International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 301-308, August-2023
https://doi.org/10.47001/IRJIET/2023.708044

ARIMA Model Application in the Detection of Future Trends of Annual Neonatal Mortality Rate for Italy

[1]Dr. Smartson. P. NYONI, [2]Thabani NYONI

[1]ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe
[2]Independent Researcher & Health Economist, Harare, Zimbabwe

***Abstract* - This study uses annual time series data on neonatal mortality rate (NMR) for Italy from 1960 to 2019 to predict future trends of NMR over the period 2020 to 2030. Unit root tests have shown that the series under consideration is an I (2) variable. The optimal model based on AIC is the ARIMA (2,2,2) model. The findings of this study indicate that neonatal mortality will remain very low throughout the out of sample period. Therefore, we encourage policy makers in Italy to formulate neonatal policies that will address disparities in neonatal and infant mortality, and attend to other factors that contribute to deaths among neonates.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

Neonatal mortality strongly indicates the quality of health care services during antenatal, delivery, and postnatal periods, therefore analyzing neonatal mortality rate (NMR) trends is important to inform policy, planning, decision making and resource allocation to maternal & Child health programs (Simeon *et al.* 2019; Reidpath & Allotey, 2003). Italy has made significant progress in the reduction of child mortality rates. According to Italian Statistics Bureau (ISTAT), there is a disparity in neonatal and infant mortality among immigrants and Italian residents. Neonatal mortality rates have been found to be higher among foreigners compared to Italian residents. The aim of this study is to model and project future trends of neonatal mortality rate for Italy using the ARIMA model. This econometric and statistical technique is suitable for analyzing linear time series data (Nyoni, 2018; Box & Jenkins, 1970). The findings of this study are expected to detect abnormal future trends of neonatal mortality and stimulate an appropriate evidence based response to the problem of mortality among neonates.

II. LITERATURE REVIEW

Regression analysis was employed by Jawad *et al.* (2021) to assess the association between conflict and maternal and child health globally. Data for 181 countries (2000–2019) from the Uppsala Conflict Data Program and World Bank were analyzed using panel regression models. The study findings showed that armed conflict is associated with substantial and persistent excess maternal and child deaths globally. Harpur *et al.*(2021) investigated trends in infant mortality rates (IMR) and stillbirth rates by socio-economic position (SEP) in Scotland, between 2000 and 2018, inclusive. Data for live births, infant deaths, and stillbirths between 2000 and 2018 were obtained from National Records of Scotland. Annual IMR and stillbirth rates were calculated and visualized for all of Scotland and when stratified by SEP. Negative binomial regression models were used to estimate the association between SEP and infant mortality and stillbirth events, and to assess for break points in trends over time. The study revealed that IMR fell from 5.7 to 3.2 deaths per 1000 live births between 2000 and 2018, with no change in trend identified. Stillbirth rates were relatively static between 2000 and 2008 but experienced accelerated reduction from 2009 onwards. When stratified by SEP, inequalities in IMR and stillbirth rates persisted throughout the study and were greatest amongst the sub-group of post-neonates. Simeoni *et al*. (2019) analyzed the infant (IMR) and neonatal (NMR) mortality rates of Italian and foreign children and evaluated if there is a disparity among geographical macro-areas. Data from 2006 to 2015 were collected by the Italian Statistics Bureau (ISTAT) and extracted from two different national databases, which considered i) underlying cause of death and ii) birth registry. The main analyses were made comparing Italian versus foreigners as a single category as well as by country origin and contrasting Northern residents versus Southern ones. Comparisons between groups were done using relative risks. The study findings indicated that Inequalities in neonatal and infant mortality are evident between Italians and immigrants and among geographical macro-areas. The effects of individual bio-demographic and socioeconomic components on infant mortality were investigated by Scalone *et al*. (2016). The study utilized micro data from births, deaths and marriages civil registers of Granarolo, an Italian rural municipality close to Bologna, from 1900 to 1939 and then reconstructed some typical bio-demographic characteristics and the socioeconomic status of parents. Cox and Piecewise constant exponential models were used

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 301-308, August-2023
https://doi.org/10.47001/IRJIET/2023.708044

to estimate the effects of the selected predictors. The study indicated that still in the first four decades of the twentieth century rural daily wagers experienced a lower level in infant survivor, whereas the upper class registered significantly higher ones.

