UNIVERSITY OF CAPE COAST

IDENTIFICATION OF OUTSTANDING PERFORMANCES IN SENIOR HIGH SCHOOL EXAMINATIONS: A PRINCIPAL COMPONENT APPROACH

LAWRENCE KWESI AWI

2010

UNIVERSITY OF CAPE COAST

IDENTIFICATION OF OUTSTANDING PERFORMANCES IN SENIOR HIGH SCHOOL EXAMINATIONS: A PRINCIPAL COMPONENT APPROACH

BY

LAWRENCE KWESI AWI

A dissertation submitted to the Department of Mathematics and Statistics of the School of Physical Sciences, University of Cape Coast in partial fulfillment of the requirements for the award of Master of Science Degree in Statistics

JULY 2010

DECLARATION

Candidate's Declaration

I hereby declare that this dissertation is the result of my own original work and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's signature: …………………………… Date: …………………..

Candidate's name: LAWRENCE KWESI AWI

Supervisor's Declaration

I hereby declare that the preparation and presentation of the dissertation were supervised in accordance with the guidelines on supervision of dissertation laid down by the University of Cape Coast.

Supervisor's signature: …………………………… Date: …………………..

Supervisor's name: BISMARK K. NKANSAH

ABSTRACT

The study looks at identification of outstanding and abysmal performances of students in examinations involving seven common subjects taken in the first year at Ghana Senior High School in Koforidua. The scores of the students, who were grouped into six different program classes, in the seven subjects constituted the data for the study. In order to achieve the objective of the study using this high dimensional data set, a multivariate data analysis technique (the Principal Components Analysis) was used as the main statistical tool to examine the variance-covariance structure of the performance in the seven subjects.

The study reveals that the first principal component is a weighted sum of all the subjects offered by the students. As a result, the first component is found to be the most appropriate index in determining the general performances of the students. Core Mathematics and French are observed to be the two most influential subjects in the formation of the first component. Thus, in the determination of general performance of students, Core Mathematics and French are the most influential of all the seven subjects.

Using the scores of the first principal component, it is discovered that the three best students are all members of the General Arts class. The worst scores are recorded by a Science student; the second worst scores are obtained by a student of Home Economics class; and the third worst student was from the Science class. The study also reveals that in general, the Visual Arts class is the strongest class whilst the Agric class is the weakest among the six classes. The performances of Business, General Arts and the Science classes are quite normal.

ACKNOWLEDGEMENTS

In writing this dissertation, I received tremendous assistance from my supervisor, Mr. B.K. Nkansah of the Department of Mathematics and Statistics, University of Cape Coast. He gave general guidance and advice to the successful organization and completion of this work. I am indebted to him for his selfless contributions.

Mention should be made here of Professor B.K. Gordor of the Department of Mathematics and Statistics, University of Cape Coast and Ms Rosemond Bampo, Headmistress of Ghana Senior High School, Koforidua, for their encouragement, advice and support they gave me. To them, I owe a debt of gratitude.

Without the care and support of Mrs. Vida Awi, Mr. Moses Amponsah of Intercontinental Bank and Mrs. Victoria Gordor little could have been achieved in my attempt to write this dissertation. Thus to them also, I am very grateful.

Finally, my sincere thanks go to my son Dallas Justin Awi, of the Regional Maritime University, who typed the manuscript.

DEDICATION

To my wife Adobea and children Juliana, Dallas, Conrad, and Litela

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Background

Years after independence, Ghana's formal education systems (particularly at the basic and secondary levels) continue to reflect the elitist and academic type bequeathed to us by our colonial masters. Various attempts were made to increase access to education by establishing more schools as demanded by 1961 Education Act.

The education system generally, did not undergo any major structural and/or curricular changes to enable majority of both elementary and secondary school leavers to either proceed to the next stage of education or become economically viable in the society. Between 1970 and 1973, only 14 percent of the 64 percent pupils who benefited from primary and middle school education could proceed further to second cycle institutions annually (Antwi, 1992 p.87). This naturally had created a scenario of over-swelling number of unemployed elementary and secondary school leavers who socially and politically were becoming a danger to the society.

In 1974, a new structure and content of education was introduced to address this trend. In the new structure, the Experimental Junior Secondary School was introduced but could not have any sustaining impact on the general education system in the country. This was due, among other factors, to lack of coordinated governmental support and parental anxiety about future prospects of the programme Thus, by the beginning of 1980's, only 118

Experimental Junior Secondary schools were functioning throughout the country.

In 1987, the PNDC Government put another new Education reform in place. The intention of the government was, among other things, to reduce the duration of the pre-university education from seventeen to twelve years. As part of the implementation, Junior Secondary School (J.S.S.) replaced middle form one classes in September 1987 and later was completely phased out in 1990. Under the reform is a system of evaluation, which at the end of the nine years basic education combines individuals' continuous assessment, and their results of performance in the Basic Education Certificate Examination (BECE) to determine which students move to the senior high school level. This led to the opening of more schools both at the primary and secondary levels.

One of the Senior High Schools in Ghana, which is the subject of this study, is Ghana Senior High School. This school, one of the leading schools in the New Juaben District of the Eastern Region of Ghana, admits students based on their performance at the B.E.C.E. to pursue courses in General Arts, Business, Science, Agriculture, Home Economics and Visual Arts. After three years of study in the school, they sit for the Senior Secondary School Certificate Examination conducted by the West African Examination Council. It is a Government Assisted Mixed, non-denominational institution.

Ghana Senior High School was established in 1943 as a private school with an initial population of sixteen students and was called Phoenix College. In 1957, Dr. Kwame Nkrumah, the first president of Ghana, changed the name of the school to Ghana Secondary School.

During 2001 Senior Secondary School Certificate examination, GHANASS, as it is known, produced the overall best candidate as well as the best in the General Arts. This is a feat, which according to the West African Examination Council, is rare.

From a humble beginning, Ghana Senior High School has been one of the leading Senior High Schools in the country. A student population of at least four hundred and fifty passed out of the school each year. Currently it has a population of 1540 students made up of 729 boys and 811 girls.

In 2006 another education policy was put in place by the NPP government that consists of eleven years of pre-university education including four years of Senior High School education. The second batch of students admitted into Ghana Senior High School are reading seven core subjects, English Language, Core Mathematics, Integrated Science, Social Studies, Physical Education, Information and Communication Technology and French. The results of the first terminal examinations are used to group them to begin with their various courses of study in the second year and to award them with various prizes during the school's Speech and Prize Giving Day celebration. After four years study, the students sit for their final year examination leading to the award of the West African Senior High School Certificate. The West African Examination Council conducts the examination. Marks from the examination and assessment that are carried out by a variety of means, including the collection of evidence of routine student performance produced under ordinary classroom conditions, are used for final grading of students.

Objective of the Study

An educational institution, like any other institution, needs some dimensions whereby it measures and evaluates the performance of its staff and students from over a period of time. These dimensions would eventually be the bases for awarding excellence in academic work over the period. The knowledge of these measures would identify the strength and weaknesses of the institution, thereby giving it focus for allocating effort and resources. This study is in this direction to identify an index for assessing the general performance of students of Koforidua Senior High School at the end of the first year. Thus, the main objective of this study is to identify outstanding and abysmal performances of first year students in end of term examinations involving seven common subjects.

Research Questions

In order to achieve the aim of the study, one needs to be guided by relevant questions. The following are the pertinent questions to direct the study of the first year examination scores of Koforidua Senior High School.

1. What is the general academic performance of the students in their first year in the school?
2. Which class is generally the strongest among the six classes?
3. Which class is generally the weakest among the six classes?
4. Who are the three outstanding students and three worst students after one year in the school?
5. Which programmes produced the best and worst students in the examinations?

Data

Data for the research was made up of first year examination scores of students of Ghana Senior High School in Koforidua who are currently in their second year. The examination scores were obtained in seven subjects, namely English Language, Core Mathematics, Integrated Science, Physical Education, Social Studies, Information and Communication Technology and French. Thus, in this work, the seven subjects constitute the variables of study. The scores of each of the students in each of the subjects generated the data for the study. Since each of the students had a score on each of the seven subjects (variables), the data obtained constitute a multivariate data set. The students are initially grouped into six programme classes. These programmes are Agriculture, Business, General Arts, Home Economics, Science and Visual Arts.

The choice of first year scores was because of the important use of these results. As it is done in most schools, the final groupings of students into the various programmes of study in their second year are based on their performance in the first year. The performance in the first year is also used for giving awards to the students for the first time in the school after their admission.

The target population is between the ages of fourteen and seventeen years. The school admits over four hundred and fifty students each year. The batch of students used in this study is the 2008/2009 group and involves four hundred and seventy nine (479) students. The groupings of the students in the various programmes are given in Table 1.

Table 1: Distribution of Students in the Various Programmes

Programme	Number of students	Percentage
Agriculture	44	9.19
Business	48	10.02
General Arts	148	30.90
Home Economics	97	20.25
Science	92	19.21
Visual Arts	50	10.43
Total	479	100.00

Table 1 shows that General Arts students form close to a third of the total number of observations for the study. Agriculture, Business, and Visual Arts programmes have about 10 percent each of the total number. Thus, in the school, the students in those three programmes can easily be put into one class for teaching. The students in classes for General Arts, Home Economics and Science may need to be regrouped into smaller class sizes that may be taught by different tutors.

Literature Review

Principal component analysis technique is applied to virtually every area of study including biology, medicine, chemistry, meteorology and geology as well as the behavioral and social sciences. Some areas where principal component analysis and other similar multivariate techniques have been applied to assess the performance of students are reviewed in this section.

Morrison (1976) conducted principal components analysis of a covariance matrix given by Birren and Morrison (1961) for 11 subscales of the Wechsler Adult Intelligence Scale (WAIS) along with age and year of education completed for a sample of 933 white males and females. The goal was to isolate the dimensions underlying the variation in the WAIS subscales and, in addition, to see how age and education were related to these dimensions. Two underlying dimensions, accounted for over 62 percent of the variation in the original 13 variables. The first principal components, which by itself explained over 51 percent of the total variance in the 13 variables, had high correlations with all 11 WAIS subtests (correlation ranged from 0.62 to 0.83) and was interpreted as a measure of general intellectual ability. Education had correlation of 0.75 with this dimension. The second principal component which accounted for 11 percent of the total variation in the 13 variables, correlated positively with the verbal subtests and negatively with performance subtest. It was interpreted as contrast between verbal and performance subtests. People who scored high on this dimension had high verbal scores and low performance scores. Age had a correlation coefficient of 0.80 with this dimension indicating that older people did better on verbal tests than on performance tests, compared with younger people.

The first principal component has high loadings on the five indicators that reflect prevalence. The component is essentially an average of the five prevalence indicators. The second principal component has a high positive loading for price and a high negative loading for purity. It was a contrast between price and purity. It was therefore interpreted as a heroin availability index reflecting illicit drug market forces.

A multivariate analysis of students' performance in engineering classes was conducted by Sullivan et al (1996) at Virginia Tech. The study was aimed at describing statistics results by the application of multiple linear regression to students records and performance. The final weighted score was the dependent variable. The independent variables include gender, academic level, grade point average, SAT Math score, SAT verbal score and high school class standing. Further delineations regarding particular engineering major and morning versus afternoon section instructor were also made in the student records data base.

Linear regression was discovered to provide the most accurate predictions of the final weighted score in the engineering economy course. Furthermore, experiments were conducted to determine whether the final examination score by itself could be used as the dependent variable. Because of the greater variability in the final scores, it was determined that the final weight score was the more appropriate dependent variable to use.

The work of Caeser (1994) is very relevant to this study. He studied into academic performance in Mathematics, English, Science and Social Studies in Uganda. The study revealed that there was a great disparity between individuals and between schools in Mathematics and Science performance. Some individuals and schools scored very high while some others score very low in Mathematics and Science related subjects. On the other hand, Uganda Government White Paper of 1992 stated categorically that the worst results in the Primary Leaving Examinations and other levels were more pronounced in Mathematics than in other subjects. This phenomenon has created negative attitudes towards Mathematics right from lower levels to University. Even most teachers and parents appear to demonstrate such negative attitude towards Mathematics.

The apparent lack of mathematical computational skills has led to poor performance in Mathematics related subjects like Core Mathematics, Elective Mathematics, Physics, Chemistry, Integrated Science, Economics, Geography, Costing, Business Management and Accounting, just to mention a few. Carpenter et al, (1975) noted from the report of the National Assessment of Education that only 55 percent of the nine-year old pupils could complete two–digit subtraction problem with regrouping.

The performance of a student in examinations has been attributed by some researchers ultimately to the competence of the class room teacher. For example, Merret and Tang (1994) noted that the nature of schooling and the way schools have been structured, that pupils are taught in groups of varying sizes, presupposes that someone should be expert in the management of such groups in order to bring about good examination results.

Outline of Dissertation

The dissertation is made up of five chapters. In the first chapter we look at the background to the study, objectives, data collection and literature review.

Chapter Two discusses the basic theories and methods. This chapter describes in detail the main statistical tools that will be employed in the analysis of the data.

Chapter Three is on preliminary analysis, which is largely exploratory. This initial step in analysis will help to clarify the structure of the data and serve as a guide to further analysis. It involves the description of the data using numbers, charts and figures.

Chapter Four is on further analysis. In this chapter, the researcher analyzes, the data using principal component analysis technique to further determine the structure of the data in order to achieve the objective of the study.

Chapter Five has three main sub-headings: summary, discussion and conclusion and recommendations. The summary part of the concluding chapter presents a summary of the study including the problem studied, the methodology employed and the results. The conclusion is derived directly from the findings. The evidence that leads to the conclusion is stated in this part. The section also contains relevant recommendations based on the findings of the research.

