

TITLE PAGE

# **COMPUTERIZATION OF REGRESSION AND CORRELATION ANALYSIS**

**CASE STUDY: INTERDEPENDENCY OF SOME STOCK MARKET  
INDICATORS TO MARKET CAPITALISATION ON THE NIGERIAN  
STOCK EXCHANGE**

**BY**

**MOHAMMED ABDULRAZAK  
PGS/MCS/2004/2005/1168**

**A PROJECT SUBMITTED TO THE DEPARTMENT OF  
MATHEMATICS/COMPUTER SCIENCE**

**IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD  
OF POSTGRADUATE DIPLOMA IN COMPUTER SCIENCE**

**FEDERAL UNIVERSITY OF TECHNOLOGY  
MINNA, NIGER STATE**

**JANUARY 2007**

## CERTIFICATION

This is to certify that this project is the original work of Mr. Mohammed Abdulrazak (PGS/MCS/2004/2005/1168). Submitted for the award of Post Graduate Diploma in Computer Science Department of Federal University of Technology Minna, Niger State.

---

Mrs. Cole A. T.  
Project Supervisor

---

Date

---

DR N. I. Akinwande  
Head of Department

---

Date

---

External Examiner

---

Date

## DEDICATION

This project is dedicated to the glory of almighty God. Who in his infinite mercy granted the wisdom, strength and finances toward the completion this tasking work.

## ACKNOWLEDGEMENT

My sincere appreciation goes to my able project supervisor, Mrs. Cole A. T., who despite her family and tight schedule, thoroughly scrutinize this project work to its completion.

My gratitude is also due to Head of Department, Mathematics and Computer Science, DR N. I. Akinwande, Abdulrahman Ndanusa, PGD coordinator, and other able lecturers of the Department.

Behind a successful man, they say, is a woman. My profound appreciation is due to my adoring wife who during the period of the course and project was patient, supportive and encouraging.

It is pertinent to mention the memorable time expended together with my colleagues (PGD part-time 2004/2005), Mr. Kayode (class captain), Mr. Ustin, Mr. John, Bros and Mrs. Blessing. I wish every one of them a wonderful break thought as we transit through life

## ABSTRACT

This project work uses BASIC programming language to write and run linear multiple regression (involving two independent variables) and correlation coefficients to determine the strength of relationship between variables involved. Data on growth in new issues and stock index was examined to ascertain their contribution to the growth in the market capitalization (dependent variable).

The multiple regression fitted with a sampled data showed that stock index contributed higher value than new issues to the growth of market capitalization. Positive and high partial correlation coefficients were also observed between the variables, though a weak and negative correlation was obtained in respect of new issues and stock index.

The test on the overall significance of regression estimates was significant at 5%.

# TABLE OF CONTENT

<b>Certification</b>	<b>i</b>
<b>Dedication</b>	<b>ii</b>
<b>Acknowledgement</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>Table of Content</b>	<b>v</b>
<b>CHAPTER ONE</b>	
<b>INTRODUCTION</b>	
Introduction	1
Data analysis	2
Computerization and data analysis	3
Historical background	3
Aims and objectives	5
Scope and limitation	6
<b>CHAPTER TWO</b>	
<b>LITERATURE REVIEW</b>	
Regression models	7
Correlation coefficient analysis	11
Significance of regression model	12
<b>CHAPTER THREE</b>	
<b>SYSTEM ANALYSIS AND DESIGN</b>	
Program design	14
Design element	14
Input description	14
Program design algorithm	15
Choice of program language	16
Program command and statements	17
<b>CHAPTER FOUR</b>	
<b>SYSTEM DEVELOPMENT</b>	
Problem implementation	20
Problem specification	20
Practical operation of the program	20
System justification	22
Output/sample results	22

**CHAPTER FIVE**  
**SUMMARY/RECOMMENDATION**

Summary/conclusion	24
Recommendations	25
References	27

Appendix A	Flowchart
Appendix B	Program
Appendix C	Sample Result

## **CHAPTER ONE**

### **1.0 INTRODUCTION**

In today's global business and economic environment, vast amount of statistical information are available. The most successful government, managers and decision makers are the ones who give priority and invest a great deal of their resources towards understanding the behaviors of data, utilizing them effectively for organizational goals.

A sensitive and most emerging segment of the financial system; the capital market for instance, is worth examining. Specifically, the swing in stock prices (on the secondary market) and the rate of expansion in the new issuance of securities are apparently recurring activities linked to the growth of the total market value. This call for a substantive statistical investigation to establish the contribution, if it exist, and to what degree the new issues and stock index are related to the market capitalization. Meanwhile, the capital market is the market that facilitates the mobilization and allocation of long-term funds for sustenance and development of industries and government programmes.

In the real world, manual applications of statistical analysis (numerical measures) to complex and diverse information generating from contemporary businesses can be too tedious and sometimes not possible at all to adopt. These processes are faster, error free, highly



accurate and utilize less personnel, space and time when executed with precise computerized procedures.

This project work is adopting Basic programming language to run regression and correlation analysis to ascertain the level of interdependency in growth of new issues and stock index to the growth of total market capitalization on the Nigeria Stock Exchange.

### **1.10 DATA ANALYSIS**

Applications of data analysis and statistical methodology are an integral part of organizations. Modern businesses have more need to predict future operations than did those of the past, when businesses were smaller. Small-business managers often can solve problems simply through personal contact. Managers in large corporations, however, have to deal with problems incorporating interactive elements or constraints or variables with distinct attributes.

Managerial decisions often are based on the relationship between two or more variable. For example, after considering the relationship between advertising expenditures and sales, a marketing manager might attempt to predict sales for a given level of expenditures. Areas of business that adopt statistical methods and techniques include: Industries (Quality control), Econometrics (trend analysis and forecasting), and demography among others.

## **1.20 COMPUTERIZATION AND DATA ANALYSIS**

A Computer is a device that works under the control of stored programs, automatically accepting, storing and processing data to produce information. Computers are very widely used for the creation, manipulation and storage of documents but there is enormous variation in the methods used and in the levels of sophistication. When the computer processes data it actually performs a number of separate functions which include input (accept data), storage (hold data internally before, during and after processing), processing (operation on data held) and output (information for use).

Every functional area of business needs to collect data to gain insight into projections for the future or information for decision making. The organization and measurement of data into useful information, for system and sub-systems are better achieved electronically. Computers play a major and central role in our contemporary economy with broad application in the area of data analysis.

A wide range variety of software, for processing statistical data (packages) is on the increase. These packages assist greatly in the task of data manipulation, storage and presentation with high level of sophistication.

## **1.30 HISTORICAL BACKGROUND**

The Nigerian Stock Exchange (NSE) first stated as the Lagos Stock Exchange (LSE) in 1960 and later became NSE in December 1977.

Thereafter branches of the exchange were established in different parts of the country. The exchange which started with only 19 securities traded on its floor in 1961 now has 288 securities as at end of December 2005.

The capital market is divided into the primary and secondary market. The stock exchange is the secondary side of the market. The primary market, also known as the new issues market is the segment of the market that facilitates the actual issuance of new securities while the secondary market is often described as the re-sale segment of the market; the market facilitates the transfer of already existing securities. On the other hand, the securities in the two markets may be either bonds or equities. Bonds are facilities for raising debts by corporate and government while equities confer ownership rights of a company on the holders.

The strength and worth of a particular stock market is measured by its total market capitalization. This in turn depends greatly on the amount of new securities created and variation in prices of the stock and bonds. The external business environments especially government reforms and policies influence greatly the creation of more securities in the market, and of course, the market forces of demand and supply as well as companies' performances play a vital role in determining the stock value at a point in time.

The NSE on regular bases publishes stock performance in terms of gains, stock index, market capitalization and price/earning ratios e.t.c. These indicators are worth examining as the market shows great potentials.

This project work will look at the extend and strength of the relationship between market capitalization, new issues and stock price index of the NSE.

#### **1.40 AIMS AND OBJECTIVES**

The aim of this project is to develop a computer program for computation and execution of empirical assessment of Regression and Correlation analysis. The developed program will be used to achieve the following objectives:

- i. To fit a multiple regression equation involving two independent variables  $Y = a + bx_1 + cx_2$ , where  $x_1$  and  $x_2$  represent Stock Price Index on the NSE and Value of New Securities floated in the primary market. Total Market Capitalization is the dependent variable (Y).
- ii. To ascertain the degree of Correlation between Total Market Capitalization, new issues and stock index.
- iii. To ascertain the significant of the overall estimated regression parameters using Analysis of Variance (ANOVA).

