

An Application of Categorical Analysis of Variance in Nested Arrangements

Iwundu M. P.^{*}, Anyanwu C. O.

Department of Mathematics and Statistics, University of Port Harcourt, Nigeria

Abstract Categorical data, having no underlying metric, are frequently encountered in experimental situations. Analysis of such data deviates from the usual methods involving continuous-type data. In this paper, the significance of main factor (Geo-political Zones) and sub-factor (States) are studied using categorical data in nested arrangements. The CATANOVA technique used is suitable for mixed design, having some factors crossed and others nested. Consideration is made for frequency data involving National Examination Council (NECO) and West African Examination Council (WAEC), from 2003 to 2010. Numerical results show that the main factor, Geo-political Zone and the sub-factor, States, are significant having p-value of less than 0.00001 in each case.

Keywords Categorical data, Categorical Analysis of Variance, Crossed factor, Nested factor, Main effects, Sub-factor effects

1. Introduction

According to Agresti (2007), categorical data analysis deals with data that classify an observation into one or more categories. Two main types of categorical variables involves are variables whose data are on nominal or ordinal scale of measurement. By nominal scale, we refer to a discrete classification of data, in which data are neither measured nor ordered but subjects are merely allocated to distinct categories. Nominal scales are used merely for labelling variables, without any quantitative value, to identify or classify the objects. Quantitative data in the case of nominal categorical variables refer to discrete data otherwise called frequency data. Nominal variable yields unordered qualitative variable and possesses no inherent ordering or nominal distance between categories. In other words, the order of listing the categories is not important. According to Weiss (2009), nominal data are inherently less informative than quantitative data, it does not convey information about the magnitude of differences and it is not also clear on how to deal with missing data. Categorical variables having ordered scales are ordinal variables. Ordinal variables give rise to natural ordering. It is common sometimes to use numerical values to denote the ordering information in an ordinal variable. However, numerical variables corresponding to

categories of ordinal variables only show the ranking order in a particular attribute.

A number of researchers have displayed categorical variable in a 2 x 2 contingency table or by cross classification. Cross-classified data are viewed as when subjects are classified by two or more types of clusters, but the clusters are not hierarchical or nested within one another. As is commonly seen, many data sets come in the form of counts of individuals, classified into one of a finite set of disjoint categories. The data consist of the numbers of observations falling into each category. Thus, two factors are crossed when every category of one factor occurs in the design with every category of the other factor. In other words, there is at least one observation in every combination of categories for the two factors. On the other hand, a factor is nested within another factor when each category of the first factor co-occurs with only one category of the other factors. In other words, an observation has to be within one category of say Factor 2 in order to have a specific category of Factor 1.

One of the simplest categorical data analyses involving nominal data in cross classification requires checking whether the categorical variables are independent. Another includes goodness of fits tests, where observed and expected frequencies are compared. Yet another area of interest where categorical data analysis is implored, particularly in medical fields, is in prospective, retrospective and cross-sectional studies, aimed at comparing incidence of diseases. As in any inferential study, estimation of parameters and test of hypothesis are key aspects to categorical analysis. Although most encountered data arrangements make use of cross-classification, in real life situation, some data are in nested arrangement. Hence, multi-factor studies can involve

^{*} Corresponding author:

mary.iwundu@uniport.edu.ng (Iwundu M. P.)

Published online at <http://journal.sapub.org/ijps>

Copyright © 2018 The Author(s). Published by Scientific & Academic Publishing

This work is licensed under the Creative Commons Attribution International

License (CC BY). <http://creativecommons.org/licenses/by/4.0/>

treatment combinations in which some factors are crossed with other factors, and some factors are nested within other factors. The need for nesting of factors arises when an experimental situation requires that unique levels of one factor occur within each level of a second factor.

Researchers are faced with the problem of developing useful methods for analysing categorical data. Although methods for analysing quantitative (continuous) variables are vast, the assumptions are completely different when the data structure is non-quantitative. This puts a big restriction on the use of such techniques for categorical data. Due to the seeming complexity with categorical data, many inferential problems still remain unanswered when considering nominal categorical data. The aim of this paper is to illustrate the analysis of variance of categorical data involving frequency data in nested arrangement, using performance of students examinations conducted by National Examination Council (NECO) and West Africa Examination Council (WAEC). The main factor shall be the Geo-political Zones and the sub-factor shall be the States. The objectives of this study include

1. Assessing the significance of the main factor.
2. Assessing the significance of the sub-factor.

In all cases, investigation shall be gender-based (sex) and Period-based or Duration-base (year of examination).

Many literatures have shown the existence of several techniques for analysing data in contingency table. Some group of researchers deal with categorical data in their original format while a second group of researchers transform categorical data to enable the data be treated with existing methods for quantitative data. For instance, Winsor (1948) introduced logit transformation for factorial analysis of multiple dichotomous data. Cochran (1950) while considering comparison of percentages in matched samples suggested assigning values of 0 or 1 to represent the responses in the two categories. Dyke and Patterson (1952) provided a method for the case in which the data are proportions, applicable whether the numbers of observations are equal or not. The data are subjected to logit transformation such that in the new scale of measurements observations can reasonably be represented as linear functions of a number of parameters. Gart and Zweifel (1967) extended the technique of Dyke and Patterson (1952) by introducing weighting function for a special case of dichotomous data. Light and Margolin (1971) reported the measure of variations for categorical data and developed a general approach of analysis of variance for a one-way table with categorical response.

The general approach to categorical data proposed by Light and Moagolin (1971) started with the computation of the total variation in the data and then partitioning this variation into specific components. The distributions of the various components are derived using a multinomial model. Onukogu (1985a) studied the effect of specification such as random, fixed and mixed models on test for significance of factors in a CATANOVA. A linear additive model for a

two-way classification was proposed, hence establishing a strong link between Analysis of variance (ANOVA) and Categorical Analysis of variance (CATANOVA). Onukogu (1985a) obtained an F-test for main effects as well as estimate of missing responses and adjusted row and column effects for balanced incomplete designs. Transformation such as log-linear and logit were not necessary rather data were analysed using their original format. According to Onukogu (1985b), a two-way ANOVA with quantal responses is equivalent to a three-way contingency table in which one of the classifications is treated as responses to the other two classifications. Knowing that sum of square in ANOVA is viewed as the departure of individual observations from their mean which is not helpful in the case of nominal data since mean is an undefined concept, Onukogu (1985b) reported that one of the hard nut to crack in any analysis of variance of categorical data is the definition and computation of sum of squares. However, he viewed sum of square (SS) of a set of data as the trace of its variance-covariance matrix. Singh (1996) obtained the adjusted sum of squares for rows and columns with restriction to situation where interaction is absent. He used the one moment approximation to derive the asymptotic null distribution of the test statistic as used by Light and Margolin (1971) and Onukogu (1985a,b).

It is believed that categorical outcomes can be analysed using ANOVA technique if the arcsine-square-root logit transformation is applied according to Cochran (1940), Rao (1960), Winer et al (1971). However, Florian (2008) argued that the use of ANOVA for the analysis of categorical data is misleading as with ANOVA over arcsine-square-root transformation, proportions of categorical outcomes can lead to spurious null result and spurious significances which go beyond normal chance of type I and type II error. Singh (2004) studied the validity of the existing CATANOVA method to analyse the nominal data for single and multi-group repeated measure designs and developed a modified test for single and multiple-repeated measure design.

Florian (2008) introduced ordinary logit model, though unfortunately, ordinary logit model do not included random effect modeling. However, Florian described a missed logit model (i.e. a generalized linear model) to address the issue, but these models are considered to be too complicated for analytical evaluations. Fortunately, logit model are built into statistical software such as package R thereby reducing computational rigor. Weiss (2009) presented a method for analyzing nominal responses using fractional design with the goal of allowing the analytic power afforded by fractional designs to be extended to studies in which nominal responses are the natural way to express a person's action or opinions. According to Weiss (2009) nominal responses in the usual sense do not exhibit variance, since there is no metric that can support measure of distance. As in Weiss (2009), nobody can say how much two nominal variable differ, but it is possible to apply the concept of disparity between variable

since two variables will either match or do not match. Onukogu (2014) provided the analytical tools for mixed design when some factors may be crossed while other are nested. In the technique, frequency data are analysed in their original format, without the need for any transformation. The work provides a CATANOVA method for testing for the significance of nested factors. In his study, a breakdown was given of the total observed variation into independent sources such that the effect of each factor can be tested and concluded separately.

