

Analytical Modelling of Human Population

Instructor: Dr. Sunil Pratap Singh

Group 3: Abee Nelson - 18007
Aiswarya Lakshmi - 18017
Derin Wilson -18082
Harsha Sudhakaran - 18100
Madhav Sharma K N -18131



Indian Institute of Science Education and Research, Bhopal

Acknowledgement

We would like to express our special gratitude to our professor Dr Sunil Pratap Singh, for providing us with this golden opportunity to do a wonderful project on Human Population Modeling and for his constant guidance and support throughout the project. We would also like to thank our family and friends for their constant support in these challenging times.

Abstract

In this project, we tried to learn about human population modelling and how different factors affect the population growth. We concentrated on exponential modelling of the human population and used it to predict the future population of three different countries taken entirely from different economical backgrounds. We have also used logistic growth model to model human population. To predict the future population we have made use of Euler and RK4 method. We selected India from developing, USA from developed and Madagascar from under developed category. Apart from all this we also used birth and death rate to model the population of the countries taken into consideration. We were able to successfully predict the population of the countries and conclude many things. We also studied the fertility growth rate of these three countries and also predicted the fertility growth rate for next 10 years using polynomial regression.

Contents

Abstract	3
1 Introduction	6
1.1 What is a Population Model ?	6
1.2 Why is it important ?	6
2 Different Types of Population Models	7
2.1 Exponential Growth Model	7
2.2 Logistic Growth Model	8
3 Human Population Growth	9
3.1 Factors affecting Human Population Growth	9
3.1.1 Birth and Death Rate	9
3.1.2 Migration	9
3.1.3 Fertility Rate	10
4 Numerical Methods Used	11
4.1 Solutions to system of Linear equations	11
4.1.1 Gauss-Elimination	11
4.2 Curve-fitting	13
4.2.1 Linear Regression-Least Squares Regression	13
4.2.2 Polynomial Regression - Least Squares Method	14
4.3 Solving Differential Equations	15
4.3.1 Euler's Method	15
4.3.2 RK-4 Method	15
5 Case Study	17
5.1 Analysing population of different countries based on Development	17
5.1.1 Population Graphs of different countries	17
5.1.2 Exponential Curve Fitting	18
5.2 Predicting population trend for next 10 years	19
5.2.1 Euler Method	19
5.2.2 RK4 Method	19
5.2.3 Observations	20
5.3 Logistic Growth Model	21
5.3.1 Using Euler	21
5.3.2 Using RK4	21
5.3.3 Observations	22
5.4 Modelling the population growth using Birth and Death Rate . .	23
5.4.1 Predicting population trend for next 10 years using the above model	24
5.5 Fertility growth rate	25
5.5.1 Predicting next 10 year fertility growth	26
6 Conclusion	27

7	Future Work	27
8	Codes	29
8.1	Curve fitting for Population (Exponential model)	29
8.2	Euler's Method and RK4 for Population Model (Exponential model)	30
8.3	Logistic Growth Model	33
8.4	Population Modelling using Birth Rate and Death Rate	36
8.5	Prediction using Birth Rate and Death Rate	38
8.6	Curve fitting for Fertility Rate (Exponential model)	39
8.7	Polynomial Regression for Fertility Rate	40
8.8	Predicting Fertility Rate with Polynomial Regression	42
9	References	43

1 Introduction

1.1 What is a Population Model ?

A model is a simplification of a real system that is used to provide some insights into the system. Population models are often (although not always) formulated as a set of rules or assumptions, expressed as mathematical equations, that describe change in population size over time as a consequence of survival and reproduction, and they include external factors that affect these characteristics. A model simplifies a system, retaining essential components while eliminating parts that are not of interest.

Ecology has a rich history of using models to gain insights into populations, often borrowing both model structures and analysis methods from demographers and engineers. Much of the development of the models has been a consequence of mathematicians and physicists seeing simple analogies between their models and patterns in natural systems. Consequently, one major application of ecological modeling has been to emphasize the analysis of dynamics of often complex models to provide insights into theoretical aspects of ecology.

1.2 Why is it important ?

Population modelling help scientists in making better predictions about future population size and growth rates. This is very essential for answering questions in areas such as biodiversity conservation (how quickly a species population is declining or whether it is at the verge of extinction) and human population growth (how fast will the human population grow and how to sustain the required resources and biodiversity).

Population models are also used for determining maximum harvests for agriculturists, environmental conservation, to understand the spread of parasites, viruses, and diseases. Human Population modelling is very important because it helps country's policy makers to make efficient policies and decision for there people.

2 Different Types of Population Models

2.1 Exponential Growth Model

A biological population with plenty of food, space to grow, and no threat from predators, tends to grow at a rate that is proportional to the population - that is, in each unit of time, a certain percentage of the individuals produce new individuals. If reproduction takes place more or less continuously, then this growth rate is represented by

$$\frac{dN}{dt} = rN,$$

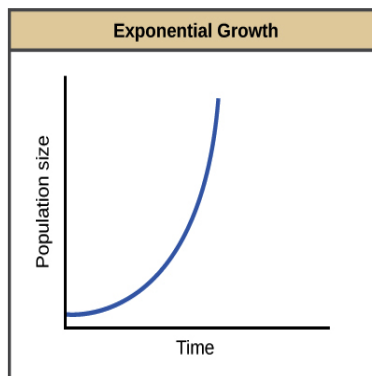
where N is the population as a function of time t , and r is the proportionality constant. We know that all solutions of this natural-growth equation have the form

$$N(t) = N_0 e^{rt},$$

where N_0 is the population at time $t = 0$. In short, unconstrained natural growth is exponential growth.

The best example of exponential growth is seen in bacteria. Bacteria are prokaryotes that reproduce by prokaryotic fission. This division takes about an hour for many bacterial species. If 1000 bacteria are placed in a large flask with an unlimited supply of nutrients, after an hour, there is one round of division and each organism divides, resulting in 2000 organisms - an increase of 1000. In another hour, each of the 2000 organisms will double, producing 4000, an increase of 2000 organisms. After the third hour, there should be 8000 bacteria in the flask, an increase of 4000 organisms.

Another example of exponential growth is that of human population. The world's population is presently growing at an exponential rate and hence the time required for the population to grow is becoming shorter.



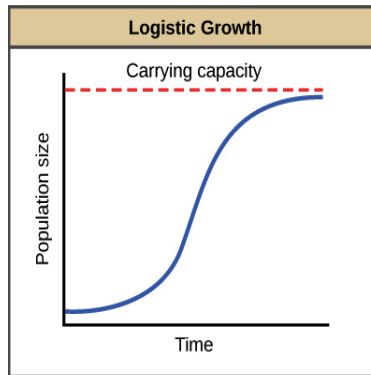
2.2 Logistic Growth Model

Exponential growth is possible only when infinite natural resources are available, this is not the case in the real world. In real world, there is limitation of resources and competition for existence among individuals. To model the reality of limited resources, ecologists developed the logistic growth model. Logistic model was first developed by Belgian mathematician Pierre Verhulst (1838) who suggested that the rate of population increase may be limited, i.e., it may depend on population density:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

where the constant r defines the growth rate and K is the carrying capacity. Population growth rate declines with population, N , and reaches 0 when $N = K$. Parameter K is the upper limit of population growth and also called as carrying capacity. It is usually interpreted as the amount of resources expressed in the number of organisms that can be supported by these resources. If population numbers exceed K , then population growth rate becomes negative and population numbers decline.

Examples of organisms showing logistic growth include yeast, a microscopic fungus, that exhibits the classical S-shaped population size versus time curve when grown in a test tube. Its growth levels off as the population depletes the nutrients that are necessary for its growth.



3 Human Population Growth

World population has been growing since the end of the Black Death, around the year 1350. Improved agricultural productivity, sanitation and medical advancements that reduced mortality are the causes of such a growth. The world population is currently experiencing an exponential growth even though human reproduction is far below its biotic potential. A consequence of exponential human population growth is that the time it takes to add a particular number of humans to the earth is becoming shorter. 123 years were required to add 1 billion humans in 1930 but it took only 24 years to add 2 billion people between 1975 and 1999.

Global human population growth amounts to around 83 million annually or 1.1% per year. The global population has grown from 1 billion in 1800 to 7.9 billion in 2020. In accordance with present growth rate, the total population is estimated to be around 8.6 billion by mid 2030, 9.8 billion by mid 2050 and 11.2 billion by 2100. Humans are unique in their ability to alter their environment with the conscious purpose of increasing its carrying capacity. This ability is a major factor responsible for human population growth and a way of overcoming density-dependent growth regulation. Much of this ability is related to human intelligence, society, and communication. Other factors in human population growth are migration and public health. Public health, sanitation, and the use of antibiotics and vaccines have decreased the ability of infectious diseases to decrease human population growth.

3.1 Factors affecting Human Population Growth

3.1.1 Birth and Death Rate

Birth and death rates are crucial factors in determining population growth. Birth rate is typically indicated as the annual births per thousand total population. Similarly, the death rate is the annual deaths per thousand total population. The average global birth rate was 18.5 births per thousand people in 2016 and the death rate was 7.8 per thousand people.

3.1.2 Migration

Migration plays an important role in population growth. It changes not only the population size but also the composition of urban and rural population in terms of age and sex composition. The net migration rate is the difference between the number of immigrants (people coming into an area) and the number of emigrants (people leaving an area) throughout the year. The most important force behind European population change is international migration.

3.1.3 Fertility Rate

Fertility rate, average number of children born to women during their reproductive years. For the population in a given area to remain stable, an overall total fertility rate of 2.1 is needed, assuming no immigration or emigration occurs.

