# ANALYSING DATA FOR EDUCATIONAL RESEARCH: ANALYSIS OF MATHEMATICS SCORES AT KENYA CERTIFICATE OF SECONDARY EDUCATION, 2011, FOR WESTERN KENYA

BY

ZACHARIAH LUSENAKA MBASU

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN APPLIED STATISTICS

DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

MASENO UNIVERSITY

## Abstract

Analysing Kenya Certificate of Secondary Education (KCSE) exam results is vital to inform the design of interventions to improve teaching and learning of Mathematics in Kenyan schools. Here, further analysis of the 2011 Kenya Certificate of Secondary Education results explores patterns in the results and factors that affected the grades scored in Mathematics by students. KCSE results for 699 schools with a total of 50,584 candidates in Western Kenya region were collected and further analysis was done. Descriptive analysis of the data involved calculation and presentation of summary statistics using tables and graphs produced using R. This study demonstrates the application of two approaches to analyse the clustered data to compare mathematics scores between the different school types and between genders. The first approach, Analysis of Variance (ANOVA), assumes independence of observations while the second approach, Residual Maximum likelihood (REML) recognises dependency of observations. Analysis of the data revealed that Mathematics was generally poorly performed by students. The results indicated that students in single gender schools performed better in Mathematics than those in mixed schools while boys performed better than girls. The type of school and gender of the student were found to have a significant effect on the candidate's Mathematics score. The parameter estimates are observed to be lower when the random term is included in the Linear mixed model that uses Residual Maximum Likelihood. The interpretation and significance of fixed and random effects in the model is discussed and areas for further research are highlighted.

# Chapter 1: Introduction

## 1.1: Overview

Public examinations in Kenyan schools like Kenya Certificate of Secondary Education (KCSE) exam are used as a selection instrument for further education and training (Lucas, 1993). The analysis of KCSE exam marks from the recent years forms a baseline, through which we should be able to assess the value of interventions to improve teaching and learning in schools. For example, to know whether action plans to improve learning had an effect on the various school types and whether the effect varies with gender of schools.

As initiatives are put in place to improve teaching and learning of mathematics, there is a need to understand the many factors that are associated with the performance of the subject in schools. Some of the factors could be the type of school or the characteristics of the students. This will contribute highly to the improvement of the design of project interventions like Strengthening of Mathematics and Science in Secondary Education (SMASSE) (Japan International Cooperation Agency, 2007) and maths camps (African Maths Initiative, 2013) that are geared to improve teaching and learning of mathematics.

This study is aimed at understanding patterns in student KCSE results with respect to factors associated with KCSE performance. When a data set has student and school level factors, we must keep in mind the multilevel (Gray, 1989) data structure;

The ANOVA method is applied to study variation in students' Mathematics scores in the KCSE exam alongside other factors including the school "type" and the student's gender. Key assumptions in the ANOVA method are that the population of values of the maths score variable associated with the school types and genders have equal variances that they follow normal distribution and that the variables are independent. However, when data such as maths scores for schools are collected, the residuals are unlikely to be independent of each other (Cohen, et al., 2003). For example, we would expect mathematics scores within a school to be more similar than in a completely random sample of students. This is because students in the same school are more likely to share a common curriculum, common mathematics textbooks, common teachers, as well as other school and community resources, than a random sample of students drawn across schools. A major concern when using ANOVA to estimate relationships on clustered data is that the estimated standard errors will be too small, leading to an overestimation of the statistical significance of regression coefficients.

Mixed model analyses provide a generally, flexible approach in these situations because they allow a wide variety of correlation patterns to be explicitly modelled (Kreft, et al., 1998). Measurements of scores per student generally result in correlated errors that break the assumptions of standard (between students) ANOVA and regression models. In recent years a general algorithm known as Restricted Maximum Likelihood (REML) has been developed for estimating variance parameters in linear mixed models (LMM). In this study, an ANOVA technique was reviewed and the use of a Linear mixed model (REML) is demonstrated.

For balanced data, REML reproduces the same parameters as those of ANOVA because the assumptions made when using a regression ANOVA (independent normally distributed errors with constant variance) fit within a LMM (REML) framework but the procedure is not dependent on balance. Unlike ANOVA, it allows for spatial correlations, changing variance when the school types have a changing variance structure.

Student scores in an exam may be influenced by a number of factors. These factors include characteristics and background of a student, characteristics of the school where the student learns and the economic and social context in which the school operates.

This study examines relationships between a student's Mathematics score in the KCSE exam and the school type and student's gender. It demonstrates how statistical techniques like ANOVA and Linear mixed model (Eleanor, et al., 2001) use population means and variances to test uniformity/homogeneity of Mathematics scores for student and school variables. The aim was to compare the maths score for school types; and to bring out any differences that may exist within the populations in such a way as to enable a statistician to make a decision that either there are significant differences, or that differences do not exist.

## 1.2: Statement of the problem

It is expected that the implementation of intervention strategies will translate into better students' Mathematics scores and better performance in the KCSE exam. These KCSE examination results are released every year by the Kenya National examination council where summaries of regions, schools, subjects and students are presented to the public. Teachers in schools do similar summaries that include calculating averages of subject scores, number of students and frequencies of grades.

These summaries are not enough to create patterns and interpret the results, creating a need to take this kind of analysis further as a means to inform and improve policy level decision making. Doing an in-depth analysis like comparing maths score by students in various types of school (single gender, mixed etc.), should enable further studies to be designed to measure the impact(s) of strategies implemented to improve teaching and learning of Mathematics and KCSE performance in different types of secondary schools.

## 1.3: Objectives

The main objective of the study was to do an in-depth analysis of KCSE results by integrating fixed and random effects in a linear mixed model to study variations in students' KCSE Mathematics scores. The specific objectives were:

- To check the degree of any relationship between KCSE mathematics scores and other subject scores in Western Kenya.
- To compare differences in KCSE mathematics scores between single gender schools and mixed schools in Western Kenya.
- To compare differences in KCSE mathematics scores between male and female candidates in Western Kenya.

## 1.4: Research Questions

1. How does the KCSE maths score relate to the aggregate point scored by a student? i.e. Does Mathematics affect the number of students getting direct entry into public university?
2. Is there any difference in the KCSE maths score between different school types? i.e. does the school type affect the mathematics grade scored by a student?

3. Is there any difference in KCSE mathematics score achieved by girls and boys? i.e. does gender affect the mathematics score achieved by a student?

## 1.5: Significance of the study

To achieve better educational outcomes, research findings are relied upon during education decision-making processes. There is, therefore, a need for ascertaining variables that affect Mathematics scores as well as identifying any interactions between these variables. This study has contributed to the availability of reliable information helpful to teachers, educational researchers and public decision makers because it recognises the multilevel nature of the KCSE results.

The use of a linear mixed model helped to generate parameters and generalize effects of school types and gender on mathematics scores. This adds value to the action plans of interventions put in place to improve Mathematics education in schools and other educational outcomes. The in-depth analysis will enable studies to be designed to measure impacts of strategies implemented to improve teaching and learning of mathematics in the different types of secondary schools.

## 1.6: Outline of the thesis

This thesis is structured into five chapters.

In this first chapter we give background information on the topic of study.

In chapter two, we review previous research related to the current problem investigated and the literature related to statistical techniques used in the study.

Chapter three provides the methodology used to achieve the aims and objectives of this study. We give a detailed description of data collection, data organisation and data analysis procedures that were used in this study.

In chapter four, we discuss the results of the study. This includes the exploration of the KCSE exam data, testing for statistical significant difference in maths scores across the different school types and gender. This chapter also contains explanations of analysis resulting from the data.

The final chapter (chapter 5) provides conclusions, recommendations and suggestions for further work.

# Chapter 2: Literature review

This chapter discusses work done by previous researchers related to the problem investigated in this thesis; and the literature related to statistical techniques used in the study. It shows the connection between factors that affect achievement in Kenyan schools whilst identifying knowledge gaps.

A lot of work has been written by a wide range of professionals about the Kenya Certificate of Secondary Education exam (Lucas, 1993). Performance in KCSE by a student is influenced by various factors such as school infrastructure, the standard of teaching, and reading and learning resources (Lewis, et al., 2012) (Maundu, 1986). There are a number of projects like SMASSE (Japan International Cooperation Agency, 2007) and Africa Mathematics Initiative (African Maths Initiative, 2013) that are geared towards improving teaching and learning Mathematics (KNEC, 2009).

Teachers often lack the statistical background, understanding and tools to help them measure the impact of these interventions on exam results (North, et al., 2010). More precisely, studies about teachers' comprehension of measures of centre like mean, mode and median (Cai, et al., 2002) (Groth, et al., 2006) revealed a lack of understanding of the algorithm for calculating the average, a difficulty in discriminating the mean with the other measures of centre and little or no understanding of the effect of outliers on the mean (Jacobbe, et al., 2011). Research involving teachers' comprehension of graphs (Monteiro, et al., 2003) revealed that many teachers have difficulties in the interpretation of statistical graphs; they often weren't able to make generalizations about the data (Gonzalez, 2011).

Data from school examinations and intervention studies are often characterised by a hierarchical structure (Ukoumune, et al., 2004). A two level hierarchy is established when measurements are repeated on the same study subjects. Measurement occasions such as different schools are referred to as level-1 units and subjects like individual students are referred to as level-2 units (Goldstein, et al., 2001). Standard regression modelling usually assumes that the errors have zero mean and are mutually independent. However, in clustered data (Donner, et al., 1994) it is expected that errors for the same school are correlated (Rabe-Hesketh, et al., 2005).

The distinguishing aspect of grouped data is that observations of students' aggregate points and Mathematics scores within a school may be correlated, and the degree of similarity of aggregate points within a school can be measured (Donner, et al., 2003) and more reliable conclusions can be drawn from the analysis (Aitkin, et al., 1981). Thus, early work on the analysis of hierarchical data in the context of education was carried out by Aitkin and Longford (Aitkin, et al., 1986). The statistical and computing techniques for hierarchical linear modelling (Douglas, 1999) (Nezlek, et al., 1998) are based on fitting the regression analyses done at various levels for each unit in a single model (Goldstein, et al., 2001). Investigators have also conducted numerous studies aimed at identifying effective schools, determining which practices are related to their effectiveness, and measuring school contributions to student performance (Good, et al., 1986) and (Heyns, 1986).

Important questions left out during KCSE exam analyses are whether there is an association between mathematics scores and other subjects, whether there are statistically significant differences in mathematics score between single gender schools and mixed schools; and whether there are statistically significant differences in mathematics scores between boys and girls. If the questions are adequately answered then the information yielded will describe which (if any) school type and / or gender require further attention. Furthermore, this information could contribute to teachers' and educational researchers' working knowledge as they design interventions to help improve teaching and learning of mathematics in various schools.

# Chapter 3: Methodology

## 3.1: Introduction

This chapter discusses the data collection, data organisation and data analysis procedures that were used in this study. First, we have a discussion of how KCSE results for 2011 candidates were collected, and how the data was prepared and organised for analysis. This is followed by a description of the data analysis techniques; exploratory analysis, analysis of variance (ANOVA) and linear mixed model techniques that were used in the study.

## 3.2: Data Collection

At the end of Kenya Certificate of Secondary Education course, a national exam is taken by students at the end of their four year learning in secondary schools in Kenya. In 2011 this exam was set, administered, scored by the Kenya National Examination Council (KNEC) and was used for selecting candidates for universities and middle level colleges. The KCSE results drawn from KNEC were supplied to this study in text files. The permission to use the results for educational purposes was granted by the Kenya National Examination Council and it is gratefully acknowledged.

This comprised the results for 50,584 candidates in 699 schools which is 12% of the total 413,492 candidates who sat for the exam that year in the whole country. The information collected included variables such as school code, school random number, school name, student name, student gender, aggregate point scores and letter grades for each subject done by a student. The subject categories were as shown in Table 1 below

| | |
|---|---|
| Group one | Mathematics, English, Kiswahili |
| Group two | Biology, Physics and Chemistry |
| Group three | History& government, Geography, Religious Education |
| Group four | Home science, Art and Design, Agriculture, Computer Studies, Aviation |
| Group five | French, German, Arabic, Music, Business Studies |

*Table 1: Table showing how the various subjects taught in secondary schools are grouped by the Kenya National Examination Council*

For grading, a candidate must have had a minimum of seven subjects. The candidate must have taken all the three subjects in group one which are compulsory subjects, at least two group two subjects (sciences), at least one subject of the group three (humanities) and at least one of group four (technical) or group five (practical) subjects. Success in this KCSE exam is very important because it determines whether a given student will get direct entry to university or not. The target population for this study was students in Western Kenya who did their KCSE exam in 2011.

Each data vector represented information about a student; thus each student in the dataset is represented by one row of data in which one can find the region, school code, school random number, school name, student gender, student mean grade, student aggregate points and grades for each subject done by the student. Due to requirements imposed for confidentiality of students, a random identification number was used to identify a student.

