Journal of Statistical and Econometric Methods, vol.3, no.2, 2014, 31-46
ISSN: 2241-0384 (print), 2241-0376 (online)
Scienpress Ltd, 2014

# Evaluating Secondary School Examination Results: Application of Principal Component Analysis

**Elizabeth W. Njoroge[1], Gladys G. Njoroge[2] and Dennis K. Muriithi[3]**

## Abstract

Results from Kenya National Examination Council (KNEC) indicate that there are schools that have had an upward trend in performance while others have continued to show a decline. This paper seeks to find out the principal components, in terms of subjects, that contribute to this performance. Principal Component Analysis (PCA), a data reduction procedure was applied to assess the performance of the national examination at the Kenya Certificate of Secondary Examination (KCSE) level for the last three years. The schools were purposively selected from Nyanza, Nairobi, Rift Valley and Eastern provinces. Secondary data from KNEC was used and analyzed using SPSS software. The PCA brought out the component loadings and the correlation structure between the different subjects; as a result one component was extracted. The results provided evidence that all the subjects are highly correlated and the first component having the highest variance. This principal component emerged to be English language. Being the subject with the highest sum of the squared loadings, it was concluded that it played the greatest role in performance of the examinations.

**Mathematics Subject Classification:** 62H25; 62G10; 97K40; 97K80
**Keywords:** Correlation structure; performance; Eigen values; principal components

[1] Chuka University, Faculty of Business studies. E-mail: elizawn@gmail.com
[2] Chuka University, Faculty of Sciences. E-mail:gg.njoroge@gmail.com
[3] Chuka University, Faculty of Business studies. E-mail:kamuriithi@yahoo.com

Article Info: *Received* : December 7, 2013. *Revised* : February 28, 2014.
*Published online* : May 15, 2014.

# 1. Introduction

## 1.1 The concept of Principal Component Analysis

Principal Component Analysis first described by Pearson in 1901 and Hotteling in 1933 prepared a fully functional method that generates a set of orthogonal axes, placed in decreasing order and determining the main directions of variability of samples. It enables researchers to reduce the number of possible groupings. It is significant due to the occurrence of some redundancy in variables. In this case, redundancy means that some of the variables are correlated with one another, because they are measuring the same construct. Therefore Principal Component Analysis replaces many original characters with only a few most significant principal components (PCs) which represent combinations of closely correlated original characters.

## 1.2 Background of the problem

The education taskforce report (February 2012) by the ministry of Education on the re-alignment of the education sector to the constitution of Kenya 2010 stated that, The vision of education service provision is to have a globally competitive quality education, training and research for Kenya’s sustainable development. This is in line with the Millennium Development Goal (MDG) number 2 discussed at the 2000 United nations millennium summit, ”Universal Education to eradicate poverty by 2015”. The report came up with specific goals to be achieved by the education sector. Some of these are:

i. To promote the socio-economic, technological and industrial skills for the country‘s development

ii. To promote individual development and self-fulfillment.

With this in mind then it is very necessary to prepare a workforce (citizens) that is confident, responsible and dependable. This can only be achieved if a thorough study is done on what issues affect performance of students.

This research is meant to bring out clearly the components in terms of subjects done especially at secondary school level that could be affecting performance of students at the same level and at the university level. Many students at these two levels seem to have a common problem of expressing their ideas especially in writing. Therefore it was found necessary to do a scientific research on KCSE (Kenya Certificate of Secondary Examinations) results for several categories of schools for better conclusions. The use of Principal component analysis intend to bring out the specific linear combinations in terms of subjects referred to as the principal components which could be the main causes of the given performance. When students perform poorly in their examinations and are not able to express them erodes their confidence in life. This can be a very desirous effect upon the economy of this country as these individuals loose enthusiasm to get jobs in places where they face challenges.

## 1.3 Objectives of the study

The main purpose of this study was to identify the principal components that affect secondary school performance in terms of subjects and make appropriate recommendations to the secondary school stakeholders.

Specific Objectives:

i. To apply PCA to reduce the dimensionality of examination results data in Kenya.

ii. To study the variation trends and reveal the relations of the components at the different levels.

iii. To identify the principal components affecting the performance in the all the schools categories combined and in one other school to compare results and make the appropriate recommendations.

