MULTIVARIATE ANALYSIS OF VARIANCE ON THE ACADEMIC PERFORMANCE OF SECONDARY SCHOOL STUDENTS IN NATIONAL EXAMINATION BODIES IN NIGER STATE, NIGERIA

[1]Abubakar USMAN, [1]Ramatu ADAMU, [1]Rasheed A. ADEYEMI, [2]Abdullahi Abubakar WUCHIN

[1]Department of Statistics
Federal University of Technology, Minna, Nigeria
[2]Department of Mathematics, Faculty of Science,
Air force Institute of Technology, Kaduna, Nigeria
Corresponding Author: - abu.usman@futminna.edu.ng

Abstract

The research is aimed at determining the academic performance of secondary school students in National Examinations in Niger State, Nigeria, using multivariate analysis of variance. Two schools were randomly selected from each geopolitical zone of the state. The results of the Senior Secondary School Certificate Examinations (SSSCE) released by the West African Examinations Council (WAEC) and the National Examinations Council (NECO) for the period of ten years from 2011 to 2020 in English Language, General Mathematics and Biology were extracted from WAEC and NECO gazettes for the study. Multivariate Analysis of Variance (MANOVA), the test procedures for analysis are Wilks' Lambda, Pillai Bartlett, Hotelling T2 and Roy's Characteristic Root and each of these four procedures involves the solution of an equation to obtain the characteristic roots of the sum of squares and sum squares cross products for the appropriate source of variation. The findings revealed that there are significant differences in the academic performance of the students in the three geopolitical zones of the state and also the examination bodies (WACE and NECO). It has been concluded that the academic performance of students in national examinations affects the zones, and also the academic performance of students in NECO is better than that of WACE.

Keywords: WAEC, NECO, MANOVA, ACADEMIC PERFORMANCE, EXAMINATION, ENGLISH, MATHEATICS AND BIOLOGY

Introduction

Education is the total process of human learning by which knowledge is acquired, faculties trained and skills developed. Secondary schools not only occupy a vantage in the educational system in Nigeria; it is also the link between the primary and the tertiary levels of education. According to Asikhai (2010), education at the secondary school level is supposed to be the bedrock and the foundation towards higher knowledge in tertiary institutions. It is an investment as well as an instrument that can be used to achieve a more rapid economic, social, political, technological, scientific and cultural development in a country. It is rather unfortunate that secondary schools today are not measuring up to the standards expected of them, as anticipated in their performance in external examinations. There have been public outcries over the persistently poor performance of secondary school students in public examinations. According to Nwokocha and Amadike (2005), the academic performance of students is the yardstick for testing the educational ability (prowess) of a nation. Hence, it is inevitable to maintain high performance in internal and mostly external examinations.

Wikipedia (2013) defines academic performance is the outcome of education, that is, the extent to which a student, teacher or institution has achieved their educational goals. Thus, performance is characterized by performance on tests associated with coursework and the performance of students on other types of examinations Japo (2014).

According to Rofikul and Zebun (2017) education is the most important weapon to bring changes in the society by removing orthodoxy and superstitions, and making people wise and rational. Education is the prime equipment to make the people of a state or country skilled and civilised, and leads the development of a nation through the individual development of its citizens. Without educated citizen no country can make progress in Science and Technology which are the prime requisites for the development of a nation.

Objectives of the study

This paper has the following objectives: - To

i. Check whether there is any significant difference between the zones
ii. know whether there is any significant difference between the examination's bodies

Literature Review

Mildin and Rubio (2021) carried out multivariate analysis on the performance of students in statistics, self-efficacy and attitudes of senior high school. The study made use of a five-point Likert scale. Likert-scale data are analysed at the interval measurement scale and are created by calculating a composite score (sum or mean) from four or more Likert-type items; therefore, the composite scores are analysed at the interval measurement scale. Descriptive statistics recommended for interval scale items include the mean for central tendency and standard deviations for variability. The mean and standard deviation were utilised to determine the level of self-efficacy beliefs, attitudes toward statistics, and performance in statistics of senior high school students. The Box's M and Levene's test were used to determine the normality and homogeneity of the data before testing the significant difference. Multivariate Analysis of Variance (MANOVA) was used to determine whether multiple levels of independent variables on their own or in combination with one another, influence the dependent variables. The findings revealed that among the demographic factors, only the type of school has a significant difference to the self-efficacy beliefs, attitudes towards Statistics, and performance of senior high students in Statistics. Okeke *et al.* (2018) studied multivariate analysis of variance of university students' academic performance using multivariate analysis of variance (MANOVA) and descriptive statistics. The data for the analysis were first tested for normality and equality of variance to see if they are suitable for an analysis of variance test. The Doormik-Hansen test gave a p-value that is greater than 0.05, which indicates that the data were approximately normal. The Box M test of equality of covariance matrices gave a p-value that is greater than 0.05, which indicated that all the covariance matrices are equal across the groups. The Wilks Lambda test shows a p-value of 0.306, which is greater than the level of significance (α= 0.05). The result indicated that the performance of 300-level students of Federal University Wukari does not significantly depend on the faculty in which they belong.