2. Methodology

The data for this research work involve m-way contingency table in which one of the classifications may be viewed as response to the others as in Onukogu (1985b). The m-way contingency table is equivalent to an (m-1)-way CATANOVA. Throughout this work, frequency data

involving the performance of students in senior category Examinations conducted by West African Examination council (WAEC) and National Examination Council (NECO) for the period of 2003 to 2010 are employed. The data satisfy the general layout for Categorical data in nested arrangement as presented in Onukogu (2014). Geopolitical Zones and States shall respectively form the main factor and the sub-factor. The performance in the examination shall be viewed as (i) Gender-based and (ii) Duration or Time-based.

The two-way Categorical analysis of variance in a nested arrangement due to Onukogu (2014) is adopted for this research. We assume no loss in generality using the method, for unequal levels of sub-factor that do not differ significantly. However, for computational ease, the average of the sub-factor levels shall be employed. The layout for the two-way CATANOVA nested classification is presented in Figure 1 as follows;

| Main factor A(i) | | | | | | | | | | | |
|------------------|-----------|------|-----------|-----------------|-----------|------|-----------|-----------------|-----------|------|-----------|
| a_1 | | | | a_2 | | | | a_i | | | |
| Sub-factor B(j) | | | | Sub-factor B(j) | | | | Sub-factor B(j) | | | |
| b_1 | b_2 | | b_J | b_1 | b_2 | | b_J | b_1 | b_2 | | b_J |
| n_{111} | n_{121} | | n_{1J1} | n_{211} | n_{221} | | n_{2J1} | n_{i11} | n_{i21} | | n_{iJ1} |
| n_{112} | n_{122} | | n_{1J2} | n_{212} | n_{222} | | n_{2J2} | n_{i12} | n_{i22} | | n_{iJ2} |
| \vdots | | | | | | | | | | | |
| n_{11k} | n_{12k} | | n_{1Jk} | n_{21k} | n_{22k} | | n_{2Jk} | n_{i1k} | n_{i2k} | | n_{iJk} |

Figure 1. Layout of two-way CATANOVA nested classification

In Figure 1, factor A and factor B have $k = 1, 2, \dots, K$ responses per unit where factor A has $i = 1, 2, \dots, I$ levels and factor B has $j = 1, 2, \dots, J$ levels. The cell totals are such that

$$\sum_{ijk} n_{ijk} = n, \sum_{ij} n_{ij} = n, \sum_i n_i = n, \sum_k n_k = n; n_{ik} = \sum_j n_{ijk}, n_{ij} = \sum_k n_{ijk},$$

$$n_k = \sum_{ij} n_{ijk}, P_{ijk} = \frac{n_{ijk}}{n_{ij}}, P_{ik} = \frac{n_{ik}}{n_i}, P_k = \frac{n_k}{n}$$

The statistical model for the two-way nested classification for the Categorical data is the three-component model.

$$P_{ijk} = P_k + (P_{ik} - P_k) + (P_{ijk} - P_{ik}) \quad (1)$$

where

P_k is a constant for the k th response

$P_{ik} - P_k$ is the effect of the i th level of factor A

$P_{ijk} - P_{ik}$ is the effect of the j th level of factor B within the i th level of factor A.

The fundamental assumptions in Categorical analysis of variance are associated with statistical distributions,

statistical independence and constant variance. It is general to assume multinomial probability model given by

$$P(n_{ij1}, n_{ij2}, \dots, n_{ijk}) = \frac{n_{ij}!}{n_{ij1}! n_{ij2}! \dots n_{ijk}!} \times \pi_{ij1}^{n_{ij1}} \times \pi_{ij2}^{n_{ij2}} \dots \pi_{ijk}^{n_{ijk}}$$

where $n_{ijk} = 0, 1, \dots, n_{ij}; 0 \leq \pi_{ijk} \leq 1$

$i = 1, 2, \dots, I; j = 1, 2, \dots, J; \text{ and } k = 1, 2, \dots, K$

n_{ijk} is binomial with probability $n_{ij}\pi_{ijk}$. The covariance is given as

$$E(n_{ijk}, n'_{ijk}) = \begin{cases} -n_{ij}\pi_{ijk}\pi'_{ijk}; i = i'; j = j', k \neq k' \\ 0 \text{ elsewhere} \end{cases}$$

The variance is given as

$$\text{var}(n_{ijk}) = n_{ij}\pi_{ijk}(1 - \pi_{ijk})$$

More details regarding the assumptions may be seen in the literature used in this work. The two hypotheses considered in a nested classification, given that the data are enumerative, are;

$$H_{OA}: \pi_{ik} = \pi_k$$

$$H_{OB}: \pi_{ijk} = \pi_{ik}$$

The interpretation of the statistical statement in H_{OA} is that there is no main factor effect. In a similar way, the

interpretation of the statistical statement in H_{OB} is that there is no sub-factor effect.

Let the model be;

$$p_{ijk} = p_k + (p_{ik} - p_k) + (p_{ijk} - p_{ik}) \quad (2)$$

$$\text{where } p_{ijk} = \frac{n_{ijk}}{n_{ij}} \quad (3)$$

P_k is a constant for the k th response

$P_{ik} - P_k$ is the effect of the i th level of factor A

$P_{ijk} - P_{ik}$ is the effect of the j th level of factor B within the i th level of factor A.

Equating (2) and (3) gives

$$\frac{n_{ijk}}{n_{ij}} = p_k + p_{ik} - p_k + p_{ijk} - p_{ik}$$

Squaring both sides

$$\frac{n_{ijk}^2}{n_{ij}^2} = (P_k + P_{ik} - P_k + P_{ijk} - P_{ik})^2 \quad (4)$$

Multiplying both sides by n_{ij} gives

$$\frac{n_{ijk}^2}{n_{ij}} = n_{ij}(P_k + P_{ik} - P_k + P_{ijk} - P_{ik})^2 \quad (5)$$

Taking summation over both sides yields

$$\sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} = \sum_{ijk} n_{ij} (P_k + P_{ik} - P_k + P_{ijk} - P_{ik})^2 \quad (6)$$

This implies

$$\sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} = \sum_{ijk} n_{ij} (P_k + (P_{ik} - P_k) + (P_{ijk} - P_{ik}))^2 \quad (7)$$

This implies

$$\begin{aligned} \sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} &= \sum_{ijk} n_{ij} (P_k^2 + (P_{ik} - P_k)^2 + (P_{ijk} - P_{ik})^2 \\ &\quad + 2P_k(P_{ik} - P_k) + 2P_k(P_{ijk} - P_{ik}) \\ &\quad + 2(P_{ik} - P_k)(P_{ijk} - P_{ik})) \end{aligned} \quad (8)$$

Since $\sum_{ijk}(P_{ik} - P_k) = 0$ and $\sum_{ijk}(P_{ijk} - P_{ik}) = 0$, (Onukogu; 2014)

$$\sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} = \sum_{ijk} n_{ij} (P_k^2 + (P_{ik} - P_k)^2 + (P_{ijk} - P_{ik})^2) \quad (9)$$

Expanding (9) gives

$$\begin{aligned} \sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} &= \sum_{ijk} n_{ij} P_k^2 + \sum_{ijk} n_{ij} (P_{ik} - P_k)^2 + \\ &\quad \sum_{ijk} n_{ij} (P_{ijk} - P_{ik})^2 \end{aligned} \quad (10)$$

which results in

$$\begin{aligned} \sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} &= \sum_{ijk} n_{ij} \frac{n_k^2}{n^2} + \sum_{ijk} n_{ij} (P_{ik} - P_k)^2 + \\ &\quad \sum_{ijk} n_{ij} (P_{ijk} - P_{ik})^2 \end{aligned} \quad (11)$$

By rearrangement, we have

$$\begin{aligned} -\sum_{ijk} n_{ij} \frac{n_k^2}{n^2} &= -\sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} + \sum_{ijk} n_{ij} (P_{ik} - P_k)^2 + \\ &\quad \sum_{ijk} n_{ij} (P_{ijk} - P_{ik})^2 \end{aligned} \quad (12)$$

Adding n to both sides gives

$$\begin{aligned} n - \sum_{ijk} n_{ij} \frac{n_k^2}{n^2} &= n - \sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} + \sum_{ijk} n_{ij} (P_{ik} - P_k)^2 + \\ &\quad \sum_{ijk} n_{ij} (P_{ijk} - P_{ik})^2 \end{aligned} \quad (13)$$

Equation (13) is a partition of the sum of squares.