It is important to distinguish this from birth rates, which are defined as the number of live births per 1,000 women in the total population. The single most important factor in population growth is the total fertility rate (TFR). If, on average, women give birth to 2.1 children and these children survive to the age of 15, any given woman will have replaced herself and her partner upon death. A TFR of 2.1 is known as the replacement rate. Generally speaking, when the TFR is greater than 2.1, the population in a given area will increase, and when it is less than 2.1, the population in a given area will eventually decrease, though it may take some time because factors such as age structure, emigration, or immigration must be considered. More specifically, if there are numerous women of childbearing age and a relatively small number of older individuals within a given society, the death rate will be low, so even though the TFR is below the replacement rate, the population may remain stable or even increase slightly. This trend cannot last indefinitely but could persist for decades.

Tracking fertility rates allows for more efficient and beneficial planning and resource allocation within a particular region. If a country experiences unusually high sustained fertility rates, it may need to build additional schools or expand access to afford child care. This occurred in the United States during the post-World War II baby boom era. During this period, the TFR peaked at about 3.8, roughly twice the average 21st century rate in the United States. The unusually high number of children born during this period left communities unprepared. Conversely, sustained low fertility rates may signify a rapidly aging population, which may place an undue burden on the economy through increasing health care and social security costs.

4 Numerical Methods Used

4.1 Solutions to system of Linear equations

4.1.1 Gauss-Elimination

Gauss elimination, also known as row reduction is an algorithm for solving system of linear equations. It consists of a sequence of operations, namely forward elimination and back substitution performed on the corresponding matrix of coefficients. The approach is to solve a set of n equations as shown below:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \quad (1)$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2 \quad (2)$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n = b_n \quad (3)$$

Forward elimination of unknowns:

Here the set of n equations is reduced to an upper triangular system. The initial step is to eliminate the first unknown, x_1 , from the second through nth equations. This is done by multiplying equation (1) by $\frac{a_{21}}{a_{11}}$.

$$a_{21}x_1 + \frac{a_{21}}{a_{11}}a_{12}x_2 + \dots + \frac{a_{21}}{a_{11}}a_{1n}x_n = \frac{a_{21}}{a_{11}}b_1 \quad (4)$$

Now this equation is subtracted from equation (2) to give

$$(a_{22} - \frac{a_{21}}{a_{11}}a_{12})x_2 + \dots + (a_{2n} - \frac{a_{21}}{a_{11}}a_{1n})x_n = (b_2 - \frac{a_{21}}{a_{11}}b_1) \quad (5)$$

or

$$a'_{22}x_2 + \dots + a'_{2n}x_n = b'_2 \quad (6)$$

Repeating the procedure for the remaining equations results in the following modified system:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \quad (7)$$

$$a'_{22}x_2 + a'_{23}x_3 + \dots + a'_{2n}x_n = b'_2 \quad (8)$$

$$a'_{32}x_2 + a'_{33}x_3 + \dots + a'_{3n}x_n = b'_3 \quad (9)$$

$$\vdots$$

$$a'_{n2}x_2 + a'_{n3}x_3 + \dots + a'_{nn}x_n = b'_n \quad (10)$$

Now to eliminate the second unknown, x_2 , we multiply equation (8) by $\frac{a'_{32}}{a'_{22}}$ and subtract the result from equation (9). Performing a similar elimination for the remaining equations we get

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \quad (11)$$

$$a'_{22}x_2 + a'_{23}x_3 + \dots + a'_{2n}x_n = b'_2 \quad (12)$$

$$a''_{33}x_3 + \dots + a''_{3n}x_n = b''_3 \quad (13)$$

\vdots

$$a''_{n3}x_3 + \dots + a''_{nn}x_n = b''_n \quad (14)$$

where the double prime indicates that the coefficients have been modified twice. The procedure can be continued using the remaining pivot equations. The final manipulation in the sequence is to use the $(n-1)^{th}$ equation to eliminate the x_{n-1} term from the n^{th} equation.

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \quad (15)$$

$$a'_{22}x_2 + a'_{23}x_3 + \dots + a'_{2n}x_n = b'_2 \quad (16)$$

$$a''_{33}x_3 + \dots + a''_{3n}x_n = b''_3 \quad (17)$$

\vdots

$$a^{n-1}_{nn}x_n = b^{n-1}_n \quad (18)$$

Back substitution:

Equation (18) can be solved for x_n as

$$x_n = \frac{b^{n-1}_n}{a^{n-1}_{nn}} \quad (19)$$

This result can be back substituted into the $(n-1)^{th}$ equation to solve for x_{n-1} . The procedure which is repeated to evaluate the remaining x_i 's can be represented by the following formula:

$$x_i = \frac{b^{i-1}_i - \sum_{j=i+1}^n a^{i-1}_{ij}x_j}{a^{i-1}_{ii}} \quad (20)$$

4.2 Curve-fitting

Curve fitting is a technique used in numerical methods to find an appropriate mathematical model that expresses the relationship between a dependent variable and an independent variable and then estimating the values of its parameters with regression techniques.

4.2.1 Linear Regression-Least Squares Regression

This the most common technique used for curve fitting. In this an approximate function is derived that fits the shape or general trend of the data without necessarily matching the individual points. One way to do this is to derive a curve that minimizes the discrepancy between the data points and the curve. A technique for accomplishing this objective is called least squares regression. For a form of linear equation i.e.

$$y = a_0 + a_1x \quad (21)$$

we use linear regression where we will minimise the squares of errors/discrepancies.

$$S_r = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - a_0 - a_1x_i)^2 \quad (22)$$

So by minimising S_r with respect to the variables a_0 and a_1 we will get the equation as follows:

$$\frac{\partial S_r}{\partial a_0} = -2 \sum (y_i - a_0 - a_1x_i) \quad \frac{\partial S_r}{\partial a_1} = -2 \sum ((y_i - a_0 - a_1x_i) x_i) \quad (23)$$

Now equating both equations to zero we get :

$$-2 \sum (y_i - a_0 - a_1x_i) = 0 \quad (24)$$

$$-2 \sum ((y_i - a_0 - a_1x_i) x_i) = 0 \quad (25)$$

And now solving above two equations we get :

$$a_1 = \frac{n \sum (x_i y_i) - \sum x_i \sum y_i}{n \sum (x_i)^2 - (\sum x_i)^2} \quad (26)$$

and

$$a_0 = \bar{y} - a_1 \bar{x} \quad (27)$$

where \bar{x} and \bar{y} are means of x_i and y_i

Fitting An Exponential Function Using Linear Regression

Let ,

$$y = \alpha e^{\lambda x} \quad (28)$$

We can linearize this function by taking natural logarithm on both sides

$$\ln(y) = \ln(\alpha) + \lambda x \quad (29)$$

So by plotting the graph of $\ln(y)$ vs x we can obtain the slope of the graph, from which we can determine the value of λ and from intercept we can deduce the value of α .

4.2.2 Polynomial Regression - Least Squares Method

In polynomial regression we try fit a polynomial of appropriate degree to given the data. The general form of an n degree polynomial is

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots + a_mx^m$$

For this case the sum of the residual is

$$S_r = \sum_{i=0}^n (y_i - (a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots + a_mx^m))^2$$

Similar to linear regression(previous section) we take derivative with respect to each of the coefficient and we get the following set of equations:

$$\frac{\partial S_r}{\partial a_0} = -2 \sum (y_i - (a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots + a_mx^m)) \quad (30)$$

$$\frac{\partial S_r}{\partial a_1} = -2 \sum x_i (y_i - (a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots + a_mx^m)) \quad (31)$$

$$\frac{\partial S_r}{\partial a_2} = -2 \sum x_i^2 (y_i - (a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots + a_mx^m)) \quad (32)$$

$$\frac{\partial S_r}{\partial a_3} = -2 \sum x_i^3 (y_i - (a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots + a_mx^m)) \quad (33)$$

\vdots

$$\frac{\partial S_r}{\partial a_n} = -2 \sum x_i^n (y_i - (a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots + a_mx^m)) \quad (34)$$

These equations are set to zero and rearranged to develop the following set of normal equations:

$$\begin{aligned} ma_0 + \left(\sum x_i\right) a_1 + \left(\sum x_i^2\right) a_2 + \left(\sum x_i^3\right) a_3 + \dots + \left(\sum x_i^m\right) a_m &= \sum y_i \\ \left(\sum x_i\right) a_0 + \left(\sum x_i^2\right) a_1 + \left(\sum x_i^3\right) a_2 + \left(\sum x_i^4\right) a_3 + \dots + \left(\sum x_i^{m+1}\right) a_m &= \sum x_i y_i \\ \left(\sum x_i^2\right) a_0 + \left(\sum x_i^3\right) a_1 + \left(\sum x_i^4\right) a_2 + \left(\sum x_i^5\right) a_3 + \dots + \left(\sum x_i^{m+2}\right) a_m &= \sum x_i^2 y_i \end{aligned}$$

\vdots

$$\left(\sum x_i^m\right) a_0 + \left(\sum x_i^{m+1}\right) a_1 + \left(\sum x_i^{m+2}\right) a_2 + \left(\sum x_i^{m+3}\right) a_3 + \dots + \left(\sum x_i^{m+m}\right) a_n = \sum x_i^m y_i$$

We can solve the above set of equation using Gauss elimination, and find out the values of the coefficients $a_0, a_1, a_2, \dots, a_m$.

The standard error for this method is

$$St_{error} = \sqrt{\frac{S_r}{n - (m + 1)}} \quad (35)$$

here n is the number of data points and m is the degree of the polynomial to be fitted.