Endris (2016) examined multivariate analysis of factors influencing academic achievements of grade 10 students at Hawassa City, Ethiopia. The secondary data on students' Ethiopian General Secondary Education Certificate Examination (EGSECE) scores were obtained from the Education Department as achievements of students in the five selected subjects: Mathematics, Biology, Physics, Chemistry and English. Descriptive analysis, factor analysis and multivariate multiple linear regression analyses were used to analyse the data. From the results, both governmental and non-governmental school students achieved the poorest in physics and best in English. However, on average, non-governmental school students' achievements were better than those of governmental school students.

Udokang and Odeyemi (2021) examined the Multivariate Analysis of Variance of the Effect of Extra Lesson on Secondary School Students' Academic Performance in English Language and Mathematics in Kwara State. The outcome of MANOVA showed that type and duration of extra lessons contributed significantly to the performance of students in English Language and Mathematics combined together but their interaction was not significant. Further study using Analysis of Variance (ANOVA) revealed that type and duration of extra lessons have a significant effect on students' academic performance in English Language as well as Mathematics.

Materials and Method

It describes the research design, population of study, sample and sampling technique, instrument used for data collection, method of data collection and method of data analysis.

Design of the study

The study employed multivariate analysis to determine the performance of students in senior secondary school by schools and zones; it will also check the degree of relationship between the independent variable (schools and zones) and the dependent variable (students' academic performance) in English Language, General Mathematics and Biology at Senior Secondary School in WAEC and NECO examinations.

Data collection and data source

Data collected for analysis include Student performance based on grade and converted to percentage in English Language, General Mathematics and Biology of students who sat for National examinations, which are WAEC and NECO.

The data for this study is secondary data collected and extracted from gazettes of the West African Examinations Council (WAEC) and National Examinations Council (NECO) of internal examination May/June and June/July respectively, from the Test and Measurement Unit, Ministry of Education, Niger State, from 2011 to 2020. Data were extracted for only six public schools, two from each geo-political zone, that is, in each zone, two local government areas of the state were selected. Niger State has 25 Local Government Areas (LGA) grouped into 3 zones: A, B and C which zone A having 8 LGA, zone B having 9 LGA and zone C having 8 LGA respectively. In each of the selected public secondary schools three (3) subjects were choosing that is English Language, Mathematics and Biology were chosen for the study.

A simple random sample is a randomly selected subset of a population. In this sampling method, each member of the population has an exactly equal chance of being selected.

Method of Analysis

Multivariate analysis of variance

Multivariate analysis of variance (MANOVA) is an extension of univeriate analysis of variance (ANOVA). In ANOVA, differences among various group means on a single-response variable are studied. In MANOVA, the number of response variables is increased to two or more. The hypothesis concerns a comparison of vectors of group means. When only two groups are being compared, the results are identical to Hotelling's T^2 procedure. The multivariate extension of the F-test is not completely direct. Instead, several test statistics are available, such as Wilks' Lambda and Lawley's trace. The actual distributions of these statistics are difficult to calculate, so we rely on approximations based on the F-distribution.

The use of MANOVA is adopted for this study because of its advantages over univariate Analysis of Variance (ANOVA) with only one response variable. The analysis could as well use performance in English Language as a response variable, performance in Mathematics as a separate response variable and performance in Biology as a separate response variable as a separate analysis under ANOVA. The calculations in MANOVA are based on matrix approach because there are more than one response variables which are students' academic performance in English Language, Mathematics and Biology. The independent variables are schools, zones, examinational bodies (WAEC and NECO) and year. The total sum of squares of and cross product matrix (SSCPT) is partitioned in terms of between sum-of-squares and cross-products matrix (SSCPB) and a within sum-of-squares and cross-products matrix (SSCPW) with the equation below