Thus

$$TSS = n - \sum_{ijk} n_{ij} \frac{n_k^2}{n^2} = n - \sum_k \frac{n_k^2}{n} = n - \sum_k \frac{n_k^2}{n} \quad (14)$$

$$USS = n - \sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} \quad (15)$$

$$SSA = \sum_{ijk} n_{ij} (P_{ik} - P_k)^2 = \sum_{ik} \frac{n_{ik}^2}{n_i} - \sum_k \frac{n_k^2}{n} \quad (16)$$

$$SSB_{(A)} = \sum_{ijk} n_{ij} (P_{ijk} - P_{ik})^2 = \sum_{ijk} \frac{n_{ijk}^2}{n_{ij}} - \sum_{ik} \frac{n_{ik}^2}{n_i} \quad (17)$$

If we let

$$C_k = \sum_k \frac{n_k^2}{n}, C_k = \sum_{ik} \frac{n_{ik}^2}{n_i}, C_k = \sum_{ijk} \frac{n_{ijk}^2}{n_{ij}},$$

then

$$TSS = n - C_k$$

$$SSA = C_{Ik} - C_k$$

$$SSB_{(A)} = C_{ijk} - C_{ik}$$

$$\text{and } USS = n - C_{ijk}$$

The total sum of squares (TSS) has $n - 1$ degree of freedom. The unit of sum of squares (USS) has $(n - I)J$ degrees of freedom. The sum of squares (SSA) of factor A has $I - 1$ degrees of freedom. The sum of squares ($SSB_{(A)}$) of the sub-factor B has $I(J - 1)$ degrees of freedom.

It has been shown in Onukogu (2014) that as n_{ij} tends to infinity, SSA and $SSB_{(A)}$ have approximate Chi-square distributions with $(k - 1)(I - 1)$ and $I(k - 1)(J - 1)$ degrees of freedom respectively. Hence in testing the hypothesis about the main factor effects, the test statistics is

$$X_A^2 = \frac{SSA(n - 1)(k - 1)}{TSS} \sim \chi_{(k-1)(I-1)}^2$$

Also, in testing the hypothesis about the sub-factor effects, the test statistics is

$$X_{B(A)}^2 = \frac{SSB_{(A)}(n - 1)(k - 1)}{TSS} \sim \chi_{(k-1)(J-1)I}^2$$

provided that TSS is independent of SSA and $SSB_{(A)}$.

Further details on asymptotic distribution may be seen in Onukogu (2014).

The CATANOVA table is summarized in Table 2.1.

Table 2.1. CATANOVA table for significance of effects

| Source of variation | Degrees of freedom | Sum of squares | Mean Squares | Test ratio | Hypothesis |
|---------------------|--------------------|----------------|--------------|--------------|--|
| A_i | $I - 1$ | SSA | MSA | X_A^2 | $H_{OA}: \pi_{ik} = \pi_k, \forall i$ |
| $B_{j(i)}$ | $I(J - 1)$ | $SSB_{(A)}$ | $MSB_{(A)}$ | $X_{B(A)}^2$ | $H_{OB}: \pi_{ijk} = \pi_{ik}, \forall j, i$ |
| U_{ijk} | $n - IJ$ | USS | UMS | | |
| Total | $n - 1$ | TSS | | | |

3. Results

With reference to the two-way nested classification for NECO 2003 in Appendix 1, the components of the sums of squares are computed as

$$C_k = \frac{131352^2 + 116085^2}{247437} = 124189.49$$

$$C_{ik} = \frac{12191^2 + 16240^2}{29431} + \frac{44794^2 + 42708^2}{87502}$$

$$+ \dots + \frac{24730^2 + 17562^2}{42292}$$

$$= 14873.44 + 43775.86 + \dots + 21753.45$$

$$= 125380.5$$

$$C_{ijk} = \frac{2209^2 + 2290^2}{4499} + \frac{1776^2 + 1839^2}{3615} + \dots + \frac{494^2 + 259^2}{753}$$

$$= 2250.23 + 1808.05 + \dots + 413.16$$

$$= 142666.05$$

The respective sums of squares are

$$TSS = n - c_k = 247437 - 124189.49$$

$$= 123247.51$$

$$SSA = C_{ik} - C_k = 125380.59 - 124189.49$$

$$= 1191.10$$

$$SS_{B(A)} = C_{ijk} - C_{ik} = 142666.05 - 125380.59$$

$$= 17285.46$$

The test statistic for assessing the significance of main factor effect is computed as

$$\chi_A^2 = \frac{SSA(n-1)(k-1)}{TSS}$$

$$= \frac{1191.10(247437-1)(2-1)}{123247.51}$$

$$= 2391.29$$

The test statistics for assessing the significance of sub-main factor effect is computed as

$$\chi_{B(A)}^2 = \frac{SS_{B(A)}(n-1)(K-1)}{TSS}$$

$$= \frac{17285.46(247437-1)(2-1)}{123247.51}$$

$$= 34702.89$$

The CATANOVA Table for the partitioning of sums of squares and testing relevant hypotheses is as in Table 3.2.

Table 3.2. CATANOVA Table for NECO 2003

| Source of Variation | Degrees of Freedom | Test Ratio | Critical -Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 2391.29 | 1.15 | < 0.00001 |
| State | 30 | 34702.89 | 18.5 | < 0.00001 |
| Within Units | 247401 | - | - | |
| Total | 2474436 | - | - | |

Table 3.3. CATANOVA Table for NECO 2004

| Source of Variation | Degrees of Freedom | Test Ratio | Critical -Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 2391.29 | 1.15 | < 0.00001 |
| State | 30 | 34702.89 | 18.5 | < 0.00001 |
| Within Units | 147767 | - | - | |
| Total | 147802 | - | - | |

Table 3.4. CATANOVA Table for NECO 2005

| Source of Variation | Degrees of Freedom | Test Ratio | Critical -Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 2284 | 1.15 | < 0.00001 |
| State | 30 | 966.36 | 18.5 | < 0.00001 |
| Within Units | 83375 | - | - | |
| Total | 83410 | - | - | |

Table 3.5. CATANOVA Table for NECO 2006

| Source of Variation | Degrees of Freedom | Test Ratio | Critical -Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 5352.71 | 1.15 | < 0.00001 |
| State | 30 | 2102.58 | 18.5 | < 0.00001 |
| Within Units | 258613 | - | - | |
| Total | 258648 | - | - | |

Table 3.6. CATANOVA Table for NECO 2007

| Source of Variation | Degrees of Freedom | Test Ratio | Critical -Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 50455.36 | 1.15 | < 0.00001 |
| State | 30 | 1648.24 | 18.5 | < 0.00001 |
| Within Units | 263773 | - | - | |
| Total | 263808 | - | - | |

Table 3.7. CATANOVA Table for NECO 2008

| Source of Variation | Degrees of Freedom | Test Ratio | Critical -Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 10839.45 | 1.15 | < 0.00001 |
| State | 30 | 18040.68 | 18.5 | < 0.00001 |
| Within Units | 635302 | - | - | |
| Total | 635337 | - | - | |