## **1.50 SCOPE AND LIMITATIONS**

The market capitalization of a company is the product of its fully paid shares and the current market price of the stock at a particular period.

Thus, the Total Market Capitalization of a stock market is the aggregate market capitalization of all listed securities at a point in time. The stock index is a measure of stock market price trend and performance.

The sources of variation in the value of market capitalization are closely tied to variation in companies' shares and well as price fluctuations.

Broadly, the two variations can be influenced by the following;

### **i. Variation in Share**

- Creation of additional shares (new issues)
- Bonus issues
- Government policy e.t.c

### **ii. Price fluctuations**

- Forces of demand and supply
- Company's performance
- Government policy e.t.c

This project study will limit its scope to analysis of regression and correlation considering the contribution of new issues and stock index performance to the growth of total market capitalization between 1985 and 2005.

## CHAPTER TWO

### 2.0 LITEATURE REVIEW

#### 2.10 REGRESSION ANALYSIS

Regression analysis is a method of studying the relationship between two (or more) variables, one purpose being to arrive at a method for predicting a value of the dependent variable. This project will examine a linear regression analysis involving two variables. But first we shall examine a general model, beginning from a simple linear regression analysis.

#### 2.11 REGRESSION MODEL

The equation that describes how the dependent variable is related to independent variable (s) and an error term is called the regression model. This describes  $Y_i = F(X_i)$  and is given as;

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_i X_i + \varepsilon \dots (1)$$

#### 2.12 SIMPLE LINEAR REGRESSION MODEL

In simple linear regression, we use only one predictor variable,  $x$ , to describe the behavior of the dependent variable,  $y$ . Also, the relationship between  $x$  and  $y$  is assumed to be basically linear. When a linear regression model is assumed, we employ a statistical model of the form;

$$y = \beta_0 + \beta_1 x + \varepsilon \dots (2)$$

Where  $\beta_0 + \beta_1x$  is the assumed line about which all values of  $x$  and  $y$  will fall (deterministic portion of the model) and  $\varepsilon$  is the error component (random part of the model). The simple linear regression equation is given by;

$$E(y) = \beta_0 + \beta_1x \dots (3)$$

To have an effective predictor and a model that will enable us to make decisions, the following assumptions are necessary;

- The mean of each error component is zero
- Each error component follows an approximate normal distribution
- The variance of the error component is the same for each value of  $x$
- The error are independent of each other

### **2.13 ESTIMATE OF LINEAR REGRESSION EQUATION (LEAST SQUARE METHOD)**

If the value of the parameters  $\beta_0$  and  $\beta_1$  were known, we could use equation (3) to compute the mean value of  $y$  for a known value of  $x$ . Unfortunately, the parameter values are not known in practice and must be estimated by using sample data. Sample statistics ( $b_0$  and  $b_1$ ) are computed as estimates of the parameters  $\beta_0$  and  $\beta_1$ . The estimated regression equation is given as;  $\hat{y} = b_0 + b_1x$

Where

$$b_1 = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2} \dots (4)$$

and

$$b_0 = \frac{\sum y}{n} - b_1 \frac{\sum x}{n} \dots (5)$$

## 2.14 THE MULTIPLE REGRESSION MODEL

Multiple linear regression offers a method of predicting or modeling the behavior of a particular dependent variable ( $y$ ) using two or more independent (predictor) variables. As in the case of simple linear regression, which uses one predictor variable, the regression coefficient are those that minimize Sum of Square Error (SSE) =  $\sum (y_i - \hat{y}_i)^2$ . In the multiple regression model,  $\beta_0, \beta_1, \beta_2 \dots \beta_i$  are the parameters and  $\epsilon$  is a random variable, from equation (1). The model depict that  $y$  is a linear function of  $x_1, x_2 \dots x_i$ . To use this technique properly, one must pay special attention to the assumptions behind it;

- The regression errors follows a normal distribution with mean ( $\mu$ ) zero and have a common variance ( $\sigma$ )
- The errors are statistically independent

Consequently, the expected value of  $y$  is given as:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_i x_i \dots (6)$$

The estimates (sample statistics) for the values of the parameters  $\beta_0, \beta_1, \beta_2 \dots \beta_i$  are  $b_0, b_1, b_2 \dots b_i$ . These sample statistics provide the following estimated multiple regression equation;



$$\hat{y} = b_0 + b_1x + b_2x_2 + \dots b_ix_i \dots (7)$$

## 2.15 ESTIMATE OF MULTIPLE REGRESSION EQUATION WITH TWO INDEPENDENT VARIABLES

The regression estimates involving to independent variable  $x_1$  and  $x_2$  is given by;  $\hat{y} = b_0 + b_1x_1 + b_2x_2 \dots (8)$

By illustration, the value of the total market capitalization ( $y$ ) is dependent on the number of new issues created ( $x_1$ ) and stock index ( $x_2$ ) i.e.  $y = f(x_1, x_2)$ . For this type of multiple regression analysis, the least squares criterion calls for minimization of Sum of Square Error (SSE) i.e.

$$SSE = \sum (y_i - b_0 - b_1x_{1i} - b_2x_{2i})^2 \dots (9).$$

To minimize (9), we must take the partial derivatives of SSE with respect to  $b_0$ ,  $b_1$  and  $b_2$ . We can then set the partial derivatives equal to zero and solve for the estimated regression coefficients  $b_0$ ,  $b_1$  and  $b_2$ .

Taking the partial derivatives setting them equal to zero provides;

$$\frac{\partial SSE}{\partial b_0} = -2 \sum (y_i - b_0 - b_1x_{1i} - b_2x_{2i}) = 0$$

$$\frac{\partial SSE}{\partial b_1} = -2 \sum x_{1i} (y_i - b_0 - b_1x_{1i} - b_2x_{2i}) = 0, \text{ and}$$

$$\frac{\partial SSE}{\partial b_2} = -2 \sum x_{2i} (y_i - b_0 - b_1x_{1i} - b_2x_{2i}) = 0$$

This leads to the following series of equations;

$$\sum y_i = n b_0 + (\sum x_{1i}) b_1 + (\sum x_{2i}) b_2$$

$$\sum x_{1i} y_i = (\sum x_{1i}) b_0 + (\sum x_{1i}^2) b_1 + (\sum x_{1i} x_{2i}) b_2$$

$$\sum x_{2i} y_i = (\sum x_{2i}) b_0 + (\sum x_{1i} x_{2i}) b_1 + (\sum x_{2i}^2) b_2$$

## 2.16 MULTIPLE COEFFICIENT OF DETERMINATION

The coefficient of determination is used to measure the goodness of fit for the estimated regression equation. This concept is applicable to simple and multiple regression equation. The Sum of Squares Total (SST) can be partitioned into two; Sum of Square due to Regression (SSR) and Sum of Squares due to Error (SSE), that is,  $SST = SSR + SSE$  where

$$SST = \sum (y_i - \bar{y})^2$$

$$SSR = \sum (\hat{y}_i - \bar{y})^2$$

$$SSE = \sum (y_i - \hat{y}_i)^2$$

The multiple coefficient of determination, denoted  $R^2 = SSR/SST$ .  $R^2$  can be interpreted as the proportion of the variability in the dependent variable that can be explained by the estimated multiple regression equation.  $R^2$  multiplied by 100 can be interpreted as the percentage of variability.

## 2.20 CORRELATION COEFFICIENT ANALYSIS

The sample coefficient of correlation denoted by  $r$ , measures the strength of the linear relationship that exists within a sample of  $n$  bivariate data. Its value is given by;

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\text{SQRT} ((x - \bar{x})^2 (y - \bar{y})^2)}$$

The possible range of  $r$  between -1 and 1

## 2.21 PARTIAL CORRELATION COEFFICIENT

A partial correlation coefficient measures the relationship existing between any two variables when all other variables connected with those two are kept constant. The partial correlation coefficient is determined in terms of simple correlation coefficient (as stated above) among the variable involved in a multiple relationship. Partial correlation coefficient involving two variable  $x_1$  and  $x_2$  can be obtained as follows;

Keeping  $x_2$  as constant

$$r_{x_1y.x_2} = \frac{r_{x_1y} - (r_{x_2y})(r_{x_1x_2})}{\text{SQRT}((1 - r_{x_2y}^2)(1 - r_{x_1x_2}^2))}$$

Keeping  $x_1$  as constant

$$r_{x_2y.x_1} = \frac{r_{x_2y} - (r_{x_1y})(r_{x_1x_2})}{\text{SQRT}((1 - r_{x_1y}^2)(1 - r_{x_1x_2}^2))}$$

Keeping  $y$  as constant

$$r_{x_1x_2.y} = \frac{r_{x_1x_2} - (r_{x_2y})(r_{x_1y})}{\text{SQRT}((1 - r_{x_1y}^2)(1 - r_{x_2y}^2))}$$

## 2.30 TESTING FOR SIGNIFICANCE OF REGRESSION STATISTICS

Multiple linear regression modeling procedure does not always result in an accurate and reliable predictor. It is highly important to determine whether the overall model is satisfactory. To determine the adequacy of regression models, two types of test exist; overall significant test (testing all set of predictor variables) and individual significant test (to ascertain the contribution of individual predictor variables).