III. METHODOLOGY

The Autoregressive (AR) Model

A process Y_t (annual NMR at time t) is an autoregressive process of order p, that is, AR (p) if it is a weighted sum of the past p values plus a random shock (Z_t) such that:

$$Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \emptyset_3 Y_{t-3} + \cdots + \emptyset_p Y_{t-p} + Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [1]$$

Using the backward shift operator, B, such that $BY_t = Y_{t-1}$, the AR (p) model can be expressed as in equation [2] below:

$$Z_t = \emptyset(B) Y_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [2]$$

where $\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \emptyset_3 B^3 - \cdots - \emptyset_p B^p$

The 1st order AR (p) process, AR (1) may be expressed as shown below:

$$Y_t = \emptyset Y_{t-1} + Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [3]$$

Given $\emptyset = 1$, then equation [3] becomes a random walk model. When $|\emptyset| > 1$, then the series is reffered to as explosive, and thus non-stationary. Generally, most time series are explosive. In the case where $|\emptyset| < 1$, the series is said to be stationary and therefore its ACF (autocorrelation function) decreases exponentially.

The Moving Average (MA) Model

A process is reffered to as a moving average process of order q, MA (q) if it is a weighted sum of the last random shocks, that is:

$$Y_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [4]$$

Using the backward shift operator, B, equation [4] can be expressed as follows:

$$Y_t = \theta(B) Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [5]$$

where $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$

Equation [4] can also be expressed as follows:

$$Y_t - \sum_{j \leq 1} \pi_j Y_{t-j} = Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [6]$$

for some constant π_j such that:

$$\sum_{j \leq 1} |\pi_j| < \infty$$

This implies that it is possible to invert the function taking the Z_t sequence to the Y_t sequence and recover Z_t from present and past values of Y_t by a convergent sum.

The Autoregressive Moving Average (ARMA) Model

While the above models are good, a more parsimonious model is the ARMA model. The AR, MA and ARMA models are applied on stationary time series only. The ARMA model is just a mixture of AR (p) and MA (q) terms, hence the name ARMA (p, q). This can be expressed as follows:

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 301-308, August-2023
https://doi.org/10.47001/IRJIET/2023.708044

$$\emptyset(B)Y_t = \theta(B)Z_t \ldots\ldots\ldots [7]$$

Thus:

$$Y_t\left(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p\right) = Z_t\left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right) \ldots\ldots [8]$$

where $\emptyset(B)$ and $\theta(B)$ are polynomials in B of finite order p, q respectively.

The Autoregressive Integrated Moving Average (ARIMA) Model

The AR, MA and ARMA processes are usually not applied empirically because in most cases many time series data are not stationary; hence the need for differencing until stationarity is achieved.

$$\left.\begin{array}{c} The\ first\ difference\ is\ given\ by: \\ Y_t - Y_{t-1} = Y_t - BY_t \\ The\ second\ difference\ is\ given\ by: \\ Y_t(1-B) - Y_{t-1}(1-B) = Y_t(1-B) - BY_t(1-B) = Y_t(1-B)(1-B) = Y_t(1-B)^2 \\ The\ third\ difference\ is\ given\ by: \\ Y_t(1-B)^2 - Y_{t-1}(1-B)^2 = Y_t(1-B)^2 - BY_t(1-B)^2 = Y_t(1-B)^2(1-B) = Y_t(1-B)^3 \\ The\ d^{th}\ difference\ is\ given\ by: \\ Y_t(1-B)^d \end{array}\right\} \ldots [9]$$

Given the basic algebraic manipulations above, it can be inferred that when the actual data series is differenced “d” times before fitting an ARMA (p, q) process, then the model for the actual undifferenced series is called an ARIMA (p, d, q) model. Thus equation [7] is now generalized as follows:

$$\emptyset(B)(1-B)^d Y_t = \theta(B)Z_t \ldots\ldots\ldots [10]$$

Therefore, in the case of modeling and forecasting NMR, equation [10] can be written as follows:

$$\emptyset(B)(1-B)^d Y_t = \theta(B)Z_t \ldots\ldots\ldots [11]$$

The Box – Jenkins Approach

The first step towards model selection is to difference the series in order to achieve stationarity. Once this process is over, the researcher will then examine the correlogram in order to decide on the appropriate orders of the AR and MA components. It is important to highlight the fact that this procedure (of choosing the AR and MA components) is biased towards the use of personal judgement because there are no clear – cut rules on how to decide on the appropriate AR and MA components. Therefore, experience plays a pivotal role in this regard. The next step is the estimation of the tentative model, after which diagnostic testing shall follow. Diagnostic checking is usually done by generating the set of residuals and testing whether they satisfy the characteristics of a white noise process. If not, there would be need for model re – specification and repetition of the same process; this time from the second stage. The process may go on and on until an appropriate model is identified (Nyoni, 2018). The Box – Jenkins technique was proposed by Box & Jenkins (1970) and is widely used in many forecasting contexts, including public health. In this paper, hinged on this technique; the researcher will use automatic ARIMA modeling for estimating equation [10].