The outcome measure used in this study is the students' Mathematics scores in the KCSE exam and their aggregate points score. Below is a section of the text files of students results - please note again that the candidates' names have been omitted to make the students anonymous.

```
CNO S MG AGP CANDIDATE'S NAMES          101 102 121 231 232 233 311 312 313 443 565
046 M C-  36                            C-  C   D   C-  D-  D-  B-  -   -   C+  -
047 F W   00                            P   P   P   P   P   P   -   P   -   -   P
048 F D+  29                            D   C-  E   D   -   D-  -   -   B   -   C
049 M D+  27                            C-  C-  D-  D   D   D   -   C-  -   -   D+
050 F C-  38                            B-  C   D   D+  -   D   C-  -   B   -   C-
051 F C-  34                            C+  C   D-  D   -   D-  -   -   B-  C   -
052 M D   24                            D+  D+  D-  D   D   D-  -   D   -   -   C-
053 F C   40                            B-  C+  D   C   -   D   C-  -   C   -   C+
054 M D+  31                            C-  C   E   C-  -   D   -   D   C-  -   C
055 M C   44                            B-  C-  C   C   -   D   B-  -   B-  B-  -
056 M D   20                            D   D   E   D-  -   E   C-  -   C-  D   -
057 M C   42                            C+  C-  D   C   D-  D+  B+  -   -   -   C+
058 M C-  33                            C-  D+  D   C+  D   D   -   C-  -   -   C
059 M C-  38                            C   C+  D   C   D+  D   -   C-  -   -   C+
060 F C-  38                            C+  C-  E   C-  -   D   C   -   B+  -   C+
061 M C-  34                            C+  D+  D-  D   -   D   C+  -   B-  -   C+
062 M D+  25                            D+  D   E   D   -   D-  D+  -   B-  -   D
063 F C   45                            C   C-  C+  B-  C-  C-  -   C+  -   -   C+
064 F D   21                            D+  D+  E   D-  -   D-  D   -   C-  D-  -
065 M D   20                            D   D-  E   D-  -   D-  D+  -   C   -   -
066 M D   20                            D   D+  E   D-  -   D-  D   -   C-  -   D
067 M D+  30                            C   D   E   D+  -   D-  -   -   B-  C   -
068 M D-  15                            D   D   E   E   E   E   D   -   -   -   D
069 M X   00                            X   X   X   X   -   X   X   -   X   -   -
070 M B-  57                            C+  B   C   C+  -   B   B   -   B+  -   C+
071 M C+  49                            B+  B-  D+  C+  -   D+  -   C+  B   -   C+
072 M C-  38                            B-  C+  E   D   -   D-  B+  -   C+  -   C
073 M D   20                            C-  D   D-  D-  D-  D-  -   D   -   -   D
074 F D   22                            C-  D+  D-  D   D-  D-  D   -   -   -   D
075 M D-  16                            D-  D   E   D-  -   D-  -   D   -   D   -
076 M D   22                            C-  D   D   D   D   D-  -   D   -   -   -
077 M B   63                            B   B-  A   B+  C+  C+  B   -   -   -   B-
078 F C-  34                            C   C+  E   D   -   D-  D+  -   B   C   -
079 M D+  30                            C-  D   E   D-  -   E   B   -   B   -   -
080 M C-  33                            C-  C-  E   C-  -   D-  C+  -   B-  -   -
081 M D   22                            D+  D   D   D   D-  D-  -   D   -   D+  -
082 F D   18                            D-  D   E   D   E   D-  -   D   -   D+  -
                  *** MEAN GRADE SUMMARY ***
   ENTRY   A   A-   B+   B   B-   C+   C   C-   D+   D   D-   E   X   Y   Z   W
    82     0   0    0    3   2    3    13  22   15   17  4    0   2   0   0   1
   X - ABSENTEES     Y - IRREGULARITIES  Z - MISSING GROUP(S)  P/W- PENDED/WITHHELD
                  *** AGGREGATE POINTS SUMMARY ***
AGP   84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59
TOTL   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  1  1  0  0  0  0
AGP   56 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33
TOTL   0  2  0  0  0  0  0  0  0  1  0  1  1  2  3  3  2  1  2  0  4  3  1  1  4  7
AGP   32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 09 08 07
TOTL   2  1  3  1  2  4  1  3  4  0  5  1  5  0  2  1  2  1  0  0  0  0  0  0  0  0
```

*Figure 1: A snapshot of the data collected from Kenya National Examination Council*

## 3.3: Data Organisation

Since this data set was acquired as a text file it was necessary to organise it into a suitable form for analysis. The first step involved a number of software packages being used in manipulating and restructuring the data into an appropriate form for analysis. These software packages included Python, Excel, GenStat 14th edition and R. The statistical software packages GenStat and R were chosen based on availability and flexibility.

### 3.3.1: Data processing

The Python programming language was used to clean the summary section in order to restructure the data into a usable format. The data acquired from KNEC was in plain text format i.e. a notepad file which is a basic text-editing program used to check over text files. The file contained information about schools such as the school code, number of students who did the KCSE exam, their grades and a summary of how the students performed in every subject as shown in Figure 2 below.

```
056 M D   20          D   D   E   D-  -   E   C-  -   C-  D   -
057 M C   42          C+  C-  D   C   D-  D+  B+  -   -   -   C+
058 M C-  33          C-  D+  D   C+  D   D   -   C-  -   -   C
059 M C-  38          C   C+  D   C   D+  D   -   C-  -   -   C+
060 F C-  38          C+  C-  E   C-  -   D   C   -   B+  -   C+
061 M C-  34          C+  D+  D-  D   -   D   C+  -   B-  -   C+
062 M D+  25          D+  D   E   D   -   D-  D+  -   B-  -   D
063 F C   45          C   C-  C+  B-  C-  C-  -   C+  -   -   C+
064 F D   21          D+  D+  E   D-  -   D-  D   -   C-  D-  -
065 M D   20          D   D-  E   D-  -   D-  D+  -   C   -   -
066 M D   20          D   D+  E   D-  -   D-  D   -   C-  -   D
067 M D+  30          C   D   E   D+  -   D-  -   -   B-  C   -
068 M D-  15          D   D   E   E   E   E   D   -   -   -   D
069 M X   00          X   X   X   X   -   X   X   -   X   -   -
070 M B-  57          C+  B   C   C+  -   B   B   -   B+  -   C+
071 M C+  49          B+  B-  D+  C+  -   D+  -   C+  B   -   C+
072 M C-  38          B-  C+  E   D   -   D-  B+  -   C+  -   C
073 M D   20          C-  D   D-  D-  D-  D-  -   D   -   -   D
074 F D   22          C-  D+  D-  D   D-  D-  D   -   -   -   D
075 M D-  16          D-  D   E   D-  -   D-  -   D   -   D   -
076 M D   22          C-  D   D   D   D   D-  -   D   -   -   -
077 M B   63          B   B-  A   B+  C+  C+  B   -   -   -   B-
078 F C-  34          C   C+  E   D   -   D-  D+  -   B   C   -
079 M D+  30          C-  D   E   D-  -   E   B   -   B   -   -
080 M C-  33          C-  C-  E   C-  -   D-  C+  -   B-  -   -
081 M D   22          D+  D   D   D   D-  D-  -   D   -   D+  -
082 F D   18          D-  D   E   D   E   D-  -   D   -   D+  -
              *** MEAN GRADE SUMMARY ***
   ENTRY   A   A-   B+   B   B-   C+   C   C-   D+   D   D-   E   X   Y   Z   W
    82     0   0    0    3   2    3    13  22   15   17  4    0   2   0   0   1
    X - ABSENTEES      Y - IRREGULARITIES  Z - MISSING GROUP(S)  P/W- PENDED/WITHHELD
              *** AGGREGATE POINTS SUMMARY ***
AGP   84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59
TOTL   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  1  1  0  0  0  0
AGP   56 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33
TOTL   0  2  0  0  0  0  0  0  0  1  0  1  1  2  3  3  2  1  2  0  4  3  1  1  4  7
AGP   32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 09 08 07
TOTL   2  1  3  1  2  4  1  3  4  0  5  1  5  0  2  1  2  1  0  0  0  0  0  0  0  0
```

*Figure 2: Data showing summary of school's grades at the bottom of the text file collected from Kenya National Examination Council*

Figure 2 above shows the summary of grades section for a school. The mean grade summary section was edited to remove it in the data file. The data processing capabilities of Python were required when using this data set where the text files were converted into comma separated values (CSV) format. The Python script removed unnecessary white space and cleaned up the data automatically for all 50,584 records. It also prevented any human errors that could have occurred if this work was done manually. The CSV format was then easily accessible with spreadsheets and other statistical packages (we used GenStat and R).

## 3.3.2: Data manipulation

### 3.3.2.1: Using Excel

The CSV (comma separated values) file format created using python was imported into Excel for further manipulation. Creating new variables such as the grade score from the raw data was necessary prior to statistical analysis and was done in Excel.

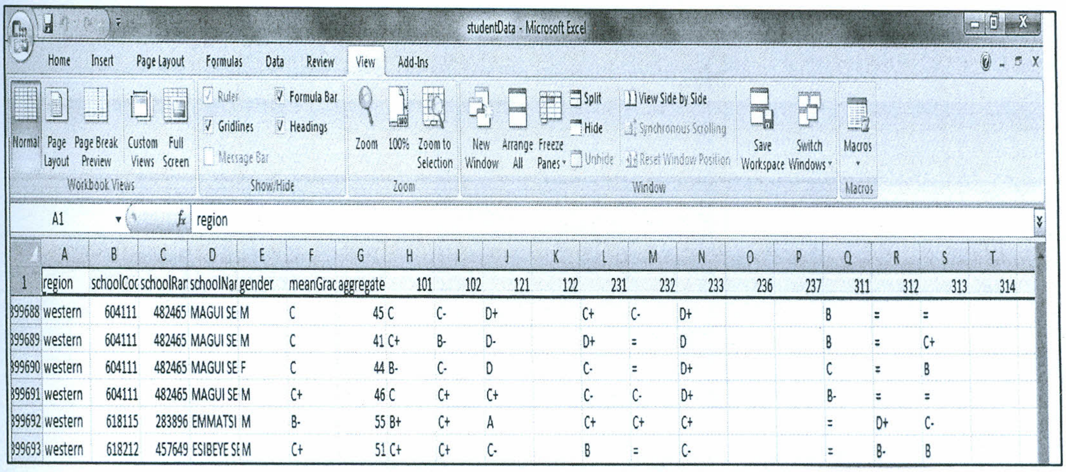



*Figure 3: CSV file created using Python*

Using the VLOOKUP function in Excel, values were assigned to the grades by looking for grades in a column in one worksheet and then returning a value in another column in the same sheet. Figure 4 below shows a VLOOKUP dialog box.

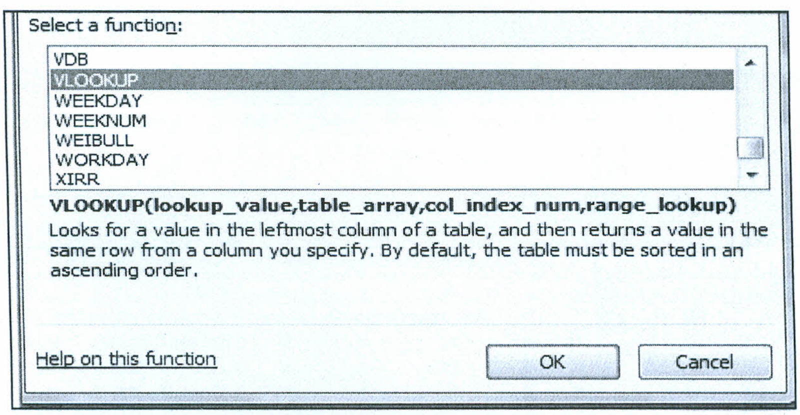



*Figure 4: Vlookup dialogue box*

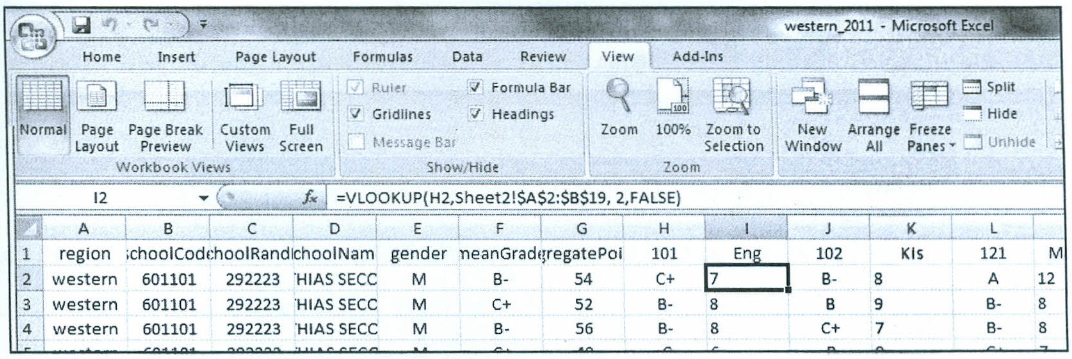
The letter grades for each subject were converted to an equivalent twelve point scale with 12 corresponding to grade A (the highest grade) and 1 corresponding to grade E, the lowest grade. : The scale assigned A = 12, A- = 11, B+ =10, B = 9, B- = 8, C+ = 7, C = 6, C- =5, D+ =4, D = 3, D- =2, and E=1. X – Absentees Y - irregularities Z - missing group(s) P/W- pended/withheld were also assigned. Table 2 below shows the worksheet from which numerical values were assigned to each grade.

| | A | B |
|---|---|---|
| 1 | Grade | Points |
| 2 | A | 12 |
| 3 | A- | 11 |
| 4 | B+ | 10 |
| 5 | B | 9 |
| 6 | B- | 8 |
| 7 | C+ | 7 |
| 8 | C | 6 |
| 9 | C- | 5 |
| 10 | D+ | 4 |
| 11 | D | 3 |
| 12 | D- | 2 |
| 13 | E | 1 |
| 14 | P | 99 |
| 15 | = | 99 |
| 16 | W | 99 |
| 17 | X | 99 |
| 18 | Y | 99 |
| 19 | Z | 99 |

*Table 2: Assigning of numerical values to grades done in excel*



The formulaic function VLOOKUP (lookup_value, table_array, col_index_num, range_lookup) was used. For example, this formula searched for the text format grades in the first column and returned the matching values from the second column in table array.




| | A | B | C | D | E | F | G | H | I | J | K | L | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | region | ;choolCod | hoolRand | choolNam | gender | neanGrad | gregatePoi | 101 | Eng | 102 | Kls | 121 | M |
| 2 | western | 601101 | 292223 | HIAS SECC | M | B- | 54 | C+ | 7 | B- | 8 | A | 12 |
| 3 | western | 601101 | 292223 | HIAS SECC | M | C+ | 52 | B- | 8 | B | 9 | B- | 8 |
| 4 | western | 601101 | 292223 | HIAS SECC | M | B- | 56 | B- | 8 | C+ | 7 | B- | 8 |

*Figure 5; Organized Excel file*

Figure 5 shows an organised excel file. Each data vector represented information of a student; thus, for each student in the dataset (each row of data), one can find the region, school code, school random number, school name, gender of the student, aggregate points, mean grade, grade and assigned value for each subject done by the student.