CHAPTER TWO

REVIEW OF BASIC METHODS

This chapter reviews the theory of the main technique used in the analysis of the data in Chapter Four. In this work, we will mainly use multivariate techniques with special focus on Principal Component Analysis methods. We first explain the basic concept of Principal Component Analysis. Then, we explain the relationship between a principal component and the original variables of study. The chapter also includes a review of the use of principal component scores in further analysis of the data. We conclude the chapter with a study of a test of multiple comparisons of the means of variables.

Basic Concept of Principal Component Analysis

A principal component analysis is concerned with explaining the variance-covariance structure, $\mathbf{\Sigma}$, of a p-dimensional variable vector $\mathbf{X} = \left(X_1, X_2, X_3, \cdots, X_p\right)'$. This is done by constructing a few linear combinations of the original variables. These linear combinations are what is referred to as the principal components. The ith principal component is given by the expression

$$Y_i = \sum_{k=1}^{p} a_{ik} X_k \tag{1}$$

The vector of coefficients $a_i = (a_{i1}, a_{i2}, a_{i3}, \cdots, a_{ip})'$, $i = 1, 2, \cdots, p$, are obtained such that they satisfy the pair of constraints

$$a_i' a_i = 1 \text{ and } a_i' a_k = 0, \; i \neq k \tag{2}$$

The first constraint in Equation (2) means that the components are formed so that they are of unit length. The second condition allows the components to be constructed such that they are uncorrelated. The variance of the component is given by

$$\mathrm{Var}(Y_i) = a_i' \Sigma a_i \tag{3}$$

The covariance between any two components, Y_i and Y_k, is given as $a_i' a_k$ or $\sum_{j=1}^{p} a_{ij} a_{jk}$. If the two components are uncorrelated, then $a_i' a_k = 0$. By the composition of the components, the first principal component, Y_1, is the linear combination with the maximum variance. Generally, the *i*th component Y_i, explains the *i*th largest variations in the data. Thus, each of the *p* components explains successively smaller variation in the data.

Although *p* components are required to reproduce the total variability in the system, often much of the variability can be accounted for by a smaller number, *k*, of the principal components. The *k* fewer components can then replace the initial *p* components. The original data set, consisting of *n* observations on *p* variables, is now reduced to one consisting of *n* observations on *k* principal components. The dimension of the data set is thus reduced but there is (almost) as much information in the reduced dimension as there is in the original. Thus, the general objectives of the principal component analysis technique are: (1) data reduction; and (2) interpretation of the extracted components.

The sum of the variance of the principal components is equal to the sum of the variances of the original variables. That is, $\sum_{i=1}^{p}\lambda_i = \sum_{i=1}^{p}\sigma^2{}_i$, where λ_i is the variance of the ith principal component and $\sigma^2{}_i$ is the variance of the ith original variable, X_i. The sum of the variances of the original variables actually gives the total variation in the data set generated on the p original variables. If the variables are standardized, then $\sum_{i=1}^{p}\lambda_i = p$.

The proportion of variance in the original p variables that k principal components accounts for can be easily calculated as

$$\frac{\sum_{i=1}^{k}\lambda_i}{\sum_{i=1}^{p}\lambda_i} \tag{4}$$

where k is less than p. The proportion of variance that any single principal component accounts for is simply $\dfrac{\lambda_i}{\sum_{i=1}^{p}\lambda_i}$. If the sum of the variance of the first few principal components is close to p, the number of original variables, then we have captured most of the information in the original variables by a few principal components that are linear transformations of the original variables.

From Equation (3), we write the generalized variance of a principal component as $\mathrm{Var}(Y) = a'\Sigma a$. If we denote this variance by λ , and using the first constraint, we obtain the matrix equation

$$\Sigma a = \lambda a \tag{5}$$

Thus, λ is the eigenvalue of the variance-covariance matrix, $\mathbf{\Sigma}$, and a is it's associated eigenvector. This matrix equation can be solved for λ and a, the basic statistics of principal components analysis. It can be shown that the largest eigenvalue λ_1 of $\mathbf{\Sigma}$, is the variance of the first or largest principal component of $\mathbf{\Sigma}$ and its associated vector $a_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1p} \end{bmatrix}$ is the set of weights for the first principal component that maximizes the variance of the component. The second largest eigenvalue of $\mathbf{\Sigma}$ is the variance of the second largest principal component and it's associated vector $a_2 = \begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2p} \end{bmatrix}$ is the set of weights for the second principal component with the next largest variance. The ith eigenvalue (λ_i) is the variance of the ith principal component and its associated vector $a_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{ip} \end{bmatrix}$ are the variable weights defining the ith principal component. If the eigenvalues are all distinct, then there are p distinct associated latent vectors.

Since the principal components are uncorrelated, each one makes an independent contribution in accounting for the variance of the original variables. If, for example, X_i has a correlation of r_{i1} with the first (largest) principal component and r_{i2} with the second largest principal component, then, since the two principal components are uncorrelated, the squared multiple correlation of x_i with the first two principal components is $r_{i1}^2 = r_{i2}^2$. The first i largest principal components maximize the sum of these squared multiple correlations across all the variables. This is a generalization of the fact that the first principal component maximizes the sum of the squared simple correlations of the variables with the largest principal component.

The vector of p principal components variables may be expressed more succinctly in the matrix algebra as $\mathbf{Y} = \mathbf{A}'\mathbf{X}$, where $\mathbf{Y}$ is a p-element vector of principal component scores, $\mathbf{A}$ is a $p \times p$ matrix of latent vectors with the *i*th row corresponding to the elements of the latent vector associated with the *i*th latent eigenvalue, and $\mathbf{X}$ is a p element column vector of the original variables. This is a linear transformation of a p-element random vector **X** into a p element random vector **Y**, the principal components.

From the definition of principal components, we have $\mathbf{A}'\mathbf{A} = \mathbf{I}$. Note that $\mathbf{A}$ is the matrix with latent vectors as columns, $\mathbf{A}'$ is the transpose of $\mathbf{A}$ with latent vectors as rows, and $\mathbf{I}$ is the $p \times p$ identity matrix with ones in the principal diagonal and zeros elsewhere. Thus $\mathbf{A}'\mathbf{A} = \mathbf{I}$ simply indicates that the cross products of any two eigenvectors are 0 and the sum of squares of the elements for a given eigenvector are equal to 1.

Since the *i*th eigenvalue and its associated eigenvectors must satisfy the matrix Equation (5), we have, premultiplying by a_i, $a_i' a_i \Sigma a_i = \lambda_i a_i' a_i = \lambda_i$ for the variance of the *i*th principal component since $a^1{}_i a_i = \sum_{j=1}^{p} a^2{}_{ij} = 1$.

We can succinctly express the fact that

$$\mathbf{\Sigma} a_1 = \lambda a_1,\ \mathbf{\Sigma} a_2 = \lambda_2 a_2, ...,\ \mathbf{\Sigma} a_p = \lambda_p a_p$$

by combining these relations in one matrix expression as $\mathbf{\Sigma A} = \mathbf{A\Lambda}$ where $\mathbf{A}$ is a matrix of eigenvectors as column vectors, and $\mathbf{\Lambda}$ is a diagonal matrix of the corresponding latent roots ordered from largest to smallest. The elements of $\mathbf{\Lambda}$, the diagonal matrix of latent roots, have to be in the same order as their associated eigenvector, the columns of $\mathbf{A}$, in order for the matrix equation $\mathbf{\Sigma A} = \mathbf{A\Lambda}$ to hold. That is, the eigenvectors in the *i*th row and column of $\mathbf{\Lambda}$ must have its corresponding latent vector in the *i*th column of $\mathbf{A}$. We can use any arbitrary ordering of the eigenvector in $\mathbf{\Lambda}$ as long as we use ordering of the associated eigenvectors in $\mathbf{A}$, but it makes more sense to order them with respect to their importance. Equation (3) can be written as

$$a_i' \Sigma a_i = \lambda_i \tag{6}$$

as the equation for the variance of the *i*th principal component using matrix algebra to obtain the covariance matrix of the principal components as. That is, $\mathbf{A'\Sigma A} = \mathbf{\Lambda}$ since $\mathbf{\Sigma A} = \mathbf{A\Lambda}$, we can premultiply both sides of this expression by $\mathbf{A}'$ to obtain $\mathbf{A'\Sigma A} = \mathbf{A'A\Lambda} = \mathbf{\Lambda}$, since $\mathbf{A'A} = \mathbf{I}$. Also, since $\mathbf{\Sigma A} = \mathbf{A\Lambda}$, we can post multiply both sides by $\mathbf{A}'$ to obtain

$$\mathbf{A\Sigma A'} = \mathbf{A\Lambda A'} \tag{7}$$

That is, in Equation (7), we can decompose $\mathbf{\Sigma}$ into a product of three matrices involving eigenvectors and eigenvalues. The goal of principal components analysis is thus, to decompose the variance-covariance matrix. That is, by the use of principal component technique, we seek to explain the variation expressed in $\mathbf{\Sigma}$ in terms of weighting vectors (eigenvectors) of the principal components and variances (eigenvalues) of the principal components. The decomposition of $\mathbf{\Sigma}$ is a key concept in principal components analysis.

Correlation Coefficient between Principal Component and a Variable

The *i*th principal component $Y_i = \sum_{k=1}^{p} a_{ik} X_k$ may be written as

$$Y_i = \mathbf{a}_i{}' \mathbf{X}$$

where $\mathbf{X} = \left(X_1, X_2, X_3, \cdots, X_p\right)'$ is the $p-$ dimensional variable vector. If we represent the *p* dimensional vector $\mathbf{a}'_\mathbf{j}$ by $\mathbf{a}'_\mathbf{j} = (0,0,\cdots,0,1,0,0,\cdots 0)$, where $a_{jj} = 1$, then we can write the component variable X_j as $X_j = \mathbf{a}'_\mathbf{j}\mathbf{X}$

Given the Variance-Covariance matrix of $\mathbf{X}$ as $\mathbf{\Sigma}$, we infer from Equation (3) that

$$\text{Var}(Y_i) = \lambda_i \qquad (8)$$

But since $a_i'\Sigma a_i = \lambda_i$, it implies that $\lambda_i a_i = \Sigma a_i$. Now given $\text{Var}\left(X_j\right) = s_j^2$ we have

$$\begin{aligned} \text{Cov}\left(Y_i, X_j\right) = \text{Cov}\left(a_i'X, a_j'X\right) = \ & a_j'\Sigma a_i \\ &= a_j'\left(\lambda_i a_i\right) \\ &= \lambda_i a_j' a_i \\ &= \lambda_i a_{ij} \end{aligned}$$

The correlation coefficient between the *i*th principal component, Y_i, and the *j*th variable, X_j, is generally given by the expression

$$\rho_{Y_i X_j} = \frac{Cov(Y_i, X_j)}{\sqrt{Var(Y_i)}\sqrt{Var(X_j)}}.$$

Making appropriate substitutions, we obtain an expression for the correlation coefficient, $\rho_{Y_i X_j}$, between Y_i, and the *j*th variable, X_j, in terms of the loading, λ_i, and the weights, a_{ij}, as

$$\rho_{Y_i X_j} = \frac{a_{ij}\sqrt{\lambda_i}}{s_j} \qquad (9)$$

Equation (9) suggests that the size of the correlation coefficient between a component and a variable depends to a large extent on the variability in the observations on the variable. If the values are dispersed, the correlation coefficient will more likely be small. On the other hand, if there is little spread in the values of the variable, the correlation coefficient is more likely to be large.

Computation and Interpretation of Component Scores

We discuss the computation of component scores in relation to the data on examination scores that we will study in this dissertation. The specific relation to the data under study is to aid easy explanation of the concept which will help in interpretation of results in Chapter Four; a generalization of the discussion may run into difficulties.

The *i*th principal component is given in Equation (1) as $Y_i = \sum_{j=1}^{p} a_{ij} X_j$.

Clearly, the value of this expression will be influenced by the observed values of the variables, X_j. The variability in the values of these variables will then be reflected in the component scores. This can distort interpretation of the score. To eliminate the effect of the variation in the individual variables, we first standardize the data. Thus, the component score corresponding to Y_i is given by

$$C_i = \sum_{j=1}^{p} a_{ij} \frac{X_j - \mu_j}{s_j} \qquad (10)$$

The standardization process determines the magnitude and the sign of the *j*th term in the expression in Equation (10). Assuming that all the weights, a_{ij}, are positive, we consider three typical scenarios: C_i could be a high positive value, a high negative value or close to zero. A high positive score is obtained if the values of most of the items $\left(X_{j,}\ j = 1, 2, \cdots, p\right)$ are higher than the average values of the respective items. Thus, in relation to the examination scores data used for this study, a high positive score C_i indicates that the student obtained scores in all the subjects $\left(X_{j,}\right)$ that are consistently much higher than the mean score in each of the subjects. Such a candidate is a typical outstanding student.

A high negative score is obtained if the values of most of the items $\left(X_{j,}\ j = 1, 2, \cdots, p\right)$ are lower than the average values of the respective items. Thus, in relation to the examination scores data used for this study, a high negative score, C_i, indicates that the student obtained scores in all the subjects $\left(X_j\right)$ that are consistently much lower than the mean score in each of the subjects. Such a candidate is a typical weak student.

A very small score (close to zero) is obtained if the values of most of the items are just about the same value as the average values of the respective items. Thus, in relation to the examination scores data, a small value of C_i indicates that the student consistently obtained scores in all the subjects that are about the same value as the mean score in each of the subjects. Such a candidate is a typical average student.

The above discussion is basically on the effect of the standardization of the data on the sign and size of the component score assuming that the loadings are all positive. However, the magnitude of the score is also determined to a large extent by the size of the loading, a_{ij}, on the variable X_j. In order to explain this point, we consider two scenarios: the loadings are almost equal; and some of the loadings are much larger than others.

