COURSE GUIDE

MAC 413
DATA ANALYSIS IN COMMUNICATION RESEARCH

Course Team Dr. Nnamdi Tobechukwu Ekeanyanwu (Course Writer) – Covenant University, Ota
Dr. Victor Ayedun-Aluma (Course Editor) – UNILAG
Dr. Chidinma Onwubere (Course Coordinator) – NOUN

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
University Village
Plot 91, Cadastral Zone
Nnamdi Azikiwe Expressway
Jabi, Abuja

Lagos Office
14/16 Ahmadu Bello Way
Victoria Island, Lagos

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Printed 2013, 2018, 2021

ISBN: 978-058-001-8

Printed by: NOUN Press

CONTENTS **PAGE**

Introduction …….. iv
What You Will Learn in this Course …… v
Course Aims ……. v
Course Objectives …… vi
Working through this Course ……... vii
Course Materials and Structure ……... vii
Study Units …….. viii
Textbooks and References ……... ix
Assessment File ……... x
Tutor-Marked Assignment (TMA) …….. x
Final Examination and Grading ……... x
Course Marking Scheme ……. xi
Presentation Schedule and Course Overview …….. xi
How to Get the Most from this Course …… xii
Tutors and Tutorials ……. xiii
Summary …….. xiv

INTRODUCTION

This is MAC 413 Data Analysis in Communication Research. It is designed to teach you how to analyse and interpret results of your research project, especially in communication sciences or studies. The course builds on the knowledge gained in introduction to statistics for social sciences/mass communication and foundation of communication research. These are introductory courses on research and the use of statistics for university students, who enter into the Mass Communication programme. In specific terms, this course is a study in data presentation, analysis and interpretation using tables, graphs, measures of central tendency and dispersion, and other statistical techniques. The use of relevant computer software programmes or statistical packages like the Statistical Package for Social Sciences (SPSS) also received adequate attention. This type of statistical packages aid data analysis, presentation, and interpretation.

Data analysis is a very vital aspect of any research effort. This is because data cannot convey the necessary information needed for decision making on their own. Raw data simply do not speak for themselves; they have to be analysed and interpreted to provide answers to research questions. Un-analysed data is as good as data not collected (Obikeze: 1990).

My experience as a teacher and supervisor of students' research projects and thesis is that the transition from the field work to the stage of analysing data is not always smooth for most of the students. They come back with enormous valid data from the field but how to handle the data to make it make sense is usually a huge challenge for most of them. This underscores the need for a proper course on data analysis which will build on the introductory courses on foundation of mass communication research and introduction to statistics.

Data Analysis in Communication Research is a two-credit-unit course with 24 study units. This Course Guide is for distance learners enrolled in the B.A. Mass Communication programme of the National Open University of Nigeria. The Course Guide provides you with the necessary information about the contents, process, and materials with which to read and understand the subject matter of the course. The guide also specifies the amount of time you are required to spend on each unit of the assignments. It also gives you some guidance on your Tutor-Marked Assignments (TMAs). You are advised to attend the tutorial classes to discuss the areas of challenge with your tutorial facilitators, who will act as guides to you. I strongly recommend that you go through this Course Guide and complete the feedback form at the end before you begin your study of the course. The feedback form must be submitted to

your tutorial facilitator along with your first assignment. This guide also provides answers to several of your questions in case you are still in doubt of anything at this stage of your study. However, do not hesitate to contact your study centre if you have further questions. It is my honest expectation that at the end of this course, you will have become familiar with the basic tools of data analysis in Social Sciences in general and Mass Communication in particular.

WHAT YOU WILL LEARN IN THIS COURSE

The course is designed primarily to teach you how to analyse research results or raw data generated in the course of an empirical investigation of a phenomenon. The whole essence is to be able to meet the objectives set for such specific investigations. Data analysis is a very vital aspect of any research effort. This is because data cannot convey the necessary information needed for decision making on their own. Raw data simply do not speak for themselves; they have to be analysed and interpreted to provide answers to research questions and meet study objectives. According to Obikeze (1990) "un-analysed data are as good as data not collected". Researchers and statisticians have therefore developed a means for making such mass of raw data meaningful through analysis to suit different situations or expectations. As a communication student, however, you should be able to make meaning out of a data set. The goal of this particular course is to get you started and to give you some of the tools you may need and how to apply them in analysing raw data gotten from scientific investigations. Beyond the acquisition of the relevant skills needed to analyse research data, this course will also help refresh the mind of the student on some of the fundamental elements of research in the social sciences in general and mass communication in particular. In summary, the student who goes through this course is expected to first acquire the skills and tools for data analysis in communication studies and second, know how to conduct a scientific investigation, the various types of investigation, the type of data expected from the various types of scientific enquiries and finally how to analyse such raw data so as to make sense of the results.

COURSE AIMS

The aim of this course is to teach you how to analyse research results in the Social Sciences in general and Mass Communication in particular. The course also hopes to help you read and understand tables in order to sharpen your abilities to analyse, interpret, understand, and present social sciences data in particular or standardised formats. This course therefore, is to get you started and to give you some of the tools you may need and how to apply them in analysing raw data gotten from scientific investigations. In summary, this course is meant to provide

you the requisite skills needed to carry out and analyse social data for various purposes. It is also a major aim of this course to refresh your mind on some of the fundamental pillars of scientific investigation so as to help you gather the right kind of data for analysis. The course also examines the use of computer and computer based statistics and analysis software in data analysis and interpretation.

COURSE OBJECTIVES

This course is an exploration in the world of social scientific research, with particular attention to how data gathered for specific studies are analysed to meet specific objectives. It is my expectation that at the end of the course, you should be able to:

- identify and explain the fundamental elements of social science research with particular attention to media or mass communication research
- identify and describe the various kinds of scientific investigation possible in Mass Communication
- describe scientific data and how to analyse them
- analyse research findings using appropriate test statistics
- present statistical data using appropriate tables, graphs and charts
- interpret data results through the help of tables, graphs and charts
- use SPSS for data analysis and interpretation
- calculate the measures of central tendency and dispersion as key tools in data analysis
- present, understand and analyse research results of different research methods
- distinguish between qualitative research techniques and quantitative research methods and the different methods of analysing the data that come from the two research gathering methods
- describe the scientific details involved in data processing and management as preludes to data analysis
- state the major differences between research questions and hypotheses and calculate the chi-square, t-test etc using appropriate text statistics
- discuss the fundamental principles of significance testing as a core component of data analysis in communication research.

WORKING THROUGH THIS COURSE

To make the most out of this course, you are expected to read the study units and other materials provided by NOUN. Each unit contains Self-Assessment Exercise (SAE). At certain points in the module, you are required to carry out the assignments, which will be marked by assessors appointed by the university. Remember, all components of this course contribute to your all-round success. So, take your time to read and study the units very well in order to successfully derive the best from the course.

I will also advise you to make concerted efforts towards locating the recommended texts listed hereunder and read them. This course material will never take the place of those recommended literature for further reading. They will help equip you for an all round experience and exposure in the exciting world of research and data analysis in communication studies in particular.

COURSE MATERIALS AND STRUCTURE

The major materials you will need for this course are:

1. A Course Guide
2. Study units that have been broken down to 24
3. Self-Assessment Exercise (SAE) file
4. Tutor-Marked Assignments (TMAs) file
5. Relevant textbooks including the ones listed under each unit

In addition to the materials listed above, you are advised to read through this Course Guide to familiarise yourself with the structure of the course; the study units as well as attempt all SAEs and TMAs. You are also expected to consult most of the recommended resource materials for further reading. Each unit contains Self-Assessment Exercises, and at points in the course you are required to submit assignments for assessment purposes. At the end of the course, there is a final examination. The course should take you about 60 hours to complete. You have to draw up your own timetable and allocate time to complete each study unit in order to complete the course successfully and on time. All the components of the course are listed and explained below.

STUDY UNITS

This course is structured in a six-module compact with 24 study units thus:

Module 1 Introduction: Research Process and the Scientific Procedure

Unit 1 Communication Research and the Scientific Process
Unit 2 Characteristics of Research
Unit 3 Research Classification
Unit 4 Measurement

Module 2 Computers and Data Analysis

Unit 1 Data Processing and Management
Unit 2 The Use of Computer in Data Processing and Analysis
Unit 3 Statistical Package for Social Sciences (SPSS) program
Unit 4 Data Analysis in Communication Research

Module 3 Qualitative Data Analysis

Unit 1 Overview of Qualitative Research and Types
Unit 2 Data Analysis in Focus Group Discussion
Unit 3 Analysing Case Study
Unit 4 Analysing Data from Observation Technique

Module 4 Quantitative Data Analysis

Unit 1 Introductory Overview of Quantitative Analysis
Unit 2 Procedure for Quantitative Analysis
Unit 3 Construction of Frequency Distribution Tables
Unit 4 Survey Data Analysis and Statistics

Module 5 Quantitative Data Analysis: Focus On Content Analysis and Experiments

Unit 1 Introductory Overview of Content Analysis
Unit 2 Processing, Analysing and Writing the Content Analysis Data and Report
Unit 3 Introductory Overview of Experimental Research
Unit 4 Processing and Analysing Experimental Data

Module 6 Statistics and Data Analysis

Unit 1 Measures of Central Tendency
Unit 2 Measures of Dispersion
Unit 3 Recommendations for Appropriate Use of Statistics
Unit 4 Significance Testing

TEXTBOOKS/REFERENCES

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Champion, D. J. (1981). *Basic Statistics for Social Research*. NY: Macmillan.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Nwabuokei, P. O. (1999). *Fundamentals of Statistics*. Enugu: Koruna Books

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Okeke, M.C. (1997). *Statistics: A Business Approach*. Onitsha: Onwubiko Prints Ltd.

Otokiti, S. et al. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martin's.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Spiegel, M. R. (S.I. Edition). *Theory and Problems of Statistics*. New York: McGraw-Hill

Wilson, D. et al. (2008). Communication Research. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth

ASSESSMENT FILE

An assessment file and a marking scheme will be made available to you. In the assessment file, you will find details of the assignments you need to do and submit to your assigned tutor for grading. The assessment for this course is divided into two namely: TMAs and the written examination. Both of them will total 100% at the end of the course. The assignments should be submitted to your tutor for formal assessment in accordance with the lifelines stated in the presentation schedule and the assessment file. The work you submit to your tutor for assessment will count for 30% of your total score while the written examination will account for the remaining 70%.

TUTOR- MARKED ASSSIGNMENTS (TMAs)

You will have to submit about 24 TMAs. This means one TMA for each of the units of the entire course structure. You are required to attempt all the questions, and you will be assessed on all of them but the best three performances from the (TMAs) will be used for your 30% grading. When you have completed each assignment, send it together with a tutor-marked assignment form to your tutor. Make sure each assignment reaches your tutor on or before the lifeline for submissions. If for any reason, you cannot complete your work on time, contact your tutor with a valid explanation on why you need an extension. Failure to do this may result in unpleasant and avoidable situations.

FINAL EXAMINATION AND GRADING

Since this is a two-unit course, the final examination will be a test of two hours. All areas of the course will be examined. It is in your own interest to read the entire units all over again before your written examination. As earlier mentioned, the final examination will attract 70% of the total course grade. This should not be taken for granted. The examination will consist of questions, which reflects the kinds of SAEs and TMAs you have previously dealt with in the course.

COURSE MARKING SCHEME

The following table lays out how the actual course mark allocation is broken down.

S/N	ASSESSMENT	MARKS
	The best in Module One (Assignments 1-4)	5%
	The best in Module Two (Assignments 5-8)	5%
	The best in Module Three (Assignments 9-12)	5%
	The best in Module Four (Assignments 13-16)	5%
	The best in Module Five (Assignments 17-20)	5%
	The best in Module Six (Assignments 21-24)	5%
	Total for TMAs	**30%**
	Final Written Examination	**70%**
	Overall Total	**100%**

COURSE OVERVIEW AND PRESENTATION SCHEDULE

Module	Unit	Title of Module	Weeks Activity	Assessment
ONE	1-4	Introduction: Research Process And The Scientific Procedure		
	Unit 1	Communication Research and the Scientific Process	Week 1	Assignment 1
	Unit 2	Characteristics of Research	Week 1	Assignment 2
	Unit 3	Research Classification	Week 2	Assignment 3
	Unit 4	Measurement	Week 2	Assignment 4
TWO	1-4	Computers And Data Analysis		
	Unit 1	Data Processing and Management	Week 3	Assignment 1
	Unit 2	The Use of Computer in Data Processing and Analysis	Week 3	Assignment 2
	Unit 3	Statistical Package for Social Sciences (SPSS) Program	Week 4	Assignment 3
	Unit 4	Data Analysis in Communication Research	Week 4	Assignment 4
Three	1-4	Qualitative data analysis		
	Unit 1	Overview of Qualitative Research and Types	Week 5	Assignment 1
	Unit 2	Data Analysis in Focus Group Discussion	Week 5	Assignment 3
	Unit 3	Analysing Case Study	Week 6	Assignment 4

	Unit 4	Analysing Data from Observation Technique	Week 6	Assignment 5
Four	1-4	Quantitative data analysis		
	Unit 1	Introductory Overview of Quantitative Analysis	Week 7	Assignment 1
	Unit 2	Procedure for Quantitative Analysis	Week 7	Assignment 2
	Unit 3	Construction of Frequency Distribution Tables	Week 8	Assignment 3
	Unit 4	Survey Data Analysis and Statistics	Week 8	Assignment 4
Five	1-4	Quantitative data analysis: focus on content analysis and experiments		
	Unit 1	Introductory Overview of Content Analysis	Week 9	Assignment 1
	Unit 2	Processing, Analysing and Writing the Content Analysis Data and Report	Week 9	Assignment 2
	Unit 3	Introductory Overview of Experimental Research	Week 10	Assignment 3
	Unit 4	Processing and Analysing Experimental Data	Week 10	Assignment 4
Six	1-4	Statistics and data analysis		
	Unit 1	Measures of Central Tendency	Week 11	Assignment 1
	Unit 2	Measures of Dispersion	Week 11	Assignment 2
	Unit 3	Recommendations for Appropriate Use of Statistics	Week 12	Assignment 3
	Unit 4	Significance Testing	Week 12	Assignment 4
		Revision	2 weeks	
		Written Examination	1 Week	
		Total Weeks	15 Wks	

HOW TO GET THE MOST FROM THIS COURSE

The Open University system is a unique system whereby the study units replace the traditional university lectures. You will therefore be required to study the units on your own. However, you may arrange to meet with your assigned tutor for tutorials on an optional basis at the study centre. You can also locate like-minded course mates and have interactive sessions with them. This interaction could be facilitated virtually.

You will have assignments at the end of every **m**odule and **u**nits. This is not different from the traditional university system where a **l**ecturer might give you some reading to do. The study units will tell you where to read, and your text materials or recommended books. You have been given enough in this course so none of you will have any excuse not to do well in this course. Each of the study units follows a common format. The first item is an introduction to the subject matter of the unit, and how a particular unit is integrated with the other units and the course as a whole. Next to this is a set of learning objectives. These objectives will help you assess your comprehensive of the unit per time. They are also meant to guide your study so as to give you an all round learning experience. The main body of the unit guides you through the required reading from other sources. This will usually be either from your recommended books or from a particular reading section.

In summary, this structure is designed to bring out the best in a distance learner. However, you may still want to interact with a tutor over controversial, confusing or any other topic or issue you are finding it difficult to comprehend yourself. NOUN has made adequate arrangement to see that this is resolved in the interest of the student. This is the major reason why you must patroni**s**e the **s**tudy **c**entres. They have been so designed to take care of your academic needs.

FACILITATION/TUTORS AND TUTORIALS

NOUN headquarters will provide you with information relating to the tutor and tutorials at the appropriate time. Your tutor will mark and comment on your assignments, keep a close academic watch on your progress as well as assist you resolve any difficulties you might encounter from time to time. You must therefore take your TMAs to the study centre well before the due date (at least two working days before the expiration of the lifeline). The TMAs will be marked by your tutor and returned to you as soon as possible.

In the NOUN guidelines, you are expected to contact your tutor if you need help over the following issues:

- you do not understand any part of the study units or the assigned readings
- you have difficulty with the exercises and assignments
- you have a question or a problem with your tutor’s comments on an assignment or with the grading of an assignment
- you have a question or problem with any part of the Course Structure which you think has not been adequately discussed in this Course Guide.

At NOUN, it is also expected (in your own interest) that you make concerted efforts to attend tutorials. This is the only chance to have face-to-face contact with your tutor and ask pending and pertinent questions which are answered immediately. You can raise any problem encountered in the course of your study during the tutorial sessions. To gain the maximum benefit from the course tutorials, it is advised you prepare a question list before attending them. You will learn a lot from being an active participant in the discussions at the study centres or during tutorial sessions.

SUMMARY

This course aims at preparing you to move from data gathering to data analysis. The primary aim is to develop the requisite skills and knowledge base to analyse and interpret empirical data gathered for a particular study or investigation. This is particularly instructive because the movement from data gathering or field work to data analysis is not usually smooth for most undergraduate and even graduate students. They return with a mass of valid data but know very little of what else to do. Furthermore, it is well observed that raw data cannot make sense until someone refines and analyse them to make sense. This is the primary focus of this course – refine and analyse data so that they become more meaningful and useful in meeting research objectives and other expectations.

I wish you success with this course and hope that at the end of it, you will be able to practically move over to the stage of data analysis after data gathering with the required ease and smoothness associated with knowledge and skill.

MAIN COURSE

CONTENTS **PAGE**

Module 1	**Introduction: Research Process And The Scientific Procedure …….**	**1**
Unit 1	Communication Research and the Scientific Process	1
Unit 2	Characteristics of Research	7
Unit 3	Research Classification	13
Unit 4	Measurement ……	18
Module 2	**Computers and Data Analysis ……..**	**23**
Unit 1	Data Processing and Management ……..	23
Unit 2	The Use of Computer in Data Processing and Analysis ……..	33
Unit 3	Statistical Package for Social Sciences (SPSS) Program …….	41
Unit 4	Data Analysis in Communication Research ……	47
Module 3	**Qualitative Data Analysis ……..**	**61**
Unit 1	Overview of Qualitative Research and Types ….	61
Unit 2	Data Analysis in Focus Group Discussion ……	72
Unit 3	Analysing Case Study ……	78
Unit 4	Analysing Data from Observation Technique …..	84
Module 4	**Quantitative Data Analysis ……..**	**91**
Unit 1	Introductory Overview of Quantitative Analysis ……..	91
Unit 2	Procedure for Quantitative Analysis ……	96
Unit 3	Construction of Frequency Distribution Tables …….	101
Unit 4	Survey Data Analysis and Statistics …….	108

Module 5	**Quantitative Data Analysis: Focus On Content Analysis and Experiments …….. 110**	
Unit 1	Introductory Overview of Content Analysis …..	110
Unit 2	Processing, Analysing and Writing the Content Analysis Data and Report …….	115
Unit 3	Introductory Overview of Experimental Research ……..	122
Unit 4	Processing and Analysing Experimental Data	128
Module 6	**Statistics and Data Analysis …….. …… 133**	
Unit 1	Measures of Central Tendency …….	133
Unit 2	Measures of Dispersion …….	144
Unit 3	Recommendations for Appropriate Use of Statistics ……..	156
Unit 4	Significance Testing ……..	160

MODULE 1 INTRODUCTION: RESEARCH PROCESS AND THE SCIENTIFIC PROCEDURE

Unit 1 Communication Research and the Scientific Process
Unit 2 Characteristics of Research
Unit 3 Research Classification
Unit 4 Measurement

UNIT 1 COMMUNICATION RESEARCH AND THE SCIENTIFIC PROCESS

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
  3.1 What is Communication Research?
  3.2 Nature and Scope of Media Research
  3.3 The Scientific Process in Research
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

This course is all about data analysis in communication research. In other words, how is data gathered for any specific research in communication analysed? Does this analysis differ from known social science models? Are analysis techniques in media or communication research distinctively unique to the discipline? These are valid questions that we will attempt to answer throughout the duration of this course. However, it will not be proper to jump into data analysis without reviewing what the research process is all about so as to bring you the student into the same frame of reference with the course administrators. In essence, this unit is primarily meant to review the research processes again so as to establish the obvious relationship between the social science or scientific research and procedure and that of media or communication research. Communication research follows the scientific procedure. What is this scientific procedure? This is partly the focus of this unit.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define research with specific emphasis on communication research
- identify the nature and operational scope
- explain the scientific process in research and the steps to follow if one is doing an empirical investigation.

3.0 MAIN CONTENT

3.1 What is Communication Research?

To properly define or explain communication research, we must first define research itself. The simple reason being that communication research is just about the type of research that has media or communication matters and issues as its main focus.

Research is simply a thorough, systematic, organised and purpose driven search for knowledge and facts to support a position or argue a phenomenon. According to Obikeze (1990:3), research may be defined as a systematic process of investigation or inquiry carried out in accordance with laid down (scientific) procedures for the purpose of finding answers or solutions to a set of defined problems or perplexing issues.

On their part, Wilson, Esiri and Onwubere (2008:20) explain research as knowledge that can be explained or verified through some procedure. According to them, "for one to engage in any research, the expected outcome of the research must be important otherwise there will be no need for the research. Consequently, all research activities start from problems that require solutions. This may sometime originate from an idea, a puzzle or simply the wish to explore our knowledge about simple issues, phenomena, situations or societies."

The *New Oxford Dictionary* also defines research as "careful, systematic, patient study and investigation in some field of knowledge undertaken to discover or establish facts and principles". This also tallies with the *Webster's Ninth New Collegiate Dictionary* definition of research thus: "investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts or practical application of such new or revised theories or law".

This brings us to a fact that Wilson, Esiri and Onwubere (2008:20) tried to explain: research has two sides thus:

- There are two sides to any scientific enterprise namely theory and research. Both are essential in fact finding. Theory is a way of making sense out of our world, a way of explaining things that seem puzzling. According to Tichenor and McLeod (1970) quoted by Stempel and Westly (1989:30) a theory is a tentative explanation invented to assist in understanding some small or large part of the reality around us. Theories arise from observation, often systematic, sometimes causal and occasionally accidental. However, to be scientific a theory must, at least in principle, be capable of being proved wrong.

From the foregoing, we can infer the meaning of communication research. Communication or media research is the type of research that deals with the functioning of the media or communication system of a country. It focuses on the structure, content, application, and influences of the communication system and how the interface with other related disciplines comes to bear on the overall systematic structure of the media. Features of the various valid definitions or research include:

- a research activity is planned and well thought – out
- a research activity is purposeful and aimed at achieving well defined and specific objectives
- a research activity is ordered, systematic and follows well known and clearly laid-out procedures. This ensures replicability and generalisability.

3.2 Nature and Scope of Media/Communication Research

Communication or media research is the type of research that deals with the functioning of the media or communication system of a country. It focuses on the structure, content, application, and influences of the communication system and how the interface with other related disciplines comes to bear on the overall systematic structure of the media.

According to Wilson, Esiri and Onwubere (2008), media research is the application of scientific method to the study of the functioning and application of the mass media. This study of the media is not isolated because media research is intrusively interwoven with other disciplines. This is because the media are linked with other phenomena of life. Every social or even pure scientific phenomenon has a human angle. Anything with a human angle definitely has a communication angle. According to Barzum and Graff (1970) as cited in Wilson, Esiri and

Onwubere (2008), the concerns of media research include the followings:

- media messages and their origins
- functions and purposes of media message
- media channels, languages and codes
- media content, references and information types
- media audiences
- effects of media messages, intended and unintended
- media noise and feedback
- media technologies
- media regulation
- media ownership and control
- media management etc.

