



Capacity Building of Students' Skills in Data Analysis and Visualization in Colleges of Education in South East Nigeria.

Akamigbo Izuchukwu Stephen

Department of Educational Psychology, (Field of Educational Research, Measurement and Evaluation), Nwafor Orizu College of Education, Nsugbe
E-mail: stevezunna@yahoo.com

ABSTRACT

The study exposed students' in colleges of education in south east Nigeria on the use of computer software and application to carry out data analysis and visualization. The study uses descriptive research design to exposed students on the use of Microsoft excel and statistical package for the social science (SPSS) to perform mean, student t-test and analysis of variance to gain insight on the data analyzed, visualized and interpreted. two research questions guided the study, all students in the colleges of education, south east constitute the population of the study while simple random technique was composed to draw two departments in each of the colleges of education in south east. A total of 450 students participated in the study. The students were exposed on how to carryout basic descriptive statistics like mean, mode and median and inferential statistics like students t-test and Analysis of variance. These statistical analyses are basis for research findings in colleges of education in Nigerian and south east in particular. The findings reveal that colleges of education students in south east uses Microsoft excel and SPSS to perform mean, t-test and analysis of variance.

Key words: Capacity building, Skills, Microsoft Excel, SPSS, Colleges of Education

Introduction

Research work are carried out in colleges of education in Nigeria. Colleges of education are higher institution in Nigeria mandated to offer teacher education programme in Nigeria and produce qualified teachers who will have a minimum teaching qualification and teach at the junior secondary education, pre-primary education or early childhood and care education, adult and non-formal education and special needs education.(FRN, 2014). Research are carried out to provide answers to research problems, research questions and test hypothesis to make inform decisions. The need to answer research questions and test hypothesis to address the early identify problem of the study gives rise to carry out data analysis. Research questions are posers for a study that guide the researchers' pursuit of appropriate sources of information that will lead to the emergence of the desired result of the study. The number, structure and content of research questions of a study are often determined generally by the problem of the study and specific purpose of that study.

Data analysis is the process that summarized data collected. It is the process of using statistical techniques to summarized data; interpret the data to gain insight on the relationship and patterns for decision making. The essence of carrying out data analysis is to provide answers to research questions, make inferences among the variables of the study, interpret the data and make the findings available for generalization. Many students don't have understanding of the data analysis in research even though they were exposed to research methodology in the course of their training. This gap stimulates the researcher to build the capacity skills of students in colleges of education on how to use computer application to carried out data analysis in educational research.

Computer applications available for data analysis are but not limited to Microsoft excel, SPSS, EVIEW and STATA. All these applications house statistical techniques for quantitative analysis. The choice of data analysis to perform is a function of the purpose of the study, design of the study, type of data collected, research question posed and hypothesis. Data visualization is the graphical representation of data to gain insight and make informed decisions. Data visualization is the use chart, graphs, maps to describe relationships, patterns and correlations among set of variables under investigations. The graphical displays are aided with the help of computer application and software. The software under study perform descriptive and inferential statistics to answer and test research questions in educational research. All these software perform different analysis but this study is limited to mean, t-test and analysis of variance which is common among the test statistics often use by researchers in colleges of education.

Microsoft office excel is a Spreadsheet software that is use to analyzed, organized, query and retrieved data. The application performs descriptive and inferential statistics. The foundation of data analyst starts in acquiring a skill or mastering the use of Microsoft excel. In excel package analysis took pack houses all the descriptive and inferential statistics excel can perform

Statistical package for the social science (SPSS) is a software program used by researcher in various disciplines for quantitative analysis and complex data. In research or educational research, the software is used for the analysis of statistical data for computation of analysis of variance, students' t-test and mean scores.

Mean is the average of the set of score in a distribution or dataset. It gives us a good idea of where the center of a dataset is located. In educational research mean is use to compare results of students achievement in a test, scored of respondents in a questionnaire. High mean score shows better achievement or positive response in a questionnaire. To calculate the mean of a dataset in excel, we can use the =Average (Range) function where range is the range of values.

Student t –test compares the means of two independent groups in order to determine whether there is statistical evidence that the associated population mean are significantly different. The independent sample t-test is a parametric test.