## 1.4 Significance of the study

Academic performance in various subjects specifically in Kenya, determines to a great extent the career opportunities an individual is exposed to. It is therefore important for the stake holders in the education sector to be well informed about the weights borne by each subject and its influence on the examination results.

# 2. Literature review

A research paper presented for the Master of Science in Statistics, Njoroge E. W (2007) applied the principal component analysis on secondary school performance at KCSE. She took one case study and found that English turned out to have the highest incidence on performance followed by Mathematics. This meant that they were the biggest contributors to performance and stake holders were advised to put more emphasis on these two subjects. This prompted research of principal components in KCSE performance to more schools across the country and of different categories.

In a paper presented to the Kenya National Academy of Science, on the “Application of Canonical Correlation Analysis (CCA) in Educational performance evaluation” Odhiambo and Weke (1988) tried to reduce the dimensionality of courses in pure Mathematics, to four, which are significantly related to corresponding dimensions of courses in statistics for second year of study during the 1980 and 1981 academic year. There exists a strong relationship

between CCA and PCA in that there is data reduction in both concepts.
In his study on Foreign exchange Market variation, Mwai (2005) was able to reach a conclusion, using PCA, that when a set of variables are highly correlated, the bulk of variation in the data can be attributed to the first Principal Component alone. In their research work ' Public Parks Aesthetic Value Index', ( Mohamad & Nurashikin, 2007) sought to find that the non-existence of this public parks aesthetic value index in Malaysia, brings the issue of the uncertainty whether the existing public parks are functioning well in giving relaxation, enjoyment and pleasuring to the community and visitors. The study aimed to introduce an index to assess the level of aesthetics value of public parks and applied it to assess the aesthetic value level of certain public parks in Malaysia, (Mohamad & Nurashikin, 2007) found that the index that can best explaining the real condition or situation is the index that focused on dominant variables in each scope or classification. In this study, it refers to dominant variables in the scope of tree, fauna, lake and flower.

PCA need to be conducted to determine the level of aesthetic level for each attribute within the benchmarking scale of 0-10. This analysis involved all respondents without separation. The study showed that the attribute of quality of flower has the highest ranking compared to the other five attributes with the value of 5.12. The other attributes were water quality 4.13, colour of water 3.99, quantity of flower 4.75, colour of flower 4.90 and odour of flower 4.28.

In recent studies they noted that the herding phenomenon (simultaneously trade the same stocks in the same direction) exists in the behavior of institutional. The study focused on the U.S. fund trading sample because herd behavior became increasingly important when large institutional investors dominated the market. To estimate herding by fund managers, the study applied the (Lakonishok, Shleifer, & Vishny, 1992) measure and (Wylie, 2005) trinomial-distribution approaches which considers buy, hold and sell positions for a given stock in a period. The study calculated conditional herding measures for stock-months that have a higher or lower proportion of fund buyers relative to the expected proportion in each month, which are the buy-herding measure and the sell-herding measure.

To estimate investor sentiment for each stock-month, this paper employed the principal component analysis as the means of extracting the composite unobserved sentiment measure, rather than just select a single indicator to proxy sentiment. (Fisher, Kenneth, & Statman, 2000) Neal and Wheatley, 1998).(Brown, Draper , & McKenzie., 2005)indicated that this estimating procedure is able to successfully extract measures of unobserved sentiment from various indicators.

The findings indicated a significantly positive association between the sentiment measure and subsequent sell-herding, after controlling the fund performance deviation, the capitalization of the stock, the number of funds trading the stock, the net mutual fund redemption of the stock, and the market-to-book ratio of the stock. However, the evidence showed no significant correlation between the composite sentiment measure and subsequent buy-herding. These findings suggested that institutional investors herd on the selling side when they observe high level of investor optimism, consistent with the intuition that rational institutional investors tend to counteract the optimistic sentiment of the investors (Wermers, 2002).