4.3 Solving Differential Equations

4.3.1 Euler's Method

The Euler Method (also called point-slope method) is a first order numerical procedure for solving ordinary differential equations with a given initial value. It is the most basic explicit method for numerical integration of ordinary differential equations and the simplest Runge-Kutta Method. It is a first-order method, which means the local error (the error per step) is proportional to the square of step size and the global error (error at a given time) is proportional to step size.

$$\frac{dy}{dx} = f(x, y) \quad (36)$$

$$y_{i+1} = y_i + f(x_i, y_i) \cdot h \quad (37)$$

where $f(x_i, y_i)$ is the slope at x_i and h is the step size.

4.3.2 RK-4 Method

In numerical analysis, the Runge-Kutta methods are a family of implicit and explicit iterative methods which are used for finding the approximate solutions of ordinary differential equations. The most widely used member from the Runge-Kutta family is 'RK4' or 'Classic Runge-Kutta Method'. The general form of RK-4 method is given by

$$y(x + h) \simeq y(x) + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) \quad (38)$$

where

$$k_1 = f(x, y)h \quad (39)$$

$$k_2 = f(x + h/2, y + k_1/2)h \quad (40)$$

$$k_3 = f(x + h/2, y + k_2/2)h \quad (41)$$

$$k_4 = f(x + h, y + k_3)h \quad (42)$$

where h is the step-size and k_1 is the slope at the beginning of the interval, k_2 is the slope at the midpoint of the interval (using y and k_1), k_3 is again the slope at the midpoint (using y and k_2), k_4 is the slope at the end of the interval (using y and k_3). The RK-4 method is basically a fourth-order method, meaning that the local truncation error is on the order of $O(h^5)$, while the total accumulated error is of the order of $O(h^4)$.

5 Case Study

5.1 Analysing population of different countries based on Development

For this purpose, we took 3 different countries each from Developed, Developing and Under-Developed category. India represented Developing category, United States Of America represented Developed category and Madagascar (an island nation) represented under developed category. We took the population data from the World Bank and used different curve fitting methods to find an appropriate mathematical model. We also plotted the population data to get the general trend.

5.1.1 Population Graphs of different countries

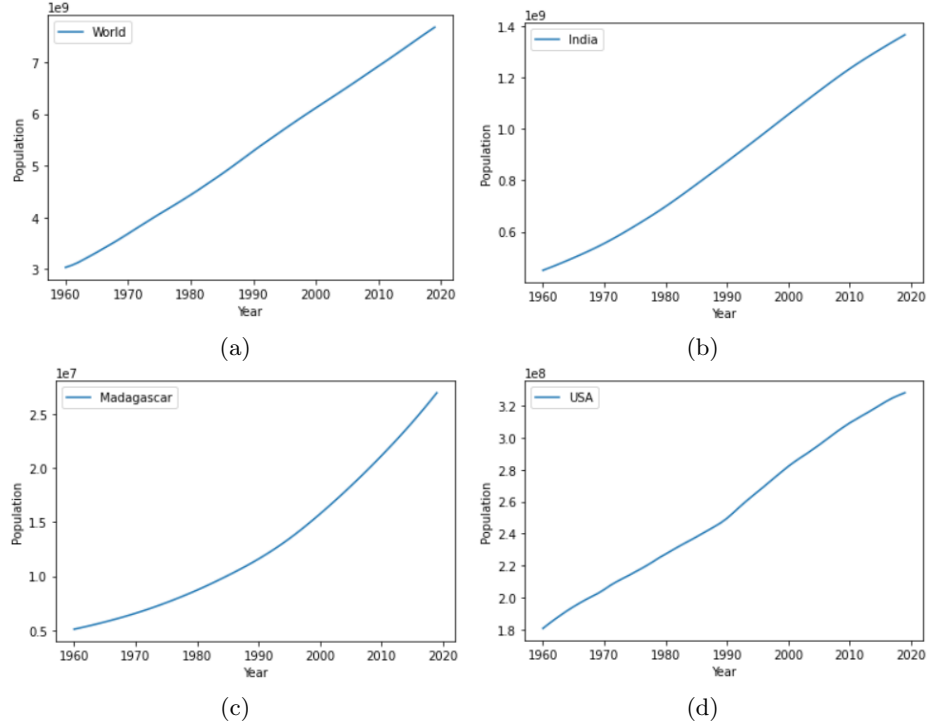


Figure 1: Population versus year graph for (a) World, (b) India, (c) Madagascar and (d) USA

From the above graphs, we can infer that all the graph follows an exponential increase especially that of Madagascar, and so in the next section we are going to try to fit an exponential curve to the given data and observe the trend.

5.1.2 Exponential Curve Fitting

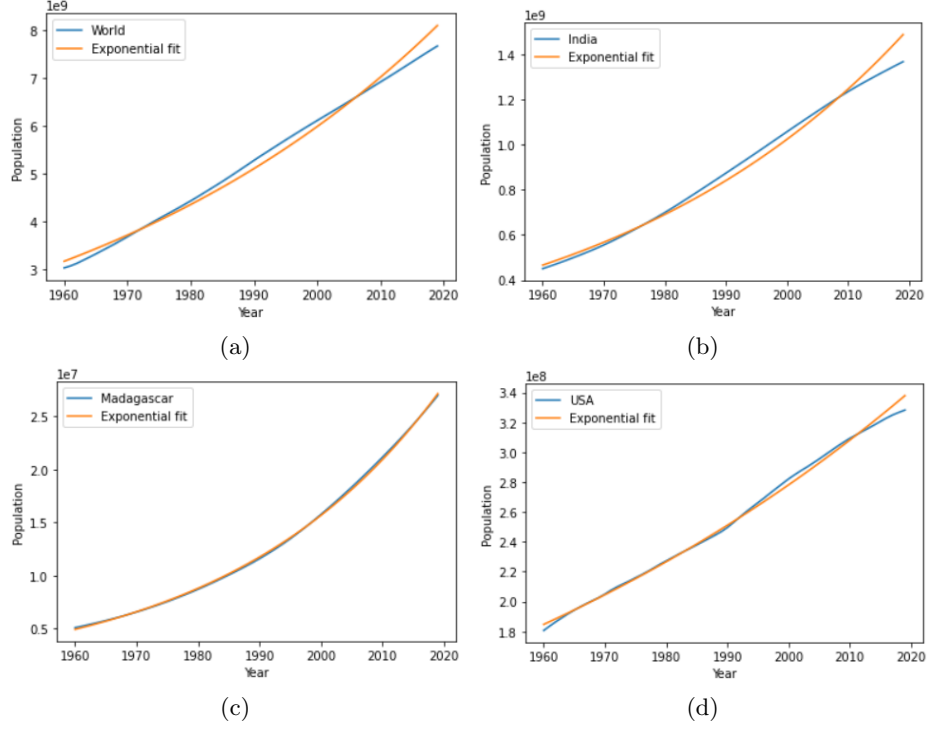


Figure 2: Population versus year graph for (a) World, (b) India, (c) Madagascar and (d) USA

In this section we fitted an exponential growth model of the form

$$\frac{dp}{dt} = r \cdot p$$

$$\implies p(t) = p_0 e^{r \cdot t}$$

By curve fitting we can find out the value of r for each country and then use it in Euler's and RK4 method to predict the future trend.

$$r_{World} = 0.015906861837129742$$

$$r_{India} = 0.01965563394203113$$

$$r_{Madagascar} = 0.02886290941170391$$

$$r_{USA} = 0.010234409905400798$$

5.2 Predicting population trend for next 10 years

5.2.1 Euler Method

In this section we used Euler's algorithm to solve the differential equation and predict the next 10 years population trend for the three countries and world.

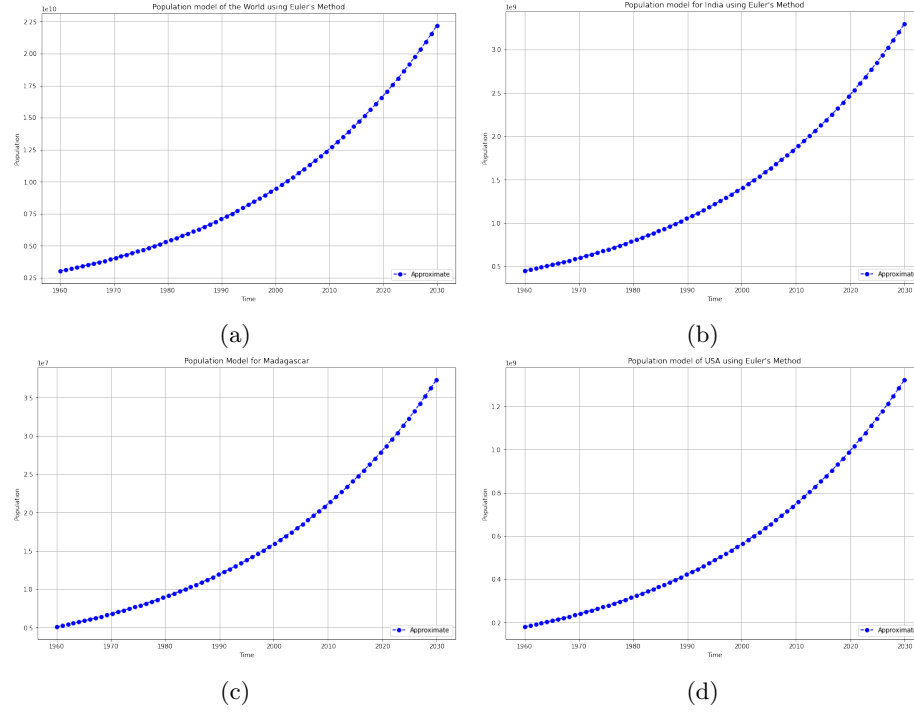


Figure 3: Population versus year graph using Euler Method for (a) World, (b) India, (c) Madagascar and (d) USA

5.2.2 RK4 Method

Under this section we used Runge-Kutta's algorithm to solve the exponential population growth model's differential equation and predict next 10 years population trend for the three countries and the world.