Analysis of variance is the analysis of independent variable with two or more levels(Eneja, 2013). It is a test of inference set out to established the weather two or more groups of independent variables are equal or not. Gamst, Meyers and Guarino (2008) noted that ANOVA is one of the most fundamental and ubiquitous univariate methodologies employed by psychologists and behavioral scientists. The design presents the foundation of experimental design including assumptions, statistical significance, and strength of effect and partitioning of the variance. Exploring the effect of one or more independent variables on a single dependent variable as well as two-way and three –way mixed design. Sawyer (2009) defined ANOVA as a statistical tool used to detect differences between experimental group means. The sources of variation are between and within the group means of independent variables. ANOVA test is used to determine the influence that independent variables have on the dependent variable. Statistically speaking analysis of variance splits an observed aggregate mean variability found inside a data set into two parts: systematic factors and random factors. The systematic factors have a statistical influence on the given data set, while the random factors do not. However, this study is tend to expose students on the use of Microsoft excel and SPSS to perform basic analysis of mean, student-t-test and analysis of variance.

Research Questions

The following research questions would guide the study:

1. What is the practical steps to analyze mean, student t score and analysis of variance in a Microsoft excel package.
2. What is the practical steps to analyze mean, student t scores and analysis of variance and visualization in SPSS application?

Research Methodology

The study adopt a descriptive survey design. This study was conducted in colleges of education in south east Nigeria to build students capacity in data analysis using Microsoft excel and Statistical package for the social science (SPSS). The population of this study comprises of students in colleges of education in south east Nigeria. The sample size of this study is made up of 450 students composed through simple random techniques. The instrument was data set extract from research work and analysis was carried out step by step using excel and SPSS to demonstrate mean, students t-test and analysis of variance

Results

Practical steps to carry out mean, students' test and analysis of variance in excel

Mean

To carry out the mean comparison in Microsoft excel do the following:

1. Input the data set in excel row and column
2. Click the data tab
3. Click data analysis
4. Select descriptive statistics and input the range
5. Click on labels in first row
6. Click on summary statistics
7. Click ok the result will appear

Student's t-test

1. Input the data set in excel column and rows
2. Click on data

3. Click on data analysis
4. Select the t-test paired two sample for mean
5. Input the independent variable scores in each range variables
6. Click on labels
7. Choose the output window
8. Click ok

Analysis of variance

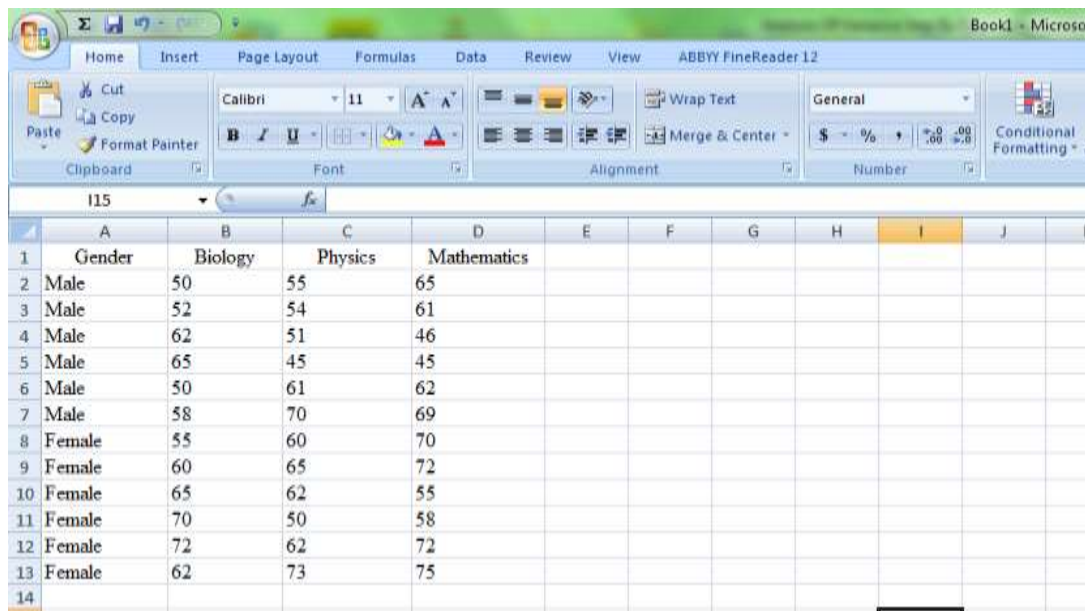
1. Click the data tab
2. Click data analysis
3. Select the ANOVA you want to run: single factors, two-factor with or without replication and click ok
4. Select the input range from the excel work file or workbook
5. Click on labels in first row (it will bring out the labels name)
6. Select the output range
7. Click ok

The practical steps are as follows. Educational researchers want to find out the influence of gender on students achievement in science process skills. The test of science process skills was administered to students and the scores are recorded according to departments and gender as follows.