This project limits its test on the overall significant test. To conduct overall significant test the F-test statistic is used. The F test statistic is computed using Analysis of Variance (ANOVA). The ANOVA table for the computation of F-test statistic is given below;

### ANOVA TABLE

Source of Variation	Degree of Freedom	Sum of Squares	Mean Squares	F-statistic
Regression	K	SSR	SSR/k	MSR/MSE
Error	n - k - 1	SSE	SSE/n-k-1	
Total	n - 1	SST		

n = number of observations and k = number of independent variables.

### 2.31 DECISION RULE

For a multiple regression analysis with two independent variables we are examining the hypotheses;

$$H_0: \beta_1 = \beta_2 = 0$$

$$H_1: \text{at least one of the } \beta\text{'s} \neq 0$$

$$F\text{-statistic} = MSR/MSE$$

F-table value =  $F_{\alpha, v_1, v_2}$  where  $v_1 = k$  and  $v_2 = n - k - 1$  having a right-tail area  $\alpha$ . At 5% significant level considering 21 observation and 2 variables  $F_{\alpha, v_1, v_2} = F_{0.05, 2, 17} = 3.59$ . We will reject  $H_0$ , if  $F > F_{\alpha, v_1, v_2}$ , accept otherwise. If  $H_0$  is rejected we can conclude that at least of the independent variables contributes significantly to the prediction of Y.

## CHAPTER THREE

### 3.0 PROGRAM DESIGN

#### 3.1 DESIGN ELEMENT

The design of a new system can be conveniently divided into the following elements.

**INPUT:** - Considering of input will be influenced greatly by the needs of output e.g the necessity for quick response from the system world determines the need for an online type of input. In this project consideration would be given to design of input layouts.

**OUTPUTS:** - It is necessary to consider what is required from the system before deciding how to set about producing it. These requirements will have become clear as this project progressed.

**FILES:** - This element is very much linked to input and output. Input is processed against the files to produce the necessary output.

**PROCEDURES:** - They are the steps, which unify the whole process, which link everything together to produce the desired output. These will involve both computer and clerical procedure. They will start with the origination with the source document and end with the output document being distributed.

#### 3.2 INPUT DESCRIPTION

A data file is a separate stored in Secondary. It is not attached to any program, but may be used by any number of programs. To avoid

confusion each buffer is numbered (often referred to as a file number) and assigning that number is also part of the opening process.

The only time a file number appears in a program is during the opening process. Once the file number is assigned, only that number refers to the file.

The opening statement will require you to state the file name, any allowable options and the buffer number (file number) assigned to it.

After finishing using a file at the end of a program the file must be close. The closing process writes the current block of the file permanently into secondary storage, and releases the buffer area for other uses. Failing to close a file will result in the buffered data being lost.

In this project four data files namely YEAR, MCAP, N\_ISSUES AND INDEX were created in order to provide a convenient means of storing data sets. Since data file can easily be read and updated by a basic program. The YEAR data was created for yearly value of market capitalization on the NSE.

### **3.3 ALGORITHM REPRESENTATION.**

In order to find computer solution to a problem, one must decide steps to be taken by the computer. This involves working algorithm for the problem. The algorithm is then presented in any of the acceptable forms.

Although other forms of presenting algorithms exist (e.g pseudo codes, N-S diagrams etc), this project employed flowchart as it provides good visual representation and easy appreciation of the logic of the algorithm its representing. After an understanding of the input and output, a rigorous description of these which covers almost all cases were drawn. The flowchart for data entry procedure and Co-efficient of regression estimate procedure were drawn as shown in Appendix 'A'.

### **3.4 CHOICE OF PROGRAMMING LANGUAGE**

Basic is defined as know as the beginner all purpose symbolic instruction codes. Those who are writing a program for the first or those who are new into the program writing mostly use it. It is choosing as the programming language for the following reasons.

- It has only view concepts to learn and digest
- It has a design that facilitates the writing of programs in a style that is accepted as a good programming profile
- It is very easy to implement and compile
- It has a good file management facility
- Is the most commonly used programming language for mini computers and microcomputers system because of it small interpreter and compiler.

### 3.5 PROGRAM COMMANDS AND STATEMENTS

In BASIC, you make three types of entries from the keyboard these are COMMANDS, DATA and PROGRAM STATEMENTS. A Command is an immediate order for the computer to do something. Once the command is given it is executed by the computer. Command do not constitute part of a program rather, they process programs by performing various tasks related the program. Some commands in Basic program are:

**LOADING QBASIC:** - Before you can enter, change or execute a BASIC program, you must load the Basic interpreter/compiler program into your computer memory. When you have booted the system and see C:/>(DOS PROMPT) on the screen then QBASIC is typed and press the Enter key. Basic statement or system command from the keyboard can only be accepted if the Basic interpreter/compiler has been loaded into the memory.

**LIST COMMAND:** - To see all of the program statements that have been entered, we type the LIST COMMAND. The list command has the following variations

LIST	list the entire program
LIST 10	list statements with statement number 10
LIST 10 - 200	list from line 10 - 200
LIST -200	list from the beginning to line to line 200



LIST 200 - list from 200 to the end of the program

The LIST directs the listing to a computer printer instead of the screen.

**DELETE COMMAND:** The delete command erases whole sections of a program from main memory. It has the same variation as for LIST command. DELETE alone erases the entire program from the main memory.

**SAVE:** Main memory is volatile, i.e as soon as you get out of BASIC environment or when the computer is turned off, the program in main memory disappears. To save the current program to a disk (secondary storage media) we use the save command.

**RUN COMMAND:** Once a program has been entered into the memory of the computer for us to follow the instruction the program statement will have to be executed. The command for program execution is RUN.

### **BASIC STATEMENTS**

A BASIC program centres on statements, each statement is written in conformity with the language syntax the pattern or structure of the word order. The order in which the statements are written is of vital importance, since the computer executes a program by embodying the ingredients of sequence. i.e. one statement after the other starting from the first selection, i.e, it skips a sequence of statements; and iteration which implies that it is directed by a loop to repeat a sequence of instructions.

BASIC statements are classified into two namely; executable and non executable statements. An executable statements command the computer to perform some actions or simply to take note of something. A non executable statement on the other hand, does not require the computer to take any action. Below is the function of Qbasic statement use in the program.

**REM** statements (short of Remark) are comment or remark statement that helps you to understand your program. You can use REM at any point in time inside your program to explain what you are doing inside the program or when you feel like giving your program a title. Its use is not compulsory but it is advisable to use it in order to describe what you are doing.

**INPUT** statement enable computer to transfer data or values to variable names. The values could be assigned to variables within the program or inputted while the program is running.

**PRINT** statements enable computer to display or print a constant or the content of a variable name.

**LOCATE** functions allow the movement of a cursor row and column wise, and

**DIMENSSION** statements are used to inform the computer to reserve cells for the element of list in the program.

## **CHAPTER FOUR**

### **4.0 PROGRAM IMPLEMENTATION**

#### **4.1 PROBLEM SPECIFICATION**

The task of this design is to develop a system or program that could accept regression analysis data and generate result on such things as the equation of the model, analyze the variance of the regression coefficient as well as estimate the correlation coefficient.

#### **4.2 THE PRACTICAL OPERATION OF THE PROGRAM**

The program was developed in BASIC programming language to compute the regression and correlation analysis to know if new issues and stock index has any effect on the value of market capitalization on the NSE. The entire program is divided into various subprograms and this can be illustrated in the following steps below.

##### **STEP I CREATING AND STORING OF INPUT DATA FILES**

The two lines (10 - 30) stated the title of the program "THIS PROGRAM USES REGRESSION AND CORRELATION TO ANALYSE THE NIGERIAN STOCK EXCHANGE INDICATORS WRITTEN BY MOHAMMED ABDULRAZAK".