Denoting the k^th^ observation at level ***i*** of zones and level ***j*** of Exams by ***Xijk***, the univariate two-way model is

$$X_{ijk} = \mu + \alpha_i + \beta_j + (\propto \beta)_{ij} + e_{ijk} \quad (3.1)$$

where

i = 1, 2, . . ., r, j = 1, 2, . . ., c and k = 1, 2, . . ., n

α_i is the zones and r is the number of the zones in Niger state

β_j is the Examination bodies and c is number of Examination bodies selected for the study

where

$\sum_{i=1}^{r} \alpha_i = \sum_{j=1}^{c} \beta_j = \sum_{i=1}^{r} \sum_{j=1}^{c} (\propto \beta)_{ij} = 0$ and e_{ijk} are independent N(0, σ^2) random variables. Here, μ represents an overall level, α_i represents the fixed effects of factor 1 (α), β_j represents the fixed effect of factor 2 (β), and $(\alpha\beta)_{ij}$ is the interaction between zones and examination bodies the expected response at the i^{th} level of zones and the j^{th} level of examination bodies is thus

$$E(X_{ijk}) = \mu + \alpha_i + \beta_j + (\propto \beta)_{ij} \quad (3.2)$$

i = 1, 2, . . ., *r and j = 1, 2, . . c*

The presence of interaction λ_{ij}, implies the factor effects are not additive

$$X_{ijk} = \bar{X}_{...} + (\bar{X}_{i..} - \bar{X}_{...}) + (\bar{X}_{.j.} - \bar{X}_{...}) + (\bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X}_{...}) + (\bar{X}_{ijk} - \bar{X}_{ij.}) \quad (3.3)$$

Where X_{ijk} is the total observation of the academic performance of students in the zones and examination bodies, $\bar{X}_{...}$ is the overall mean academic performance of students in the zones and examination bodies, $\bar{X}_{i..}$ is the mean of academic performance of students in the zones for the i^{th} level $\bar{X}_{.j.}$ is the mean of academic performance of students in the examination bodies for the j^{th} level and $\bar{X}_{ij.}$ is the mean of the academic performance of students in the zones and examination bodies for the i^{th} level and j^{th} level

After using the matrix approach in computations to get F-ratio and the p – value which in turn are used to decide on which of the following hypotheses to accept.

Hypothesis

The paper will investigate the following hypotheses

Hypothesis 1

H_{01}: There is no significant difference between the academic performance of students and zones Vs

H_{11}: There is a significant difference between the academic performance of student's zones

$H_0: \alpha_1 = \alpha_2 = \alpha_3$ Vs $H_1: \alpha_1 \neq \alpha_2 \neq \alpha_3$

Hypothesis 2

H_{02}: There is no significant difference between academic performance of students and examination bodies

Vs

H_{12}: There is significant difference between academic performance of students and examination bodies

$H_0: \beta_1 = \beta_2$ Vs $H_1: \beta_1 \neq \beta_2$

Sum of squares cross product

$$SS_{zones} = cn\sum_{i=1}^{r}(X_{i..} - \bar{X}_{...})^2 = cn\sum_{i}^{r}(X_{i..} - \bar{X}_{...})(X_{i..} - \bar{X}_{...})' \quad (3.4)$$

$$SS_{Exams\ bodies} = rn\sum_{j=1}^{c}(X_{.j.} - \bar{X}_{...})^2 = rn\sum_{j}^{c}(X_{.j.} - \bar{X}_{...})(X_{.j.} - \bar{X}_{...})' \quad (3.5)$$

$$SS_{int} = n\sum_{i=1}^{r}\sum_{j=1}^{c}(X_{ij.} - X_{i..} - \bar{X}_{.j.} - \bar{X}_{...})^2 = n\sum_{i=1}^{r}\sum_{j=1}^{c}(X_{ij.} - X_{i..} - \bar{X}_{.j.} - \bar{X}_{...})(X_{ij.} - X_{i..} - \bar{X}_{.j.} - \bar{X}_{...})' \quad (3.6)$$

$$SS_{res} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(X_{ijk} - \bar{X}_{ij.})^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(X_{ijk} - \bar{X}_{ij.})(X_{ijk} - \bar{X}_{ij.})' \quad (3.7)$$

$$SS_{Total} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(X_{ijk} - \bar{X}_{...})^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(X_{ijk} - \bar{X}_{...})(X_{ijk} - \bar{X}_{...})' \quad (3.8)$$

Table 3.1: MANOVA Table for comparing effects of two factors and their interaction

Source of variation (SV)	**Sum of squares (SS)**	**Degrees of freedom (df)**
Zones	$SS_{zones} = cn\sum_{i=1}^{r}(X_{i..} - \bar{X}_{...})(X_{i..} - \bar{X}_{...})'$	
Exams Bodies	$SS_{Exams} = rn\sum_{j=1}^{c}(X_{.j.} - \bar{X}_{...})(X_{.j.} - \bar{X}_{...})'$	c – 1

Interaction	$SS_{int} = n\sum_{i=1}^{r}\sum_{j=1}^{c}(X_{ij.} - X_{i..} - \bar{X}_{.j.} - \bar{X}_{...})(X_{ij.} - X_{i..} - \bar{X}_{.j.} - \bar{X}_{...})'$	(r – 1)(c – 1)
Residual (error)	$SS_{res} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(X_{ijk} - \bar{X}_{ij.})(X_{ijk} - \bar{X}_{ij.})'$	rc(n – 1)
Total (corrected)	$SS_{cor} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(X_{ijk} - \bar{X}_{...})(X_{ijk} - \bar{X}_{...})'$	rnc – 1

Source: Johnson and Wichern, (1944)

The F-ratio of the mean squares, $SS_{zones}/(r-1)$, $SS_{Exams}/(c-1)$ and $SS_{int}/(r-1)\,(c-1)$ to the mean square, $SS_{res}/[rc(n-1)]$ can be used to test for the effects of zones, exams and zones – exams interaction respectively.