#### 3.3.2.2: GenStat

Calculations to restructure the data into appropriate form for analysis as well as the actual analysis were done using GenStat 14th edition and R. The Excel file was imported into GenStat which was able to handle the rectangle of data where each row referred to a student record and each column referred to a measurement for that student. It contained 50,584 records and was the main data file used. This package is very good at manipulating e.g. sorting, selecting and counting the many rows of data. Figure 6 shows data imported into GenStat and ready for manipulation.

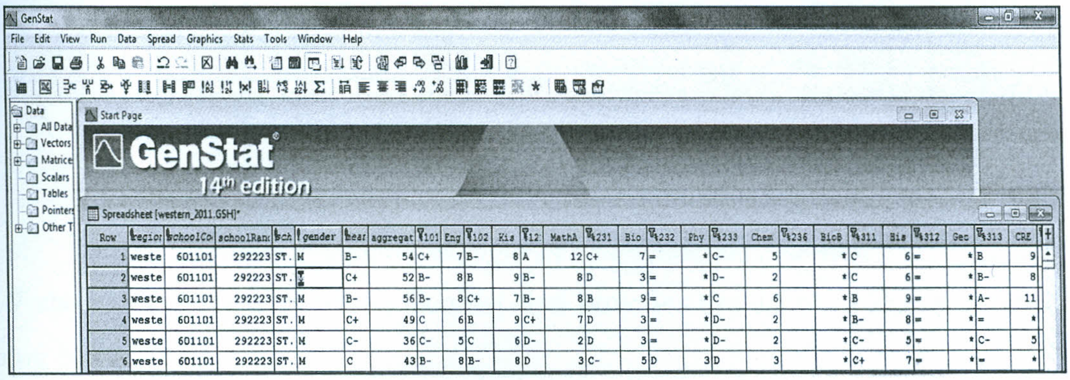



*Figure 6: Data imported into GenStat:*

GenStat had data manipulation features that allowed the formation of new variables from the existing variables and factors. Most of the calculations needed were done within the dialog box facilities of GenStat. The information in the table was essential for analysis, but was incomplete. Additional information such as school type was needed. For example, to have the aggregate points for 6 graded subjects minus Mathematics we use the spread menu in GenStat as follows.

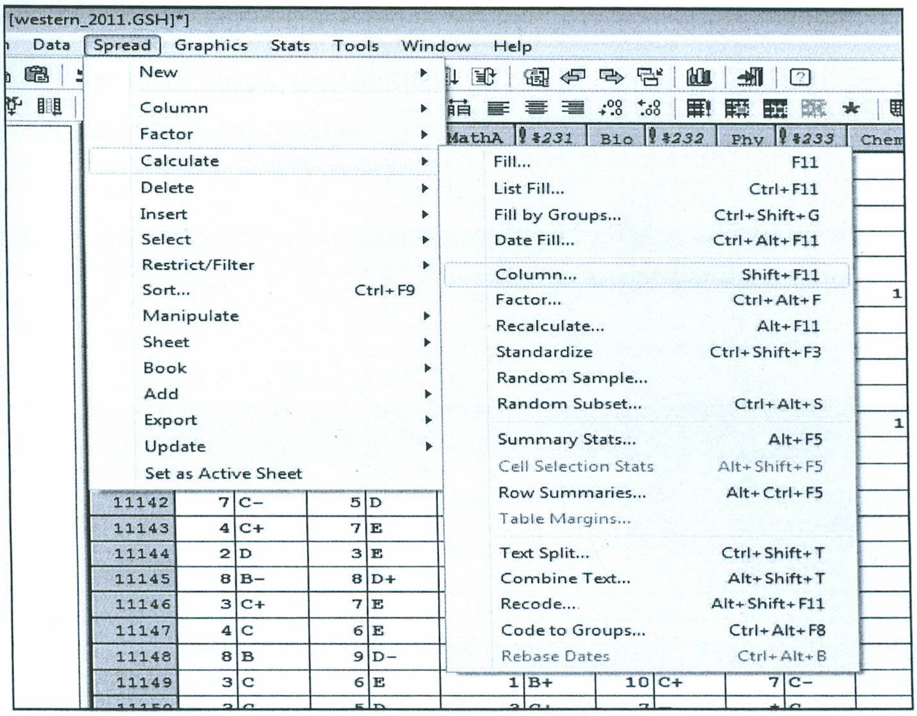


*Figure 7: GenStat dialogue box manipulation facilities:*

Figure 7 shows the Spread menu in GenStat that provides the different manipulation facilities and dialog boxes.

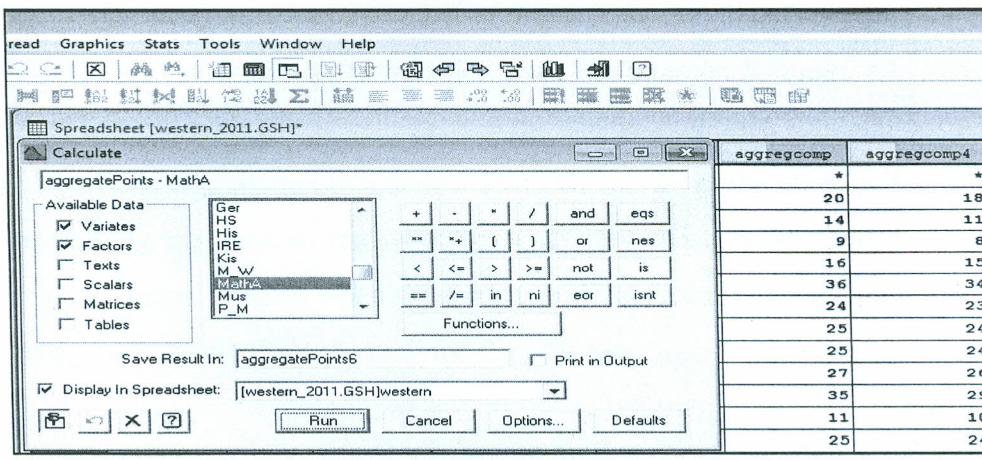
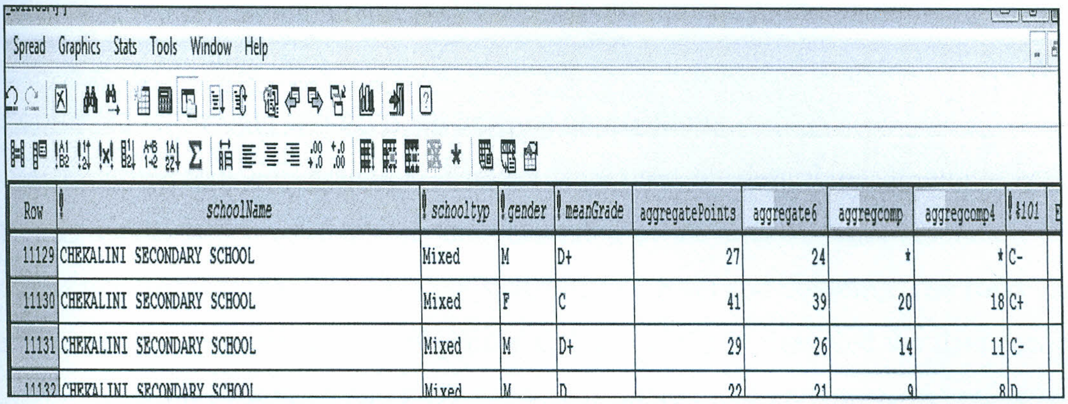


*Figure 8: An example of data manipulation screen depicting a form of calculation*

Figure 8 shows an example of a data manipulation procedure in GenStat that was used to create new variables like aggregate6, aggrecomp, aggrecomp4 and many others. Aggregate6 involved taking the total aggregate score a student gets then subtracting his or her Mathematics score, aggrecomp involved summing the scores for Maths, English, Kiswahili and any two sciences taken by a student while aggrecomp4 involved subtracting a student Maths score from the aggrecomp. In all the three cases, the original scores are combined to form the new variables. Figure 9 below shows a few of the additional variables created.




| Row | schoolName | schooltyp | gender | meanGrade | aggregatePoints | aggregate6 | aggregcomp | aggregcomp4 | \$101 |
|---|---|---|---|---|---|---|---|---|---|
| 11129 | CHEKALINI SECONDARY SCHOOL | Mixed | M | D+ | 27 | 24 | * | * | C- |
| 11130 | CHEKALINI SECONDARY SCHOOL | Mixed | F | C | 41 | 39 | 20 | 18 | C+ |
| 11131 | CHEKALINI SECONDARY SCHOOL | Mixed | M | D+ | 29 | 26 | 14 | 11 | C- |
| 11132 | CHEKALINI SECONDARY SCHOOL | Mixed | M | D | 22 | 21 | 9 | 8 | D |

*Figure 9: New variables and additional information formed using existing variables:*

### 3.3.2.3: Using R

R is a powerful statistical package where all the code that supported a wide a range of computations, statistical procedures and most of the exploratory plots was accessible. It is a free open source package that is easily downloaded from the internet. RStudio, a system and a text editor designed to work with R was very helpful for this large dataset. Calculations were done in the console. The code was written in a script (.r) and then sent to the console. R made a file that contained everything that was in the work space. It was then possible to save the workspace when quitting and import it again when we started working again. From Excel, where the data set was saved as a 'studentData.CSV' format, the data file was imported into R using the command

```
dat<-read.table("studentData.csv",header=T,sep=";",na.strings=c("","="))
```

and further manipulation to the data was carried out on the data set.

## 3.4: Data Analysis

### 3.4.1: Exploratory analysis

Exploratory analysis included generating summaries such as frequencies, percentages, means, standard deviations and representations for the scores in the exam. Exploratory data analysis gave insights into the variability of maths scores; provided explanation for surprising patterns observed in the data and revealed unusual observations.

Numerical methods and graphical displays were used for describing the important aspects of the KCSE results which gave a basic understanding of the data. A cross tabulation classified the data into two dimensions. The table consisted of two rows and columns. The rows classified the data according to one dimension of gender and the columns classified the data according to the second dimension of school types. The row percentages and column percentages helped to quantify relationships.

Graphical methods were used to study relationships between variables. For example, scatter plots were used to study the relationship between maths scores and aggregate scores in different school types. Vertical bar charts and histograms were used to depict the frequencies and distribution of mathematics grades and aggregate grades for the region. A graph of cumulative % aggregate points was used to show the relative cumulative frequency distribution of the aggregate points and relative cumulative percentages. Box and whiskers displays sometimes called box-plots were used to describe candidates' aggregate scores by using a five number summary within a school type. The five number summary consisted of;

- the smallest aggregate score,
- the first quartile,
- the median for the school type,
- the third quartile,
- the largest aggregate score and indications of outliers.

In describing the distribution of the population's grades, we described the dataset's measures of central tendency. Population means for scores were calculated by adding the population scores and then dividing the resulting sum by the number of students.

$$\mu = \frac{\sum_{i=1}^{n} X_i}{n}$$

The population variance and standard deviations were also calculated. Since this was a complete enumeration of the candidates from schools and school types, the complete population of candidates was used to calculate the standard deviation, that is:

$$\sigma_j = \sqrt{\sum_{i=1}^{n_j} \frac{(x_{ij} - \mu_j)^2}{n_j}}$$

$\sigma_j$ is the population standard deviation aggregate score for school type j,

$x_{ij}$ is the aggregate score for a candidate *is* in a school type j,

$\mu_j$ is the mean aggregate score for school type j, and

$n$ is the total number of candidates who were graded after the school type j

The degree of asymmetry, or departure from symmetry of the score distribution in mathematics and other subjects was also checked. For skewed distributions the subject mean grades tended to lie on the same side as the mode grade. The formula for skewness in the grades was given by:

$$Skewness = \frac{\sum_{i=1}^{N} (X_{ij} - \mu_j)^3}{(n-1)S_j^3}$$

Where $X$ is a subject score of student *j* in subject *i*

$\mu$ is the mean of subject *j*.