SELF-ASSESSMENT EXERCISE

i. Describe the two sides of research you are familiar with.
ii. Explain two valid definitions of research.

3.3 The Scientific Process in Research

The scientific method is a process of deliberate and controlled observation with the distinctive criterion of objectivity. It emphasises quantification, logical exposition, controlled empirical testing, replicability of findings, and inter-subjectivity. It, therefore, formally rejects the influence of authority precedent, degree of current acceptance, and other such influence internal selective factors might have (Cohen & Nigel: 1934).The scientific process in research includes:

1. **Controlled Experimentation**, preferably in a laboratory with a view to systematic reduction of sources of variation and error
2. **Objectivity** and lack of personal or emotional attachment with the study subjects (ethical neutrality)
3. **Sound theoretical foundation** and construction of complex (grand) theories
4. **Predictability** and generalisability of results (from sample to population)
4. **Systematic documentation of procedures**. The essence of a systematic documentation of research procedures is to make it easy and possible for other researchers to replicate such procedures used in a particular study and aim at similar or same result.

The scientific method also has some related characteristics that distinguish it from non scientific methods. These characteristics as identified by Cohen and Nigel (1934) include:

1. it is **empirical** which means it is based on scientific observation
2. It is **theoretical,** which means it attempts to summarise complex scientific observations using grand theories in a logical manner
3. It is also **cumulative,** which means theories build upon one another sometimes making new ones, using theories to correct others, and retaining the older ones
4. It is also **non-ethical,** which means it does not judge a particular social action as good or bad; it merely explains them in the light of theory and data analysed.

4.0 CONCLUSION

The scientific process in research is core to research outcomes and to what use research efforts are put. It also helps to validate the processes involved. This is the major reason Obikeze's (1990:3) description of research as a systematic process of investigation carried out in accordance with laid down (scientific) procedures for the purpose of finding answers or solutions to a set of defined problems or confusing issues becomes very instructive here. It must also be pointed out here that the establishment of a sound research base is vital to data analysis. A well carried out data analysis procedure cannot change the status of a poorly organised investigation.

5.0 SUMMARY

This unit discussed the definitions and meanings of research, the scientific procedure, communication research and the concerns of media research. The major objective is to set the stage for a proper exploration into the realm of data analysis in communication research.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss any two commonly accepted procedures involved in scientific research.
ii. Discuss five major concerns of media research in a developing world.

7.0 REFERENCES/FURTHER READING

Maner, Martin. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins

Sobowale, Idowu. (2009). *Scientific Journalism* (2nd ed.) Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

Onabajo, O. (2004). *Introduction to Broadcasting*. Lagos: Gabi Concepts.

UNIT 2 CHARACTERISTICS OF RESEARCH

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Characteristics of Research
    3.2 An Overview of the Development of Communication Research
        3.2.1 Major Events or Social Forces that Influenced the Growth of Media Research
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

In this unit, we continue our introductory treatment of the subject matter of communication research and how data gathered for it are analysed. Two vital information we must acquire before getting neck deep into data analysis and which are critical to analysis of data in communication research are core features of research and the development of communication/media research. A clear understanding of these two issues will help us know how to handle research, especially communication research and every issue related to it.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- list, identify and explain the characteristics of research
- discuss (in brief) the development of communication/media research.

3.0 MAIN CONTENT

3.1 Characteristics of Research

Research in general follows a particular pattern and is usually guided by a set of standards and features. We have partly discussed some of these in the first unit of this module. We will complete their discussion in the remaining two units of this introductory module. However, for this unit, we are interested in the characteristics of research. There are many such

characteristics of research but for a communication/social science student, about six core characteristics are vital to your further exploration in data analysis. This includes:

a. **Research is Systematic and Procedural**
Research follows a definite set of procedures. These procedures are standards that are generally accepted the world over. Anyone who engages in research is therefore expected to adhere strictly to such procedures. This adherence has made research acquire the notion of being systematic and organised in a peculiar and particular way.

b. Research is Logical
Wilson, Esiri and Onwubere (2008) see logic as a system of constructing proofs to get reliable confirmation of the truth of a set of hypotheses. Generally, they see it as the rational way of drawing or arriving at a reasonable conclusion on any subject matter through research. Hence in research, according to Wilson, Esiri and Onwubere (2008), the following logical laws are maintained:

- hypotheses derived are expressed in a formal language
- the allowable steps of inference are codified formally so that well formed proofs are obtained
- the permitted inferences and conclusions are sound.

c. A Research Activity is Purposeful, Well Planned and Well Thought-Out
Research has the characteristics of being purposeful and aimed at achieving well defined and specific objectives. A research activity is ordered, systematic and follows well known and clearly laid-out procedures. This ensures replicability and generalisability.

d. Research is Reductive
The reductive nature of research makes it possible to summarise complex observations in logically related propositions which attempt to explain a subject matter. According to Wilson, Esiri and Onwubere (2008), observations are converged in a way that irrelevant variables are excluded while relevant variables are included. Hence, research has the characteristics of controlling the flux of things and establishing facts (Wilson, Esiri & Onwubere, 2008).

e. **Research is Empirical**
Empiricism suggests that research is objective, observed, experimental, experiential, and pragmatic in nature. According to Wilson, Esiri and Onwubere (2008), "A common image of 'research' is a person in a laboratory wearing a white overcoat, mixing chemicals or looking through a microscope to find a cure for an exotic illness. Research ideas are accepted and rejected

based on evidence. Hence, one of the most outstanding feature or characteristics of research is its empirical nature".

f. Research is Replicable and Generalisable

Wilson, Esiri and Onwubere (2008) see replication as a critical step in validating research to build evidence and to promote use of findings in practice. Replication involves the process of repeating a study using the same methods with the assurances that you will get same or similar results. According to Wilson, Esiri and Onwubere (2008), replication is important for a number of reasons. These include:

- assurance that results are valid and reliable
- determination of generalisability or the role of extraneous variables
- application of results to real world situations
- inspiration of new research combining previous findings from related studies; and
- assurances that the schema used is available, reusable and reliable in terms of producing same or similar results.

On the other hand, generalisability ensures that the findings of a particular research conducted using a selected sample are easily implied or generalised to the larger population. To meet the demands of this characteristic, researchers make sure their research samples are representative of the main population. This representativeness ensures generalisability of findings to the population.

SELF-ASSESSMENT EXERCISE

i. Identify and discuss four core features of research in the 21st century.
ii. Discuss two valid reasons why you think generalisability should be a core feature of research.

3.2 An Overview of the Development of Communication Research

Wimmer and Dominick (2011:6-7) note that mass media research has evolved over the years by taking definable steps to reach what we now regard as media research. The authors therefore identified four stages in the development of communication research thus:

a. **Stage one:** this stage featured the early interest in the medium. Researchers started asking pertinent questions about what the medium is all about, how the medium operates, which technology is involved in the medium, what roles or functions it plays or performs, how much will it eventually cost users and how it is

different or similar to what was already in existence. These were the valid posers that came up in the first phase or stage of the development of mass media research.

b. **Stage two:** the focus of research at this stage changed from the initial interest in the medium to the uses and the users of the medium. So, the research questions that defined the studies at this stage also changed to issues about how and why people use the medium. The audience profile was also a major feature of this phase in the development stage.

c. **Stage three:** the development of media research at this stage took a more dramatic and elevated turn when researchers began detailed enquiry on the social, psychological, and physical effects of the medium. Again, the research questions changed to how much time people spent with the medium, and how the medium affects them generally and in other specific ways.

d. **Stage four:** stage four actually marked a remarkable phase in the development of media research. During this stage, there were remarkable empirical evidences to show that the medium actually has effects on the audience and that beyond effects; the medium also plays unique roles and performs critical functions to the society. This prompted researchers to change the research focus to issues on how to improve the medium so that the medium will offer more and acceptable services to more diverse audiences.

3.2.1 Major Events or Social Forces that Influenced the Growth of Media Research

Wimmer and Dominick (2011:8-9) identified four major events and social forces that influenced the growth of mass media research thus:

- the need to understand war time propaganda as the fallout of the First World War
- the interest shown by advertisers and potential advertisers alike in the early 1950s and 1960s that media research data are useful tools in persuading customers and potential consumers to buy their products and services
- the increasing interest of citizens in the effects of the media on the various publics especially the children. This was the era when research on media violence and sexual content became a research trend all over the world.
- of course, increased competition among the media for the ever increasing advertising dollars was the fourth social force that influenced the growth of research in the mass media.

4.0 CONCLUSION

To get the ultimate benefits from research, a researcher must make sure the characteristics of research are adequately taken note of. This is very vital even in data analysis. Any aspect of the core research features that is underplayed or worse still ignored, comes back to haunt the researcher at the analysis stage. Be sure to also know that peers do not waste time dismissing the outcome of such so-called research that underplayed significant characteristics of research.

5.0 SUMMARY

This unit specifically dealt with the characteristics of research. We discussed six that are most relevant to the communication researcher. We also had a brief overview of the growth of media research and that factors and social forces that influenced the development.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss two core characteristics of research and justify their place in social science and mass media research.
ii. Briefly explain how the issue of propaganda affected the growth of mass media research.
iii. Research characteristics could change from discipline to discipline, true or false? Please, justify your position.
iv. Highlight one major event that happened in each of the identified stages in the development of communication research.

7.0 REFERENCES/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martin.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd ed.) Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). "Communication Research". Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 3 RESEARCH CLASSIFICATION

CONTENTS

1.0 Introduction
2.0 Objective
3.0 Main Content
    3.1 Research Classification
        3.1.1 Classification by Practice
        3.1.2 Classification by Measurement
        3.1.3 Classification by Method Used in Gathering Data
        3.1.5 Classification by Discipline
        3.1.5 Other Research Classification Types
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

In the previous two units, we discussed the meaning of research and identified some key research characteristics. In this unit, we will identify the various types of research and the labels given to them. These labels are usually influenced by the nature of the study and the purpose the researcher wants to achieve with such research. Researchers are not all agreed on these labels or groupings. However, I must warn here that these categories are just convenient groupings that are generally accepted because of the common traits found in their structure, otherwise there is really no hard and fast rule in arriving at the classifications presented in this unit as one type of research can easily overlap and fit into more than one (Wilson, Esiri & Onwubere, 2008).

2.0 OBJECTIVE

At the end of this unit, you should be able to:

- mention and explain the different types/methods of research used in communication research.

3.0 MAIN CONTENT

3.1 Research Classification

3.1.1 Classification by Practice

One way scholars have classified research in communication is to look at it from the perspective of practice. From this school of thought, communication research could be divided into two thus:

- Academic research
- Applied research.

A. Academic Research

Academic research by practice is the type of research faculty and students of higher institutions of learning conduct in their institutions to meet certain career and graduation demands respectively. It is usually theoretical in nature. In other words, it is conducted for academic purpose rather than for its intrinsic values for the society as a whole (Wilson, Esiri & Onwubere, 2008).

Wilson, Esiri and Onwubere (2008) further note that notwithstanding their description of this type of research, that the assumption that it has no intrinsic value is not entirely true because situations have arisen when at later dates someone who obviously consults such studies for literature for a current study have found something of value in such academic research.

B. Applied Research

Applied Research is the opposite of academic research. This research usually sets out to address specific needs in the society at large. It is research conducted for its intrinsic values rather than to meet some obscure academic expectations. This type of research is also enterprising in nature because of the availability of funds to support it. When compared to academic research, most companies and funding agencies will always have a preference for applied research than academic research in their sponsorship budgeting. In essence, there are usually enough grants, funds and resources to support of fully sponsor applied research because of the value it holds.

SELF-ASSESSMENT EXERCISE

Justify the various classification of research.

3.1.2 Classification by Measurement

The way data for research is generated and measured could also be used to classify it. This leads us to the classification by measurement. In this category, we have two types as presented below:

- Quantitative Research
- Qualitative Research

A. Quantitative Research

Quantitative research takes numerical values and uses serious statistical tools for its measurement. The research is therefore designed to yield numerical data or expected to turn the variables into numbers (Wilson, Esiri & Onwubere, 2008). According to Wilson, Esiri and Onwubere (2008), quantitative research "is concerned with how often a variable is present and generally uses figures to communicate this amount. In other words, the quantitative approach involves the collection of numerical data in order to explain, predict and/or control the phenomena of interest. Data analysis in quantitative research is mainly statistical or deductive process". Quantitative research techniques include field experiment, survey, and content analysis.

B. Qualitative Research

Qualitative research is the opposite of quantitative research. While quantitative takes numerical values, qualitative does not. Rather, it gives a more detailed and in-depth analysis of the subject as a result of the closer interaction with the subject of investigation. Wilson, Esiri & Onwubere (2008) observe that qualitative research involves the collection of extensive narrative data in order to gain insights into the phenomena of interest. According to them, data analysis in qualitative research "involves the coding of data and production of a verbal synthesis or inductive process. In other words, it does not depend on the measurement of variables or research elements." Qualitative research methods include focus group discussion (FGD), in-depth interview, field observation, case study approach, historical analysis and ethnography.

3.1.3 Classification by the Method Used in Gathering Data

Based on actual practice and literature review, this type of classification is the most popular in the field of mass communication and most other disciplines in the social sciences and humanities. This is because no matter how research is classified, it is fundamentally identified by the method used in collecting the data for the study (Wilson, Esiri &

Onwubere, 2008). By the method used in data gathering for a particular research, the following methods, already identified as either quantitative or qualitative suffice as examples: surveys, field experiments, content analysis, Interviews, focus group discussion (FGD), field observation, ethnography, historical research and case study approach.

3.1.4 Classification by Discipline

Research could also be classified by the discipline or subject orientation. Here, the subject matter becomes the focus and label for such research endeavour. Using this perspective, we have the following types of research: communication research, media research, public relations research, advertising research, social science research, clinical research, marketing research, operations research, legal research, population research, psychological research, political research, and opinion research etc.

3.1.5 Other Research Classification Types

There are other research classification types that do not fall under the categories already discussed in this unit. Some do and even overlap to more than two different levels of classification. However, we will merely list them here because this course is not about communication research *per se*; it is about data analysis in communication research.
The other research classifications that we are yet to identify under any of the existing categories include: longitudinal research, administrative research, critical research, exploratory research, ethno methodological research, primary research, social research, cultural research and secondary research amongst others.

4.0 CONCLUSION

Research could be classified in different ways by different people and still be valid. The key is to understand what one wants to achieve. The labels discussed in this unit are mere nomenclatures to properly differentiate one type of research from the other. Otherwise, research itself has one ultimate goal: to find answers to a set of perplexing questions. These classifications therefore do not really make a difference. They are mere guides to any budding researcher to follow for particular reasons.

5.0 SUMMARY

In this unit, we identified five core ways researchers classify research in the social sciences and mass communication thus: classification by practice, classification by measurement, classification by data gathering method, classification by discipline and other types of classification not mentioned in the first four. It is my expectation that you have grabbed the core gist about these research types and could possibly explain their meaning, similarities and differences.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss three major differences between primary research and secondary research.
ii. Discuss any two differences between exploratory research and survey research.
iii. Research classification by the method of data gathering sounds duplicative. Please, discuss with valid argument.

7.0 REFERENCE/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd Ed.) Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 4 MEASUREMENT

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Communication Research and Measurement
    3.2 Levels of Measurement
    3.3 Scales
    3.4 Indexes
    3.5 Reliability and Validity
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

Measurement, according to Wimmer and Dominick (2011:62), "is the assignment of numerals to objects, events, or properties according to certain rules". The four levels of measurement are nominal, ordinal, interval, and ratio. This is the focus of this unit. The importance of this unit to the data analysis business cannot be overemphasised. Without the proper measurement in place, data analysis will not be smooth. Measurement is therefore a prerequisite for a valid data analysis.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- list the levels of measurement
- describe measurement scales
- explain reliability and validity.

3.0 MAIN CONTENT

3.1 Communication Research and Measurement

Measurement is the mathematical process of assigning numerals to objects, events, or properties following some set standard rules. According to Wilson, Esiri and Onwubere (2008), "Broadly speaking, measurement can be described as the process of determining dimensions, values or degrees. This includes a person's height, weight, temperature, academic performance, wealth, etc. This means that certain measuring devices are required to carry out measurement".

Wilson, Esiri and Onwubere (2008) also see measurement as a logical rule for assigning numbers to observations to represent the quality of a trait of characteristic possessed. Wimmer and Dominick (2011) identify three important terms to note in the various definitions of measurement thus: **numerals**, **assignment** and **rules**. According to them, a numeral has no implicit quantitative meaning. It only takes quantitative meaning when it becomes a number and used in mathematical calculations (Wimmer & Dominick, 2011). Assignment on the other hand is the designation of numerals or numbers to particular objects or events while rules specify the way such numerals or numbers should be assigned (Wimmer & Dominick, 2011:49).

With regards to the apparent relationship between measurement and communication research, Wilson, Esiri and Onwubere (2008) note thus: In social science research (communication research inclusive) scholars are very interested in identifying and defining **concepts**, **variables** and **constructs**. After this has been done, researchers have the responsibility of measuring these research elements in real life. For instance the concept, "Newspaper readership" could severally be measured in the following ways: Yes or No response (to ascertain whether respondents read newspapers). Number of times a week or month a respondent reads newspaper. The number of minutes or hours the respondents spends reading newspapers.

These are all different measures of a particular concept, "newspaper readership". A concept that is frequently measured in communication research is **attitude** which can be defined as "an enduring, learned predisposition to behave in a consistent way towards a person or given class of objects". Attitudes are usually measured in terms of **direction**, **intensity** and **stability**.

3.2 Levels of Measurement

According Wimmer and Dominick (2011), scientists have identified four different ways to measure things, which is also regarded as the four levels of measurement thus: Nominal, Ordinal, Interval and Ratio. Continuing, Wimmer and Dominick (2011:51-52) explain these different levels thus:

- The **nominal level** is the weakest form of measurement. In nominal measurement, numerals or other symbols are used to classify people, objects, or characteristics.
- Objects measured at the **ordinal level** are usually ranked along some dimension, such as from smallest to largest.

- When a scale has all the properties of an ordinal scale and the intervals between adjacent points on the scale are of equal value, the scale is **interval level**.
- Scales at the ratio level of measurement have all the properties of interval scales plus one more: the existence of a true zero point.

3.3 Scales

A scale, according to Wimmer and Dominick (2011), represents a composite measure of a variable. It is based on more than one item. They can also be described as composite, multiple measures about a particular aspect of a theoretical concept. Scales and Indexes are generally used with complex variables that do not easily lend themselves to single-item or single-indicator measurement. Scales in particular are used to measure theoretical concerns which are generally measures of independent variables. Scales typically have formulated rules for developing the multiple indicators and assembling them into one composite value (Wimmer & Dominick, 2011:53).

Some simple specialised rating scales include transforming scales, thurstone scales, guttman scaling, likert scales and semantic differential scales.

3.4 Indexes

The term index refers to an indirect measure of a particular variable under investigation. It is a composite measure of multiple dimensions of a particular theoretical concept (Wilson, Esiri & Onwubere, 2008). Indexes do not have detailed construction rules but are generally used to measure dependent variables. “Both scales and indexes are constructed to provide ordinal (ranking) measures of a given variable. Both are ordinal measures because they rank people or other units of analysis in terms of specific variables such as religiosity, prejudice, intelligence. In essence, a respondent’s score on a scale or index of religiosity gives an indication of his or her religiosity vis-à-vis other people” (Wilson, Esiri & Onwubere, 2008).

3.5 Reliability and Validity

Wimmer and Dominick (2011) note that using a newly developed scale without preliminary testing in the form of a pilot study is poor research. To be valid, therefore, a measurement must have reliability and validity. According to Wimmer and Dominick (2011:57), a measure is reliable if it consistently gives the same or similar results each time it is applied in similar contexts while validity implies that a measuring device actually measures what it purports or sets out to measure.

SELF- ASSESSMENT EXERCISE

Discuss some valid reasons why measurement is important in mass media research.

4.0 CONCLUSION

In this unit, we considered measurement from the perspective of Wimmer and Dominick (2011) as the assignment of numerals to objects, events, or properties according to certain rules. We also considered four levels of measurement to include nominal, ordinal, interval, and ratio and note that measuring devices will be useless if they are not reliable and valid.

5.0 SUMMARY

We focused on the place of measurement in media and communication research. This is vital considering the fact that understanding empirical research requires a basic knowledge of concepts, constructs, variables and measuring devices. This again is a prerequisite knowledge to data analysis.

6.0 TUTOR-MARKED ASSIGNMENT

Develop a measurement technique to test the newspaper reading habits of undergraduates of NOUN.

7.0 REFERENCE/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd Ed.) Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

MODULE 2 COMPUTERS AND DATA ANALYSIS

Unit 1 Data Processing and Management
Unit 2 The Use of Computer in Data Processing and Analysis
Unit 3 Statistical Package for Social Sciences (SPSS) Program
Unit 4 Data Analysis in Communication Research

UNIT 1 DATA PROCESSING AND MANAGEMENT

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Data Editing
        3.1.1 Objectives of Editing
        3.1.2 Levels of Editing
    3.2 Data Coding
        3.2.1 Basic Data Coding Operations
        3.2.2 Purpose of Coding
        3.2.3 Types of Coding
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

Data processing is a preliminary procedure that must take place for data analysis to make sense. If this does not happen and correctly too, data analysis will merely be “garbage in, garbage out” syndrome. Therefore, data processing and management is the all encompassing and the entirety of the processes by which raw data or information gathered from the field of investigation are refined, cleaned, organised and structured so as to facilitate or enhance data analysis and interpretation of the research findings.

Data processing is essentially a preliminary procedure to data analysis and interpretation of the results. It therefore involves a number of processes which includes primarily data editing and data coding.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- describe data processing and management
- identify and describe data editing processes
- identify and describe data coding processes.

3.0 MAIN CONTENT

3.1 Data Editing

Editing refers to the process of examining and thoroughly checking the raw data or information that had been collected for possible incorrect entries and other similar errors. The essence of data editing primarily is to detect any possible incorrect entries or errors and rectify them (Obikeze, 1990). Please note that editing may be carried out on any type of data irrespective of the instrument through which they were collected or the expertise of the researcher(s) who carried out the data gathering exercise.