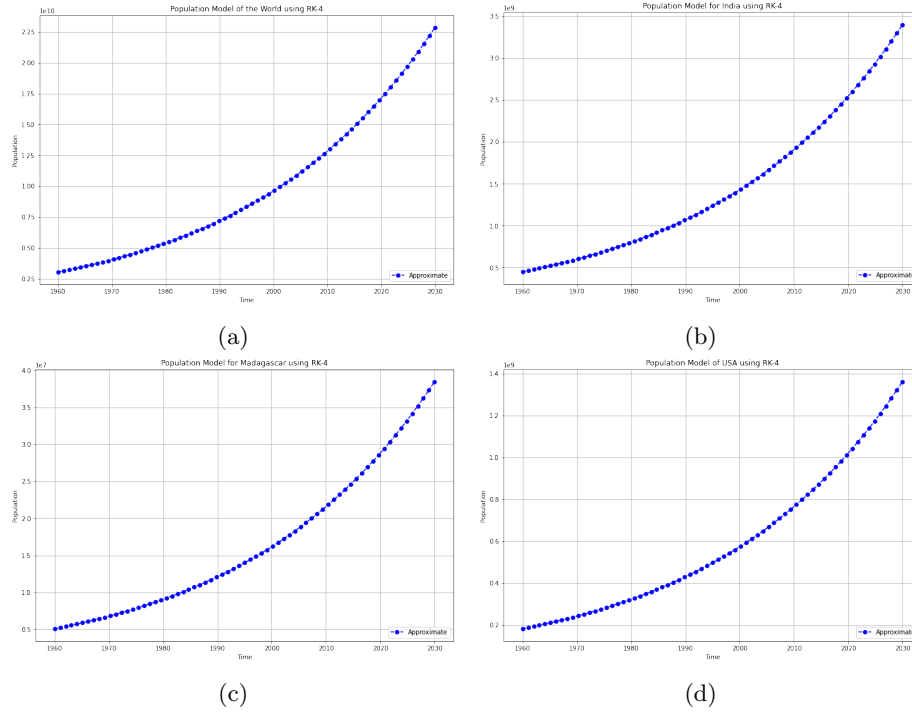


Figure 4: Population versus year graph using RK4 Method for (a) World, (b) India, (c) Madagascar and (d) USA

5.2.3 Observations

Using both the RK4 and Euler method to plot differential equation, our exponential model of population predicts that population of India will touch approximately 3.5 billion, Madagascar will touch 36 million, USA will touch 1.4 billion and whole world population will touch 27.5 billion. But the prediction has not taken into factors like fertility rate, mortality rate etc. of a country. Hence there is huge margin of error in this prediction. In the next section we will see fertility rate of different countries and try to make a better qualitative prediction.

5.3 Logistic Growth Model

To logistically model population of a country we need the carrying capacity (K) and population growth rate (r) of a country. For each country and for world we have assumed the carrying capacity K to be double the maximum population of the country. The r value for each country has been decided by trial and error method, also by using prior knowledge of population growth rate from the previous section.

5.3.1 Using Euler

Under this section we are going to use Euler's algorithm to solve the logistic growth differential equation and obtain the parameters r and K of the equation for observing the population trend for over 200 years and to find the flattening of the curve.

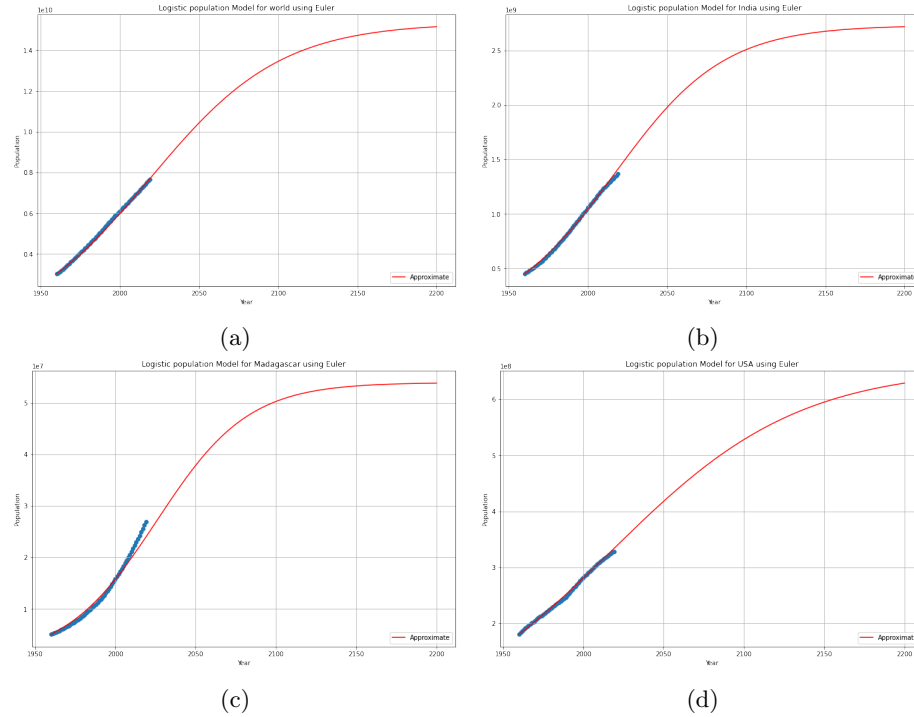


Figure 5: Population growth via logistic model versus year graph using Euler Method for (a) World, (b) India, (c) Madagascar and (d) USA

5.3.2 Using RK4

Under this section we are going to use Runge-Kutta's algorithm to solve the logistic growth differential equation and obtain the parameters r and K of the

equation for observing the population trend for over 200 years and to find the flattening of the curve.

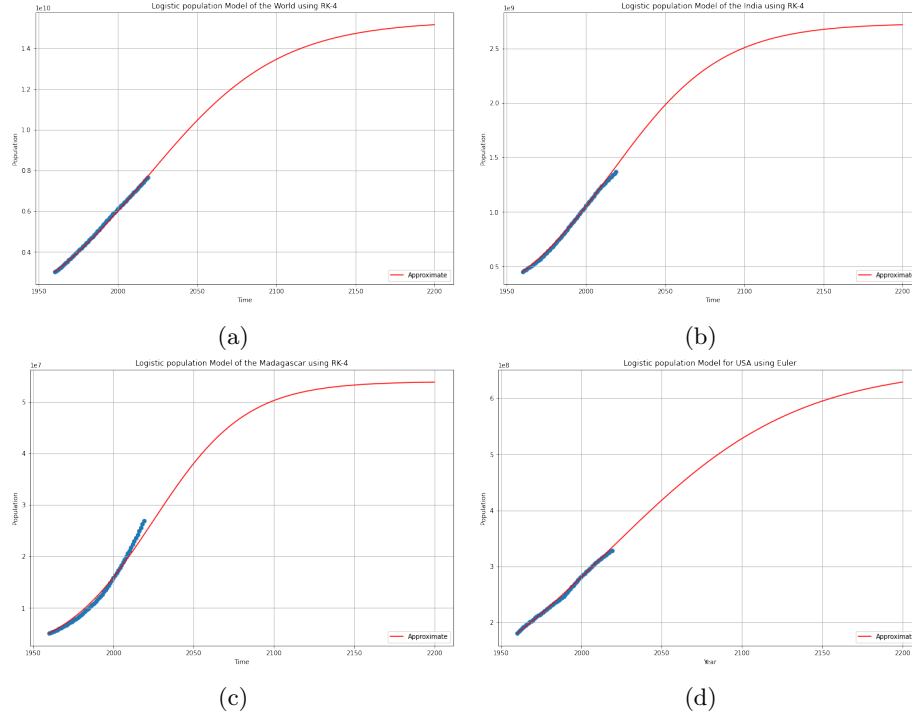


Figure 6: Population growth via logistic model versus year graph using RK4 Method for (a) World, (b) India, (c) Madagascar and (d) USA

5.3.3 Observations

As expected from Logistic growth model we got an S shaped graph. For most of the countries, the graph flattens after the year 2200 because the population reaches the carrying capacity (K) of the country.

$$r_{World} = 0.024$$

$$r_{India} = 0.02879$$

$$r_{Madagascar} = 0.03478$$

$$r_{USA} = 0.017$$

5.4 Modelling the population growth using Birth and Death Rate

In this section, we take the average birth rate and death rate of the countries from 1960 to 2018 to model the population growth of the country using the formula

$$\frac{dp}{dt} = B - D$$

where B = average birth rate and D = average death rate

We use RK4 method to plot this differential equation once and calculate B and D for each country.

