TIME SERIES ANALYSIS OF POPULATION DATA: A CASE STUDY OF FEDERAL GOVERNMENT GIRLS' COLLEGE, ZARIA

BY

CHIDI JONATHAN OGOCHUKWU
PGD/MCS/1997/439

DEPARTMENT OF MATHEMATICS / COMPUTER SCIENCE
FEDERAL UNIVERSITY OF TECHNOLOGY,
MINNA

NOVEMBER, 2003.

TIME SERIES ANALYSIS OF POPULATION DATA: A Case Study of Federal Government Girls' College, Zaria.

BY

CHIDI JONATHAN OGOCHUKWU
PGD/MCS/1997/439

A PROJECT REPORT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF A POST GRADUATE DIPLOMA IN COMPUTER SCIENCE

DEPARTMENT OF MATHEMATICS / COMPUTER SCIENCE,
FEDERAL UNIVERSITY OF TECHNOLOGY,
MINNA.

NOVEMBER, 2003.

CERTIFICATION

This is to certify that the project titled "Time Series Analysis of Population Data: A case study of Federal Government Girls' College, Zaria." is an original work undertaken by Chidi Jonathan Ogochukwu, PGD/MCS/1997/439

______________________ ______________________

Mallam Audu Isah Date
Project Supervisor

______________________ ______________________

Date

______________________ ______________________

External Examiner Date

DEDICATION

This work is dedicated to my loving wife, Mrs. Patience Chidi and our lovely children, Jehoram and Uchechukwu Osanwuta for their support.

ACKNOWLEDGEMENT

To the Almighty God, Jehovah goes my profound gratitude for seeing me through this work.

My immense appreciation and gratitude goes to Mallam Audu Isah, my able supervisor, for painstakingly going through the scripts and offering useful suggestions and advice. I thank Dr. Aiyesimi of the Dept. of Maths & Computer Science, FUT, Minna; for his constructive criticisms, which proved to very helpful. I also thank Mr. Sunday Onwu of F.G.G.C., Zaria for typing the manuscripts.

I wish to also express my appreciation to my loving wife, Mrs. Patience Chidi, our little boys, Jehoram and Uchechukwu and other family members for being supportive during this course.

ABSTRACT

This project work focuses on the use of modern scientific approach to projection or forecasting. Particular attention was paid to time series analysis approach using the students' population of F.G.G.C., Zaria as a case study.

Chapter 1 gives a general introduction to the study, its background and statement of the problem. Research questions were posed with the aim of finding answers to them later. The significance, justification and scope of the study were discussed.

Chapter 2 has to do with a brief review of literature relevant to the study. It highlights certain concepts commonly encountered in time series analysis.

Chapter 3 deals with the design of the study, procedure for data collection and the result of graphing the data.

The software development and implementation were discussed in chapter four, while chapter five covered the summary and conclusion.

TABLE OF CONTENTS

TIME SERIES ANALYSIS OF POPULATION DATA: A CASE STUDY OF FEDERAL GOVERNMENT GIRLS' COLLEGE, ZARIA

BY

CHIDI JONATHAN OGOCHUKWU
PGD/MCS/1997/439

DEPARTMENT OF MATHEMATICS / COMPUTER SCIENCE
FEDERAL UNIVERSITY OF TECHNOLOGY,
MINNA

NOVEMBER, 2003.

CHAPTER I

1.0 Introduction

Planning for the future is one of the main tasks facing the Principal of a secondary school. He has the responsibility of looking ahead and making forecasts of future activities upon which intelligent planning must depend. Thus a large part of the effort devoted to running a school is spent on planning for the future. The school management must, therefore, spend a great deal of time in thinking about the past, in analyzing statistical information about the past, in order to gain an understanding of phenomena that can be projected into the future. In making many decisions one is persuaded to believe that the future follows the past with some degree of regularity. Hence the first step in forecasting the future consists of gathering observations from the past. In this regard, one usually deals with statistical data, which are collected, observed and recorded at successive intervals of time, which are generally referred to as time series.

