DESIGN AND DEVELOPMENT OF A DATA-MINING MODEL FOR CENSUS ANALYSIS WITH GEOGRAPHICAL INFORMATION SYSTEM (GIS)

Ogochukwu Clementina Okeke[1*], Esther Nneka Okonkwo[2]
[1*] Department of Computer Science, Chukwuemeka Odumegwu Ojukwu University, Uli, Nigeria
[2]Department of Computer Science, Chukwuemeka Odumegwu Ojukwu University, Uli, Nigeria.

Abstract: ***A new model for National Population Commission and Survey in Nigeria was developed to obtain geo-spatial distribution of predicted demographic information. Model used in census analysis gives prediction without searching for existing patterns or trends and without geo-spatial distribution of predicted demographic information. The system is a desktop application based on states and behaviours of objects. Object-Oriented Analysis and Design Methodology was adopted in this study and Microsoft Visual Basic.NET was used to implement the system. This was supported with SQLite database, a free distributed database engine. The new model developed automated the process of searching for patterns in the census data. The system is a multi-line user input interface which allows users to input specific data input to make predictions. The output of this paper is shown on the Nigerian map comprising of thirty six states, table data and bar chart simultaneously.***

Keywords: Data-mining, Geographical Information System (GIS),Nigeria, Population.

1.0 Introduction

Nigeria has had a long history of census taking. The first census was in 1866. The censuses of 1866, 1871, 1896, were restricted to Lagos Island and parts of Lagos Mainland. The censuses of 1901, 1911 and 1921 covered, in addition to Lagos a few more urban towns in the colony. Most of these censuses were actually population estimates. Although the census of 1952/53 was elaborate in organization, its non-simultaneity, which has implications for possible double counting is considered its weak points.The first post-independence census in Nigeria was carried out in 1962. This was out rightly cancelled and another conducted in 1963. The 1963 census results became the official figure and were used until the 1991 census was accepted. There was a census in 1973 but the results were declared unacceptable on account of massive inaccuracies. The 1991 census broke the myth of failed censuses in Nigeria. In 2006, Nigeria had its last census [1]. It provides the total characteristics in every town, village or /and locality. This information provides data for planning programmes in education.Information on housing is dispensable to planers and policy makers in evaluating housing conditions, estimating housing needs and formulating housing policies [2]. Most countries of the world conduct population and economic censuses at regular intervals. Population census information is of great value in planning public services for governments at all levels, such as cities, counties, provinces, and states. Both population and economic census data are also used by private companies or community organizations for various purposes, such as marketing studies, situating new factories or shopping malls, developing social service programs. Therefore, the application of data mining techniques to census data has great potential both in underpinning good public policy and in supporting business developments. However, mining census data is not straight forward and requires challenging methodological research [3].Census analyses in Nigeria do not search for existing patterns or trends during prediction to bring out hidden and important attributes in census data. A lot of data are not discovered during analysis due to the model in existence. National Population Commission has conducted out a lot of predictions in past, having much data that requires mining for further analysis. In this paper a model that predicts and gives predicted information on the map of Nigeria using dot density was developed. The new model developed searches for existing patterns using selected number of data input to carry forecast. The data-mining algorithm is the mechanism that creates mining models. To create a model, an algorithm first analyzes a set of data, looking for specific patterns and trends [4].The algorithm then uses the result of this analysis to define the parameters of the mining model to give geo-spatial distribution. The data-mining model this algorithm creates takes various forms including; Marital status, gender, unemployment, employment, number of registered births, number of registered deaths and total population in each States and give geo-spatial distribution of predicted demographic information. The data-mining extract these attributes out from pool of census database. Data-mining techniques are the result of a long process of research and product development [5].This evolution began when business data was first stored on computers, continued with improvements in data access, and more recently, generated technologies that allow users to navigate through their data in real time. Data-mining takes evolutionary process beyond retrospective data access and navigation to prospective and proactive information delivery. Data-mining is ready for application in the business community because it is supported by three technologies that are now sufficiently mature, which are massive data collection, powerful multiprocessors computers and data-mining algorithms [5]. The concept of data-mining is becoming increasingly popular as a business information management tool where it is expected to reveal knowledge structures that can guide decisions in conditions of limited certainty. Recently, there has been increased interest in developing new analytic techniques specifically designed to address the issues relevant to business data-mining (e.g Classification Trees), but data-mining is still based on the conceptual principles of statistics including the traditional exploratory data analysis and modeling and it shares with them both some components of its general approaches and specific techniques.

Data mining is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems [6]. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Data mining has two tasks as follows: Prediction method: Use some variables to predict unknown or future values of other variables [7]. Description method: Find human-interpretable patterns that describe the data

[7].The first and simplest analytical step in data mining is to describe the data, summarize its statistical attributes (such as means and standard deviations), visually review it using charts and graphs, and look for potentially meaningful links among variables (such as values that often occur together). A good model should never be confused with reality but it can be a useful guide to understanding your business [8]. Data mining is increasingly popular because of the substantial contribution it can make. It can be used to control costs as well as contribute to revenue increases.Data-mining process consists of three stages: the initial exploration, model building or pattern identification with validation and verification and deployment (i.e. the application of the model to new data in order to generate predictions) [9]. A geographic information system or geographical informationsystem (GIS) is a system designed to capture, store, manipulate, analyze, manage, and present all types of spatial or geographical data [10]. A GIS is an organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information [11].

1.1 Statement of the Problem

The problems identified with census analyses in Nigeria are;

- **a.** Census map is visual representation of States, local governments and towns and there is no prediction of demographic information on it.
- **b.** Census demographic analysis is calculated without mining any pattern or trend of existing demographic information.
- **c.** Census figures are not stored in a data-mining warehouse; population distribution decisions, finding hidden patterns and relationship become a problem.
- **d.** Census maps do not contain any chart thatshows how predictions are narrowed down to a particular State.

1.2 Objectives of the Paper

- **a.** To develop a model that accepts given probabilistic demographic information outputs a better prediction of future values.
- **b.** To develop a population density map of Nigeria using dot density.
- **c.** To develop a data-mining warehouse and obtain geo- spatial distribution attributes of a population.
- **d.** To make population distribution decisions, find hidden patterns and relationships from census database.
- **e.** To develop a model that makes use of chart to illustrate population predictions in each State in Nigeria.

1.3 Importance of the Paper

With the predictive capability of data-mining and geographical information system that gives geo-spatial distribution of population, Nigerian government will be more proactive in formulating economic and other developmental plans. The work will provide support for tactical and strategic business decisions.Geographical information system has the ability to separate information in layers, and then combine it with other layers of information which is used in research and for decision making tools.