3.1.1 Objectives of Editing

Generally, the primary objective of editing is to:

a. *Enhance the quality of field data* by making sure that all information are complete, unambiguous, legible, uniform, accurate and true.
   Any incomplete data could render analysis useless. At this stage, the appropriate research assistants or data editors are expected to ensure information supplied are detailed enough depending on what the researcher wants to achieve and that the information gathered is not ambiguous, illegible, inaccurate and untrue before it becomes difficult to reach the participants or respondents for further clarification.
   Data analysis is essentially analysis of what was brought in as information. Thus, any wrong information that creeps in at this stage unnoticed will surely make analysis also wrong. The implication again is that data interpretation or interpretation of the so-called results also becomes very deceptive because of the faulty premise it was based.

b. *Eliminate Coding Difficulties*
   Eliminating coding difficulties is another very important objective of data editing. At this stage, the investigators are really interested in making sure any potential difficulty that may creep

in at the stage of coding is eliminated now before that stage. The reason is simple. At the coding stage, any unresolved difficulty may engender analysis because it may be difficult to get back to the field to seek clarification. Coding, as we will later find out in this unit must take place for a meaningful data entry and subsequent analysis to be effectively carried out.

c. *Facilitate Data Analysis and Interpretation of Results*
Data editing is primarily meant to facilitate data analysis. The research process is so linked together that every of the procedures or processes are interrelated and aiming at a single primary purpose: find answers to the research questions. So, data editing helps the next procedural step by making sure its own loose ends are tight against possible errors that could be traced back to poor editing.

d. *Detect Errors/ Incorrect Entries and Correct Them*
Editing in normal or other situations is done to identify errors and to correct them. This is not different in data editing. Humans are naturally prone to errors. This should not be a problem at all if a proper editorial team is in place to proofread materials before they are moved to the next level. Therefore, detecting errors, incorrect entries, or incomplete entries with the aim of correcting them before any form of analysis is performed on the data should be a primary objective of data editing.

e. *Facilitate Interpretation of Results*
Interpretation of research results is a fundamental purpose of research which is meant to address specific needs. The outcome of researches cannot really make any sense if such outcomes are not properly put in context by way of interpretation. Interpretation will help the researcher to find out if the objectives have been met and to find out if the research questions have been answered. This makes interpretation of research findings very vital in the research business. Having established the importance of result interpretation, every effort should be made at the editing stage to facilitate this.

3.1.2 Levels of Editing

Generally, editing is carried out at three different levels by three different categories of research personnel. This is supported by the findings of Obikeze (1990) thus:

a. **Interviewer's Editing**
This is carried out in the field by interviewers or any other instrument administrator or supervisor. Here the personnel in question is expected to review and check each questionnaire and /or field notes or interview schedule to ensure that every item has

been properly completed, that the responses are clear, unambiguous and legible. At this level of editing, any inconsistencies or omissions discovered are promptly referred to the subject or respondent for clarification, verification and resolution (Obikeze, 1990).

Please, note that the temptation for the interviewer to supply the missing link by him or herself should be avoided. Where it is no longer feasible to get back to the subject, the best advice is not to change anything but to be more alert to the occasions or error in subsequent interviews (Obikeze, 1990).

b. Supervisor's Editing

This is also carried out in the field by the project supervisors. This level of editing is one of the primary assignments of supervisors and it is carried out while the interviewer is still around so that he/she could be reached to provide further clarification, missing details and explanation. The purpose of a supervisor's editing is to ensure legibility and to detect and correct apparent errors while in the field (Obikeze, 1990).

c. Research Assistant's Editing

This is carried out at the research office by research assistants. Data editing at this level consists of two closely related operations, namely: final-check editing and judgmental editing. In the first operation, the office staff undertakes a quick but systematic check of the questionnaire or interview schedules. In this check, each instrument must be checked for the following issues:

- **Completeness** – to ensure that every item in the questionnaire has been filled.
- **Inconsistencies and Inaccuracies** so as to correct them before instrument administration.
- **Uniformity** to make sure the response pattern is similar in format from questions of similar expectation.

Judgmental editing is the second operation which requires greater skill and experience than a research assistant would normally possess. It is therefore generally carried out by a few specialist members of the research team. Judgmental editing, according to Obikeze (1990) involves making expert judgment and technical decisions on a number of intricate or complicated questionnaire items so as to facilitate coding and analysis. Judgmental editing, according to Obikeze (1990), concentrates on the following:

- Making decisions regarding the appropriateness or otherwise of certain recorded responses. e.g. an unemployed person responding to question on job satisfaction.

- Effecting necessary modifications where information is given in units other than the one specified. e.g. salary given in monthly rather than annual basis.
- Developing appropriate categories and producing complex variables such as occupation.
- Preparing summaries and computing indexes for intricate issues like income and quality of life.
- Note that judgmental editing is carried out to remove all obstacles to successful coding.

SELF-ASSESSMENT EXERCISE

i. Discuss the major difference(s) between final-check editing and judgmental editing.
ii. Discuss two major objectives associated with editing and how this is achieved.

3.2 Data Coding

Coding is a process by which data gathered from the field are put in categories and given numbers or letters of the alphabet to enhance storage, retrieval, tabulation and data analysis (Obikeze, 1990).

3.2.1 Basic Data Coding Operations

A Establishment of Categories or Classes

One basic coding instruction is to establish a proper and near exhaustive category system or classes to reflect expected responses. The classes should be exhaustive as earlier indicated and should not also overlap. It is expected that a well defined and delineated class can only take in similar type of content. In essence, if a response could fit into two different classes per time, then something is wrong with such category system. Therefore, it is expected that a good category system for research must be mutually exclusive, exhaustive and one-dimensional in type and sequence.

For example, a question asked to determine the highest educational qualification of a research participant could have the following classes:

- FSLC
- JSSC
- SSCE/WASSCE O'LEVEL
- OND/DIPLOMA

- HND/BSC
- Master's degree
- Doctoral degree
- Postdoctoral qualification
- No formal education
- Others, please specify

B Assigning of Numbers or Labels to these Categories/Classes

The next step you must take in a typical basic coding operation is to assign certain or specific numbers to the categories you have developed in 3.2.1.(A) above. For example:

FSLC……...1
JSSC……....2
SSCE/WASSCE O'LEVEL……3
OND/DIPLOMA...……..4
HND/BSC……....5
Master's degree……...6
Doctoral degree……...7
Postdoctoral qualification…….....8
No formal education…….…….….9
Others, please specify…….…….….10

C Sorting the Raw Data into these Classes

After you have established the category system and finished assigning numbers to the classes, you can then commence the process of sorting your raw data into the classes. Again, I need to reiterate here that the raw data should only fit into a single class per time in a well defined and exhaustive classification arrangement or system.

3.2.2 Purpose of Coding

The objectives of coding differ according to the type of study, the nature of data gathered and aims of carrying out such a study. For qualitative based studies, coding is carried out to ensure orderliness and structure in data processing so as to enhance analysis and interpretation. In quantitative based studies like surveys, coding is carried out to properly classify all possible range of responses into limited categories to enhance analysis and interpretation (Obikeze, 1990).

Coding is one of the vital ways to organise a mass of data into manageable components so that the arduous task of analysing such mass of data does not scare the researcher. In essence, the ultimate purpose of coding is to ensure orderliness and proper organisation while the ultimate objective is to enhance data analysis and result interpretation.

3.2.3 Types of Coding

- Qualitative Coding
- Quantitative Coding

1. **Qualitative Coding**

Coding in qualitative studies is a process of actual data analysis and interpretation rather than the assumed data processing or preparation for processing exercise. What this means is that an attempt to code qualitative data leads you to almost analyse and probably interpret such data. Goode and Hatt (1952) cited in Obikeze (1990) summarised the procedure for qualitative coding thus:

- have the objectives of a given research as well as the important research questions to be answered and issues to be resolved clearly spelt out and clarified
- decide the way of organising, approaching or classifying the subject matter that best suits the research objectives
- set up or develop appropriate classes, categories or units for classifying or grouping the contents of the materials in accordance with the chosen approach
- go through the material to identify other alternative words and phrases that serve as indicators, pointers or cues to the various established categories or classes
- work through the materials carefully and allocate the contents to the categories.

At the end of this process, it is then possible to establish the frequency of occurrence of each class of phenomenon or event in the study material (Obikeze, 1990).

Uses of Qualitative Coding

Qualitative coding has many uses. Some of these uses are summarised by Obikeze (1990) thus:

- helps the researcher to clarify, sharpen and specify the key concepts, issues and overall research objectives
- helps the researcher to detect errors, omissions, inconsistencies etc in the construction and administration of research instruments
- helps the researcher to replicate data analysis procedures and verifications of research findings and interpretations
- helps the researcher to apply statistics in the analysis of qualitative data
- helps in the development of new and specialised social research techniques and methodologies.

2. Quantitative Coding

This type of coding involves survey based studies and generally follows the same three basic code operations identified earlier. In essence, quantitative coding is assignment of raw data gotten from quantitative based studies into an organised coding system and operations.

A Establishment of Code Categories or Classes

This basic quantitative code operation is governed by three basic principles as identified by Obikeze (1990) thus:

- **The categories or classes must be completely exhaustive:** this means that they must cover all possible range of responses for a given question or issue.
- **The categories must be mutually exclusive:** this means that the classes must not overlap. A given category must fit into one and only one category.
- **The categories must be one-dimensional:** this means that the categories must be derived from a single variable or aspect of one variable that is in line with what the researcher has in mind. For instance "**Where were you raised up?"** Could refer to:
  - Place of residence
  - Type of institution.

Obikeze (1990) therefore notes that the objectives of the study will determine the appropriate variable to use to avoid any potential confusion from creeping up at the sorting stage.

B Assigning of Numbers to Category

The assignment of codes (numerical values) is usually done arbitrarily, but a school of thought in research, especially in the social sciences suggests that certain standard coding conventions should be followed when doing this sorting. The primary objective of doing this is to maintain uniformity in standards across disciplines and world divides. These suggested standards include:

- **In Yes or No responses, assign fixed codes to the two classes throughout the study.** If Yes=1 and No=2, let it be so throughout the study.
- **Also, assign fixed codes to responses like "Don't know", "Can't say", "No response/answer" etc categories.** Usually, a **No response** could be **0** or the last number after every class has taken codes

C Sorting Responses into the Categories

This could be carried out by the researcher himself if the questionnaire is well structured and pre-coded, otherwise some

trained coders are responsible for this particular operation. What is expected at this level is for whoever that is responsible for the sorting out process to take the coding instructions/guide, check out the developed category system and the assigned numbers and commence the process of getting the raw data into an organised structure ready for data analysis.

Note that from the manual operation of this, the sorting could also take place directly on a computer system that has any of the software application for statistical delivery we will be discussing at a later stage in this module.

4.0 CONCLUSION

What we do with our data before they are analysed is as important as analysing the data to answer specific research questions. This is why data processing and management procedures are very vital in the data analysis. The data analysis procedure is in such a way that if you feed in the wrong type of data into any analysis software or platform, it brings you a wrong result. Efforts should therefore be taken to process and properly manage the raw data gathered from the field before any kind of meaningful analysis is carried out on such data.

5.0 SUMMARY

This unit focused on data processing and management with particular attention to data editing and coding. Under these two sub-headings, we discussed the objectives and purposes of both data editing and coding; various types of editing and who is responsible for each type; how to develop coding guides and instructions; the different types of coding and how each is done; and finally noted how to assign numerals to developed categories and sort the raw data preparatory to data analysis.

6.0 TUTOR-MARKED ASSIGNMENT

i. Distinguish between editing and coding.
ii. How would a Likert type of measurement scale be coded?
iii. Use a self developed coding guide to explain what we mean by a category should be mutually exclusive, one-dimensional and exhaustive.

7.0 REFERENCES/FURTHER READING

Bruning, J.L., & Kintz, B.L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Champion, D.J. (1981). *Basic Statistics for Social Research*. NY: Macmillan.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Nwabuokei, P.O. (1999). *Fundamentals of Statistics*. Enugu: Koruna Books.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Okeke, M.C. (1997). *Statistics: A Business Approach*. Onitsha: Onwubiko Prints Ltd.

Otokiti, S. *et al*. (2007). *Contemporary Statistical Methods*. Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd Ed.) Lagos: Idosa Konsult.

Spiegel, M.R. (S.I. Edition). *Theory and Problems of Statistics*. New York: McGraw-Hill.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 2 THE USE OF COMPUTER IN DATA PROCESSING AND ANALYSIS

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
  3.1 The Use of Computer in Data Processing and Analysis
  3.2 Task Areas in the Use of Computer for Research
  3.3 Abuse of the Computer in Communication/Social Science Research
  3.4 Advantages of the Use of Computer in Communication Research
  3.5 Disadvantages of the Use of Computer in Communication Research
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

The computer technology is one that revolutionalised the research industry, especially in the 21st century. Today, one can start a research project on the computer and complete the entire process on it without stepping out. The software packages that support data analysis are all computer based. Beyond data analysis, the computer technology has also revolutionalised data storage and retrieval making data processing and management fascinating. We do not have to carry mass of papers and files all around us anymore; that era is gone for good. We now have many friendly and portable computer based technologies to help us manage our mass of data.

There are many other areas of research that the emergence of computer technology has affected both positively and negatively. However, our major focus in this unit is on research in general with specific interest on data analysis.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- explain the use of computer in data processing and analysis
- identify the task areas in the use of computer for research

- describe the advantages of the use of computer in communication research
- describe the disadvantages of the use of computer in communication research
- list the different ways the computer is abused in communication/social science research.

3.0 MAIN CONTENT

3.1 The Use of Computer in Data Processing and Analysis

The 21st century is aptly described as the computer age which means citizens of this century must be familiar with computer applications or sink into irrelevances in the scheme of things. In developed societies like the Western capitalist nations, computer literacy has become an aspect of primary education and computer application has permeated virtually every aspect of the nation's life and culture. Even in developing societies, computer literacy is becoming fashionable (Obikeze, 1990).

The use of the computer in social science research is revolutionary. According to Obikeze (1990), the development of such computer programmes as the general inquirer (a set of programmes designed specifically for content analysis of verbal materials and also flexible enough to be applied to a variety of qualitative information) has greatly simplified and revolutionised the processing and analysis of qualitative material or data.

The revolutionary nature of the computer is felt or seen more in the analysis of quantitative data. On account of its speed, efficiency and accuracy, the computer is able in a matter of seconds to carry out complex mathematical computations and manipulations on large volumes of mass data which hitherto would have taken months to accomplish using worksheets and desk calculators (Obikeze, 1990). The use of the computer has also made the application of a variety of complex statistical techniques like regression analysis in social science research possible. Third, the use of the computer in social science research has facilitated the process of inductive social research by enhancing the interactive process of moving back and forth between the researcher's ideas and the findings from the data (Obikeze, 1990). In essence, the findings of the study suggest new ideas which in turn suggest new analysis.

3.2 Task Areas in the Use of Computer for Research

Generally, computer technology could be applied in media or social sciences researches to perform five major. These tasks, as identified by Obikeze (1990), make the computer a wonder working machine:

a. Data Storage and Retrieval

The computer has the capacity to store any kind of data and produce such information on demand. The stored information/data are also safe wherever it is stored until when needed, unless of course attacked by viruses or system breakdown or collapse, which could be handled by adequate protection of the computer system.

Another issue here is the ability of the computer to store any amount of data, which of course is dependent upon the memory and storage capacity of the particular system.

b. Data Processing

The computer also has the ability to process any kind of data if properly entered. It could verify information, edit data, sort, replicate, and tabulate information in whatever form as demanded by the researcher.

c. Mathematical / Statistical Calculations

The computer is capable of manipulating any kind of statistical or mathematical problem or calculations. Chi-square, co-efficient correlation, t-test, ANOVA, etc are easily performed using some computer software or programs. The amount of time required to manipulate these sophisticated test statistics is also reduced using the computer technology.

d. Simulations

The computer also has the capacity to generate data which is supposed to represent the actual behaviour of an object or a temporal process.

e. Data Analysis and Interpretation / Presentation of Results

Apart from processing data or research information, the computer has the capacity to analyse the results of the research as well as interprets the outcome for easy comprehension of the user. The researcher only has to present the analysed data in the form his readers could understand. Do not forget that the computer could also help in the presentation through appropriate tables, charts, graphs etc.

3.3 Abuse of the Computer in Communication/Social Science Research

Generally, the computer has developed enough statistic based software applications which enable it perform any kind of statistical calculations or manipulations. This is often taken advantage of by abusive users. The ability of the computer to multitask also opens it up to abuse by neophyte researchers. Some of these abuses, according to Obikeze (1990) confuse rather than clarify. They include:

a. Wrong Application of Statistical Techniques for Data Analysis

Evidence abounds to support the argument that available statistical software or packages are being wrongly applied by both students and researchers who understand little or nothing about the assumptions or implications of such statistical software or packages. This is often leads to the problem of "garbage in, garbage out" syndrome in social science research (Obikeze, 1990).

To avoid this particular abuse, Nie, et al (1975) warns that "a user should never attempt to use a statistical procedure unless he understands both the appropriate procedure for the type of data and the meaning of the statistics produced."

b. The Tendency to Overproduce Statistical Tables

The second type of abuse is the tendency for the inexperienced researcher to overproduce tables, graphs, charts, etc without consideration to main issues or variables under study. This is seen more in the overproduction of cross tabulation tables and correlation matrixes. The result of this type of abuse is a truckload of computer print-outs which will take a decade to go through (Obikeze, 1990).

This particular abuse also impedes understanding of research efforts rather than promote if the analysis of the statistical calculations is part of the promotion. The readers are confused, the policy makers who are supposed to make use of the results are lost within the results and the supervisor for the students' project just cannot read through within the stipulated time so he lets it go. Meaningful application of research results for developmental or growth purposes suffer as a result.

Furthermore, the researcher, by engaging in this particular abuse, alienates himself from his products or findings. He has no defense for his reported findings either theoretically or practically because he is lost in a midst of enormous statistical or mathematical manipulations he least understands (Obikeze, 1990). Another danger related to this is that the

researcher may even play around these figures just to protect his ego and hide the fact that he does not really understand the issues at stake.

One practical way to avoid this form of abuse is to ensure that only variables which have theoretical and/or practical relevance to the research objectives and hypotheses are included in computer packages for analysis (Obikeze, 1990).

SELF-ASSESSMENT EXERCISE

Discuss the practical ways the computer is abused by the inexperienced user.

3.4 Advantages of the Use of Computer in Communication Research

a. It Makes Data Storage and Retrieval Easy
One critical advantage of the computer is its capacity to store large amount of data and produce same on request many days or years after storing such data.

b. It Makes Data Processing Easy to Perform
Data processing and management are easily performed using the computer technology unlike when humanity used manual efforts to attempt same. Then, the job will take days with so much errors and omissions. Today, the computer has speed, accuracy and efficiency in processing and managing large amount of data.

c. It Facilitates Simulation
The computer could generate data which could take the place of actual behaviour of an object or a temporal process especially for research procedures.

d. It Increases the Accuracy of Statistical/Mathematical Calculations as Well as Maintains Speed in Doing Such Calculations
The computer has the advantage of speed, accuracy and efficiency in its application. Unless there is a mechanical fault which is not usually the regular case, a computer computation cannot be wrong. This is without prejudice to the garbage in, garbage out syndrome the technology is associated with.

e. It Enhances Research Results and Subsequent Interpretation of such Results
A good data analysis will definitely lead to a good interpretation of the result. The computer obviously has the capacity to generate valid results which will help the researcher or user make valid inferences and interpretation from the results.

f. **It Facilitates Data Presentation through the Application of Appropriate Techniques**

The computer helps the researcher or presenter with diverse presentation applications that enhance delivery and skill development. The computer also has other presentation techniques and modules that make data presentation appear in a much more appealing and comprehendible formats.

3.5 Disadvantages of the Use of Computer in Communication Research

a. It is Subject to Abuse

The computer is often abused by inexperienced users who just get excited about the many things they could do with the computer. I have seen a situation where an undergraduate simply instructed the computer to print out simple frequency tables, bar charts and pie-charts of a 30-item questionnaire constructed to address five research questions only. So, she had 30 FDTs, 30 similar bar charts and 30 pie-charts in the same chapter four of her work. This is a typical abuse found with undergraduate using computer to do their projects.

b. Virus Attack can Wipe off your Entire Data

Virus and other types of attack could completely wipe out stored data. What is the need of saving data that will never be used again because it has been deleted by a virus that attacked the system? This is a core disadvantage of the computer technology. The earlier researchers begin addressing this issue, the better for all of us because once people get the conviction that their private, confidential and sensitive information are no longer safe in the computer they will begin reinventing the wheel.

c. System Crash could also Cause the Loss of Data

Apart from virus attack which wipes out data; a worse case scenario exists: system crash! A system could crash as a result of various forms of virus attack. This will result in the loss of data.

d. Data could also Easily fall into the Hands of Wrong Persons who have Access to the Computer System in which such Data are Stored

Such wrong persons could take advantage to go through your confidential and personal data, or even steal your personal data for personal use.

4.0 CONCLUSION

The development of computer technology is both phenomenal and revolutionary. It changed the course of many disciplines. For this particular course on data analysis, the emergence of the computer has made it easier, faster, more convenient, accurate and efficient to do any type of statistical analysis on any type of data.

5.0 SUMMARY

This unit focused on the use of computer in data analysis and management, with particular attention to task areas in the use of computer for research; abuse of the computer in communication/social science research; advantages of the use of computer in communication research and disadvantages of the use of computer in communication research.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss a typical way could simulate behaviour using the computer.
ii. Could you justify the deployment of computer systems in major research institutes all over the world?
iii. Discuss six critical ways the computer technology has helped the growth of mass media research.

7.0 REFERENCES/FURTHER READING

Bruning, J.L. & Kintz, B.L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Champion, D.J. (1981). *Basic Statistics for Social Research*. NY: Macmillan.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Nwabuokei, P.O. (1999). *Fundamentals of Statistics*. Enugu: Koruna Books.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Okeke, M.C. (1997). *Statistics: A Business Approach*. Onitsha: Onwubiko Prints Ltd.

Otokiti, S. *et al*. (2007). *Contemporary Statistical Methods*. Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd Ed.) Lagos: Idosa Konsult.

Spiegel, M.R. (S.I. Edition). *Theory and Problems of Statistics*. New York: McGraw-Hill.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note Developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 3 STATISTICAL PACKAGE FOR SOCIAL SCIENCES (SPSS) PROGRAM

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Statistical Package for Social Sciences (SPSS) Program
    3.2 Steps in Using SPSS to Test Hypothesis
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

The Statistical Package for Social Sciences (SPSS) program software is primarily designed to facilitate data analysis in the social and behavioural sciences. There are other statistical packages but the SPSS is mostly recommended and used in the social and behavioural sciences. The SPSS is a user friendly package that you can learn while using it. The tutorials are also learner friendly and the output is always accurate to the extent the researcher has keyed-in the accurate data.

Data analysis in communication research using SPSS is a made-easy process even for the inexperienced researcher and user. This unit will therefore merely highlight some of the basic information about the SPSS program package. To properly understand the software, one must attempt to install it in a computer system and actually begin to use it. So, this lecture material cannot replace the actual practical exercise you must carry out. It will merely complement it. So, you are advised to attempt practicing using the program to analyse some miniature studies you may be involved in.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- explain the meaning of SPSS and why it is a valuable resource for data analysis in communication or social science research
- describe the processes involved in the application of SPSS for data analysis
- state some advantages inherent in the use of the SPSS program package.

3.0 MAIN CONTENT

3.1 Statistical Package for Social Sciences (SPSS) Program

The Statistical Package for Social Sciences (SPSS) program was developed specifically to meet the needs of researchers in the social and behavioural sciences which mass communication belongs to. The SPSS is an integrated system of programmes designed to enable the social or communication scientist carry out many different forms of data processing, management and analysis in a most simple and convenient manner. It does not require any sophistication in statistics or prior experience in computer programming to be able to use the SPSS system software (Obikeze, 1990).

SPSS came to address the inconveniences which hitherto existed when working on different computer packages. For instance, when working in MS Word, one may need to go to EXCEL to do some mathematical manipulations which are not easily accessible on MS WORD. SPSS is therefore a single package that has come to address this. Its popularity amongst students and researchers, according to Obikeze (1990), is based on the followings:

- the SPSS control-statements or instructions are written in a quasi-natural language thus making it possible for the social science researcher to carry out his data analysis in his natural language
- the SPSS package does not require sophistication in statistics or sophistication in computer programming before application and detailed use in any of the social sciences
- the SPSS is able to handle data files containing large or very large amount of data and variables thus making it possible for the researcher to test complex social theories which hitherto could have been impossible
- the SPSS package has a detailed set of procedures for data storage, transformation and manipulation. These procedures are also easy to understand and follow especially for new users.