Data Issues

This study is based on annual NMR in Italy for the period 1960 to 2019. The out-of-sample forecast covers the period 2020 to 2030. All the data employed in this research paper was gathered from the World Bank online database.

Evaluation of ARIMA Models

Criteria Table

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 301-308, August-2023

https://doi.org/10.47001/IRJIET/2023.708044

Table 2: Criteria Table

Model Selection Criteria Table
Dependent Variable: D(Y, 2)
Date: 01/22/22 Time: 14:45
Sample: 1960 2019
Included observations: 58

Model	LogL	AIC*	BIC	HQ
(2,2)(0,0)	41.930703	-1.238990	-1.025840	-1.155964
(2,3)(0,0)	42.236082	-1.215037	-0.966363	-1.118174
(3,2)(0,0)	42.197797	-1.213717	-0.965043	-1.116854
(3,3)(0,0)	42.392943	-1.185964	-0.901765	-1.075262
(4,2)(0,0)	42.354998	-1.184655	-0.900456	-1.073954
(2,4)(0,0)	42.265132	-1.181556	-0.897357	-1.070855
(2,5)(0,0)	42.988122	-1.172004	-0.852280	-1.047465
(3,4)(0,0)	42.923698	-1.169783	-0.850059	-1.045244
(4,3)(0,0)	42.541431	-1.156601	-0.836877	-1.032062
(5,2)(0,0)	42.496116	-1.155038	-0.835315	-1.030500
(5,0)(0,0)	40.471047	-1.154174	-0.905500	-1.057310
(4,4)(0,0)	43.416753	-1.152302	-0.797053	-1.013925
(3,5)(0,0)	43.269553	-1.147226	-0.791977	-1.008849
(5,1)(0,0)	41.177126	-1.144039	-0.859840	-1.033338
(3,0)(0,0)	37.962691	-1.136645	-0.959020	-1.067456
(4,0)(0,0)	38.721330	-1.128322	-0.915172	-1.045296
(5,5)(0,0)	44.647384	-1.125772	-0.699473	-0.959720
(5,3)(0,0)	42.595721	-1.123990	-0.768742	-0.985614
(4,1)(0,0)	39.417376	-1.117841	-0.869166	-1.020977
(5,4)(0,0)	42.938969	-1.101344	-0.710570	-0.949130
(0,4)(0,0)	37.188061	-1.075450	-0.862301	-0.992424
(4,5)(0,0)	42.103798	-1.072545	-0.681771	-0.920331
(2,1)(0,0)	36.096389	-1.072289	-0.894665	-1.003101
(1,4)(0,0)	37.659916	-1.057238	-0.808564	-0.960375
(0,5)(0,0)	37.493068	-1.051485	-0.802811	-0.954622
(1,3)(0,0)	36.050646	-1.036229	-0.823080	-0.953203
(3,1)(0,0)	35.755809	-1.026062	-0.812913	-0.943036
(2,0)(0,0)	33.132560	-1.004571	-0.862472	-0.949220
(1,5)(0,0)	36.958012	-0.998552	-0.714353	-0.887851
(1,2)(0,0)	33.786543	-0.992639	-0.815015	-0.923451
(0,1)(0,0)	30.866169	-0.960902	-0.854328	-0.919389
(0,3)(0,0)	32.801840	-0.958684	-0.781060	-0.889496
(0,2)(0,0)	31.542234	-0.949732	-0.807633	-0.894382
(1,1)(0,0)	31.002664	-0.931126	-0.789027	-0.875776
(1,0)(0,0)	29.796218	-0.924008	-0.817433	-0.882495
(0,0)(0,0)	25.412402	-0.807324	-0.736274	-0.779649

International Research Journal of Innovations in Engineering and Technology (IRJIET)