1.4 Related Literatures

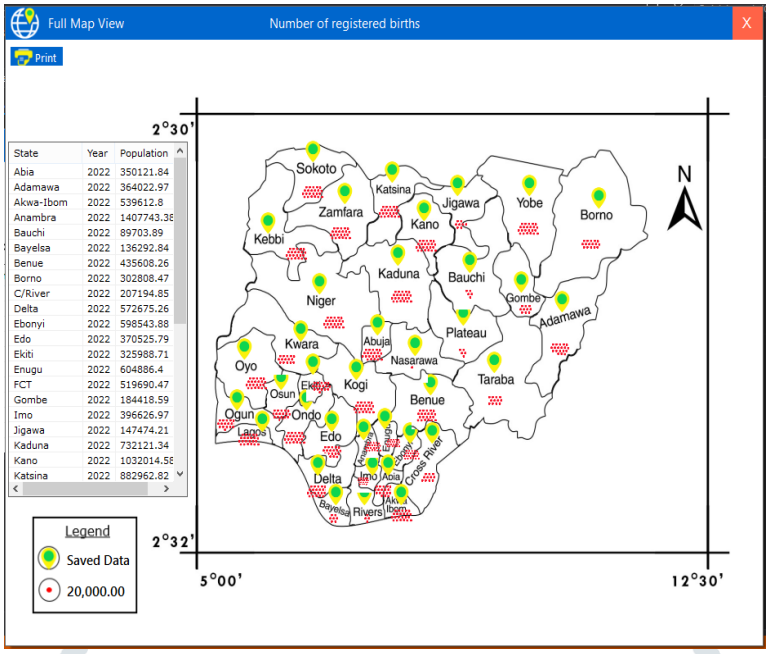
Ramezankhani et al. [12] Identified a low risk population for type 2 diabetes, Tehran lipid and glucose study using decision tree algorithm. The output was to find out number of people with Type 2 diabetes, Tehran and glucose level in future. Manoj and Madhu [13] applied a time series ARIMA model for predicting sugarcane production in India; their output was to forecast sugar cane production in future years. Md and Syed [14] used cohort-component method in forecasting future levels of fertility, mortality, sex composition, migration and other parameters and their output was to forecast factors that cause population change. David, Alan and Bob [15] applied Hamilton –Perry Method to forecast population using data of two recent censuses and the result showed how population increased in years to come.Kayode and Jimoh [16] used regression analysis with time series data to periodically forecast stock market prices stock market prices and have been proven complement to other numeric forecasting method and the result was to predict stock market prices in future and financial institution. Friedrich and Weixin [17] predicted literacy rate based on age disaggregated literacy data and demographic data collected from UN population using Global Age Specific Literacy Projection (GALP). Saigal and Mehrotra [18] predicted total birth using autoregressive integrated moving average (ARIMA) and output was predict number of births over given years. Rajasekhar and Karth [19] proposed hybrid support vector machine technique for weather prediction using Artificial Neural Network (ANN) and result yielded Euclidean distance on monthly mean of each year average temperature.

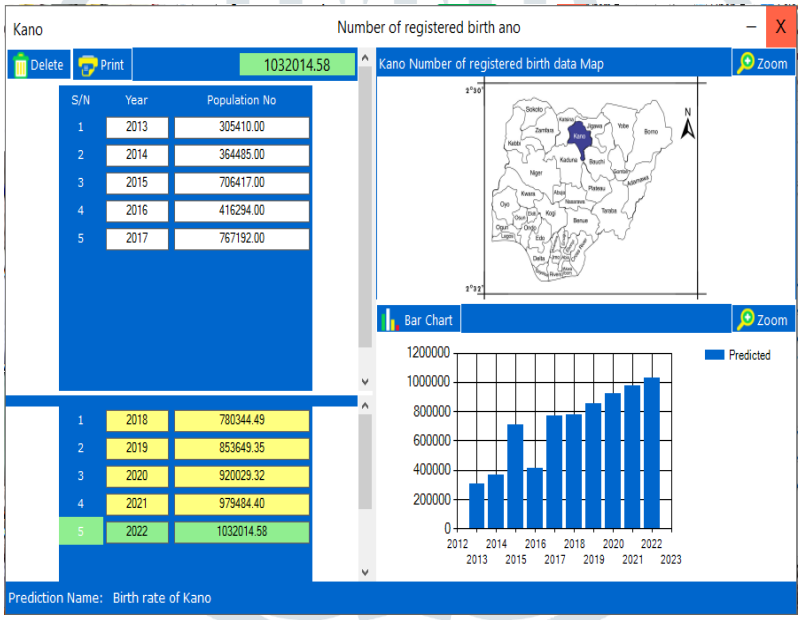
2.0 Methodology

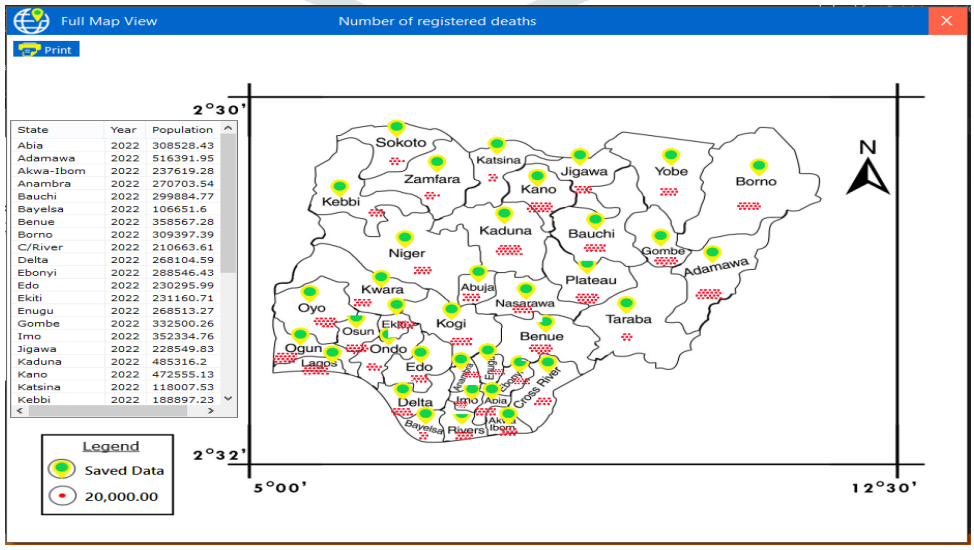
Object-Oriented Analysis and Design Methodology (OOADM) was used. The model used in analysis was $Pt= (X'X + \ln KI)^{-1}X'Y$. Where P_t is number of people at a future time, X'X is a matrix with rows and columns; a three by three matrix, lnKI is an identity matrix with three by three matrix with a value 0.000099995 and Y is acentered n-vector that is three by one matrix. The calculations were done using Microsoft Excel to obtain a constant, and time coefficients which is $P_t = \beta_0 + \beta_1 t + \beta_2 t^2$

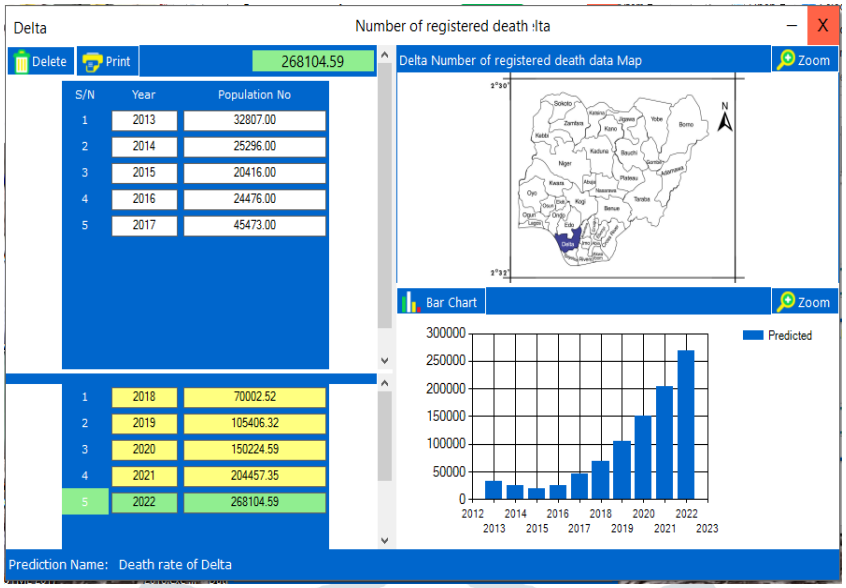
3.0 Results

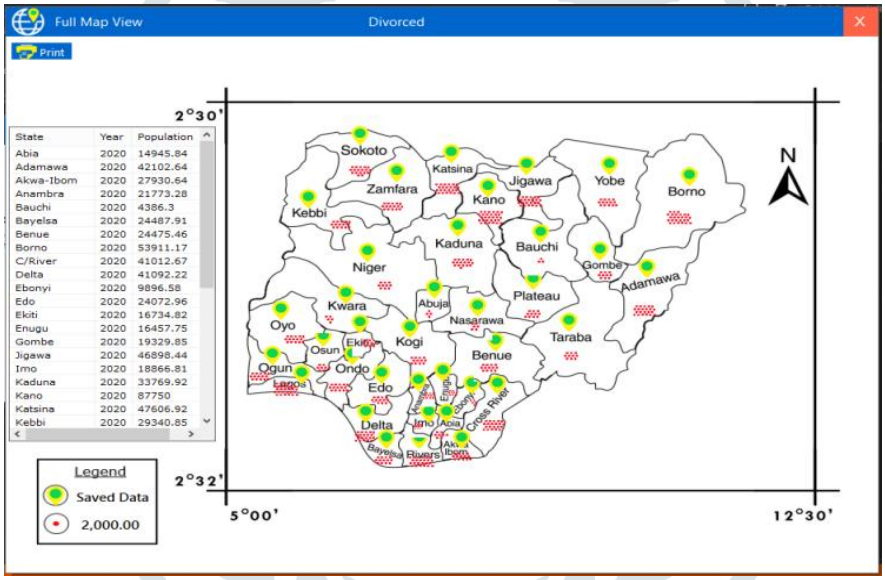
The output of this paper are in three forms; dot density map, table data showing number of data inputs and predicted data with year and bar chart illustrating prediction based on the predicted figures.