The next two lines (40 -50) dimensioned the variables into their appropriate sizes. These were followed by lines (60 - 250) for creating and storing the four input data files; YEAR for duration from 1985 - 2005, MCAP for the total market capitalization between 1985 and 2005, N\_ISSUES for new issues during the period and INDEX for stock index.

## **STEP 2 THE MULTIPLE REGRESSION LINE FOR THE DATA**

The multiple regression line for the data is given by  $Y = b_0 + b_1x_1 + b_2x_2$ . The line numbers (260 - 1310) were meant to compute the multiple regression line for the data. In achieving this, the original data was sorted out and their means were obtained and the values (constant  $b_0$ ,  $b_1$ ,  $b_2$ ) were equally computed and regression equation is printed using the appropriate formulae.

## **STEP 3 TESTING FOR THE SIGNIFICANCE OF THE PARAMETER ESTIMATES USING ANOVA.**

The test aims at finding out whether the explanatory variables ( $x_1$ ,  $x_2$  i.e New issues and Stock Index) do actually have any significant influence on the dependent variable  $y$  (Market Capitalization). Line numbers 1320 - 2000 were meant to test if there is any significance different between the data. To achieve these, the data obtained were sorted using the binary sorting method. We calculate the sum of squares and mean sum of squares using the appropriate formulae. Lastly, a decision was taken at a 5% level of significance.

## **STEP 4 THE CORRELATION CO-EFFICIENTS BETWEEN Y, X1 AND X2**

The line numbers 2010 - 2200 were meant to find the partial correlation coefficients between  $Y$ ,  $X_1$  and  $X_2$ . We first compute the regression for  $Y$  and  $X_1$  ( $r_{12}$ ),  $Y$  and  $X_2$  ( $r_{13}$ ) and  $X_1$  and  $X_2$  ( $r_{23}$ ) and there relationship were obtained.

Finally, line number 2210 and 2220 were files created are closed and the END of the entire program. The end statement tells the computer that that there is no more statement to execute. The entire program is saved using the name ABDUPROJ.BAS and can be accessed at the QBASIC environment. It is equally listed in Appendix 'B'.

### **4.3 SYSTEM JUSTIFICATION**

The computation procedures for evaluating regression and correlation analysis are necessarily tedious manually, even with the assistance of a hand-operated calculator. One can overcome this with the use of digital computer. The computer only save time and energy but also prevent the inevitable cascade of errors that occur whenever one error such as mis-punching the key of a calculator is committed and provides greater level of accuracy. Although there are full-featured statistical package that can handle regression and correlation analysis problem apart from being expensive they imposes heavy hardware requirements. For example statistical package for social sciences (SPSS) require a 16-bit microcomputer 7million bytes of fixed disc for programs and data. Moreover, in most cases only users with sound statistical background can make use of the available features. This program is design with a view to getting rid of these problems.

### **4.4 OUTPUT EXPECTED/SAMPLE RESULT**

The program is expected to output a set of results as follows: -

- The worksheet for regression equation, products, squares, deviations, squares of deviations, covariance, the Regression model for future forecast, estimate of Y, residuals, sums and sum of squares.
- Statement of the Hypothesis, Decision rule, ANOVA table and Conclusion for the significance of the regression estimates.
- The strength of linear relationship that may exist between new issues, stock index recorded and the value of market capitalization for the period under study.

The results obtained from the execution of the program are given in the sample results at Appendix "C"

## CHAPTER FIVE

### 5.0 SUMMARY/CONCLUSION

In our present industrial society, efficiency of production of wealth depends heavily on various kinds of automation. Computers are special in this regard as they automate many methods of processing information. Statistical tools, like the regression and correlation, involving two or more independent variable cannot be easily modeled and evaluated manually. Consequently, this project adopt BASIC programming to fit a multiple regression equation (involving two independent variable), obtained the partial coefficient of correlations and conducted test on the overall significant of regression statistics.

The multiple regression fitted shows that the second independent variable  $x_2$  (stock index) contribute higher to the value of predictor  $Y$  (market capitalization) given it positive coefficient value of 1.3832 while new issues reduce the growth in market capitalization with 0.0262. The estimate of  $Y$  was thus given by;  $y = -29.6203 - 0.0262x_1 + 1.3832x_2$ .

Similarly the partial coefficients of correlation obtained suggest a positive and high correlation between new issues and market capitalization (0.9967), positive correlation of 0.3759 between new issues and stock index while the strength of relationship between new issues and market capitalization was negative and weak (-0.1766).

Though the test on the overall coefficient of regression carried out was significant, the estimates obtained may not satisfactory serve as a model for market capitalization, because this study did not exhaustively capture the entire contributory variables of the market capitalization. A more extensive study that encompasses more variable may be required in order to obtain a highly reliable model for the system.

### **RECOMMENDATION**

Round the globe, a shift to a capital market base economy from a money market one is intensifying; Nigeria capital market inclusive. To sustain and strengthen the current pace of the Nigerian capital market, continuous assessment and exploration of every element and structure are essential. Though, this project work, considered only, the contributions of new issues and stock index to the growth of the market capitalization, several other factor can be explore with numerous other statistical tools. This will better be achieved if stakeholders, government and corporate individual exploit the following among others;

- Creation and maintenance of well structured databanks on the market
- Involvement of expert system (man and machine) in decision making processes



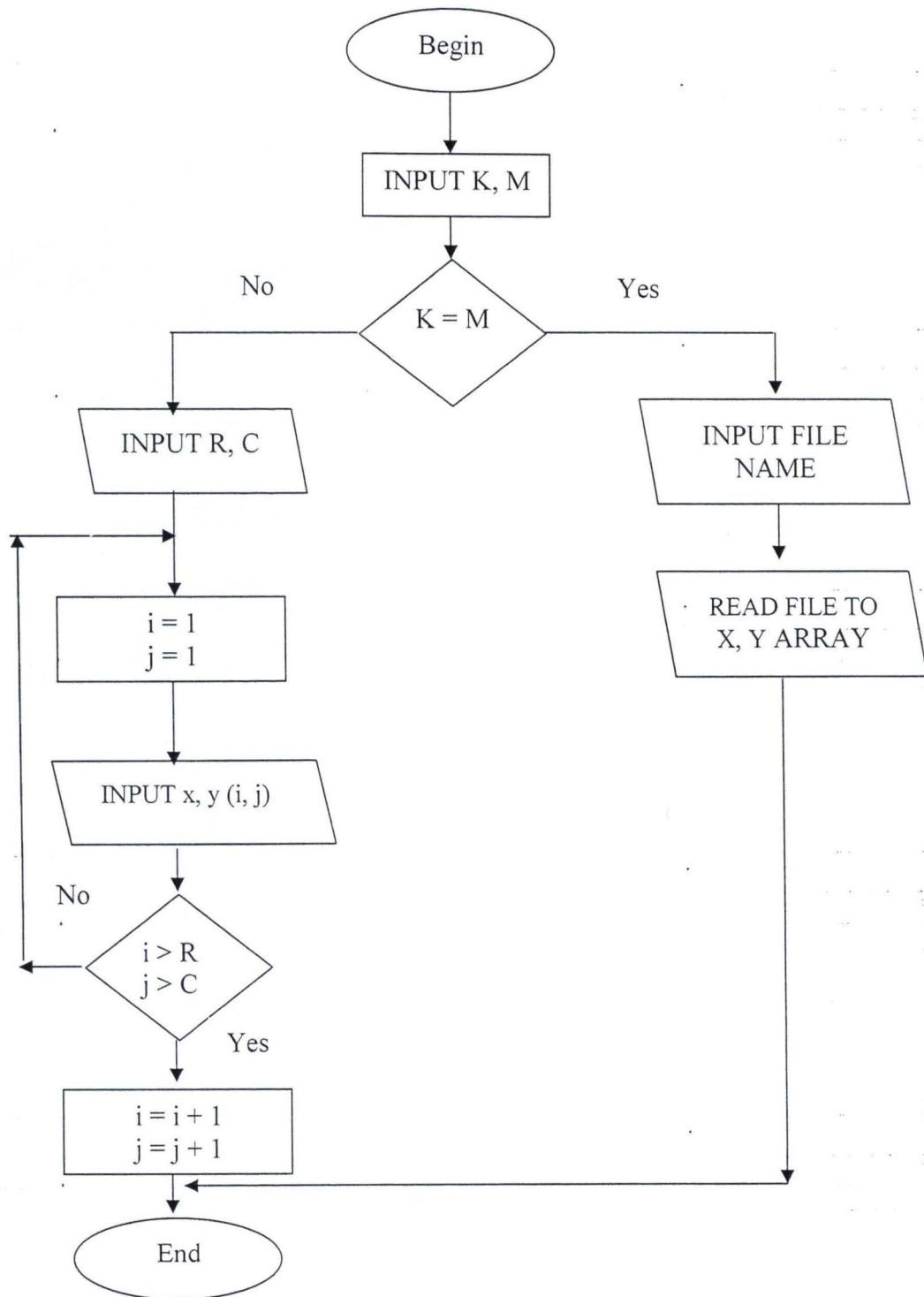
- Incorporating statistical tools and machines for the attainment of a more accurate and reliable projection and estimates
- Strengthening the nation's planning and budgeting processes by aligning the numerical inferential findings with empirical evidence.