The principal of a school needs to know the projected number of students in the school in order to prepare his budget, which must be submitted to the ministry in advance. This will enable him to estimate, with a fair degree of accuracy, the expected revenue to be generated in terms of levies, tuition (where obtainable), dues, etc. The amount of money to be expended on the total number of students in a particular period, say one academic session, can also be inferred. The population of a school is usually not static, and if measured, constitutes a time series. (Freund, 1976:383). The study and measurement of seasonal patterns constitute a very important part of the analysis of a time series.

The school's yearly population for 12 years was collected from the Administrative Office of Federal Government Girls' College, Zaria. There was no need to adjust the data for comparability as it could not have been affected by factors such as leap years and holidays. The graph of the data was drawn, noting qualitatively the presence of seasonal variations and of long-term trend and cyclic variations. Attempt was made to select the equation that best describes the gradual and consistent pattern of growth. Graphing and differencing aided in deciding which kid of trend is right for this series.

1.1 Background To The Study

Since its inception in 1989, Fed. Govt. Girls' College, Zaria has been experiencing a yearly increase in students' population. With this growth, it becomes more challenging to manage the affairs of the institution. An essential aspect of managing the school has to do with projecting the students' population figure. This is necessary for planning against the future in terms of budgeting, feeding, provision of social facilities, etc. In making this projection, some principals rely on past information, experience and intuition.

Presently, scientific and statistical methods of forecasting are seldom used in Nigerian secondary schools to predict the student population for a future date. The complex nature of the mathematical calculations involved in scientific and statistical methods of forecasting tends to discourage the use of such methods. With the introduction of computers that can handle many complex calculations at a time, Managers of institutions can make `informed' and `calculated' guesses of future events.

1.2 Statement Of The Problem

Planning is an important aspect of managing an institution like Fed, Govt. Girls' College, Zaria. As of now, scientific and statistical method of planning is not being used in the school. So the problem remains that of:

1. Identifying information and data to be sought or generated;
2. How to analyze, store and retrieve such data and build time series from it as necessary;
3. Using the data to make projections.

1.3 Research Questions

The research questions have to do with the effectiveness of the present method of students' population projection.

1. Should the current method of population forecasting be simply improved upon or completely replaced by a more scientific and statistical approach?
2. Would the school Authority be willing to adopt this new approach?
3. Is the new approach practicable and cost effective?
4. What type of curve is most appropriate for describing the population data of Fed. Govt. Girls' College, Zaria?

1.4 Objectives Of The Study

The objectives of this study are:

(i) To build a 'time series' for students' population;

(ii) To identify the most appropriate curve to be fitted and to examine how well the curve 'fits' the existing data;

(iii) To build a model to project the school's student population for the next ten years.

1.5 Significance Of The Study

The significance of this study stems from the fact that a different perspective is being considered for forecasting student population. It is hoped that the method being considered will improve on the existing one.

1.6 Justification Of The Study

To run a school successfully, it is imperative to plan for the future. Reliable and effective method of harnessing data can lead to accurate projections. The method of planning in institutions that should be in vogue involves scientific and statistical analysis of data. This study tends to popularize this modern approach.

1.7 Scope Of The Study

This study seeks to use a time series approach to forecast students' population figures for ten years. It does not aim to exhaust all possible methods for achieving this. The range is kept short to minimize errors. Thus, this research cannot claim to have constructed the best possible model that describes population data. It mainly examines how well the Gompertz curve models the population growth in Fed. Govt. Girls' College, Zaria.

1.8 Summary

Being an integral part of the managerial decision making process, forecasting provides an aid to management. It also provides a more scientific approach to making decisions. In a growing institution, like Fed. Govt. Girls' College, Zaria , one needs to select an equation or method of measuring trend which best describes the gradual and consistent pattern of growth. To operate effectively, a knowledge of the major types of curves and methods used is necessary as well as a considerable amount of judgment and experience. The Gompertz curve is chosen because it serves to describe growth of series which, while increasing seen to approach some maximum value as a limit. According to Freund and Williams (1976), "the growth may continue but does so at a decreasing rate". Graphing and the second difference test aided in deciding that the Gompertz curve is the right kind of trend for this series.