In the case where the loadings are almost equal, the principal component, Y_i, is usually referred to as a *weighted sum* of the original variables. In this case all the variables have about the same influence in the formation of the component. In relation to the examination scores data, a high component score means that the candidate is generally good in all subject, and hence, an outstanding student.

On the other hand, if a_{ij} are large for some variables, say X_k and X_t, and a_{ij} are low on all others, then the size of the component score would have been influenced by the loadings of X_k and X_t, a_{ik} and a_{it}. If the component score is high and positive, then it implies that the candidate is a good student in the subjects represented by the variables X_k and X_t. If the component score is high and negative, then it implies that the candidate is a poor student in the two subjects, X_k and X_t. If the component score is small (close to zero) then it implies that the candidate is an average student in the two subjects, X_k and X_t.

The Number of Principal Components to Retain

The number of principal components to retain depends principally on the goals of the analysis. Thus, if the interpretation assigned to a principal component meets the objective of the study, the component may be considered relevant for extraction. However, a number of researchers in the area have given various suggestions to serve as a rule to component extraction. Kaiser (1960) recommends dropping those principal components of a covariance matrix with eigenvalues less than one. He argues that the eigenvalues less than one contain less information than a single standardised variable whose variance is one.

Joliffe (1972) has suggested that Kaiser's rule tends to throw away too much information and the basis of simulation studies has suggested a cut off of 0.7 for correlation matrices.

Chattel (1966) proposed the use of a Scree graph to help decide on how many principal components to retain. The Scree graph involves plotting the eigenvalues and finding a point where the line joining the points is steep to the left of the point k, and not steep to the right of k. One then retains k principal components. If many principal components are retained relative to the number of variables, the description of the data is less parsimonious. In addition, fewer principal components are, in general, easier to interpret.

Multiple Comparisons of Means

To make multiple comparisons of a set of treatment means, a number of procedures are used which, under various assumptions, ensure that the overall confidence level associated with all the comparisons remains at or above 100(1 - α)% level. One of the widely used techniques is the Bonferroni method.

The method is used for pair-wise comparisons or contrasts among k treatment means. If a set of contrast has been specified a priori but not orthogonal, the *t*-test is used but with a Bonferoni correction. In this case, if g tests (number of comparisons) are involved and the overall Type 1 error is to be held at α, then the significance level for individual test (each of the g tests) is set at $\alpha' = \frac{\alpha}{g}$. For example, to maintain $\alpha = 0.05$ in making five comparisons, we use a significance level of $\alpha' = \frac{0.05}{5} = 0.01$.

The following are the steps followed to perform Bonferroni's Multiple Comparison test:

1. For each contrast L, calculate the estimate L' and the standard error $Se(L')$;
2. Calculate the value $B = t_{\alpha/(2g),N-K}\, Se(L')$, where N is the number of observations and k is the number of treatments under consideration.
3. Bonferroni's method declares L to be significant if $|L'| > B$.

A $(1-\alpha/g)100\%$ confidence intervals for g comparisons can then be constructed as

$$L' \pm t_{\alpha/(2g),\, N-K}\, Se(L')$$

If a confidence interval for L does not contain 0, then we reject $H_o: L = 0$ in favor of $H_1: L \neq 0$.

CHAPTER THREE

PRELIMINARY ANALYSIS

In this chapter, we explore the examination scores data to identify some basic features that will be necessary to guide the next chapter. We will make use of routine statistical procedures such as descriptive statistics and the distribution of the performance in each subject for each programme. We will also consider an analysis of variance in order to determine the main source that contributes to the performance of the students.

The Variance-Covariance Matrix of the Study Variables

Table 2 gives the variance-covariance matrix of the variables (subjects) under study. The diagonal elements are the variances of the respective variables. For example, the amount of variation in the performance in English is given as 157.42, and that in Physical Education (PE) is 135.87. An off-diagonal element is the amount of covariance between a pair of subjects. For example, the covariance between Core Mathematics and Science is 125.67. We recall from the review chapter that the higher the amount of covariance between a pair of variables, the higher the correlation coefficient between the variables. We see from Table 2 that the highest variation in performance is recorded in Core Mathematics. The second most varied performance is observed in French. It means that in these two subjects there are a number of students whose scores are very high whilst some others have very low scores.

Table 2: Variance-Covariance Matrix of the Study Variables

Var	English	Maths	Science	Soc. Stds	PE	ICT	French
English	157.42						
Maths	71.73	247.72					
Science	67.91	125.67	157.58				
Soc. Stds.	-7.81	42.26	28.66	111.04			
PE	24.46	55.48	44.22	14.67	135.87		
ICT	65.28	51.16	54.79	-10.84	25.86	149.94	
French	69.75	92.39	66.34	-0.25	6.00	23.70	238.73

On the other hand, scores in Social Studies is the least varied among the subjects. This means that the disparity in performance of students in Social Studies is less wide than in any other subject. The spread in performance in English Language, Science, Physical Education and ICT is almost the same.

We see from Table 2 that the highest covariance (of 125.67) is between Core Mathematics and Science. It means that among the seven subjects, the highest correlation (of 0.636) exists between these two subjects. Thus, knowledge in Mathematics is more reflected in Science than in any other subject. The correlation matrix corresponding to the Variance- Covariance matrix in Table 2 can be seen in Appendix B. The value of covariance is least (–0.25) between Social Studies and French. This suggests that performance in these two subjects is the least related.

Adding the diagonal elements in Table 2, we obtain a total variance of 1198.30 in the scores data. The percentage contribution of the variation in each subject to the total variation is given in Table 3. The subjects have been re-ordered in the table to reflect the amount of respective contribution to the total variation in the data.

Table 3: Percentage Variation in Each of the Study Subjects

Subject	Variance	% of Total Variation	Cumulative % of Total Variation
Maths	247.72	20.67	20.67
French	238.73	19.92	40.59
Science	157.58	13.15	53.74
English	157.42	13.14	66.88
ICT	149.94	12.51	79.39
PE	135.87	11.34	90.73
Soc. Stds	111.04	9.27	100.00
Total	1198.30	100.00	

We see from Table 3 that the variation in each subject contributes substantially to the total variation. Particularly, the percentage contribution of Core Mathematics and French are almost the same and the two together contribute about 41 percent of variability in the data. The variance figures have a number of implications in this context. One relevant implication is that they may reflect the disparity in perception on the level of difficulty (to the students under study) associated with the respective subjects. The higher the variation, the wider the disparity in perception on the level of difficulty of the subject.

Distribution of Performances in each of the Study Subjects for all Programmes

In this section we examine the distribution in the scores obtained by the students in each of the six programmes for each of the seven subjects under study. We make use of the box-and- whiskers plot in each case. The box-and-whiskers plot, like the histogram, enables us to see the spread in the scores. In addition, it enables us to obtain estimates of some descriptive statistics about the data, such as the median and the quartile values. It is also constructed to identify observations that are generally considered as extreme within the range of the particular data set.

In Figure 1 we have the box plot of the scores in English Language for the six programmes. We see that the distribution of the performance in this subject differ widely from programme to programme. The scores range from about 22 percent to about 92 percent. It can be seen that the highest score was obtained by students from Visual Arts programme. The lowest score was obtained by a student of Science. All the students of Business are scoring between about 55 percent to about 70 percent, with 50 percent of the students scoring between 60 and 65 percent. Thus, the performance of Business students in English Language is the least variable among all the programmes. We observed this in the descriptive statistics that we have already encountered. In contrast to the performance of Business students is that of the Visual Arts students. The scores are widely varied and by the standard of that class, three students performed extremely below expectation. The worst, however, was obtained by students in Science.

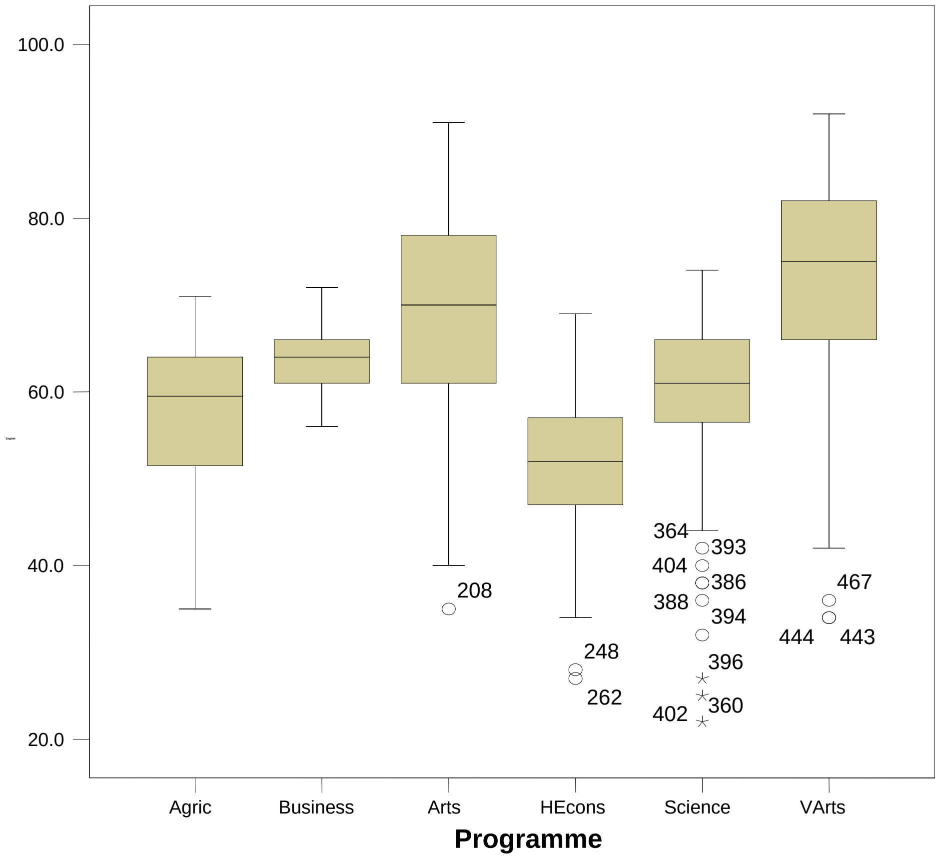


Figure 1: Distribution of scores in English Language for all programmes

In the Science class there are several students who performed extremely below the average of the class. Generally, it appears that in English Language, students in Visual Arts perform better than any other programme, whilst Science students are the weakest.

The distribution of scores in Core Mathematics for all the programmes is given in Figure 2. The scores range from about 20 percent to about 98 percent. In this subject, we see that the distribution of performance is very similar for all the programmes: the variability in performance is very wide for all the programmes. In particular, the performance is almost the same for Arts and Science. It can be seen that the highest score was obtained by students from the Arts and Science programmes. The lowest score was obtained by a Home Economics student. By the standards of the classes, two students in Agriculture are outstanding in that class, whilst the performances of two students in Visual Arts are abysmal for their class. Generally, it appears that in Core Mathematics, performance of Arts students is almost normal and better than any other programme. Students of Home Economics and Agriculture are, however, the weakest in the subject.

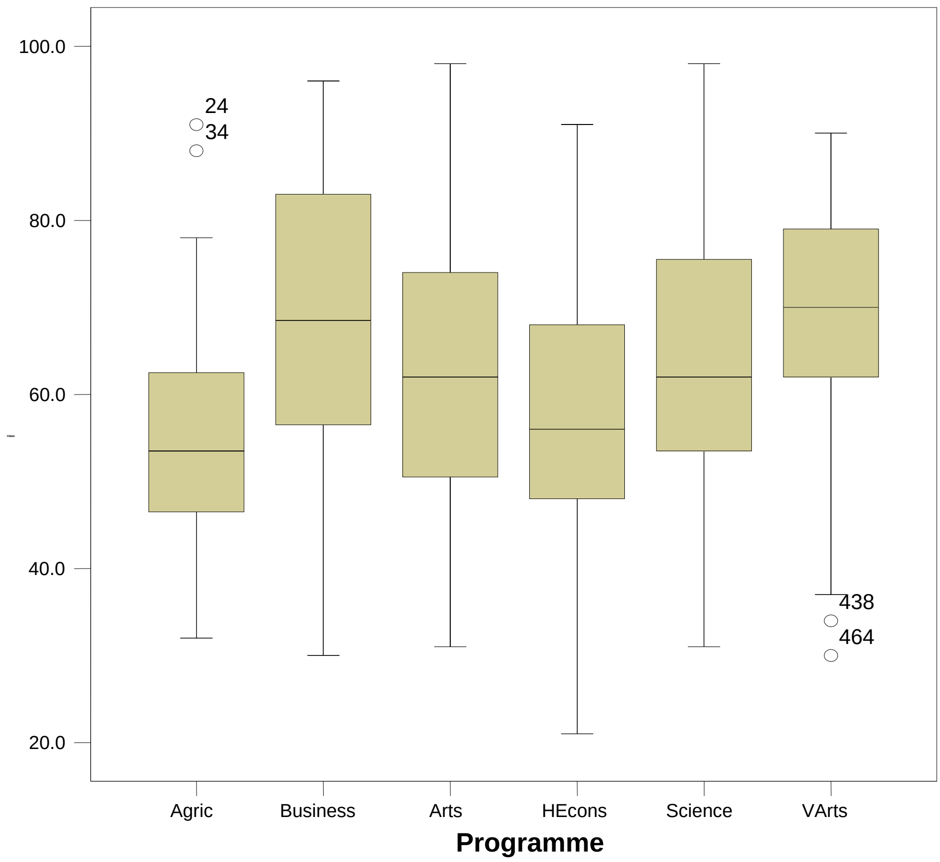


Figure 2: Distribution of scores in Core Mathematics for all programmes

In Figure 3, we see the distribution of scores in Integrated Science for the programmes. The scores range from about 20 percent to about 93 percent. We see that the distribution of performances for students of Home Economics and Visual Arts differs markedly from those of the other programmes. It appears that the highest score was obtained by students from Arts, Science and Visual Arts. The student who obtained the highest score in Visual Arts is an outstanding student in his/her class. The lowest score (of about 20 percent) was obtained by a student in Home Economics.