## REFERENCES

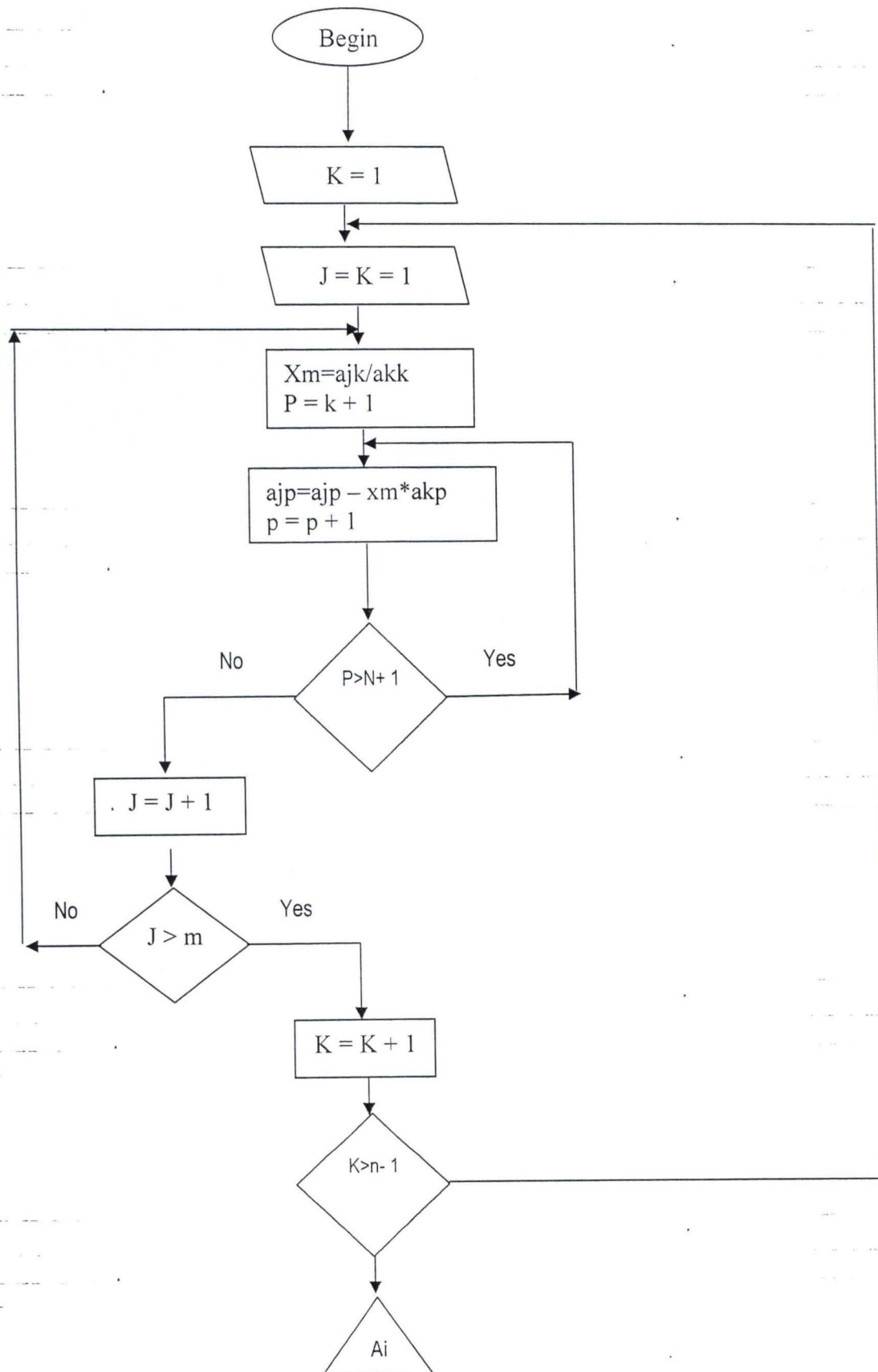
1. Allen H. Kvanli, Robert J. C. Stephen Guynes (2000)  
University of North Texas  
INTRODUCTION TO BUSINESS  
STATISTICS (Fifth edition)  
South-Western College Publishing
2. D. R. Anderson, D. J. Sweeney  
T. A. Williams (1996)  
STATISTICS FOR BUSINESS AND  
ECONOMICS (Sixth edition)  
West Publishing Company
3. C. S. French (1999)  
DATA PROCESSING AND  
INFORMATION TECHNOLOGY  
(10<sup>th</sup> edition)  
Letts Educational Ltd
4. Jiya Mohammed (2004)  
Dept. of Math/Computer  
Science, FUT, Minna  
COMPUTER PROGRAMMING  
(BASIC LANGUAGE)  
Unpublished
5. Badmos R. O.  
Dept. of Math/Computer  
Science, FUT, Minna  
SYSTEM ANALYSIS AND DESIGN  
Unpublished
6. Research and Market  
Development Department  
A PRIMER ON THE CAPITAL  
MARKET (Reviewed edition)
7. COLLECTION OF DIRECTOR GENERAL'S PAPERS  
Securities and Exchange Commission