CHAPTER TWO

2.0 Review of Related Literature

2.1 Introduction

The ability to handle, interpret and use figures is a prerequisite for today's modern manager. Such ability helps to plan for the future, looking ahead and making the forecasts of future levels of activity upon which all intelligent planning necessarily depends. Normally, the first step in forecasting the future consists of gathering observations from the past. The observations are statistical data which are collected observed or recorded at successive intervals of time, generally referred to as time series, as noted earlier on.

It is common to see graphs of time series showing the behaviour of stocks and bonds, reports on weekly or monthly sales or charts of the daily attendance in a school. Some of these graphs look like straight lines, others look like smooth curves but many, especially those representing economic data, give the impression of haphazard movements. Consequently, special statistical techniques have been devised to help bring some order into the irregular patterns and the seemingly erratic appearance of time series data.

2.2 Time Series

What is time series? According to Spiegel and Stephens (1999), "a time series is a set of observations taken at specific times, usually at equal intervals." Mathematically, a time series is defined by the values Y_1, Y_2 of a variable (temperature, closing price of share, annual population growth, etc) at times t_1, t_2,......... Thus, Y is a function of t, which can be written as Y = f(t). Examples of time series are total number of barrels of crude oil produced yearly in Nigeria over the years, the monthly temperatures announced by the Meteorological center in Ikeja, etc.

Time series contain certain characteristic movements or variations, some or all of which are present to varying degrees. It is customary to classify the variations into four basic types, often called components of a time series, namely:

(a) Secular trend

(b) Seasonal variations

(c) Cyclical variation and

(d) Irregular variation or Random movements.

2.2.1 Secular Trend

Secular trends are also known as long-term movements and they refer to the smooth or regular movement of the series over a fairly long period of time. The trend of a time series display the general sweep of its development or it characterizes the gradual and consistent pattern of its changes. Some series recorded over a given period of time show an upward trend, some series show a downward trend, while others remain more or less at a constant level.

2.2.2 Seasonal Variation

This type of variation consists of regularly repeating patterns whose movement is due to recurring events that take place annually, such as a sudden increase of Department Store sales before Christmas or Sallah. It is also used to indicate any kind of variation which is of a periodic nature and where the period is not longer than one year. An example of seasonal pattern is the number of passengers traveling on Nigerian roads where increase in traffic is experienced during public holidays. Another example of seasonal variation is the effect of seasonal factors on the demand for currency. One of the main objectives of the Central Bank of Nigeria is to provide the country with a supply of currency, which would expand and contract in accordance with public needs. In a typical year, the amount of currency in circulation changes frequently and substantially. The demand for currency increases just before Sallah and Christmas when people want money for shopping and gifts. After the holidays, the traders and others who have collected money deposit it in the banks, which in turn deposit it with the Central Bank of Nigeria.

2.2.3 Cyclical Variations

These refer to the long-term oscillations, or swings, about a trend line or curve. These cycles, as they are sometimes called, may or may not be periodic; that is, they may or may not follow exactly similar patterns after equal intervals of time. In business and economic activities, movements are considered cyclic only if they recur after intervals of more than one year. An important example of cycle movements are the so-called business cycles representing intervals of prosperity, recession, depression, and recovery.

2.2.4 Irregular Variations

Irregular variations or erratic, fluctuations of a time series are those variations which are either completely unpredictable or which are caused by isolated special occurrences as floods, strikes, etc. Although it is assumed that such events produce variations lasting only a short time, it is conceivable that they may be so intense as to result in new cyclic or other movements.

2.2.5 Forecasting

According to Harrison (1983) forecasting is to determine or predict when an event will occur or is likely to occur so that appropriate actions may be taken. Forecasting provides an aid to management in that it acts as an integral part of the managerial decision – making process. It also provides a tool which enable management to adopt a more scientific approach to decision – making .

The techniques of forecasting can be applied to a wide variety of situations which include:

(i) Planning of Current Resources: It is necessary to plan for production, distribution, labour requirements, machine capacities, etc in order to utilize resources efficiently. Such forecasting is usually concerned with the disposition of resources on short term basis.

(ii) Material Requirements: The future requirement of materials can be determined using forecasting techniques. For example, when a component is used on a number of products which are subject to random fluctuations, it is not easy to calculate how many components are likely to be used over the next few months or so.

(iii) Long Term Planning: All organizations are required to determine their long term goals, at least to some extent.