In their paper, Principal component estimation for generalized linear regression, (Marx & Smith, 1990) developed and presented an asymptotically biased Principal component parameter estimation technique, as an option to traditional maximum likelihood estimation for parameter estimation technique, as an option to traditional maximum likelihood estimation for generalized linear regression. PCR utilizes explanatory variables that are standardized so that X'X is proportional to correlation matrix. Thomas et al (2003) in their paper "Assessment of water quality for human consumption" applied Chemomatric approaches (Cluster analysis and PCA) to a portable water monitoring system and showed that the data classification by cluster analysis and data structure by principal component reveals similar results. He concluded that water quality is influenced by location. Do Kim-Anh.et al (1990) in their paper "Discriminant analysis of event-related potential curves using smoothed principal component" used PCA, enhanced by the use of smoothing in conjunction with Discriminant analysis technique to device a statistical classification method for analysis of events related potential data. A training set for pre-medication potentials collected adolescents with Attention deficit disorder (ADHD) which they used to predict whether adolescents from independent subject group will respond to the long term medication.(Thaddeus , Apponi, McCarthy, & Gordon, 1999)in his journal "self consistency and principal component analysis," examined the self consistency of a principal axis and found that when a distribution is centered about principal component axis, a PCA axis of a random vector X is self consistence if each point on the axis corresponds to the mean of X given that projects orthogonally onto that point. Lindsay (2002) gave a tutorial on PCA as he stated that it is a statistical technique that has found applications in field such as face recognition and image compression and is a common technique for finding patterns in data of high dimension. (Wulder, 2005) compared and contrasted PCA and Factor Analysis. He gave their direct and indirect uses, and applied KMO test to assess the input variables and defined communalities concluding that high communality means most important and vice versa.

# 3. Principal component analysis procedures

## 3.1 Derivation of Principal Component

Suppose $X^{'} = \left[ x_1, x_2, \ldots, x_p \right]$ is a $p$ dimensional random variable with mean $\mu$ and covariance matrix $\Sigma$. We need to find a new set of variables say $Y_1, Y_2, \ldots, Y_p$ which are uncorrelated and whose variances

$$X \to Y = Q^{'}\left(X - \mu\right) \qquad (3.1)$$

where $Q$ is an Orthogonal matrix and $Q^{'}\sum Q = G$ such that $G = dia\left(\lambda_1, \lambda_2, \dots, \lambda_P\right)$ where $\lambda_1 > \lambda_2 > \dots > \lambda_P > 0$. Each $Y_j$ is taken to be a linear combination of $X_{i's}$ so that

$$Y_j = \beta_{1j}X_1 + \beta_{2j}X_2 + \dots + \beta_{pj}X_pY_j = \beta_j^{'}X \qquad (3.2)$$

Where $Q = \left(\beta_1, \beta_2, \dots, \beta_p\right)^{'}$ is the set of eigenvectors corresponding to Eigen values $\lambda_1, \lambda_2, \dots, \lambda_P$ and $\beta_j^{'}\beta_j = \sum_{k=1}^{p}\beta_{kj}^2 = 1$ (3.3)

The positivity of $\lambda_i$ is guaranteed if $\Sigma$ is positive definite. The representation $\Sigma$ follows from the spectral decomposition theorem. The first PC is found by choosing $\beta_1$ so that $Y_1$ has the largest possible variance. That is we choose $\beta_i$ so as to maximize the variance of $\beta_1^{'}x$ subject to the constraint $\beta_1^{'}\beta = 1$. The 2nd PC is found by choosing $\beta_2$ so that $Y_2$ has the 2nd largest possible variance for all combinations of the form of equation (3.2) which are uncorrelated with $Y_1$ and similarly for $Y_3, \dots, Y_p$ so as to be uncorrelated and have decreasing variance. To find the first PC, we choose $\beta_1$ in order to maximize variance of $Y_1$ subject to the normalization constraint $\beta_1^{'}\beta_1 = 1$, implying that

$$Var\left(Y_1\right) = Var\left(\beta_1^{'}x\right) = \beta_1^{'}\Sigma\beta_1 \qquad (3.4)$$

Which is the objective function. Principal components may also be obtained from standardized variables.
Consider

$$Z_j = \frac{x_i - \mu_i}{\sqrt{\sigma_{ii}}} \qquad (3.5)$$

In matrix notation, $Z = \left(V^{1/2}\right)^{-1}\left(x - \mu\right)$. The diagonal standard deviation matrix $V^{1/2}$ is defined as

$$V^{1/2} = \begin{pmatrix} \sqrt{\sigma_{11}} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{\sigma_{22}} & 0 & \dots & 0 \\ 0 & 0 & . & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & \dots & \sqrt{\sigma_{nn}} \end{pmatrix}$$