Data on Students' Achievement on Science Process Skills According To Gender and Departments

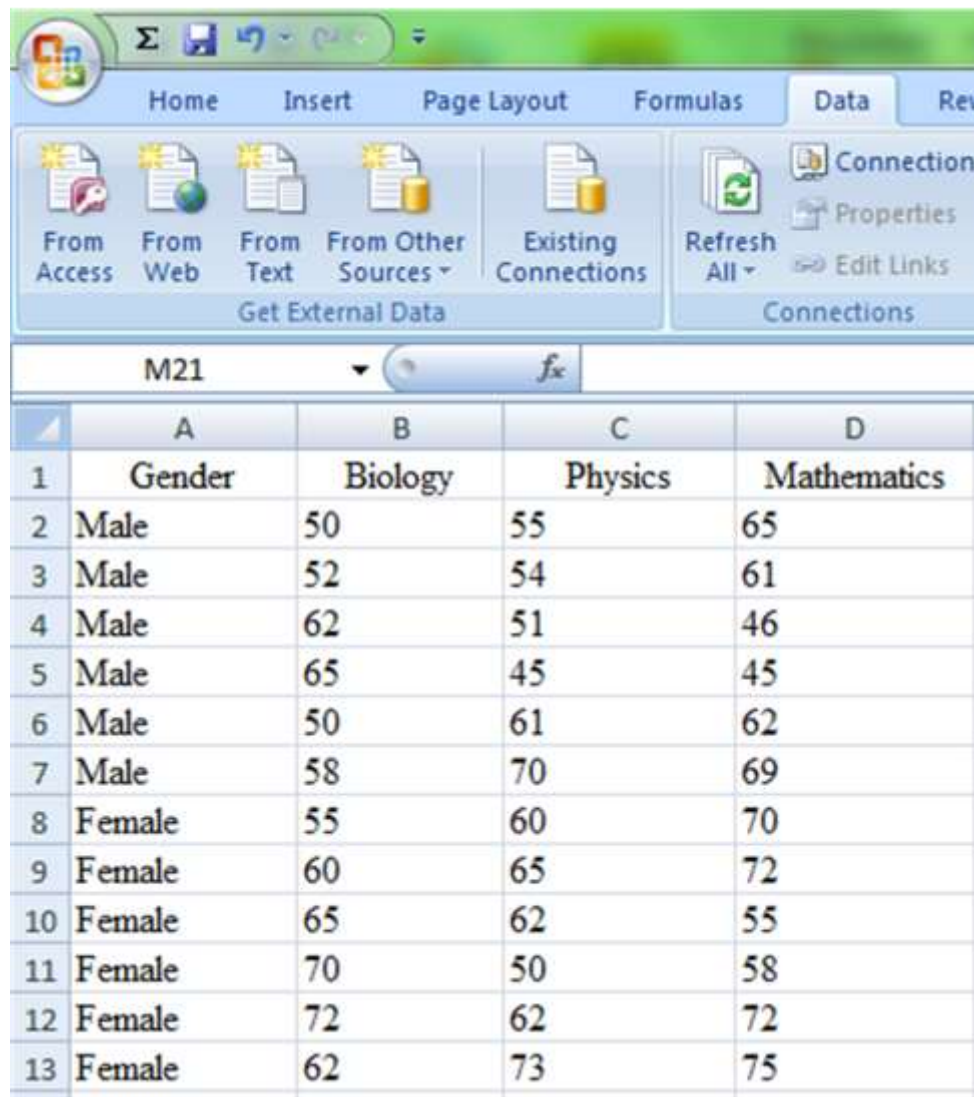
Gender	Biology	Physics	Mathematics
Male	50	55	65
Male	52	54	61
Male	62	51	46
Male	65	45	45
Male	50	61	62
Male	58	70	69
Female	55	60	70
Female	60	65	72
Female	65	62	55
Female	70	50	58
Female	72	62	72
Female	62	73	75