$S$ is the standard deviation of subject *j*

$n$ is the number of data points



Simple correlations measured the linear relationship between mathematics and other subjects like chemistry, physics and the languages; and if the value of the coefficient of determination r was near zero, it meant that there was no linear correlation. The degree of relationship of Mathematics and other subjects was checked using Karl Pearson correlation coefficient $r_{xy}$ given by the ratio of the covariance between mathematics (X) and other subjects ($Y_i$), to the product of the standard deviations of X and $Y_i$.

Symbolically

$$r_{xy} = \frac{Cov(X,Y)}{\sigma_x \quad . \sigma_y}$$

Where, $(X_1, Y_{j1})$; $(X_2, Y_{j2})$; .......... $(X_n, Y_{jn})$ are n pairs of observations of Mathematics score and a given subject j score.

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \mu_x)(Y_i - \mu_y)}{n}$$

$$\sigma_x = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \mu_x)^2}{n}}$$

$$\sigma_y = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \mu_y)^2}{n}}$$

The exploratory data analysis was undertaken because not so much about the data was known; therefore preliminary work was done to understand the nature of the KCSE examination results. The main drawback was that the data could only be scrutinised in "parts". For example, it was difficult to look at several components of the exam results pattern at the same time. We therefore needed a method of apportioning variability of math scores across the school types and gender all at once.

### 3.4.2: Two way Analysis of Variance (ANOVA)

In this section we examined the effect of two factors (school type and gender) on a response variable (maths score). Factor 1 had two levels and factor 2 had two levels. For two factors that might affect a response variable, an interaction may exist if the relationship between the mean response and one factor depends on the other factor. For this reason, an interaction between the two factors was also examined.

Analysis of variance (ANOVA) was carried out to find if there was a significant difference in Mathematics achievement among students from the different school types as well as gender. Mathematics scores were classified on the basis of candidates being in the different school types and on the basis of gender of the candidates. Each factor had two categories; the school type factor has two categories that are single gender schools and mixed schools, gender had two categories that are male or female. ANOVA model was used to determine how much of the variation in the response can be attributed to different factors.

First, the effects of school type and gender are estimated (Stern, et al., 2004), and then an interaction term which is the product of school type and gender is included in the model. Using the analysis of variance technique provided an opportunity to look at the whole pattern e.g. the two factors together and the part of the data that cannot be explained by the model.

The ANOVA table (Table 3) below gives summary elements called sum of square, degrees of freedom and mean squares for both the components of the pattern and the residual. The sums of square allowed us to see what proportion of the variation is explained by the different parts of the pattern in the data and the residual sum of squares shows us what remains unexplained.

| ANOVA summary table | | | | | Sum of squares | What the letters stand for |
|---|---|---|---|---|---|---|
| Source of variation | Sum of squares | Degrees of freedom | Mean square | Test statistic | $SS_T=\sum\sum\sum X_{ijk}^2 - \frac{T^2}{n}$<br>$SS_A=\sum_i \frac{T_i^2}{n_i} - \frac{T^2}{n}$ | $a$ = number of levels of schooltype2<br>$b$ = number of levels of gender<br>$r$ = number of data values for each combination of levels of schooltype and gender |
| Schooltype2 | $SS_A$ | a-1 | $MS_A$ | $F=\frac{MS_A}{MS_R}$ | | $n$= total sample size |
| Gender | $SS_B$ | b-1 | $MS_B$ | $F=\frac{MS_B}{MS_R}$ | $SS_B=\sum_j \frac{T_j^2}{n_j} - \frac{T^2}{n}$ | $X_{ijk}$ =Kth member of sample at level $i$ of schooltype2 & level $j$ of gender<br>T= $=\sum\sum\sum X_{ijk}$ overall total |
| A×B Interaction | $SS_{AB}$ | (a-1)(b-1) | $MS_{AB}$ | $F=\frac{MS_{AB}}{MS_R}$ | $SS_G=\sum_j \frac{T_g^2}{n_g} - \frac{T^2}{n}$ | $T_i$ =total of data at level of $i$ schooltype2 |
| Residual | $SS_R$ | n-ab | $MS_R$ | | | |
| Total | $SS_T$ | n-1 | | | $SS_R= SS_T -(SS_A+ SS_B + SS_{AB})$ | $n_i$ =number of data items at level i of schooltype2<br>$T_j$ = total of data at level j of gender<br>$n_j$ =number of data items at level j of gender<br>$T_g$ = total of all data at a particular level of schooltype2 and of gender (There are *ab* such groups)<br>$n_g$ = total of data values at a particular level of schooltype2 and of gender |

*Table 3: ANOVA table containing a summary of components of pattern and the residual*

Where

$X_{ijk=}$the $K$th value of the response variable observed when using level $i$ of factor 1 and level $j$ of factor 2

The ANOVA procedure for two factors partitioned the total sum of squares $SS_T$ into four components: the factor 1 sum of squares $SS_A$, the factor 2 sum of squares $SS_B$, the interaction sum of squares $SS_{AB}$ and the error sum of squares $SS_R$. The formula for this portioning was as follows:

$$SS_T = SS_A + SS_B + SS_R$$

*a-1, b-1, (a-1)(b-1) and n-ab* gave the degrees of freedom for factor 1, factor 2, interaction and unexplained error. In addition, we obtained the mean square associated with the factors and finally calculated the F ratio which tested the effects of factors. The process was done at the same time with the testing of the effects of factors. The significance of both factors and interaction was tested by comparing the F values that were obtained through computation and the expected. The F-probabilities that are given in the final column of the ANOVA table above enabled us to test the hypotheses about the components of the pattern part of the data. The assumptions made in the ANOVA technique are that

i. The population of values of the response variable associated with the factors had equal variances.

ii. The population of values of the response variable associated with the factors all had normal distributions.

iii. The students associated with the factors were randomly selected.

In secondary schools, students existed within a hierarchical structure that included classroom, class level, school and district. Usually, regression analysis (Gelman, et al., 2008) is used in examining the relationships among the student's dependent variables such as KCSE aggregate points and one or more independent variables. These assume that the maths score of any student is not related to any other student. However, this assumption is violated if some of the students are from the same family, class, school, school type, and district. Thus, the KCSE exam results are structured hierarchically since students are nested within a class, school, school type which are then nested within districts.

Among the key assumptions in ANOVA is the independence of candidates' scores under study. This means (for example) that the mathematics score of any one student is not related to the observations of any other student. The assumption is however violated when results are from the same school. When assumption of independency among observations do not hold, the estimated coefficients can be biased and result in misleading analysis (Bryk, et al., 1992).

### 3.4.3: Linear Mixed Model

Using Residual Maximum Likelihood (REML) algorithm with fixed effects and a random effect (Laird, et al., 1982) (Lee, et al., 2006) the study was no longer restricted to independent data. This becomes an extremely flexible estimating tool that demonstrates how to analyse data that is successively correlated (Allan, et al., 2001).

A linear mixed model which included a random effect term-school was fitted, to take into consideration the dependency of mathematics scores within a school (Venables, et al., 1999). The variables "school type" and "gender" were the fixed effects. School type was fixed because it only had one of the two values; single gender or mixed school. Gender was also fixed because a student can only be either male or female. The influence of the school on the maths score was considered a random effect. The linear mixed model was given by

$$Y_{ij} = \beta_o + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \Upsilon_i Z_i + \varepsilon_{ij}$$

Where

$Y_{ij}$ is the value of the response variable which is the maths score for the $j_{th}$ student in the $i_{th}$ group.

$\beta_o$ is the overall maths score mean

$\beta_1$ *and* $\beta_2$ are the fixed-effects coefficients

$X_{1i}$ *and* $X_{2i}$ are the fixed effects variables for observation in group *i*

$\Upsilon_i$ is the random-effect coefficient for each school. This is thought of a random variable and not as a parameter. It is assumed to be independent and identically distributed N $(0,\delta^2)$

$Z_i$ is the random effect variable which is the school

$\varepsilon_{ij}$ is the error for the observation j in group i. The errors are assumed to be normally distributed and multivariate normal. It is assumed to be independent and identically distributed N $(0,\delta^2)$

The results of the analysis are presented in the following chapter.

# Chapter 4: Results from analysis and Discussion

## 4.1: Introduction

This chapter describes how data was analysed to answer the objectives of the study. Findings from the study are presented, interpreted and discussed. In the first section of data exploration below, data is presented in tables and numerical summaries such as means, standard deviations, five number summaries and correlations. Graphical summaries such as box-plots, histograms and scatter plots are presented in the subsequent sub-sections to explain the patterns observed in the results. The section gives ideas and examples of patterns while also giving us an insight into variability contained in the exam results.



## 4.2: Examining the results

Out of 50,584 candidates who had registered for the KCSE exam in the region, 22,178 (44%) were female candidates while 28,406 (56%) were male candidates. The western region had 106 (15%) and 55 (8%) single gender girl and boy schools respectively and 538 (77%) mixed schools that had students who sat for the KCSE examination. From a total of 50,584 candidates, 49,815 (98.5%) were graded. The remaining 769 students did not receive grades either because their results were withheld for not providing proper documentation, they were absent, had a missing group/subject or were involved in exam malpractice and had their results cancelled. The number of students who were graded are summarised in Table 4 by gender and the type of school.

| | Single gender schools | Mixed schools | Total |
|---|---|---|---|
| Female candidates | 9006 (53.5%) | 12908 (39.2%) | 21914 (44.0%) |
| Male candidates | 7843 (46.5%) | 20058 ( 60.8%) | 27901 (56.0%) |
| *Total* | 16849 (33.8%) | 32966 (66.2%) | 49815 (100.0%) |

*Table 4: Cross tabulation of counts of students by gender and type of school*

The row totals in Table 4 above provide the total number candidates by gender while the column totals provide number of candidates by school types. One good way to investigate relationships such as these was to compute row percentages and column percentages. The row percentages are computed by dividing each cell's value by its corresponding row total and expressing the resulting fraction as a percentage.

For instance, the row percentage for the lower left hand cell (candidates in single gender schools) in the table above is (16849/49815) × 100% = 33.8%. Similarly, column percentages are computed by dividing each cell's frequency by its corresponding column total and expressing the resulting fraction as a percentage. For example, the column percentage for the upper right hand cell in the table above is (21914/49815) ×100% = 44.0%. The table also gives a percentage distribution of candidates in different school types and helps to quantify relationships. The average number of candidates in mixed schools was 61 which was much lower than that in single girls and boys schools (85 and 143 respectively). The high number of candidates in mixed schools could be attributed to the greater number of mixed schools (538).

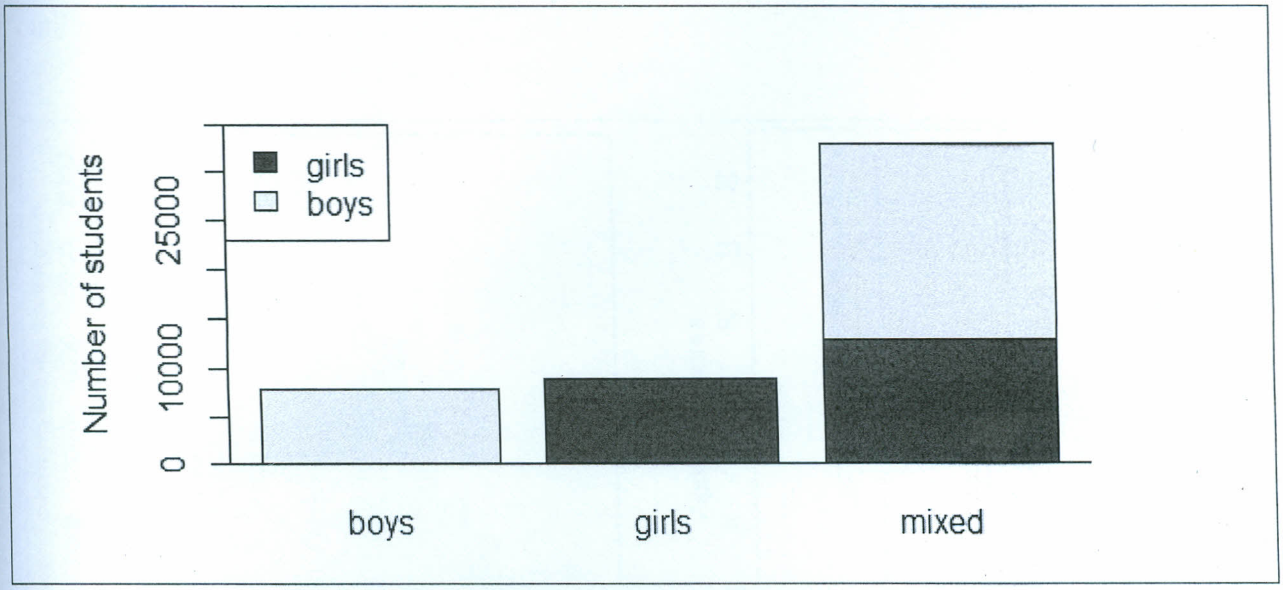



*Figure 10: A bar chart presentation of the information in Table 4*

### 4.2.1: A look at Aggregate scores

Aggregate student scores are important since they determine whether a student can be admitted to university or not. The calculation entails the summation of all the scores for the seven subjects taken by a student (The choice of subjects is explained in section 3.2. The maximum that a student can score in a subject is 12 points and aggregately is 84 points. The Joint Admissions Board (JAB) is the Kenya body responsible for setting the admission criteria for entry into public university in Kenya. For KCSE 2011 candidates, male and female candidates needed to have an aggregate of at least 63 and 62 points respectively for them to gain direct entry into a public university.

