, Al-Sabir A, Jamil K, and Raggers H: (2006). How should we measure maternal mortality in the developing world? A comparison of household deaths and sibling history approaches. *Bulletin of the World Health Organization*, 84:173-180.

Hussein, M. A. (1993). Time series analysis for forecasting both fertility and mortality levels in Egypt until year 2010. The Egyptian Population and Family Planning Review, 27(2), 67–81.

Kirk, D. (1996). Demographic Transition Theory. Population Studies, 50(3), 361–387.

Lee, R. (1992). Stochastic demographic forecasting. International Journal of Forecasting, 8, 315–327.

Lee, R., & Tuljapurkar, S. (1994). Stochastic population forecasts for the United States: Beyond high, medium, and low. Journal of the American Statistical Association, 89, 1175–1189

Madise, N., Nyarko, P., & Diamond, I. (2016). Child Loss and Fertility Behaviour in Ghana. Retrieved from https://www.researchgate.net/publication/279685779

Rau, R. (2007). Seasonality in human mortality: a demographic approach. Demographic Research Monographs, Springer.

Palloni, A., &Rafalimanana, H. (1999). The Effects of Infant Mortality on Fertility Revisited: New Evidence from Latin America. Demography, 36(1), 41–58.

Shegaw, A. (2012). *Application of data mining technology to predict child mortality patterns:* the case BRHP, MSc. Thesis, Addis Ababa University, Ethiopia.

Soest, A. Van, &Saha, U. R. (2018). Relationships between infant mortality, birth spacing and fertility in Matlab, Bangladesh. PLoS ONE, 13(4), 1–21.

Tayman, J., Smith, S. K., and Lin, J. (2007). Precision, bias, and uncertainty for state population forecasts: an exploratory analysis of time series models. Population Research and Policy Review 26(3): 347–369