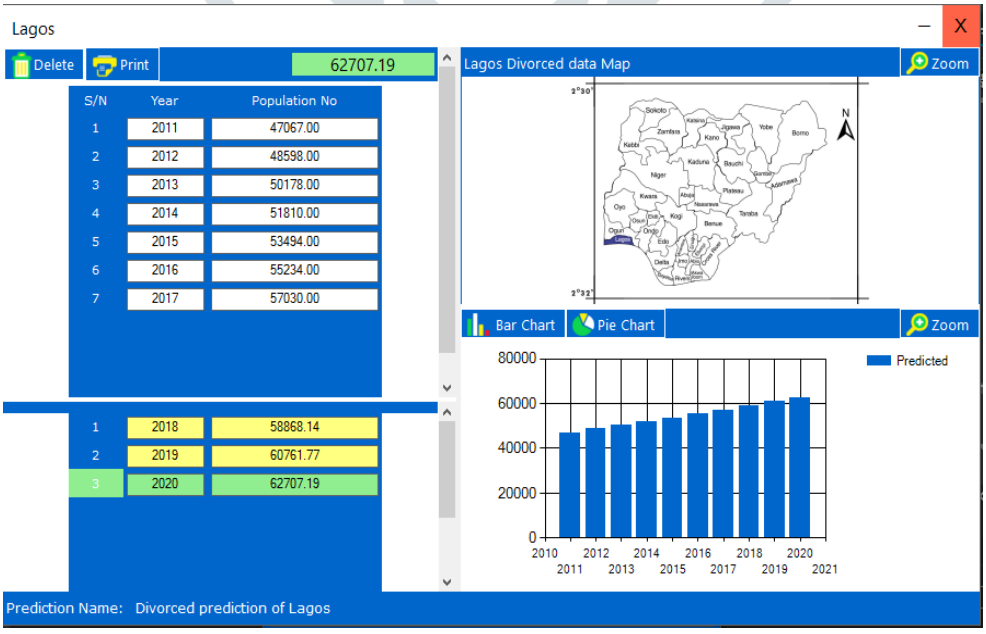

Full Map View
Number of registered births
Print
State Year Population
Abia 2022 350121.84
Adamawa 2022 364022.97
Akwa-Ibom 2022 539612.8
Anambra 2022 1407743.38
Bauchi 2022 89703.89
Bayelsa 2022 136292.84
Benue 2022 435608.26
Borno 2022 302808.47
C/River 2022 207194.85
Delta 2022 572675.26
Ebonyi 2022 598543.88
Edo 2022 370525.79
Ekiti 2022 325988.71
Enugu 2022 604886.4
FCT 2022 519690.47
Gombe 2022 184418.59
Imo 2022 396626.97
Jigawa 2022 147474.21
Kaduna 2022 732121.34
Kano 2022 1032014.58
Katsina 2022 882962.82
Legend
Saved Data
20,000.00
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Katsina
Zamfara
Jigawa
Yobe
Borno
Kebbi
Kano
Kaduna
Bauchi
Niger
Gombe
Adamawa
Plateau
Abuja
Kwara
Nasarawa
Oyo
Taraba
Kogi
Osun
Ekiti
Benue
Ogun
Ondo
Lagos
Edo
Delta
Imo
Abia
Cross River
Bayelsa
Rivers
Akwa Ibom

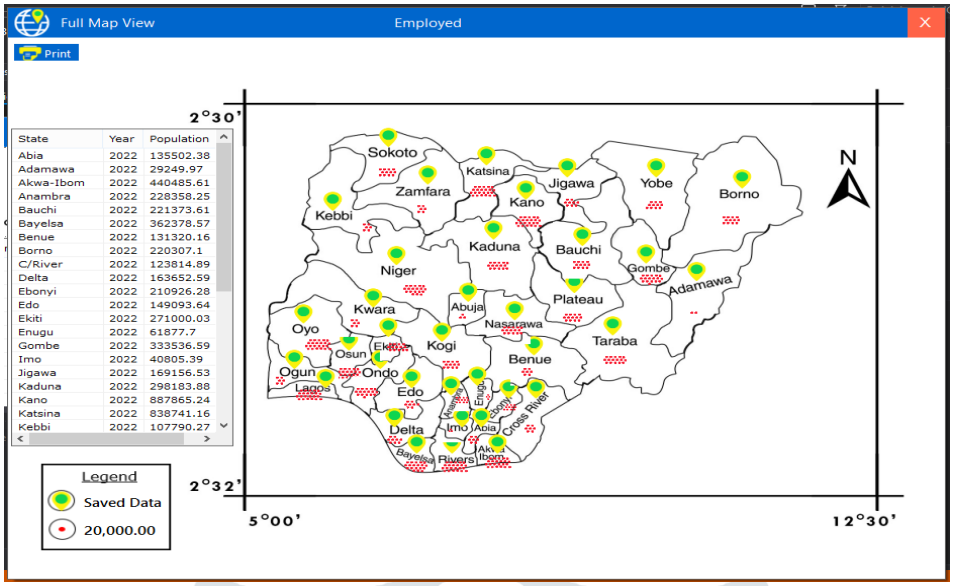

Kano
Number of registered birth ano
Delete
Print
1032014.58
S/N Year Population No
1 2013 305410.00
2 2014 364485.00
3 2015 706417.00
4 2016 416294.00
5 2017 767192.00
1 2018 780344.49
2 2019 853649.35
3 2020 920029.32
4 2021 979484.40
5 2022 1032014.58
Kano Number of registered birth data Map
Zoom
Bar Chart
Zoom
Predicted
1200000
1000000
800000
600000
400000
200000
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023
Prediction Name: Birth rate of Kano