In Figure 11 below, a histogram is presented to show the general distribution of the aggregate scores and a box plot distinguishing the distributions between the three school types involved. Since most candidates were from mixed schools, the distribution of the aggregate scores for mixed schools has had a tremendous effect on the overall distribution. For instance, the highest frequencies in the histogram are between the aggregate scores of 20 and 40, which is where 50% of the candidates in mixed schools were. There were only 550 candidates in mixed schools who got an aggregate of 70 points and above. This was only 1.7% of the total candidates in mixed schools in the region. Therefore, it is not surprising to see them being pointed out as outliers in the box plot.

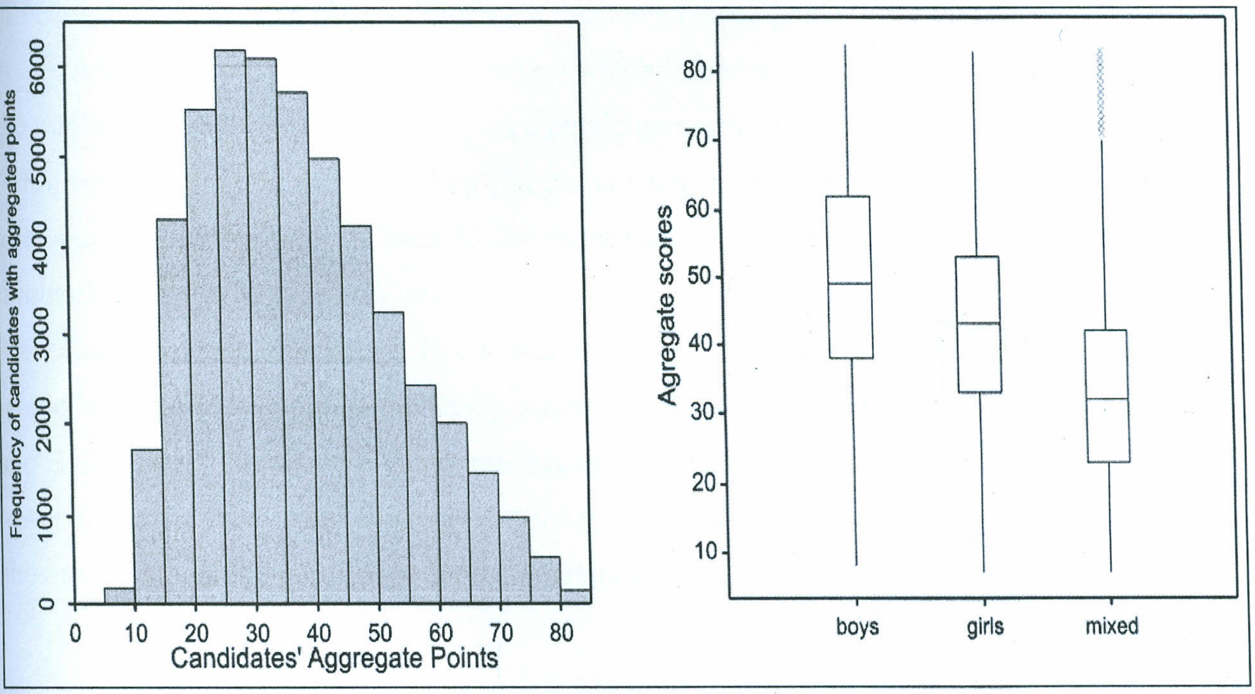



*Figure 11: The distribution of overall candidates' aggregate scores (histogram) and by school type (box plot). As can be seen in the box plot, around 25% of the candidates in boys schools would make it to the university through direct entry.*

| Grade | boys | girls | mixed |
|---|---|---|---|
| E | 4 | 3 | 171 |
| D- | 63 | 190 | 3020 |
| D | 292 | 619 | 6200 |
| D+ | 630 | 1133 | 6797 |
| C- | 978 | 1477 | 6002 |
| C | 1157 | 1705 | 4295 |
| C+ | 1329 | 1471 | 2880 |
| B- | 1088 | 1075 | 1631 |
| B | 975 | 716 | 1091 |
| B+ | 757 | 404 | 640 |
| A- | 464 | 190 | 220 |
| A | 106 | 23 | 19 |

*Table 5: Table showing the aggregate grade distribution in the various school types and information presented in Figure 11*

The best male candidate from boys' school got an aggregate of 12 out of 12 in all seven subjects while the best female candidate from girls' school scored 12 out of 12 in six of the seven subjects, just like the best male student from mixed schools. The best female candidate from mixed schools had an aggregate of 82. The majority of the candidates (63.8%) scored a C plain or below, and 19,134 candidates (38.4%) of them got a mean grade of D+ and below.

JAB has a policy on affirmative action which improves the chances of female candidates, the minimum requirement for one to join a public university is an aggregate of 42 points but due to limited facilities this was moved up (for direct entry) to a minimum of 62 aggregate points for girls and 63 for boys who did their KCSE exam that year. 90.7% of the students did not get the minimum grade to join a public university. Only 9% had a minimum grade of 62 points or above proceeded to public university. If C+ was the minimum requirement then 19,013 candidates would have made it to public university however only 4,397 made it from the region 1,327 girls and 3,070 boys. Despite the minimum aggregate points for joining public universities being lower for girls, there were fewer girls joining public universities than boys. Figure 12 below shows the cumulative percentage for the aggregate points scored by the candidates.

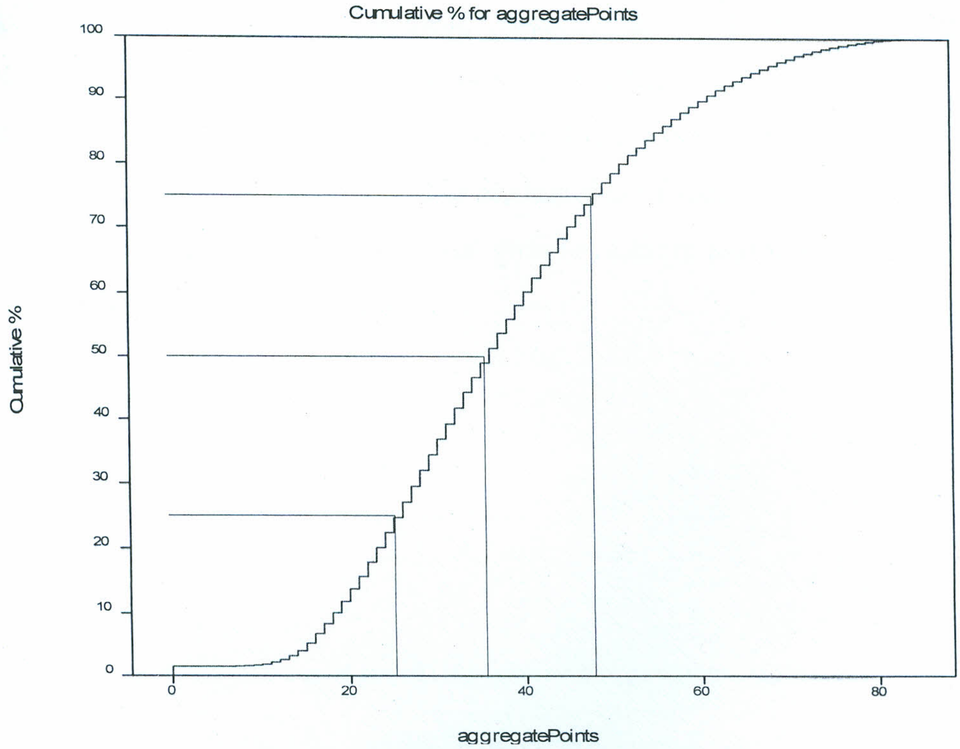


*Figure 12: A graph of cumulative percentage for the aggregate points in Western region*

From these, we can calculate the aggregate score of a school by taking the total sum of aggregate scores for all the students in a school and divide by the total number of students to get the average score of a school. This was done again for categories of the schools. The standard deviation was also calculated. Since this was a complete enumeration of the candidates from schools in Western Kenya, the complete population of candidates was used to calculate the standard deviation.

The mean aggregate point for purely girl's schools ($\mu_{Fgirls}$) was 43.62 points with a standard deviation ($\sigma_{Fgirls}$) of 14.30, for purely boy's schools ($\mu_{Mboys}$) it was 50.06 points with a standard deviation ($\sigma_{Mboys}$) of 15.44. There were approximately 1,200 more female candidates in girl schools than male candidates in boy schools. 95% of male candidates in girl and boy schools got between ($43.62 - 1.96\times14.30$, $43.62 + 1.96\times14.30$) and ($50.06 - 1.96\times15.44$, $50.06 + 1.96\times15.44$) respectively. This can be simplified to be (15.59, 71.65) and (19.80, 80.32) points respectively.

The girls in mixed schools had a mean ($\mu_{Fmixed}$) of 31.33 with a standard deviation ($\sigma_{Fmixed}$) of 12.12 while the boys in mixed had a mean ($\mu_{Mmixed}$) of 35.64 with a standard deviation ($\sigma_{Mmixed}$) of 14.74. ($\sigma_{Fmixed}$) is unexpectedly lower than ($\sigma_{Mmixed}$) despite the number of male candidates in mixed schools exceeding that of females by approximately 6,000. One plausible explanation for this could be that girls have more uniform scores or because the girls in mixed schools had a lower average.

### 4.2.2: Subject summaries

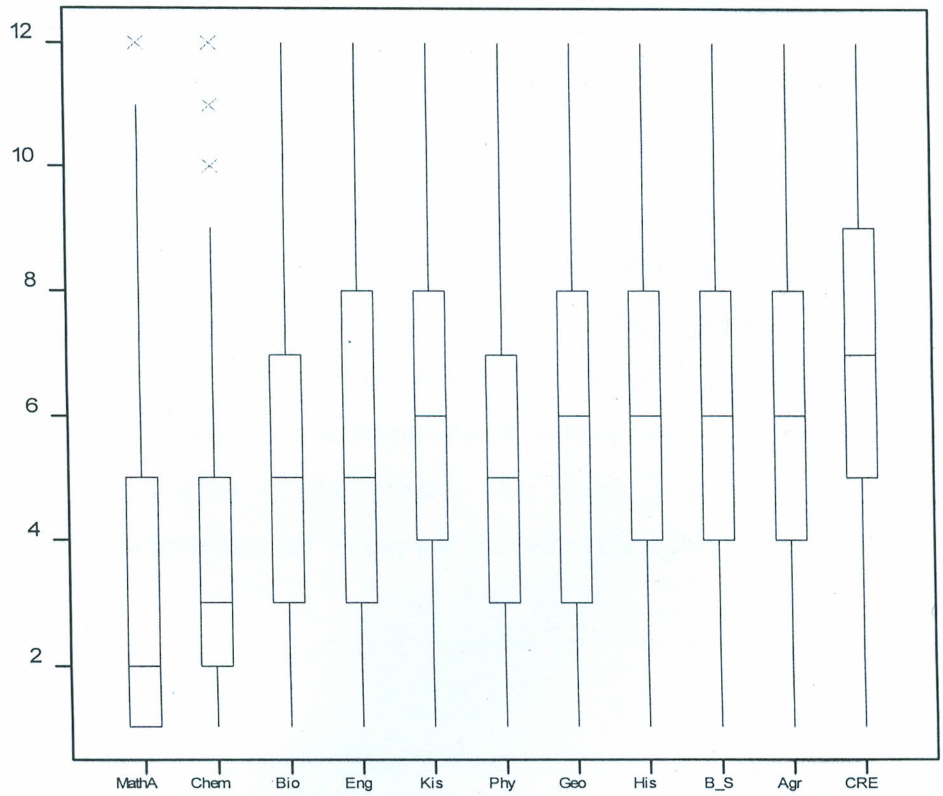



*Figure 13: Box plots showing the distribution of scores in various subjects*

Figure 13 above shows box plots that illustrate how students' performed in various subjects in the province. It reveals that all the other subjects appear to have a higher median than Mathematics which has an unusual distribution. It illustrates the differences in variation in the subject score and indicates outliers for Mathematics and Chemistry. The spread of the grades in the subjects is fairly symmetrical except for Mathematics as indicated by the relative positions of the medians within respective boxes. The lack of symmetry in the box plots indicates skewness in the grades in the various subjects.

| Subject | Mat | Chem | Bio | Eng | Kis | Phy | Geo | His | B/S | Agri | CRE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | 3.525 | 3.972 | 5.174 | 5.501 | 6.287 | 5.483 | 5.846 | 6.014 | 6.061 | 6.323 | 7.109 |
| Median | 2 | 3 | 5 | 5 | 6 | 5 | 6 | 6 | 6 | 6 | 7 |
| SD | 3.00 | 2.552 | 2.582 | 2.385 | 2.468 | 2.804 | 2.703 | 2.483 | 2.514 | 2.585 | 2.437 |
| Skewness | 1.379 | 1.338 | 0.640 | 0.279 | 0.210 | 0.581 | 0.325 | 0.229 | 0.235 | 0.111 | -0.233 |
| No. of candidates | 49815 | 48781 | 48617 | 49815 | 49815 | 12698 | 33719 | 32693 | 19251 | 21849 | 35182 |

*Table 6: A Summary of students' performance in Western region per subject. Only CRE had more candidates scoring above the 50%.*

The results in Table 6 above show the means, medians and standard deviations of KCSE examination scores for some subjects for the whole of the Western region. Studying the means further show that Christian Religious Education (CRE) had the highest mean while Mathematics had the lowest mean. Amongst the subjects, Mathematics scores had the highest variability while English had the lowest variability. The asymmetrical form for the Mathematics box plot indicates skewness in the distribution of the grades. The bigger positive value for skewness in Mathematics grades indicate that the grades are skewed to the right i.e. the tail to the right is long relative to the left tail.