Many interrelated factors like market opportunities, financial considerations and technological advancements have effect on long term planning decisions. These factors require ability, not only in the manipulation of the mathematics of forecasting techniques, but more importantly, in the ability to understand and use the predictive powers of the many different techniques available to management. The three areas mentioned above are typical of the short, medium and long term requirements faced by many of today's business and commercial organizations.

2.3.1 Choice Of Forecasting System

The results of a forecasting system can never be 100% accurate. Hence, according to Harrison (1983), "little is gained by evaluating forecasting systems simply by their results". A good forecasting system should have the following three qualities:

(i) Small Error Size: A key measure of how suitable a forecasting system is for a particular application is the average size of the error. Instead of using the standard deviation as the measure of error, it is more usual to calculate a mean absolute deviation, which is simply the average of the errors between the forecasted and actual results, taken without regard to their signs.

(ii) Robustness: A good forecasting system should not be sensitive to the changes exhibited by the variability of the pattern of the data.

(iii) Easy to Use and Understand: Management may not to know the details of the mathematical calculations involved in designing a forecasting system, it is essential that an appreciation of the principles upon which the forecasts are based is known and understood. Not much is gained in having a sophisticated forecasting system that is not easy to use or difficult to understand and interpret.

2.3.2 Forecasting Techniques

Several forecasting techniques have been developed to handle different types of situations which may vary in many dimensions like different time horizons, different causes and effects, and different data patterns. These techniques can be divided into two main groups, namely: Qualitative and Quantitative methods of forecasting.

(i) Qualitative method of forecasting: Qualitative methods of forecasting are used when (a) the pattern of data indicated by the past cannot be assumed to continue into the future (b) there is no information relating to the past. The use of qualitative forecasting models does not always provide a step-by-step procedure. Rather, the use of such models varies with each situation and requires a considerable commitment of time on the part of the model –builder to understand the situation and to adapt both the method and the forecasting model so that it accurately reflects the problem. The processor of facts, knowledge and information is the decision-maker rather than a set of mathematical rules Harrison (1983). Some of the most commonly used qualitative forecasting models are: Subjective assessment method and the Delphi approach.

In subjective assessment, little detail is given to the steps taken by the individual or group in arriving at decisions. Usually, the emphasis is on obtaining predictions conforming to certain formats which can then be added and integrated into the overall decision-making process. The basic idea of subjective assessment is to make efficient use of the knowledge and experience of the decision-maker in processing very diverse pieces of information.

On the other hand, the Delphi approach is a method for obtaining a refined consensus from a group, whether it is for a forecast or for some other type of estimate. The objective is to obtain a reliable consensus of opinion from the group of forecasters. The main disadvantage of the Delphi method include low level of reliability and possible ambiguity in the questionnaire.

(ii) Quantitative Methods of Forecasting: Quantitative methods of forecasting can be used when (a) past information is available (b) It is possible to quantify the past information into suitable data form (c) When the assumption is made that the pattern of the past will continue into the future.

Quantitative techniques can be divided into two types, namely: Naïve or intuitive method and formal quantitative methods based on statistical principles.

The Naïve or intuitive method is likely to be based on empirical experience and uses trend or seasonal extrapolation (Harisson 1983:99). This method, though simple to use, have the disadvantage of being significantly less robust. The actual methods and techniques used are not transportable and are most likely to be industry or product –specific.

Formal quantitative methods are based on a systematic approach which attempts to minimize the error inherent within the forecast.

2.2.5 Time Series Analysis

According to Spiegel and Stephens (1999:435) time series analysis consists of a mathematical description of the component movements present. The time series variable Y is an operation of the variables T,C,S, and I that produce the trend, cyclic, seasonal and irregular movements respectively. Symbolically, Y=T*C*S*I. Thus, time series analysis amounts to investigating the factors C,S and I and is often referred to as decomposition of a time series into its basic component movements. Spiegel and Stephen (1999) assumed the operation to be a product. Some statisticians however prefer to consider * as a sum. In practice, the interpretation of the operation (*) preferred depends on the degree of success achieved in applying it.

Estimation of the trend: Spiegel and Stephens (1999) discussed 4 ways of estimating the trend. They are: least squares, freehand, moving average and semi-averages.