Please, note that these four factors discussed above also constitute the advantages inherent in the use of this package. The disadvantages in the use of computer for data analysis are also applicable here. In essence, the four factors listed above are also the advantages associated with the use of the program. For the disadvantages, the disadvantages of the use of computer in communication research as highlighted below:

- it is subject to abuse
- virus attack can wipe off entire data

- system crash could also cause the loss of valuable data and
- data could easily fall into the hands of wrong persons who have access to the computer system in which the data are stored; are also disadvantages in the use of SPSS for communication research purposes.

How to Use SPSS System or Program Package
Note: This aspect of the lecture will be carried out practically in the computer room. Areas of practical exposure will include:

- **Descriptive Statistics**: frequency distributions, cross tabulations, computation of means, standard deviations etc.
- **Chi-Square computations**, correlation measures, analysis of variance tests (ANOVA), regression analysis etc.
- **Tables, graphs** and charts construction.

NB: All of the above could be computed using the SPSS program package. It is therefore highly recommended that you make every effort to avail yourself of any opportunity to have a hands-on experience trying to compute any of the test statistics listed above.

3.2 Steps in Using SPSS to Test Hypothesis

The SPSS program software package is a standard test kit that follows a set of rules to achieve desired output for any level of researcher. It is a popular statistical package for most researchers in the social and behavioural sciences. For a typical study in communication or media research, a researcher could compute the significance of a hypothesis by following the underlisted procedures (Note that this is for SPSS 13.0/14.0, newer versions have updated procedures to achieve same result):

- Open the **SPSS** package on your computer system. This is the first step assuming you have already installed the software on your computer system
- When the program comes up, it will request you to **OPEN AN EXISTING WORK (FILE) OR START A NEW ANALYSIS**. Of course, if you are continuing an existing work, you scroll to the exiting file, otherwise, instruct it to commence something new if you are just opening to begin a new SPSS data analysis procedure
- Go to **VARIABLE VIEW** to construct your **CODING CATEGORIES** or guide. The variable view is like your coding guide in a manual set up
- Go to **DATA VIEW** to enter the data generated from the survey. You are to manually enter your raw data into this data view after

you must have given specific instructions in the variable view. The data view is like the coding sheet in a manual arrangement

- Go to the **TOOLS BAR** and wait for the icons there to appear. Click on **ANALYSE** on the list
- On **ANALYSE**, click **DESCRIPTIVE STATISTICS**, and then wait for the icons there to show up; click on **CROSS TABULATION** from the list of icons.
- Pick a **VARIABLE** for the **ROW** entry. Pick another **VARIABLE** for the **COLUMN** entry. Press **ENTER** after each pick. The variables you are picking are not randomly done. They must be the Dependent and Independent Variables implied in the particular Hypothesis been tested or analysed
- Click on **STATISTICS** on the **CROSSTABULATION** entry display. Select **CHI-SQUARE** (if this is the result you are interested in) and click continue
- Go back to **CROSSTABULATION** entry display click **CELLS** and then select items needed and then click continue. The items selected here will be dependent on what the researcher wants to achieve plus the research objectives he/she has set at the beginning of the study.
- **CROSSTAB** display appears again, click on **FORMAT** and select **ROW ORDER** and click continue. On the formatting platform, you could give any type of instructions in terms of how you want the output to come out.
- **CROSSTAB** display appears again, click **OK** and then go to the **OUTPUT VIEWER** for the **RESULTS.**

SELF-ASSESSMENT EXERCISE

Discuss four reasons why you think SPSS comes highly recommended.

4.0 CONCLUSION

The 21st century is the computer age; the age of information society driven by increased sophistication in information and communication technology. Data analysis in communication or media related researches can therefore afford to lag behind in updates in the social science research industry. The computer technology is one critical innovation of the information society age and the computer technology cannot run without specific software packages. This is where SPSS program software fits in. Students of data analysis in communication must therefore train to use the statistical application which regularly receives updates from feedbacks gotten from users.

5.0 SUMMARY

This unit discussed the Statistical Package for the Social Sciences (SPSS) as a highly recommended statistical software for media/communication researchers who are desirous of achieving accurate and speedy analysis of their research findings. It is our understanding that you now know the meaning of SPSS, how to use it and the situations you can use it. Above all, we believe we have challenged you to take another critical step: purchasing the SPSS program software, install it on your system and begin regular practical and guided practice so that you could master the art of this scientific process.

6.0 TUTOR-MARKED ASSIGNMENT

i. Explain in details how you could test a given hypothesis at the 5% alpha level of significance using the SPSS.
ii. Discuss five core advantages in the use of SPSS for the mass communication researcher.

7.0 REFERENCE/FURTHER READING

Bruning, J.L., & Kintz, B.L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Champion, D.J. (1981). *Basic Statistics for Social Research*. NY: Macmillan.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Nwabuokei, P.O. (1999). *Fundamentals of Statistics*. Enugu: Koruna Books.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Okeke, M.C. (1997). *Statistics: A Business Approach*. Onitsha: Onwubiko Prints Ltd.

Otokiti, S. *et al.* (2007). *Contemporary Statistical Methods*. Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd Ed.) Lagos: Idosa Konsult.

Spiegel, M.R. (S.I. Edition). *Theory and Problems of Statistics*. New York: McGraw-Hill.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 4 DATA ANALYSIS IN COMMUNICATION RESEARCH

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
  3.1 Meaning of Data
  3.2 Types of Data
  3.3 Sources of Data
  3.4 Measurement of Data
  3.5 Meaning of Data Analysis
  3.6 Types of Data Analysis
  3.7 Major Descriptive Statistical Tools and how to Apply Them
  3.8 Inferential Statistical Tools and How to Apply Them
  3.9 Post Analysis Evaluation
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

This unit introduces you to the remaining critical concepts you must grapple with before a meaningful data analysis in communication research could take place successfully. At this stage, you may have acquired some basic information in this area but we have designed this unit to formally introduce you to some of the basic tools you may need for data analysis.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define the meaning of data
- explain the nature of statistical data
- list types of data
- mention sources of data
- describe measurement of data
- state meaning of data analysis
- mention types of data analysis
- itemise major descriptive statistical tools and how to apply them

- explain inferential statistical tools and how to apply them
- do a summary of statistical tools and conditions for using them
- describe post analysis evaluation of any specific research project you undertake.

3.0 MAIN CONTENT

3.1 Meaning of Data

The Chamber Dictionary of English Language defines data as "available facts from which conclusions may be drawn". McGraw-Hill Dictionary of Scientific and Technical Terms (2nd ed.) also defines data as the:

- General term for numbers, letters, symbols and analogue quantities that serve as input for computer processing. Any representation of characters or analogue quantities to which meaning, if not information may be assigned: Numerical or qualitative values derived from scientific experiments.

Gupta (1973) also defines data as:

- Aggregate of facts, affected to a certain extent by multiplicity of factors, numerically expressed, enumerated or estimated according to a reasonable standard of accuracy, collected in a systematic manner for a predetermined purpose and placed in relation to each other.

Wilson, Esiri and Onwubere (2008), however, see data in the following explanations:

- Data are unprocessed raw information for studying a phenomenon. This means that figures, words (e.g. news stories, photographs, colours, shapes etc.), all constitute what is referred to as data.
- Data are collection of facts or figures from which conclusions may be drawn. Data provide information which eventually leads to knowledge about the phenomenon upon which the data was gathered.

3.2 Types of Data

Statistical investigations come in different types. They include economic and social inquiries such as human population census, family budget inquiries, housing survey, public opinion polls, rural economic surveys etc. The conduct of these inquiries demands the collection of relevant

data (Nwabuokei, 1999). The data we collect for any statistical study may be qualitative or quantitative.

Qualitative Data
Qualitative data are those used for describing characteristics, which cannot be defined in numerical terms e.g. color of hair, color of eyes, performances grades as excellent, good, poor, average etc. These characteristics are referred to as ATTRIBUTES (Nwabuokei, 1999). Qualitative data can also be described as categorical data.

Quantitative Data
Quantitative data, on the other hand, are data that are capable of numerical description. E.g. data on the wages of workers in Naira, weights of students in kilograms, heights of students in meters, scores of students in a course exams etc. Quantitative data can also be called measurement data.

3.3 Sources of Data

The sources of data also come in as the types. In order words, the various sources of statistical data also function as type. So, we derive the type of data from the sources from which such data are gathered. In this line, we shall now discuss the two sources of data from where the two types of data emanate from thus:

a. **Primary Data**
Primary data are the type of data gathered directly from the field by the researcher. In order words, the researcher himself collects primary data. He collects them through interviews, experiments, direct observations, surveys, experiments etc. Primary data are collected for a specific purpose. After this need has been met, the data may become useless for any other purpose(s). In conclusion, one can comfortably say that primary data come in raw for processing and some even end up as secondary data in certain situations or circumstances.

b. Secondary Data
Secondary data already exist in a database or any other storage type but put there for a different purpose other than for research purposes. They are called secondary because the researcher himself does not directly obtain them. The researcher merely secures permission of the custodian of such data to use them for a different purpose for which it was originally obtained and stored.

Secondary data, however, come in first time as primary data except that they are data collected for use, for a purpose different from that for which they were originally collected. They are usually obtained from

existing records. For instance, medical records of patients in a particular area could now be used to determine the cause of maternal mortality of that particular area in a related study. Those medical records were merely kept there as a hospital policy but now a medical researcher who is interested in maternal mortality could make use of them with appropriate approvals.

SELF-ASSESSMENT EXERCISE

i. Discuss three major differences between primary data and secondary data.
ii. Discuss any two reasons why quantitative data may be referred to as measurement data.

3.4 Measurement of Data

As explained in unit 4 of module 1, there are four levels of measurement in social and behavioural sciences research. These are nominal, ordinal, interval and ratio measurement scales. Wimmer and Dominick (2011:51-52) explain these different levels thus:

- The **Nominal Level** is the weakest form of measurement. In nominal measurement, numerals or other symbols are used to classify people, objects, or characteristics.
- Objects measured at the **Ordinal Level** are usually ranked along some dimension, such as from smallest to largest.
- When a scale has all the properties of an ordinal scale and the intervals between adjacent points on the scale are of equal value, the scale is **Interval Level**.
- Scales at the **Ratio Level** of measurement have all the properties of interval scales plus one more: the existence of a true zero point.

In addition, Wilson, Esiri and Onwubere (2008) noted that the elements that are measured in communication research are usually variables, concepts and occasionally constructs.

3.5 Meaning of Data Analysis

Wilson, Esiri and Onwubere (2008) citing Williams (1968) explain that data analysis "involves the process of treating data with statistical tools so that a mass of data can be summarised, simplified and interpretable". This is based on the fact that data do not make any meaning without a form of analysis. They have to be scientifically processed for such data to make sense. The process involved in analysing data scientifically in order for such data to make sense is called data analysis.

Data analysis, according to Wilson, Esiri and Onwubere (2008), begins with the coding of the data preparatory to data entry using data entry software. The coding of the data is to make them amenable for analysis. After the coding, the data is then entered into the computer for analysis. This preliminary process involved is called data processing and management. This has been duly treated in unit 1 of this course material.

3.6 Types of Data Analysis

Two broad types of data analyses, derivable from the two types of statistics, are possible in the social and behavioural sciences. These are:

a. Descriptive data analysis
b. Inferential data analysis

a. Descriptive Data Analysis

Descriptive data analysis is a type of data analysis that occurs when data are analysed in such a way as to describe and summarise the content therein. The tools which are used in descriptive data analysis are descriptive statistics. Descriptive statistics is an aspect of statistics that studies a body of statistical data and no generalisations are made from the results obtained. Descriptive statistics only seeks to describe and analyse a given set of data without drawing any conclusions or inferences about the population. Population in statistics refers to any finite or infinite collection of objects under study (Nwabuokei, 1990).

In other words, everything dealing with collection, processing, analysis, presentation and interpretation of numerical data belongs to this aspect of statistics. Descriptive statistics include tabulation, graphical representation of data (e.g. bar chart, histogram, pie chart, etc) and measures of central tendency (Wilson, Esiri & Onwubere, 2008).

b. Inferential Data Analysis

Inferential data analysis happens when the researcher is interested in doing more than just description of the data. At this stage, parametric or inferential statistics is used. Inferential statistics is a branch of statistics that studies a group of data in order to use the results obtained in making generalisation on a larger group of data. In other words, statistical inference is the use of sample results to reach conclusions about populations from which the samples have been drawn (Nwabuokei, 1990).

Inferential statistics, according to Wilson, Esiri and Onwubere (2008), involves making generalisations about the whole population based on information or data obtained from a sample. Inferential statistics include estimation theory, hypothesis testing, parametric tests etc. Citing Olaitan, et al, (2000), Wilson, Esiri and Onwubere (2008), also note that "analyses done on this basis are used for testing hypotheses and making inferential decisions, based on some sample data." Thus, on the basis of analysed sample data, generalisation can be made about the overall population from which the sample was originally drawn.

3.7 Major Descriptive Statistical Tools and How to Apply Them

In a regular research project, the raw data usually come in a large set. To do proper analysis of such raw data, the data have to be organised in some sort of way so that the analyst will find his or her way around the mass of data easily. One way to do this is through tables, charts etc. However, to actually analyse data various types of frequency distribution tables are used. Some of these are presented below:

1. Frequency Distribution Tables (FDT)

There are basically three FDT namely:

- Univariate FDT
- Bivariate FDT
- Multivariate FDT

UNIVARIATE FDT means that only one variable or questionnaire item is being considered. See Table 4.1 below:

Table 4.1: Gender Distribution of Mass Communication Students in Covenant University

Gender	Frequency
Male	150
Female	180
Total	330

BIVARIATED FDT means that only two variables are being considered together. More specifically, it is the cross-tabulation of responses to two questionnaire items or two variables simultaneously. See Table 4.2 below:

Table 4.2: Distribution of Media Professionals According to Level of Education and Religious Beliefs

	Level Of Education					
Religious Affiliation	**University**	**College**	**Secondary**	**Primary**	**No School**	**Total**
Christian	7	10	13	2	1	33
Moslem	1	2	4	3	2	12
Others	3	2	3	5	2	15
Total	11	14	20	10	5	60

Multivariate FDT helps to describe, explain or understand relationships among three or more variables (containing one dependent and two or more independent and intervening variables). See Table 4.3 below:

Table 4.3: Distribution of Media Professionals According to Level of Education, Sex and Religious Beliefs

Level of Education and Sex								
	HIGH		MEDIUM		LOW		TOTAL	
Religion	M	F	M	F	M	F	M	F
Christian	10	7	5	8	1	2	16	17
Moslem	3	0	3	1	2	3	8	4
Others	3	2	2	1	2	5	7	8
TOTAL	**16**	**19**	**10**	**10**	**5**	**10**	**31**	**29**

2. Construction of Statistical Tables

The construction of a statistical table follows well defined procedures as highlighted below:

- a table must be numbered; this number is used to make reference to the table
- a table must also have a title which is brief, but self explanatory.
- a table should also have a row(s) with a row heading called a STUB and column heading called the CAPTION

- a table should also show the units and the signs of recognising the items in the table e.g. the income level in dollars or Naira
- a table should also have a footnote when some items on the table require further explanation
- a table should also have a source from which the data displayed are reported from. This is always placed below the table as the last footnote.

3. Description of Statistical Tables

Tables ordinarily do not speak for themselves. They need to be described to bring out the features of the data presented. Tables are best described using descriptive statistics and graphic presentations. Descriptive statistics could be used to explain FDT in the form of Relative Measures, Measures of Central Tendency and Measures of Dispersion. Please note that the principles and theories behind these statistics will not be discussed here. The assumption is that we have already treated this in earlier courses in statistics for social sciences. We shall therefore be more concerned with their application and place in data analysis.

Relative Measures

Relative measures also called relative statistics is used to measure the relative frequency of the different items of value of a variable presented. With this, the researcher could easily make comparisons within variables or categories. Relative measures that is commonly applied in describing FDT are proportion, percentage, ratio, and rates. For us in the media studies, percentages are the most commonly used relative measures. See Table 4.4 below:

Table 4.4:

Male	**Frequency**	**Percentages**	**Proportion**
Female	90	45	0.45
Total	100	55	0.55
	200	100%	1.00

Ratio on the other hand is used to compare two independent values by expressing one as a quotient of the other. A ratio measure used very often in describing research data is Sex Ratio. This is defined as the ratio of females to males in a population; sex ratio is usually expressed to a base of 100. In that case, it is interpreted as the number of males per 100 females in the population (Obikeze, 1990). The sex ratio is illustrated with the Table 4.5 below:

Table 4. 5: Sex Ratio of 60 Respondents in Trans-Ekulu

Male	Female	Sex Ratio
6	9	67
9	11	82
14	11	127
29	**31**	**94**

Source: Obikeze, 1990

Table 4.5 shows that while there were 94 (29/31*100) males to every 100 females in the population selected from Trans-Ekulu, there were 67 (6/9*100) males with primary education or less for every 100 females with the same level of education but 127 (14/11*100) males with college or higher for every 100 females with similar level of education in the population.

Rates, on its part, measure the average incidence or occurrence of an event over a total population that is exposed to the risk of that event. It may also be considered as the average unit of one variable (the criterion variable) per unit value of another variable (the norming variable) [Obikeze, 1990]. Thus, we can talk about a speed rate of 67 Kilometres per hour. Other types of rates are Birth rates, Marriage rates, Death rates, Crime rates etc. Rates could be computed over a base of 100, more often 1000 or 10,000. The formula for rate is presented below:

X/N * 1000/1 (if calculating with a base of 1000). We will illustrate this using Table 6 below**:**

Table 4.6: Marriage Rate among Christians and Non-Christians at Trans-Ekulu

Religion	Population	N0.of Marriages	Marriage Rate(‘000)
Christians	35	6	171
Non-Christians	25	5	200
Total	60	11	183

Table 4.6 shows that the marriage rate was 171 per 1000 among Christians of Trans-Ekulu and 200 per 1000 among Non-Christians.

Inferential Statistical Tools and How to Apply Them

We have already identified inferential statistics is a branch of statistics that studies a group of data in order to use the results obtained in making generalisation on a larger group of data. In other words, statistical inference is the use of sample results to reach conclusions about

populations from which the samples have been drawn. Some of the more commonly needed inferential statistical tools for research in communication are the chi-square (fairly weak), t-test of independent/non-independent means, Analysis of Variance (ANOVA), Tukey test, multiple analysis of variance, analysis of covariance, multiple analysis of covariance, and so on (Wilson, Esiri & Onwubere, 1999). The most popular inferential statistic tool commonly found in research projects of graduating students in the Mass Communication Departments is the chi-square test statistic. I shall illustrate application of inferential statistics with a typical example of a chi-square computation as presented below:

Example

A researcher is interested in finding out if the perception of Nigerian media experts on the potential influence of globalisation and ICT on local cultural values is directly related to the knowledge of such experts in the application/use of ICT. Table 1 below shows the observed frequency for the hypothesis.

Table 4. 7: Observed Frequencies for the Hypothesis

Perception on Cultural Influence	Knowledge in the Use of ICT		
	Positive	**Negative**	**Not Sure**
Strongly Disagree	04	02	01
Disagree	-	-	-
Undecided	23	04	02
Agree	190	08	02
Strongly Agree	29	04	01

Use the formula $X^2 = \sum \frac{(O^2)}{(E)} - N$

Test the null hypothesis that the perception of Nigerian media experts on the potential influence of globalisation and ICT on local cultural values is not directly related to the knowledge of such experts in the application/use of ICT at the 5% alpha level of significance. Do not forget to properly state the hypothesis (H_0 and H_1) for this enquiry and present your answer to two decimal places.

Hypotheses

H_0: The perception of Nigerian media experts on the potential influence of globalisation and ICT on local cultural values is not directly related to the knowledge of such experts in the application/use of ICT.

H_1: The perception of Nigerian media experts on the potential influence of globalisation and ICT on local cultural values is directly related to the knowledge of such experts in the application/use of ICT.

To compute for Expected E, the formula is E is usually gotten from **CM x RM**

N
Where
CM = Column Marginal
RM = Row marginal
N = Total number of frequencies

X^2 Computations

0	**E**	**0^2**	**$0^2/E$**
04	6.38	16	2.51
02	0.47	04	8.51
01	0.16	01	6.25
23	26.42	529	20.02
04	1.93	16	8.29
02	0.64	04	6.25
190	182.22	36100	198.11
08	13.33	64	4.80
02	4.44	04	0.90
29	30.98	841	27.15
04	2.27	16	7.05
01	0.76	01	1.32
			291.16

Applying the formula $\mathbf{X^2} = \sum \frac{(O^2)}{(E)} - N$

$X^2 = 291.16 - 270$
$= 21.16$

D.F. = (C-1) (R-1) = (3-1) (5-1)
= 2 x 4 = 8
:. X^2 (alpha= 0.05, 8) = 15.507

Since X^2 computed (21.16) is greater than X^2 tabulated (15.507), i.e. lies in the rejection region, we reject the H_0 and accept H_1 (There is a significant difference in observed and expected data or frequencies).

[Please, Note: A step by step procedure on how to calculate Chi-square is found in Module Six, Unit 4.]

We have listed many statistical test tools for data analysis but not all of them apply in every situation. When and how do we know the appropriate test statistics to use? Wilson, Esiri and Onwubere (1999) attempted an answer thus:

- The hypothesis to be tested should be stated – null or nondirectional, alternative – non-directional or directional. Statistical test to use (see the presentation or statistical tools and their uses) must be appropriate to the type of data to be analysed and the kind of comparison being made. Level of significance must be stated: the researcher usually decides this. Usually, the level is 5% or 1%. Sampling distribution is usually stated as a basis for determining the degrees of freedom. For t, the df is n1 + n2 – 2; for ANOVA, the df is (between – groups is k – 1 and within group is N – K). Critical region is determined from the critical value table against which the calculated value is compared for making a decision.

3.9 Post Analysis Evaluation

Post data analysis evaluation in communication research is a very important stage in the research process. It affords the researcher the opportunity to take another look at the entire research process and ask some vital questions that could bring closure to some gaps that were found during the data processing and management process. According to Wilson, Esiri and Onwubere (1999).

Some researchers carry out the evaluation consciously while others do so unconsciously. Post data analysis evaluations are aimed at finding answers to some questions arising during the analysis of the study data. Among such questions are: was there a possible wrong entry? For example, in a situation where the age of an informant was written as 198 could it have been 98 or 19? In few cases of no results, what factors are responsible for this? In most cases, the magnitude of these errors cannot be assessed before the post analysis stage, post analysis evaluation provide an opportunity to assess the magnitude of the errors (Andrew et al, 1974).

4.0 CONCLUSION

Data analysis is the focus of this course and this unit proper. We identified the need for data analysis to include the attempt to make raw data make more meaning to the readers and users. Data analysis is therefore very important in making inferences from research.

5.0 SUMMARY

This unit focused on data analysis with particular reference to the meaning of data, types of data, sources of data, measurement of data, meaning of data analysis, types of data analysis, major descriptive statistical tools and how to apply them, inferential statistical tools and how to apply them, and post data analysis evaluation.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss three descriptive statistics tools you can use in data analysis.
ii. Discuss how inferential statistics helps data analysis.
iii. Discuss the need for post data analysis evaluation.