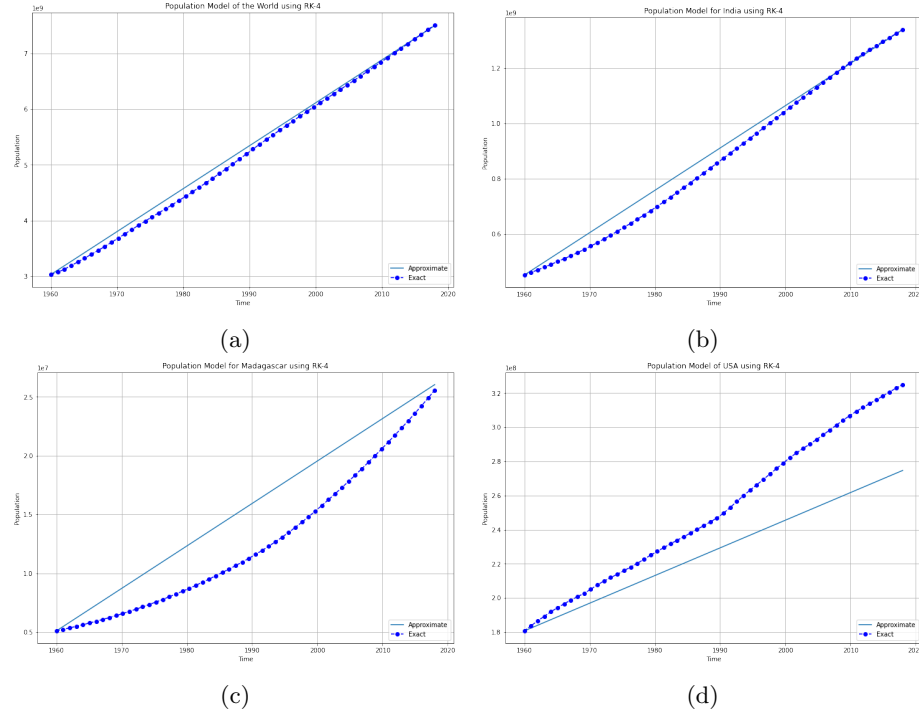


Figure 7: Population versus year graph using RK4 method for (a) World, (b) India, (c) Madagascar and (d) USA

So from the above figure it is clear that this model predicts the overall world population and India's population more accurately but miserably fails for USA and there is a significant error in the case of Madagascar. This huge error can be accounted if we take into consideration other factors like immigration, emigration etc, because USA is country that has a huge number of immigrations every year. Not taking into account the immigration factor is why this model

fails miserably to predict USA's growth model. This also explains the reason why the model predicts the world population growth with very little error.

5.4.1 Predicting population trend for next 10 years using the above model

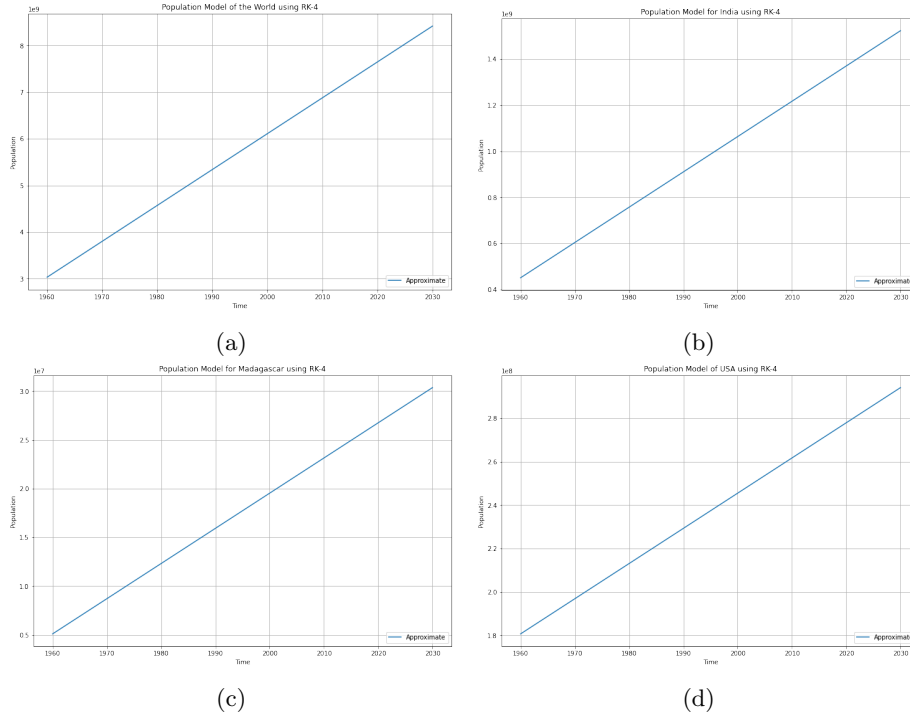


Figure 8: Population versus year graph using RK4 method for (a) World, (b) India, (c) Madagascar and (d) USA

According to this model the world population will reach more than 8 billion by the end of 2030. At the same time India will reach a population of 1.5 billion approx, USA will reach 280 million and Madagascar will reach 30 million. From these numbers we can clearly say that by 2030 around 19% of the world population will be Indians.

5.5 Fertility growth rate

In this section we try to fit polynomial function (solid line) over the fertility growth rate data of a country obtained from the World Bank. We use polynomial regression technique for this purpose.

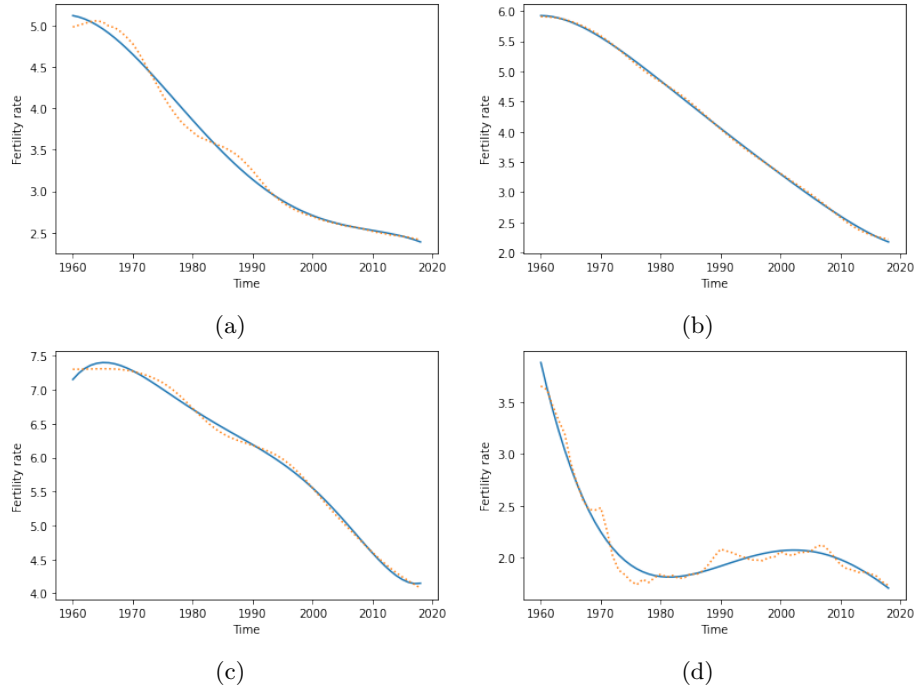


Figure 9: Fertility versus year graph using polynomial regression for (a) World, (b) India, (c) Madagascar and (d) USA

So from the above graph it is clear that that India's fertility rate over the years have come down to 2 children per women in 2019 from 6 children per women in 1960. Similarly for USA it has come down from 4 to 1(approx) and for Madagascar it has come down from 7.5 to 4 children per women, which is still quiet high. Overall the world population fertility rate has come down from 5 to 2.5, which is a good sign. We can see all the countries are approaching the replacement rate of 2.1. As you can see from the graph, the fertility rate of USA has gone down from the replacement rate and hence the population is expected to decline. On the other hand India it is going to reach the replacement rate very soon and currently it is at 2.5, therefore India's population will grow at a more stable rate and will not see drastic increase. Whereas in the case of Madagascar the fertility rate is pretty high and hence it will see drastic increase in its population in the coming years.

So, from the graph we can conclude that Under Developed Countries have high

fertility growth rate, whereas Developed and Developing countries have comparatively low fertility growth rate.

5.5.1 Predicting next 10 year fertility growth

Using polynomial regression in the above section we found the polynomial that approximately mimics the fertility growth of different countries, using the same polynomial obtained in the previous section we are going to predict the fertility growth of different countries for next 10 years in this section.