Test statistic for Multivariate Analysis of Variance (MANOVA) and hypotheses

There are commonly four test statistics used in multivariate analysis of variance, these statistics are detail below.

1. **Wilks' lambda (∧)**: - The product of the unexplained variance on each of the variates.

   $\wedge = \prod_{i=1}^{n}\frac{1}{1+\lambda_i}$ (1.9)

   i =1, 2, …, n

   The symbol is similar to the summation symbol (∑) except that it means multiply rather than add up. This is the most common traditional test when there are more than two groups formed by the independent variables. It is a measure of the differences between groups of the centriod (vector) of means on the independent variables. The smaller the lambda (∧) the greater the differences, the Bartlett's transformation of lambda (∧) is then used to compute the significance of the lambda. Lambda. It represents the ration of error variance to total variance $\frac{SS_E}{SS_T}$ for each variate. Or

   Wilks Lambda

   $\wedge^* = \frac{|E|}{|H+E|}$ (1.10)

   Where E is the error sum of square and H is the hypothesis, and then H + E is the total sum of square

   Here, the determinant of the error sums of squares and the cross-products matrix E is divided by the determinant of the total sum of squares and cross-product matrix T = H + E. If H is large relative to E, then $|H+E|$ will be large relative to $|E|$. Thus, we reject the null hypothesis if Wilks lambda is small (close to zero).

2. **Hotelling-Lawley Trace**: The Hotelling-Lawley trace is the sum of eigenvalues for each variate and computed by the equation $= \sum_{i=1}^{n}\lambda_i$. This is the most common traditional test for two independent groups. This statistic represents the proportion of explained variance to unexplained variance $\left(\frac{SSCP_H}{SSCP_T}\right)$for each of the variates, and so it compares directly to the F ratio in ANOVA. Or

   Hotelling-Lawley Trace

   $T_0^2 = trace(HE^{-1})$ (1.11)

Here, H is multiplying the inverse of E; then it takes the trace of the resulting matrix. If H is large relative to E, then the Hotelling-Lawley trace will take a large value. Thus, we will reject the null hypothesis if this test statistic is large.

3. **Pallai – Bartlett trace**: - Pillai trace expressed by the equation $V = \sum_{i=1}^{n} \frac{\lambda_i}{1+ \lambda_i}$ In which represents the eigenvalues for each of the discriminant variates, and n represents the number of variates. Pillai's trace is the sum of explained variances on the discriminant variants, which are the variables which are computed based on the canonical coefficients for a given set of roots, a large value, by convention, indicates a significant difference and is similar to the ratio of explained variance to total variance.
   Similarly, represented **as**

   $$\textbf{Pallai Trace } \boldsymbol{V = trace(Hse(H + E)^{-1})} \quad (1.12)$$

   Here, H is multiplying by the inverse of the total sum of squares and cross-products matrix $T = H + E$. If H is large relative to E, then the Pillai trace will take a large value. Thus, we will reject the null hypothesis if this test statistic is large.
4. **Roy's greatest characteristic roots**: - Roy's largest root is the Eigen value for the first variate. In a sense, according to Michael (2010), it is the same as the Hotelling-Lawley trace, except for the first variate only. This statistic represents the proportion of explained variance to unexplained vary $\left(\frac{SSCP_H}{SSCP_E}\right)$ for the first discriminant function. This is similar to the Pillai Bartlett trace, but based only on the first (and hence most important) root. Roy's largest root sometimes equated with the largest Eigen value. This value is conceptually the same as the F- ratio in univariate ANOVA and represents the maximum possible between-group difference given the data collected. The test statistic is less robust than the other tests in the face of violations of the assumption of multivariate normality.
   or
   Roy's Maximum Root: Largest eigenvalue of HE^{-1}
   Here, H is multiplied by the inverse of E, and then the largest eigenvalue of the resulting matrix is computed. If H is large relative to E, then Roy's root will take a large value. Thus, the null hypothesis is rejected if this test statistic is large.