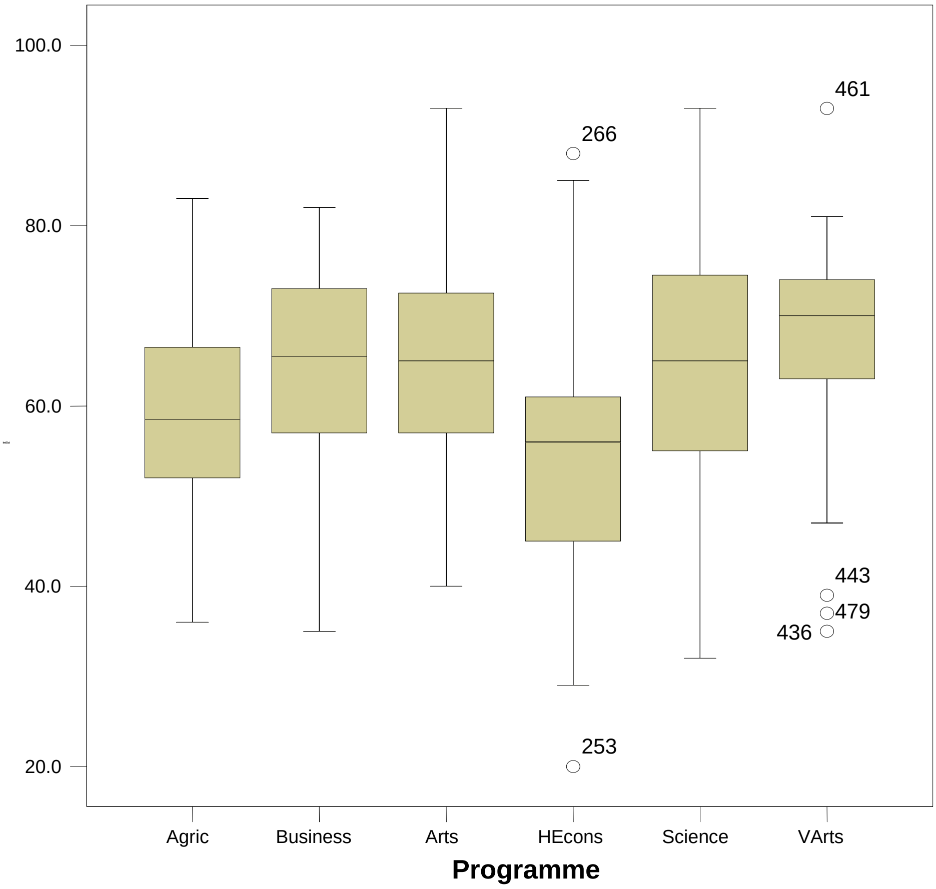


Figure 3: Distribution of scores in Integrated Science for all programmes

This lowest score is an abysmally low performance. Generally, in Integrated Science, it appears that the performance of the Arts students is the best among the programmes. Students of Home Economics are, however, the weakest in the subject.

The distribution of scores in Social Studies is shown in Figure 4. The scores in this subject range from about 30 percent to about 92 percent. It can be seen that the distribution of performance of Visual Arts students has the widest variability and is markedly different from those of the other programmes. Particularly, the distribution of performance of Visual Arts students is in contrast to that of Agriculture students, who have the least variable performance. Another

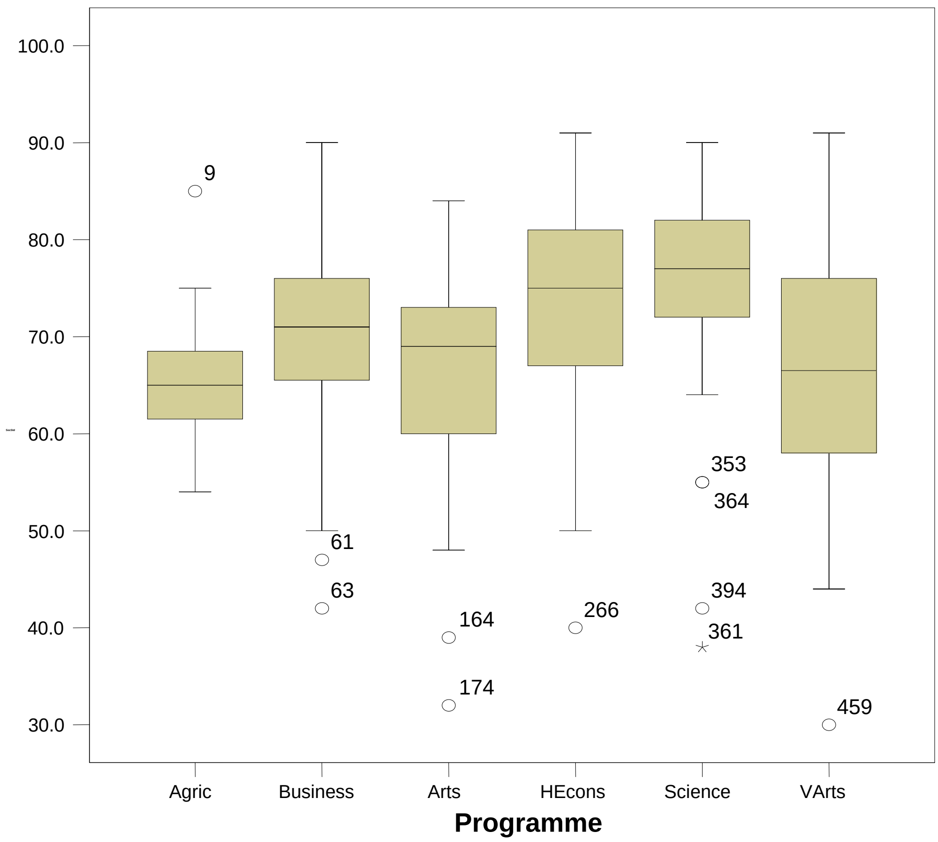


Figure 4: Distribution of scores in Social Studies for all programmes

Observation from the diagram is that in all the classes there are extreme performances. The extreme performance in Agriculture is an outstanding one in that class. Apart from this, all the other extreme performances are abysmal by the standards of the respective classes. The worst of all these extreme performances is the one from the Visual Arts class. Incidentally, Visual Arts students scored one of the highest marks in Social Studies, with another highest score from Home Economics. On the whole, it appears that in Social Studies, the best class is Science, whilst the worst is the Agriculture class.

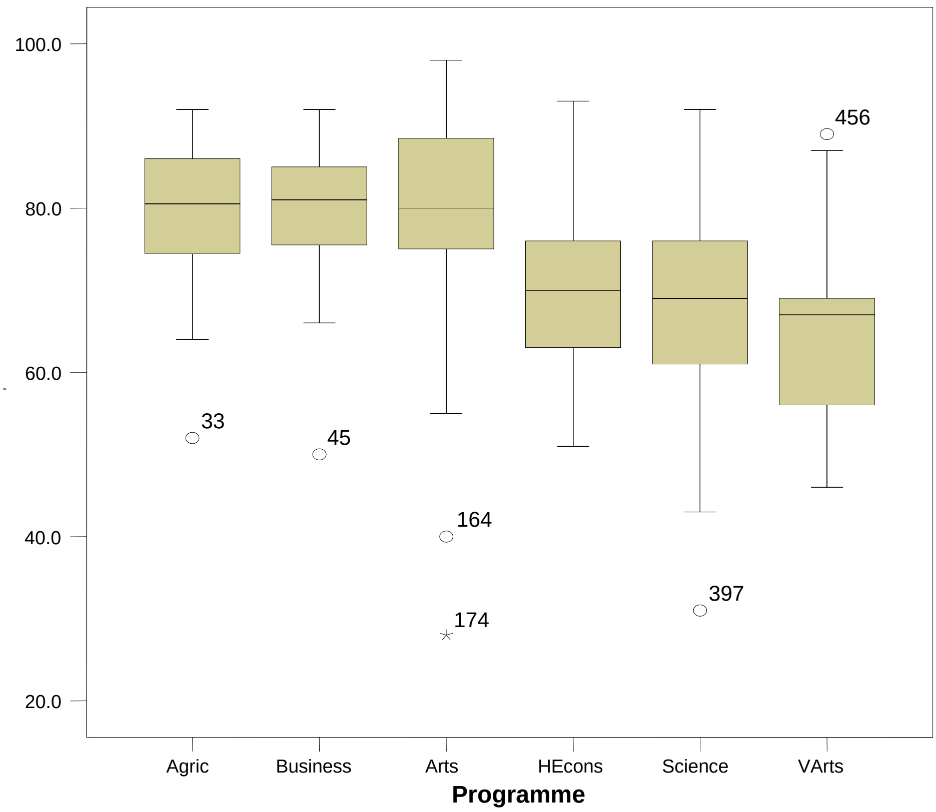


Figure 5: Distribution of scores in Physical Education for all programmes

The distribution of scores in Physical Education is given in Figure 5. The scores range from about 30 percent to about 98 percent. Both the highest and the lowest scores were recorded by students from the General Arts programme. Thus, the performance in the General Arts is the most dispersed in the subject. The performances in the Agriculture and the Business classes are similar: they are the least variable and have almost the same average performance. The distribution of performance in Visual Arts is markedly different from the rest and appears to have the lowest average performance among the programmes. The Visual Arts class appear to be the weakest in the Physical Education subject. Generally, the General Arts class appears to be the best class in the subject.

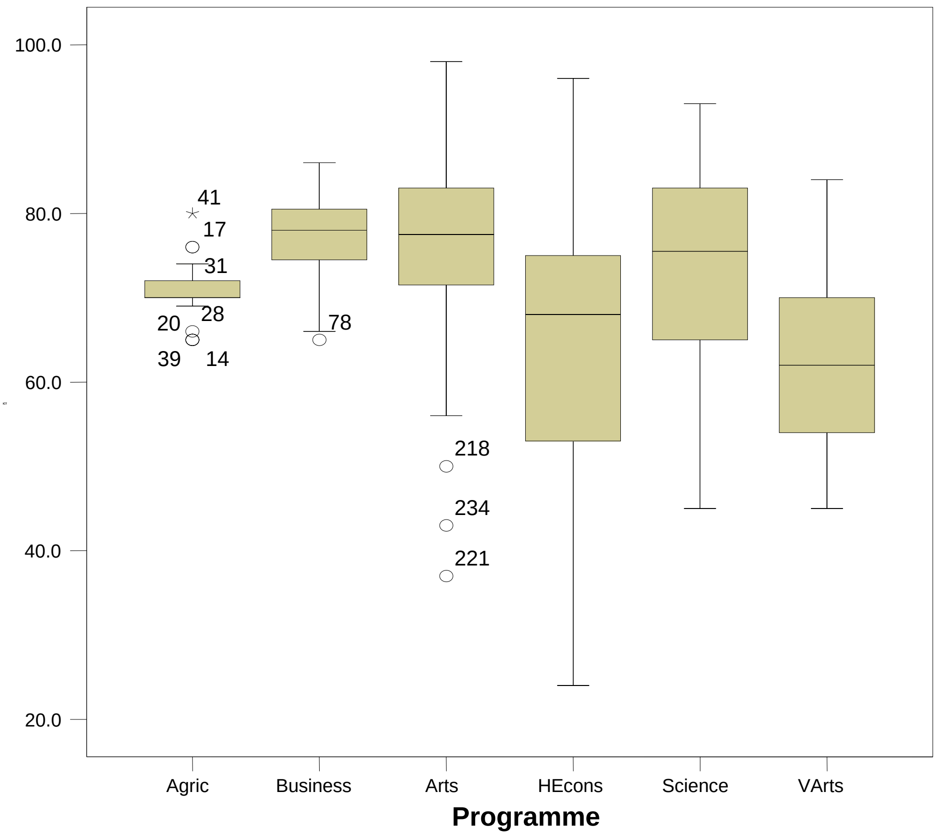


Figure 6: Distribution of scores in ICT for all programmes

In Figure 6 we have the distribution of scores in ICT for all the programmes. The scores range from about 24 percent to about 98 percent. The highest score is obtained by a General Arts student whilst the lowest score is obtained by a Home Economics student. We see that the distribution of performance differs markedly from programme to programme. Particularly, the average performance in Visual Arts appears to be much lower than those of the rest. The performance of Home Economics is the most dispersed. In sharp contrast, the scores in Agriculture are almost uniform with a few extreme performances. The scores of Business are also very much crowded in a small range. Thus, in Business and Agriculture performance in ICT is very competitive. Generally, the General Arts class appears to be the best class in ICT, whilst the Visual Arts appears to be the weakest.

The distribution of scores for each programme in French is what is shown in Figure 7.