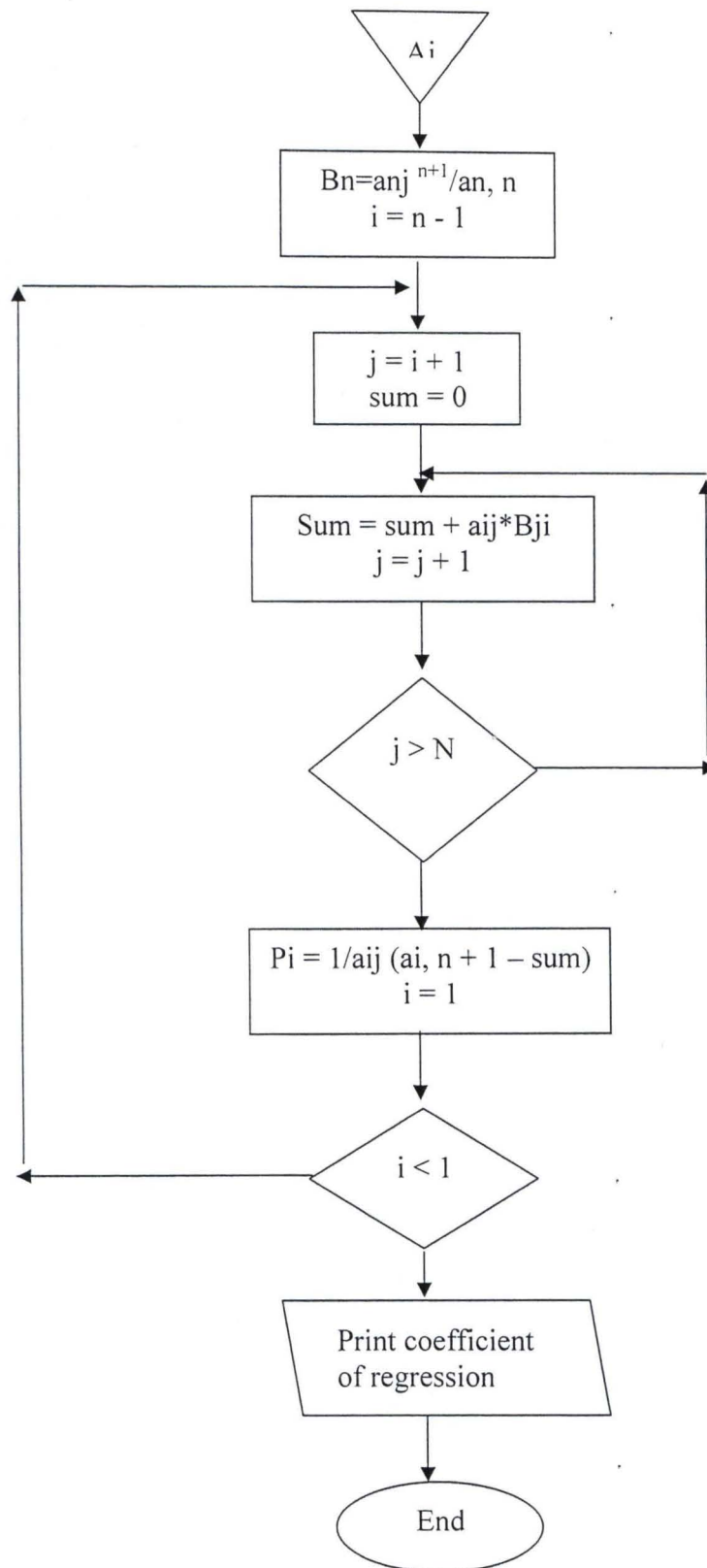
APPENDIX A

DATA ENTRY PROCEDURE FLOWCHART



# COEFFICIENT OF REGRESSION ESTIMATE PROCEDURE





## APPENDIX B

```

10 REM THIS PROGRAM USES REGRESSION AND CORRELATION TO ANALISE THE NIGERIAN
20 REM STOCK MARKET INDICATORS
30 REM WRITEN BY MOHAMMED ABDULRAZAK
40 DIM X1(500), Y(500), X2(500), S(500), YR(500), E(500), F(500), G(500)
50 DIM L(500), M(500), N(500), X1Y(500), X2Y(500), X12(500), X22(500), X1X2(500)
   DIM P(500), Q(500), R(500), YCAP(500), U(500), V(500)
60 KEY OFF: CLS : LOCATE 12, 35: INPUT "ENTER NUMBER OF YEARS"; YN
70 OPEN "YEARS" FOR OUTPUT AS #1
80 FOR I = 1 TO YN
90 INPUT YR(I)
100 NEXT I
110 CLOSE #1: OPEN "MCAP" FOR OUTPUT AS #2
120 CLS : LOCATE 23, 1: PRINT "READING THE YEARLY MARKET CAPITALISATION DATA
   WAIT..."
130 FOR J = 1 TO YN
140 INPUT Y(J)
150 NEXT J
160 CLOSE #2: OPEN "N_ISSUES" FOR OUTPUT AS #3
170 CLS : LOCATE 23, 1: PRINT "READING THE YEARLY NEW ISSUES DATA WAIT..."
180 FOR K = 1 TO YN
190 INPUT X1(K)
200 NEXT K
210 CLOSE #3: OPEN "INDEX" FOR OUTPUT AS #4
220 CLS : LOCATE 23, 1: PRINT "READING THE YEARLY STOCK INDEX DATA WAIT..."
230 FOR I = 1 TO YN
240 INPUT X2(I)
250 NEXT I
260 REM FIND THE MULTIPLE REGRESSION EQUATION  $Y = b_0 + b_1x_1 + b_2x_2$ 
270 SUMX1 = 0: SUMX2 = 0: SUMY = 0:
280 FOR I = 1 TO YN
290 SUMX1 = SUMX1 + X1(I): SUMX2 = SUMX2 + X2(I): SUMY = SUMY + Y(I)
300 NEXT I
310 SUMX1Y = 0: SUMX2Y = 0: SUMX1X2 = 0: SUMX12 = 0: SUMX22 = 0
320 FOR I = 1 TO YN
330 X1Y(I) = X1(I) * Y(I): X2Y(I) = X2(I) * Y(I): X1X2(I) = X1(I) * X2(I)
340 X12(I) = X1(I) ^ 2: X22(I) = X2(I) ^ 2
350 SUMX1Y = SUMX1Y + X1Y(I): SUMX2Y = SUMX2Y + X2Y(I)
360 SUMX1X2 = SUMX1X2 + X1X2(I): SUMX12 = SUMX12 + X12(I):
   SUMX22 = SUMX22 + X22(I)
370 NEXT I
380 AB = SUMX1 * SUMX2Y - SUMX2 * SUMX1Y
390 AC = SUMX1 * SUMX22 - SUMX2 * SUMX1X2
400 AD = SUMX1 * SUMX1X2 - SUMX2 * SUMX12
410 REM LET THE MEANS OF Y, X1 AND X2 BE DENOTED BY A, B, AND C RESPECTIVELY
420 A = SUMY / YN: B = SUMX1 / YN: C = SUMX2 / YN: SUME = 0: SUMF = 0: SUMG = 0:
   SUML = 0: SUMM = 0: SUMN = 0
430 FOR I = 1 TO YN
440 P(I) = Y(I) - A: Q(I) = X1(I) - B: R(I) = X2(I) - C: E(I) = P(I) * P(I)
   F(I) = Q(I) * Q(I): G(I) = R(I) * R(I): L(I) = P(I) * Q(I):
   M(I) = P(I) * R(I): N(I) = Q(I) * R(I)
450 SUME = SUME + E(I): SUMF = SUMF + F(I): SUMG = SUMG + G(I):
   SUML = SUML + L(I): SUMM = SUMM + M(I): SUMN = SUMN + N(I)
460 NEXT I
470 b1 = ((SUMN * AB) - (SUML * AC)) / ((SUMN * AD) - (SUMF * AC))
480 b2 = (SUML - (SUMF * b1)) / SUMN
490 b0 = A - (B * b1) - (C * b2)
500 PRINT : PRINT "-----"
510 CLS : LOCATE 1, 3: PRINT "WORKSHEET FOR REGRESSION EQUATION"
520 PRINT "-----"
530 PRINT TAB(5); "YEARS"; TAB(25); "Y"; TAB(45); "X1"; TAB(65); "X2"
540 YR = 1985
550 FOR I = 1 TO YN
560 YR = YR + 1
570 NEXT I

```

```

580 FOR I = 1 TO YN
590 PRINT TAB(3); YR(I); USING "###.##"; TAB(25); Y(I); TAB(45); X1(I); TAB(65);
X2(I)
600 NEXT I
610 PRINT
"-----"
620 PRINT TAB(2); "SUM"; USING "##,###.##"; TAB(25); SUMY; TAB(45); SUMX1; TAB
(65); SUMX2
630 PRINT TAB(2); "MEAN"; USING "###.###"; TAB(25); A; TAB(45); B; TAB(65); C
640 PRINT : PRINT "-----"
650 PRINT "WORKSHEET FOR REGRESSION EQUATION CONTINUED"
660 PRINT "-----"
670 PRINT TAB(5); "YEARS"; TAB(25); "X1Y"; TAB(45); "X2Y"; TAB(65); "X1X2"
680 YR = 1985
690 FOR I = 1 TO YN
700 YR = YR + I
710 NEXT I
720 FOR I = 1 TO YN
730 PRINT TAB(3); YR(I); USING "#,###,###.####"; TAB(25); X1Y(I); TAB(45); X2Y(I);
TAB(65); X1X2(I)
740 NEXT I
750 PRINT
"-----"
760 PRINT TAB(2); "SUM"; USING "#,###,###.####"; TAB(25); SUMX1Y; TAB(45); SUMX2Y;
TAB(65); SUMX1X2
770 PRINT : PRINT "-----"
780 PRINT "WORKSHEET FOR REGRESSION EQUATION CONTINUED"
790 PRINT "-----"
800 PRINT TAB(5); "YEARS"; TAB(25); "(X1)2"; TAB(45); "(X2)2"
810 YR = 1985
820 FOR I = 1 TO YN
830 YR = YR + I
840 NEXT I
850 FOR I = 1 TO YN
860 PRINT TAB(3); YR(I); USING "#,###,###.####"; TAB(25); X12(I); TAB(45); X22(I)
870 NEXT I
880 PRINT
"-----"
890 PRINT TAB(2); "SUM"; USING "##,###,###.####"; TAB(25); SUMX12; TAB(45); SUMX22
900 PRINT : PRINT "-----"
910 PRINT "WORKSHEET FOR REGRESSION EQUATION CONTINUED"
920 PRINT "-----"
930 PRINT TAB(5); "YEARS"; TAB(25); "DIVY"; TAB(45); "DIVX1"; TAB(65); "DIVX2"
940 YR = 1985
950 FOR I = 1 TO YN
960 YR = YR + I
970 NEXT I
980 FOR I = 1 TO YN
990 PRINT TAB(3); YR(I); USING "#,###,###.####"; TAB(25); P(I); TAB(45); Q(I); TAB
(65); R(I)
1000 NEXT I
1010 PRINT
"-----"
1020 PRINT TAB(2); "SUM"; USING "#,###,###.####"; TAB(25); SUMP; TAB(45); SUMQ;
TAB(65); SUMR
1030 PRINT : PRINT "-----"
1040 PRINT "WORKSHEET FOR REGRESSION EQUATION CONTINUED"
1050 PRINT
"-----"
1060 PRINT TAB(5); "YEARS"; TAB(25); "(DIVY)2"; TAB(45); "(DIVX1)2"; TAB(65);
"(DIVX2)2"
1070 YR = 1985
1080 FOR I = 1 TO YN
1090 YR = YR + I
1100 NEXT I
1110 FOR I = 1 TO YN