Table 3.8. CATANOVA Table for NECO 2009

| Source of Variation | Degrees of Freedom | Test Ratio | Critical -Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 771.37 | 1.15 | < 0.00001 |
| State | 30 | 429.45 | 18.5 | < 0.00001 |
| Within Units | 126434 | - | - | |
| Total | 126469 | - | - | |

Table 3.9. CATANOVA Table for NECO 2010

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 909.77 | 1.15 | < 0.00001 |
| State | 30 | 538.30 | 18.5 | < 0.00001 |
| Within Units | 105946 | - | - | |
| Total | 105981 | - | - | |

Table 3.10. CATANOVA Table for WAEC 2003

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 5352.71 | 1.15 | < 0.00001 |
| State | 30 | 2102.58 | 18.5 | < 0.00001 |
| Within Units | 200111 | - | - | |
| Total | 200146 | - | - | |

Table 3.11. CATANOVA Table for WAEC 2004

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 1503.01 | 1.15 | < 0.00001 |
| State | 30 | 615.05 | 18.5 | < 0.00001 |
| Within Units | 191902 | - | - | |
| Total | 191937 | - | - | |

Table 3.12. CATANOVA Table for WAEC 2005

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 1266.48 | 1.15 | < 0.00001 |
| State | 30 | 897.48 | 18.5 | < 0.00001 |
| Within Units | 203955 | - | - | |
| Total | 203990 | - | - | |

Table 3.13. CATANOVA Table for WAEC 2006

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 782.51 | 1.15 | < 0.00001 |
| State | 30 | 919.75 | 18.5 | < 0.00001 |
| Within Units | 105464 | - | - | |
| Total | 105499 | - | - | |

Table 3.14. CATANOVA Table for WAEC 2007

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 489.18 | 1.15 | < 0.00001 |
| State | 30 | 335.57 | 18.5 | < 0.00001 |
| Within Units | 98121 | - | - | |
| Total | 98156 | - | - | |

Table 3.15. CATANOVA Table for WAEC 2008

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 541.19 | 1.15 | < 0.00001 |
| State | 30 | 453.56 | 18.5 | < 0.00001 |
| Within Units | 127111 | - | - | |
| Total | 127146 | - | - | |

Table 3.16. CATANOVA Table for WAEC 2009

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 782.51 | 1.15 | < 0.00001 |
| State | 30 | 919.75 | 18.5 | < 0.00001 |
| Within Units | 37859 | - | - | |
| Total | 37894 | - | - | |

Table 3.17. CATANOVA Table for WAEC 2010

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 1751 | 1.15 | < 0.00001 |
| State | 30 | 70779 | 18.5 | < 0.00001 |
| Within Units | 307287 | - | - | |
| Total | 307322 | - | - | |

Table 3.18. CATANOVA Table for WAEC (Duration-based)

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 73997.52 | 1.15 | < 0.00001 |
| State | 30 | 158297.52 | 18.5 | < 0.00001 |
| Within Units | 1272063 | - | - | |
| Total | 1272098 | - | - | |

Table 3.19. CATANOVA Table for NECO (Duration-based)

| Source of Variation | Degrees of Freedom | Test Ratio | Critical –Value at 5% | p-value |
|---------------------|--------------------|------------|-----------------------|-----------|
| Zone | 5 | 250943.03 | 1.15 | < 0.00001 |
| State | 30 | 177780.34 | 18.5 | < 0.00001 |
| Within Units | 1868863 | - | - | |
| Total | 1868898 | - | - | |

With reference to Appendices 1-4, the computational requirements for the nested data arrangements are summarized in CATANOVA Tables 3.3-3.19.

4. Discussion of Results

CATANOVA for data in nested arrangements has been applied in studying the significance of performance in WAEC and NECO. The need for nesting of factors arises

when an experimental situation requires that unique levels of one factor occur within each level of a second factor. Two-way nesting due to Onukogu (2014) was considered where Geopolitical Zone served as the main factor and State served as the sub-factor. In one instance, Gender was viewed as the “response” factor, having two levels. In another instance, “Duration of Examination” was viewed as the “response” factor, having eight levels. Six levels of Geopolitical Zone were considered. Although the States within each Geopolitical Zone are not constant, thus creating imbalance to sub-factor levels, an average of sub-factor levels was used throughout the research. It is assumed that for large sample size, the CATANOVA technique employed is robust against slight imbalance in sub-factor levels. In examining the performances in WAEC and NECO in relation to Geopolitical Zones and States within Zones, two aspects, namely, Gender-based assessment and Duration-based assessment were considered. The categorical data analysis revealed significant difference in the performance of students in WAEC and NECO within the Geo-political Zones (main factor) and the States (sub-factor). The p-value for each test revealed that the test

is significant at an Observed Significance Level (OSL) less than 0.00001. When sub-factor levels, nested within the main factor, are not constant, an average of sub-factor levels may be used in the analysis of the data having unbalanced structure, particularly if the sub-factor levels are not grossly different and if the sample size is large.

5. Conclusions

Frequency data are increasingly encountered in real-life situations. Unfortunately, some of such data have been erroneously analysed without the consciousness that there is no underlying metric with such data. Fortunately, attention is gradually being drawn to the need for methods of handling frequency (count, categorical or nominal) data. In this work, an application of categorical analysis of variance has been presented for data in nested arrangement, in view of establishing the significance of main factor and sub-factor effects. It provides a practical guide for determining the significance of categorical factors when data are in nested arrangements.

Appendix 1: Two-way Nested Classification for NECO (Gender-based)

| NECO2003 | | | | | | | | | | | | | |
|----------|------|------|------|-------|------|----------|-------|------|------|-------|-------|-------|----------|
| | Z1 | | | | | | | Z2 | | | | | |
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} |
| male | 2209 | 1776 | 2508 | 3872 | 2826 | 13191 | 8218 | 4691 | 3779 | 6940 | 7700 | 13466 | 44794 |
| female | 2290 | 1839 | 2222 | 6161 | 3738 | 16250 | 9701 | 4320 | 2734 | 7099 | 6901 | 11953 | 42708 |
| n_{ij} | 4499 | 3615 | 4730 | 10033 | 6564 | | 17919 | 9011 | 6513 | 14039 | 14601 | 25419 | |
| total | | | | | | | | | | | | | |
| n_i | | | | | | 29441 | | | | | | | 87502 |

| | Z3 | | | | | | | Z4 | | | | | | | |
|----------|------|-------|-------|------|------|------|----------|------|------|------|-----|------|------|-----|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} |
| male | 2107 | 12064 | 8765 | 3803 | 4726 | 2793 | 34258 | 1393 | 5349 | 1329 | 165 | 1213 | 1630 | 252 | 11331 |
| female | 2160 | 11115 | 8059 | 2975 | 4530 | 2387 | 31226 | 241 | 3964 | 1010 | 186 | 411 | 55 | 621 | 6488 |
| n_{ij} | 4267 | 23179 | 16824 | 6778 | 9256 | 5180 | | 1634 | 9313 | 2339 | 351 | 1624 | 1685 | 873 | |
| Total | | | | | | | | | | | | | | | |
| n_i | | | | | | | 65844 | | | | | | | | 17819 |

| | Z5 | | | | | | | Z6 | | | | | | | Total | |
|----------|-----|------|------|-----|-----|-----|----------|------|------|-------|------|------|------|-----|----------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | n_k |
| male | 373 | 721 | 1185 | 163 | 420 | 186 | 3048 | 5615 | 5232 | 8310 | 973 | 2096 | 2010 | 494 | 24730 | 131352 |
| female | 145 | 389 | 1013 | 108 | 190 | 16 | 1861 | 3599 | 3792 | 6814 | 557 | 1170 | 1371 | 259 | 17562 | 116095 |
| n_{ij} | 518 | 1110 | 2198 | 271 | 610 | 202 | | 9214 | 9024 | 15124 | 1530 | 3266 | 3381 | 753 | | |
| Total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 4909 | | | | | | | | 42292 | 247447 |