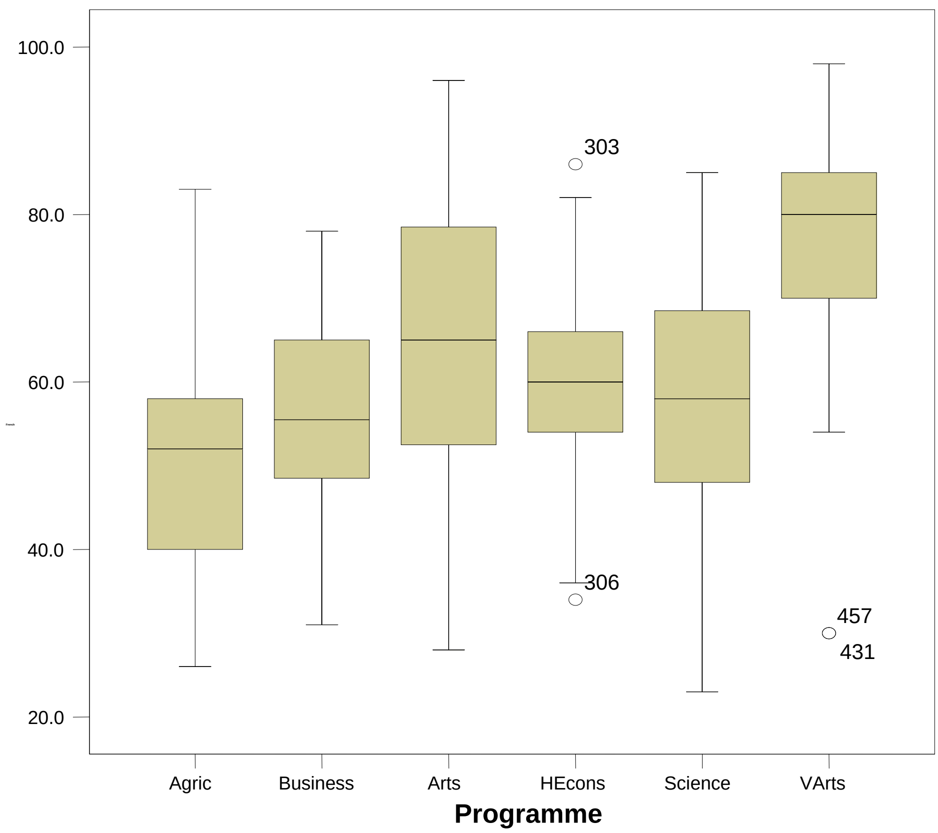


Figure 7: Distribution of scores in French for all programmes

The scores range from about 23 percent to about 98 percent. The highest score is obtained by a Visual Arts student whilst the lowest score is obtained by a Science student. We observe that the distribution in the Visual Arts class is very different from the rest, with the highest average (around 80 percent) performance. The average performances in the other six classes are all in the neighbourhood of 60 percent. Thus, when it comes to French, the Visual Arts is a much more superior class. It is the least competitive in General Arts class in the subject, whilst the Agriculture class, generally, appears to be the weakest in the subject.

General Comparison Between the Programmes

With exception of French, the performance of Agriculture class in the other six subjects was the least dispersed among all the seven programmes. The performance distribution in the Business has been similar to that of the Agriculture. The Business class has been consistently less dispersed in performance in all the seven subjects.

Visual Arts is a far more superior class when it comes to French, and the class also has the highest average performance in English Language. Thus, it appears that when it comes to the languages, Visual Arts has the best students. Generally, in Visual Arts, the distribution of performance has been different from the other classes: it may have the highest average performance (as in the languages); it may have the lowest average performance (as in the PE and ICT); and it records extreme performances in all subjects, except in the ICT.

In the General Arts, performance has been very dispersed in all the subjects. It produced the best results in four out of the seven subjects. These are the ICT, Physical Education, Integrated Science and Core Mathematics. It also produced the second highest scores in the two languages, English Language and French. It, however, recorded the lowest score in Physical Education. Particularly, we observed that the 174th student (in Arts class) who obtained the lowest score in Physical Education also obtained the lowest score in Social Studies in his/her class and the over all second lowest in that subject.

Similar to the General Arts, the performance in the Home Economics has been very dispersed in all the subjects, and most extremely so in the ICT. The Home Economics class is particularly noted for recording the least scores in a number of subjects. These are the ICT, Integrated Science and Core Mathematics.

The performance of the Science class is quite similar to that of the Home Economics: it exhibits a lot of variability; and recorded the least scores in the two languages, English and French. In contrast to the Visual Arts class, the Science class appears to be the weakest in the two languages.

Analysis of Variance of Students' Performance in each Subject

Table 4 gives the results of the analysis of variance of the performance in each of the seven subjects offered in the first year in the school. It will be recalled that a total of 478 students were involved in the study. These students were grouped into six programmes. Thus, the performance in any subject may be as a result of the particular programme offered by the students. The performance may also be as a result of the peculiar nature of the subject itself. Thus, the source of variation in performance may be attributable to differences between programmes (Between Groups) and the peculiar nature of the subject (Within Groups). The value of the F statistic as well as the corresponding p–values are respectively given in the six and seventh columns of Table 4.

Table 4: Analysis of Variance of the Performance in each Subject

Subject	Source of Variation	Sum of Squares	Deg. of Freedom	Mean Square	F	Sig.
English	Between Groups	22043.108	5	4408.622	39.196	0.0000
	Within Group	53201.210	473	112.476		
	Total	75244.317	478			
Core Maths	Between Groups	8215.969	5	1643.194	7.053	0.0000
	Within Group	110193.900	473	232.968		
	Total	118409.800	478			
Integrated Science	Between Groups	8502.795	5	1700.559	12.038	0.0000
	Within Group	66820.641	473	141.270		
	Total	75323.436	478			
Social Studies	Between Groups	8249.002	5	1649.800	17.408	0.0000
	Within Group	44827.829	473	94.773		
	Total	53076.831	478			
Physical Education	Between Groups	15507.761	5	3101.552	29.675	0.0000
	Within Group	49437.508	473	104.519		
	Total	64945.269	478			
ICT	Between Groups	17122.798	5	3424.560	29.695	0.0000
	Within Group	54548.864	473	115.325		
	Total	71671.662	478			
French	Between Groups	21981.922	5	4396.384	22.571	0.0000
	Within Group	92130.057	473	194.778		
	Total	114112.000	478			

We see from Table 4 that the *p*–values of the tests of equality of means of scores in each subject for all six programmes are very negligible compared to a significance level of five percent. Thus, all the tests are significant. This means that the average performance in each of the subjects is not the same for some pairs of the six programmes. The result of the analysis of variance shows that students performance in each of the seven subjects is largely attributable to the respective programmes that they offer.

We observe from Table 4 that even though all the tests of equality of means are significant, the value of the F statistic for Core Mathematics is small. This means that, in Core Mathematics, we expect the performance of students in only a few pairs of programmes to be different. In other words, for most of the programmes the average performance of students is the same in Core Mathematics.

Appendix A gives the results of the Least Squares Difference Multiple Comparison test. In that table we see that, for example in French, the average performance of students in Visual Arts is entirely different from the performance in all other programmes. Again, in English Language, the average performance of Home Economics students is entirely different from those of the other six programmes. We produce in Table 5 from Appendix A a portion of the results of least squares difference test in Mathematics between Agriculture and the other programmes and in French between Visual Arts and the other programmes.

Table 5 gives but just two examples of sets of results of the Least Squares Difference Multiple Comparison test.

Table 5: LSD Results in Mathematics and French for Agriculture and Visual Arts

Difference between Agriculture (i) and other programmes in Core Mathematics		
Programme (j)	Mean Difference ($i - j$)	Significance
Business	-13.2424	0.0000
Arts	-7.6726	0.0040
HEcons	-2.7029	0.3300
Science	-8.2243	0.0030
VArts	-13.6691	0.0000
Difference between Visual Arts (i) and other programmes in French		
Programme (j)	Mean Difference ($i - j$)	Significance
Agriculture	25.8473	0.0000
Business	20.2033	0.0000
Arts	10.5254	0.0000
HEcons	16.2643	0.0000
Science	18.6743	0.0000

We see from the table that when it comes to Core Mathematics, the performance of Agriculture class is only comparable to that of the Home Economics class. The average performance in Agriculture is significantly different from the average performances in the other programmes. Another observation is that all the differences are negative. This means that the average performance of the Agriculture class is below all the performances of the other programmes.

The table also shows that in French, the performance of the Visual Arts class is comparable to none of the other groups. The differences between the average performance in the Visual Arts class and the other classes are all positive. This indicates that in French, the Visual Arts class is far superior to the other classes. This observation is a confirmation of the representation in Figure 7. The other results in Appendix A are interpreted in a similar manner. It is clear that the results of the Least Squares Difference Multiple Comparison tests are much in support of the representations in the box plots in the previous section.

Summary of Preliminary Analysis

The preliminary analysis of the data examined the Variance-Covariance matrix of the study variables (subjects) and the distribution of performance of students in each of the six programmes in each subject. The chapter also studied the main sources of variation in the performance in each of the subjects. The Variance-Covariance matrix showed that performance was most dispersed in Core Mathematics, and then in French. Performance was least dispersed in Social Studies. It was also realized that performance in Core Mathematics and Science were more related than any other pair of subjects.

A study of the distribution of performance in the various programmes identified Visual Arts to be generally the most superior in the two language subjects, English Language and French. However, students in Visual Arts have the tendency of recording extreme scores, with exception of ICT.

Performance of students in General Arts was typically widely dispersed in all the subjects. It produced the best results in four out of the seven subjects. These are the ICT, Physical Education, Integrated Science and Core Mathematics. It also produced the second highest scores in the two languages, English Language and French. It, however, recorded the lowest score in Physical Education. Similar to the General Arts, the performance in the Home Economics is characterized by wide disparity in performance in all the subjects. The Home Economics class is particularly noted for recording the least scores in three of the subjects. These are the ICT, Integrated Science and Core Mathematics.

The Science class appears to be the weakest in the languages, English and French, recording the least scores in the two subjects. Generally, the class also exhibits a lot of variability in a manner quite similar to that of the Home Economics class. The performance in the Agriculture and Business classes were the least dispersed in all the subjects. Thus, the students in these two classes are the most competitive within their respective classes. They are, however, not competitive among the entire groups of programmes.

The result of the analysis of variance revealed that the average performance in each of the subjects differ from programme to programme. This shows that students' performance is largely attributable to the respective programmes that they offer.

CHAPTER FOUR

FURTHER ANALYSIS

In the preliminary analysis, a number of observations were made about the performance of the students in the various programmes in each of the subjects under study. One of these observations was that no particular programme consistently produced the best or the worst performance in all the subjects. This pattern makes it rather difficult to determine the particular class (programme) that may be regarded as generally the best among the six programmes. It is even more difficult to identify the particular students from the six programmes that may be considered as the best (or the worst) in specific subjects. It was also observed that the disparity in performance was not the same in all the subjects. In that chapter, we could not examine the individual subjects to identify the best (or the worst) students among all the students together. To achieve this objective in this chapter, we further analyze the data by making use of a multivariate technique that combines the scores of a student in all the subjects. This technique is the Principal Component Analysis method. The technique also takes into account the relative importance of the variation in individual subject in determining the overall best (or worst) performance.

Principal Component Analysis of Data

We now analyze the score data using the Principal Component Analysis method. As outlined in the review method, since we have seven original variables, we should have seven principal components. First, we consider the variance explained by the each of these components. The next is to determine these components and find labels for them. The labels will enable us interpret these components. Using the interpretation assigned to each of the components, we will then be able to identify which of them could further be used to identify the overall best and worst students among the set of students under consideration.

The Percentage of Variance Explained by Components

Table 6 gives the variation in the data accounted for by each of the seven components and the corresponding percentage of total variation explained. The last column gives the cumulative percentage variations explained by the components. The components have been numbered from 1 to 7, where the first component accounts for the largest variation in the data and successive components account for successively smaller variations. The extraction of the variations is done using the Variance-Covariance matrix in Table 2. The Variance-Covariance matrix is found to be more appropriate here than the Correlation matrix (see Appendix B). This is because the variance of the individual variables (subjects) reflects the relative "importance" in explaining the performance of the students. As a result, the extraction of the components must take into account these variances. In the section where the relevant component will be used to identify specific performances, we will use standardization procedures to nullify the effect of the variations in the data.

Table 6: Percentage of Variance Explained

Component	Variance Explained	% Variance explained	% Cumulative Variance Explained
1	502.857	41.965	41.965
2	193.699	16.165	58.130
3	168.753	14.083	72.213
4	105.761	8.825	81.038
5	83.724	6.987	88.025
6	80.460	6.714	94.739
7	63.037	5.261	100.000

As noted in the previous chapter, the total variation in the data is 1198.30. We see from Table 6 that the first principal component alone accounts for much more than a third of the total variation in the data. The contribution of the first component is more than twice that of the second component. The last four components account for less than a third of the variation. The results show that the first component can be used to explain some identifiable dimension in the performance of the students under study with a very high degree of precision. The other components (e.g. 2 and 3) may be used to determine some more specific dimension of interest about the data. We consider in the next section the extraction and labeling of the components.

Extraction of the Principal Components

Table 7 is the component matrix that shows the weights for the various components. The weights in the table are constructed such that the conditions of Equation (2) are satisfied. (The standardized weights or coefficients are given in Appendix C.)

Table 7: Principal Component Matrix

Variable	Component						
	1	2	3	4	5	6	7
English	0.3190	-0.1134	0.4644	-0.0497	0.4316	-0.6370	0.1897
Maths	0.6554	0.3004	-0.4141	-0.3092	-0.6223	-0.1934	0.4554
Science	0.3945	0.1754	-0.0247	-0.1466	0.0894	0.0036	-0.8292
Soc. Stds	0.0660	0.2055	-0.3607	-0.1677	0.6393	0.2204	0.1927
PE	0.1619	0.3324	0.0308	0.8468	0.0414	0.0299	0.0497
ICT	0.2263	0.1105	0.6340	-0.1778	-0.0325	0.6291	0.1694
French	0.4809	-0.8374	-0.2815	0.3219	0.0835	0.3341	0.0270

We observe from Table 7 that the coefficients of the first principal component are all positive. The loading on Mathematics is the highest followed by that of French. With the exception of Social Studies, all the loadings are quite high on the variables. Thus, all the variables contribute fairly significantly to the formation of the first component. We can therefore label this component as a *weighted sum* of the original variables. Thus, the first component can be used to identify the first few overall best (or worst) performances in the examinations. The magnitudes of the loadings suggest that in ranking the general performance of the students, one’s performance in Core Mathematics and French, in particular, will be very influential. The performance in another two subjects will be important, but to lesser extents, in deciding the rankings of students. These subjects are Science and then English Language.