```

```

1120 PRINT TAB(3); YR(I); USING "#,###,###.####"; TAB(25); E(I); TAB(45); F(I);
TAB(65); G(I)
1130 NEXT I
1140 PRINT
"-----"
1150 PRINT TAB(2); "SUM"; USING "#,###,###.####"; TAB(25); SUME; TAB(45); SUMF;
TAB(65); SUMG
1160 PRINT : PRINT "-----"
1170 PRINT "WORKSHEET FOR REGRESSION EQUATION CONTINUED"
1180 PRINT
"-----"
1190 PRINT TAB(5); "YEARS"; TAB(25); "COV(X1Y)"; TAB(45); "COV(X2Y)"; TAB(65);
"COV(X1X2)"
1200 YR = 1985
1210 FOR I = 1 TO YN
1220 YR = YR + I
1230 NEXT I
1240 FOR I = 1 TO YN
1250 PRINT TAB(3); YR(I); USING "#,###,###.####"; TAB(25); L(I); TAB(45); M(I);
TAB(65); N(I)
1260 NEXT I
1270 PRINT
"-----"
1280 PRINT TAB(2); "SUM"; USING "#,###,###.####"; TAB(25); SUML; TAB(45); SUMM;
TAB(65); SUMN
1290 PRINT
1300 PRINT
1310 PRINT "THE REGRESSION EQUATION Y = "; b0; "+"; b1; "X1"; "+"; b2; "X2"
1320 REM THE SIGNIFICANCE OF REGRESSION EQUATION USING ANALYSIS OF VARIANCE
1330 SUMV = 0
1340 FOR I = 1 TO YN
1350 YCAP(I) = b0 + (b1 * X1(I)) + (b2 * X2(I))
1360 U(I) = Y(I) - YCAP(I)
1370 V(I) = U(I) ^ 2
1380 SUMV = SUMV + V(I)
1390 NEXT I
1400 SSE = SUMV: SST = SUME: SSR = SST - SSE
1410 REM FINDING THE ASSOCIATE DEGREE OF FREEDOM
1420 REM LET DF1, DF2 AND DF3 REPRESENT THE DEGREE OF FREEDOM
1430 REM FOR SST, SSR AND SSE RESPECTIVELY
1440 K = 2
1450 DF1 = YN - 1: DF2 = K: DF3 = YN - K - 1
1460 MSR = SSR / DF2: MSE = SSE / DF3
1470 F1 = MSR / MSE
1480 PRINT : PRINT "-----"
1490 PRINT "ESTIMATE OF Y (MARKET CAPITALISATION)"
1500 PRINT
"-----"
1510 PRINT TAB(5); "YEARS"; TAB(15); "Y"; TAB(30); "X1"; TAB(50); "X2"; TAB(65);
"YCAP"
1520 YR = 1985
1530 FOR I = 1 TO YN
1540 YR = YR + I
1550 NEXT I
1560 FOR I = 1 TO YN
1570 PRINT TAB(3); YR(I); USING "#,###.####"; TAB(15); Y(I); TAB(30); X1(I); TAB
(50); X2(I); TAB(65); YCAP(I)
1580 NEXT I
1590 PRINT
"-----"
1600 PRINT TAB(2); "SUM"; USING "#,###,###.####"; TAB(15); SUMY; TAB(30); SUMM;
TAB(50); SUMN;
1610 PRINT : PRINT "-----"
1620 PRINT "ESTIMATE OF Y AND ERROR RESIDUAL"
1630 PRINT
"-----"

```



```

1640 PRINT TAB(5); "YEARS"; TAB(15); "Y"; TAB(30); "YCAP"; TAB(50); "DIV(ERROR)";
TAB(65); "(DIV(ERROR))2"
1650 YR = 1985
1660 FOR I = 1 TO YN
1670 YR = YR + I
1680 NEXT I
1690 FOR I = 1 TO YN
1700 PRINT TAB(3); YR(I); USING "##,###.####"; TAB(15); Y(I); TAB(30); YCAP(I);
TAB(50); U(I); TAB(65); V(I)
1710 NEXT I
1720 PRINT
"-----"
1730 PRINT TAB(2); "SUM"; USING "#,###,###.####"; TAB(15); SUMY; TAB(65); SUMV
1740 PRINT
"-----"
1750 PRINT "SIGNIFICANT TEST OF THE OVERALL REGRESSION ESTIMATES"
1760 PRINT
"-----"
1770 PRINT "HYPOTHESIS"
1780 PRINT "HO: Bi's = 0 (X1 AND X2 DO NOT CONTRIBUTE TO THE VALUE OF Y)"
1790 PRINT "H1: NOT ALL Bi's ARE EQUAL TO ZERO"
1800 PRINT
1810 PRINT "DECISION RULE"
1820 PRINT "REJECT HO IF FCRITICAL >= F*"
1830 PRINT TAB(10);
"-----"
1840 PRINT TAB(9); "ANALYSIS OF VARIANCE TABLE"
1850 PRINT TAB(3);
"-----"
1860 PRINT TAB(3); "SOURCE OF"; TAB(18); "SUM OF"; TAB(33); "DEGREE OF"; TAB(48);
"MEAN"; TAB(63); "F*"
1870 PRINT TAB(3); "VARIATION"; TAB(18); "SQUARES"; TAB(33); "FREEDOM"; TAB(48);
"SQUARE"
1880 PRINT TAB(3);
"-----"
1890 PRINT TAB(3); "REGRESSION"; TAB(17); SSR; TAB(37); K; TAB(47); MSR; TAB(62);
F1
1900 PRINT TAB(3); "ERROR"; TAB(17); SSE; TAB(37); YN - K - 1; TAB(47); MSE;
1910 PRINT TAB(3);
"-----"
1920 PRINT TAB(3); "TOTAL"; TAB(17); SST; TAB(37); YN - 1
1930 PRINT TAB(3);
"-----"
1940 PRINT
1950 PRINT "CONCLUSION"
1960 IF F1 >= FCRITICAL THEN
1970 PRINT "WE REJECT HO AND CONCLUDE THAT THE VALUE OF MARKET CAPITALISATION IS
SIGNIFICANTLY RELATED TO NEW ISSUES AND STOCK INDEX"
1980 ELSEIF F1 < FCRITICAL THEN
1990 PRINT "WE ACCEPT HO AND CONCLUDE THAT NEW ISSUES AND STOCK INDEX DO NOT
CONTRIBUTE TO THE VALUE OF MARKET CAPITALISATION"
2000 END IF
2010 REM CORRELATION ANALYSIS
2020 REM TO DETERMINE THE DEGREE OF RELATIONSHIP BETWEEN Y, X1 AND X2
2030 r12 = SUML / ((SUMF * SUME) ^ .5)
2040 r13 = SUMM / ((SUMG * SUME) ^ .5)
2050 r23 = SUMN / ((SUMF * SUMG) ^ .5)
2060 REM FIND THE PARTIAL CORRELATION COEFFICIENTS BETWEEN Y, X1 AND X2
2070 r123 = (r12 - (r13 * r23)) / (((1 - (r13) ^ 2) * (1 - (r23) ^ 2)) ^ .5)
2080 r132 = (r13 - (r12 * r23)) / (((1 - (r12) ^ 2) * (1 - (r23) ^ 2)) ^ .5)
2090 r231 = (r23 - (r12 * r13)) / (((1 - (r12) ^ 2) * (1 - (r13) ^ 2)) ^ .5)
2100 PRINT : PRINT "-----"
2110 PRINT "THE CORRELATION COEFFICIENTS REPORTS"
2120 PRINT "-----"
2130 PRINT
2140 PRINT