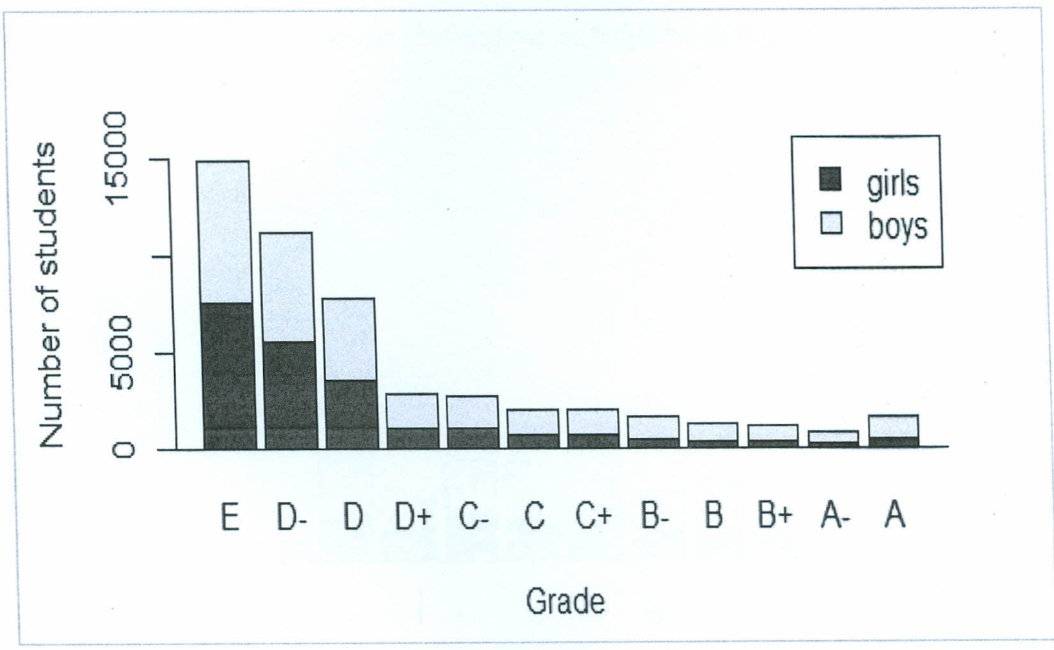



*Figure 14: Bar graph showing the distribution of MathA grades*

Figure 14 above shows the distribution of Mathematics mean grades per gender which is clearly skewed. From the bar plot, it can be seen that most students, irrespective of their gender, got a mean grade E in maths and the numbers reduce in comparison as the grades get better. But it is also interesting that the number of A's were more than the number of A- for both genders.

Like Mathematics, English and Kiswahili are the other subjects that are done by all candidates. Together, they comprise the compulsory subjects. For this reason, we also look at the performance in English and Kiswahili by the candidates.

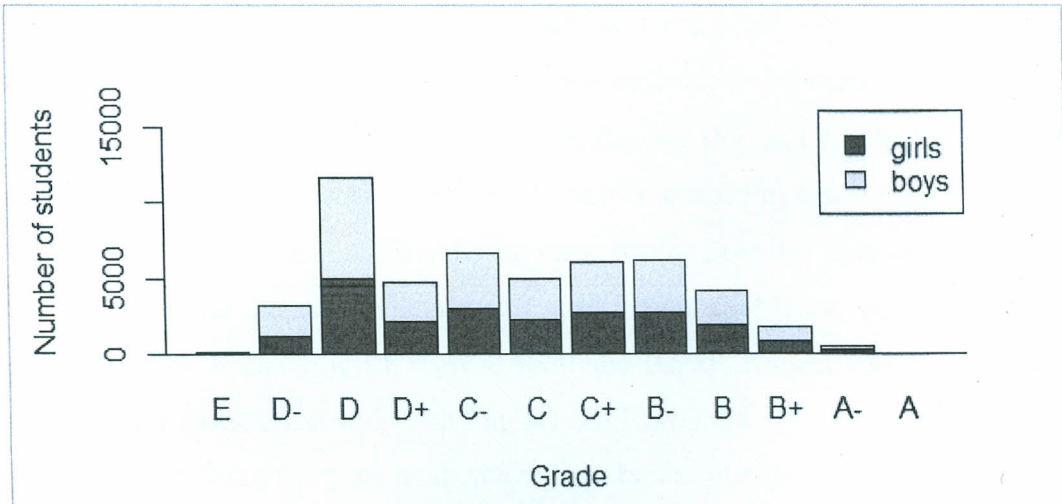



*Figure 15: Bar graph showing the distribution of English grades*

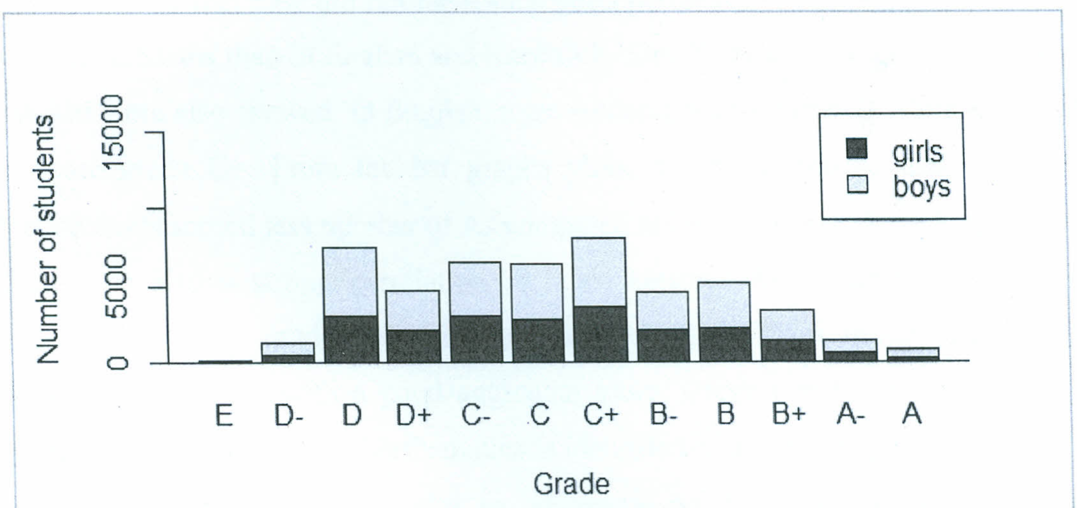



*Figure 16: Bar graph showing the distribution for Kiswahili grades*

Figure 15 and Figure 16 above summarise the performance of English and Kiswahili respectively. It is quite visually clear that the distribution of grades is not the same as that of Mathematics. Most of the students in English have a D while in Kiswahili have a C+. Unlike Mathematics where most of the students have an E, it is the complete opposite in English and Kiswahili where the grade E has the least number of students.

We also check on the association of Mathematics and other subjects using Karl Pearson correlation coefficient $r_{xy}$ given by the ratio of the covariance between mathematics (X) and other subjects ($Y_i$), to the product of the standard deviations of X and $Y_i$. The Pearson correlation coefficients between Maths A and Chemistry and Maths A and Physics were 0.8376 and 0.8189 respectively. Mathematics has sufficiently a very strong positive association with the two subjects Chemistry and Physics. The Pearson correlation coefficient between Maths A and Kiswahili and Maths A and English were 0.6060 and 0.6006 respectively. Mathematics has a moderately positive association with both English and Kiswahili.

It was not surprising for poor math students to be discouraged from doing Physics, hence the low number of students who sat for the Physics examination. Generally students who performed well in Maths also performed well in Chemistry. There was a deviation when it came to other "non-mathematical" subjects like Languages. There were some who performed well in them despite the fact that they did not have such good performances in Maths. Notice that there were more As in Maths than in English and Kiswahili. The distribution of grades in both English and Kiswahili were also skewed. In English, most students scored grade D while Kiswahili most students score grade C+. From the bar graphs plots, it can be noticed that in English and Kiswahili, students scored less number of As compared to the number of As in Mathematics.

There was also a strong correlation (0.7606) between Mathematics and the students' aggregate score of the six graded subjects minus Mathematics. This is an identification of factor "Mathematics", associated with a good aggregate score. Given that there are multiple factors influencing the aggregate scored, Mathematics is identified as an important factor. Furthermore, the distributions of Mathematics grades vs. aggregate points by gender and type of school suggest a pattern between aggregate points and Mathematics score.

| Mean Grade | E | D- | D | D+ | C- | C | C+ | B- | B | B+ | A- | A |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| E | 178 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D- | 3111 | 156 | 5 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D | 5073 | 1688 | 295 | 37 | 14 | 3 | 1 | 0 | 0 | 0 | 0 | 0 |
| D+ | 3907 | 3215 | 1114 | 191 | 74 | 43 | 13 | 2 | 1 | 0 | 0 | 0 |
| C- | 1956 | 3378 | 2096 | 490 | 285 | 138 | 72 | 34 | 3 | 4 | 0 | 1 |
| C | 532 | 2017 | 2386 | 760 | 642 | 366 | 241 | 112 | 68 | 23 | 9 | 1 |
| C+ | 87 | 667 | 1498 | 833 | 868 | 622 | 523 | 287 | 149 | 87 | 34 | 25 |
| B- | 8 | 76 | 423 | 393 | 544 | 529 | 617 | 476 | 321 | 219 | 112 | 76 |
| B | 0 | 9 | 51 | 75 | 189 | 261 | 446 | 469 | 440 | 378 | 218 | 246 |
| B+ | 0 | 0 | 0 | 0 | 15 | 28 | 103 | 179 | 260 | 343 | 300 | 573 |
| A- | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 15 | 34 | 91 | 132 | 599 |
| A | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 141 |

Mathematics grade

*Table 7: Counts of final grade vs Mathematics grade scored by candidates*

Table 7 is intended to show the relationship between the grade scored in Mathematics and the aggregate grade scored by a student. By visual inspection, it is apparent that poor grades in Mathematics are correlated with poor mean grades by the student. For instance, most of the students with a grade E in Maths got a mean grade of D. There are only very few, 8 students, who got a better grade of B-. Most of the students who did extremely well with a mean grade of A also had an A in Mathematics. This explains the strong correlation between the performance of a student in Maths and the other subjects and as a result the overall mean grade.

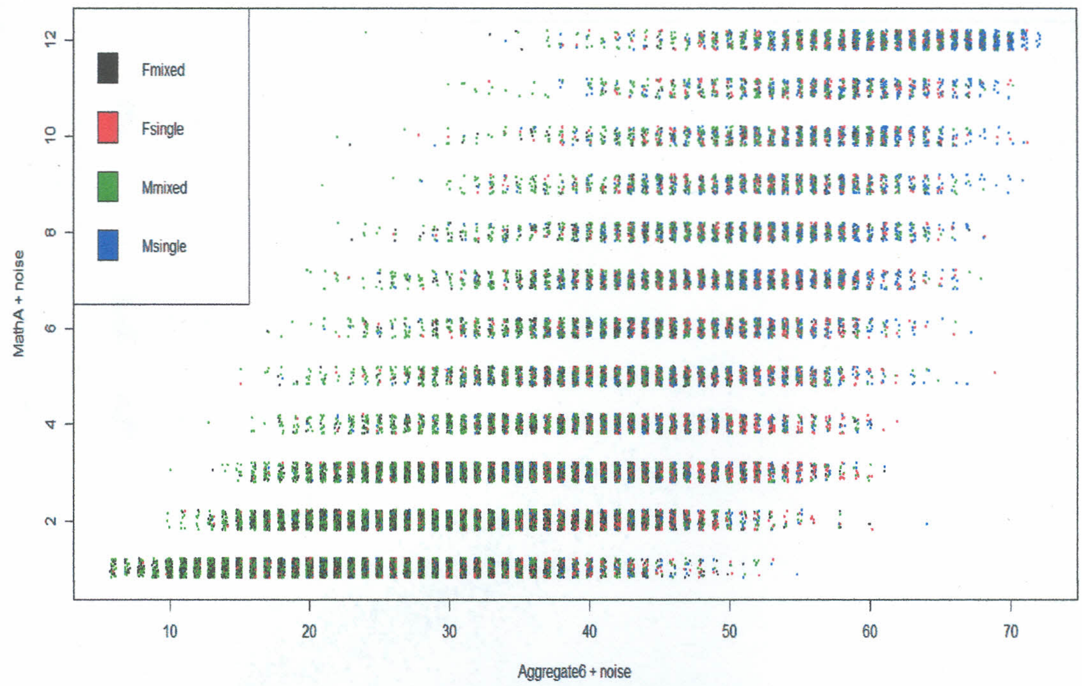



*Figure 17: A Scatter diagram of Mathematics scores versus Aggregate 6 in the different school types*

Figure 17 above is a scatter plot with maths score on the vertical axis and aggregate6 on the horizontal axis. It shows that the variables are strongly related, the crosses in the scatter lie closely to some straight line. Not all of the points are exactly on a straight line but the crosses extend from bottom left to top right indicating a positive relationship. This suggests that Mathematics score could be a factor in determining the aggregate score of a student in KCSE. Furthermore, the distributions of Mathematics grades vs. aggregate points by gender and type of school suggest a pattern between aggregate points and Mathematics score.