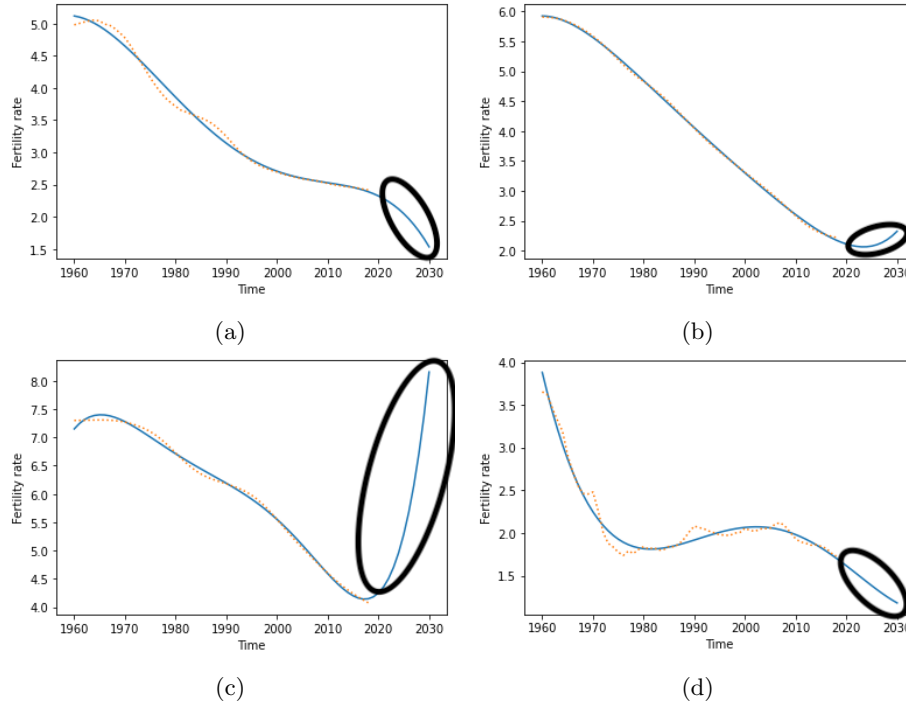


Figure 10: Fertility versus year graph using polynomial regression for (a) World, (b) India, (c) Madagascar and (d) USA

Now in the above graphs the black encircled region is the predicted fertility growth of different countries for next 10 years. We can see that for Madagascar there is an abrupt increase in fertility growth which is a worst case scenario, and hence our model is not accurate. But for the rest of the 2 countries and even for the world it does not show this kind of abrupt behaviour and hence we can take those predictions into account. So, the polynomial model for India predicts that by the next 10 years the fertility growth will experience a slight increase, whereas for USA it will further fall down from 1.5 to 1 (approx). Overall, the world's fertility growth is going to fall down and reach the mark of 1.5 by the end 2030.

6 Conclusion

So we can conclude following things :

- Population of developing countries like India is going to grow with a fertility rate of 2 children's per women and the total population will reach 1.5 billion by 2030.
- Population of developed countries like USA is going to grow with a fertility rate of 1 child per women and the total population will reach 280 million by 2030.
- Population of under developed countries like Madagascar will grow with a fertility rate of 4 children's per women and is going to reach 36 million by 2030.
- The total world population is going to grow with fertility rate of 2.5 children's per women and the total world population will reach 8 billion people by 2030.
- We can also interpret from above figures that under developing countries like Madagascar are going to experience exponential population growth, which is problem that needs to be addressed urgently. The people of this countries must be educated about their rising population.
- On the other hand developed and developing countries are doing good they brought down their fertility rate significantly and are managing their population growth efficiently.
- Also, according to logistic growth model it may flatten after the year 2200. That is population may stabilise by 2200 and it will not grow beyond that. This is true for all the three countries and world population.
- Fertility growth prediction also agree with the above observations because for every single country fertility growth is decreasing except Madagascar which we believe is an error associated with model.

7 Future Work

In this project we modeled population model with simple exponential growth model, logistic growth model and birth and death rate model. In the future we would like to work more on this project and come up with a mathematical model that takes into account different factors like economic status of a country, health facility, educational status etc. before predicting the population. We also wish to a make model that predicts the population more accurately with very less error because this will help countries to devise their future policies, that benefits all the people. We would like to inculcate machine learning techniques

to predict more accurate mathematical models. We would also like to learn non-linear regression techniques so that we can fit the logistic curve more accurately to the data and find the parameters involved with more precision.

8 Codes

8.1 Curve fitting for Population (Exponential model)

```
1 import numpy as np
2 import math
3 import pandas as pd
4 import matplotlib.pyplot as plt
5
6 from google.colab import files
7 uploaded = files.upload()
8
9 data=pd.read_csv("population data - Sheet1.csv",header=None)
10
11 data
12
13 d=pd.DataFrame(data)
14
15 x_=pd.to_numeric(d.iloc[0,1:])
16
17 y_=pd.to_numeric(d.iloc[2,1:])
18
19 plt.plot(x_, y_, label = "line 1")
20
21 """Curve fitting-Exponential"""
22
23 def estimate_coef(x, y):
24     n = np.size(x)
25
26     m_x, m_y = np.mean(x), np.mean(y)
27
28     SS_xy = n*np.sum(y*x) - n*n*m_y*m_x
29     SS_xx = n*np.sum(x*x) - n*n*m_x*m_x
30
31     a_1 = SS_xy / SS_xx
32     a_0 = m_y - a_1*m_x
33
34     return(a_0, a_1)
35
36 def plot_regression_line(x, y, a):
37     plt.scatter(x, y, color = "g",
38                 marker = "o", s = 30)
39
40     y_pred = a[0] + a[1]*x
41
42     plt.plot(x, y_pred, color = "r")
43
44     plt.xlabel('x')
45     plt.ylabel('y')
46
47     plt.show()
48
49 x=x_
50 y=np.log(y_)
51 a = estimate_coef(x, y)
52 print(a)
53 plot_regression_line(x, y, a)
```

```

54
55 N0=math.exp(a[0])
56 l=a[1]
57
58 y=N0*np.exp(l*x)
59 plt.plot(x_, y_)
60 plt.plot(x_, y)

```

8.2 Euler's Method and RK4 for Population Model (Exponential model)

```

1  """## Euler's Method"""
2
3  def euler(f,a,b,n,No):
4      h = (b-a)/(n-1)
5      ts = a + np.arange(n)*h
6      ns = np.zeros(n)
7      n = No
8      for j,t in enumerate(ts):
9          ns[j] = n
10         n += h*f(t, n)
11     return ts, ns
12
13  """## RK-4 Method"""
14
15  def rk4(f,a,b,n,No):
16      h = (b-a)/(n-1)
17      ts = a + np.arange(n)*h
18      ns = np.zeros(n)
19      n = No
20      for j,t in enumerate(ts):
21          ns[j] = n
22          k0 = h*f(t, n)
23          k1 = h*f(t+h/2, n+k0/2)
24          k2 = h*f(t+h/2, n+k1/2)
25          k3 = h*f(t+h, n+k2)
26          n += (k0 + 2*k1 + 2*k2 + k3)/6
27     return ts, ns
28
29  """## Population Model (Exponential Model)
30  ### $$dN/dt=rN$$ N= population as a function of time, r=
31     proportionality constant
32  ### $$
33  N(t) = N_0 e^{\{rt\}},
34  $$ No= population at time t=0
35  ### 1. Madagascar
36  r=0.02886290941170391, No=5099373
37  """
38
39  def f(t,n):
40      r=0.02886290941170391
41      return r*n #we are defining the differential equation here
42
43  """### Euler's Method"""
44

```

```

45 t, n= euler(f, 1960, 2030, 69,5099373 )
46
47 plt.figure(figsize = (12, 8))
48 plt.plot(t, n, 'bo--', label='Approximate')
49 plt.title('Population Model for Madagascar')
50 plt.xlabel('Time')
51 plt.ylabel('Population')
52 plt.grid()
53 plt.legend(loc='lower right')
54 plt.show()
55
56 """### RK-4 Method"""
57
58 t1, n1= rk4(f, 1960, 2030, 69,5099373 )
59
60 plt.figure(figsize = (12, 8))
61 plt.plot(t1, n1, 'bo--', label='Approximate')
62 plt.title('Population Model for Madagascar using RK-4')
63 plt.xlabel('Time')
64 plt.ylabel('Population')
65 plt.grid()
66 plt.legend(loc='lower right')
67 plt.show()
68
69 """## India
70 r=0.01965563394203113, No=450547679
71 """
72
73 def f(t,n):
74     r=0.01965563394203113
75     return r*n
76
77 """### Euler's Method"""
78
79 t, n= euler(f, 1960, 2030, 69, 450547679 )
80
81 plt.figure(figsize = (12, 8))
82 plt.plot(t, n, 'bo--', label='Approximate')
83 plt.title("Population model for India using Euler's Method")
84 plt.xlabel('Time')
85 plt.ylabel('Population')
86 plt.grid()
87 plt.legend(loc='lower right')
88 plt.show()
89
90 """### RK-4 Method"""
91
92 t1, n1= rk4(f, 1960, 2030, 69, 450547679 )
93
94 plt.figure(figsize = (12, 8))
95 plt.plot(t1, n1, 'bo--', label='Approximate')
96 plt.title('Population Model for India using RK-4')
97 plt.xlabel('Time')
98 plt.ylabel('Population')
99 plt.grid()
100 plt.legend(loc='lower right')
101 plt.show()

```

```

102
103 """## USA
104 r=0.010234409905400798, No=180671000
105 """
106
107 def f(t,n):
108     r=0.010234409905400798
109     return r*n
110
111 """### Euler's Method"""
112
113 t, n= euler(f, 1960, 2030, 69, 180671000 )
114
115 plt.figure(figsize = (12, 8))
116 plt.plot(t, n, 'bo--', label='Approximate')
117 plt.title("Population model of USA using Euler's Method")
118 plt.xlabel('Time')
119 plt.ylabel('Population')
120 plt.grid()
121 plt.legend(loc='lower right')
122 plt.show()
123
124 """### RK-4 Method"""
125
126 t1, n1= rk4(f, 1960, 2030, 69, 180671000 )
127
128 plt.figure(figsize = (12, 8))
129 plt.plot(t1, n1, 'bo--', label='Approximate')
130 plt.title('Population Model of USA using RK-4')
131 plt.xlabel('Time')
132 plt.ylabel('Population')
133 plt.grid()
134 plt.legend(loc='lower right')
135 plt.show()
136
137 """## World
138 r=0.015906861837129742, No=3031437768
139 """
140
141 def f(t,n):
142     r=0.015906861837129742
143     return r*n
144
145 """### Euler's Method"""
146
147 t, n= euler(f, 1960, 2030, 69, 3031437768)
148
149 plt.figure(figsize = (12, 8))
150 plt.plot(t, n, 'bo--', label='Approximate')
151 plt.title("Population model of the World using Euler's Method")
152 plt.xlabel('Time')
153 plt.ylabel('Population')
154 plt.grid()
155 plt.legend(loc='lower right')
156 plt.show()
157
158 """### RK-4 Method"""