As indicated earlier, the other principal components may be used to determine performance in specific dimensions. For example, the second component has negative loadings on only the two languages, English Language and French with the loading on French being the highest in absolute terms. We may label this component as *language* component. We can therefore use the second component to assess the performance of the students in the languages. However, this is not the objective of this work. The other components may be labeled in a similar way.

In order to ascertain the extent of relationship between the first principal component and each of the seven variables, we compute the correlation coefficient between the component and each of the subjects. Using Equation (9) in the review chapter, we obtain the correlations in Table 8. The table also shows the loadings of the first component in Table 7.

Table 8: Correlation between the First Principal Component and the Variables

Variable	Loading	Correlation Coefficient
English	0.3190	0.5701
Maths	0.6554	0.9340
Science	0.3945	0.5690
Soc. Stds	0.0660	0.1405
PE	0.1619	0.3114
ICT	0.2263	0.4144
French	0.4809	0.6982

From Table 8, we see that the higher the loading, the higher the correlation coefficient between the component and the variable. There is a strong positive correlation coefficient between Core Mathematics and the component. This buttresses the extent of dominance of Mathematics in the formation of the component. Just like its loading, the correlation between the component and French is the second highest. However, English Language and Science now have almost the same amount of correlation coefficient with the component.

In the next section, we make use of the first principal component to determine outstanding (and abysmal) performances in the examinations.

Determination of Outstanding and Abysmal Performances

The main objective of this work is the determination of outstanding and abysmal performances in the examination scores data. The results of the components extraction in the previous section revealed that it is only the first principal component that is relevant for achieving this objective. In this section, we compute the scores of the first component and use it to identify the extreme performances.

As explained in the review chapter, the computation of the component scores makes use of standardized data. In Table 9, we have the first principal component scores for some selected observations. The component scores for all the observations can be found in Appendix D. The computations of these scores were carried out using the standardized component coefficients in Appendix C, rather than the coefficients in Table 7.

Table 9: First Principal Component Score for Selected Observations

Student	Programme	Component Score	Student	Programme	Component Score
			103	Arts	-0.65835
⋮	⋮	⋮	104	Arts	1.25911
			105	Arts	2.77529
34	Agriculture	1.38540	⋮	⋮	⋮
35	Agriculture	-0.19798			
36	Agriculture	-1.00025	155	Arts	2.63602
37	Agriculture	0.37810	⋮	⋮	⋮
38	Agriculture	-0.78702			
39	Agriculture	0.38833	203	Arts	2.51586
40	Agriculture	0.03448	⋮	⋮	⋮
41	Agriculture	-0.56273			
42	Agriculture	-0.54910	213	Arts	-1.29866
43	Agriculture	-0.04600	214	Arts	-1.14427
44	Agriculture	-1.70272	215	Arts	-2.22323
⋮	⋮	⋮	216	Arts	-1.17784
			⋮	⋮	⋮
47	Business	0.00648			
48	Business	-0.82126	246	HEcons	0.01379
49	Business	-0.01282	247	HEcons	-0.41083
50	Business	-1.03106	248	HEcons	-2.44240
51	Business	0.90694	249	HEcons	-1.68550
52	Business	0.53216	250	HEcons	-1.39692
53	Business	-1.05026	⋮	⋮	⋮
54	Business	1.77270			
55	Business	-0.38734	386	Science	-1.52690
56	Business	0.88109	387	Science	-1.38830
57	Business	0.16533	388	Science	-2.63130
58	Business	-1.46207	⋮	⋮	⋮
59	Business	-0.19218	474	VArts	1.84660
60	Business	-0.47494	475	VArts	1.50063
⋮	⋮	⋮	476	VArts	0.56250
			⋮	⋮	⋮

As already pointed out, even though all the coefficients of the first component are positive, the standardization of the original scores may introduce negative signs in the expression for the component. The effect on the magnitude and sign of the final component score would depend on the magnitude of the coefficient of the variable (subject) and whether or not the score of the student in that subject is higher or lower than the mean score in that subject. (See Chapter Two for detailed derivation of component score.)

The component scores range from -2.6313 to 2.7753. The highest component score is obtained by an Arts student whilst the smallest score is obtained by a Science student. A typical score that is close to zero is obtained by a Business student. The relationship among the scores is illustrated in Figure 8.

In Figure 8, we have a scatter plot of the first principal component scores against corresponding number of observations. The figure distinguishes the component score with the particular programme of the student. As noted earlier, the component scores range from -2.6313 to 2.7753 (the derivation of component scores has been treated in the review chapter.) A high positive score indicates that, in almost all the subjects, the student obtained scores that are higher than the mean scores in the respective subjects. Thus, the student with the highest component score is the best student. A high negative score indicates that, in almost all the subjects, the student obtained scores that are lower than the mean scores in the respective subjects. Thus, the student with the highest negative component score is the weakest student. A score that is close to zero indicates that, in almost all the subjects, the student obtained scores that are just about the mean scores in the respective subjects.

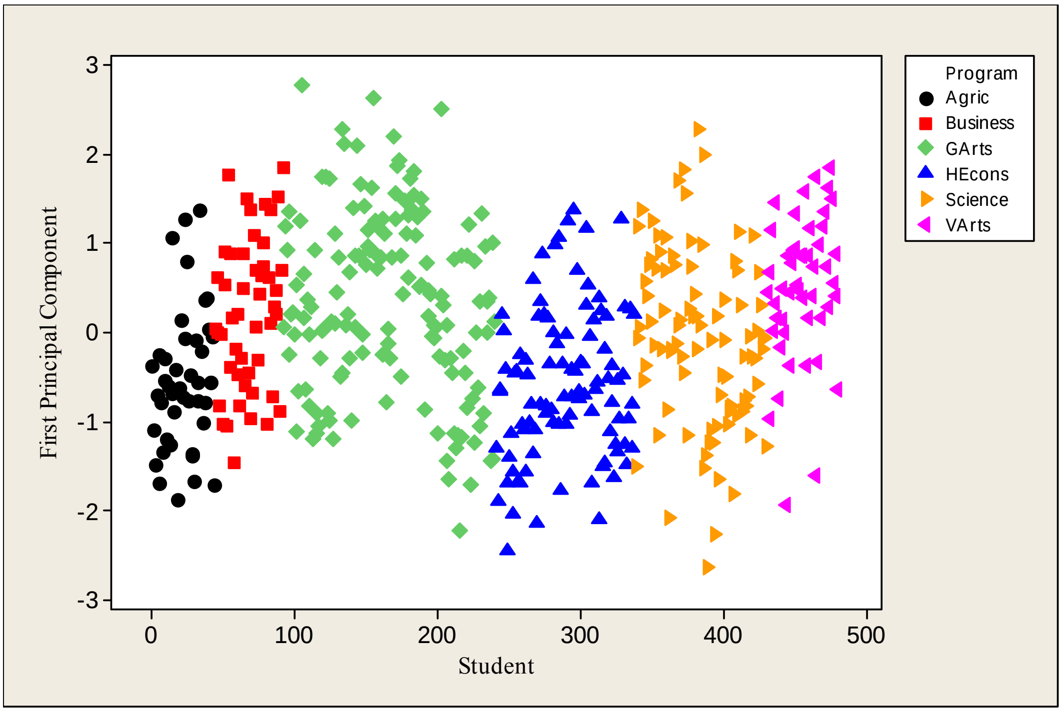


Figure 8: First Principal Component Score against Corresponding Student

From Figure 8, we see that the first three best students are all from the General Arts class. We also see that the worst student is from Science, the second worst student is from Home Economics and the third worst student is from Science. In Appendix E, we have the same plot as in Figure 8, with each plot labeled with the corresponding number of the student. From that table, we compile in Table 10 the list of the best three and the worst three performances with their respective programmes and their scores in the seven subjects.

Table 10 shows that the best student is not the one who consistently obtained the highest score in all the subjects; the student actually obtained the highest score (or one of the highest scores) in only Integrated Science. It can also be seen that the worst student did not consistently obtain the lowest scores in all the subjects; in fact, the worst student obtained the lowest score in only French (see also Figure 7). The same pattern of performance could be said of the other performances in the table. Consistent with the observation in the previous section, we note that the third best student obtained the highest scores in the two most influential subjects, Core Mathematics and French. However, the best student, who also obtained almost equally high scores in the two important subjects, performed much better in the remaining two important subjects, Science and English Language. Similar explanation could be given to the performance of the worst student. We observe that the student is not the poorest in five of the seven subjects. However, his/her score in the second most important subject, French, was much less than the mean performance in the subject. The effect of this low score is that it leads to a small component score which ranked as the last.

From Figure 8, the distribution of the component scores appears to be normal for Business, General Arts, and Science. We see that the Agriculture class is typically below average compared to the rest of the students. By this feature, the Agriculture class appears to be the weakest class among the six classes. The performance of the Home Economics class is similar to the Agriculture class. In the Visual Arts class, the distribution of performance is negatively skewed, indicating that there are more students in that class who are above the general average performance. It appears that it is only the Visual

Arts class that has this feature. Thus, the Visual Arts class appears to be the strongest class among the six classes.

Summary of Further Analysis

This chapter made use of the Principal Components method to further analyze the scores data. A number of interesting results have been revealed in this chapter. Principal component analysis identified the first component as a weighted sum of all the subjects offered by the students. As a result, the first component was found to be the most appropriate index in determining the general performances of the students. Core Mathematics and French were observed to be the two most influential subjects in the formation of the first component. Thus, in the determination of general performance of students, Core Mathematics and French are the most influential of all the seven subjects. The first principal component alone explains about 42 percent of the total variation in the scores data.

The first principal component scores show that the three best students were members of the General Arts class. In order of decreasing performance, the three best students were those with numbers 105, 155 and 203. The three worst performances were recorded by Science and Home Economics students: the worst score was recorded by a Science student with number 388; the second worst score was obtained by a student of Home Economics class with number 248; and the third worst student was from the Science class with number 393. A plot of the component scores suggested that in general, the Visual Arts class is the strongest, the Agriculture class is the weakest, whiles the performances by the Business, General Arts, and the Science classes were average.

CHAPTER FIVE

SUMMARY, DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

In this chapter we take a second look at the summaries that we noted after the preliminary and the further analysis. We will then have a discussion on some results that were obtained in the analysis. The final section of this chapter is devoted to the conclusion that we have drawn from this study along with some recommendations based on the conclusion.

Summary

This study sought to determine an index that could be used to identify best and abysmal performances of students in examinations involving seven common subjects taken at the first year in Ghana Senior High School in Koforidua. The scores of the students, who were grouped into six different programme classes, in the seven subjects constituted the data for the study. Since each of the students had a score on each of the seven subjects (variables), the data obtained constituted a multivariate data set.

Preliminary analyses of the data made use of routine exploratory methods such as a study of the Variance-Covariance matrix of the seven subjects and the distributions of the scores in the subjects for each of the six programme classes in the school. Preliminary analysis also studied the main sources of variation in the performance in each of the subjects. The Variance-Covariance matrix showed that performance was most dispersed in Core Mathematics, and then in French, whilst it was least dispersed in Social

Studies. It was also realized that performance in Core Mathematics and Science were more related than any other pair of subjects.

A study of the distribution of performance in the various programmes identified Visual Arts to be generally the most superior in the two language subjects, English Language and French. However, students in Visual Arts have the tendency of recording extreme scores, with exception of ICT.

Performance of students in General Arts was typically widely dispersed in all the subjects. It produced the best results in four out of the seven subjects. These are the ICT, Physical Education, Integrated Science and Core Mathematics. It also produced the second highest scores in the two languages, English Language and French. It, however, recorded the lowest score in Physical Education. Similar to the General Arts, the performance in the Home Economics is characterized by wide disparity in performance in all the subjects. The Home Economics class is particularly noted for recording the least scores in three of the subjects. These are the ICT, Integrated Science and Core Mathematics.

The Science class appears to be the weakest in the languages, English and French, recording the least scores in the two subjects. Generally, the class also exhibits a lot of variability in a manner quite similar to that of the Home Economics class. The performance in the Agriculture and Business classes were the least dispersed in all the subjects. Thus, the students in those two classes are the most competitive within their respective classes. They are, however, not competitive among the entire groups of programmes.

The result of the analysis of variance revealed that the average performance in each of the subjects differ from programme to programme.

This shows that students performance is largely attributable to the respective programmes that they offer.

In order to analyze this high dimensional data set, a multivariate data analysis technique, the Principal Components Analysis, was used as the main statistical tool to examine the Variance-Covariance matrix of the performance in the subjects. The result of the study of the Variance-Covariance matrix and the main objective of the study further informed the choice of the technique.

A number of interesting results were revealed in the further analysis of the data using the Principal Component analysis method. The technique identified the first component as a weighted sum of all the subjects offered by the students. As a result, the first component was found to be the most appropriate index in determining the general performances of the students. Core Mathematics and French were observed to be the two most influential subjects in the formation of the first component. Thus, in the determination of general performance of students, Core Mathematics and French are the most influential of all the seven subjects. The first principal component alone explained about 42 percent of the total variation in the scores data.

Using the scores of the first principal component, it was realized that the three best students were all members of the General Arts class. In order of decreasing performance, the three best students were those with numbers 105, 155 and 203. The three worst performances were recorded by Science and Home Economics students: the worst score was recorded by a Science student with number 388; the second worst score was obtained by a student of Home Economics class with number 248; and the third worst student was from the Science class with number 393.

A plot of the component scores suggested that in general, the Visual Arts class is the strongest class. However, the Agriculture class is the weakest among the six classes. The distribution of scores of the Business, General Arts, and the Science classes suggested that their performances were average.

Discussion

A couple of results obtained in this work need some amount of discussion. These discussions will focus on issues that are concerned with the disparity in results in the use of Principal Component method, and what would have been achieved if ordinary sums of scores were used. Another issue worth discussing is the identification of the Visual Arts class as the strongest among the six classes.