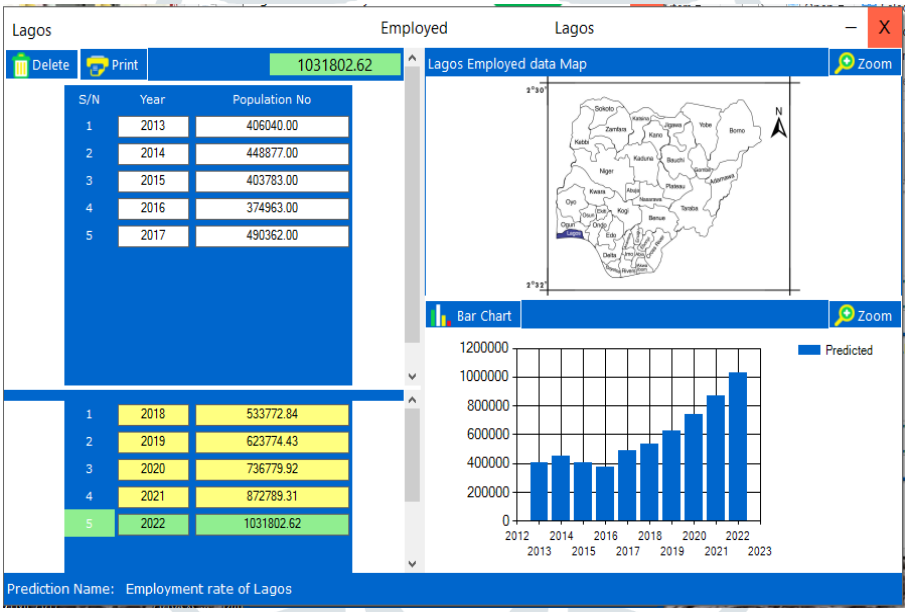

Full Map View
Number of registered deaths
Print
State Year Population
Abia 2022 308528.43
Adamawa 2022 516391.95
Akwa-Ibom 2022 237619.28
Anambra 2022 270703.54
Bauchi 2022 299884.77
Bayelsa 2022 106651.6
Benue 2022 358567.28
Borno 2022 309397.39
C/River 2022 210663.61
Delta 2022 268104.59
Ebonyi 2022 288546.43
Edo 2022 230295.99
Ekiti 2022 231160.71
Enugu 2022 268513.27
Gombe 2022 332500.26
Imo 2022 352334.76
Jigawa 2022 228549.83
Kaduna 2022 485316.2
Kano 2022 472555.13
Katsina 2022 118007.53
Kebbi 2022 188897.23
Legend
Saved Data
20,000.00
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Katsina
Zamfara
Jigawa
Yobe
Borno
Kebbi
Kano
Kaduna
Bauchi
Niger
Gombe
Adamawa
Plateau
Abuja
Kwara
Nasarawa
Oyo
Taraba
Kogi
Osun
Ekiti
Benue
Ogun
Ondo
Lagos
Edo
Delta
Imo
Abia
Cross River
Bayelsa
Rivers
Akwa Ibom

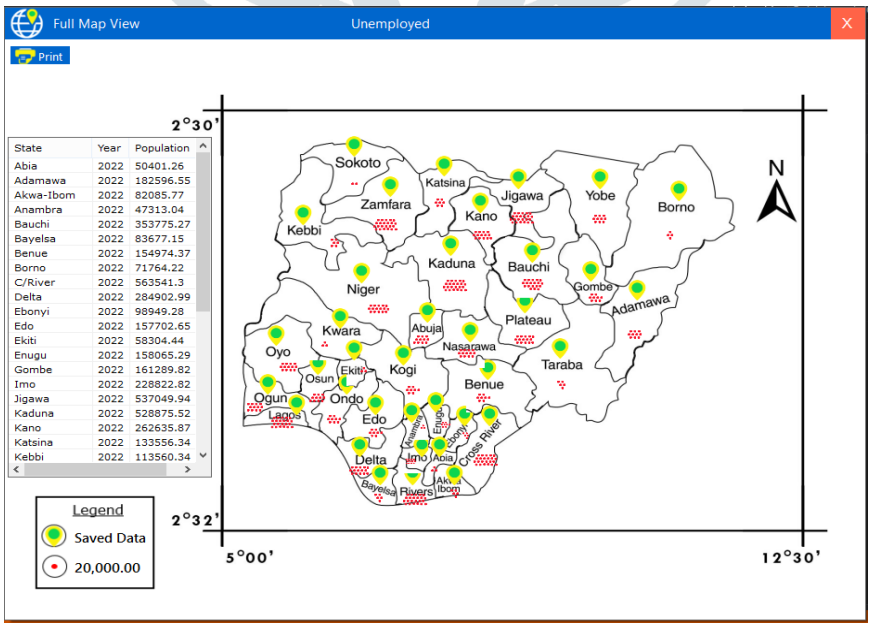

Delta
Number of registered death :lta
Delete
Print
268104.59
S/N Year Population No
1 2013 32807.00
2 2014 25296.00
3 2015 20416.00
4 2016 24476.00
5 2017 45473.00
1 2018 70002.52
2 2019 105406.32
3 2020 150224.59
4 2021 204457.35
5 2022 268104.59
Delta Number of registered death data Map
Zoom
Bar Chart
Zoom
Predicted
300000
250000
200000
150000
100000
50000
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023
Prediction Name: Death rate of Delta

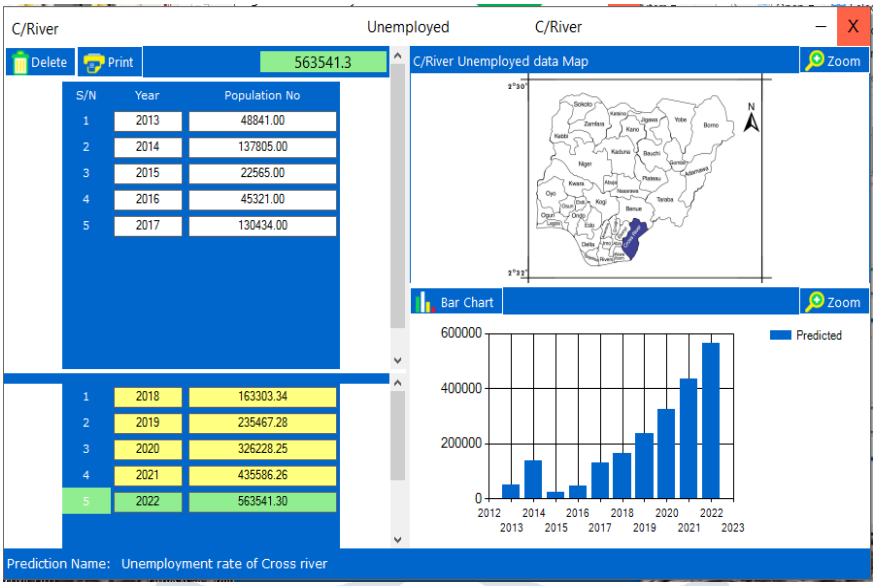

Full Map View
Divorced
Print
State Year Population
Abia 2020 14945.84
Adamawa 2020 42102.64
Akwa-Ibom 2020 27930.64
Anambra 2020 21773.28
Bauchi 2020 4386.3
Bayelsa 2020 24487.91
Benue 2020 24475.46
Borno 2020 53911.17
C/River 2020 41012.67
Delta 2020 41092.22
Ebonyi 2020 9896.58
Edo 2020 24072.96
Ekiti 2020 16734.82
Enugu 2020 16457.75
Gombe 2020 19329.85
Jigawa 2020 46898.44
Imo 2020 18866.81
Kaduna 2020 33769.92
Kano 2020 87750
Katsina 2020 47606.92
Kebbi 2020 29340.85
Legend
Saved Data
2,000.00
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Zamfara
Katsina
Kano
Jigawa
Yobe
Borno
Kebbi
Kaduna
Bauchi
Gombe
Adamawa
Niger
Abuja
Plateau
Kwara
Nasarawa
Taraba
Oyo
Osun
Ekiti
Kogi
Benue
Ogun
Ondo
Lagos
Edo
Delta
Imo
Abia
Cross River
Rivers
Akwa Ibom