| 2004 NECO | | | | | | | | | | | | | |
|-----------|------|------|------|------|------|----------|------|------|------|-------|------|-------|----------|
| | Z1 | | | | | | | Z2 | | | | | |
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} |
| male | 1801 | 1475 | 2479 | 3093 | 1753 | 10601 | 4398 | 1563 | 1957 | 6036 | 2635 | 9279 | 25868 |
| female | 2130 | 1699 | 2257 | 4281 | 2252 | 12619 | 5507 | 1639 | 1491 | 5998 | 2532 | 8301 | 25468 |
| n_{ij} | 3931 | 3174 | 4736 | 7374 | 4005 | | 9905 | 3202 | 3448 | 12034 | 5167 | 17580 | |
| Total | | | | | | | | | | | | | |
| n_i | | | | | | 23220 | | | | | | | 51336 |

| | | | | | | | | | | | | | | | | |
|-----------------------|------|-------|-------|-------|-------|-----------------------|-----------------------|-------|-------|------|-------|------|-------|-----------------------|-----------------------|----------------------|
| | Z3 | | | | | | | Z4 | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | <i>n_{ij}</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | <i>n_{ij}</i> | |
| male | 1699 | 7255 | 4403 | 2782 | 4959 | 3383 | 24481 | 982 | 2529 | 227 | 230 | 301 | 55 | 27 | 4351 | |
| female | 1736 | 6681 | 4416 | 2395 | 4781 | 3167 | 23176 | 271 | 1815 | 117 | 380 | 193 | 78 | 26 | 2880 | |
| <i>n_{ij}</i> | 3435 | 13936 | 8819 | 5177 | 9740 | 6550 | | 1253 | 4344 | 344 | 610 | 494 | 133 | 53 | | |
| Total | | | | | | | | | | | | | | | | |
| | | | | | | | 47657 | | | | | | | | 7231 | |
| | | | | | | | | | | | | | | | | |
| | Z5 | | | | | | | Z6 | | | | | | | Total | |
| | 1 | 2 | 3 | 4 | 5 | 6 | <i>n_{ij}</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | <i>n_{ij}</i> | <i>n_k</i> |
| male | 239 | 660 | 1248 | 171 | 370 | 74 | 2762 | 2907 | 1513 | 995 | 923 | 1022 | 594 | 520 | 8474 | 76537 |
| female | 132 | 205 | 838 | 30 | 180 | 34 | 1419 | 1802 | 1014 | 814 | 462 | 712 | 474 | 426 | 5704 | 71266 |
| <i>n_{ij}</i> | 371 | 865 | 2086 | 201 | 550 | 108 | | 4709 | 2527 | 1809 | 1385 | 1734 | 1068 | 946 | | |
| total | | | | | | | | | | | | | | | | |
| | | | | | | | 4181 | | | | | | | | 14178 | 147803 |
| | | | | | | | | | | | | | | | | |
| 2005NECO | | | | | | | | | | | | | | | | |
| Total | Z1 | | | | | | | Z2 | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | <i>n_{ij}</i> | | 1 | 2 | 3 | 4 | 5 | 6 | <i>n_{ij}</i> | | |
| male | 1019 | 615 | 1039 | 1456 | 624 | 4753 | | 1365 | 672 | 931 | 2034 | 1591 | 6112 | | 12705 | |
| female | 1271 | 742 | 850 | 2215 | 779 | 5857 | | 1535 | 461 | 747 | 2117 | 1387 | 5455 | | 11702 | |
| <i>n_{ij}</i> | 2290 | 1357 | 1889 | 3671 | 1403 | | | 2904 | 1133 | 1678 | 4151 | 2978 | 11567 | | | |
| Total | | | | | | | | | | | | | | | | |
| <i>n_i</i> | | | | | | | 10610 | | | | | | | | | 24411 |
| | | | | | | | | | | | | | | | | |
| | Z3 | | | | | | | Z4 | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | <i>n_{ij}</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | <i>n_{ij}</i> | |
| male | 1425 | 3973 | 1988 | 577 | 1898 | 2145 | 12006 | 2800 | 71 | 736 | 644 | 537 | 905 | 387 | 6080 | |
| female | 1434 | 3848 | 1936 | 498 | 1872 | 1782 | 11370 | 349 | 81 | 492 | 179 | 347 | 365 | 245 | 2058 | |
| <i>n_{ij}</i> | 2859 | 7821 | 3924 | 1075 | 3770 | 3927 | | 3149 | 152 | 1228 | 823 | 884 | 1270 | 632 | | |
| total | | | | | | | | | | | | | | | | |
| | | | | | | | 23376 | | | | | | | | | 8138 |
| | | | | | | | | | | | | | | | | |
| | Z5 | | | | | | | Z6 | | | | | | | Total | |
| | 1 | 2 | 3 | 4 | 5 | 6 | <i>n_{ij}</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | <i>n_{ij}</i> | <i>n_k</i> |
| male | 657 | 845 | 984 | 158 | 291 | 58 | 2993 | 1820 | 1443 | 1007 | 1693 | 566 | 595 | 374 | 7498 | 46035 |
| female | 311 | 358 | 902 | 75 | 125 | 38 | 1809 | 1028 | 1069 | 671 | 854 | 288 | 366 | 300 | 4576 | 37372 |
| <i>n_{ij}</i> | 968 | 1203 | 1886 | 233 | 416 | 96 | | 2848 | 2512 | 1678 | 2547 | 854 | 961 | 674 | | |
| total | | | | | | | | | | | | | | | | |
| | | | | | | | 4802 | | | | | | | | 12074 | 83411 |
| | | | | | | | | | | | | | | | | |
| 2006 NECO | | | | | | | | | | | | | | | | |
| | Z1 | | | | | | | Z2 | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | <i>n_{ij}</i> | | 1 | 2 | 3 | 4 | 5 | 6 | <i>n_{ij}</i> | | |
| male | 3465 | 2406 | 2802 | 4583 | 4100 | 17356 | | 5199 | 1711 | 3619 | 5852 | 3772 | 8823 | | 28976 | |
| female | 3766 | 2604 | 2573 | 5827 | 5457 | 20227 | | 5729 | 1629 | 2771 | 5840 | 3495 | 8197 | | 27661 | |
| <i>n_{ij}</i> | 7231 | 5010 | 5375 | 10410 | 9557 | | | 10928 | 3340 | 6390 | 11692 | 7267 | 17020 | | | |
| Total | | | | | | | | | | | | | | | | |
| <i>n_i</i> | | | | | | | 37583 | | | | | | | | | 56637 |
| | | | | | | | | | | | | | | | | |
| | Z3 | | | | | | | Z4 | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | <i>n_{ij}</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | <i>n_{ij}</i> | |
| male | 4035 | 8148 | 5792 | 3972 | 5888 | 6503 | 34338 | 2238 | 7557 | 2202 | 2285 | 1051 | 5459 | 673 | 21465 | |
| female | 3828 | 8152 | 5623 | 2979 | 5396 | 5783 | 31761 | 335 | 5566 | 1246 | 856 | 307 | 2872 | 381 | 11563 | |
| <i>n_{ij}</i> | 7863 | 16300 | 11415 | 6951 | 11284 | 12286 | | 2573 | 13123 | 3448 | 3141 | 1358 | 8331 | 1054 | | |
| total | | | | | | | | | | | | | | | | |
| | | | | | | | 66099 | | | | | | | | | 33028 |
| | | | | | | | | | | | | | | | | |
| | Z5 | | | | | | | Z6 | | | | | | | Total | |
| | 1 | 2 | 3 | 4 | 5 | 6 | <i>n_{ij}</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | <i>n_{ij}</i> | <i>n_k</i> |
| male | 1815 | 3328 | 2106 | 775 | 1651 | 989 | 10664 | 7928 | 6714 | 3514 | 5459 | 2213 | 2485 | 928 | 29241 | 142040 |
| female | 901 | 1501 | 1126 | 403 | 912 | 647 | 5490 | 5019 | 5431 | 2709 | 2872 | 1370 | 1727 | 779 | 19907 | 116609 |
| <i>n_{ij}</i> | 2716 | 4829 | 3232 | 1178 | 2563 | 1636 | | 12947 | 12145 | 6223 | 8331 | 3583 | 4212 | 1707 | | |
| total | | | | | | | | | | | | | | | | |
| <i>n_i</i> | | | | | | | 16154 | | | | | | | | 49148 | 258649 |