Table 3.1 below gives details of the F approximations for the four test statistics commonly used in multivariate analysis of variance

Table 3.2: Test Statistics used to compare sample mean vectors with Approximate F-test

Test	Statistic	F	*df₁*	*df₂*	Comment
Pillai's Trace	V	$(n-m-p+s)V\big(d(s-V)\big)^{-1}$	*Sd*	S(n-m-p+s)	$d = \max(p, m-1)$ $s = \min(p, m-1)$ = number of positive eigen values
Wilks's Lambda	^	$\left\{\left(1-\wedge^{\frac{1}{t}}\right)\wedge^{\frac{1}{t}}(df_2 \times df_1^{-1})\right\}$	P(m -1)	wt − (df_1x0.5) + 1	$w = n-1(p+m)0.5$ $t = [(df_1^2-4)\times(p^2+(m-1)^2-5)^{-1}]^{0.5}$ If *df_1 = 2,* set t = 1
Hotelling's Trace	U	$df_2U \times (sdf_1)^{-1}$	S(2A+s+1)	2(sB + 1)	*s is as for pillai's trace* $A = (\lvert m-p-1\rvert - 1)\times 0.5$ $B = (n-m-p-1)\times 0.5$ The significance level obtained is a lower bound
Roy's Largest Root	λ_1	$(df_2 \times df_1^{-1})\lambda_1$	D	n − m − d − 1	$d = \max(p, m-1)$ $s = \min(p, m-1)$ = number of positive eigen values

Source: Manly (2005

Data Analysis
The multivariate analysis of variance is the statistical method use for the analysis. Due to the complicated nature of the data, a statistical software package known as SPSS version 21 was used for the analysis.

The matrices of the appropriate sum of squares and cross-products were calculated (see the SPSS Statistical software output in Appendix I) leading to the matrices below:

$$SSCP_{ZONES} = \begin{pmatrix} 25664.227 & 35605.209 & 24189.398 \\ 35605.209 & 118362.984 & 47232.619 \\ 24189.398 & 47232.619 & 25510.287 \end{pmatrix}$$

The sum of squares and cross product for the zones ($SSCP_{ZONES}$) in Niger State

$$SSCP_{EXAMS} = \begin{pmatrix} 17285.444 & -9231.023 & 10301.014 \\ -9231.023 & 4929.685 & -5501.097 \\ 10301.014 & -5501.097 & 6138.743 \end{pmatrix}$$

The sum of squares and cross product for the Examination bodies ($SSCP_{EXAMS}$)

$$SSCP_{Error} = \begin{pmatrix} 2122819 & 885393.114 & 844003.525 \\ 885393.114 & 2780950.653 & 921317.664 \\ 844003.525 & 921317.664 & 1965926.210 \end{pmatrix}$$

The sum of squares error and cross product for the zones and examinations ($SSCP_{ERROR}$)

Calculation of the MANOVA test statistics

i. **Zones:** the test statistics for the zones (academic performance base on zones) are as computed below.

(a) Wilks's lambda

$$\wedge = \frac{|E|}{|H+E|}$$

$$|E| = \begin{vmatrix} 2122819.665 & 885393.114 & 844003.525 \\ 885393.114 & 2780950.653 & 921317.664 \\ 844003.525 & 921317.664 & 1965926.210 \end{vmatrix}$$

$$2122819.665 \times \begin{vmatrix} 2780950.653 & 921317.664 \\ 921317.664 & 1965926.210 \end{vmatrix} - 885393.114 \times \begin{vmatrix} 885393.114 & 921317.664 \\ 844003.525 & 1965926.210 \end{vmatrix} + 844003.525 \times \begin{vmatrix} 885393.114 & 2780950.653 \\ 844003.525 & 921317.664 \end{vmatrix}$$

$$2122819.665 \times (5.4671 \times 10^{12} - 8.4883 \times 10^{11}) - 885393.114 \times (1.7406 \times 10^{12} - 7.7760 \times 10^{11}) + 844003.525 \times (8.1573 \times 10^{11} - 2.3471 \times 10^{12})$$

$$9.8038 \times 10^{18} - 8.5263 \times 10^{17} - 1.2925 \times 10^{18}$$

$$|E| = 7.6587 \times 10^{18}$$

$$|H+E| = \begin{vmatrix} 2148483.892 & 920998.323 & 868192.923 \\ 920998.323 & 2899313.637 & 968550.283 \\ 868192.923 & 968550.283 & 1991436.497 \end{vmatrix}$$

$$2148483.892 \times \begin{vmatrix} 2899313.637 & 968550.283 \\ 968550.283 & 1991436.497 \end{vmatrix} - 920998.323 \times \begin{vmatrix} 920998.323 & 968550.283 \\ 868192.923 & 1991436.497 \end{vmatrix} + 868192.923 \times \begin{vmatrix} 920998.323 & 2899313.637 \\ 868192.923 & 968550.283 \end{vmatrix}$$