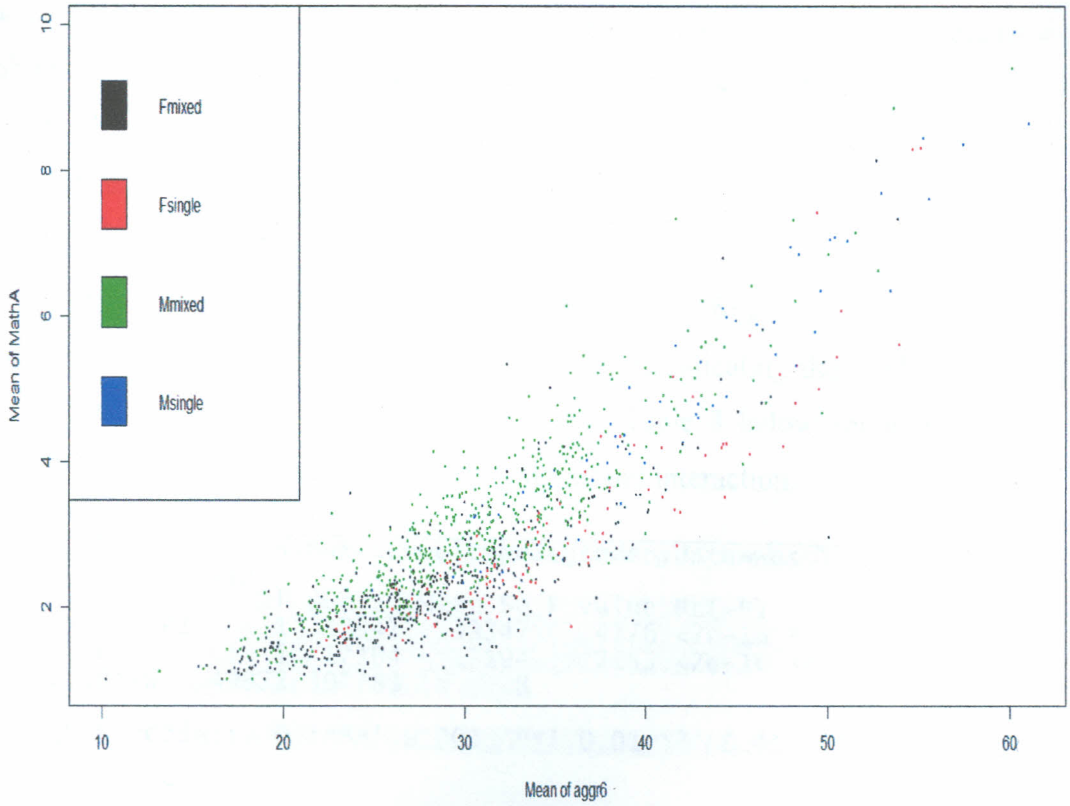



*Figure 18: A scatter plot of school Mathematics means versus means of aggregate6 for the different school types*

Figure 18 above demonstrates a possible linear relationship between the mean aggregate points of six subjects without Mathematics and the mean Mathematics score in schools. Not all the points are exactly on a line. Nevertheless, because the relationship between mathematics means and average aggregate6 appears to be approximately linear, it seems reasonable to represent the general relationship between the two variables using a straight line. This suggests that on the average, schools with low Mathematics achievement tend to have low aggregate points. The plot shows evidence of schools doing well in Mathematics having a higher aggregate and clustering is also clearly seen.

## 4.3: School type and gender effects

In this section, ANOVA was used for modelling the relationship between the students' Mathematics score which is the dependent variable and the explanatory variables that included the gender of a student and school type. This was with the interest of assessing the effect of explanatory variables on the response. To determine how much the variation in maths scores can be attributed to school type and gender, first we consider separate effects for schooltype2 (single gender school or mixed school) and gender. We have two factors that are school type at school level and gender at students' level.

The model considered is

Maths score$_{ijk}$ = $\beta_o + \beta_1$ schooltype2$_j$ + $\beta_2$ gender$_k$ + $\varepsilon_{ijk}$

Running this model, we obtain the analysis of variance indicating that both the type of school and gender are significant. The ANOVA output in Table 8 below shows the results for the analysis of variance and parameter estimates but without interaction.

```
> fit1<-aov(MathA.num~schooltype2+gender,data=west)
> summary(fit1)
               Df Sum Sq Mean Sq F value Pr(>F)
schooltype2     1  33347   33347    4176 <2e-16 ***
gender          1  17194   17194    2153 <2e-16 ***
Residuals   49812 397783       8
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> coef(fit1)
      (Intercept) schooltype2single              genderM
         2.213473          1.900174             1.194719
```

*Table 8: ANOVA model output when schooltype2 and gender included in the model*

We conclude from the small p-value for the F-statistic that there is significant difference between the single gender schools and mixed schools. There is also a significant difference between the boys and girls. The mixed school is the reference level for school type while female is the reference level for gender. We now include the interactions as a measure of interrelationships between schoolype2 and gender.

We have the model as

Maths score$_{ijk}$ = $\beta_o + \beta_1$ schooltype2$_j$ + $\beta_2$ gender$_k$ + $\beta_3$ schooltype2. Gender$_{jk}$ +$\varepsilon_{ijk}$

In the output in Table 9 below, the significant interaction implies that schooltype2 and gender do not act independently. However, the F-ratios for the significant main effects are larger than that of the interaction. This means that schooltype2 and gender maybe dominant over the interaction and calls for further analysis such as profile analysis.

```
> fit1.1<-aov(MathA.num~schooltype2*gender,data=west)
> summary(fit1.1)
                      Df Sum Sq Mean Sq F value Pr(>F)    
schooltype2            1  33347   33347  4185.2 <2e-16 ***
gender                 1  17194   17194  2157.9 <2e-16 ***
schooltype2:gender     1    894     894   112.2 <2e-16 ***
Residuals          49811 396888       8                   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> coef(fit1.1)

(Intercept)          schooltype2single                gendeRM
2.3345987                    1.6054413              0.9956436
schooltype2single:genderM
                  0.5720303
```

*Table 9: ANOVA model output when schooltype2, gender and their interactions are included in the model*

A close inspection of Table 9 above shows that both the main effects (school type and gender) and their interactions are significant.

When we model using the schooltype4 which takes into account gender,

$$\text{Maths score}_{ij} = \beta_o + \beta_1 \ \text{schooltype4}_j + \varepsilon_{ij}$$

It can be seen that the output in Table 10 below is the same as the previous output in which interaction is taken into account.

```
> fit2<-aov(MathA.num~schooltype4,data=west)
> summary(fit2)
               Df Sum Sq Mean Sq F value Pr(>F)    
schooltype4     3  51436   17145    2152 <2e-16 ***
Residuals   49811 396888       8                   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> coef(fit2)
        (Intercept) schooltype4Fsingle  schooltype4Mmixed
          2.3345987          1.6054413          0.9956436

        schooltype4Msingle
               3.1731152
```

*Table 10: Table displaying ANOVA model output when schooltype4 variable with four levels is used*

Although a significant F-test with a schooltype4 factor is obtained, the schooltype4 effect cannot be interpreted directly. The ANOVA can only tell that the 4 levels in schooltype4 differ but we can't tell which levels differ. The output says we have strong evidence that at least two of the means differ.

A Tukey's test was then used after ANOVA in additional exploration of the differences among the pairs of means. This is to provide information on which means were different from each other and we were then able to estimate how large the differences were.

```
> TukeyHSD(fit2)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = MathA.num ~ schooltype4, data = west)

$schooltype4
                      diff        lwr        upr p adj
Fsingle-Fmixed   1.6054413  1.5058764  1.7050062     0
Mmixed-Fmixed    0.9956436  0.9138158  1.0774714     0
Msingle-Fmixed   3.1731152  3.0692931  3.2769373     0
Mmixed-Fsingle  -0.6097977 -0.7017810 -0.5178144     0
Msingle-Fsingle  1.5676739  1.4556732  1.6796747     0
Msingle-Mmixed   2.1774716  2.0808963  2.2740469     0
```

*Table 11: Table showing the Tukey's test output in exploration of the differences among the pairs of means*

From Table 11 above, the differences are different from 0 and all the bands do not include 0, so we conclude that all these pairs of means are significantly different. The following can be summarised from the above ANOVA outputs in Table 9 and Table 10.

- Candidates from single gender schools had a higher score in Mathematics compared to candidates from mixed schools.
- Female candidates from purely girls school had a significant higher maths score by 1.605 than females from mixed schools (95% CI for the difference is 1.51 to 1.71)
- Male candidates from mixed schools had a significant higher maths score by 0.996 than females from mixed schools (95% CI for the difference is 0.91 to 1.08).
- Male candidates from purely boys school had a significant higher maths score by 3.173 than females from mixed schools (95% CI for the difference is 3.07 to 3.28).
- The gender of the student was also significant. The male candidates had higher scores in Mathematics.

## 4.4: Integrating fixed and random effects

The objective was to determine the effects of the school types and the gender on maths score and there was a difference from the ANOVA outputs in the previous section. With the linear mixed model, we include the additional random effect term (school represented by the schoolRandId in Table 12 below to represent the clustered and dependent maths score arising from students having the same experiences in the same school. A mixed model description of the data set is given by

$$Mathscore_{ij} = \beta_o + \beta_1 schooltype2_j + \beta_2 gender_k + \Upsilon_i \; school_i + \varepsilon_{ij}$$

Where:

$\beta_o$ is the overall maths score mean

$\beta_1$ is the coefficient of the of school type effect

$\beta_2$ is the coefficient of gender effect on the maths score

$\Upsilon_i$ is the random variable for each school is identified by schoolRandId

$\varepsilon_{ij}$ is the random error

The fixed effects are schooltype2, gender and their interaction whereas the school effect and the error are random effects.

```
>library("lme4", lib.loc="C:/Users/Mbasu/Documents/R/win-
library/2.15")
>fit2<-lmer(MathA.num~schooltype2*gender+(1|schoolRandId),west)
>print(fit2,digits=6,corr=FALSE)
Linear mixed model fit by REML
Formula: MathA.num ~ schooltype2 * gender + (1 | schoolRandId)
   Data: west
    AIC    BIC  logLik deviance REMLdev
 234953 235006 -117471   234928  234941
Random effects:
 Groups       Name        Variance Std.Dev.
 schoolRandId (Intercept) 1.1720   1.0826
 Residual                 6.3198   2.5139
Number of obs: 49815, groups: schoolRandId, 699

Fixed effects:
                           Estimate Std. Error t value
(Intercept)               2.1592124  0.0524866 41.1384
schooltype2single         0.9287495  0.1225554  7.5782
genderM                   1.0121117  0.0291337 34.7402

schooltype2single:genderM 0.9496943  0.1882502  5.0449
```

*Table 12: Table showing the Linear Mixed Model output with school type and gender as fixed effects while school as a random effect*

From the output in Table 12 above, we state our prediction equation in light of the estimated coefficients and have our equation as;

$$Mathscore_{ij} = 2.1592 + 0.9287_1 schooltype2_j + 1.0121_2 gender_k + 1.1720_{i1} \ school_{i1} + \varepsilon_{ij}$$

In the output, we have the Bayesian information criterion (BIC) and Akaike information criterion (AIC) that are used to measure the model adequacy. The absolute BIC and AIC computed values have no interpretation in this case however, the values can be computed for two or more models, and the values can then be compared. A smaller BIC/AIC indicates a better model. Deviance is another way to compare models.

The deviance value 234,928 gives an overall measure of how well the model fits the data. While the method of least squares is the method usually adopted when fitting models involving fixed effects only, the method of maximum likelihood is the method used by REML. This method calculates an expression known as the likelihood, which measures how well the model fits the data. The better the fit to the data the smaller is the value of the $-2$ log likelihood. By comparing the deviance values derived from separate models one can determine which model provides a better fit to the data.

The estimate of variance parameter for the random effect was 1.1720. This variance component provides a measure of variation directly associated with the schools random effect showing that there was variation among the schools with a corresponding standard deviation of 1.0826. The residual variance 6.3198 provides a measure of variation that cannot be explained with a corresponding standard deviation of 2.5139. The number of candidates on which observations are made is stated as 49,815 to which the model fits and the number of schools is stated as 699 for the random factor.