| 2007NECO | | | | | | | | | | | | | | |
|----------|------|------|------|-------|-------|----------|-------|------|------|-------|-------|-------|----------|--|
| | Z1 | | | | | | Z2 | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | |
| male | 3892 | 3116 | 4218 | 5206 | 4913 | 21345 | 6739 | 3012 | 4693 | 7144 | 6558 | 11232 | 39378 | |
| female | 4101 | 3463 | 3681 | 6567 | 6845 | 24657 | 7501 | 2688 | 3711 | 7072 | 6247 | 10686 | 37905 | |
| n_{ij} | 7993 | 6579 | 7899 | 11773 | 11758 | | 14240 | 5700 | 8404 | 14216 | 12805 | 21918 | | |
| Total | | | | | | | | | | | | | | |
| n_i | | | | | | 46002 | | | | | | | 77283 | |

| | Z3 | | | | | | Z4 | | | | | | | | | |
|----------|------|-------|-------|------|------|-------|----------|------|-------|------|------|------|-----|------|----------|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | |
| male | 4817 | 11839 | 7572 | 4837 | 4030 | 8335 | 41430 | 2180 | 8048 | 4997 | 3875 | 1838 | 115 | 655 | 21708 | |
| female | 4627 | 11480 | 7143 | 3754 | 3645 | 7502 | 38151 | 488 | 5465 | 2346 | 1419 | 584 | 35 | 571 | 10908 | |
| n_{ij} | 9444 | 23319 | 14715 | 8591 | 7675 | 15837 | | 2668 | 13513 | 7343 | 5294 | 2422 | 150 | 1226 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 79581 | | | | | | | | 32616 | |

| | Z5 | | | | | | Z6 | | | | | | | | | Total |
|----------|------|------|------|------|------|-----|----------|-------|-------|-------|------|------|------|------|----------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | n_k |
| male | 1458 | 2471 | 1787 | 1063 | 1267 | 73 | 8119 | 8662 | 6972 | 5634 | 5806 | 3464 | 4397 | 1450 | 36385 | 147020 |
| female | 739 | 884 | 962 | 461 | 720 | 49 | 3815 | 5443 | 5332 | 4390 | 3719 | 2314 | 3328 | 1484 | 26010 | 116789 |
| n_{ij} | 2197 | 3355 | 2749 | 1524 | 1987 | 122 | | 14105 | 12304 | 10024 | 9525 | 5778 | 7725 | 2934 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 11934 | | | | | | | | 62395 | 263809 |

| 2008NECO | | | | | | | | | | | | | | | | |
|----------|-------|-------|-------|-------|------|----------|-------|-------|-------|-------|-------|-------|----------|--|--|--|
| | Z1 | | | | | | Z2 | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | | | |
| male | 4948 | 5948 | 9376 | 10399 | 6601 | 37272 | 12061 | 7250 | 5850 | 13359 | 11511 | 19661 | 69692 | | | |
| female | 5947 | 6716 | 8812 | 13197 | 2426 | 37098 | 13199 | 5131 | 4655 | 12714 | 11590 | 17709 | 64998 | | | |
| n_{ij} | 10895 | 12664 | 18188 | 23596 | 9027 | | 25260 | 12381 | 10505 | 26073 | 23101 | 37370 | | | | |
| Total | | | | | | | | | | | | | | | | |
| n_i | | | | | | 74370 | | | | | | | 134690 | | | |

| | Z3 | | | | | | Z4 | | | | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|----------|------|-------|-------|-------|------|------|------|----------|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | |
| male | 7244 | 20889 | 11893 | 8684 | 11015 | 14810 | 74535 | 4877 | 19761 | 18801 | 10305 | 3661 | 2691 | 6601 | 66697 | |
| female | 6522 | 20282 | 11323 | 7504 | 9708 | 421 | 55760 | 784 | 13745 | 7895 | 3755 | 1536 | 671 | 2426 | 30812 | |
| n_{ij} | 13766 | 41171 | 23216 | 16188 | 20723 | 15231 | | 5661 | 33506 | 26696 | 14060 | 5197 | 3362 | 9027 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 130295 | | | | | | | | 97509 | |

| | Z5 | | | | | | Z6 | | | | | | | | | Total |
|----------|-------|-------|-------|-------|------|------|----------|-------|-------|-------|-------|-------|-------|------|----------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | n_k |
| male | 9676 | 10272 | 8973 | 9816 | 5229 | 3991 | 47957 | 19340 | 13684 | 6475 | 15893 | 9287 | 9072 | 2096 | 75847 | 372000 |
| female | 5689 | 4353 | 5849 | 4222 | 2379 | 1378 | 23870 | 12451 | 10391 | 4848 | 9008 | 5239 | 6973 | 1890 | 50800 | 263338 |
| n_{ij} | 15365 | 14625 | 14822 | 14038 | 7608 | 5369 | | 31791 | 24075 | 11323 | 24901 | 14526 | 16045 | 3986 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 71827 | | | | | | | | 126647 | 635338 |

| 2009NECO | | | | | | | | | | | | | | | | |
|----------|------|------|------|------|------|----------|------|------|------|-------|------|-------|----------|--|--|--|
| | Z1 | | | | | | Z2 | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | | | |
| male | 2347 | 1327 | 1125 | 1448 | 1762 | 8009 | 2463 | 1256 | 2337 | 5186 | 2566 | 5607 | 19415 | | | |
| female | 3061 | 1621 | 1065 | 1680 | 2391 | 9818 | 2763 | 1143 | 2038 | 5148 | 2626 | 5392 | 19110 | | | |
| n_{ij} | 5408 | 2948 | 2190 | 3128 | 4153 | | 5226 | 2399 | 4375 | 10334 | 5192 | 10999 | | | | |
| Total | | | | | | | | | | | | | | | | |
| n_i | | | | | | 17827 | | | | | | | 38525 | | | |

| | Z3 | | | | | | Z4 | | | | | | | | | |
|----------|------|-------|------|------|------|------|----------|----|------|------|------|-----|-----|-----|----------|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | |
| male | 1071 | 10049 | 3554 | 1554 | 2241 | 3775 | 22244 | 42 | 2619 | 1065 | 997 | 67 | 168 | 119 | 5077 | |
| female | 1212 | 10049 | 3554 | 1464 | 2146 | 3613 | 22038 | 20 | 2397 | 782 | 398 | 63 | 48 | 70 | 3778 | |
| n_{ij} | 2283 | 20098 | 7108 | 3018 | 4387 | 7388 | | 62 | 5016 | 1847 | 1395 | 130 | 216 | 189 | | |
| Total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 44282 | | | | | | | | 8855 | |

Appendix 2: Two-way Nested Classification for WAEC (Gender-based)