The least squares method can by used to find the equation of an appropriate trend line or trend curve from which the trend values T can be computed.

The Freehand method consists of fitting a trend line or curve simply by looking at the graph from which the trend, T can be estimated. An obvious drawback of this method is that it depends too much on individual judgment.

The Method of Moving –Average uses moving averages of appropriate orders to eliminate cyclic, seasonal, and irregular patterns, thus leaving only the trend movement. One disadvantage of this method is that the data at the beginning and end of a series are lost. Secondly, moving averages may generate cycles or other movements that were not present in the original data. A third disadvantage is that moving averages are strongly affected by extreme values. A weighted moving average with appropriate weights could be used to overcome this problem. In such a case, the central item or items are given the largest weight, and extreme values are given small weights.

In using the Method of Semi-averages, the data is separated into two parts (preferably equal) and averaging the data in each parts, thus obtaining two points on the graph of the time series. The trend value can be determined from the trend line after drawing a line between the points. Although this method is simple to apply, it may lead to pro results especially when indiscriminately used. It is applicable only where the trend is linear or approximately linear.

2.4.2 Estimation of the Seasonal Variations

Estimating how the data in a time series vary from month to month throughout a typical year can enable one to determine the seasonal factor. A set of numbers showing the relative values of a variable during the months of the year is called a seasonal index for the variable. For example, if it is known that sales during January, February, March, etc, are 50, 120, 90.... percent of the average monthly sales for the whole year, then the numbers 50, 120,90,.. . provide seasonal index for the year, they are sometimes called seasonal index numbers (Spiegel and Stephens: 1999).

Seasonal index can be computed using any of the following methods:

(a) The Average – percentage Method: In this method, the data for each month is expressed as a percentage of the average for the year. The percentages for corresponding months of different years are then averaged using either a mean or a median. If the mean is used, it is best to avoid any extreme values that may occur. The resulting 12 percentages give the seasonal index. If their mean is not 100% (i.e. if the sum is not 1200%), they should be adjusted by multiplying them by a suitable factor.

(b) The percentage Trend or Ratio – to-Trend Method. In this method, the data for each month is expressed as a percentage of monthly trend values. An appropriate average of the percentages for corresponding months then gives the required index.

(c) The percentage Moving-Average, or Ratio-to-Moving Average, Method: In this method a 12 – month moving average is computed. Since the results thus obtained fall between successive months instead of in the middle of the month (which is where the original data fall), we compute a 2 – month moving average of this 12-month moving average. The results is often called a 12-month centered moving average.

Deseasonalization of Data: Dividing the original monthly data by the corresponding seasonal index numbers results in a data said to be deseasonalized, or adjusted for seasonal variation. Trend, cyclic and irregular movements may still be included in the data.

2.4.3 Estimation of Cyclic Variations

Data that have been deseasonalized can also be adjusted for trend simply by dividing the data by the corresponding trend values. The process of adjusting the equation Y =TCSI for seasonal variation and trend corresponds to dividing Y by ST which yields CI (the cyclic and irregular variations). To smooth out the irregular variations I and to leave only the cyclic variations C, an appropriate moving average of a few months' duration is computed. The cyclic variation can then be studied in detail after isolation.

2.4.4 Estimation Of Irregular Variations

The estimation of the irregular (or random) variations can be done by adjusting the data for the trend, seasonal and cyclic variations. This can be achieved by dividing the original data Y = TCSI by T, S and C to yield I. Irregular movements tend to have a small magnitude and often tend to follow the pattern of a normal distribution in practice.

2.4.5 Comparability Of Data

The need arises at times for the Investigator to adjust the data using his discretion. Hence care must be exercised when comparing, for example, March data with February data since March has 31 days, while February has 28 or 29 days. Also, the number of working days during various months of the same or different years may differ because of holidays, strikes, layoffs, etc.

2.5 Summary

The forecasting of time series involves generalizations, predictions, estimations and decisions. The Chief Executive of an institution must project past experience. Thus, he has to make predictions of future levels of activity basing his forecasts, at least in part, on what has happened in the past. If the methods are intelligently used – with full appreciation of the possibilities and in full recognition of their limitations – they can provide the Chief Executive with invaluable aids in reaching well – informed and wise decisions.