The other part represents the estimates for the fixed effects and their standard errors. The fixed-effect "intercept" = 2.1592 represents an estimate of the average level of maths score with a standard error of 0.0525. The coefficient labelled schooltype2single = 0.9287 represents the difference between the average level of maths score in single gender schools and the mixed schools with a standard error of 0.1226. Single gender schools had an average maths score higher by 0.9287(±0.2403) than mixed schools. The coefficient labelled genderM = 1.0121 represents the difference between the average maths score in boys and girls with a standard error of 0.0291. Male candidates had an average maths score higher by 1.0121(±0.0570) than female candidates. The coefficient for the interaction labelled schooltype2single.genderM= 0.9497, gives the average change in the within single gender school type associated with the change in the gender with a standard error of 0.1883.

## 4.5: Discussion of findings

In Kenyan schools, students existed within a hierarchical structure such as schools, school types and districts. The performance in the KCSE exam was at multiple levels. Student achievement in the KCSE exam was viewed at as an individual occurrence of a school type effect. There were individual student factors, school or school type effects that affected achievement in Mathematics. Policies are put in place to address improvement of student performance as well as school performance.

When analysis is done for the KCSE results at any level, they tend to ignore the gender and school type differences or they do not address it adequately. A mixed model allowed us to estimate the effects of individual characteristics like Mathematics score in KCSE performance and then better account for the variability in student achievement between different school types and gender. However, the observations made on a students' mathematics score results are not independent i.e. the maths scores within a school or school type are more similar to each other than to other schools due to the characteristics of the school.

The linear mixed model allowed us to model differences in mathematics scores associated with school types and gender while bearing in mind that Mathematics scores within a given school are not necessarily independent. By applying a mixed model with fixed and random effects, we compare the linear mixed model with that obtained by ANOVA. For this case, gender and school types were considered as fixed effects since they were not a random sample from the population of all possible levels. We can only have a male or female candidate and again we can only have a single gender school or a mixed school thus fixed effects parameters tell how the means differ between school types and gender.

A student scored any grade in Mathematics irrespective of the school and this was thought of as a random selection from a much larger collection of schools thus representing the general variability among schools accounting for school to school differences. The linear mixed model used more of the information contained within the results than when we use ANOVA. This was because it dealt with and combined information from different data layers; providing appropriate and correct standards errors consequently improving the precision of fixed effects comparison.

With the random term specified in the mixed model, the estimate of the constant is reduced from 2.33 to 2.16. Candidates from single gender schools have an average Mathematics score higher by 0.93 than those from mixed schools. A male candidate has now an average Mathematics score higher by 1.0 point than a female candidate. This is due to taking into account the variations between schools in maths scores whereas the ANOVA assumed variation to be the same in different schools. When the random term was included in the linear mixed model, the relationship between maths scores and the independent variables school type and gender changed. Therefore, the linear mixed model with fixed and random effects describes more accurately the different layers of variation associated with the hierarchical data providing a more appropriate and correct analysis.

### 4.5.1: Mathematics scores differences

Naturally, some students were more intelligent than others or they understood what they were taught in Mathematics better than others leading to variability between students' performance in the mathematics scores. Mathematics is important and the foundation for science and technological education at higher levels. For this reason, it is a compulsory subject in secondary school however, a great majority performed poorly at the 2011 KCSE examination not knowing its value in their subsequent university or tertiary education.

Universities expect students to have scored a certain grade in Mathematics; this grade determines whether these students are to be scientists, accountants, technicians, engineers or Mathematics teachers (http://jab.uonbi.ac.ke/cluster_information). With the kind of patterns seen in the KCSE results many students fail to get direct admission to public universities or fail to get on courses they had wished to pursue as a result of not performing well in Mathematics.

### 4.5.2: School type and gender differences

It was noted that there were more girls in purely girls' schools than there are boys in purely boys schools. The region also had more purely girls' schools than purely boys' schools. Conversely, we have more boys in mixed schools than girls. This could be because of the perception that girls do better if they learn in a purely girls school than when they learn in a mixed school environment.

The variability in the Mathematics scores in KCSE between the school types could have been caused by differences between the backgrounds of the students. Students tend to be more similar to each other if they are from the same school than when they are randomly sampled from different schools. This is because students are not randomly assigned to schools from the population, but rather are assigned to different types of schools based on their KCPE marks. Thus students within a particular school type tend to have scored certain range of marks in their KCPE exam. Certain school types had a fairly selective group of students and the schools tend to be homogenous in terms of educational preparation and experiences.

The estimates furnished by the linear mixed model are adjusted according to how reliably they have been approximated which is done by borrowing strength from the data of the full sample and shrinking each estimate towards the school's average. Students in a particular school are more similar to each other because they share the same teachers, environment and experience.

The performance in purely single gender schools differed significantly from the mixed schools because the school types vary in their characteristics especially with respect to the students they admit. The school membership is determined by a non-random process of Form one selection.

# Summary, Conclusions and Recommendations

## Summary

This thesis sought to analyse the differences in mathematics scores between different school types and between genders and the correlation between these scores and students' aggregate point scores. In addressing the research questions, an ANOVA model containing fixed effects for school type and gender was fitted into the data. Using a linear mixed model we extended the model by introducing random effects for school and used the method of restricted maximum likelihood (REML) to fit the linear mixed model. The outputs produced by REML are described and compared with outputs produced by the ANOVA method.

Although the presentations of results are different, analyses of variance and parameter estimates and standard errors are shown to be the same when no random terms are included in the model. Random term for school is then added in the linear mixed model and the output compared with that obtained by ANOVA. Having fitted the linear mixed model, the interpretation and significance of the effects are discussed. The study explored the multilevel structure of the data by dealing effectively with layers in the data and gave more valid approximate significance tests and standard errors.

**Conclusion**

In conclusion, this thesis has demonstrated how further analysis of the KCSE examination results can be done. This thesis begins with a simple question, how does the KCSE mathematics achievement relates to the aggregate points scored by candidates. As is true of many other investigations into educational phenomena, the answer is much more complex than expected. As the investigation was correlational but not experimental in design, assigning of a causal status to any particular variable has to be done with due caution.

Notwithstanding this, the analysis suggests that the poor aggregate points in KCSE could be due to poor mathematics score achievement by candidates. There was a strong correlation between mathematics scores and aggregate points scored by candidates. Maths was identified as an important factor in getting a good KCSE aggregate score. Maths scores were strongly correlated with Physics and moderately positively correlated with the languages i.e. English and Kiswahili.

There was a significant difference in KCSE maths score between single gender schools and mixed schools in Western Kenya. Candidates from single gender schools had a higher maths scores than those from mixed schools. Partly this may be due to the academic status of the single gender schools which are mostly boarding schools, have more learning resources and students in these schools have more time at their disposal for academic pursuits.

There was a significant difference in KCSE maths score between boys' schools and girls' schools in Western Kenya. Partly this may be due to the strong motivation in boys, social in nature, to do well in maths, could be a contributing factor as reflected in boys' positive attitudes towards the study of mathematics unlike girls. Going by the findings reported here, KCSE mathematics achievement as is true with any other educational achievement, is a multifaceted phenomenon.

**Recommendations**

Critical studies should be carried out to determine why there was a significant difference in Mathematics scores between the different types of schools and between genders as evidenced from the data. Studies should also be carried out to verify if various action plans put in place in institutions increase statistics, mathematics and pedagogical knowledge among the participating students and if so, in what ways does this happen? For example, the study could involve integrating qualitative techniques to investigate the perception of using resources as such GeoGebra, Computer Assisted Statistics Textbooks (CAST) and GTL in teaching and learning mathematics and statistics in schools.

The aggregate points' scored by students depends on complex relationship of variables apart from maths score, gender and school type. Therefore, there is need for future and further research on these and more variables for better understanding of the KCSE results. Further, application of the research findings will consequently inform policy makers for informed decision making.

# References


African Maths Initiative. (2013). *African Maths Initiative*. Retrieved from Kenyan Maths Initiative: http://kenya.africanmathsinitiative.net/

Aitkin, M., & Longford, N. (1986). Statistcal modeling in school effectiveness studies . *J. R. Stat. Soc. A149* , 1-43.

Aitkin, M., Anderson, D., & Hinde, J. (1981). Statistical modelling of data on teaching styles. *J. R. Stat. Soc. A148* , 144-161.

Allan, E., & Rowlands, J. (2001). Mixed Models and Multilevel Data Structures in Agriculture.

Bryk, A. S., & Raudenbush, S. W. (1992). Hieriarchical linear models: Applications and data analysis methods. Newbury Park, CA: Sage Publications.

Cai, J., & Gorowara, C. C. (2002). Teachers' conceptions and constructions of pedallogical representations in teaching arithmetic average. *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a statistically literate society.* Cape Town, South Africa.

Cohen, J., Cohen, P., West, S., & Aiken, L. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* NJ: Erlbaum: Hillsdale.

Donner, A., & Klar, N. (1994). Cluster randomisation Trials in Epidemiology: Theory and Application . *Journal of Statistical Planning and Inference* , 42, 37-56.

Donner, A., Piaggio, G., & Villar, J. (2003). *Cluster randomization trials: Power considerations* (Vol. 26(3)). Evaluation and theHealth Proffessionals.

Douglas, W. J. (1999). *Basic Concepts in Hierarchical Linear Modelling with Applications for Policy Analysis.*

Eleanor, A., & John, R. (2001). *Mixed Models and Multilevel Data Structures in Agriculuture.* Reading: The University of Reading Statistical Services Centre.

Gelman, A., & Hill, J. (2008). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York: Cambridge University Press.

Goldstein, H., & Woodhouse, G. (2001). *Modelling repeated measurements in multilevel modeling of health statistics.* John Wiley & sons, Ltd.

Gonzalez, M. e. (2011). *Teachers' Graphical Competence. In C. Batanero, G. Burril, and C.Reading (eds.), Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICME/IASE Study,Springer Science+Business Media B.V. 2011. .*

Good, T. L., & Brophy, J. E. (1986). *School effects; Handbook of research on teaching.* New York: Macmillan.

Gould, R. (2010). "Statistics and the Modern Student". *International Statistical Review* , 78,97–315.

Gray, J. (1989). Multilevel models: issues and problems emerging from recent application in British studies of school effectiveness. In *Multilevel analyses of educational of educational data* (pp. 127-145). Chigago: University of Chigago press.

Groth, R. E., & Bergner, J. A. (2006). Preservice elementary teachers conceptual and procedural knowledge of mean, median and mode. In *Mathematical Thinking and Learning* (pp. 37-63).

Heyns, B. (1986). Educational effects: Issues in conceptualisation and measurement. In J. G. Richardson, *Handbook of theory and research for the sociology of education* (pp. 305-340). Westport CT: Greenwood Press.

Jacobbe, T., & Carvalho, C. (2011). *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A joint ICME/IASE Study.*

Japan International Cooperation Agency. (2007). *Analysis from a Capacity Development Perspective - SMASSE.* JICA.

KNEC. (2009). *K.C.S.E. Annual Report.* Nairobi: Kenya National Examiniation Council.

Kreft, I., & de Leeuw, J. (1998). *Introduction to Multilevel Modeling.* CA: Sage: Thousand Oaks.

Laird, N. M., & J, H. W. (1982). "Random-Effects Models for Longitidinal Data" Biometrics.

Lee, Nelder, & Pawitan. (2006). Generalized Linear models with Random effects: Unified Analysis via H-likelihood .

Lewis, N., Lydia, W., Beth, M., & Nyabisi, E. (2012). Critical determinants of Poor Perfomance in KCSE among Girls in Arid and Semi-Arid (ASAL) Regions in Kenya. *Journal of African Studies in Educational Management and Leadership , Vol:2 No.1.*

Lucas, A. O. (1993). *A study of predictive Validity of the Kenya Certificate of Primary Education: Application of Hierarchical Linear Model.* Vancouver, Canada: The University of British Columbia.

Maundu, J. N. (1986). *Student achievement in science and mathematics : a case study of extra-provincial, provincial, and Harambee secondary schools in Kenya.* McGill University Libraries.

Mbunzi, Stephen, Nagda, & Sonal. (2009). *Mixed model analysis using R.* Nairobi: ILRI.

Monteiro, C., & Ainley, J. (2003). Developing Critical Sense in Graphing. . *Proceedings of III Conference of the European Society for Research in Mathematics Education.* .

Nezlek, J. B., & Zyzniewski, L. E. (1998). Using hierarchical linear modelling to analyse grouped data: Group dynamics.

North, D., Scheiber, J., & Ottaviani, G. (2010). Training teachers to teach Statistics in South Africa: Realities and attitudes. *Proceedings of 8th International Conference on Teaching Statistics.* Ljubljana, Slovenia.

Rabe-Hesketh, S., & Skrondal, A. (2005). *Multilevel and longitudinal modelling using Stata.* StataCorp LP.

Stern, R. D., Coe, R., Allan, E. F., & Dale, I. (2004). *Statistical Good Practice for Natural Resources Research.* Wallingford, UK: CABI Publishing.

Ukoumune, O., Gulliford, M., & Chinn. (2004). On the distribution of random effects in a population-based multi- stage cluster sample survey. *Journal of Official Statistics* , 481-493.

Venables, W. N., & B, D. R. (1999). *Modern Applied Statistics with S-PLUS* (3rd Edition ed.). New York: Springer.