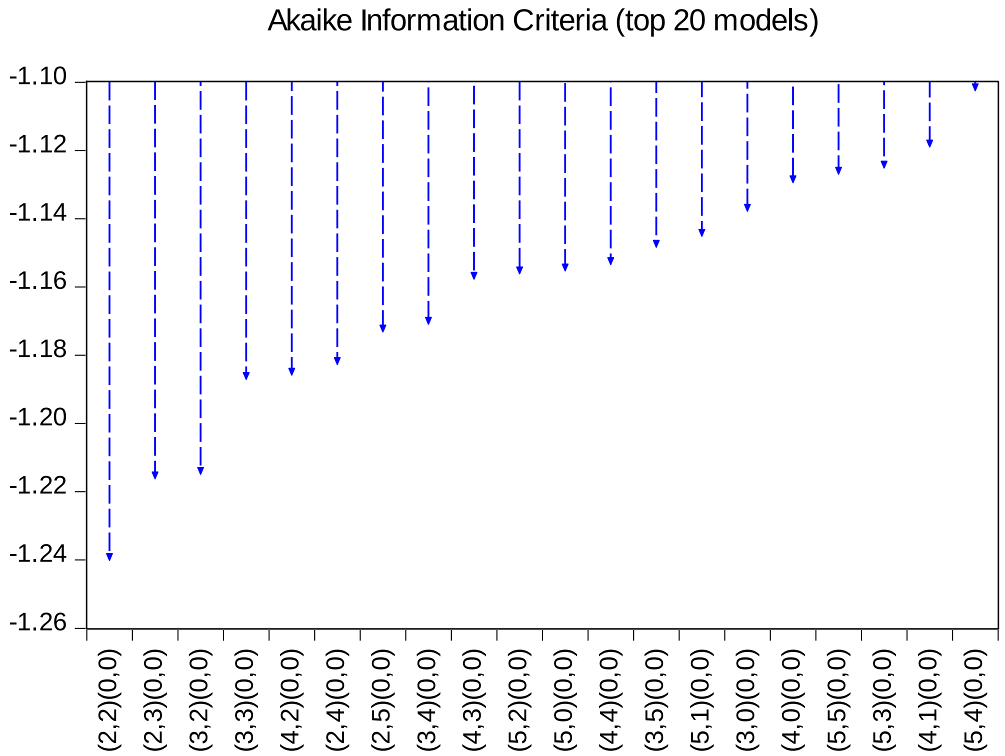
ISSN (online): 2581-3048

Volume 7, Issue 8, pp 301-308, August-2023

https://doi.org/10.47001/IRJIET/2023.708044

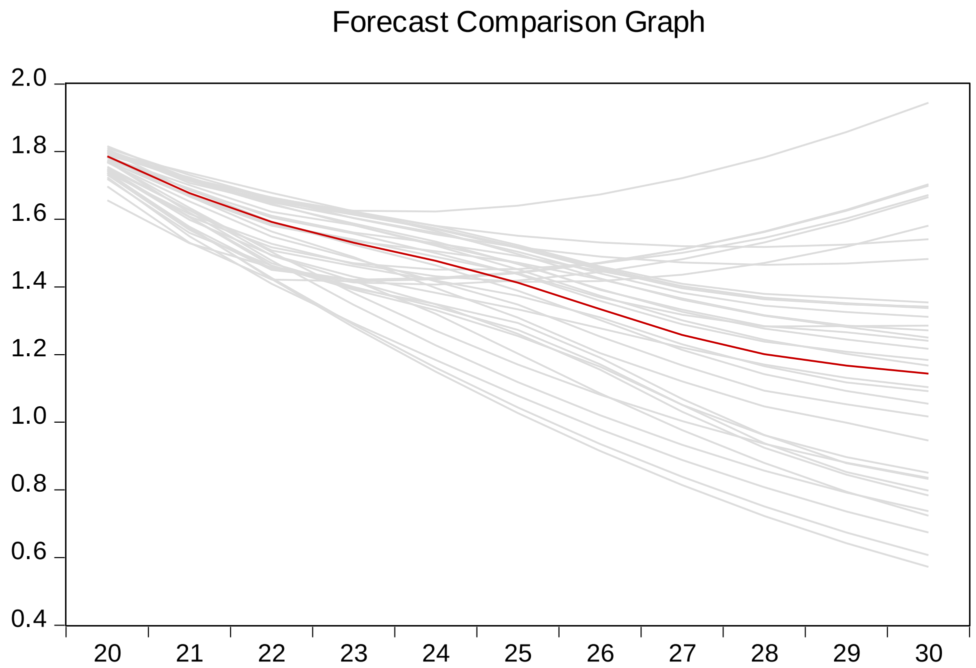
Criteria Graph

Figure 1: Criteria Graph



Forecast Comparison Graph

Figure 2: Forecast Comparison Graph



ISSN (online): 2581-3048
Volume 7, Issue 8, pp 301-308, August-2023
https://doi.org/10.47001/IRJIET/2023.708044

` Table 2 and Figure 1 indicate that the optimal model is the ARIMA (2,2,2) model. Figure 2 is a combined forecast comparison graph showing the out-of-sample forecasts of the top 25 models evaluated based on the AIC criterion. The red line shows the forecast line graph of the optimal model, the ARIMA (2,2,2) model.