7.0 REFERENCE/FURTHER READING

Bruning, J.L., & Kintz, B.L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Champion, D.J. (1981). *Basic Statistics for Social Research*. NY: Macmillan.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Nwabuokei, P.O. (1999). *Fundamentals of Statistics*. Enugu: Koruna Books.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Okeke, M.C. (1997). *Statistics: A Business Approach*. Onitsha: Onwubiko Prints Ltd.

Otokiti, S. *et al.* (2007). *Contemporary Statistical Methods*. Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd Ed.) Lagos: Idosa Konsult.

Spiegel, M.R. (S.I. Edition). *Theory and Problems of Statistics*. New York: McGraw-Hill.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

MODULE 3 QUALITATIVE DATA ANALYSIS

Unit 1 Overview of Qualitative Research and Types
Unit 2 Data Analysis in Focus Group Discussion
Unit 3 Analysing Case Study
Unit 4 Analysing Data from Observation Technique

UNIT 1 OVERVIEW OF QUALITATIVE RESEARCH AND TYPES

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
  3.1 Overview of Qualitative Research and Types
  3.2 Preparing Data for Analysis
  3.3 Techniques of Qualitative Analysis
  3.4 Reliability and Validity in Qualitative Techniques
  3.5 Elements of Qualitative Data Analysis
  3.6 Writing the Qualitative Research Report
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

This unit focuses on qualitative analysis of data with a summary overview of qualitative research techniques and types. Qualitative data come in various forms and so have various approaches to analyse such data. The unit also considered the two best known techniques used in qualitative analysis of data and presents the format for reporting the results or findings.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- present an overview of qualitative research techniques or types
- identify focus group discussion, ethnography, cultural studies, reception studies, in-depth interview, and textual analysis as common qualitative types
- discuss how to prepare qualitative data for analysis
- discuss the techniques of qualitative analysis

- evaluate issues of reliability and validity in qualitative techniques
- describe elements of qualitative data analysis; and how to write the qualitative research report.

3.0 MAIN CONTENT

3.1 Overview of Qualitative Research and Types

Scholars have disagreed on a common definition of the term qualitative. However, Wimmer and Dominick (2011) conclude that the term refers to:

- a broad philosophy and approach to research
- a research methodology
- a specific set of research techniques.

Before we discuss specific types of qualitative research techniques, we will first discuss qualitative data and techniques in general.

Qualitative data are gathered or generated in various forms. These include:

- notes made while observing in the field
- interview transcripts
- documents
- diaries
- journals.

Unlike the quantitative technique which waits until all the numbers are in before analysis begins, data analysis in qualitative researches is done early in the data collection process and continues throughout the project (Wimmer & Dominick, 2011).

Furthermore, qualitative researchers use an inductive approach or method by which data relevant to the topic(s) are collected and grouped into appropriate and meaningful categories. Explanations consequently emerge from the data themselves (Wimmer & Dominick, 2011). Examples of qualitative research methods include focus group discussion, ethnography, cultural studies, reception, in-depth interview, and textual analysis.

3.2 Preparing Data for Analysis

To enhance working with the large amounts of data generated by qualitative technique, Wimmer and Dominick (2011:119-120) recommend that the following preparations should be carried out:

- the data should be arranged in chronological order according to the sequence of events that occurred during the study
- furthermore, each piece of information should be coded to identify the source
- multiple photocopies of notes, transcripts and other documents generated during the investigation should be made to avoid loss (when they are in single copies)
- organise the data into a preliminary category system. These categories might arise from the data themselves, or they might be suggested by prior research or theory. E.g. a qualitative study of teenage radio listening might produce many pages of interview transcript. The researcher could read through and might categorise factors that influence teenage radio listening as "peer pressure", "escape", and "timing" etc. These categories, although, temporary comes from the research itself
- other researchers might prefer to have a room to themselves where they could easily display data visually on bulletin boards etc. This "analytical wallpaper" method is particularly helpful when there are several members of the research team working on the project because it is an efficient way to display the data to several people at once

Finally, the researcher, who is the main instrument in qualitative data collection and analysis, must do some preparation on himself before beginning the investigation. Maykut and Morehouse (1994) as cited in Wimmer and Dominick (2011) call this preparation 'Epoche'.

Epoche is the process by which the researcher tries to remove or at least become aware of prejudices, viewpoints, or assumptions that might interfere with the analysis. It helps the researcher put away personal viewpoints or prejudices so that the phenomenon under study may be seen the way it is.

SELF-ASSESSMENT EXERCISE

From the perspective of Wimmer and Dominick (2011) identify and discuss six preliminary steps a qualitative researcher could undertake to ensure a valid data analysis in any of the qualitative type studies.

3.3 Techniques of Qualitative Data Analysis

Many techniques are available to analyse qualitative data. We shall, however, discuss only two of the best know techniques as presented by Wimmer and Dominick (2011: 120-121).

1. The Constant Comparative Technique (CCT)

a. This technique according to Wimmer and Dominick (2011), was first formulated by Glaser and Strauss (1967) and was modified by Lincoln and Cuba (1985). At a general level, the process consists of four steps:

- comparative assignment of incidents to categories
- elaboration and refinement of categories
- searching for relationships and themes among categories
- simplifying and integrating data into a coherent theoretical structure.

b. **Comparative Assignments of Incidents**: to categorise here entails putting each unit of analysis into a set of provisional categories. It is possible that some initial categories may have only one or two incidents assigned to them while others may have many. Also, if some incidents do not fit into existing categories, more are created. Till the end of the analysis, the categories are subject to continuous refinement until the exhaustive, valid and mutually exclusive categories are identified (Wimmer & Dominick, 2011).

c. **Elaboration and Refinement of Categories:** this entails modification to accommodate complex incidents. At this stage, the researcher establishes rules that will guide what should or should not be in a particular category. These rules help to focus the study as well as allow the researcher to start the exploration of theoretical dimensions of the emerging category system. These rules ultimately help to reveal what you are learning about your chosen topic and also help you to determine your research outcome (Wimmer & Dominick, 2011).

d. **At the Stage of Searching for Relationships and Themes among Categories**, the researcher examines the propositional statement and looks for any meaningful connections. Some propositions are probably strong enough to stand alone; others might be related in several basic ways. Whatever the situation, the goal of this stage is to generate assertions that can explain and further clarify the phenomenon under study (Wimmer & Dominick, 2011).

e. **At this Stage, all the Results of the Foregoing Analysis are Integrated into Some Coherent Explanation of the Phenomenon**. The goal at this stage of the study is to arrive at an understanding of the people and events being studied (Wimmer & Dominick, 2011).

2. The Analytic Induction Strategy (AIS)

This technique blends hypothesis construction and data analysis. According to Wimmer and Dominick (2003) citing Stainback and Stainback (1988), this technique consists of the following steps:

a. **Define a topic of interest and develop a hypothesis**. This is usually the first step in the analysis process. The hypothesis will help to focus the research.

b. **Study a case to see whether the hypothesis works**. If it does not work, reformulate the hypothesis. You need to try out the hypothesis on a case to determine its validity. This trial will help you to determine if you could go on with the current hypothesis or if you need a new or amended version.

c. **Study other cases until the hypothesis is in refined form**. One case study may not be enough to decide if the hypothesis is in order. So, you need to study more cases to again determine that the hypothesis is in order. The result here may also lead to more refinement of the reformulated hypothesis.

d. **Look for "negative cases" that might refute the hypothesis and then reformulate again.** It is also important to consciously look for negative cases that may disprove the hypothesis. The attempt to get to this stage will see you further refine the hypothesis to solid state.

e. **Continue until the hypothesis is adequately tested.** You continue the refinement and reformulation process until you think the hypothesis has been adequately tested.

Note very carefully that in this method, an explanation for the phenomenon, in the form of a hypothesis, is generated at the beginning of the study. This contrasts with the constant comparative technique, in which an explanation is derived as the result of the research.

3.4 Reliability and Validity in Qualitative Techniques

A study is reliable when repeated measurement of the same material results in similar decisions or conclusions. Validity, on the other hand, is usually defined as the degree in which an instrument actually measures what it sets out to measure. These two concepts both help readers determine how much confidence can be placed on outcomes of

qualitative studies and how credible the results are truly are (Wimmer & Dominick, 2011).

Factors that can Affect the Credibility of Qualitative Studies

a. **The Issue of the Completeness of the Data Collected**
If there was laziness and sloppy attitude in the data gathering process, then obviously, the results will never pass the test of credibility. The data should be complete in the sense of not leaving out any vital information in the data gathering process. Anything less than complete information in the data gathering process injures the entire process.

b. Selective Perception is Another Major Factor
A qualitative researcher should not dismiss data that do not fit his/her favoured interpretation of the subject. He must analyse all cases and give reasons as to why some data do not fit. This is the major reason it is advocated that every researcher must rise to the level of being regarded as a professional. He must purge himself of all prejudices that may bias the outcome of the study.

c. The Question of Reactivity
Reactivity comes up when the object being observed change because they are being observed. Some schools of thought believe that subjects may behave differently if they know that they are not being observed. Reactivity is a very difficult challenge for researchers but they must try to minimise it whenever possible so as to increase the credibility of the study.

Factors that Help Increase the Credibility of Qualitative Data

a. The Use of Multiple Methods of Data Collection
This factor is similar to the notion of triangulation. The use of interviews along with field observation and analysis of existing documents suggests that the topic was examined holistically. This helps to build credibility in the findings as well as introduces other perspectives that a single method may have ignored.

b. Audit Trail
This is essentially a permanent record of the original data used for analysis and the researcher's comments and analysis methods. Audit trail allows others to examine the thought process involved in the researcher's work and allows them to assess the accuracy of his/her conclusions. This is basically about getting an outside view point on the procedure that may have influenced the position taken.

c. Member Checks
This allows research participants access to the notes and conclusions of the researcher. This helps the participants to

determine whether the researcher has accurately described what he/she was told. This is quite effective and corrective in approach. The research participants have an opportunity to determine if the researchers have adequately captured what they were thinking or what they actually meant in responses they gave. Any error of judgment or misconception or misinterpretation identified during this process could be handled at this stage making this a practical way of increasing the credibility and acceptability of a research finding. Member check is a solid validation process in qualitative studies.

d. Research Team

This method or factor assumes that team members keep one another honest and on target when describing and interpreting their data. Sometimes an outside person is asked to observe the process and raise questions of possible bias or misinterpretation where appropriate.

e. Debriefing

This consists of having an individual outside the project question the meanings, methods, and interpretations of the researcher. The researcher is obligated to report the entire process to this individual who is at liberty to ask all manner of question regarding how the research was conducted and the final outcome. His or her views on the entire process are very vital to the credibility of the research outcome.

3.5 Elements of Qualitative Data Analysis

Many elements make up the data analysis process in qualitative studies. However, the following points discussed below generally situate qualitative data analysis in the parlance of empiricism. These elements are:

a. **Analysis is Holistic and Multi-Dimensional:** rather than concentrate on specific key variables, qualitative analysis attempts to grasp the totality of the socio-cultural condition, the historical development, and the entirety of what that research subject represents. In essence, all of these help the researcher to address the problems or objectives holistically.

b. **Analysis is Descriptive, In-Depth and Longitudinal:** the analysis is also very descriptive in terms of specifying what transpired; in-depth in terms of going deep in revealing insightful details; and longitudinal in terms of the length of time and period the analysis covers so as to situate the research as well as the findings in a period in history.

c. **Analysis is Naturalistic:** this actually implies that the analysis deals with direct observations under the natural settings of the

subject under investigation. The implication again is that the data gathered are natural behaviors of people or the subject and so the analysis produces findings that relate to the real world; not something imagined in the rarefied mind of the researcher. It is devoid of artificialities and computational manipulations of the original information to meet some standardised way of doing research.

d. **Analysis is Humanistic:** qualitative data analysis permits the researcher to interact closely with the data and to look at the problems from the insider perspective. By making human beings and their behavior become only statistical figures or numbers, quantitative analysis dehumanises human beings. In contrast, qualitative analysis is dialectical and interactive.

e. **Data for Analysis may be Recorded and Presented in Oral or Written Forms, in Audio Forms or in Any other Visual or Art Forms:** this enriches the accessibility mode for all those who are interested in the outcome. Its ability to be produced in all of these forms is also a plus in terms of acceptability and credibility.

f. **Analysis is generally not aimed at Testing Specific Hypothesis and the Data usually cover a Few Numbers of Cases:** this is another unique element of qualitative data analysis. The analysis does not concentrate on testing any particular hypothesis as is the case in quantitative data analysis. Here, the so-called hypothesis is never a finished outcome as it continues a refinement process even while the analysis remains ongoing.

3.6 Writing the Qualitative Research Report

Writing the qualitative research report is a very complicated process compared to the straight forwardness involved in quantitative research reports. Qualitative data come in form of sentences, extended quotes, paragraphs of descriptions, and even pictures and diagrams. So it is very difficult to condense these into numerical tables and charts and then begin to subject human actions and reasoning to mere numbers.

The point being made here is not to lampoon or simplify quantitative data analysis; rather, attention needs to be pointed to the fact that there are human actions that cannot be adequately represented in figures or percentages or any other form of quantitative representation. In this regard, the fact that 90 percent said so is not the issue but that there is a unique phenomenon coming out of observing a subject for a long time in a particular research situation.

Qualitative research reports is written in a loose narrative style that employs many devices used in novels and short stories. Keyton (2001)

as cited in Wimmer and Dominick (2011) describes three separate styles that can be adopted by authors of qualitative research reports thus:

a. **Realist**: a dispassionate third-person point of view that removes himself from the subject of analysis
b. **Confessional**: this is a first-person point of view that reveals much about the author alongside issues of the analysis
c. **Impressionist:** in this situation, the writer uses metaphor and vivid images to get his point across to his audience. This is a somehow flamboyant qualitative data analysis reporter.

General Format for Structuring Qualitative Report

The Introduction
The introduction of qualitative studies is similar to that of quantitative studies. The introduction provides an overview of the project, the precise research questions, problems statement, objectives, justification, and why the reader should be interested in reading the research report. There is no difference with the structure of quantitative reports except with the content of each report.

The Literature Review
Unlike the quantitative research reports, the literature review section of a qualitative research report is not usually extensive because the qualitative researcher prefers not to have an exhaustive literature review for fear of unduly influencing his/her perception of the research situation and thus bias the environment of the research.

The Method
This section includes a number of issues that concern how the research was done. It is expected that this particular section deals with the followings issues or areas of the research:

- The method or methodology used to collect data and why that particular method was preferred over other useful methods
- The research setting must also be explained
- The sampling technique must be described and its appropriateness explained
- The sample size must also be justified and reasons for selection further explained.
- The data collection methods must also be mentioned here in greater details to enable replication of the entire process for other follow up studies in the area.

The Findings
This is the longest section in the qualitative research report. The writer should avoid the report being too thick and voluminous. He should also avoid very thin report. Two guiding principles, according to Wimmer and Dominick (2011) are:

a. **First,** remember that it is impossible to say everything in a research report. Try to select those vignettes, quotes, or an example that most vividly illuminates the findings.
b. **Second,** choose the data that illustrates the variety of information collected, including situations that are common or atypical.

An overall organisational scheme will make the report more understandable. According to Chenail (1995):

- organise the material (data) chronologically
- present the most important findings (results) first
- use a powerful presentation and save the most important points till the end
- arrange the data according to some theoretical or conceptual scheme.

Note: The findings section should strike a balance between description and analysis. Detailed quotes or examples should be followed by analysis and generalisations, with the primary aim to offer further explanations.

The Discussion
This section is about the implications of the research findings with regard to the research objectives, research questions, hypotheses and the problem. It should also include a summary, additional implication of the study for further studies, and a discussion of strengths and weaknesses of the study.

4.0 CONCLUSION

The framework for qualitative research is an interesting area that researchers must pay greater attention to. Qualitative research can never be overwhelmed by quantitative techniques because of its uniqueness and dealing with research subjects in all its nuances. Following from this, the way or methods of analysing qualitative researches, which have preoccupied our attention in this unit becomes pertinent.

5.0 SUMMARY

This unit is a detailed discussion on analysing qualitative research data with specific attention paid to issues of qualitative research techniques or types; how to prepare qualitative data for analysis; discussion on the various techniques of qualitative data analysis; an overview of reliability and validity in qualitative data analysis and techniques; discussion on the elements of qualitative data analysis; and how to write the qualitative research report so that the outcome of the study is not lost in midst of poorly constructed statements and confusing words.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss the schema for reporting the analysed outcome of qualitative research.
ii. Discuss the core elements considered in qualitative analysis of data.
iii. Mention and explain four qualitative research techniques you are familiar with.

7.0 REFERENCES/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd ed.) Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 2 DATA ANALYSIS IN FOCUS GROUP DISCUSSION

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Overview of Focus Group Discussion
    3.2 How to Analyse Focus Group Data
    3.3 Writing the Focus Group Research Report
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

This unit discusses Focus Group Discussion, popularly referred to as FGD, which is a core type of qualitative research technique. A focus group is a small discussion on a specific subject that is moderated by a moderator who may be the principal researcher or investigator. Focus groups rely on the participants' words, action, inaction and other body gestures during interaction to provide information, insight, fill in the texture, and explain the reasons behind a subjects' behaviour in the light of the research (Poindexter & McCombs, 2000). This shall be the focus of this unit. However, particular attention will be paid to how FGD data are analysed.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define Focus Group Discussion
- describe how FGD data are analysed
- explain how to present the analysed results from FGD.

3.0 MAIN CONTENT

3.1 Overview of Focus Group Discussion

A focus group, according to Poindexter and McCombs (2000) "is a qualitative research method that uses open-ended, follow-up and probing questions to search below the surface of a small group of participants'

attitudes, opinions, and behaviours to understand motivations, feelings and reactions". Focus groups rely on the participants, words to provide insight, fill in the texture, capture the nuances, and explain why participants behave and respond the way they do (Poindexter & McCombs, 2000).

A focus group is a small discussion on a specific subject that is moderated by a moderator, who may be the principal researcher or investigator. Focus groups rely on the participants' words, action, inaction and other body gestures during interaction to provide information, insight, fill in the texture, and explain the reasons behind a subjects' behaviour in the light of the research (Poindexter & McCombs, 2000).

Purposes and Benefits of Focus Group Discussion

Poindexter and McCombs (2000:241) note that when used appropriately, a Focus Group Discussion could benefit the following or perform the underlisted purposes thus:

- A Focus Group Discussion can be an effective research tool and useful device for generating new insights and ideas about a research phenomenon
- Focus Groups can also be a helpful first step in developing an effective survey research instrument or questionnaire which could be used to explore new or novel topics
- It is also an effective approach to following up on specific survey results
- It can also be an effective way of soliciting consumer reactions to concepts, campaigns, creative products and services

Focus Group Discussion is a powerful means (if properly handled) of learning how consumers talk about issues, candidates, products, services, commercials, celebrities, news anchors, TV programmes, companies, institutions, government establishments and parastatals etc and the reason(s) why they do so.

Pre-Research Activities of a Focus Group Discussion Procedure

Poindexter and McCombs (2000:241) identify the following as the expected pre-research activities to be engaged by the Focus Group investigator:

- monitor the environment and literature
- identify problem, trend, opportunity, issue, phenomenon, or theory

- decide if a Focus Group is the most effective method for understanding the problem, trend, opportunity, issue, phenomenon, or theory
- specify Focus Group topic in research question form
- design Focus Group study using ethical standards
- decide type of Focus Group participants to recruit
- establish budget and timetable
- determine who will conduct Focus Group.

The Core Characteristics of a Focus Group Moderator

The focus group moderator is core to a successful focus group discussion. He or she is central to how effective and efficient focus group could turn out to be in gathering relevant and valid data that could provide a rich insight on a research subject. According to Poindexter and McCombs (2000:243), the following are the expected characteristics or features of a good Focus Group moderator:

- must be experienced and versatile in organising and analysing Focus Group activities
- must be thoroughly briefed about the subject of the focus group and should therefore have a clear and detailed understanding of the subject matter
- must be very confident in his or her abilities and capacity to deliver on expectation
- must be able to establish a rapport and control the group dynamics.
- must also be able to make the participants feel very relaxed, at home and very comfortable
- must be an active listener with particular attention to details
- he or she must also be a very flexible, quick thinker who is able to quickly respond confidently to unexpected events
- should have the capacity to ask follow up questions that were not scripted
- a focus group moderator should have and respect people, especially recognise and respect individual differences
- he is also highly professional and ethical
- he or she is also very considerate of people.

SELF-ASSESSMENT EXERCISE

i. Discuss extensively five solid qualities expected of any researcher who moderates Focus Group Discussion
ii. Discuss five preliminary activities that a Focus Group investigator should engage in before the research proper.

3.2 How to Analyse Focus Group Data

According to Poindexter and McCombs (2000:262), analysis of focus group discussion data varies significantly from the analysis of survey data. “Within days of the completion of the focus groups, the research expert should review the videotapes and take notes on general themes and verbatim quotes. By using the discussion guide as an outline, the research expert will be able to easily pick out the more important findings”.

In summary, Poindexter and McCombs (2000:243), note the following as effective process of analysing focus group discussion Data (there is no form of chronological display here).

Review your videotapes and take notes on general themes and verbatim quotes

Use the discussion guide as an outline to pick out vital details from the data

Transcribe your tapes and analyse the transcript. This should not in any way replace the viewing of the videotapes which will help you “capture the texture, the tone, the nuances, the mood, the hearts and the minds of real people and how they feel and talk about the products, services, issues, ideas, and images of concern to the decision maker”. It is therefore better and preferable to view the videotape than rely solely on cold transcripts of what was said

Make sure you summarise the biographical details of the participants. This will help place the results in the proper context

However, do not struggle to analyse your focus group data to become generalisable to the population. This is not possible and so not necessary. Your attention is to make sure all nuances, insights and details etc that are displayed by the participants on the subject matter are adequately captured

Make sure ethical standards are not compromised. It is very easy to dismiss views and comments that do not reflect preconceived notions of the researcher because of the qualitative nature of focus group discussion. This should be avoided in all its ramifications. Furthermore, ethical standards dictate that analysis and interpretation of results should accurately and fairly report what the Focus Group participants said and not what the decision maker wanted to hear (Poindexter & McCombs, 2000:243).

3.3 Writing the Focus Group Research Report

The focus group discussion report contains almost similar elements like the survey research report. The components of the reports are presented below as conceptualised by Poindexter and McCombs (2000:244).

- Cover Letter
- Report Binder
- Title and Author Page
- Table of Contents
- Executive Summary
- Introduction
- Background/Context
- Description of Method
- Results
- Discussion, Limitation, and Recommendations
- Conclusions
- Endnotes
- References
- Appendix

4.0 CONCLUSION

Ideally, Focus Group Discussion is a qualitative research method that provides texture, context and deeper insights that cannot easily be gleaned from normal survey research method. This places it at a very high premium. However, as earlier noted, the inability to generalize Focus Group results from sample to the population remains a limiting factor in its use. Otherwise, it is a valid method of getting greater details in an investigation that has human angle or interest.

5.0 SUMMARY

This unit focused on one of viable qualitative research methods: Focus Group Discussion, which is a research technique that brings together participants who may have same or similar characteristics with regards to the subject of investigation. We did a comprehensive overview of the method with emphasis on how to analyse the data from the technique as well as present the analysed results.