$V^{1/2}\Lambda V^{1/2} = \Sigma$ ,Where $\Lambda = \left(p_{ij}\right)$ is the correlation matrix,

$$\Rightarrow \Lambda = \left(V^{1/2}\right)^{-1} \Sigma V^{1/2^{-1}}$$

If $Z = \left(z_1, z_2, \ldots z_p\right)'$ is the vector of standardized variables, then

$$E\left(Z\right) = E\left[V^{-1/2}\left(x-\mu\right)\right] = V^{1/2}E\left(x-\mu\right) = 0 \tag{3.6}$$

and

$$Cov\left(Z\right) = E\left[V^{-1/2}\left(x-\mu\right)\left(x-\mu\right)' V^{-1/2}\right] = V^{-1/2}E\left(x-\mu\right)\left(x-\mu\right)' V^{-1/2}\Big]$$

$$= V^{-1/2}V^{-1/2} = \Lambda \tag{3.7}$$

It follows that Principal Components of $Z$ may be obtained from the eigenvectors of the correlation matrix $\Lambda$ of $X$ .

Note, the proportion of the total sample variation due to $k^{th}$ Principal Component is

$$\frac{\hat{\lambda}_k}{\hat{\lambda}_1 + \hat{\lambda}_2 + \ldots + \hat{\lambda}_p} \quad \text{for} \quad k = 1, 2, \ldots, p \, .$$

Where total variation is equal to $tr\Sigma = \sum_{i=1}^{p} \lambda_i$.

# 4. Empirical Results

## 4.1 Descriptive Statistics

Table 4.0: Descriptive Statistics

| Subject | Mean | Std. Deviation | Analysis N |
|---|---|---|---|
| S101 | 7.920875 | 1.8762226 | 54 |
| S102 | 7.959627 | 1.9530585 | 54 |
| S121 | 6.447214 | 2.6382587 | 54 |
| S231 | 7.657015 | 2.1664033 | 54 |
| S232 | 6.989899 | 2.2234639 | 54 |
| S233 | 6.217062 | 2.1300728 | 54 |
| Humanities | 8.140385 | 1.8791143 | 54 |

| Subject | Mean | Std. Deviation | Analysis N |
|---|---|---|---|
| S101 | 7.920875 | 1.8762226 | 54 |
| S102 | 7.959627 | 1.9530585 | 54 |
| S121 | 6.447214 | 2.6382587 | 54 |
| S231 | 7.657015 | 2.1664033 | 54 |
| S232 | 6.989899 | 2.2234639 | 54 |
| S233 | 6.217062 | 2.1300728 | 54 |
| Humanities | 8.140385 | 1.8791143 | 54 |
| Applied | 8.099893 | 1.5317434 | 54 |

From Table 4.0 above it is evident that the performance in the science subjects is lower than others. The means are lower with a wider spread. Further analysis is required to determine the component.

## 4.2 Correlation Analysis

Table 4.1: Correlation Matrix

| | Subjects | S101 | S102 | S121 | S231 | S232 | S233 | Humanities | Applied |
|---|---|---|---|---|---|---|---|---|---|
| Correlation | S101 | 1.000 | .889 | .885 | .854 | .812 | .882 | .859 | .858 |
| | S102 | .889 | 1.000 | .887 | .920 | .854 | .907 | .911 | .830 |
| | S121 | .885 | .887 | 1.000 | .959 | .919 | .955 | .939 | .881 |
| | S231 | .854 | .920 | .959 | 1.000 | .919 | .947 | .957 | .874 |
| | S232 | .812 | .854 | .919 | .919 | 1.000 | .935 | .904 | .854 |
| | S233 | .882 | .907 | .955 | .947 | .935 | 1.000 | .945 | .883 |
| | Humanities | .859 | .911 | .939 | .957 | .904 | .945 | 1.000 | .865 |
| | Applied | .858 | .830 | .881 | .874 | .854 | .883 | .865 | 1.000 |
| Sig.<br>(one-tailed) | S101 | | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | S102 | .000 | | .000 | .000 | .000 | .000 | .000 | .000 |
| | S121 | .000 | .000 | | .000 | .000 | .000 | .000 | .000 |
| | S231 | .000 | .000 | .000 | | .000 | .000 | .000 | .000 |
| | S232 | .000 | .000 | .000 | .000 | | .000 | .000 | .000 |
| | S233 | .000 | .000 | .000 | .000 | .000 | | .000 | .000 |
| | Humanities | .000 | .000 | .000 | .000 | .000 | .000 | | .000 |
| | Applied | .000 | .000 | .000 | .000 | .000 | .000 | .000 | |