$$2148483.892 \times (5.7738 \times 10^{12} - 9.3809 \times 10^{11}) - 920998.323 \times (1.8341 \times 10^{12} - 8.4089 \times 10^{11}) + 868192.923 \times (8.9203 \times 10^{11} - 2.5172 \times 10^{12})$$

$$1.0389 \times 10^{19} - 9.1474 \times 10^{17} - 1.4110 \times 10^{18}$$

$$|H+E| = 8.0638 \times 10^{18}$$

$$\wedge = \frac{|E|}{|H+E|} = \frac{7.6587 \times 10^{18}}{8.0638 \times 10^{18}}$$

$$= 0.950$$

b) **Pallai Trace**

$$\boldsymbol{V = trace(H(H+E)^{-1})}$$

$$trace\left(\begin{pmatrix} 25664.227 & 35605.209 & 24189.398 \\ 35605.209 & 118362.984 & 47232.619 \\ 24189.398 & 47232.619 & 25510.287 \end{pmatrix}\begin{pmatrix} 2148483.892 & 920998.323 & 868192.923 \\ 920998.323 & 2899313.637 & 968550.283 \\ 868192.923 & 968550.283 & 1991436.497 \end{pmatrix}^{-1}\right)$$

$$trace\begin{pmatrix} 0.0061 & 0.0086 & 0.0053 \\ 0.0104 & 0.0400 & 0.0225 \\ 0.0035 & 0.0136 & 0.0046 \end{pmatrix}$$

$$V = 0.0061 + 0.0400 + 0.0046$$

$=0.051$

(c) **Hotelling Lawley Trace**

$$\cup = HE^{-1}$$

$$\begin{pmatrix} 25664.227 & 35605.209 & 24189.398 \\ 35605.209 & 118362.984 & 47232.619 \\ 24189.398 & 47232.619 & 25510.287 \end{pmatrix} \times \begin{pmatrix} 2122819.665 & 885393.114 & 844003.525 \\ 885393.114 & 2780950.653 & 921317.664 \\ 844003.525 & 921317.664 & 1965926.210 \end{pmatrix}^{-1}$$

$$= trace\begin{pmatrix} 0.0062 & 0.0090 & 0.0054 \\ 0.0105 & 0.0416 & 0.0226 \\ 0.0035 & 0.0143 & 0.0048 \end{pmatrix}$$

$$\cup = 0.0062 + 0.0416 + 0.0048$$

$$= 0.052$$

(d) **Roy's Characteristic Root**
Largest eigenvalue of HE^{-1}

$\lambda_{max} = 0.052$

The computed values are the same as the values obtained from the SPSS output in Appendix II (see Zones (A, B and C))

ii. Examination Bodies: - The test statistics for the examination bodies (WAEC and NECO) are as computed below.

(a) Wilks's lambda

$\wedge = \frac{|E|}{|H+E|}$

$$|E| = \begin{vmatrix} 2122819.665 & 885393.114 & 844003.525 \\ 885393.114 & 2780950.653 & 921317.664 \\ 844003.525 & 921317.664 & 1965926.210 \end{vmatrix}$$

$$2122819.665 \times \begin{vmatrix} 2780950.653 & 921317.664 \\ 921317.664 & 1965926.210 \end{vmatrix} - 885393.114 \times \begin{vmatrix} 885393.114 & 921317.664 \\ 844003.525 & 1965926.210 \end{vmatrix} + 844003.525 \times \begin{vmatrix} 885393.114 & 2780950.653 \\ 844003.525 & 921317.664 \end{vmatrix}$$

$$2122819.665 \times (5.4671 \times 10^{12} - 8.4883 \times 10^{11}) - 885393.114 \times (1.7406 \times 10^{12} - 7.7760 \times 10^{11}) + 844003.525 \times (8.1573 \times 10^{11} - 2.3471 \times 10^{12})$$

$$9.8038 \times 10^{18} - 8.5263 \times 10^{17} - 1.2925 \times 10^{18}$$

$$|E| = 7.6587 \times 10^{18}$$

$$|H+E| = \begin{vmatrix} 17285.444 & -9231.023 & 10301.014 \\ -9231.023 & 4929.685 & -5501.097 \\ 10301.014 & -5501.097 & 6138.743 \end{vmatrix} + \begin{vmatrix} 2122819.665 & 885393.114 & 844003.525 \\ 885393.114 & 2780950.653 & 921317.664 \\ 844003.525 & 921317.664 & 1965926.210 \end{vmatrix}$$

$$= \begin{vmatrix} 2140105.109 & 876162.091 & 854304.539 \\ 876162.091 & 2785880.338 & 915816.567 \\ 854304.539 & 915816.567 & 1972064.953 \end{vmatrix}$$