To identify best and worst students in examinations involving multiple subjects, the usual practice has been to find the sum of scores a student obtains in all the subjects. The student that obtains the highest sum is adjudged the best in the examination. On the other hand, the student who obtains the least sum is the poorest student. The use of this approach implies that all the subjects are given equal importance in the overall ranking of the students. However, in this study, by using the first principal component as an index, we do not consider the seven subjects to be equally important in the assessment of the students. The loadings of the first principal component on the subjects indicated that Core Mathematics weighted as the most influential in the method used in assessing students. The next most influential subject is French. Science and English Language had the third and fourth highest weights. The subject that contributes the least in performance assessment is Social Studies.

This means that the best students in the examinations are those who are generally strong in the Core Mathematics, and French, and then to a lesser extent, Science and English Language. The scores for the best three students from Table 10 shows that, from the best student, the sums are 619, 620 and 596. That is, the sum of scores of the best student is just one less than that of the second best student. However, the principal component scores of these students, from the best student, are 2.77529, 2.63602 and 2.51586. The disparity between the ranks of the two best students by the two methods is as a result of the unequal weighting of the subjects in the construction of the principal component. The best student has been so adjudged since he/she consistently obtained high scores in the four most influential subjects mentioned earlier, which translated in the highest component score.

The study identified the Visual Arts as the strongest class in the examinations. This is because in the preliminary analysis we identified this class to be much more superior in French. The class was also the best in English Language and Integrated Science. Besides, it was one of the best classes in Mathematics. Thus, we realize that when it comes to the subjects that matter in performance assessment in this study, the Visual Arts class performed consistently well.

The Agriculture class, in contrast to the Visual Arts class, was the weakest in French. In addition, the class also was one of two weakest classes in Core Mathematics and Integrated Science. Thus, the Agriculture class is rightly considered as the poorest in the examinations.

Currently, performance in Mathematics, English Language and Science after secondary school is given a priority consideration in admission requirements into tertiary institutions, and in job placements. This practice is in line with the findings of this research. However, these three subjects appear to be given equal weights in such requirements. The study shows that although all these three subjects are important, differential weights could be assigned to each of them. The study also brings to light the importance of considering French as a fourth subject in such candidate evaluation. Since French has been identified as the second most influential subject in this work, its inclusion may enhance the effectiveness of assessment procedures of candidates' competence.

Conclusions and Recommendations

The study looked at identification of outstanding and abysmal performances of students in examinations involving seven common subjects taken at the first year in Ghana Senior High School. The scores of the students, who were grouped into six different programme classes, in the seven subjects constituted the data for the study. Since each of the students had a score on each of the seven subjects (variables), the data obtained constituted a multivariate data set. The objective of this study has been to determine an index for identifying the best and abysmal performances in the examination scores. In order to achieve this objective using this high dimensional data set, a multivariate data analysis technique, the Principal Components Analysis, was used as the main statistical tool to examine the Variance-Covariance matrix of the performance in the subjects.

The study revealed that the first principal component is a weighted sum of all the subjects offered by the students. As a result, the first component was found to be the most appropriate index in determining the general performances of the students. Core Mathematics and French were observed to be the two most influential subjects in the formation of the first component. Thus, in the determination of general performance of students, Core Mathematics and French are the most influential of all the seven subjects.

Using the scores of the first principal component, it was realized that the three best students were all members of the General Arts class. In order of decreasing performance, the three best students were those with numbers 105, 155 and 203. The worst scores were recorded by a Science student with number 388; the second worst scores were obtained by a student of Home Economics class with number 248; and the third worst student was from the Science class with number 393. The study also revealed that in general, the Visual Arts class is the strongest class whilst the Agriculture class is the weakest among the six classes. The performances of Business, General Arts, and the Science classes were quite average.

The findings of this dissertation shows that in determining relative performances of students in examinations, one need not consider all the subjects of examination as equally important. The consideration of Mathematics, English Language and Science as priority subjects currently in admission requirements into tertiary institutions, in particular, is in line with the findings of this work. This study, however, recommends that the inclusion of French as a fourth subject may enhance the effectiveness of assessment procedures of candidates.

REFERENCES

Antwi, M. K. (1992). *Education, Society and Development in Ghana*. Accra: Unimax Publications.

Birren, J. E., & Morrison, D. F. (1961). Analysis of WAIS subsets in relation to age and education. *Journal of Gerontology*, *16*, 363-369.

Carpenter, T. P., Cobum, T. G., Reys, R. E. & Wilson, J. W. (1976). Research implications and questions from the year 04 NAEP Mathematics Assessment. *Journal for Research in Mathemaics Education. 7*(5), 327-336.

Cattell, R. B. (1996). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Gordor, B. K., & Howard, N. K. (2000). *Elements of Statistical Analysis*. Accra: Mathematics Group.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*, 417-441, 498-520.

Jolliffe, I. T. (1986). *Principal Component Analysis*. New York: Springer-Verlag.

Johnson, R. A & Wichern, D. W. (1992). *Applied Multivariate Statistical Analysis.* (3rd ed.). London: Prentice-Hall International, Inc.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 23*, 187-200.

Malhotra N. K. (2004). *Applied Orientation In Marketing Research.* (4th ed.). London: Pearson Prentice-Hall.

Merret, F., & Tang, W. M. (1994). The attitudes of British primary school pupils to praise, rewards, punishments, and reprimands. *British Journal of Educational Psychology, 64*, 91-103.

Morrison, D. F. (1974). *Multivariate Statistical Methods.* New York: McGraw-Hill.

Sharma, S. (1996). *Applied Multivariate Techniques*. New York: John Willey & Sons, Inc.

Sullivan, J. H., & Woodall, W. H. (1996). A Control Chart for Preliminary Analysis of Individual Observations. *Journal of Quality Technology, 28*, 265-278.

APPENDICES

APPENDIX A: Least Significant Difference Results

A: LSD Results for English Language

Difference between Agriculture (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Business	-6.0492	0.007
Arts	-11.2193	0.000
HEcons	5.4194	0.005
Science	-1.5474	0.426
VArts	-13.4709	0.000

Difference between Business (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Agriculture	6.0492	.007
Arts	-5.1700	.003
HEcons	11.4686	.000
Science	4.5018	.018
VArts	-7.4217	.001

Difference between General Arts (i) and other programmes

Agriculture	11.2193	0.000
Business	5.1700	0.003
HEcons	16.6387	0.000
Science	9.6719	0.000
VArts	-2.2516	0.195

Difference between Home Economics (i) and other programmes

Agriculture	-5.4194	.005
Business	-11.4686	.000
Arts	-16.6387	.000
Science	-6.9668	.000
VArts	-18.8903	.000

Difference between Science (i) and other programmes

Agriculture	1.5474	.426
Business	-4.5018	.018
Arts	-9.6719	.000
HEcons	6.9668	.000
VArts	-11.9235	.000

A continued

Difference between Visual Arts (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Agriculture	1.5474	.426
Business	-4.5018	.018
Arts	-9.6719	.000
HEcons	6.9668	.000
VArts	-11.9235	.000

B: LSD Results for Core Mathematics

Difference between Agriculture (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Business	-13.2424	.000
Arts	-7.6726	.004
HEcons	-2.7029	.330
Science	-8.2243	.003
VArts	-13.6691	.000

Difference between Business (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Agriculture	13.2424	.000
Arts	5.5698	.029
HEcons	10.5395	.000
Science	5.0181	.065
VArts	-.4267	.890

Difference between General Arts (i) and other programmes

Agriculture	7.6726	.004
Business	-5.5698	.029
HEcons	4.9697	.013
Science	-.5517	.786
VArts	-5.9965	.017

B continued

Difference between Home Economics (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Agriculture	2.7029	.330
Business	-10.5395	.000
Arts	-4.9697	.013
Science	-5.5214	.013
VArts	-10.9662	.000

Difference between Science (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Agriculture	8.2243	.003
Business	-5.0181	.065
Arts	.5517	.786
HEcons	5.5214	.013
VArts	-5.4448	.043

Difference between Visual Arts (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Agriculture	13.6691	.000
Business	.4267	.890
Arts	5.9965	.017
HEcons	10.9662	.000
Science	5.4448	.043

C: LSD Results for Integrated Science

Difference between Agriculture (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Business	-5.2652	.034
Arts	-5.8102	.005
HEcons	3.8800	.073
Science	-4.9209	.024
VArts	-9.0818	.000

Difference between Business (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Agriculture	5.2652	.034
Arts	-.5450	.783
HEcons	9.1452	.000
Science	.3442	.871
VArts	-3.8167	.113

Difference between General Arts (i) and other programmes

Agriculture	5.8102	.005
Business	.5450	.783
HEcons	9.6902	.000
Science	.8892	.573
VArts	-3.2716	.093

Difference between Home Economics (i) and other programmes

Agriculture	-3.8800	.073
Business	-9.1452	.000
Arts	-9.6902	.000
Science	-8.8010	.000
VArts	-12.9619	.000

Difference between Science (i) and other programmes

Agriculture	4.9209	.024
Business	-.3442	.871
Arts	-.8892	.573
HEcons	8.8010	.000
VArts	-4.1609	.047

C continued

Difference between Visual Arts (*i*) and other programmes

Programme (*j*)	Mean Difference ($i-j$)	Significance
Agriculture	9.0818	.000
Business	3.8167	.113
Arts	3.2716	.093
HEcons	12.9619	.000
VArts	4.1609	.047

D: LSD Results for Social Studies

Difference between Agriculture (*i*) and other programmes

Programme (*j*)	Mean Difference ($i-j$)	Significance
Business	-4.7936	.019
Arts	-1.7727	.289
HEcons	-8.6825	.000
Science	-11.1749	.000
VArts	-1.0427	.605

Difference between Business (*i*) and other programmes

Programme (*j*)	Mean Difference ($i-j$)	Significance
Agriculture	4.7936	.019
Arts	3.0208	.062
HEcons	-3.8890	.024
Science	-6.3813	.000
VArts	3.7508	.057

Difference between General Arts (*i*) and other programmes

Agriculture	1.7727	.289
Business	-3.0208	.062
HEcons	-6.9098	.000
Science	-9.4022	.000
VArts	.7300	.647

D continued

Difference between Home Economics (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Agriculture	8.6825	.000
Business	3.8890	.024
Arts	6.9098	.000
Science	-2.4924	.079
VArts	7.6398	.000

Difference between Science (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Agriculture	11.1749	.000
Business	6.3813	.000
Arts	9.4022	.000
HEcons	2.4924	.079
VArts	10.1322	.000

Difference between Visual Arts (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Agriculture	1.0427	.605
Business	-3.7508	.057
Arts	-.7300	.647
HEcons	-7.6398	.000
Science	-10.1322	.000

E: LSD Results for Physical Education

Difference between Agriculture (*i*) and other programmes

Programme (*j*)	Mean Difference ($i - j$)	Significance
Business	.2121	.921
Arts	.0590	.973
HEcons	9.6821	.000
Science	11.0237	.000
VArts	14.3955	.000

Difference between Business (*i*) and other programmes

Programme (*j*)	Mean Difference ($i - j$)	Significance
Agriculture	-.2121	.921
Arts	-.1532	.928
HEcons	9.4699	.000
Science	10.8116	.000
VArts	14.1833	.000

Difference between General Arts (*i*) and other programmes

Agriculture	-.0590	.973
Business	.1532	.928
HEcons	9.6231	.000
Science	10.9647	.000
VArts	14.3365	.000

Difference between Home Economics (*i*) and other programmes

Agriculture	-9.6821	.000
Business	-9.4699	.000
Arts	-9.6231	.000
Science	1.3417	.368
VArts	4.7134	.008

Difference between Science (*i*) and other programmes

Agriculture	-11.0237	.000
Business	-10.8116	.000
Arts	-10.9647	.000
HEcons	-1.3417	.368
VArts	3.3717	.061

E continued

Difference between Visual Arts (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Agriculture	-14.3955	.000
Business	-14.1833	.000
Arts	-14.3365	.000
HEcons	-4.7134	.008
VArts	-3.3717	.061

F: LSD Results for ICT

Difference between Agriculture (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Business	-6.5587	.004
Arts	-6.5491	.000
HEcons	7.1640	.000
Science	-3.0632	.120
VArts	8.1455	.000

Difference between Business (i) and other programmes

Programme (j)	Mean Difference ($i - j$)	Significance
Agriculture	6.5587	.004
Arts	.0096	.996
HEcons	13.7227	.000
Science	3.4955	.068
VArts	14.7042	.000

Difference between General Arts (i) and other programmes

Agriculture	6.5491	.000
Business	-.0096	.996
HEcons	13.7132	.000
Science	3.4859	.015
VArts	14.6946	.000

F continued

Difference between Home Economics (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Agriculture	-7.1640	.000
Business	-13.7227	.000
Arts	-13.7132	.000
Science	-10.2273	.000
VArts	0.9814	.600

Difference between Science (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Agriculture	3.0632	.120
Business	-3.4955	.068
Arts	-3.4859	.015
HEcons	10.2273	.000
VArts	11.2087	.000

Difference between Visual Arts (i) and other programmes

Programme (j)	Mean Difference ($i-j$)	Significance
Agriculture	-8.1455	.000
Business	-14.7042	.000
Arts	-14.6946	.000
HEcons	-.9814	.600
Science	-11.2087	.000

G: LSD Results for French

Difference between Agriculture (*i*) and other programmes

Programme (*j*)	Mean Difference ($i - j$)	Significance
Business	-5.6439	.053
Arts	-15.3219	.000
HEcons	-9.5829	.000
Science	-7.1729	.005
VArts	-25.8473	.000

Difference between Business (*i*) and other programmes

Programme (*j*)	Mean Difference ($i - j$)	Significance
Agriculture	5.6439	.053
Arts	-9.6779	.000
HEcons	-3.9390	.110
Science	-1.5290	.539
VArts	-20.2033	.000