[illegible]

| | | | | | | | | | | | | | | | | |
|------------|------|-------|------|------|------|----------|----------|------|------|-------|------|-------|----------|------|----------|--------|
| 2004WAEC | | | | | | | | | | | | | | | | |
| | Z1 | | | | | | Z2 | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | | | |
| male | 2118 | 1992 | 1808 | 3097 | 3985 | 13000 | 6391 | 1289 | 116 | 11203 | 2643 | 9280 | 30922 | | | |
| femal | 2443 | 2683 | 1771 | 4678 | 5102 | 16677 | 7639 | 1321 | 96 | 13151 | 2274 | 9274 | 33755 | | | |
| n_{ij} | 4561 | 4675 | 3579 | 7775 | 9087 | | 14030 | 2610 | 212 | 24354 | 4917 | 18554 | | | | |
| Total | | | | | | | | | | | | | | | | |
| n_i | | | | | | 29677 | | | | | | | | | 64677 | |
| | | | | | | | | | | | | | | | | |
| | Z3 | | | | | | Z4 | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | |
| male | 1025 | 21894 | 3319 | 4757 | 1446 | 2901 | 35342 | 461 | 2795 | 1630 | 174 | 137 | 192 | 675 | 6064 | |
| female | 913 | 21374 | 3056 | 3836 | 1378 | 2610 | 33167 | 126 | 2078 | 1494 | 68 | 136 | 166 | 505 | 4573 | |
| n_{ij} | 1938 | 43268 | 6375 | 8593 | 2824 | 5511 | | 587 | 4873 | 3124 | 242 | 273 | 358 | 1180 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 68509 | | | | | | | | 10637 | |
| | | | | | | | | | | | | | | | | |
| | Z5 | | | | | | Z6 | | | | | | | | | Total |
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | n_k |
| male | 227 | 2010 | 635 | 63 | 133 | 153 | 3221 | 2571 | 1743 | 462 | 815 | 645 | 1024 | 459 | 7719 | 96268 |
| female | 131 | 1546 | 487 | 26 | 82 | 43 | 2315 | 1669 | 1323 | 351 | 440 | 446 | 675 | 279 | 5183 | 95670 |
| n_{ij} | 358 | 3556 | 1122 | 89 | 215 | 196 | | 4240 | 3066 | 813 | 1255 | 1091 | 1699 | 738 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 5536 | | | | | | | | 12902 | 191938 |
| | | | | | | | | | | | | | | | | |
| 2005WAEC | | | | | | | | | | | | | | | | |
| | Z1 | | | | | | Z2 | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | | | |
| male | 3220 | 2902 | 1575 | 3556 | 3828 | 15081 | 5264 | 1272 | 226 | 12628 | 2118 | 13080 | 34588 | | | |
| female | 3684 | 3063 | 1356 | 4734 | 4401 | 17238 | 5841 | 1078 | 145 | 13271 | 1819 | 12611 | 34765 | | | |
| n_{ij} | 6904 | 5965 | 2931 | 8290 | 8229 | | 11105 | 2350 | 371 | 25899 | 3937 | 25691 | | | | |
| Total | | | | | | | | | | | | | | | | |
| n_i | | | | | | 32319 | | | | | | | | | 69353 | |
| | | | | | | | | | | | | | | | | |
| | Z3 | | | | | | Z4 | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | |
| male | 1492 | 25904 | 2852 | 1847 | 1418 | 2288 | 35801 | 262 | 3413 | 2253 | 307 | 132 | 489 | 1319 | 8175 | |
| female | 1279 | 26068 | 2665 | 1504 | 1327 | 2070 | 34913 | 255 | 2742 | 1835 | 127 | 91 | 205 | 282 | 5537 | |
| n_{ij} | 2771 | 51972 | 5517 | 3351 | 2745 | 4358 | | 517 | 6155 | 4088 | 434 | 223 | 694 | 1601 | | |
| total | | | | | | | | | | | | | | | | |
| | | | | | | | 70714 | | | | | | | | 13713 | |
| | | | | | | | | | | | | | | | | |
| | Z5 | | | | | | Z6 | | | | | | | | | Total |
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | n_k |
| male | 370 | 1600 | 825 | 73 | 334 | 203 | 3405 | 1903 | 2024 | 718 | 750 | 485 | 901 | 547 | 7328 | 104378 |
| female | 173 | 1189 | 608 | 37 | 153 | 67 | 2227 | 1113 | 1476 | 631 | 387 | 384 | 593 | 349 | 4933 | 99613 |
| n_{ij} | 543 | 2789 | 1433 | 110 | 487 | 270 | | 3016 | 3500 | 1349 | 1137 | 869 | 1494 | 896 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 5632 | | | | | | | | 12261 | 203991 |
| | | | | | | | | | | | | | | | | |
| 2006WAE(Z1 | | | | | | | | | | | | | | | | |
| | Z1 | | | | | | Z2 | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | | | |
| male | 1993 | 1216 | 404 | 1850 | 2373 | 7836 | 3974 | 111 | 260 | 8324 | 1573 | 8408 | 22650 | | | |
| female | 1800 | 845 | 329 | 1607 | 2441 | 7022 | 3726 | 543 | 180 | 6084 | 1036 | 6217 | 17786 | | | |
| n_{ij} | 3793 | 2061 | 733 | 3457 | 4814 | | 7700 | 654 | 440 | 14408 | 2609 | 14625 | | | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | 14858 | | | | | | | | | 40436 | |
| | | | | | | | | | | | | | | | | |
| | Z3 | | | | | | Z4 | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | |
| male | 1370 | 11190 | 1966 | 1085 | 1448 | 1945 | 19004 | 176 | 2540 | 1266 | 194 | 22 | 52 | 427 | 4677 | |
| female | 1071 | 7947 | 1462 | 742 | 1205 | 1494 | 13921 | 19 | 1531 | 820 | 100 | 16 | 17 | 132 | 2635 | |
| n_{ij} | 2441 | 19137 | 3428 | 1827 | 2653 | 3439 | | 195 | 4071 | 2086 | 294 | 38 | 69 | 559 | | |
| toal | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 32925 | | | | | | | | 7312 | |

| | Z5 | | | | | | | | Z6 | | | | | | | | Total |
|----------|-----|------|-----|-----|-----|-----|----------|------|------|-----|-----|-----|-----|-----|----------|--------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | n_k | |
| male | 199 | 942 | 566 | 173 | 103 | 124 | 2107 | 996 | 961 | 449 | 587 | 603 | 593 | 523 | 4712 | 60986 | |
| female | 119 | 501 | 337 | 51 | 34 | 39 | 1081 | 320 | 516 | 352 | 255 | 235 | 315 | 76 | 2069 | 44514 | |
| n_{ij} | 318 | 1443 | 903 | 224 | 137 | 163 | | 1316 | 1477 | 801 | 842 | 838 | 908 | 599 | | | |
| total | | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 3188 | | | | | | | | 6781 | 105500 | |

| 2007WAE(Z1 | | | | | | | | | Z2 | | | | | | | | |
|------------|------|------|------|------|------|----------|-------|------|-----|-----|-------|------|-------|----------|--|--|--|
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | | | |
| male | 2246 | 1308 | 745 | 1817 | 1134 | 7250 | | 2687 | 504 | 169 | 8598 | 1810 | 8150 | 21918 | | | |
| female | 2273 | 1132 | 568 | 1712 | 1365 | 7050 | | 2644 | 355 | 102 | 7020 | 1329 | 6583 | 18033 | | | |
| n_{ij} | 4319 | 2440 | 1313 | 3529 | 2499 | | | 5331 | 859 | 271 | 15618 | 3139 | 14733 | | | | |
| Total | | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 14100 | | | | | | | 39951 | | | |

| | Z3 | | | | | | | Z4 | | | | | | | |
|----------|------|-------|------|------|------|------|----------|-----|------|------|-----|----|-----|-----|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_k |
| male | 1043 | 8502 | 1461 | 1575 | 971 | 1332 | 14884 | 65 | 2449 | 1291 | 426 | 42 | 135 | 252 | 4660 |
| female | 959 | 6188 | 1211 | 1236 | 842 | 990 | 11426 | 37 | 1689 | 857 | 265 | 13 | 18 | 130 | 3009 |
| n_{ij} | 2002 | 14690 | 2672 | 2811 | 1813 | 2322 | | 102 | 4138 | 2148 | 691 | 55 | 153 | 382 | |
| Total | | | | | | | | | | | | | | | |
| n_i | | | | | | | 26310 | | | | | | | | 7669 |

| | Z5 | | | | | | | Z6 | | | | | | | Total | |
|----------|-----|------|-----|-----|-----|-----|----------|------|------|-----|-----|-----|------|------|----------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | n_k |
| male | 168 | 1355 | 384 | 70 | 202 | 181 | 2360 | 843 | 659 | 505 | 359 | 334 | 677 | 630 | 4007 | 55079 |
| female | 68 | 741 | 239 | 36 | 109 | 82 | 1275 | 324 | 419 | 424 | 195 | 187 | 384 | 452 | 2385 | 43178 |
| n_{ij} | 236 | 2096 | 623 | 106 | 311 | 263 | | 1167 | 1078 | 929 | 554 | 621 | 1061 | 1082 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 3635 | | | | | | | | 6492 | 98157 |