```

```
2150 PRINT "THE CORRELATION COEFFICIENT BETWEEN Y AND X1 DENOTED r12 = "; r12
2160 PRINT "THE CORRELATION COEFFICEINT BETWEEN Y AND X2 DENOTED r13 = "; r13
2170 PRINT "THE CORRELATION COEFFICIENT BETWEEN X1 AND X2 DENOTED r23 = "; r23
2180 PRINT "THE PARTIAL CORRELATION COEFFICIENT BETWEEN Y AND X1 r123 = "; r123
2190 PRINT "THE PARTIAL CORRELATION COEFFICIENT BETWEEN Y AND X2 r132 = "; r132
2200 PRINT "THE PARTIAL CORRELATION COEFFICIENT BETWEEN X1 AND X2 r231 = "; r231
2210 CLOSE #1, #2, #3, #4
2220 END
```

OUTPUT

APPENDIX C

-----  
 WORKSHEET FOR REGRESSION EQUATION  
 -----

YEAR	Y	X1	X2
1996	625.32	857.6	506.04
1997	624.08	943.26	498.17
1998	617.01	981.31	486.22
1999	631.69	961.47	479.16
2000	689.96	1,004.41	533.05
2001	729.1	1,120.57	568.19
2002	744.56	1,185.30	578.96
2003	821.62	1,379.16	644.78
2004	866.94	1,387.68	663.21
2005	897.93	1,570.55	664.26
SUM	7,248.21	11,391.31	5,622.04
MEAN	724.821	1,139.131	562.204

-----  
 WORKSHEET FOR REGRESSION EQUATION CONTINUED  
 -----

YEAR	X1Y	X2Y	X1X2
1996	536,274.4320	316,436.9328	433,979.9040
1997	588,669.7008	310,897.9336	469,903.8342
1998	605,478.0831	300,002.6022	477,132.5482
1999	607,350.9843	302,680.5804	460,697.9652
2000	693,002.7236	367,783.1780	535,400.7505
2001	817,007.5870	414,267.3290	636,696.6683
2002	882,526.9680	431,070.4576	686,241.2880
2003	1,133,145.4392	529,764.1436	889,254.7848
2004	1,203,035.2992	574,963.2774	920,323.2528
2005	1,410,243.9615	596,458.9818	1,043,253.5430
SUM	8,476,735.1787	4,144,325.4164	6,552,884.5390

-----  
 WORKSHEET FOR REGRESSION EQUATION CONTINUED  
 -----

YEAR	(X1)2	(X2)2
1996	735,477.7600	256,076.4816
1997	889,739.4276	248,173.3489
1998	962,969.3161	236,409.8884
1999	924,424.5609	229,594.3056
2000	1,008,839.4481	284,142.3025
2001	1,255,677.1249	322,839.8761
2002	1,404,936.0900	335,194.6816
2003	1,902,082.3056	415,741.2484
2004	1,925,655.7824	439,847.5041
2005	2,466,627.3025	441,241.3476
SUM	13,476,429.1181	3,209,260.9848

-----  
 WORKSHEET FOR REGRESSION EQUATION CONTINUED  
 -----

YEAR	DIVY	DIVX1	DIVX2
1996	-99.5010	-281.5310	-56.1640
1997	-100.7410	-195.8710	-64.0340
1998	-107.8110	-157.8210	-75.9840
1999	-93.1310	-177.6610	-83.0440
2000	-34.8610	-134.7210	-29.1540
2001	4.2790	-18.5610	5.9860
2002	19.7390	46.1690	16.7560
2003	96.7990	240.0290	82.5760
2004	142.1190	248.5490	101.0060
2005	173.1090	431.4190	102.0560
SUM	0.0000	0.0000	0.0000

OUTPUT

WORKSHEET FOR REGRESSION EQUATION CONTINUED

YEAR	(DIVY)2	(DIVX1)2	(DIVX2)2
1996	9,900.4490	79,259.7040	3,154.3949
1997	10,148.7491	38,365.4486	4,100.3532
1998	11,623.2117	24,907.4680	5,773.5683
1999	8,673.3832	31,563.4309	6,896.3059
2000	1,215.2893	18,149.7478	849.9557
2001	18.3098	344.5107	35.8322
2002	389.6281	2,131.5766	280.7635
2003	9,370.0464	57,613.9208	6,818.7958
2004	20,197.8102	61,776.6054	10,202.2120
2005	29,966.7259	186,122.3536	10,415.4271
SUM	101,503.6027	500,234.7665	48,527.6086

WORKSHEET FOR REGRESSION EQUATION CONTINUED

YEAR	COV(X1Y)	COV(X2Y)	COV(X1X2)
1996	28,012.6160	5,588.3742	15,811.9071
1997	19,732.2404	6,450.8492	12,542.4036
1998	17,014.8398	8,191.9110	11,991.8709
1999	16,545.7466	7,733.9708	14,753.6801
2000	4,696.5088	1,016.3376	3,927.6560
2001	-79.4225	25.6141	-111.1061
2002	911.3299	330.7467	773.6078
2003	23,234.5672	7,993.2742	19,820.6347
2004	35,323.5353	14,354.8717	25,104.9403
2005	74,682.5117	17,666.8121	44,028.8975
SUM	220,074.4732	69,352.7616	148,644.4918

THE REGRESSION EQUATION  $Y = 20.5024 + 0.1701X1 + 0.9082X2$

ESTIMATE OF Y (MARKET CAPITALISATION)

YEAR	Y	X1	X2	YCAP
1996	625.32	857.6	506.04	625.9315
1997	624.08	943.26	498.17	633.3535
1998	617.01	981.31	486.22	628.9726
1999	631.69	961.47	479.16	619.1866
2000	689.96	1,004.41	533.05	675.4308
2001	729.1	1,120.57	568.19	727.1004
2002	744.56	1,185.30	578.96	747.8907
2003	821.62	1,379.16	644.78	840.6381
2004	866.94	1,387.68	663.21	858.8247
2005	897.93	1,570.55	664.26	890.8812
SUM	7,248.21	11,391.31	5,622.04	

ESTIMATE OF Y AND RESIDUAL

YEAR	Y	YCAP	DIV(ERROR)	(DIVERROR)2
1996	625.32	625.9315	-0.6115	0.3739
1997	624.08	633.3535	-9.2735	85.9982
1998	617.01	628.9726	-11.9626	143.1043
1999	631.69	619.1866	12.5034	156.3360
2000	689.96	675.4308	14.5292	211.0989
2001	729.1	727.1004	1.9996	3.9986
2002	744.56	747.8907	-3.3307	11.0936
2003	821.62	840.6381	-19.0181	361.6891
2004	866.94	858.8247	8.1153	65.8588
2005	897.93	890.8812	7.0488	49.6854
SUM	7,248.21			1,089.2367

OUTPUT

SIGNIFICANCE TEST OF THE OVERALL REGRESSION ESTIMATES

HYPOTHESIS

HO:  $B_i$ 's = 0 (X1 AND X2 DO NOT CONTRIBUTE TO THE VALUE OF Y)

H1: NOT ALL  $B_i$ 's ARE EQUAL TO ZERO

DECISION RULE

REJECT HO IF FCRITICAL  $\geq$  F\*

ANALYSIS OF VARIANCE TABLE

SOURCE OF VARIATION	DEGREE OF FREEDOM	SUM OF SQUARES	MEAN SQUARES	F*
REGRESSION	2	100,414.3660	50,207.1830	322.6574
ERROR	7	1,089.2367	155.6052	
TOTAL	9	101,503.6027		

CONCLUSION

WE REJECT HO AND CONCLUDE THAT THE VALUE OF MARKET CAPITALISATION IS SIGNIFICANTLY RELATED TO NEW ISSUES AND STOCK INDEX

THE CORRELATION COEFFICIENT REPORTS

THE CORRELATION COEFFICIENT BETWEEN Y AND X1 DENOTED  $r_{12}$  = 0.9767  
 THE CORRELATION COEFFICIENT BETWEEN Y AND X2 DENOTED  $r_{13}$  = 0.9882  
 THE CORRELATION COEFFICIENT BETWEEN X1 AND X2 DENOTED  $r_{23}$  = 0.9540  
 THE PARTIAL CORRELATION COEFFICIENT BETWEEN Y AND X1  $r_{123}$  = 0.7376  
 THE PARTIAL CORRELATION COEFFICIENT BETWEEN Y AND X2  $r_{132}$  = 0.8760  
 THE PARTIAL CORRELATION COEFFICIENT BETWEEN X1 AND X2  $r_{231}$  = -0.3355