1. Performing the analysis of variance input the scores in excel as shown below



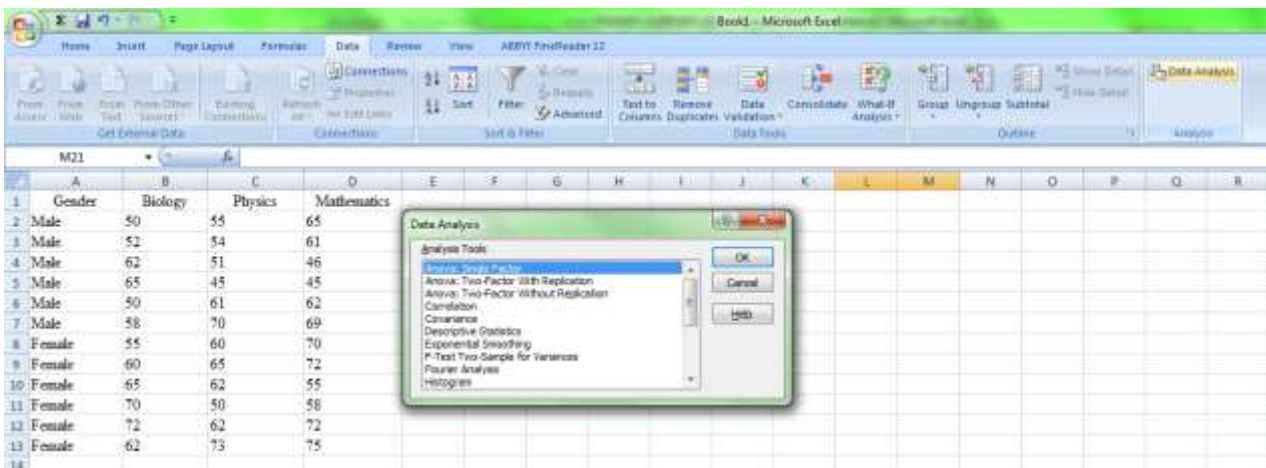
	A	B	C	D	E	F	G	H	I	J
1	Gender	Biology	Physics	Mathematics						
2	Male	50	55	65						
3	Male	52	54	61						
4	Male	62	51	46						
5	Male	65	45	45						
6	Male	50	61	62						
7	Male	58	70	69						
8	Female	55	60	70						
9	Female	60	65	72						
10	Female	65	62	55						
11	Female	70	50	58						
12	Female	72	62	72						
13	Female	62	73	75						
14										

2. Next click on data



	A	B	C	D
1	Gender	Biology	Physics	Mathematics
2	Male	50	55	65
3	Male	52	54	61
4	Male	62	51	46
5	Male	65	45	45
6	Male	50	61	62
7	Male	58	70	69
8	Female	55	60	70
9	Female	60	65	72
10	Female	65	62	55
11	Female	70	50	58
12	Female	72	62	72
13	Female	62	73	75

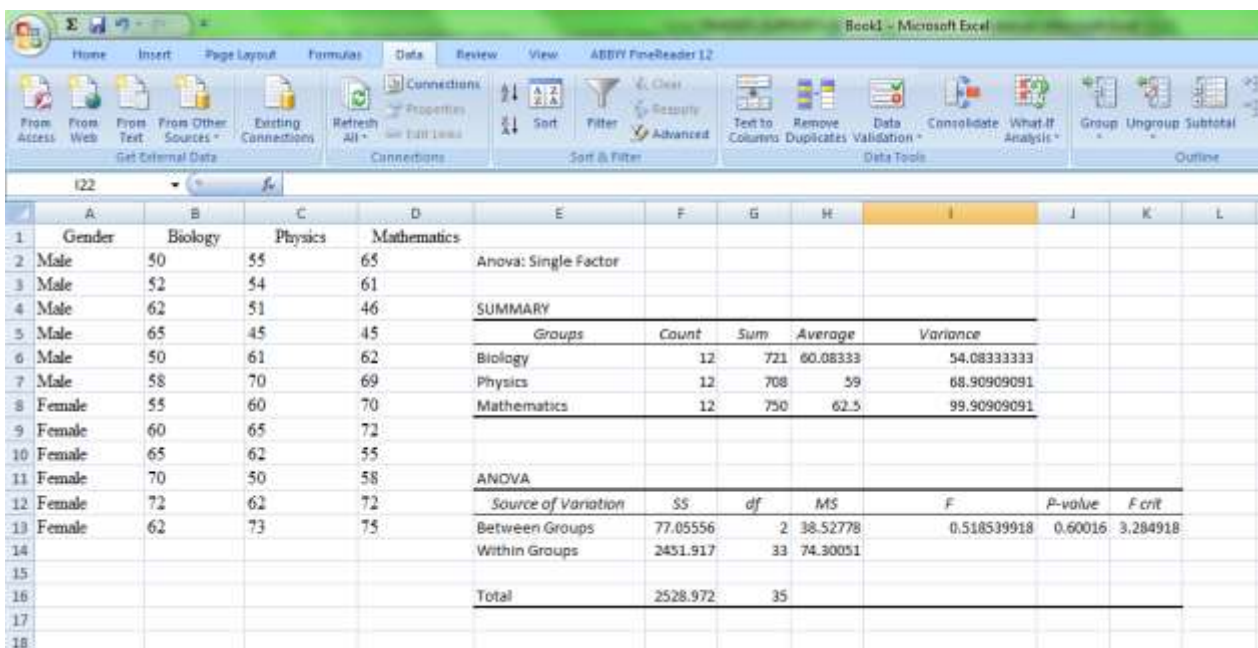
3. Click on data analysis. The Data analysis tool pack dialog box will appear choose ANOVA single factor