6.0 TUTOR-MARKED ASSIGNMENT

i. You are asked to determine the views of the youth on dress culture in the universities. Discuss the groups you will target for

the Focus Group and why. Develop a moderator's outline and discussion guide for the Focus Group.

ii. Discuss the processes involved in the analysis of Focus Group data.

7.0 REFERENCES/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd ed.) Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 3 ANALYSING CASE STUDY

CONTENTS

1.0 Introduction
2.0 Objective
3.0 Main Content
    3.1 Overview of the Case Study Approach
    3.2 Purpose of a Case Study
    3.3 Types of Case Studies
    3.4 Conducting a Case Study
    3.5 Data Analysis in Case Study
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

Case studies, according to Paula and Maxwell (2000:289), can be thought of as a study of a sample of one subject matter. Continuing, they noted that "case studies can be used to study an individual, institution, organisation, event, issue, or some type of phenomenon. The case study researcher examines his or her subject in-depth, conducting hundreds of interviews and reviewing hundreds of contemporary and historical records in order to understand everything about the research topic." This is the focus of this particular unit meant to present more issues around the area of qualitative studies and data analysis.

2.0 OBJECTIVE

At the end of this unit, you should be able to:

- analyse data from case studies.

3.0 MAIN CONTENT

3.1 Overview of the Case Study Approach

According to Poindexter and McCombs (2000:289), case studies can be thought of as a study of a sample of one while Wimmer and Dominick (2011:140-141) see case study as a qualitative research technique which uses as many data sources as possible to systematically investigate individuals, groups, organisations, or events. They further note that case studies are conducted when a researcher needs to understand or explain

a phenomenon as frequently used in medicine, anthropology, clinical psychology, management science, and history.

Yin (2003) cited in Wimmer and Dominick (2011:141) defines a case study as: "an empirical inquiry that uses multiple sources of evidence to investigate a contemporary phenomenon within its real-life context, in which the boundaries between the phenomenon and its context are not clearly evident."

Merriam (1998) cited in Wimmer and Dominick (2011:141) lists four essential features of a case study research thus:

- **Particularistic:** this means that the case study focuses on a particular situation, event, program, or phenomenon, making it a good platform to study real life situations
- **Descriptive:** this means the outcome of a case study is a detailed description of the topic under study
- **Heuristic:** this means that case studies help people understand what's being studied as well as help produce new meanings and fresh insights
- **Inductive:** this means most case studies depend on inductive reasoning whereby principles and generalisations emerge from an examination of the data. In this case, many case studies are carried out to discover new insights rather than verify existing knowledge.

In contrast to a survey technique that tries to define the phenomenon under study narrowly enough to limit the number of variables to be examined, case study research includes both single cases and multiple cases. It is the preferred strategy when "how and why" questions are asked. Likewise, it is the preferred method when the researcher has little control over events, and when the event is contextualised within a real life situation. That explains why case study is often used interchangeably with ethnography, field and participant observation (Wimmer & Dominick, 2011:141).

Wilson, Esiri and Onwubere (n.d.) note that the underlying philosophical assumptions of case study are similar to the above mentioned types of qualitative research methods because they all take place in a natural setting (such as classroom, neighbourhood, or private home) and strives for a more holistic interpretation of the event or situation under study. This they further observed is unlike more statistically-based studies which search for quantitative data.

SELF-ASSESSMENT EXERCISE

i. Discuss four characteristics associated with case studies.
ii. Do an in-depth analysis of two valid definitions of case studies.

3.2 Purpose of Case Study

According Wilson, Esiri and Onwubere (n.d.) citing Osuala (2005:186-187), the purposes of a case study research include the followings:

- **It is valuable as preliminaries to major investigations**. Because they are so intensive and generate rich subjective data they may bring to light variables, phenomena, processes and relationships that deserve more intensive investigation. In this way a case study may be a source of hypotheses for future research.
- **A case study fits many purposes,** but most case studies are based on the premise that a case can be located that is typical of many other cases. Once such a case is studied it can provide insights into the class of events from which the case has been drawn.
- **A case study may provide anecdotal evidence** that illustrates more general findings.
- **A case study may refute a universal generalisation.** A single case can represent a significant contribution to theory building and assist in refocusing the direction of future investigations in the area.
- **A case study is preferred when the relevant behaviours cannot be manipulated**.
- **A case study may be valuable in its own right as a unique case**. Gruber's (1974) study of Darwin and the processes by which he arrived at the theory of evolution is an example of this.

3.3 Types of Case Studies

Wilson, Esiri and Onwubere (n.d.) citing Osuala (2005:187-188), identify types of case studies as follows:

- **Historical Case Study:** these studies trace the development of an organisation/system over time.
- **Observational Case Study:** these studies often focus on a classroom, group, teacher and pupil, often using a variety of observation and interview methods as their major tools.

- **Oral History:** these are normally first person narratives that the researcher collects using extensive interviewing of a single individual.
- **Situational Analysis:** particular events are studied in this form of case study, for example, an act of student vandalism could be studied by interviewing the concerned, the parents, the teacher, the chairman, witnesses, etc.
- **Clinical Case Study:** this approach aims to understand in depth a particular individual, such as a child having problem with reading, a newly transferred student in his first term at school or a teacher with disciplinary difficulties.
- **Multi-Case Studies:** a collection of case studies, that is the multi-case study, is based on the sampling logic of multiple subjects in one experiment.

3.4 Conducting a Case Study

Wimmer and Dominick (2011:142) observe that the precise method of conducting a case study has not been as well documented as the more traditional methods like the survey and the experiment. Wimmer and Dominick (2011:142), however, identify five steps in the method of conducting case study research to include:

- **Design:** the design of a case study concentrates on what to ask in terms of how, why and what to ask.
- **Pilot:** this is usually used to refine both the research and the field procedures.
- **Data Collection**: a case study relies on (a) documents (b) Archival records (c) interviews (d) direct observation (e) participant observation (f) physical artefacts
- **Data Analysis**: this is the focus of the next section.
- **Report Writing**: according to Wimmer and Dominick (2011), "the case study report can take several forms. The report can follow the traditional research study format-problem, methods, findings, and discussion- or it can use a non-traditional technique.... No matter what form is chosen, the researcher must consider the intended audience of the report."

3.5 Data Analysis in Case Study

Wimmer and Dominick (2011:143-144) note that the data analysis stage is the most difficult stage to deal with in case studies because there are no specific formulas or techniques for analysing case study data. However, Yin (2003) cited in Wimmer and Dominick (2011:143-144) identifies three broad analytic strategies thus:

- **Pattern Matching Strategy**: in pattern matching empirically based pattern is compared with one or more predicted patterns. Wimmer and Dominick (2011:144) give a typical example thus:
    - Suppose a newspaper is about to initiate a new management tool: regular meetings between top management and reporters, excluding editors. Based on organisational theory, a researcher might predict certain outcomes – namely, more stress between editors and reporters, increased productivity, and weakened supervisory links. If analysis of the case study data indicates that these results do in fact occur, some conclusions about the management change can be made.
- **Explanation Building**: in the analytic strategy of explanation building, Wimmer and Dominick (2011) note that "the researcher tries to construct an explanation about the case by making statements about the course or causes of the phenomenon under study."
- **Time-Series Analysis**: in time series analysis, the researcher attempts to compare a series of data points to some theoretic trend that was predicted before the research, or to some alternative trend (Wimmer & Dominick, 2011:144). Wilson, Esiri and Onwubere (2008) give a typical example thus: "If for example several Nigerians cities have experienced newspaper strikes, a case study investigator might generate predictions about the changes in information seeking behaviours of residents in these cities and conduct a case study to see whether these predictions are supported."

4.0 CONCLUSION

Case studies draw from as many data sources as possible to give in-depth investigation to an event, issue or phenomenon under research scrutiny. Case studies are more preferred when a researcher is really interested in understanding a phenomenon beyond the ordinary level. Yes, case studies may lack the rigor associated with the scientific method (this is debatable) and so not easily generalisable from the case to the population: however, it still remains a valid means to investigate certain phenomenon that requires in-depth analysis.

5.0 SUMMARY

This unit has exposed you to the case study research approach with specific attention on the purpose, types, how to conduct it and how to analyse the data gathered through it.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss the three identified processes to analyse case study data.
ii. Discuss the major reason why Case studies are difficult to analyse.

7.0 REFERENCES/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd ed.) Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 4 ANALYSING DATA FROM OBSERVATION TECHNIQUE

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Overview of the Observation Research Technique
    3.2 Types of the Observation Research Technique
    3.3 Analysis of the Observation Research Technique
    3.4 Interpretation of the Observation Research Technique
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

Since the beginning of this module, we have been discussing qualitative studies and the various methods of doing it. This unit is an extension of such discourse as we again consider the method of observation which is considered a useful way of collecting data from the field. Like all qualitative technique, observation is concerned more with description and explanation than measurement and quantification.

According to Wimmer and Dominick (2011), field observation is classified along two major dimensions namely:

- The degree to which the researcher participates in the behaviour under observation
- The degree to which the observation is concealed.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- explain the observation research technique
- list the various types of the observation research technique
- describe the procedure and method of analysis of the observation research technique
- interpret the results gotten from observation.

3.0 MAIN CONTENT

3.1 Overview of the Observation Research Technique

There are various aspects of human behaviour that cannot be comprehended without an in-depth understudy of such situation. Such understanding could come in the form of observation of the subject with the intent of understanding the phenomenon under question. In other words, observation is one of the methodological tools that are applied in communication research to elicit evidence and subsequently explain certain behaviour traits associated with the subject of the observation.

According to Wimmer and Dominick (2011) field observation is basically concerned with the planned watching, recording and analysis of observed behaviour as it occurs in its natural setting. Field observation has long been used by mass communication researchers to deal with some research issues or subjects. For example, Gieber's (1956) seminal study of gatekeeping in the newsroom and Epstein's (1974) description of network news operations are classic examples of studies carried out through the observation technique.

Field observation as a qualitative research technique has many advantages and disadvantages. Some of these have been detailed by Wimmer and Dominick (2011:125-126) thus:

Advantages of Field Observation

a. Helps the researcher to define basic background information necessary to frame a hypothesis and to isolate independent and dependent variables.
b. Since data are gathered first hand, observation is not dependent on the subjects' ability or willingness to report their behavior.
c. Observation also provides access to groups that would otherwise be difficult to examine. E.g. producers of x-rated movies may not fill and return questionnaires but mutual trust could be developed with them by an observer and persuaded to respond to rigorous questioning.
d. It is also inexpensive in terms of human and material resources put in to it.
e. The study takes place in the natural setting of the activity being studied and this can provide data rich in detail and subtlety.

Disadvantages of Field Observation

a. Field observation is a poor choice if the researcher is concerned with external validity (generalisability of findings to the population)
b. In field observation, experimental bias is very high
c. Field observation also suffers from **Reactivity**- this means that the behavior under study may be influenced to react to a particular and usually expected way as a result of being observed.

As noted by Wilson, Esiri and Onwubere (2008), observation as a research method has a long history in psychology, anthropology and sociology. Its early use was in these three disciplines in the social sciences. It was rarely used in mass media research before1980. Recently, they continue, field observations have become commonplace in communication research. Citing Berger (2000:162), Wilson, Esiri and Onwubere (2008) argue thus:

- It is important, however, to state that scientific observation or observation as a research technique is different from everyday observation. The latter is random and fugitive i.e. quickly moving on to other matters. The former (scientific observation) is focused-on what the observer wants to find out – and it is objective and systematic

There are six stages in a typical field observation research. These according to Wimmer and Dominick (2011:126-127) include:

- Choosing the Research Site
- Gaining Access
- Sampling
- Collecting Data
- Analysing Data
- Exiting.

Our major concern in this course is stage five, which is analyzing research data gotten from field observation.

SELF-ASSESSMENT EXERCISE

i. Describe four advantages associated with the observation research technique.
ii. Discuss the procedure in observation research method.

3.2 Types of Field Observation

Field observation is basically of two types:

a. Participant observation and
b. Non participant observation

a. Participant Observation

Participant observation, according to Sobowale (1983:11), is an intensive and more involved way of gathering information through observation. However, Berger (2000:161) as cited in Wilson, Esiri and Onwubere (2008) defines it as a qualitative research technique that provides the opportunity to study people in real life situations. "It is a form of "field research" in which observations are carried out in real settings and where there is a lack of the kind of control and structure present in experiments. In participant observation, researchers become involved in some group or organisations or entity that they are studying."

According to Wilson, Esiri and Onwubere (2008), "the active participant observer is a member of the group he is trying to observe. He normally seeks the membership of the group and when accepted, he accepts the norms of the group and obeys its rules and regulations." However, care must be taken to conceal the real intention of joining the group because the success or failure of this method depends on this.

b. Non Participant Observation

The non participant observation is the second type of the field observation technique whereby the investigator takes himself or herself away from the research subject in the observation process. In other words, the investigator does not need to become an active member of the group he or she is observing to observe the group. This type of observation is therefore called observation from a distance. The investigator is not involved or engaged in the activity of the group being observed.

SELF-ASSESSMENT EXERCISE

i. Why will a researcher think twice about using participant type of field observation?
ii. Discuss the implications of being discovered in a participant type of field observation.

3.3 Analysis of the Observation Research Technique

According to Wimmer and Dominick (2011), the overall goal of data analysis in field observation is to arrive at general understanding of the phenomenon under study. They, however, note that data analysis in field observation consists mainly of the following:

- field observation primarily consists of filing the information and analyzing its content
- constructing the filing system. The purpose of the filing system is to arrange raw field data in an orderly format that is amenable to systematic retrieval later
- an observation may be placed in more than one category. It is preferable to make multiple copies of notes, periodic filing of notes during the observation to save time and confusion at this stage of analysis
- once all the notes have been assigned to their proper files, a rough content analysis is performed to search for consistent patterns
- flexibility is a key word in the analysis of field observation data. In field observation, the researcher or investigator can analyse the data during the course of the field study and later change the research design accordingly to suit emergent situations or needs.

3.4 Interpretation of the Observation Research Technique

According to Wilson, Esiri and Onwubere (2008:132), three approaches are available in interpreting field observation outcomes. The three approaches are:

a. Descriptive
b. Inferential and
c. Evaluation.

According to Wilson, Esiri & Onwubere (2008:132)

a. **Descriptive observational** variables require no inference making on the part of the researcher. You see something and write it down.
b. **Inferential observational** variables require the researcher to make inferences about what is observed and underlying emotion. For example, you may observe a journalist in the newsroom frowning. From this observation you may assume (correctly) that he or she is not happy.

c. **Evaluative observational** variables require the researcher to make an inference and a judgment from the behaviour. For example, you may question whether the frowning journalist and job satisfaction have a positive relationship. "Positive" is an evaluative judgement. It is noteworthy however that, when writing field note, the researcher should include descriptive as well as inferential data. It is important to describe the setting and the mood in a detailed manner. All such things that may change behaviour need to be noted.

4.0 CONCLUSION

Field observation method simply requires an observer-human or electronic, to observe and record the required items. Observation is a qualitative research technique, which is more concerned with description and explanation rather than measurement and quantification. This is what makes this research approach unique.

5.0 SUMMARY

This unit dealt with the observation research technique as one of the popular techniques used in qualitative studies. It is expected that you now the various types field observation, advantages and disadvantages of field observation, how to analyse the data from field observation and the methods of interpretations.

6.0 TUTOR-MARKED ASSIGNMENT

i. Distinguish between participant observation and non participant observation.
ii. Justify the use of field observation in communication research.
iii. How would you analyse data obtained from field observation?

7.0 REFERENCES/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D.O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, Idowu (2009). *Scientific Journalism* (2nd ed.) Lagos: Idosa Konsult.

Wilson, D.*et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

MODULE 4 QUANTITATIVE DATA ANALYSIS

Unit 1 Introductory Overview of Quantitative Analysis
Unit 2 Procedure for Quantitative Analysis
Unit 3 Construction of Frequency Distribution Tables
Unit 4 Survey Data Analysis and Statistics

UNIT 1 INTRODUCTORY OVERVIEW OF QUANTITATIVE ANALYSIS

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Ingredients/Elements of Quantitative Analysis
    3.2 Understanding Independent and Dependent Variables
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

As qualitative methods focus on describing the data in terms of depth and breadth, quantitative methods of analysis are also interested in describing the data but this time in terms of numerical values or quantifiable variables. Quantitative methods of data analysis, according to Obikeze (1990:87), require that the data be presented in numerical form. The process of analysis involves subjecting the data to varying degrees of mathematical and statistical manipulations so as to bring out the basic features, characteristics, trends, and relationships inherent in the body of data. This process, Obikeze (1990) continues, also facilitates the accurate identification, description and explanation of the nature of interrelationships in the study population.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- identify and describe the ingredients/elements of quantitative analysis
- show clear understanding of independent and dependent variables.

3.0 MAIN CONTENT

3.1 Ingredients/Elements of Quantitative Analysis

Ingredients or elements of quantitative analysis of data refer to those factors that must play up in a typical analysis of data using this platform. Quantitative methods of data analysis therefore imply that the following must be in place for one to comfortably say he or she has done the analysis using the quantitative platform.

Depending on the author, other elements could be identified; however, Obikeze (1990:87) is of the view that the following four elements constitute the ingredients of quantitative analysis of data:

- the data are numerical in form or otherwise quantified through the process of coding. In essence, before any analysis could be said to have taken place quantitatively, the data must be in numerical form or coded numerically
- it must be possible to differentiate among the independent, dependent and intervening variables and also to categorise such variables in terms of nominal, ordinal, interval and ratio scales of measurements. In other words, quantitative data must be identified or classifiable in terms of types and levels of measurement
- quantitative data are usually and easily presented in the form of tables and illustrated by the means of graphs and charts. In essence, graphical and tabular presentations of data are possible and in fact encouraged in quantitative analysis of data
- descriptive statistics are usually employed in the description of quantitative data while inferential statistics are utilised to achieve explanation, prediction and some level of generalisation. Quantitative data could easily be applied to statistical measures and tools.

3.2 Understanding Independent and Dependent Variables

A variable is any entity that can take on different values. In essence, anything that can vary can be considered a variable. For instance, the age of students in your class, weight of students, scores of students in data analysis course, income of workers etc are all considered as variables because the values can change from one individual to the other. Variables could be classified using different platforms. However, the classification according to relationship to one another is the most popular and often used in mass media research. In this regard, variables could be classified into two thus:

- Independent Variables
- Dependent Variables

According to Kerlinger (1973:35), cited in Wilson, Esiri and Onwubere (2008), the most important and useful way to classify variables is to see them as either independent or dependent variables. This classification, according to him, is highly useful because of its general applicability, simplicity and special importance in conceptualising and designing research and in communicating the results of the research.

Kerlinger (1973:35) defines an independent variable (IV) as the presumed cause of the dependent variable (DV), the presumed effect. The independent variable is the antecedent, while the dependent variable is the consequent (Wilson, Esiri & Onwubere, 2008).

According to Wilson, Esiri and Onwubere (2008), independent variable could be distinguished thus:

- The independent variable is observed and its value presumed to depend on the effects of the dependent variable. In other words, the dependent variable is what the researcher wishes to explain. The independent variable may be manipulated or it may just be measured. In contrast, the dependent variable is what we are studying, with respect to how it is related to or influenced by the independent variable or how it can be explained or predicted by the independent variable. It is sometimes called the response variable or the criterion variable. It is never manipulated as a part of the study. DVs are the things we measure about people.

Wilson, Esiri and Onwubere (2008) went ahead to further explain the concept of independent variable and dependent variable using the following example:

- Suppose two investigators are studying the relationship between criminal behaviour in adolescents and parental guidance to determine what kinds of advice to give parents. The two investigators may have the same data. This data includes: (1) the police records of a group of adolescents, giving data about the number of times the child has entered the criminal justice system (such as by being arrested, questioned by the police, etc.), and (2) information from a questionnaire about the kinds of information or advice that each adolescent has received from his or her parents. One investigator might be examining whether parents who give advice focusing on walking away from interpersonal conflicts differ from parents who give advice to the child to "stand up for yourself". The independent variable is the kind of

advice the parents give and the dependent variable is whether the child has criminal record or not.

Continuing, Wilson, Esiri and Onwubere (2008) further note that a second investigator might be asking a different question like "what types of parental advice and guidance distinguishes adolescents who get into the criminal system from those that don't? In this case, whether or not the child has a criminal record or not is the independent variable and the type of parental advice is the dependent variable. From this example, it should be clear that the distinction between the independent and dependent variable is based not on manipulation but on the questions one is asking of the data".

4.0 CONCLUSION

Variables are very important in any research activity, mass media research inclusive. Researchers, investigators and scientists use variables to investigate cause and effect relationships in nature or other subjects or phenomena. In other words, variables help researchers, investigators and scientists to design their research in such a way that changes to one item could cause something else to vary in a predictable way.

5.0 SUMMARY

This unit focused on elements of quantitative data analysis and how to understand and explain independent and dependent variables. It is our expectation that you can now comfortably distinguish between independent and dependent variables as well as identify and discuss the elements of quantitative analysis of data.

6.0 TUTOR-MARKED ASSIGNMENT

i. Distinguish between independent and dependent variables citing copious examples.
ii. Identify and discuss four vital elements of quantitative analysis of data.

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics*. New York: Longman.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Otokiti, *et al*. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Wilson, D.et al.(2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 2 PROCEDURE FOR QUANTITATIVE ANALYSIS

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Procedure for Quantitative Data Analysis
        3.1.1 Preliminary Analysis of Data
        3.1.2 Summarising Sample Characteristics
        3.1.3 Thematic Analysis
        3.1.4 Interpretation of Result and Inferences
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

In this unit, we continue our discussion on quantitative analysis of data with particular attention this time to the procedures and processes. The procedure for quantitative data analysis comprises four main steps each of which maybe sub-divided into a number of sub steps thus:

a. preliminary analysis of data
b. summarising sample characteristics
c. thematic analysis
d. interpretation of result and inferences.

These four constitutes the core processes involved in analysing quantitative based data and will therefore constitute the core of our discussion in this unit.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- describe the preliminary analysis of data
- summarise sample characteristics
- explain thematic analysis
- interpret results and inferences.

3.0 MAIN CONTENT

3.1 Procedure for Quantitative Data Analysis

3.1.1 Preliminary Analysis of Data

This provides a rough overview of the pattern of response for all the variables in the study. The preliminary analysis of data in quantitative data analysis does not aim at answering any of the substantive research questions (Obikeze, 1990). It usually includes the calculation of response rate, a check for response bias, and the compilation for simple frequency distribution for all the variables. The response rate measures the proportion of the total respondents who successfully filled their questionnaires and returned or were available for interview and who were actually interviewed (Obikeze, 1990).

For the Response rate = X/Y x 100/1,

Where:

X= those who properly filled and returned copies of the questionnaire and

Y= sample size or the total number of respondents for the study.

The check for Response bias is to determine whether those who responded appropriately to the interview differed in significant ways from those who failed to in respect of the crucial variables of the study. This, however, will not be easy to calculate because of the non-availability of information (Obikeze, 1990).

The third preliminary analysis for quantitative based studies is the compilation of simply frequency distributions of summaries for all the variables. It provides a quick check of the response pattern which helps the researcher to assess the relative importance of a variable for explaining relationships and trends (Obikeze, 1990).

SELF-ASSESSMENT EXECISE

i. Discuss the difference(s) between Response Rate and Response Bias.
ii. With the aid of a typical example, how would you calculate the response rate?