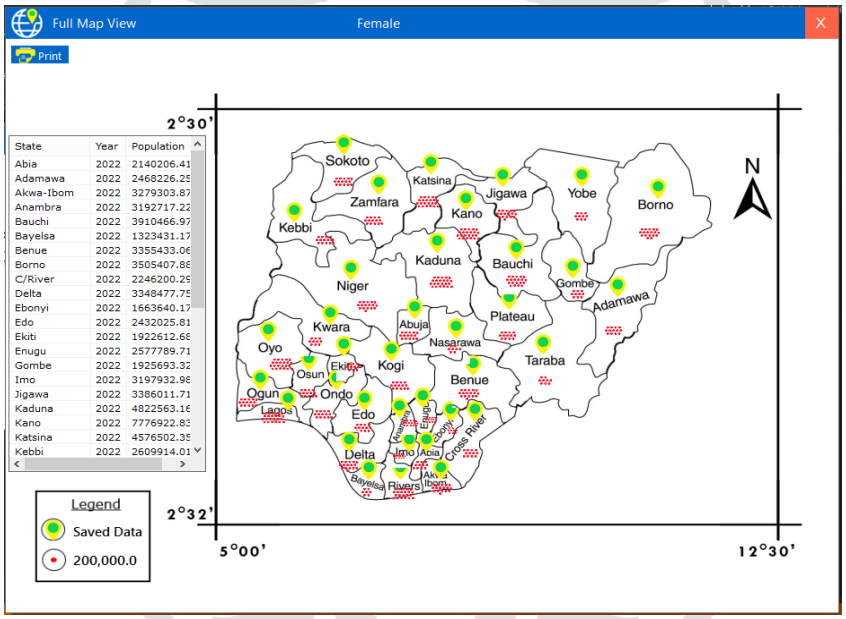

Lagos
Delete
Print
62707.19
S/N Year Population No
1 2011 47067.00
2 2012 48598.00
3 2013 50178.00
4 2014 51810.00
5 2015 53494.00
6 2016 55234.00
7 2017 57030.00
1 2018 58868.14
2 2019 60761.77
3 2020 62707.19
Lagos Divorced data Map
Zoom
Bar Chart
Pie Chart
Zoom
Predicted
80000
60000
40000
20000
0
2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
Prediction Name: Divorced prediction of Lagos

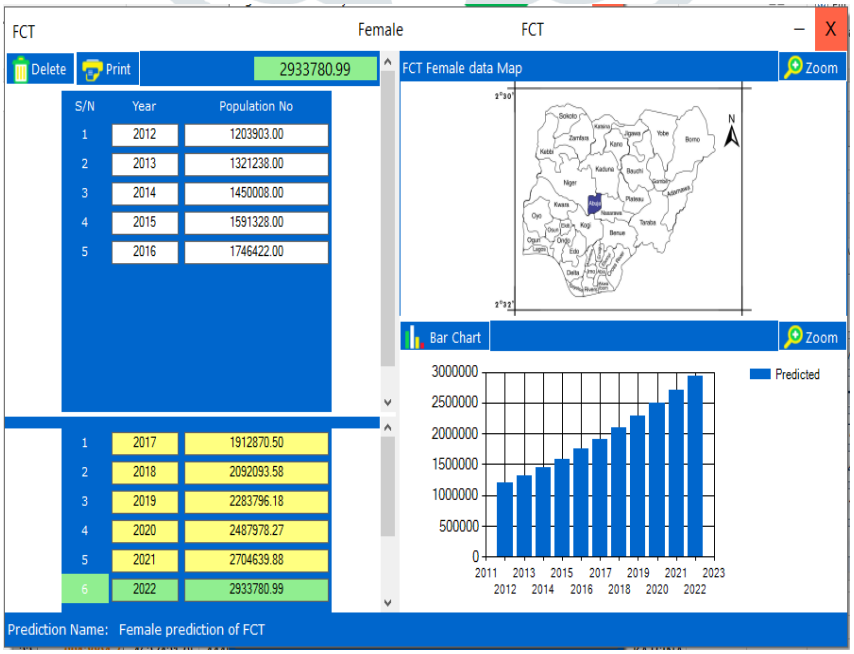

Full Map View
Employed
Print
State Year Population
Abia 2022 135502.38
Adamawa 2022 29249.97
Akwa-Ibom 2022 440485.61
Anambra 2022 228358.25
Bauchi 2022 221373.61
Bayelsa 2022 362378.57
Benue 2022 131320.16
Borno 2022 220307.1
C/River 2022 123814.89
Delta 2022 163652.59
Ebonyi 2022 210926.28
Edo 2022 149093.64
Ekiti 2022 271000.03
Enugu 2022 61877.7
Gombe 2022 333536.59
Imo 2022 40805.39
Jigawa 2022 169156.53
Kaduna 2022 298183.88
Kano 2022 887865.24
Katsina 2022 838741.16
Kebbi 2022 107790.27
Legend
Saved Data
20,000.00
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Zamfara
Katsina
Kebbi
Kano
Jigawa
Yobe
Borno
Kaduna
Bauchi
Niger
Gombe
Adamawa
Plateau
Abuja
Kwara
Nasarawa
Oyo
Kogi
Taraba
Osun
Ekiti
Benue
Ogun
Ondo
Lagos
Edo
Delta
Cross River
Imo
Abia
Rivers

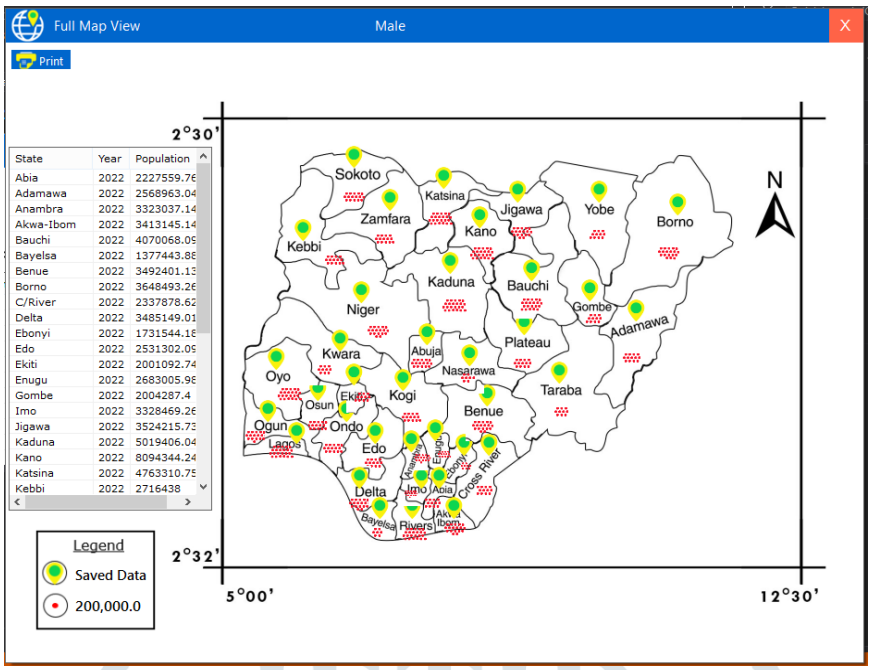

Lagos
Employed
Lagos
Delete
Print
1031802.62
S/N Year Population No
1 2013 406040.00
2 2014 448877.00
3 2015 403783.00
4 2016 374963.00
5 2017 490362.00
1 2018 533772.84
2 2019 623774.43
3 2020 736779.92
4 2021 872789.31
5 2022 1031802.62
Lagos Employed data Map
Zoom
Bar Chart
Zoom
Predicted
1200000
1000000
800000
600000
400000
200000
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023
Prediction Name: Employment rate of Lagos