On the analysis tool pack select the anova single factor, click the input range (\$B1:\$D13), click on labels in first row and output range (\$E\$2)



Click ok the summary result will display



For mean the result and students t-test data set and results output in excel

	<i>Gender</i>	<i>Biology</i>	<i>Physics</i>	<i>Mathematics</i>	
Mean	Male	50	55	65	output
	Male	52	54	61	
	Male	62	51	46	
	Male	65	45	45	
	Male	50	61	62	
	Male	58	70	69	
	Female	55	60	70	
	Female	60	65	72	
	Female	65	62	55	
	Female	70	50	58	
	Female	72	62	72	
	Female	62	73	75	

	<i>Biology</i>		<i>Physics</i>		<i>Mathematics</i>
Mean	60.08333	Mean	59	Mean	62.5
Standard Error	2.122957	Standard Error	2.396336	Standard Error	2.885439
Median	61	Median	60.5	Median	63.5
Mode	50	Mode	62	Mode	72
Standard Deviation	7.354137	Standard Deviation	8.30115	Standard Deviation	9.995454
Sample Variance	54.08333	Sample Variance	68.90909	Sample Variance	99.90909
Kurtosis	-0.98076	Kurtosis	-0.56601	Kurtosis	-0.60486
Skewness	0.054928	Skewness	0.034328	Skewness	-0.65413
Range	22	Range	28	Range	30
Minimum	50	Minimum	45	Minimum	45
Maximum	72	Maximum	73	Maximum	75
Sum	721	Sum	708	Sum	750
Count	12	Count	12	Count	12

Students –test output

t-Test: Paired Two Sample for Means

	<i>Biology</i>	<i>Physics</i>
Mean	60.08333	59
Variance	54.08333	68.90909
Observations	12	12
Pearson Correlation	-0.09531	

Hypothesized Mean Difference	0
df	11
t Stat	0.323433
P(T<=t) one-tail	0.376217
t Critical one-tail	1.795885
P(T<=t) two-tail	0.752433
t Critical two-tail	2.200985

Practical steps to carry out mean, students' test and analysis of variance in SPSS

Mean

1. Click analyze
2. Descriptive statistics
3. Drag the variables of interest to the dialog box
4. Select mean and standard deviation
5. Click ok result will appear

Student t-test

1. Click analyze > compare means > independent samples t-test
2. Compare means
3. Drag the Independent –sample t-test from the dialog box move the variable /dependent variable to test variable and independent variable to grouping variables
4. Defined the variable (1 ,2)
5. Click ok the result will appear

Analysis of variance

1. Analyze >compare means> one way ANOVA
2. Move dependent variable to dependent list
3. Move independent variable to factor list
4. Click options for visual display and other test
5. Click ok result will appear.