CHAPTER THREE

System Design And Analysis

3.0 Introduction

The Mathematical analysis of data, coupled with experience and other factors, has proved valuable both in long-range and short – range forecasting. In the course of this work, necessary steps were taken to follow the fundamental steps in time-series analysis, namely: collection of data and making sure that it is reliable, graphing the data, constructing the long-term trend, curve or line and making forecasts.

3.1 Design Of The Study

A close look at the managerial aspect of the principal's duties shows that "being the Chief Financial Controller of the College", she is directly in charge of "Budgeting, Revenue Collection, Expenditures, purchases and supplies, etc" Senior Staff Manual of FGGC, Zaria (1992:93). She is expected to send the school's budget in advance to the Federal Ministry of Education, Headquarters in Abuja. To present a budget she needs to work with a reliably projected population of students for the incoming year, predicting the expected total number of students, expected revenue to be collected (in terms of tuition fees, boarding fees, school levies, etc), expected amount to be expended on each student (in terms of feeding, medical care, etc.) and so on. Thus, a knowledge of the students' yearly population for the past years is necessary to be able to project into the future.

3.2 Procedure For Data Collection

The main source of data for this project was the Administrative Office of the College. Going through available records, the student yearly population for ten years (1994-2003) was recorded. To ensure the reliability of the data, I compared them with the principal's copy displayed in the office – no difference.

3.3 Graph of The Time Series

After obtaining the data, the values were plotted on a suitable graph sheet with a scale of 2 cm to 50 units on the Y-axis and 1cm to 1 unit on the x-axis.

The graphical presentation was intended to bring out relevant features or to convey specific ideas. Time (year) was measured along the horizontal axis, letting equal subdivisions represent successive years, while population was measured along the vertical axis.

3.4 Summary

Statistical information can reach the target audience more easily if presented in tables, graphs and pictorial form. The exact number of years to be used, vary from one study to the next, depending on the availability of data and the particular period that happens to be of interest.

CHAPTER FOUR

Software Development And Implementation

4.1 Introduction

This chapter deals with the presentation, organization and analysis of data in order to obtain answers to the research questions posed in chapter one. Kerliger (1964:134) stated that: "A primary purpose of statistics, is to manipulate and summarise numerical data and to compare the obtained results with chance expectations". The analysis of data refers to the techniques of extracting from the data information that was not previously obvious. In this chapter, it is also intended to illustrate how scientific method can be applied to modern management.

4.2 Demographic Data

The table below shows the yearly population of students in Fed. Govt. Girls' College, Zaria, for the past ten academic sessions.

X (Year)	Y (Population)
1994	807
1995	890
1996	989
1997	1081
1998	1203
1999	1302
2000	1365
2001	1396
2002	1406
2003	1450

Table 4.2

4.3 Data Plot

The curve obtained, seem to have the shape of an elongated S, a shape which is neither linear nor parabolic. The curve seems to describe the growth of the series which, while increasing, seem to approach some maximum value as a limit. Although, the growth continues, it does so at a decreasing rate. The Gompertz curve seem to describe the above characteristics (Freud, 1976:240).

4.4 **Equation of Curve**

The Gompertz equation is given by $y = ka^{b^x}$ ------------------ (1)

The equation has 3 constants, k, a and b, where x is the independent variable and y the dependent variable. Thus if we let $w = a^b$, then equation (1) becomes the more familiar exponential equation:

$y = kw^x$ --------------------------- (2)

According to Harrison (1983:103), exponential smoothing places most emphasis on the latest data and a decreasing emphasis on past data. If this principle is followed, then the immediate past two years' data could be used to predict the next year's data. For example, if we are to predict the student population for the year 1996, we could use the population figures for 1995 and 1994 as follows:

$Y = kw^x$

When x = 1994, y, 807 and when x = 1995, y = 890

$Y = kw^x$

$807 = kw^{1994}$ ------------------- (3)

Also $890 = kw^{1995}$ ------------------- (4)

From (3), $k = \frac{807}{W^{1994}}$

Substituting into equation (4)

$890 = \frac{809}{w^{1994}} . w^{1995}$

$890 = 807\,w$

$$w = \frac{890}{807}$$

$$k = \frac{807}{(890/807)^{1994}} = 807 \times (807/890)^{1994} = \frac{807 \times 807^{1994}}{(890)^{1994}} = \frac{807^{1995}}{890^{1994}}$$

To forecast the population for 1996,

$$Y = \frac{807^{1995}}{890^{1994}} \text{ X } (890/807)^{1996}$$

$$= \frac{890^2}{807}$$

$$= 982$$

The actual population for 1996 was 989.