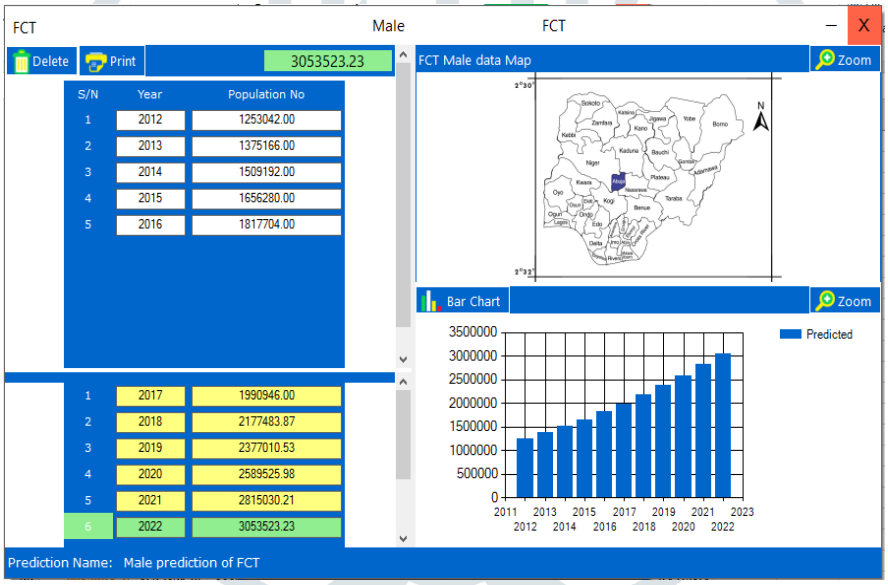

Full Map View
Unemployed
Print
State Year Population
Abia 2022 50401.26
Adamawa 2022 182596.55
Akwa-Ibom 2022 82085.77
Anambra 2022 47313.04
Bauchi 2022 353775.27
Bayelsa 2022 83677.15
Benue 2022 154974.37
Borno 2022 71764.22
C/River 2022 563541.3
Delta 2022 284902.99
Ebonyi 2022 98949.28
Edo 2022 157702.65
Ekiti 2022 58304.44
Enugu 2022 158065.29
Gombe 2022 161289.82
Imo 2022 228822.82
Jigawa 2022 537049.94
Kaduna 2022 528875.52
Kano 2022 262635.87
Katsina 2022 133556.34
Kebbi 2022 113560.34
Legend
Saved Data
20,000.00
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Zamfara
Katsina
Kebbi
Kano
Jigawa
Yobe
Borno
Kaduna
Bauchi
Niger
Gombe
Adamawa
Plateau
Abuja
Kwara
Nasarawa
Oyo
Kogi
Taraba
Osun
Ekiti
Benue
Ogun
Ondo
Lagos
Edo
Delta
Cross River
Imo
Abia
Rivers

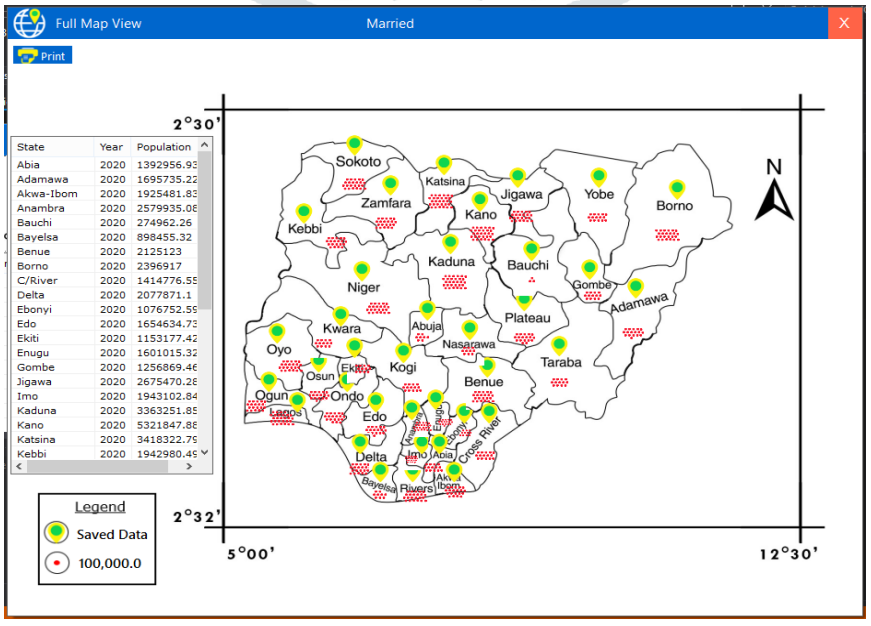

C/River
Unemployed
C/River
Delete
Print
563541.3
C/River Unemployed data Map
Zoom
S/N Year Population No
1 2013 48841.00
2 2014 137805.00
3 2015 22565.00
4 2016 45321.00
5 2017 130434.00
1 2018 163303.34
2 2019 235467.28
3 2020 326228.25
4 2021 435586.26
5 2022 563541.30
Bar Chart
Zoom
Predicted
600000
400000
200000
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023
Prediction Name: Unemployment rate of Cross river

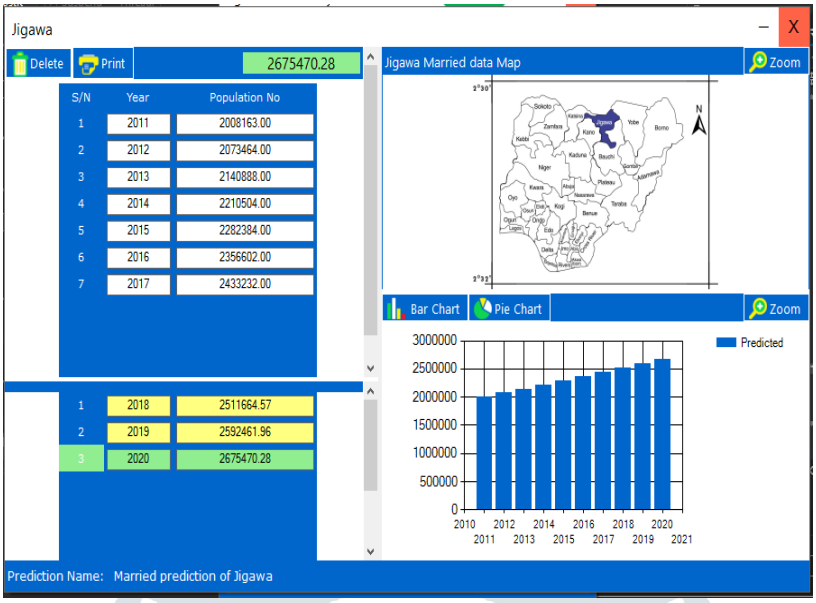

Full Map View
Female
Print
State Year Population
Abia 2022 2140206.41
Adamawa 2022 2468226.25
Akwa-Ibom 2022 3279303.87
Anambra 2022 3192717.22
Bauchi 2022 3910466.97
Bayelsa 2022 1323431.17
Benue 2022 3355433.06
Borno 2022 3505407.88
C/River 2022 2246200.25
Delta 2022 3348477.75
Ebonyi 2022 1663640.17
Edo 2022 2432025.81
Ekiti 2022 1922612.68
Enugu 2022 2577789.71
Gombe 2022 1925693.32
Imo 2022 3197932.98
Jigawa 2022 3386011.71
Kaduna 2022 4822563.16
Kano 2022 7776922.83
Katsina 2022 4576502.35
Kebbi 2022 2609914.01
Legend
Saved Data
200,000.0
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Katsina
Zamfara
Jigawa
Yobe
Borno
Kano
Kebbi
Kaduna
Bauchi
Niger
Gombe
Adamawa
Plateau
Abuja
Kwara
Nasarawa
Oyo
Taraba
Kogi
Ekiti
Osun
Benue
Ogun
Ondo
Lagos
Edo
Delta
Cross River
Imo
Abia
Rivers
Akwa Ibom
Bayelsa

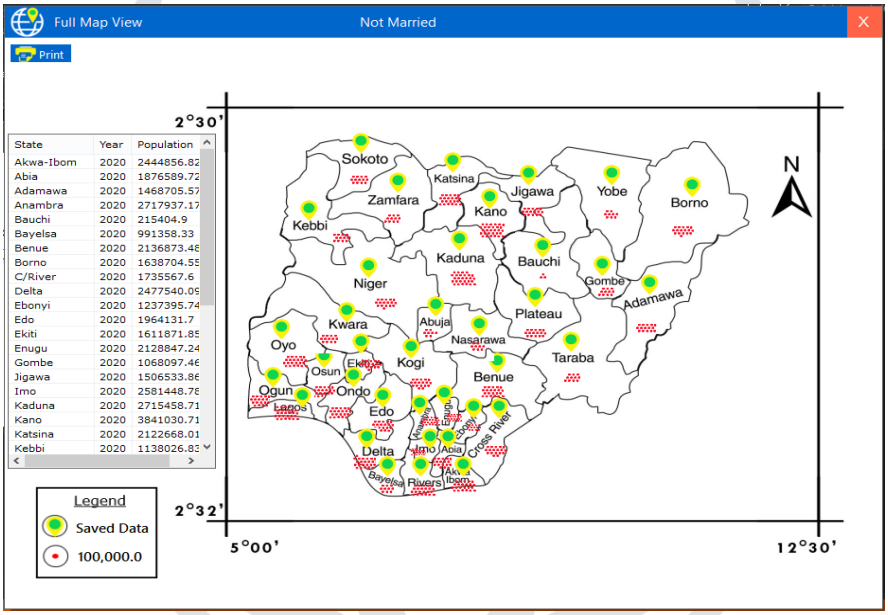

FCT
Female
FCT
Delete
Print
2933780.99
FCT Female data Map
Zoom
S/N Year Population No
1 2012 1203903.00
2 2013 1321238.00
3 2014 1450008.00
4 2015 1591328.00
5 2016 1746422.00
1 2017 1912870.50
2 2018 2092093.58
3 2019 2283796.18
4 2020 2487978.27
5 2021 2704639.88
6 2022 2933780.99
Bar Chart
Zoom
Predicted
3000000
2500000
2000000
1500000
1000000
500000
0
2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023
Prediction Name: Female prediction of FCT