| 2008WAE(Z1 | | | | | | | | | Z2 | | | | | | |
|------------|------|------|------|------|------|----------|-------|------|------|-----|-------|------|-------|----------|--|
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | |
| male | 2868 | 2320 | 890 | 2874 | 1835 | 10787 | | 4122 | 1496 | 242 | 10331 | 2741 | 9243 | 28175 | |
| female | 2637 | 1830 | 736 | 2541 | 2063 | 9807 | | 4073 | 1023 | 147 | 7832 | 1779 | 7393 | 22247 | |
| n_{ij} | 5505 | 4150 | 1626 | 5415 | 3898 | | | 8195 | 2519 | 389 | 18163 | 4520 | 16636 | | |
| Total | | | | | | | | | | | | | | | |
| n_i | | | | | | | 20594 | | | | | | | 50422 | |

| | Z3 | | | | | | | Z4 | | | | | | | |
|----------|------|-------|------|------|------|------|----------|-----|------|------|-----|----|-----|-----|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} |
| male | 797 | 11916 | 1917 | 1521 | 1295 | 1642 | 19088 | 70 | 2867 | 2065 | 532 | 75 | 104 | 311 | 6024 |
| female | 718 | 8792 | 1555 | 1171 | 1188 | 1262 | 14686 | 35 | 1934 | 1324 | 232 | 24 | 31 | 213 | 3793 |
| n_{ij} | 1515 | 20708 | 3472 | 2692 | 2483 | 2904 | | 105 | 4801 | 3389 | 764 | 99 | 135 | 524 | |
| n_i | | | | | | | 33774 | | | | | | | | 9817 |

| | Z5 | | | | | | | Z6 | | | | | | | Total | |
|----------|-----|------|-----|-----|-----|-----|----------|------|------|-----|-----|-----|------|------|----------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | 1 | 2 | 3 | 4 | 5 | 6 | 7 | n_{ij} | n_k |
| male | 232 | 1145 | 472 | 75 | 211 | 125 | 2260 | 1316 | 1351 | 369 | 322 | 622 | 1001 | 803 | 5784 | 72118 |
| female | 104 | 619 | 335 | 49 | 85 | 65 | 1257 | 563 | 769 | 320 | 112 | 298 | 624 | 553 | 3239 | 55029 |
| n_{ij} | 336 | 1764 | 807 | 124 | 296 | 190 | | 1879 | 2120 | 689 | 434 | 920 | 1625 | 1356 | | |
| total | | | | | | | | | | | | | | | | |
| n_i | | | | | | | 3517 | | | | | | | | 9023 | 127147 |

| 2009WAE(Z1 | | | | | | | | | Z2 | | | | | | |
|------------|-----|-----|----|-----|-----|----------|-----|------|------|-----|------|------|------|----------|--|
| | 1 | 2 | 3 | 4 | 5 | n_{ij} | | 1 | 2 | 3 | 4 | 5 | 6 | n_{ij} | |
| male | 254 | 131 | 28 | 64 | 114 | 591 | | 1344 | 1413 | 75 | 2551 | 946 | 3786 | 10115 | |
| female | 157 | 71 | 13 | 50 | 92 | 383 | | 1266 | 830 | 36 | 1612 | 658 | 2980 | 7382 | |
| n_{ij} | 411 | 202 | 41 | 114 | 206 | | | 2610 | 2243 | 111 | 4163 | 1604 | 6766 | | |
| Total | | | | | | | | | | | | | | | |
| n_i | | | | | | | 974 | | | | | | | 17497 | |

Appendix 3: Two-way Nested Classification for WAEC (Duration-based)

| WAEC | | | | | | | | | | | | | |
|------------|-------|-------|-------|-------|-------|----------|-------|-------|------|--------|-------|--------|----------|
| | Z1 | | | | | | Z2 | | | | | | |
| yearly/per | 1 | 2 | 3 | 4 | 5 | n_{ik} | 1 | 2 | 3 | 4 | 5 | 6 | n_{ik} |
| 2003 | 3273 | 2572 | 2547 | 5733 | 6139 | 20264 | 15733 | 4798 | 299 | 19785 | 5558 | 18689 | 64862 |
| 2004 | 4561 | 4675 | 3579 | 7775 | 9087 | 29677 | 14030 | 2610 | 212 | 24354 | 4917 | 18554 | 64677 |
| 2005 | 6904 | 5965 | 2931 | 8290 | 8229 | 32319 | 11105 | 2350 | 371 | 25899 | 3937 | 25691 | 69353 |
| 2006 | 3793 | 2061 | 733 | 3457 | 4814 | 14858 | 7700 | 654 | 440 | 14408 | 2609 | 14625 | 40436 |
| 2007 | 4319 | 2440 | 1313 | 3529 | 2499 | 14100 | 5331 | 859 | 271 | 15618 | 3139 | 14733 | 39951 |
| 2008 | 5505 | 4150 | 1626 | 5415 | 3898 | 20594 | 8195 | 2519 | 389 | 18163 | 4520 | 16636 | 50422 |
| 2009 | 411 | 202 | 41 | 114 | 206 | 974 | 2610 | 2243 | 111 | 4163 | 1604 | 6766 | 17497 |
| 2010 | 18284 | 18444 | 4471 | 11737 | 10828 | 63764 | 8641 | 10105 | 6313 | 13715 | 14065 | 14886 | 67725 |
| n_{ij} | 47051 | 40511 | 17244 | 46054 | 45705 | | 73346 | 26140 | 8409 | 136109 | 40354 | 130586 | |
| Total | | | | | | | | | | | | | |
| n_i | | | | | | 196550 | | | | | | | 414923 |

REFERENCES

- [1] Agresti, A. (2007). *Categorical data analysis*. New York: Wiley Interscience.
- [2] Cochran W. G. (1940). The analysis of variances when experimental errors follow the Poisson or binomial laws. *The Annals of Mathematical Statistics*, II, 335-347.
- [3] Cochran, W. G (1950). "The comparison of percentages in matched samples," *Biometrika*, 37: 256-55.
- [4] Dyke, G. V. and Patterson, H. D. (1952). Analysis of factorial arrangements when the data are proportions. *Biometrics*, 8, 1-12.
- [5] Florian T. J. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journals of memory and Language* 59, 434-446.
- [6] Gart, J. J. and Zweifel, J. R. (1967). "On the bias of various estimators of the logit and its variance with applications to quantal bioassay," *Biometrika*, 52, 181-7.
- [7] Light, R. J. and Margolin, B. H. (1971). An analysis of variance for categorical data. *Journal of the American Statistical Association*, 66: 534-544
- [8] Onukogu, I. B. (1985a). Reasoning by Analogy from ANOVA to CATANOVA. *Biom. Journal*, 27: 839-849.
- [9] Onukogu, I. B. (1985b). An analysis of variance of nominal data. *Biom. Journal*, 27: 375-383.
- [10] Onukogu I. B. (2014). Analysis of variance of Categorical data-Nested Designs. *Journal of Statistics: Advances in Theory and Applications* 12: 109-116.
- [11] Rao, M. M. (1960). Some asymptotic results on transformations in the analysis of variance. ARL Technical Note. 60-126. Aerospace research laboratory, Wright-Patterson Air Force Base.
- [12] Singh, B. (1996). On CATANOVA method for analysis of two-way classified nominal data. *Sankhya*, B58(3), 379-388.
- [13] Singh, B. (2004). CATANOVA for analysis of nominal data from Repeated Measures Design *J.ind.Soc. Agril. Statist.* 58(3): 257-268.
- [14] Weiss, D. J. (2009). Nominal analysis of "variance" *Behavior Research Methods*, 41(3), 901-908.
- [15] Winer, B. J., Brown, D. R., and Michel, K. M. (1971). *Statistical principles in experimental design*. New York: Graw-Hill.
- [16] Winsor, C. P. (1948). Factorial Analysis of a multiple dichotomy. *Human Biology*. 20: 195-204.