Residual = 982 – 989 = -7

The table below shows the fitted population for the other years and their residuals:

Year	Population	Fitted Value	Residual
1994	807		
1995	890		
1996	989	982	- 7
1997	1,081	1099	- 8
1998	1,203	1182	- 21
1999	1,302	1339	+ 37
2000	1,365	1409	+ 44
2001	1,396	1431	+ 35
2002	1,406	1428	- 78
2003	1,450	1416	- 34

Table 4.4

4.5 Forecasting

Having found the equation of a trend, the forecast of future events can be made; a process known as extrapolating from trends, depends on many factors. The basic question is whether the forces that have operated in the past will continue to operate in the future in the same way. The crucial problem is how to assess the impact of such factors as scientific research, government policy, etc. The more distant the future we are trying to predict, the more likely it is that changes will become operative and the more difficult it will be to evaluate their effect.

Using the facts established in 4.4 above, the predicted population figures of the school for the next seven years are displayed in the table.

Year	Fitted Population
2004	1495
2005	1541
2006	1588
2007	1636
2008	1685
2009	1735
2010	1786

CHAPTER FIVE

Summary, Discussions And Conclusion

5.1 Introduction

The main purpose of analyzing data is to extract information and use such information to guide or regulate the activities of an institution. A summary of what was done in this project is given in this chapter. After the discussion, recommendations were made together with suggestions for further project work.

5.2 Summary

The mathematical treatment of data does not in itself solve all problems. However, when coupled with an investigator's common sense, experience, ingenuity and good judgment, mathematical analysis can be useful in forecasting. The over-all problem of forecasting involves a number of disciplines including Statistics, Mathematics, Economics, etc. There are many different forecasting methods and techniques available and the best method cannot be found automatically. As with this project, it is usually necessary to try different approaches with actual data so that the method that works best can be identified.

5.3 Discussions

Each series of data presents its individual problems and characteristics. Fortunately, methods of handling them are available. Generally speaking, if the 'second differencing' of a series is more or less the same, a parabola will provide a good fit (Freud, 1996:415). A 'second order differencing' of the school's yearly population does not yield the same number hence fitting a parabola was not considered.

According to Freund (1976), if a trend departs obviously or subtly from linearity, you should consider fitting a curve of some kind other than a parabola. The graph of the series (students' Population) obviously departs from linearity; it looked like an elongated S. Hence, an exponential curve was considered. In particular, the Gompertz curve was considered because it is one of the most widely used curve to describe growth (Freund 1976). In fact, it has been used extensively by the National Industrial Conference Board of U.S.A. in its important studies on growth patterns of industries, cities and states (Freund, 1976:420).

The population growth of a school can be compared to the growth pattern of a city, at least to an extent. The main activities affecting the growth of a city are birth, immigration and death, which are analogous to fresh admission, mid stream admission and withdrawal / graduation respectively in a school. Hence one could be justified to fit an exponential trend to the population growth of a School. The population of a school depends on many factors like Government policy, environment, awareness as to the need for education, cost, etc. In forecasting the population of a school, it should be noted that, as with other businesses, the more distant the future we are trying to predict, the more likely it is that unforeseen changes will be introduced and the more difficult it will be to assess their effect on data. The Government, for instance, might decide to admit more students in order to score political points, without paying much attention to the feasibility of such a move in terms of educational facilities and infrastructure.

5.4 Conclusion

Time series, according to Melnyk (1974), can be analyzed for three main purposes, namely:

1. Discovery of possible casual relationship e.g. Relationship between output and investment;
2. Comparison of the shapes of fluctuations of time series e.g. the time series production of automobiles has to be considered a leading series with respect to the production of a certain part. If the production of cars have cyclical fluctuations, the parts which have to be replaced like batteries will also have cyclical fluctuations;

3. One does regression analysis for the purpose of forecasting.

This project tows the line of the third purpose of analysis mentioned above. Reliable forecasts cause people to act more rationally and prevent them from overreacting, contributes to more rational decision, helps in decision making concerning expansion. Thus in a school system, time series aids description, predicts estimation, control and simulation.