3.1.2 Summarising Sample Characteristics

This stage of analysis highlights the main structural features of the study population which a likely to impinge on the behaviors of, and relationships among its members thus providing necessary background for proper understanding and explanation of social realities. Covered in this analysis are the so called background variables like age, sex, marital status, occupation and income level, etc. Descriptive statistics in the form of simple frequency distribution etc are usually used here (Obikeze, 1990).

Obikeze (1990:90) further noted that this is the first procedure in the actual analysis of data. According to him, data is summarised and briefly described so that the main characteristics of the study sample with respect to the major analytical variables could easily be highlighted.

3.1.3 Thematic Analysis

Obikeze (1990:92) identifies this stage of the quantitative data process as the crux of the data analysis procedure because at this stage the researcher or investigator "attempts to find answers to the various research issues and questions, as well as test specific hypothesis. The outcome of this particular process determines the extent to which the study objectives have been met."

There is no generally accepted role or format or standardised forms for ordering issues for data analysis and presentation, however, Obikeze (1990:92-93) developed seven procedures for this which are adopted here for proper guidance and future improvement.

a. Make a list of all relevant research issues, research questions, and research problems to be considered
b. Sequential ordering of the issues and problems
c. Constructions of frequency distribution for each topic, or issue in turn
d. Description of the tables using appropriate statistics and graphic presentation
e. Identifying trends, variations, relationships or any peculiar features of the data using appropriate statistics
f. Testing any explicit hypothesis on the issues using appropriate statistics
g. Making the necessary decisions, generalisations and interpretations of result using theory, other studies and specifying the implications (practical and theoretical) of the research findings.

3.1.4 Interpretation of Result and Inferences

Interpretation of results and drawing inferences happen after the data have been analysed using descriptive statistics. If the research has followed the scientific processes diligently, it will be easy to generalise at this level or draw inferences from the analysed data. Below is an interpretation by the present author in a key variable of a recent study:

However, a very serious issue raised in the result analyses is the fact that most professionals are not adequately trained in the application of ICT and those that are trained are not given regular opportunities to update their knowledge. ICT and globalisation are on-going issues. They change very fast that what was regarded as new media technologies yesterday are now regarded as old media technologies. This dynamism has to be considered in organisations that are highly enmeshed in globalisation issues and ICT application. Professionals need to be adequately trained in the use of these technologies and retrained so that they can positively use the technologies for the benefit of their indigenous societies. Lack of depth in knowledge and application of these technologies could result in misuse, underutilisation and misapplication. These are the real issues that could lead to cultural erosion or imperialism.

4.0 CONCLUSION

The procedure for quantitative analysis of data is a commonplace method because of the popularity of the quantitative technique. It is important therefore that every communication student takes time to master the process so as to efficiently apply in related situations.

5.0 SUMMARY

This unit continued our discussion on quantitative data analysis with particular attention on the procedure for analysing data in quantitative based studies or researches. The procedure, as detailed in this unit, is composed of four main steps thus:

- preliminary analysis of data
- summarising of sample characteristics
- thematic analysis of the data to bring out the relationships in the variables
- interpretation of result and inferences to make meaning to the reason or rationale behind study.

6.0 TUTOR-MARKED ASSIGNMENT

i. Do an elaborate discussion on the procedure for quantitative data analysis.
ii. Discuss the significance of thematic analysis in quantitative based studies.

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics*. New York: Longman.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Otokiti, S. et al. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 3 CONSTRUCTION OF FREQUENCY DISTRIBUTION TABLES

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Frequency Distribution Tables (FDT)
    3.2 How to Construct Frequency Distribution Tables
    3.3 Description of Frequency Distribution Tables
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

Statistics are the basic tools of data analysis in quantitative based studies and analysis. For each issue or questions raised in a quantitative study, one of the initial things to do is to construct a frequency distribution table so that statistical issues connected with the analysis could be dealt with. This therefore is a continuation of the ongoing discourse on quantitative data analysis.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- construct and explain frequency distribution tables and cross tabulation
- explain how to construct tables
- describe a typical frequency distribution table.

3.0 MAIN CONTENT

3.1 Frequency Distribution Tables (FDT)

There are basically three types of frequency distribution tables namely:

a. univariate frequency distribution table
b. bivariate frequency distribution table
c. multivariate frequency distribution table

a. **Univariate frequency distribution** table means that only one variable or questionnaire item is being considered. **For example:**

Table 3.1: Gender Distribution of Sociology Students of NOUN

Gender	**Frequency**	%
Male	150	45.45
Female	180	54.55
Total	330	100.0

b. **Bivariate frequency distribution** means that two variables are being considered together. More specifically, it is the cross-tabulation of responses to two questionnaire items or two variables simultaneously.

Table 3.2: Level of Education by Religious Affiliation of Respondents in Trans-Ekulu

	Level Of Education					
Religious Affiliation	**University**	**College**	**Secondary**	**Primary**	**No School**	**Total**
Christian	7	10	13	2	1	**33**
Moslem	1	2	4	3	2	**12**
Others	3	2	3	5	2	**15**
Total	**11**	**14**	**20**	**10**	**5**	**60**

Table 3. 2 shows that out of the 33 Christians in Trans-Ekulu who participated in the survey, 7 have university education, 10 are college graduates, and 13 have secondary education. Only two of the respondents stopped at the primary school level while one of the respondents did not attend any formal school at all.

c. **Multivariate frequency distribution** table helps to describe, explain or understand relationships among three or more variables (compromising of one dependent and two or more independent and intervening variables).

Table 3. 3: Level of Education and Sex by Religious Affiliation of Respondents in Trans-Ekulu

Level of education and sex								
	High		**Medium**		**Low**		**Total**	
Religion Affiliation	M	F	M	F	M	F	**M**	**F**
Christian	10	7	5	8	1	2	**16**	**17**
Moslem	3	0	3	1	2	3	**8**	**4**
Others	3	2	2	1	2	5	**7**	**8**
Total	**16**	**19**	**10**	**10**	**5**	**10**	31	29

SELF-ASSESSMENT EXERCISE

i. Distinguish between the different types of frequency distributions.
ii. Discuss two major instances where cross tabulation must be used.

3.2 Construction of Frequency Distribution Tables

To construct a frequency distribution table, the researcher must follow well-defined procedures as highlighted below:

- a table must be **numbered**: this number is used to make reference to the table each time it is used to explain any aspect of the research or result
- a table must also have a **title** which is brief, but self explanatory.
- a table should also have a row(s) with a row heading called a **STUB** and column heading called the **CAPTION**
- a table should also show the **units and the signs** of recognising the items in a table e.g. the income level of Dollars or Naira
- a table should also have a **footnote** when items on the table require further explanation
- a table should also have a **source** from which the data displayed are reported from. This is always placed below the table as the last footnote.

3.3 Description of Frequency Distribution Tables

Tables ordinarily do not speak for themselves. They need to be described to bring out the features of the data presented. Tables are best described using descriptive statistics and graphic presentations.

Descriptive statistics could be used to explain frequency distribution tables in the form of relative measures, measures of central tendency and measure of dispersions. Please note that the principles and theories behind these statistics will not be discussed here. The assumption is that we have already treated this in earlier courses in statistics for social science. We shall therefore be more concerned with their application (Obikeze, 1990; Wimmer & Dominick, 2011).

Relative Measures

Relative measure also called relative statistics is used to measure the relative frequency of the different items of value of a variable presented. With this, the researcher could easily make comparisons within variables or categories. Relative measure that is commonly applied in describing frequency distribution table is proportion, percentage, ratio, and rates. For us in media studies, percentages are the most commonly used relative measure (Obikeze, 1990; Wimmer & Dominick, 2011).

Table 3. 4: Percentages Sample of Relative Measure

Sex of Respondent	Frequency	%	Proportion
Male	90	45	0.45
Female	100	55	0.55
Total	200	100%	1.00

Ratio on the other hand is used to compare two independent values by expressing one as a quotient of the other. **Rates**, on its part, measure the average incidence or occurrence of an event over a total population that is exposed to the risk of that event. It may also be considered as the average unit of one variable per unit value of another variable. How to calculate these relative measures have already been discussed in details in the preceding module.

Identification of Associations and Relationship in Variables

After the data have been described using descriptive statistics as earlier discussed, associations and relationships within the variables need to be identified so as to achieve the objective of settling out for study. Obikeze (1990) did a comprehensive work in this area and came out with the following steps:

- construction of contingency tables or cross tabulation tables to present the data
- analysis of the tables using percentage or other descriptive measures to identify trends and relationships
- measurement of the degree of identified relationship using statistics appropriate for the different types of data

- testing any formal hypothesis relationships or where no hypothesis were formulated assessing the statistical significance of identified relationships
- making necessary decisions, inferences, and interpretation of the test result
- explaining the research result in the light of the theory and practical experience.

Please, note that these steps are all very vital in identifying associations and relationships in variables. However because of the relative importance of step 3, it shall be discussed further here:

STEP 3

Appropriate test statistics are usually applied to test the relationships between variables in a study. Many of such test statistics exist and they are data specific (Obikeze, 1990: Wimmer & Dominick, 2011). But please note that it is not likely that you may use most of these tests at this stage. The vital ones that are commonly used in media research have discussed in details with worked examples.

- **For nominal data** (i.e. tests of significance of difference between proportions) we have: the t-test,x^2 z-TEST, Mc Nemar test and the Chochran Q-test.
- **For ordinal data** (Non-par metric test) we have: the kolmogorov-smirnov-test, the Man-Whitney test, the sign test, and the Wilcoxon Mat-Paris Test.
- **For interval data** (i.e. differences between means test) we have: the Z-test and the T-test.

The outcomes of these tests are usually measured to assess the strength of such relationship. This is done by using appropriate statistical measures to derive numerical indices called ***coefficient of corelation*** or coefficient of association. This gives the numerical expression of degrees of relationship between variable and range from 0.00-1.00 (note that the SPSS could help you with this value with any of the test). A coefficient of 0.00 implies no significant relationship, while a coefficient of 1.00 implies the highest or perfect relationship, between these two extremes the relationship may be described as very low, high or very high depending on the magnitude of the coefficient.

Furthermore, a negative sign (-) on the value or indices indicates negative or inverse relationship while the positive (+) indicates a positive or direct relationship. For coefficient of correlation, test statistics also exist for the different kinds of data.

- **For Nominal Data**: the PHI (ϕ) applicable to 2x2 (2 by 2) contingency tables only, the lambda (λ) and the chi square (x^2)
- **For Ordinal Data**: the spearman's Rho (p), the Kendall's Tau (T)
- **For Interval Data**: the Pearson's product-moment correlation

Note that the SPSS as earlier discussed could be used to compute all of these tests.

4.0 CONCLUSION

The understanding of how to use and when to use the various types of statistical tables discussed is very important to the overall outcome of data analysis in communication research. It is important therefore that you make every effort to familiarise yourself with what truly a statistical table stands for, know what each part does and when to use the different types for your different data types.

5.0 SUMMARY

In this unit, we discussed how to construct statistical tables for data analysis purposes. In summary, the unit focused on how to construct frequency distribution tables, the various types of frequency distribution tables as well as how to describe such tables in order to make the meaning much more meaningful to those who may use the information therein.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss in details how to identify associations and relationships in variables.
ii. Discuss some of the statistical tools you could use to describe a table so that it serves a better purpose to the investigator.

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Otokiti, S. et al. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 4 ANALYSING DATA FROM OBSERVATION TECHNIQUE

CONTENTS

1.0 Introduction
2.0 Objective
3.0 Main Content
    3.1 Overview of the Observation Research Technique
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

Since the beginning of this module, we have been discussing qualitative studies and the various methods of doing it. This unit is an extension of such discourse as we again consider the method of observation which is considered a useful way of collecting data from the field. Like all qualitative technique, observation is concerned more with description and explanation than measurement and quantification.

2.0 OBJECTIVE

At the end of this unit, you should be able to:

- explain the observation research technique

3.0 MAIN CONTENT

3.1 Overview of the Observation Research Technique

There are various aspects of human behaviour that cannot be comprehended without an in-depth understudy of such situation. Such understanding could come in the form of observation of the subject with the intent of understanding the phenomenon under question. In other words, observation is one of the methodological tools that are applied in communication research to elicit evidence and subsequently explain certain behaviour traits associated with the subject of the observation.

4.0 CONCLUSION

Field observation method simply requires an observer-human or electronic, to observe and record the required items. Observation is a

qualitative research technique, which is more concerned with description and explanation rather than measurement and quantification. This is what makes this research approach unique.

5.0 SUMMARY

This unit dealt with the observation research technique as one of the popular techniques used in qualitative studies. It is expected that you now the various types field observation, advantages and disadvantages of field observation, how to analyse the data from field observation and the methods of interpretations.

6.0 TUTOR-MARKED ASSIGNMENT

i. Distinguish between participant observation and non participant observation.
ii. Justify the use of field observation in communication research.
iii. How would you analyse data obtained from field observation?

7.0 REFERENCE/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

MODULE 5 QUANTITATIVE DATA ANALYSIS: FOCUS ON CONTENT ANALYSIS AND EXPERIMENTS

Unit 1 Introductory Overview of Content Analysis
Unit 2 Processing, Analysing and Writing the Content Analysis Data and Report
Unit 3 Introductory Overview of Experimental Research
Unit 4 Processing and Analysing Experimental Data

UNIT 1 INTRODUCTORY OVERVIEW OF CONTENT ANALYSIS

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
  3.1 Definition of Content Analysis
  3.2 Uses of Content Analysis
  3.3 Limitations in the Use of Content Analysis
  3.4 Steps in Content Analysis
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

This module and unit introduce us to quantitative research techniques beginning with Content Analysis which is one of the familiar methods of doing research in communication sciences. Content Analysis focuses on the message and this makes it a familiar method in the mass media research.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define Content Analysis
- explain the Uses of Content Analysis
- describe the Limitations in the Use of Content Analysis
- explain the Steps in Content Analysis.

3.0 MAIN CONTENT

3.1 Definition of Content Analysis

This is one of the most common methods of data analysis in communication studies. Walizer and Wiener (1978) define it as any systematic procedure devised to examine the content of recorded information. Kerlinger (2002) cited in Wimmer and Dominick (2011) defines it as a method of studying and analysing communication in a systematic, objective, and quantitative manner for the purpose of measuring variables.

Berelson (1952:18) on its part views Content Analysis as a research technique for the objective, systematic and quantitative description of the manifest content of communication. Wright (1986:125) cited in Wilson, Esiri and Onwubere (2008) on its own part defines content analysis as a research technique for the systematic classification and description of communication content according to certain usually predetermined categories. This is not different from Budd et al (1967:30) definition cited in Wilson, Esiri and Onwubere (2008) which also sees Content Analysis as a systematic technique for analysing message content and message behaviour of selected communicators.

Krippendorf (1980:51) is also of the view that Content Analysis is a technique for making replicable and valid inferences from data to their context. For example, you can decide to study the content of two Nigerian newspapers to determine the portrayal of women politicians in partisan political participation.

From these definitions three issues about content analysis play up. The three issues are the core features of Content Analysis:

- it is systematic
- it is objective
- it is quantitative

3.2 Uses of Content Analysis

Content Analysis has been used in media research to achieve a lot of research purposes. According to Wimmer and Dominick (2011:157-158), Content Analysis is usually used for the following purposes:

- **To describe communication content**: Content Analysis is used often to describe communication content of various types of media.

- **To test hypothesis of message characteristics**: Many analyses attempt to relate certain characteristics of the source of a given body of message content to the characteristics of the messages that are produced.
- **To compare media content to the real world**: In this sense, many Content Analyses are reality checks in which the portrayal of a certain group, trend, trait or feature is assessed against a standard taken from real life.
- **Assessing the image of particular group in society**: Most Content Analysis focus on the portrayal of particular groups in society. In a related sense, Content Analysis attempts to determine how particular groups in a society are represented.
- **Establishing a starting point from studies of media effects**: Content Analysis could lead to new insights into new research areas.

3.3 Limitations in the Use of Content Analysis

Content Analysis is a common research method in mass communication research. Notwithstanding its popularity in mass media studies, it has many limitations in its application. Some of these limitations are summarised below as detailed by Wimmer and Dominick (2011:159-160) thus:

- content analysis alone cannot serve as a basis for making statement about the effect of content on the audience
- the findings of a particular content analysis are limited to the framework of the categories and definition used in that analysis
- a content study may not produce adequate messages relevant to the research and so produce thin data that may not be adequate for any valid analysis to take place
- content analysis is very cumbersome and time consuming. Sometimes a researcher may have to read through a mass of content to determine category and placement. This is not an easy work considering the fact that many editions or issues of the particular medium may be up for analysis.
- content analysis is a very expensive way to analyse media content for research purposes.

3.4 Steps in Content Analysis

Wimmer and Dominick (2011:160) note that Content Analysis is conducted in several discrete stages as presented below. (Note that the steps as presented below are in sequential order but this is not cast in iron. A lot of flexibility could be exercised at this level to achieve

maximum utility in the research agenda.) The steps in Content Analysis are outlined below:

- formulate the research question or hypothesis
- define the population in question
- select an appropriate sample from the population
- select and define a unit of analysis
- construct the categories of content to be analysed
- establish a quantification system
- train coders and conduct a pilot study
- code the content according to established definitions
- analyse the collected data
- draw conclusions and search for indications.

SELF-ASSESSMENT EXERCISE

Analyse three different definitions of Content Analysis and showcase the common features.

4.0 CONCLUSION

Content Analysis is a specific research approach that is frequently used in mass media research. Over the years, researchers have come to the conclusion that it is viable, reliable and efficient way of determining how media see and report the real world.

5.0 SUMMARY

This unit discussed some valid definitions of content analysis, uses, steps and some of the limitations associated with it. However, from these definitions three issues about content analysis play up. These issues are regarded as the core features of content analysis and they include:

- it is systematic
- it is objective
- it is quantitative

Wimmer and Dominick (2011:160) also note that content analysis is conducted in several discrete stages with a lot of flexibility. This flexibility is usually determined by the research objectives.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss the steps involved in Content Analysis
ii. Some limitations are associated with the use of Content Analysis in communication research. Please, identify and discuss four.

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Otokiti, *et al*. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Wilson, D.et al.(2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 2 PROCESSING, ANALYSING AND WRITING THE CONTENT ANALYSIS DATA AND REPORT

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
  3.1 Reliability in Content Analysis
  3.2 Validity in Content Analysis
  3.3 Calculating Inter-Coder Reliability
  3.4 Improving Inter-Coder Reliability
  3.5 Analysing the Content Analysis Data
  3.6 Writing the Content Analysis Data and Report
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

In this unit, we continue our discussion on Content Analysis which we have identified as one of the frequently used research methods in mass media/communication research. In this unit, we will continue the overview of this important research method in communication studies and pay more particular attention to how content analysis is analysed and presented for better comprehension of the analysed report.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- explain Reliability in Content Analysis
- discuss Validity in Content Analysis
- calculate Inter-Coder Reliability
- explain how to Improve Inter-Coder Reliability
- analyse and Write the Content Analysis Data and Report.

3.0 MAIN CONTENT

3.1 Reliability in Content Analysis

Reliability is very crucial in content analysis. Reliability is about the repeatability or consistency of a measure. Research experts always

advise that as much as possible coding should have some degree of reliability which is necessary to have accurate measures. For you as a communication undergraduate, who may require some research assistants to code any content analysis research effort, the one most needed is intercoder reliability which refers to levels of agreement among independent coders who code the same content using the same coding instrument.

Wimmer and Dominick (2011:170) confirm that the concept of reliability is crucial to content analysis as betrayed in the following statement:

- If a content analysis is to be objective, its measure and procedures must be reliable. A study is reliable when repeated measurement of the same material results in similar decisions or conclusions. Intercoder reliability refers to levels of agreement among independent coders who code the same content using the same coding instrument.

Wimmer and Dominick (2011:170) therefore note that if the results of a content analysis fail to achieve reliability, any of the following could be responsible:

- something may be wrong with the coders and so they need to be trained properly
- something may be wrong with the coding instruction and so need to be reevaluated.
- something may be wrong with the category definitions and so need to be clarified
- something may be wrong with the unit of analysis and/or coding categories and so may need a redefinition
- a combination of any of the above may have occurred.

To achieve acceptable levels of reliability, Wimmer and Dominick (2011:171) argue that the following steps come highly recommended:

- define the category boundaries with maximum details and ensure that their exclusivity is maintained
- train the coders on the subject matter, methodology, coding categories and specific instructions to follow
- conduct a pilot study to establish the workability and validity of the research design and ensure that errors are detected early in the research process.

3.2 Validity in Content Analysis

This also raises special concerns in content analysis because it is intimately connected with the procedures used in the analysis. According to Wimmer and Dominick (2011:175-6), validity is usually defined as the degree to which an instrument actually measures what it sets out to measure.

Issues that Affect Validity Negatively

- Faulty Sampling Design
- Overlapping Categories
- Low Reliability Status
- Inadequate Definitions of Categories

Types of Validity

a. **Face Validity**: this validity assumes that an instrument adequately measures what it purport to measure if the categories are rigid and satisfactorily defined and if the procedure of the analysis have been adequately conducted (Wimmer & Dominick, 2011:176).

b. **Concurrent Validity:** this shows that the actual measure is concurrent with the stated intent on guides (e.g. using TV guide and actual viewing of the program to establish validity) (Wimmer & Dominick, 2011:176). Concurrent validity, according to Wilson, Esiri and Onwubere (2008) "is closely related to predictive validity. In this method, however, the measuring instrument is checked against some present criterion. For example, it is possible to validate a test of proofreading ability by administering the test to a group of professional proofreaders and a group of non-professional proofreaders. If the test discriminates well between the two groups, it can be said to have concurrent validity."

c. **Construct Validity:** this involves the use of semantic differentials and factor analysis in an attempt to isolate its underlying dimensions in the content being analysed. This type of validity is related to message characteristics. In essence, the characteristics of a message are used to determine this type of validity. Construct validity has been used to test sensationalism in news stories (Wimmer & Dominick, 2011:176). It involves relating a measuring instrument to some overall theoretic framework to ensure that the measurement is logically related to other concepts in the framework.

d. **Predictive Validity**: this is the use of some content attribute to predict which items a newspaper will publish (Wimmer &

Dominick, 2011:176). Checking a measurement instrument against some future outcome assesses predictive validity**.** For example, scores on a test to predict whether a person will vote in an upcoming election can be checked against actual voting behavior (Wilson, Esiri and Onwubere, 2008).

SELF-ASSESSMENT EXERCISE

i. How does Face Validity practically work in media research amongst undergraduate students?
ii. Describe three other types of validity you are familiar with in dealing with content analysed study.

3.3 Calculating Inter-Coder Reliability

According to Poindexter and McCombs (2000), inter-coder reliability measures the consistency of coders in coding the content. There are many ways of calculating inter-coder reliability. The widely used formula computes a coefficient of reliability by stating the ratio of decisions that coders agreed on to the total number of decisions made by each coder (Holsti, 1969).

$$\text{C.R.} = \frac{2M}{N_1 + N_2}$$

Where:

CR = Coefficient of Reliability
M = Number of Coding decisions agreed on
N = Total Number of Coding Decisions made by Each Coder

Note
Coefficient of reliability is reported as a percentage. To determine the M, or the number of decisions that coders agree on, the pretest coding must be compared code by code. Any discrepancies that are found should be flagged for follow-up discussion. The number of coding decisions that match will be used in the formula. Although M represents the number of decisions that the coders agreed on, the formula doubles this number to represent the agreed upon decisions of two coders. In the denominator, the total number of decisions, N, made by each coder is summed (Poindexter & McCombs, 2000).