## 4.3 Test statistics (Bartlett’s test)

This is a special $x^2$ test. It is used to test if $k$ samples have equal variances. Equal variances across samples are called homogeneity of variance. Some statistical tests for example analysis of variance, assumes that variance are equal across groups or samples therefore Bartlett’s test is used to verify that assumption.

The Bartlett’s test is defined as: $H_\circ : \sigma_1 = \sigma_2 = \cdots = \sigma_k$ vs. $H_a : \sigma_i \neq \sigma_j$

$$T = \frac{(N-K)lnS_p^2 - \sum_{i=1}^{k}(N_i - 1)lnS_i^2}{1 + \frac{1}{3(k-1)}(\sum_{i=1}^{k}\frac{1}{N_i - 1} - \frac{1}{N-k})}$$

where, $T$ is the test statistic, $S_i^2$ is variance of the $i^{th}$ group, $N$ is the total sample, $N_i$ is sample size of the $i^{th}$ group, $K$ is the number of groups and $S_p^2$ is the pooled variance i.e. weighted average of the group variance and is defined as

$$S_p^2 = \sum_{i=1}^{k} \frac{(N_i - 1) S_i^2}{N - K}.$$

The significance level is $\alpha$ and the critical region is where $T > x^2(\alpha, k-1)$. This test was derived by Snedecor and Cochran (1983).

Table 4.2: KMO and Bartlett's Test[a]

| | | |
|---|---|---|
| | Kaiser-Meyer-Olkin | .929 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 772.563 |
| | Df | 28 |
| | Sig. | .000 |

Since the significant value (0.000)>0.05, there exist sufficient statistical evidence that the different in the subject is significant at 5% level. Also from the correlation matrix given on table 4.2 above, the one tailed test has shown that the tests for variable correlation is significant.

## 4.4 Extraction

According to Kaiser those P.C. whose eigenvalues are less than average, (less than1) should be excluded if the correlation matrix has been used. The first component accounts for approximately 92% of the total variance which is very high. This implies that the first P.C has a magnificent implication on overall results.

## 4.5 Component

Both Kaiser and Catelli Scree plots extraction procedures showed one component.

$$Y_i = 0.915X_1 + 0.940X_2 + 0.979X_3 + 0.977X_4 + 0.949X_5 + 0.980X_6 + 0.968X_7 + 0.912X_8$$

Table 4.3: Components loadings

| | Raw | Rescaled |
|---|---|---|
| Subject | Component | Component |
| | 1 | 1 |
| S101 | 1.717 | .915 |
| S102 | 1.835 | .940 |
| S121 | 2.584 | .979 |
| S231 | 2.118 | .977 |
| S232 | 2.110 | .949 |
| S233 | 2.086 | .980 |
| Humanities | 1.819 | .968 |
| Applied | 1.398 | .912 |

Table 4.5 above shows the component loadings (eigenvectors). It shows that all the subjects are highly correlated and the first principal component is a roughly weighted sum of the eight subjects.

## 4.6 Communalities

The $i^{th}$ communality is the sum of the squares of the loading of the $i^{th}$ variable on the $m$ common factors. Communality measures the percentage of variance in a given variable explained by all components jointly. Part of variance explained by common factors namely

$\sum_{k=1}^{m} \lambda_{jk}^2$ is called the communality of the $j^{th}$ variable which is the amount of variance explained by all components jointly, the maximum value it can reach is 1. E.g.Score in English $0.915^2 = 0.837$. Scores in Mathematics, $0.94^2 = 0.884$

In Table 4.4 below, it can be seen that communality of each subject is >0.8, implying that all the subjects have a similar pattern, hence are highly correlated, and that each contributes highly to the performance of schools under study.