```



```

159
160 t1, n1= rk4(f, 1960, 2030, 69, 3031437768)
161
162 plt.figure(figsize = (12, 8))
163 plt.plot(t1, n1, 'bo--', label='Approximate')
164 plt.title('Population Model of the World using RK-4')
165 plt.xlabel('Time')
166 plt.ylabel('Population')
167 plt.grid()
168 plt.legend(loc='lower right')
169 plt.show()
170
171 c=0.0035429
172 d=-3.61298*10**-11
173
174 def p(t):
175     return p
176
177 def f(t,p):
178     c=0.0035429
179     d=-3.61298*10**-11
180     k= (p*c*(c-d*p))/c
181     return k
182
183 t1, p1=rk4(f, 1960, 2019, 59, 450547679)
184
185 plt.figure(figsize = (12, 8))
186 plt.plot(t1, p1, 'bo--', label='Approximate')
187 plt.title("Population model of the World using Euler's Method")
188 plt.xlabel('Time')
189 plt.ylabel('Population')
190 plt.grid()
191 plt.legend(loc='lower right')
192 plt.show()

```

8.3 Logistic Growth Model

```

1 data=pd.read_csv("population data - Sheet1.csv",header=None)
2
3 data
4
5 d=pd.DataFrame(data)
6
7 x_=pd.to_numeric(d.iloc[0,1:])
8
9 y_=pd.to_numeric(d.iloc[4,1:])
10
11 def euler(f,a,b,n,No):
12     h = (b-a)/(n-1)
13     ts = a + np.arange(n)*h
14     ns = np.zeros(n)
15     n = No
16     for j,t in enumerate(ts):
17         ns[j] = n
18         n += h*f(t, n)
19     return ts, ns
20

```

```

21 def rk4(f,a,b,n,No):
22     h = (b-a)/(n-1)
23     ts = a + np.arange(n)*h
24     ns = np.zeros(n)
25     n = No
26     for j,t in enumerate(ts):
27         ns[j] = n
28         k0 = h*f(t, n)
29         k1 = h*f(t+h/2, n+k0/2)
30         k2 = h*f(t+h/2, n+k1/2)
31         k3 = h*f(t+h, n+k2)
32         n += (k0 + 2*k1 + 2*k2 + k3)/6
33     return ts, ns
34
35 """World"""
36
37 x_=pd.to_numeric(d.iloc[0,1:])
38 y_=pd.to_numeric(d.iloc[4,1:])
39
40 def f(t,n):
41     r=0.024
42     k=2*max(y_)
43     return r*n*(1-(float)(n/k)) #we are defining the differential
44     equation here
45
46 t1,n1 = euler(f,1960,2200,240,y_[1])
47
48 plt.figure(figsize = (12, 8))
49 plt.scatter(x_, y_)
50 plt.plot(t1,n1,'r',label = 'Approximate')
51 plt.title('Logistic population Model for world using Euler')
52 plt.xlabel('Year')
53 plt.ylabel('Population')
54 plt.grid()
55 plt.legend(loc='lower right')
56 plt.show()
57
58 t1, n1= rk4(f, 1960, 2200, 240,y_[1])
59
60 plt.figure(figsize = (12, 8))
61 plt.plot(t1, n1,'r',label='Approximate')
62 plt.scatter(x_,y_)
63 plt.title('Logistic population Model of the World using RK-4')
64 plt.xlabel('Time')
65 plt.ylabel('Population')
66 plt.grid()
67 plt.legend(loc='lower right')
68 plt.show()
69
70 """India"""
71
72 x_=pd.to_numeric(d.iloc[0,1:])
73 y_=pd.to_numeric(d.iloc[1,1:])
74
75 def f(t,n):
76     r=0.02879
77     k=2*max(y_)

```

```

77     return r*n*(1-(float)(n/k))  #we are defining the differential
    equation here
78
79 t1,n1 = euler(f,1960,2200,240,y_[1])
80
81 plt.figure(figsize = (12, 8))
82 plt.scatter(x_, y_)
83 plt.plot(t1,n1,'r',label = 'Approximate')
84 plt.title('Logistic population Model for India using Euler')
85 plt.xlabel('Year')
86 plt.ylabel('Population')
87 plt.grid()
88 plt.legend(loc='lower right')
89 plt.show()
90
91 t1, n1= rk4(f, 1960, 2200, 240,y_[1])
92
93 plt.figure(figsize = (12, 8))
94 plt.plot(t1, n1,'r',label='Approximate')
95 plt.scatter(x_,y_)
96 plt.title('Logistic population Model of the India using RK-4')
97 plt.xlabel('Time')
98 plt.ylabel('Population')
99 plt.grid()
100 plt.legend(loc='lower right')
101 plt.show()
102
103 """Madagascar"""
104
105 x_=pd.to_numeric(d.iloc[0,1:])
106 y_=pd.to_numeric(d.iloc[3,1:])
107
108 def f(t,n):
109     r=0.03478
110     k=2*max(y_)
111     return r*n*(1-(float)(n/k))  #we are defining the differential
    equation here
112
113 t1,n1 = euler(f,1960,2200,240,y_[1])
114
115 plt.figure(figsize = (12, 8))
116 plt.scatter(x_, y_)
117 plt.plot(t1,n1,'r',label = 'Approximate')
118 plt.title('Logistic population Model for Madagascar using Euler')
119 plt.xlabel('Year')
120 plt.ylabel('Population')
121 plt.grid()
122 plt.legend(loc='lower right')
123 plt.show()
124
125 t1, n1= rk4(f, 1960, 2200, 240,y_[1])
126
127 plt.figure(figsize = (12, 8))
128 plt.plot(t1, n1,'r', label='Approximate')
129 plt.scatter(x_,y_)
130 plt.title('Logistic population Model of the Madagascar using RK-4')
131 plt.xlabel('Time')

```

```

132 plt.ylabel('Population')
133 plt.grid()
134 plt.legend(loc='lower right')
135 plt.show()
136
137 """USA"""
138
139 x_=pd.to_numeric(d.iloc[0,1:])
140 y_=pd.to_numeric(d.iloc[2,1:])
141
142 def f(t,n):
143     r=0.017
144     k=2*max(y_)
145     return r*n*(1-(float)(n/k)) #we are defining the differential
                                #equation here
146
147 t1,n1 = euler(f,1960,2200,240,y_[1])
148
149 plt.figure(figsize = (12, 8))
150 plt.scatter(x_, y_)
151 plt.plot(t1,n1,'r',label = 'Approximate')
152 plt.title('Logistic population Model for USA using Euler')
153 plt.xlabel('Year')
154 plt.ylabel('Population')
155 plt.grid()
156 plt.legend(loc='lower right')
157 plt.show()
158
159 t1, n1= rk4(f, 1960, 2200, 240,y_[1])
160
161 plt.figure(figsize = (12, 8))
162 plt.plot(t1, n1,'r', label='Approximate')
163 plt.scatter(x_,y_)
164 plt.title('Logistic population Model of the USA using RK-4')
165 plt.xlabel('Time')
166 plt.ylabel('Population')
167 plt.grid()
168 plt.legend(loc='lower right')
169 plt.show()

```

8.4 Population Modelling using Birth Rate and Death Rate

```

1 data_p=pd.read_excel("population data.xlsx",header=None)
2 data_p
3
4 data_b=pd.read_excel("Birth_Rate.xlsx",header=None)
5 data_b
6
7 data_d=pd.read_excel("Death_Rate.xlsx",header=None)
8 data_d
9
10 z=4
11 d1_d=data_d.iloc[z]
12 d1_p=data_p.iloc[z]
13 d1_b=data_b.iloc[z]

```

```

14
15 sum = 0
16
17 for i in range(1,np.size(d1_d)):
18     sum = sum + ((d1_d[i]*d1_p[i])/1000.0)
19 d_avg = sum/np.size(d1_d)
20 d_avg
21
22 sum = 0
23
24 for i in range(1,np.size(d1_b)):
25     sum = sum + ((d1_b[i]*d1_p[i])/1000.0)
26 b_avg = sum/np.size(d1_b)
27 b_avg
28
29 def rk4(f,a,b,n,No):
30     h = (b-a)/(n-1)
31     ts = a + np.arange(n)*h
32     ns = np.zeros(n)
33     n = No
34     for j,t in enumerate(ts):
35         ns[j] = n
36         k0 = h*f(t, n)
37         k1 = h*f(t+h/2, n+k0/2)
38         k2 = h*f(t+h/2, n+k1/2)
39         k3 = h*f(t+h, n+k2)
40         n += (k0 + 2*k1 + 2*k2 + k3)/6
41     return ts, ns
42
43 def f(t,n):
44     return b_avg-d_avg
45
46 """India"""
47
48 t1, n1= rk4(f, 1960, 2018, 58, 450547679 )
49
50 plt.figure(figsize = (12, 8))
51 plt.plot(t1, n1, linestyle='solid', label='Approximate')
52 plt.plot(t1,d1_p[1:59], 'bo--', label='Exact')
53 plt.title('Population Model for India using RK-4')
54 plt.xlabel('Time')
55 plt.ylabel('Population')
56 plt.grid()
57 plt.legend(loc='lower right')
58 plt.show()
59
60 """USA"""
61
62 t1, n1= rk4(f, 1960, 2018, 58, 180671000 )
63
64 plt.figure(figsize = (12, 8))
65 plt.plot(t1, n1, linestyle='solid', label='Approximate')
66 plt.plot(t1,d1_p[1:59], 'bo--', label='Exact')
67 plt.title('Population Model of USA using RK-4')
68 plt.xlabel('Time')
69 plt.ylabel('Population')
70 plt.grid()