$$2140105.109 \times \begin{vmatrix} 2785880.338 & 915816.567 \\ 915816.567 & 1972064.953 \end{vmatrix} - 876162.091 \times \begin{vmatrix} 876162.091 & 915816.567 \\ 854304.539 & 1972064.953 \end{vmatrix} + 854304.539 \times \begin{vmatrix} 876162.091 & 2785880.338 \\ 854304.539 & 915816.567 \end{vmatrix}$$

$$2\quad 140105.109 \times (5.4939 \times 10^{12} - 8.3872 \times 10^{11}) - 876162.091 \times (1.7278 \times 10^{12} - 7.8239 \times 10^{12}) + 854304.539 \times (8.0240 \times 10^{11} - 2.380 \times 10^{12})$$

$$|H+E| = 7.7865 \times 10^{18}$$

$$\wedge = \frac{|E|}{|H+E|} = \frac{7.6587 \times 10^{18}}{7.7865 \times 10^{18}}$$

$$= 0.984$$

(b) P**allai Trace**

$$V = trace(H(H+E)^{-1})$$

$$trace\left(\begin{pmatrix} 17285.444 & -9231.023 & 10301.014 \\ -9231.023 & 4929.685 & -5501.097 \\ 10301.014 & -5501.097 & 6138.743 \end{pmatrix}\begin{pmatrix} 2140105.109 & 876162.091 & 854304.539 \\ 876162.091 & 2785880.338 & 915816567 \\ 854304.539 & 915816.567 & 1972064.953 \end{pmatrix}^{-1}\right)$$

$$trace\begin{pmatrix} 0.0094 & -0.0078 & 0.0048 \\ -0.0050 & 0.0042 & -0.0026 \\ 0.0056 & -0.00\,47 & 0.0029 \end{pmatrix}$$

$V = 0.0094 + 0.0042 + 0.0029$
=0.016

(c) **Hotelling Lawtey Trace**

$\cup = HE^{-1}$

$$\begin{pmatrix} 17285.444 & -9231.023 & 10301.014 \\ -9231.023 & 4929.685 & -5501.097 \\ 10301.014 & -5501.097 & 6138.743 \end{pmatrix} \times \begin{pmatrix} 2122819.665 & 885393.114 & 844003.525 \\ 885393.114 & 2780950.653 & 921317.664 \\ 844003.525 & 921317.664 & 1965926.210 \end{pmatrix}^{-1}$$

$$= trace\begin{pmatrix} 0.0095 & -0.0080 & 0.0049 \\ -0.0051 & 0.0043 & -0.0026 \\ 0.0057 & -0.0047 & 0.0029 \end{pmatrix}$$

∪= 0.017

(d) **Roy's Characteristic Root**
Largest eigenvalue of HE^{-1}

$\lambda_{max} = 0.017$

The computed values are the same with the values obtained from the SPSS out in Appendix II (see Examination bodies (WAEC and NECO))

Calculation of F – Statistics for the test Statistics

The statistics values obtained above is used to calculate the F- values for the four multivariate procedures.

1. Academic performance base on Zones (A B C)
   a) Wilks's Lambda (∧)
      Where n = 10286; m = 3, p = 3, d = 3, s = 2 and ∧ = 0.950
      $$F = \left(\frac{1-\wedge}{\wedge}\right)^{\frac{1}{t}}\left(\frac{df_2}{df_1}\right) = \left(\frac{1-0.950}{0.950}\right)^{\frac{1}{2}} \times \left(\frac{20562}{6}\right) = 89.102$$
   b) Pillai's trace (V)
      $$F = \frac{(n-m-p+s)V}{d(s-V)} = \frac{(10286-3-3+2)0.051}{3(2-0.051)} = 89.684$$
   c) Hotelling Lawtey Trace
      $$F = \frac{df_2 \cup}{Sdf_1} = \frac{20560 \times 0.052}{2 \times 6} = 89.093$$
   d) Roy's Characteristics root
      $$F = \left(\frac{df_2}{df_1}\right)\lambda_{max} = \frac{10279}{3} \times 0.043 = 147.332$$

The decision rule is to reject the null hypothesis H_{01} if the computed F is greater than the tabulated $F_{3,\ 10286,\ 0.05} = 2.61$ at 5% level of significant. The calculated values of the test are 89.102, 89.684, 89.093 and 147.332 respectively are greater than the tabulated value. Therefore, the null hypothesis is rejected and concluded that the academic performance of the zones (A B and C) are not the same.