Table 4.4: Communalities

| | Raw | | Rescaled | |
|---|---|---|---|---|
| | Initial | Extraction | Initial | Extraction |
| S101 | 3.520 | 2.948 | 1.000 | .837 |
| S102 | 3.814 | 3.368 | 1.000 | .883 |
| S121 | 6.960 | 6.677 | 1.000 | .959 |
| S231 | 4.693 | 4.484 | 1.000 | .955 |
| S232 | 4.944 | 4.450 | 1.000 | .900 |
| S233 | 4.537 | 4.353 | 1.000 | .959 |
| Humanitie<br>- | 3.531 | 3.309 | 1.000 | .937 |
| Applied | 2.346 | 1.953 | 1.000 | .833 |

## 4.7 Catelli Scree Graph

The Catelli Scree, plots components in the X axis and corresponding eigenvalues in the y axis. The curve makes an elbow towards less steep decline. This indicates where large eigenvalues cease and small eigenvalues start.

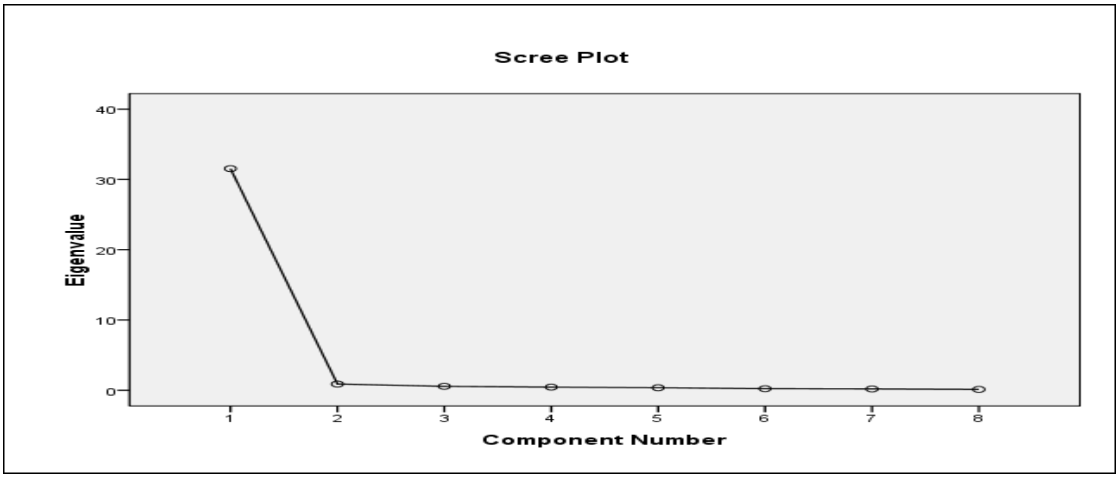

Figure 4.0: Cattelli Scree Graph

From the plot it is clear that the first component accounts for the highest variance.

## 4.8 Nairobi school results

Results for one randomly chosen school were run through the tests. This was just to confirm if there is some relationship of some kind with the analyzed results of all the categories of schools under study.

Table 4.5: Nairobi School Component Matrix

| | **Raw Component** | **Rescaled Component** |
|---|---|---|
| Subject | 1 | 1 |
| S101 | .248 | .932 |
| S102 | .323 | .655 |
| S121 | .227 | .660 |
| S231 | .114 | .435 |
| S232 | .991 | .968 |
| S233 | .593 | .919 |
| Humanities | .046 | .931 |
| Applied | -.200 | -.963 |

Table 4.6: Nairobi School Communalities

| Subject | Raw Initial | Rescaled Extraction |
|---|---|---|
| S101 | .071 | .062 |
| S102 | .244 | .105 |
| S121 | .119 | .052 |
| S231 | .068 | .013 |
| S232 | 1.047 | .982 |
| S233 | .417 | .352 |
| Humanities | .002 | .002 |
| Applied | .043 | .040 |