IV. RESULTS

Summary of the Selected ARIMA () Model

Table 3: Summary of the Optimal Model

Automatic ARIMA Forecasting
Selected dependent variable: D(Y, 2)
Date: 01/22/22 Time: 14:45
Sample: 1960 2019
Included observations: 58
Forecast length: 11

Number of estimated ARMA models: 36
Number of non-converged estimations: 0
Selected ARMA model: (2,2)(0,0)
AIC value: -1.2389897486

Main Results of the Selected ARIMA () Model

Table 4: Main Results of the Optimal Model

Dependent Variable: D(Y,2)
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/22/22 Time: 14:45
Sample: 1962 2019
Included observations: 58
Convergence achieved after 23 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.005275	0.016997	0.310346	0.7575
AR(1)	1.034058	0.038995	26.51805	0.0000
AR(2)	-0.962467	0.032337	-29.76405	0.0000
MA(1)	-0.749224	0.115297	-6.498183	0.0000
MA(2)	0.734540	0.147480	4.980608	0.0000
SIGMASQ	0.013217	0.003499	3.777900	0.0004

R-squared	0.457773	Mean dependent var	0.010345
Adjusted R-squared	0.405636	S.D. dependent var	0.157491
S.E. of regression	0.121418	Akaike info criterion	-1.238990
Sum squared resid	0.766596	Schwarz criterion	-1.025840
Log likelihood	41.93070	Hannan-Quinn criter.	-1.155964
F-statistic	8.780167	Durbin-Watson stat	1.990680
Prob(F-statistic)	0.000004		

Inverted AR Roots	.52+.83i	.52-.83i

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 301-308, August-2023

https://doi.org/10.47001/IRJIET/2023.708044

Inverted MA Roots	.37+.77i	.37-.77i

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

2020	1.786303657575311
2021	1.677572674009463
2022	1.592055850022722
2023	1.530662235670341
2024	1.476767973707193
2025	1.412308098455631
2026	1.334602298415332
2027	1.258265908346891
2028	1.20099174551709
2029	1.167008434871596
2030	1.143659868334912

Table 2 clearly indicates that neonatal mortality will remain very low throughout the out of sample period.

V. POLICY IMPLICATION & CONCLUSION

European countries have made tremendous progress in the control of maternal and child mortality, however neonatal mortality still remains an important public health problem. Several previous studies conducted globally have shown that neonatal deaths are as a result of prematurity, birth asphyxia, neonatal sepsis and congenital anomalies. Studies that have been carried out in Italy have revealed a disparity in neonatal and infant mortality among immigrants and Italian residents. It was established that the rates were higher among foreigners when compared to Italian residents. This study proposes the popular Box-Jenkins ARIMA technique to predict future trends of neonatal mortality rate for Italy and the findings indicate that neonatal mortality will remain very low throughout the out of sample period. Therefore, we encourage policy makers in Italy to formulate neonatal policies that will address disparities in neonatal and infant mortality, and attend to other factors that contribute to deaths among neonates.

REFERENCES

[1] Box, D. E., and Jenkins, G. M. (1970). Time Series Analysis, Forecasting and Control, Holden Day, London.

[2] Nyoni, T. (2018). Box-Jenkins ARIMA Approach to Predicting net FDI Inflows in Zimbabwe, *University Library of Munich*, MPRA Paper No. 87737.

[3] Istat (2014). “La mortalità dei bambini ieri e oggi in Italia” in Statistiche Focus, 2014. Disponibile sul sito: www.istat.it/it/archivio/109861.

[4] Simeoni S., Frova L., and De Curtis M (2019). Inequalities in infant mortality in Italy, Italian Journal of Pediatrics (2019) 45:11

[5] Reidpath D. D., and Allotey P (2003). Infant mortality rate as an indicator of population health. J Epidemiol Community Health.

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 301-308, August-2023
https://doi.org/10.47001/IRJIET/2023.708044

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "ARIMA Model Application in the Detection of Future Trends of Annual Neonatal Mortality Rate for Italy" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 301-308, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708044

****