References

Akin Fapohunda (1996) Introduction To Computer Science. Aflon Limited, Abuja.

Chapman R. J, Oliver E.C. (1990) Data Processing and Information Technology. Dp publications Ltd, London.

Freund J.E., Williams F.J. (1976) Modern Business Statistics Prentice-Hall, New York.

Hall J.(1979) How To Pass Examinations In Data Processing. Cassel Ltd, London.

Harrison P. (1983) Operational Research Mitchell Beazly Publishers, London.

Lucey T. (1991) Management Information Systems. DP Publications Limited, London.

Melnyk M (1974) Principles Of Applied Statistics Pergamon Press Inc. New York.

Murray R.S, Larry J.S. (1999) Statistics Schaum's Outline Series, McGraw-Hill, New York.

APPENDIX

PROGRAM TO EXTRAPOLATE POPULATION

```
CLEAR
USE A1
? " This is student population projection of
F.G.G.C, Zaria"
?
"===================================================
==="

? " "
? "  YEAR   POPULATION   FITTED VALUE
RESIDUAL  "
SET TALK OFF
FVL=0
STORE 807 TO A1
STORE 890 TO A2
STORE 989 TO A3
STORE 1081 TO A4
STORE 1203 TO A5
STORE 1302 TO A6
STORE 1365 TO A7
STORE 1396 TO A8
STORE 1506 TO A9
STORE 1450 TO A10
*A11=A10^2/A9
DO WHILE .NOT. EOF()
MYEAR=YEAR
MPOPULATION=POPULATION

IF MYEAR=1994
   FVL=((A2)^2)/A1
    RS=FVL-A1
    RS=" "
    FVL=" "
ELSE
IF MYEAR=1995
   FVL=((A3)^2)/A2
    RS=FVL-A2
    RS=" "
    FVL=" "
ELSE
```

```
IF MYEAR=1996
   FVL=((A2)^2)/A1
    RS=FVL-A3
ELSE
IF MYEAR=1997
    FVL=((A3)^2)/A2
    RS=FVL-A4
ELSE
IF MYEAR=1998
    FVL=((A4)^2)/A3
    RS=FVL-A5
ELSE
IF MYEAR=1999
    FVL=((A5)^2)/A4
    RS=FVL-A6
ELSE
IF MYEAR=2000
    FVL=((A6)^2)/A5
    RS=FVL-A7
ELSE
IF MYEAR=2001
    FVL=((A7)^2)/A6
    RS=FVL-A8
ELSE
IF MYEAR=2002
    FVL=((A8)^2)/A7
    RS=FVL-A9
ELSE
IF MYEAR=2003
    FVL=((A9)^2)/A8
    RS=FVL-A10
ELSE
IF MYEAR=2004
    A11=((A10)^2)/A9
    MPOPULATION=A11
    FVL= " "
    RS=" "
ELSE
IF MYEAR=2005
    A12=((A11)^2)/A10
    MPOPULATION=A12
    FVL=" "
    RS= " "
```

```
ELSE
IF MYEAR=2006
    A13=((A12)^2)/A11
    MPOPULATION=A13
    FVL=" "
    RS=" "
ELSE
IF MYEAR=2007
    A14=((A13)^2)/A12
    MPOPULATION=A14
    FVL=" "
    RS =" "
ELSE
IF MYEAR=2008
    A15=((A14)^2)/A13
    MPOPULATION=A15
    FVL=" "
    RS=" "
ELSE
IF MYEAR=2009
    A16=((A15)^2)/A14
    MPOPULATION=A16
    FVL=" "
    RS=" "
ELSE
IF MYEAR=2010
    A17=((A16)^2)/A15
    MPOPULATION=A17
    FVL=" "
    RS=" "
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
```

```
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
? MYEAR,  MPOPULATION,          FVL,          RS

SKIP

ENDDO
```