The presentation of the output of the analysis and data in SPSS

<i>Gender</i>	<i>scores</i>	<i>Subject code</i>	<i>subject</i>	<i>Gender code</i>
Male	50	1	biology	1
Male	52	1	Biology	1
Male	62	1	Biology	1
Male	65	1	Biology	1
Male	50	1	Biology	1
Male	58	1	Biology	1
Female	55	1	Biology	2
Female	60	1	Biology	2
Female	65	1	Biology	2

Female	70	1	Biology	2
Female	72	1	Biology	2
Female	62	1	Biology	2
male	55	2	Physics	1
Male	54	2	Physics	1
Male	51	2	Physics	1
male	45	2	Physics	1
male	61	2	Physics	1
Male	70	2	Physics	1
Female	60	2	Physics	2
Female	65	2	Physics	2
Female	62	2	Physics	2
Female	50	2	Physics	2
Female	62	2	Physics	2
Female	73	2	Physics	2
Male	65	3	Maths	1
Male	61	3	Maths	1
Male	46	3	Maths	1
Male	45	3	Maths	1
Male	62	3	Maths	1
Male	69	3	Maths	1
Female	70	3	Maths	2
Female	72	3	Maths	2
Female	55	3	Maths	2
Female	58	3	Maths	2
Female	72	3	Maths	2
Female	75	3	Maths	2

The data will be coded like this in SPSS. However we defined the variable in SPSS. I = biology

2 = physics

3 = maths

1= male

2= female

Mean result

T-TEST GROUPS=gender(1 2)

/MISSING=ANALYSIS

/VARIABLES=scores

/CRITERIA=CI(.9500).

T-Test

[DataSet0]

Group Statistics

gender		N	Mean	Std. Deviation	Std. Error Mean
Scores	male	18	56.7222	8.05719	1.89910
	female	18	64.3333	7.29222	1.71879

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Scores	Equal variances assumed	.455	.505	-2.971	34	.005	-7.61111	2.56141	-12.81652	-2.40570
	Equal variances not assumed			-2.971	33.667	.005	-7.61111	2.56141	-12.81842	-2.40380

ONEWAY scores BY subject

/STATISTICS DESCRIPTIVES

/MISSING ANALYSIS.

Oneway

[DataSet0]

Descriptives

scores								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
biology	12	60.0833	7.35414	2.12296	55.4107	64.7559	50.00	72.00
physics	12	59.0000	8.30115	2.39634	53.7257	64.2743	45.00	73.00
maths	12	62.5000	9.99545	2.88544	56.1492	68.8508	45.00	75.00
Total	36	60.5278	8.50037	1.41673	57.6517	63.4039	45.00	75.00

ANOVA

scores					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	77.056	2	38.528	.519	.600
Within Groups	2451.917	33	74.301		
Total	2528.972	35			

Summary and Conclusion

Students copying research findings from other work to present as their findings threatened the validity of the result presented and leads to false findings, answer to research question, hypothesis and objective of the study. This trend poses serious treat in education and research community hence the need to educate students' on how to carry out data analysis using Microsoft excel package and SPSS is important. However many students don't have technical knowledge on the use statistical software for analysis , this paper was written for students of college of education to gain the skills in analyzing their research data step by step using excel and SPSS. Mean, student t-test and Analysis of variance was chosen to demonstrate the steps, other analysis can be performed in excel and SPSS using similar procedures. The analysis excel can performed include independent sample t test, multiple regression, analysis of covariance, correction analysis, mean, mode, median, median , standard deviation, Z test and many more. Microsoft excel is a necessary software every students or researcher should learn and master, the application is synchronized with other statistical software like statistical package for SPSS. SPSS perform more advance parametric and non-parametric test for researcher in education

References

- Eneja , R. U. (2013). Influence of cognitive ability, gender and location on students' achievement in senior secondary school financial accounting. Unpublished M.ED Thesis University of Nigeria, Nsukka.
- Federal Republic of Nigeria (2014). *National Policy on Education (2014 Edition)*: Lagos: NERDC Press.
- Gamst, G., Meyers, L., & Guarino, A. (2008). *Analysis of Variance Designs* ([edition unavailable]). Cambridge University Press. Retrieved from <https://www.perlego.com/book/1693645/analysis-of-variance-designs-a-conceptual-and-computational-approach-with-spss-and-sas-pdf> (Original work published 2008)
- Israel, P.C & Israel, H.C (2014). Admission guidelines in colleges of education in south east, Nigeria: an analysis of implementation. *Mediterranean Journal of Social Science* 5(17),123-130
- Sawyer, S. F (2009). Analysis of variance: the fundamental concepts . *Journal of manual and manipulative therapy*. Volume 17 (2), 27-38