Difference between General Arts (*i*) and other programmes

Agriculture	15.3219	.000
Business	9.6779	.000
HEcons	5.7389	.002
Science	8.1489	.000
VArts	-10.5254	.000

Difference between Home Economics (*i*) and other programmes

Agriculture	9.5829	.000
Business	3.9390	.110
Arts	-5.7389	.002
Science	2.4100	.236
VArts	-16.2643	.000

Difference between Science (*i*) and other programmes

Agriculture	7.1729	.005
Business	1.5290	.539
Arts	-8.1489	.000
HEcons	-2.4100	.236
VArts	-18.6743	.000

APPENDIX B: Correlation Matrix of the Study Variables

Var	English	Maths	Science	Soc. Stds	PE	ICT
Maths	0.363					
Sig.	0.000					
Science	0.431	0.636				
Sig.	0.000	0.000				
Soc. Stds.	-0.059	0.255	0.217			
Sig.	0.197	0.000	0.000			
PE	0.167	0.302	0.302	0.119		
Sig.	0.197	0.000	0.066	0.000		
ICT	0.425	0.265	0.356	-0.084	0.181	
Sig.	0.000	0.000	0.000	0.066	0.000	
French	0.360	0.380	0.342	-0.002	0.033	0.125
Sig.	0.000	0.000	0.000	0.000	0.000	0.000

APPENDIX C: Standardised Component Score Coefficient Matrix

Variable	Component						
	1	2	3	4	5	6	7
English	0.203	-0.117	0.452	-0.059	0.584	-0.896	0.309
Maths	0.417	0.310	-0.403	-0.367	-0.842	-0.272	0.742
Science	0.251	0.181	-0.024	-0.174	0.121	0.005	-1.351
Soc. Stds	0.042	0.212	-0.351	-0.199	0.865	0.310	0.314
PE	0.103	0.343	0.030	1.005	0.056	0.042	0.081
ICT	0.144	0.114	0.617	-0.211	-0.044	0.885	0.276
French	0.306	-0.864	-0.274	0.382	0.113	0.470	0.044

APPENDIX D: First Principal Component Scores

Student	Score	Stud	Score	Stud	Score	Stud	Score
1	-0.36967	33	-0.76065	65	0.49164	97	-0.25375
2	-1.09998	34	1.38540	66	-0.59150	98	0.20348
3	-1.47343	35	-0.19798	67	1.49469	99	0.21782
4	-0.70289	36	-1.00025	68	-0.45128	100	-0.01961
5	-0.25057	37	0.37810	69	-0.97363	101	0.53226
6	-1.69289	38	-0.78702	70	1.37090	102	-1.10392
7	-0.78482	39	0.38833	71	-0.66910	103	-0.65835
8	-1.32886	40	0.03448	72	1.08219	104	1.25911
9	-0.28893	41	-0.56273	73	0.06974	105	2.77529
10	-0.53777	42	-0.54910	74	0.69609	106	0.64845
11	-1.18917	43	-0.04600	75	-0.29969	107	0.16219
12	-0.59048	44	-1.70272	76	0.42792	108	-0.63847
13	-1.24855	45	0.03865	77	0.64544	109	0.37367
14	1.06179	46	0.60998	78	0.74317	110	-0.82262
15	-0.68910	47	0.00648	79	1.00367	111	-0.02946
16	-0.88567	48	-0.82126	80	1.43402	112	0.28246
17	-0.41215	49	-0.01282	81	-1.02069	113	-1.19410
18	-0.65719	50	-1.03106	82	0.62397	114	0.92194
19	-1.87585	51	0.90694	83	1.37327	115	-0.90674
20	-0.62663	52	0.53216	84	0.10867	116	-1.12308
21	0.14249	53	-1.05026	85	-0.71914	117	-1.04563
22	-0.71201	54	1.77270	86	0.28403	118	-0.28217
23	-0.05429	55	-0.38734	87	0.20041	119	1.74751
24	1.26734	56	0.88109	88	0.47002	120	0.02690
25	0.80821	57	0.16533	89	1.53122	121	0.05066
26	-0.75745	58	-1.46207	90	-0.88081	122	1.74018
27	-0.46352	59	-0.19218	91	0.69850	123	-0.90735
28	-1.38206	60	-0.47494	92	1.85993	124	1.72202
29	-1.35116	61	0.20344	93	0.06024	125	-0.98842
30	-1.67665	62	-0.82037	94	1.19136	126	0.12615
31	-0.08793	63	-0.29234	95	0.93121	127	-1.19067
32	-0.54977	64	0.87752	96	1.35252	128	1.10434

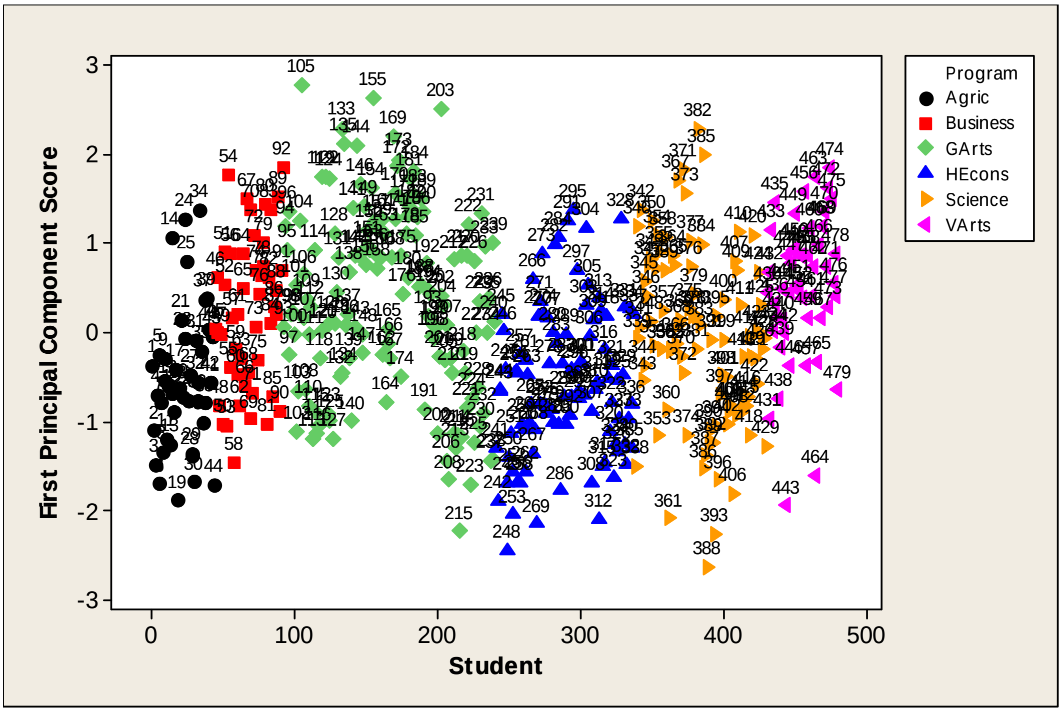
Student	Score	Stud	Score	Stud	Score	Stud	Score
129	0.09799	161	1.28620	193	0.18195	225	-1.22612
130	0.45938	162	-0.25454	194	0.47009	226	0.81182
131	0.84421	163	1.10690	195	0.43711	227	-0.00242
132	-0.48922	164	-0.77863	196	-0.04879	228	-0.61770
133	2.29500	165	0.04834	197	0.08980	229	0.34943
134	-0.45637	166	-0.11702	198	-0.06297	230	-1.05713
135	2.12241	167	-0.28765	199	0.06063	231	1.34303
136	0.08921	168	0.85369	200	-1.13364	232	-0.90319
137	0.20456	169	2.19790	201	-0.27257	233	0.96391
138	0.68559	170	1.56912	202	0.41119	234	0.00676
139	-0.29117	171	1.27287	203	2.51586	235	0.35580
140	-0.99429	172	1.87648	204	0.31797	236	0.39113
141	1.40629	173	1.94473	205	-0.31404	237	-1.43124
142	0.85976	174	-0.48554	206	-1.44202	238	-1.41349
143	0.06565	175	0.85834	207	0.08836	239	1.01668
144	2.10460	176	0.43869	208	-1.65653	240	0.12515
145	0.87103	177	1.48333	209	-0.27904	241	-1.28828
146	1.67514	178	1.11837	210	-0.45244	242	-1.89009
147	-0.22675	179	1.31838	211	-1.13736	243	-0.65827
148	-0.01433	180	0.61782	212	0.81359	244	-0.64111
149	1.41804	181	1.72382	213	-1.29866	245	0.21518
150	0.75478	182	1.37863	214	-1.14427	246	0.01379
151	0.96773	183	1.54544	215	-2.22323	247	-0.41083
152	1.14887	184	1.82008	216	-1.17784	248	-2.44240
153	0.91888	185	1.10104	217	0.87185	249	-1.68550
154	1.63190	186	1.28977	218	-0.21564	250	-1.39692
155	2.63602	187	0.52437	219	-0.44883	251	-1.14168
156	0.89341	188	0.53345	220	0.86108	252	-1.57220
157	1.26445	189	1.49833	221	-0.84609	253	-2.03991
158	0.73027	190	1.36757	222	1.20928	254	-0.44831
159	1.17861	191	-0.86342	223	-1.70801	255	-0.43002
160	0.84493	192	0.77224	224	-0.73274	256	-1.66042

Student	Score	Stud	Score	Stud	Score	Stud	Score
257	-0.24171	289	-0.02596	321	-0.37172	353	-1.15693
258	-1.67916	290	-1.03342	322	-0.78256	354	1.09126
259	-1.02903	291	1.24684	323	-1.63495	355	0.90081
260	-1.09863	292	-0.92755	324	-1.25232	356	-0.19393
261	-0.31343	293	-0.71791	325	-0.53663	357	0.25523
262	-1.56235	294	-0.40314	326	-1.33051	358	1.07941
263	-0.47256	295	1.37978	327	-0.95888	359	0.70129
264	-1.00813	296	-0.43086	328	1.28013	360	-0.87109
265	-0.79373	297	0.69996	329	-0.47570	361	-2.08625
266	0.60571	298	-0.66294	330	-1.25653	362	-0.19562
267	-1.35598	299	-0.75096	331	0.28075	363	-0.18093
268	-1.09776	300	-0.33734	332	-1.47585	364	0.86570
269	-2.14630	301	-0.35137	333	-0.96745	365	0.75727
270	0.18731	302	-0.69210	334	0.26016	366	-0.12932
271	0.34809	303	0.30629	335	-1.29733	367	1.70007
272	-0.81288	304	1.17757	336	-0.80932	368	0.09028
273	0.88131	305	0.53544	337	0.20028	369	0.14721
274	0.20757	306	-0.04404	338	-1.49639	370	-0.27162
275	-0.90137	307	-0.88380	339	-0.07182	371	1.83987
276	-0.83257	308	-1.67833	340	1.20399	372	-0.45256
277	0.17309	309	0.14073	341	0.06316	373	1.56808
278	-0.35935	310	-0.64080	342	1.38880	374	-1.15228
279	-0.86999	311	-0.54674	343	-0.54184	375	0.26343
280	-0.00212	312	-2.09025	344	-0.37547	376	0.73326
281	-1.01259	313	0.38170	345	0.56971	377	1.02806
282	0.98198	314	0.23885	346	0.40930	378	0.18701
283	-0.11389	315	-1.50049	347	0.75271	379	0.43165
284	1.06451	316	-0.19291	348	0.12546	380	0.19357
285	-1.01996	317	-1.45747	349	0.82043	381	-0.18833
286	-1.77406	318	0.18037	350	1.24805	382	2.28799
287	-0.34496	319	-0.51045	351	-0.14669	383	0.07547
288	-0.71334	320	-1.10545	352	0.72774	384	0.98125

Student	Score	Stud	Score	Stud	Score	Stud	Score
385	1.98719	413	-0.27425	441	0.48942	469	1.19964
386	-1.52690	414	-0.75586	442	-0.00763	470	1.35498
387	-1.38830	415	-0.81365	443	-1.94467	471	0.74795
388	-2.63130	416	-0.71355	444	0.86464	472	1.63037
389	-1.23986	417	-0.04706	445	0.46160	473	0.29092
390	-1.08274	418	-1.14951	446	-0.37222	474	1.84660
391	-0.07557	419	-0.26245	447	0.78459	475	1.50063
392	-1.23359	420	1.09052	448	0.93291	476	0.56250
393	-2.25706	421	-0.27868	449	1.34017	477	0.40635
394	-1.04535	422	-0.58040	450	0.94106	478	0.88415
395	0.18474	423	0.01882	451	0.53245	479	-0.64539
396	-1.65276	424	0.67407	452	0.46346		
397	-0.70093	425	0.31217	453	0.86093		
398	-0.47860	426	-0.17613	454	0.38256		
399	-0.07206	427	-0.03622	455	0.86109		
400	0.37439	428	-0.07101	456	1.57850		
401	-0.49397	429	-1.27316	457	-0.38048		
402	-1.03210	430	0.44693	458	0.87457		
403	-0.83421	431	-0.96074	459	0.17381		
404	-0.88554	432	0.67615	460	1.16634		
405	-0.82817	433	1.16169	461	0.40238		
406	-1.81058	434	0.01041	462	0.73734		
407	0.79955	435	1.47135	463	1.74959		
408	-0.92748	436	0.32355	464	-1.59779		
409	0.70677	437	0.17227	465	-0.33049		
410	1.13128	438	-0.74268	466	0.99568		
411	0.30082	439	-0.15762	467	0.15958		
412	-0.89007	440	0.13478	468	1.19548		

APPENDIX E

APPENDIX D


First Principal Component Score
Student
Program
Agric
Business
GArts
HEcons
Science
VArts