```

```

71 plt.legend(loc='lower right')
72 plt.show()
73
74 """Madagascar
75
76 """
77
78 t1, n1= rk4(f, 1960, 2018, 58,5099373 )
79
80 plt.figure(figsize = (12, 8))
81 plt.plot(t1, n1, linestyle='solid', label='Approximate')
82 plt.plot(t1,d1_p[1:59], 'bo--', label='Exact')
83 plt.title('Population Model for Madagascar using RK-4')
84 plt.xlabel('Time')
85 plt.ylabel('Population')
86 plt.grid()
87 plt.legend(loc='lower right')
88 plt.show()
89
90 """World"""
91
92 t1, n1= rk4(f, 1960, 2018, 58, 3031437768)
93
94 plt.figure(figsize = (12, 8))
95 plt.plot(t1, n1, linestyle='solid', label='Approximate')
96 plt.plot(t1,d1_p[1:59], 'bo--', label='Exact')
97 plt.title('Population Model of the World using RK-4')
98 plt.xlabel('Time')
99 plt.ylabel('Population')
100 plt.grid()
101 plt.legend(loc='lower right')
102 plt.show()

```

8.5 Prediction using Birth Rate and Death Rate

```

1  """
2  India
3  """
4
5  t1, n1= rk4(f, 1960, 2030, 69, 450547679 )
6
7  plt.figure(figsize = (12, 8))
8  plt.plot(t1, n1, label='Approximate')
9  plt.title('Population Model for India using RK-4')
10 plt.xlabel('Time')
11 plt.ylabel('Population')
12 plt.grid()
13 plt.legend(loc='lower right')
14 plt.show()
15
16 """USA
17
18 """
19
20 t1, n1= rk4(f, 1960, 2030, 69, 180671000 )
21
22 plt.figure(figsize = (12, 8))

```

```

23 plt.plot(t1, n1, 'bo--', label='Approximate')
24 plt.title('Population Model of USA using RK-4')
25 plt.xlabel('Time')
26 plt.ylabel('Population')
27 plt.grid()
28 plt.legend(loc='lower right')
29 plt.show()
30
31 """Madagascar
32
33 """
34
35 t1, n1= rk4(f, 1960, 2030, 69,5099373 )
36
37 plt.figure(figsize = (12, 8))
38 plt.plot(t1, n1, 'bo--', label='Approximate')
39 plt.title('Population Model for Madagascar using RK-4')
40 plt.xlabel('Time')
41 plt.ylabel('Population')
42 plt.grid()
43 plt.legend(loc='lower right')
44 plt.show()
45
46 """World"""
47
48 t1, n1= rk4(f, 1960, 2030, 69, 3031437768)
49
50 plt.figure(figsize = (12, 8))
51 plt.plot(t1, n1, 'bo--', label='Approximate')
52 plt.title('Population Model of the World using RK-4')
53 plt.xlabel('Time')
54 plt.ylabel('Population')
55 plt.grid()
56 plt.legend(loc='lower right')
57 plt.show()

```

8.6 Curve fitting for Fertility Rate (Exponential model)

```

1 d1=pd.read_csv("Fertility_rate _data - Sheet1.csv",header=None)
2
3 d1
4
5 d1=pd.DataFrame(d1)
6
7 y_=pd.to_numeric(d1.iloc[3,1:])
8 x_=pd.to_numeric(d1.iloc[0,1:])
9 plt.plot(x_, y_, label = "line 1")
10
11 def estimate_coef(x, y):
12     n = np.size(x)
13
14     m_x, m_y = np.mean(x), np.mean(y)
15
16     SS_xy = n*np.sum(y*x) - n*n*m_y*m_x
17     SS_xx = n*np.sum(x*x) - n*n*m_x*m_x
18
19     a_1 = SS_xy / SS_xx

```

```

20     a_0 = m_y - a_1*m_x
21
22     return(a_0, a_1)
23
24 def plot_regression_line(x, y, a):
25     plt.scatter(x, y, color = "g",
26                marker = "o", s = 30)
27
28     y_pred = a[0] + a[1]*x
29
30     plt.plot(x, y_pred, color = "r")
31
32     plt.xlabel('x')
33     plt.ylabel('y')
34
35     plt.show()
36
37 x=x_
38 y=np.log(y_)
39 a = estimate_coef(x, y)
40 print(a)
41 plot_regression_line(x, y, a)
42
43 N0=math.exp(a[0])
44 l=a[1]
45
46 y=N0*np.exp(l*x)
47 plt.plot(x_, y_)
48 plt.plot(x_, y)

```

8.7 Polynomial Regression for Fertility Rate

```

1 df=pd.read_excel("Fertility_rate_data.xlsx",header=None)
2 print(df)
3
4 df1=df.iloc[4]
5 df1
6
7 fertility=df1.to_numpy()
8 fertility
9
10 year=np.zeros((59))
11 i=0
12 for i in range(59):
13     year[i]=1960+i
14 print(year)
15
16 P=np.zeros((59))
17 for i in range(59):
18     P[i]=fertility[i+1]
19 print(P)
20
21 Errortable=np.zeros((30))
22 equation=np.zeros((30,30))
23 for n in range(1,30):
24     AUGM=np.zeros((n+1,n+1))
25     AUGM2=np.zeros(n+1)

```



```

26 i=0
27 j=0
28 for i in range (n+1):
29     for j in range (n+1):
30         sum=0
31         for k in range(np.size(P)):
32             sum= sum+((year[k])**j+i)
33
34         AUGM[j][i]= sum
35
36 for h in range(n+1):
37     sum=0
38     for k in range(np.size(P)):
39         sum=sum+(P[k]*(year[k]**h))
40
41     AUGM2[h]=sum
42 m = n+1
43 x=[0]*m
44 for k in range(0,m-1):
45     for i in range(k+1,m):
46         factor = AUGM[i][k]/AUGM[k][k]
47         for j in range(k+1,m):
48             AUGM[i][j]=AUGM[i][j]-factor*AUGM[k][j]
49             AUGM2[i]=AUGM2[i]-factor*AUGM2[k]
50
51 x[m-1]=AUGM2[m-1]/AUGM[m-1][m-1]
52 for i in range(n-1,-1,-1):
53     sum = AUGM2[i]
54     for j in range(i+1,m):
55         sum = sum - AUGM[i][j]*x[j]
56     x[i] = sum/AUGM[i][i]
57
58 Sr=0
59 for q in range(59):
60     sum2=0
61     for e in range(m):
62         sum2=sum2+ (x[e]*(year[q]**e))
63
64     Sr=Sr+((P[q]-(sum2))**2)
65 Sterror= (Sr/(60-(n+1))**(1/2))
66 print(Sterror)
67 Errortable[n]=Sterror
68 for i in range(m):
69     equation[n][i]=x[i]
70
71 print(AUGM[0][0],AUGM[1][0])
72
73 dummy =Errortable[1]
74 degree=1
75 for i in range(2,30):
76     if Errortable[i]<=dummy:
77         dummy=Errortable[i]
78         degree=i
79 print("degree=",degree,"standard error=", dummy,)
80
81 plt.plot(year,P,linestyle='dotted')
82

```

```

83 equation[degree]
84
85 y=0
86 for i in range(degree+1):
87     y=y+(equation[degree][i]*(year**i))
88
89 plt.plot(year,y,linestyle='solid')
90 plt.plot(year,P,linestyle='dotted')
91
92 for i in range(1,30):
93     v=0
94     for j in range(30):
95         v=v+equation[i][j]*(year**j)
96
97     plt.plot(year,v)

```

8.8 Predicting Fertility Rate with Polynomial Regression

```

1 y=0
2 yearnew=np.zeros((71))
3 for i in range(71):
4     yearnew[i]=1960+i
5 print(yearnew)
6 for i in range(degree+1):
7     y=y+(equation[degree][i]*(yearnew**i))
8
9 plt.plot(yearnew,y,linestyle='solid')
10 plt.plot(year,P,linestyle='dotted')
11 plt.xlabel('Time')
12 plt.ylabel('Fertility rate')
13
14 for i in range(1,30):
15     v=0
16     for j in range(30):
17         v=v+equation[i][j]*(year**j)
18
19     plt.plot(year,v)

```

9 References

- <https://data.worldbank.org/indicator/SP.POP.TOTL>
- Numerical Methods for Engineers by Steven C. Chapra and Raymond P. Canale.
- moak, Natalie. "Fertility rate". Encyclopedia Britannica, 12 May. 2016, <https://www.britannica.com/topic/fertility-rate>. Accessed 16 April 2021.
- Ecology and the environment-lumen biology
<https://courses.lumenlearning.com/biology2xmaster/chapter/human-population-growth/>
- Vandermeer, J. (2010) How Populations Grow: The Exponential and Logistic Equations. Nature
<https://www.nature.com/scitable/knowledge/library/how-populations-grow-the-exponential-and-logistic-13240157/>
- Population ecology-Khan academy
<https://www.khanacademy.org/science/ap-biology/ecology-ap/population-ecology-ap/a/exponential-logistic-growth>
- Population growth and Regulation-Concepts of Biology
<https://opentextbc.ca/conceptsofbiologyopenstax/chapter/population-growth-and-regulation/>
- <https://en.wikipedia.org/wiki/Populationgrowth>