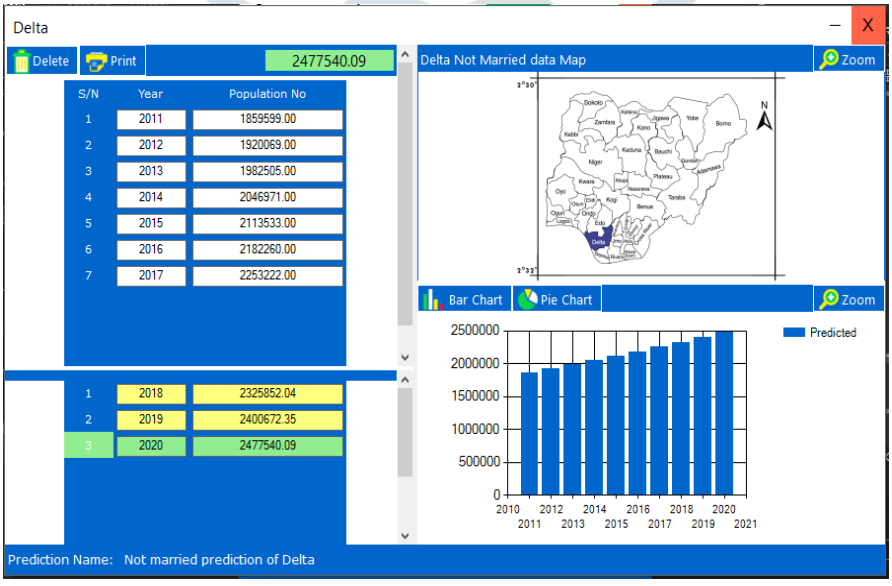

Full Map View
Male
Print
State Year Population
Abia 2022 2227559.76
Adamawa 2022 2568963.04
Anambra 2022 3323037.14
Akwa-Ibom 2022 3413145.14
Bauchi 2022 4070068.09
Bayelsa 2022 1377443.88
Benue 2022 3492401.13
Borno 2022 3648493.26
C/River 2022 2337878.62
Delta 2022 3485149.01
Ebonyi 2022 1731544.18
Edo 2022 2531302.09
Ekiti 2022 2001092.74
Enugu 2022 2683005.98
Gombe 2022 2004287.4
Imo 2022 3328469.26
Jigawa 2022 3524215.73
Kaduna 2022 5019406.04
Kano 2022 8094344.24
Katsina 2022 4763310.75
Kebbi 2022 2716438
Legend
Saved Data
200,000.0
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Katsina
Zamfara
Jigawa
Yobe
Borno
Kebbi
Kano
Kaduna
Bauchi
Niger
Gombe
Adamawa
Plateau
Abuja
Kwara
Nasarawa
Oyo
Kogi
Taraba
Osun
Ekiti
Benue
Ogun
Ondo
Lagos
Edo
Delta
Cross River
Imo
Abia
Akwa Ibom
Rivers
Bayelsa
Ebonyi
Enugu

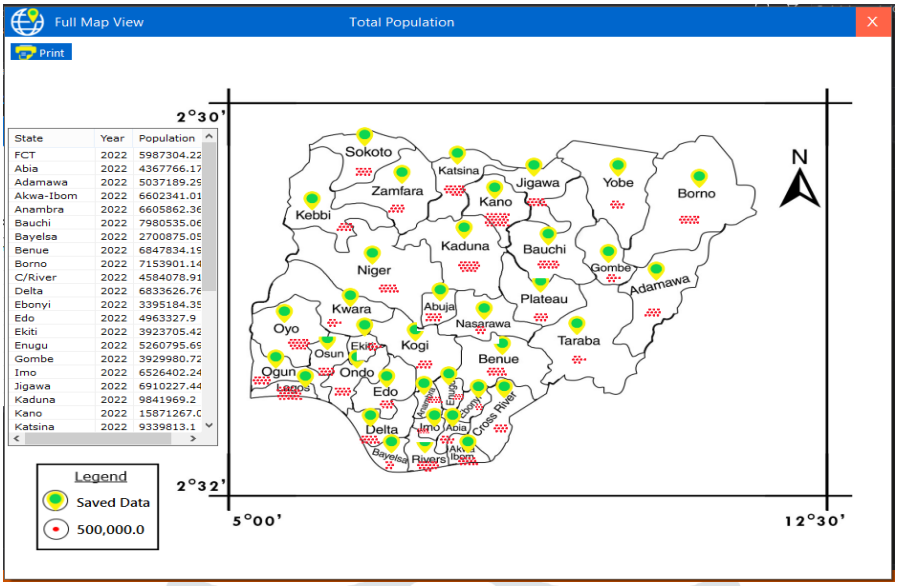

FCT
Male
FCT
Delete
Print
3053523.23
FCT Male data Map
Zoom
S/N Year Population No
1 2012 1253042.00
2 2013 1375166.00
3 2014 1509192.00
4 2015 1656280.00
5 2016 1817704.00
1 2017 1990946.00
2 2018 2177483.87
3 2019 2377010.53
4 2020 2589525.98
5 2021 2815030.21
6 2022 3053523.23
Bar Chart
Zoom
Predicted
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
2011 2013 2015 2017 2019 2021 2023
2012 2014 2016 2018 2020 2022
Prediction Name: Male prediction of FCT