2. Academic performance base on Examination Bodies (WAEC and NECO)
   a) Wilkes's Lambda (∧)
   Where n = 10286; m = 3, p = 2, d = 3, s = 2 and ∧ = 0.984

$$F = \left(\frac{1-\wedge}{\wedge}\right)^{\frac{1}{t}}\left(\frac{df_2}{df_1}\right) = \left(\frac{1-0.984}{0.984}\right)^{\frac{1}{2}} \times \left(\frac{20564}{4}\right) = 41.627$$

   b) Pillai's trace (V)

$$F = \frac{(n-m-p+s)V}{d(s-V)} = \frac{(10286-3-2+2)0.016}{3(2-0.016)} = 27.632$$

   c) Hotelling Lawtey Trace

$$F = \frac{df_2 \cup}{Sdf_1} = \frac{20562 \times 0.017}{2 \times 4} = 43.694$$

   d) Roy's Characteristics root

$$F = \left(\frac{df_2}{df_1}\right)\lambda_{max} = \frac{10280}{2} \times 0.017 = 87.380$$

The decision rule is to reject the null hypothesis H_{02} if the computed F is greater than the tabulated $F_{2,\ 10286,\ 0.05} = 3.00$ at 5% level of significance. The calculated values of the test 41.627, 27.632, 43.694 and 87.380, respectively, are greater than the tabulated value. Therefore, the null hypothesis is rejected and concluded that the academic performance of the examination bodies (WAEC and NECO) is not the same.

Conclusion

The results of the academic performance in zones and examination bodies considered for this study were statistically significant in both the zones and the examination bodies. It has been concluded zones do not affect the academic performance of students in the state. So also, the examination bodies do not affect the academic performance of students either.

Based on the findings in the study that there are significant differences in zones, LGAs/schools, examination bodies and years, as well as significant interactions between them in students' academic performance in (English Language, Mathematics and Biology) several major contributions can be outlined:

Policy implications: - policymakers can use these findings to design targeted interventions that consider the unique combinations of zones, LGAs/school, examination bodies, and years. Policies could be tailored to address specific needs and leverage strengths identified in the interactions between these factors.

Resource allocation: - Identifying significant differences between zones and LGAs/schools can inform better allocation of resources, ensuring that schools with fewer resources or lower performance receive additional support. This targeted resource allocation can help in bridging the gap between different zones and LGAs/schools and promoting equity in education.

Educational Research and Development: - The findings can guide further research into the specific causes of differences in performance, leading to the development of more effective educational practices and interventions. Continuous research based on these findings can contribute to the progressing understanding of educational dynamics and improve overall educational quality.

Reference

Bryan F. J. Manly (2004). Multivariate Statistical Methods. A Primer Third Edition. Florida 33431: Chapman and Hall/CRC

C. Chatfield A. J. Collins (1984). Introduction to Multivariate Analysis first Edition. London: Chapman & Hall

Usman A. (2014). Analysis of ChlorophyII "A" and "B" in Randomly selected Varieties of Surgarcane Leaves Using Multivariate Analysis of Variance. *Journal of Mathematics* 10(2) pp 26 – 32 www.iosrjournals.org e-ISSN: 2278-3008, p- ISSN:2319-7676.

Anthony E. U & Ekemini G. (2020) Multivariate Analysis of Academic Performance of Undergraduate Students Admitted Through Unified Tertiary Matriculation Examination and Direct Entry. *Journal of Mathematics* 7(2) pp 1 – 10 www.researchjournali.com

Adeleke B. L., Yahaya W. B. & Usman A. (2015). A comparison of Some Test Statistics for Multivariate Analysis of Variance Model With Non – Normal Responses. *Journal of Natural Science Research.* 5(15) pp 1 – 9 www.iiste.org p-ISSN 2224-3186 e-ISSN 2225-0921

Udokang A. E. & Odeyemi J. B. (2021) Multivariate Analysis of Variance of Effect of Extra Lesson on Secondary School Students' Academic Performance in English Language and Mathematics in Kwara State *International Journal of Engineering and Applied Sciences (IJEAS)*, 8(2) pp 7-12 www.ijeas.org

Paul M., Peter M., Emmanuel G. & Baah-Korang K. (2015) Multivariate Analysis of Academic Performance of Ghanaian Secondary School Students *International Invention Journal of Education and General Studies* 1(1) pp 9-17. http://internationalinventjournals.org/IIJEGS

Anthony E. U & Ekemini G. (2020) Multivariate Analysis of Academic Performance of Undergraduate Students Admitted Through Unified Tertiary Matriculation Examination and Direct Entry. *Researchjournali's Journal of Mathematics* 7(2) pp 1 – 10 www.researchjournali.com

Bashir M. I. & Luka K. M. (2022) Comparative Analysis of Students' Performance in WAEC and NECO English Language and General Mathematics Examinations of Senior Secondary Schools in Kano State, Nigeria AJSTME, 8(1), pp 480-487, https://www.ajstme.com.ng ISSN: 2251-0141