Table 4.7: Nairobi School Total Variance Explained

| | | Initial Eigen values | | | Extraction Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|
| | Component | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| Raw | 1 | 1.607 | 79.877 | 79.877 | 1.607 | 79.877 | 79.877 |
| | 2 | .405 | 20.123 | 100.000 | | | |
| | 3 | 1.768E-16 | 8.789E-15 | 100.000 | | | |
| | 4 | 4.163E-17 | 2.070E-15 | 100.000 | | | |
| | 5 | 4.310E-18 | 2.143E-16 | 100.000 | | | |
| | 6 | -5.269E-18 | -2.620E-16 | 100.000 | | | |
| | 7 | -9.470E-18 | -4.708E-16 | 100.000 | | | |
| | 8 | -2.487E-17 | -1.236E-15 | 100.000 | | | |
| Rescaled | 1 | 1.607 | 79.877 | 79.877 | 5.497 | 68.717 | 68.717 |
| | 2 | .405 | 20.123 | 100.000 | | | |

| | | | |
|---|---|---|---|
| 3 | 1.768E-16 | 8.789E-15 | 100.000 |
| 4 | 4.163E-17 | 2.070E-15 | 100.000 |
| 5 | 4.310E-18 | 2.143E-16 | 100.000 |
| 6 | -5.269E-18 | -2.620E-16 | 100.000 |
| 7 | -9.470E-18 | -4.708E-16 | 100.000 |
| 8 | -2.487E-17 | -1.236E-15 | 100.000 |

Figure 4.1: Nairobi School Scree plot

# 5 Conclusions and Recommendation

## 5.1 Conclusions

From the analysis above, it has been shown that the subjects are highly correlated. Therefore the variation in the data can be attributed to the first principal component alone, meaning that the first P.C provides a good measure of variance in performance of the schools. It accounts for 92% of the total variation. We can therefore conclude that the general performance in these schools is affected by **English** which has the greatest effect on all the other subjects. For Nairobi school, the first component accounts for approximately 80% of the total variance. While the second component accounts for 20% put together. This implies that the first and the second principal components have magnificent implication on overall results. Here the component loadings show that Physics, English, Humanities and Chemistry have high loadings.

From the communalities table it is clear that Biology has very little influence on the overall performance while Physics has a significant effect on the performance. The results of the one school showed that English has the greatest effect on all the other subjects while it is quite the opposite with Physics. In general the outcomes of the two sets of data are highly similar with the second set of data showing that Kiswahili also has some significant influence on the overall performance.

## 5.2 Recommendations

As lecturers, we have witnessed the weakness of language application, as we mark the examinations especially in Mathematics. We therefore recommend that the educators urgently re-evaluate the teaching of language subjects as they seem to affect the overall performance.

It is therefore also advisable that all stakeholders in the education sector re-evaluate the curriculum of the language subjects in general. As there could be other issues that affect schools' overall performance, it is recommended that further research be carried out in order to establish other factors that do negatively or positively influence the student's performance.

## References

[1] Mohamad , M. R., and Nurashikin, M., *Public Parks Aesthetic Value Index, Principal Component.* Universiti Putra Malaysia: InTech, (2012).

[2] Mwai, *Application of Principal Component analysis on, forex market variation.* University of Nairobi, Kenya, (2005).

[3] Njoroge, E. W., *Application of Principal Component analysis to secondary school results in Kenya.* Unpublished Masters project work, University of Nairobi, Kenya, (2007).

[4] Brown, G., Draper , C., and McKenzie., E., Consistency of UK pension fund investment performance. *Journal of Business Finance and Accounting*, 679-698, (2005).

[5] Fisher, J., and Levinson, J., *Environmental Aesthetics. In: the Oxford Handbook of Aesthetics ed.* . London: Oxford University Press, (2003).

[6] Fisher, G., Kenneth, L., and Statman, M., Investor Sentiment and Stock Returns. *Financial Analysts Journal*, 16-23, (2000).

[7] Lakonishok, J., Shleifer, A., and Vishny, R., The impact of institutional trading on stock prices*. *Journal of Financial Economics*, (1992).

[8] Marx, B., and Smith, E., Weignted Multicolliearity in Logistic Regression. *Canadian Jounal of Fisheries and Aquatic Science*, 70-85, (1990).

[9] Thaddeus, P., Apponi, A., McCarthy, M., and Gordon, V., *Laboratory Detection of HC6N, a Carbon Chain with a Triplet Electronic Ground State.* New York: Harvard-Smithsonian Center for Astrophysics, (1999).

[10] Wermers, R., Mutual Fund Herding and the Impact on Stock Prices. *The Journal of Finance*, 581–622, (2002).

[11] Wulder, H., Baker investor sentiment. *Journal of Financial Investiments*, 50-63, (2005).

[12] Wylie, S., Fund manager herding: A test of the accuracy of empirical results using U.K. data. *The Journal of Business*, 381-403, (2005).