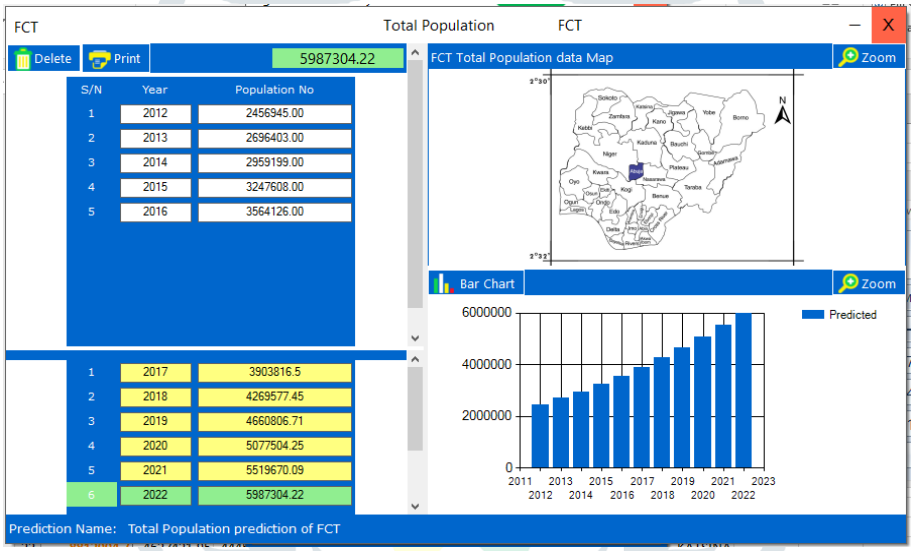

Full Map View
Married
Print
State Year Population
Abia 2020 1392956.93
Adamawa 2020 1695735.22
Akwa-Ibom 2020 1925481.83
Anambra 2020 2579935.08
Bauchi 2020 274962.26
Bayelsa 2020 898455.32
Benue 2020 2125123
Borno 2020 2396917
C/River 2020 1414776.55
Delta 2020 2077871.1
Ebonyi 2020 1076752.59
Edo 2020 1654634.73
Ekiti 2020 1153177.42
Enugu 2020 1601015.32
Gombe 2020 1256869.46
Jigawa 2020 2675470.28
Imo 2020 1943102.84
Kaduna 2020 3363251.85
Kano 2020 5321847.86
Katsina 2020 3418322.79
Kebbi 2020 1942980.45
Legend
Saved Data
100,000.0
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Katsina
Zamfara
Jigawa
Yobe
Borno
Kebbi
Kano
Kaduna
Bauchi
Niger
Gombe
Adamawa
Plateau
Abuja
Kwara
Nasarawa
Oyo
Kogi
Taraba
Osun
Ekiti
Benue
Ogun
Ondo
Lagos
Edo
Delta
Cross River
Imo
Abia
Akwa Ibom
Rivers
Bayelsa
Ebonyi
Enugu


Jigawa
Delete
Print
2675470.28
Jigawa Married data Map
Zoom
S/N Year Population No
1 2011 2008163.00
2 2012 2073464.00
3 2013 2140888.00
4 2014 2210504.00
5 2015 2282384.00
6 2016 2356602.00
7 2017 2433232.00
1 2018 2511664.57
2 2019 2592461.96
3 2020 2675470.28
Bar Chart
Pie Chart
Zoom
Predicted
3000000
2500000
2000000
1500000
1000000
500000
0
2010 2012 2014 2016 2018 2020
2011 2013 2015 2017 2019 2021
Prediction Name: Married prediction of Jigawa


Full Map View
Not Married
Print
State Year Population
Akwa-Ibom 2020 2444856.82
Abia 2020 1876589.72
Adamawa 2020 1468705.57
Anambra 2020 2717937.17
Bauchi 2020 215404.9
Bayelsa 2020 991358.33
Benue 2020 2136873.48
Borno 2020 1638704.55
C/River 2020 1735567.6
Delta 2020 2477540.09
Ebonyi 2020 1237395.74
Edo 2020 1964131.7
Ekiti 2020 1611871.85
Enugu 2020 2128847.24
Gombe 2020 1068097.46
Jigawa 2020 1506533.86
Imo 2020 2581448.78
Kaduna 2020 2715458.71
Kano 2020 3841030.71
Katsina 2020 2122668.01
Kebbi 2020 1138026.83
Legend
Saved Data
100,000.0
2°30'
2°32'
5°00'
12°30'
N
Sokoto
Zamfara
Katsina
Jigawa
Yobe
Borno
Kebbi
Kano
Kaduna
Bauchi
Gombe
Adamawa
Niger
Abuja
Plateau
Kwara
Nasarawa
Taraba
Oyo
Osun
Ekiti
Kogi
Ogun
Ondo
Benue
Lagos
Edo
Delta
Imo
Abia
Cross River
Bayelsa
Rivers
Akwa Ibom


Delta
Delete
Print
2477540.09
Delta Not Married data Map
Zoom
S/N Year Population No
1 2011 1859599.00
2 2012 1920069.00
3 2013 1982505.00
4 2014 2046971.00
5 2015 2113533.00
6 2016 2182260.00
7 2017 2253222.00
1 2018 2325852.04
2 2019 2400672.35
3 2020 2477540.09
Bar Chart
Pie Chart
Zoom
Predicted
2500000
2000000
1500000
1000000
500000
0
2010 2012 2014 2016 2018 2020
2011 2013 2015 2017 2019 2021
Prediction Name: Not married prediction of Delta





4.0 Contributions to Knowledge

National Population Commission has conducted a lot of censuses and surveys in the past. Therefore have a lot of repository of data that have not been mined for further analysis. This new model will assist in mining these repositories of data and bring out certain socio-economic trends, relationships and characteristics.

5.0 Conclusion

The new model is more effective in census analysis. It gives a better prediction than the existing ones not considering outliers and unknown parameters. The potentials of data mining for census results are enormous and it was harnessed in this research. This paper has made a huge difference to what has been in existence and thereby have added to the body of knowledge. Hence a call for the adoption of the new model is a right step in the right direction. Over estimation and under estimation which is prone to existing systems has been taken care of by adopting a narrow range smoothening constant. An appreciable level of computer knowledge to enable users successfully make prediction with the new system is required. Set of inputs which are known as observations (data input) are required by the system before it can make any prediction. Additionally, the number of observations cannot be less than five (5). These inputs can be entered via the keyboard by the user. Another significance of this study is that predicted demographic information is shown on the map.

REFERENCES

[1] Census Enumerator's Manual (2006). National Population Commission Abuja: Nigeria.

[2] Gaughan, A. E., Stevens, F. R., Linard, C., Jia, P., & Tatem, A. J. (2013). High resolution population distribution maps for Southeast Asia in 2010 and 2015. PloS one, 8(2), e55882.

[3] Mohamed, M.G (2009).Scientific Data mining and Knowledge Discovery: Principles and Foundations. Springer: New York.

[4] Micheal, B and Gordon, L (2012).Data mining Techniques.3rd Edition. Wiley: United States.

[5] Kurt, T. (2012). Introduction to Data Mining, Wiley: Asia.

[6] Acm, S. (2006).Machine Learning Algorithm and Data-mining. Cambridge University Press. London.

[7] Ian. H.W , Eibe, F and Mark, A, H (2011). Data mining Practical Machine Learning Tools and Techniques.3rd Edition. Morgan Kaufmann: United States.

[8] Micheal, R. B and David .J. H (2007).Intelligent Data Analysis.2nd Edition. Springer: New York.
[9] Witten, Ian, H and Frank, E. (2011). Data mining: Practical Machine Learning Tools and Techniques.3rd Edition. CC Press: Elsevier.
[10] Kenneth E. F. and Margaret, L.(2015).The Geographer's Craft Project, Department of Geography. The University of Colorado: Boulder.
[11] GoodChild, M. F (2010).Twenty years of progress: GIScience: Journal of spatial Information science.Vol 2, pp 10-12.
[12] Ramezankhani. A, Pournik, O, Shahrabi, A. Khalili, D. Azizi, F and Hadaegh, F. (2014).Applying decision tree for identification of a low risk population for type 2 Diabetes. Tehran lipid and Glucose study. International Journal on Diabetes Research and Clinical Practice. Elsevier Ireland Ltd. Vol 105, pg 391-398.
[13] Manoj, K and Madhu, A. (2012).International Journal of Studies in Business and Economics.Pg 81- 94.
[14] Md. M and Syed, S. H. (2012). Population Forecasts for Bangladesh, Using a Bayesian Methodology, Journal of Health, Population and Nutrition (JHPN) , 30(4): 456– 463.
[15] David, A. S, Alan, S and Bob, S. (2010). Forecasting the Population of Census tracts by age,and sex. Springer Population Research and Policy Review, Volume 29, Issue 1, pp 47-63.
[16] Kayode, J and Jimoh, M (2011).Used regression analysis with times series to periodically forecast stock. Advances in Information Mining Volume 4,Issue 1, pp57-66
[17] Friedrich, H and Weixin L. (2013).ADULT AND YOUTH LITERACY- National, Regional and Global trends, 1985-2015.UIS Information Paper.
[18] Saigal, S. and Mehrotra D. (2012). Performance Comparison of time series data using predictive data mining techniques. Advances in Information Mining Volume 4, Issue 1,pp.-57-66.
[19] Rajasekhar,N and Karth T. V. R. (2014).Hybrid SVM Data Mining Techniques for Weather Data Analysis of Krishna District of Andhra Region. International Journal of Research in Computer and Communication Technology, Vol 3, Issue 7.