For example, two coders matched on 180 coding decisions. Each coder made 240 coding decisions. The inter-coder reliability is

$$\text{C.R.} = \frac{2(180)}{240+240} = \frac{360}{480} = 75\%$$

A rule of thumb for an acceptable coefficient of reliability is 80% or above. If the coefficient of reliability is less than 80%, the researcher should find means to increase it before proceeding with the content analysis coding proper (Poindexter & McCombs, 2000).

3.4 Improving Inter-Coder Reliability

According to Poindexter and McCombs (2000), the underlisted points have been tested as valid ways to increase intercoder reliability:

- regardless of the percentages to coefficient of reliability, the researcher should discuss the discrepancies in coding with the coders
- the researcher should also identify the coder who is consistent in making coding errors and replace him/her
- if errors consistently appear because of the coding categories or descriptions that are not mutually exclusive, the researcher must revise the codebook
- once the codebook has been revised, the researcher will then ask the coders to code another sample of content. After this, the coefficient of reliability will be calculated again. Hopefully, this should increase the coefficient of reliability to at least 80%.

3.5 Analysing the Content Analysis Data

According to Poindexter and McCombs (2000), since content analysis is quantitative in nature, it takes after survey data analysis in many respects. Some are itemised here:

- content analysis data are analysed the same way the survey data are analysed
- however, before focusing on the specific research question, the research expert would run a frequency print out to produce frequency distributions of all variables. The researchers must verify that the variables are being read in the correct columns and then check for and clean any dirty data or out-of-range codes
- once the researcher confirms that the data is clean, focus can shift to answering the research questions
- descriptive statistics such as percentages, means, modes, and medians, are also used to analyse content analysis. However, as noted by Wimmer and Dominick (2011), if hypothesis tests are

planned, then common inferential statistics are acceptable. Usually, the Chi-square test is the most commonly used in content analysis because content analysis data tend to be nominal in form. If the content analysis data meet the requirements of interval or ratio levels, then a t-test, or ANOVA, or Pearson r may be used.

3.6 Writing the Content Analysis Data and Report

a. The report writing stage of content analysis follows the same process as survey
b. Check the draft for organisation and transitions as well as correct grammar, sentence structure, and punctuation
c. The tone should be professional and objective
d. Persuasion is permitted in the recommendation section only
e. Proof all numbers in the text and tables to make sure they match the original printouts
f. Ethics must be observed. Data should not be omitted or buried because they do not conform to expectations.

4.0 CONCLUSION

Content analysis is a popular and frequently used research methodology in mass communication. It also follows that some of the steps in a typical survey are applicable in content analysis. Since this is a popular research methodology, the expert researcher should make sure his or her coders are adequately trained and mentally focused to fulfill that mandate. In fact, the entire research design must be thoroughly evaluated to achieve a valid content analysis.

5.0 SUMMARY

This unit focused on the concluding part of the content analysis discourse. We considered reliability and validity in content analysis as well as how the data gotten from such a research method are analysed and reported.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss reliability and validity in content analysis.
ii. Two different coders coded 250 items and agreed on 210. Calculate the coefficient of reliability.
iii. Discuss how to practically analyse data gotten from content analysis.

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Otokiti, S. *et al*. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 3 INTRODUCTORY OVERVIEW OF EXPERIMENTAL RESEARCH

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Experimental Research
    3.2 Experimental Concepts, Designs and Symbols
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

This unit discusses the oldest research approach in mass media research-experimental research. Experimental research is one of the earliest research methods ever used to answer certain questions in mass media research and practice. It has continued to provide a wealth of information for researchers and critics of the media notwithstanding its dwindling use most mass communication research efforts in recent years.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- give an overview of experimental research
- discuss some experimental concepts, designs and symbols.

3.0 MAIN CONTENT

3.1 Experimental Research

An experiment**,** a methodology that cuts across the physical, biological, social and communication sciences, is a research method in which an independent variable is manipulated and its effects on the dependent variable are observed. When an experiment is conducted scientifically, the researcher is able to attribute any change in the dependent variable directly to the independent variable and not to extraneous variables or factors unrelated to the study. In fact, the controlled experiment has been described as the most powerful method available for finding out what causes what (Westley, 1981:196).

According to Wilson, Esiri and Onwubere (2008),

- Whenever the word, "experiment" is mentioned, what immediately comes to mind is the work of physical or natural scientists and psychologists among social scientists? We tend to think of experiments in terms of test tubes, thermometers and chemicals in a laboratory. We also tend to equate experimentation with the natural sciences. No doubt the reason for this stems from the fact that the initial steps in setting up experimentation as the ultimate in research were taken in the physical or natural sciences. Also, experiments are used much more often in the physical sciences and psychology than in mass media research.

According to Berger (2000:201), as cited in Wilson, Esiri and Onwubere (2008), an experiment is a procedure or kind of tests that:

- demonstrates that something is true
- examines the validity of an hypothesis or theory or
- attempts to discover new information.

Continuing, Wilson, Esiri and Onwubere (2008), still citing Berger (2000) explain the three procedures thus:

- In the first case, the researcher tries to show that what is held to be true about something is actually true. For instance, it is possible to transmit sound (speech, music, etc) from one point to another via radio receivers. This might involve replicating an experiment to see whether the findings in the first experiment are true. In the second instance, the researcher tests a hypothesis or theory to determine whether it is valid. For instance, the researcher might postulate that there is a relationship between heavy television viewing and violent behaviour in young people. In the third case, the researcher will want to discover something we did not already know. For instance, the researcher might try to find out whether one sided messages are more effective with people of lower education than those with higher education.

Another classic definition of experimental research is offered by Severin and Tankard (2001:37) and reported in Wilson, Esiri and Onwubere (2008) thus:

- Experimental research is the classic method of dealing with the question of causality (i.e. the relationship between something that happened and the reason for it happening i.e. cause and effect). It involves the control or manipulation of a variable (condition) by

the experimenter and an observation or measurement of the result in an objective and systematic way. A scientific experiment will answer questions of whether and to what degree a variable (the experimental or independent variable) affects another variable (the dependent variable).

Experimental, especially the laboratory research has evidence of causality, control, cost, and replication as advantages while issues of artificiality, researcher bias and limited scope or depth of the study remain disadvantages this method of research grapples with.

SELF-ASSESSMENT EXERCISE

i. Cite and explain two definitions of experimental research.
ii. Discuss the disadvantages associated with experimental research
iii. 3.2 Experimental Concepts, Designs and Symbols

Experimental Validity
The researcher designs the experiment to ensure internal validity, which answers the question: did the independent variable cause change in the dependent variable? And external validity, which answers the question: can the results be generalised to a larger population beyond the actual participants in the experiments (Poindexter & McCombs, 2000)?

Experimental Designs
The design of an experiment specifies the framework for how the study will be conducted. Experimental research scholars Donald Campbell and Julian Stanley (1963:1) identified 16 such frameworks that have been used for conducting experiments (Poindexter & McCombs, 2000). Four basic ones are presented below:

Four Experiment Design

Description	Random Assignment	Pretest	Condition	Post Test	Interval Validity
One-group pretest-posttest design		O	X	O	Weak
Pretest –Posttest Control Group Design	R R	O O	X	O O	Strong
Solomon Four Group Design	R R R R	O O	 X X	O O O O	Strong
Posttest –Only ControlGroup Design	R R		X	O O	Strong

Symbols: **X = Independent Variable**
O = Dependent Variable
R = Random Assignment

Experiment Symbols
To interpret an experimental design, it is first necessary to understand the symbols X, O, and R that are displayed below:

X = Independent Variable
O = Dependent Variable
R = Random Assignment

By knowing how to read this experiment shorthand, you can interpret the type of design used to conduct an experiment. Apart from knowing what these symbols mean, we must also understand and interpret their placement. For example, when O is placed before X, it means that the dependent variable has been measured before manipulation of the independent variable. In other words, there has been a pretest. When O is placed after the X, the dependent variable has been measured after the manipulation of the independent variable. In other words, there has been a posttest (Poindexter & McCombs, 2000).

The placement of the symbols also reveals whether or not a control group was used. Wherever an O in the posttest condition is not preceded by an X, it means that condition represents the control group (Poindexter & McCombs, 2000).

3.2 Threats to an Experiment's Internal Validity

According to Poindexter and McCombs (2000:217), the following are threats that could affect an experiment's internal validity issues:

- **History:** any event that occurs between the pretest and posttest-measurements
- **Maturation:** any changes that occur in the subjects between the pretest and posttest. Changes might include becoming older, more knowledgeable, and more confident, etc
- **Testing:** any effects of taking the pretest on the scores of the posttest
- **Instrumentation:** any effects resulting from changing the measuring instruments from the pretest to the posttest
- **Statistical Regression:** when subjects are selected on the basis of extreme scores, the posttest scores tend to regress towards the mean of the distribution of scores

- **Differential Selection of Subjects:** when subjects in the experimental group differ significantly from subjects in the control group
- **Experimental Mortality:** when subject drop-out from pretest to posttest in the experimental and control groups is not equal.

Threats to External Validity

1. Pretesting
2. Subject Selection
3. Experimental Setting

1. Pretesting

The experimental design that includes a pretest cannot be generalised beyond the experimental group because the larger population to which the subjects belong has not experienced the effects of the pretest (Poindexter & McCombs, 2000).

2. Subject Selection

This could threaten external validity because experiments are often conducted among subjects, such as college students, who are not representative of the population as a whole (Poindexter & McCombs, 2000).

3. Experimental Setting

This relates to the artificiality of the experimental setting. Many experiments are conducted in a classroom that serves as a laboratory rather than the real world or setting (Poindexter & McCombs, 2000).

4.0 CONCLUSION

The experimental research protocol has remained a popular research approach for media researchers since the inception of mass communication research many decades ago. Though criticised by many researchers as being artificial, the method offers a number of advantages that make it particularly useful to some researchers (Wimmer & Dominick, 2011:263).

5.0 SUMMARY

In this unit, you learnt the meaning of experimental research citing different definitions to support your knowledge. We also discussed some experimental concepts, designs and symbols.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss why an understanding of experimental symbols is important in conducting an experiment.
ii. Discuss five advantages that are associated with laboratory experiments.

7.0 REFERENCES/FURTHER READING

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth

UNIT 4 PROCESSING AND ANALYSING EXPERIMENTAL DATA

CONTENTS

1.0 Introduction
2.0 Objective
3.0 Main Content
    3.1 Analysing Experimental Data
        3.1.1 Processing and Analysing Experimental Data
        3.1.2 The t-test
        3.1.3 Analysis of Variance
    3.2 The Written Report for an Experiment
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

This unit is the concluding part of the discourse on experimental research. In this unit, greater attention is paid to some statistical tools to analyse experimental data. How the outcome of experimental research is reported is also a primary concern of the unit.

2.0 OBJECTIVE

At the end of this unit, you should be able to:

- analyse the experimental data and report its findings.

3.0 MAIN CONTENT

3.1 Analysing Experimental Data

3.1.1 Processing and Analysing Experimental Data

According to Poindexter and McCombs (2000:232), the following could suffice as the process for analysing data from experimental research:

- the researcher processes the result from the questionnaire that measured the dependent variable in the same manner that the survey questionnaire was processed

- a codebook, which specifies variable and columns as well as codes for the open-ended questions, would be developed
- after coding on a coding spreadsheet, the researcher runs data using a statistical program such as SPSS
- the relevant analysis for the experiment is a comparison of the responses on the subjects who were in the experimental group and those in the control group. Means are calculated and compared for the experimental group that saw the variable under test and the control group that did not see the variable.

3.1.2 The t-test

According to Poindexter and McCombs (2000:233), "a t-test, a significance test that is often used in two-group comparisons, is used to determine whether or not the mean in the experimental group is significantly different from the mean in the control group. If you use the t-test formula, you can calculate a t-value from the scores of the experimental group and control group." SPSS could be used to also calculate the t-value and determine its significance.

When the t-value that is calculated from the data from the experiment is compared to the t-value at the appropriate degree of freedom and significance level in a t-distribution table, you can determine whether the independent variable had any effect on the dependent variable. The t-test is usually for small samples$\langle n < 30 \rangle$ (Poindexter & McCombs, 2000:232).

The formula is:

$$t = \frac{\bar{x} - \mu}{s\bar{x}}$$

Where:

$\left(s\bar{x} = s/\sqrt{n}\right)$, has a t-distribution with n-1 degree of freedom. Hence, for a two-tailed test, the null hypothesis H_o is rejected if the computed t-value $\langle t \geq t\alpha/2 \rangle$ or if $t \leq -t\alpha/2$ for $n - 1$ degrees of freedom

Example
A random size of 16 from a normal population has a mean $\bar{x} = 20$ and a standard deviation $S = 4$. Does this suggests that the assumed population mean of 22 has changed? Use $\alpha = 0.05$ level of significance.

Solution (See Wimmer & Dominick (2011) and Poindexter & McCombs (2000) for further details)

Follow the following steps:

$Ho: \mu = 22;\ H1: \mu \neq 22$
$\alpha = 0.05$

Since n=16, the number of degrees of freedom is: $n - 1 = 16 - 1 = 15$. The critical value $t\ ^{\alpha}/_{2}$ for 15 degrees of freedom are $t_{0.025} = \pm 2.131$ (Use Statistics Table found in Four Figure tables. You can also find these tables inside typical statistics textbooks).

Sample mean: $\bar{x} = 20$

The test statistic $t = \frac{\bar{x} - \mu}{s\bar{x}} = \frac{20 - 22}{\frac{4}{\sqrt{16}}} =$

$$\frac{-2}{\frac{4}{4}} = \frac{12}{1} = -2$$

$$-2 > -2.131$$

we accept $Ho: \mu = 22$, and conclude that the population mean has not changed.

3.1.3 Analysis of Variance (ANOVA)

We use a t-test because it is appropriate for comparing the means from two groups such as experimental and control group or a sample and population. But often in the circles, more than two groups are studied and compared. When more than two means are compared, ANOVA can be used and SPSS could help do the multi-variable analysis.

ANOVA is used to analyse between groups' and within groups' variance to determine if different presentations contributed to significantly different levels of recall among the experiment subjects. According to Wimmer and Dominick (2011), ANOVA is essentially an extension of t-test. The advantage of ANOVA is that it can be used to simultaneously investigate several independent variables called factors. Continuing, Wimmer and Dominick (2011:315) note that "a one-way ANOVA investigates one independent variable; a two-way ANOVA investigates two independent variables, and so on."

Kerlinger (1973), as cited in Poindexter and McCombs (2000:232), however, emphasised that ANOVA is not a statistic but an approach – a way of thinking. ANOVA calculates a ratio of variance between groups and variance within groups. The resulting ratio is called the F-ratio. The F-ratio is compared to an F-ratio that can be found in an F-table in a comprehensive statistic text to determine if the results are statistically significant (Kerlinger 1973).

3.2 The Written Report for an Experiment

Many of the issues discussed in how to write the quantitative/survey report are relevant here except the method section is properly described. The following constitutes the components of the method section for a written report of an experiment:

- description of experimental design, including subjects, setting and independent and dependent variable measures
- description of procedures
- description of data processing and analysis

SELF-ASSESSMENT EXERCISE

Discuss the major difference between a t-test and ANOVA.

4.0 CONCLUSION

The experiment continues to be regarded as a valid methodology that cuts across the physical, biological, social and communication sciences. It is a research method in which an independent variable is manipulated and its effects on the dependent variable are observed therefore forming the basis of determining relationships, correlations and connections. The disadvantages notwithstanding, it does not look like mass media researchers are yet ready to dump this method of enquiry as a recent analysis in mass media-oriented journals suggest that about 30% of articles published in such journals were written based on experimental data.

5.0 SUMMARY

In this unit, you learnt about how to analyse experimental research data using samples of statistical tools like t-test, ANOVA etc. You have also been exposed to the components of a typical report in experimental research.

6.0 TUTOR-MARKED ASSIGNMENT

i. Identify and discuss how to analyse experimental data.
ii. What advantages does ANOVA have over t-test?
iii. How are Experimental results reported?

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Otokiti, S. *et al*. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Wilson, D. *et al.* (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

MODULE 6 STATISTICS AND DATA ANALYSIS

Unit 1 Measures of Central Tendency
Unit 2 Measures of Dispersion
Unit 3 Recommendations for Appropriate Use of Statistics
Unit 4 Significance Testing

UNIT 1 MEASURES OF CENTRAL TENDENCY

CONTENTS

1.0 Introduction
2.0 Objective
3.0 Main Content
    3.1 Measures of Central Tendency
        3.1.1 The Mode
        3.1.2 The Median
        3.1.3 The Mean
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

Measures, which are used for the location of the centre of a distribution, are known as measures of location or measures of central tendency. The most important measure of location includes the mode, the median and the arithmetic mean. They are used to describe the characteristics of a frequency distribution table.

2.0 OBJECTIVE

At the end of this unit, you should be able to:

- calculate the common measures of location and when to use them.

3.0 MAIN CONTENT

3.1 Measures of Central Tendency

3.1.1 The Mode

The mode of a set of variant-value is the value that occurs most frequently; that is, the value that has the highest frequency in the distribution.

For example, suppose we have the following values: 50, 61, 42, 75, 75, and 64. The mode here is 75, because the value appears twice while others appear once. For the data above, we have one mode and when a distribution has one mode, it is said to be a Unimodal Distribution. Using a diagram, a unimodal distribution is one with one peak.

Consider the following values: 50, 75, 65, 75, 50, and 81. Here, there are two modes; 50 and 75. Such a distribution is said to be a Bimodal Distribution.

A distribution that has three modes is called a Trimodal Distribution. A distribution may have no mode at all. This occurs when all the variant-values have the same frequency. A distribution that has no mode is called Rectangular or Uniform Distribution.

Consider the values 19, 21, 15, 29, 30, 42. Here there is no mode.

The Mode for Group Data
The mode for a grouped frequency distribution can be found either:

- By estimating it from the frequency histogram of the distribution
- By using a formula

1. Estimation of the Mode from a Frequency Histogram

In order to estimate the mode from a frequency histogram, it would be necessary to locate on the histogram, the following:

- The modal class- class with the highest frequency
- The pre-modal class- the class before the modal class
- The post-modal class- the class after the modal class

We shall now illustrate how this estimation is done by using the following distribution of the scores of 100 students in a statistics exam.

Table 1.0: Scores of Students in a Statistics Exam

Scores	Frequency
0-10	2
10-20	10
20-30	8
30-40	8
40-50	13
50-60	23
60-70	18
70-80	10
80-90	4
90-100	4

The mode is located by dropping a perpendicular from the intersection of the two diagonals AC and BD to the X- axis. The point where this perpendicular cuts the X-axis is the mode. The estimated value of the mode from the diagram is 56.65%.

2. Calculation of the Mode Using Formula

A more accurate value of the mode for grouped data is found by using the following formula:

$$MO = L + \frac{(d1)}{(d1+d2)}i$$

Where:

- MO- Mode
- L- Lower Limit of Modal interval
- D1- difference between modal frequency and frequency of next lower interval
- D2- difference between the modal frequency and difference of the next higher interval
- i-class interval

For our statistics examination reported on Table 1.0 above, the following figures emerge:

L= 50, d1=23-13=10, d2=23-18=5, i=10

Substituting these values in the formula

$$MO = 50 + \frac{((23-13)}{(23-13) + (23-18))}*10$$

$$= 50+ (10/10+15)* 10$$

= 50+ (10/15)* 10= 50+ 6.67
= 56.67
= 56.67%

The mode is very popular in solving practical statistical problems of market and opinion research. For example, when the problem is that of planning for large-scale production of clothes and shoes, it is necessary to establish the sizes that are most frequently demanded otherwise called modal sizes (Nwabuokei, 1999).

The mode is also particularly useful in the study of qualitative data (attributes). For instance, as a result of an opinion survey, we may make a statement that the average smoker prefers Benson and Hedges. This statement simply means that more smokers prefer Benson and Hedges to other types of cigarettes. In the above examples, the use of the arithmetic mean would simply be meaningless. It would make no sense, for instance, to speak of the mean size of shorts, mean preference of smokers (Nwabuokei, 1999).

3.1.2 The Median

The median of a series is the value that divides the series into two equal halves. In order to find the median of a group of values, the values have to be arranged in an ascending or descending order of magnitude. If we have an odd number of values, the median will simply be the middle value. For instance, the daily pay of unskilled workers in NOUN is given as 500, 400, 650, 350, and 300. To find the median, we must first arrange the values in an ascending order of magnitude as follows-
300, 350, ***400***, 500, 650

The third value is the middle item, which is 400. Therefore, the median is 400.

When we have an even number of distribution, then the median will be half of the sum of the two middle numbers. For example: 500, 400, 650, 350, 300, and 450. As usual, we arrange in ascending order thus: 300, 350, 400, 450, 500, and 650

The two middle values are 400 and 450

$$\frac{400+450}{2} = \frac{850}{2} = 425$$

The Median for Grouped Data

In practice, we are often required to find the median for grouped frequency distributions. Like the mode, the median for grouped data could be found either:

- By estimating it graphically from the Ogive
- By Calculating it from a formula

Estimating the Median from the Ogive

To estimate the Ogive, you need to use the "less than" graph. In the graph, the curve AB is the Ogive of the distribution. The estimated value of the median is found by drawing a horizontal line from the 50% level (1/2 way along the Y-axis) to cut the curve AB at a point C. from the point C, we drop a perpendicular to cut the X-axis at the point M. The point M is the median we seek. The estimated value of the median is 54%.

Calculating the Median Using Formula

The median for grouped data is found more accurately by using the following formula:

$$Me = L + \frac{(N/2 - CF)i}{F}$$

Where:

- Me= Median
- L= Lower limit of the median case
- CF= Cumulative frequency of the interval immediately lower than the median interval
- F= Frequency of the median interval
- N= Total frequency
- I= Class interval

Substituting

L=50
CF= 41
F= 23
N=100
I=10 = 50+ (90/23) = 50+3.91
= 53.9%

The result is very close to the estimated value of the median obtained from the Ogive.

One obvious advantage of the median is that it is not affected by extreme values, and hence, it is preferred to the arithmetic mean in extremely skewed distributions, such as in income distributions. However, the median is only a positional average and not a mathematical concept, and hence, it is not adaptable to algebraic treatment like the arithmetic mean (Nwabuokei, 1999).

3.1.3 The Arithmetic Mean

The arithmetic mean (or simply the mean) of a group of values is the sum of these values divided by their number. Suppose we have a variable X (say, monthly earnings of a group of workers) (Nwabuokei, 1999). Let X assumes the values X1, X2, X3, X4, Xn. Then, the arithmetic mean of these values denoted as X is given as

$$X = \frac{\text{Sum of all the n values}}{n}$$

Where n is the number of values; that is

$$X = \frac{X1+X2+X3\ldots.Xn}{n}$$

Using the sigma notation, we write

$$\underline{X} = \frac{\sum xi}{1-1/n}$$

Where $n/\sum xi / x-1 = X1+X2+X3\ldots. +Xn$

If it is clear which items we are summing, we can simply write $X= \sum x/n$

Suppose we have data on the monthly income (in Naira) of five Covenant University workers as follows: X1= 50,000, X2= 65,000, X4= 45,000, X5= 70,000

Then the mean of the above values is

$$X = \frac{50,000+65,000+60,000+45,000+70,000}{5}$$

$$= \frac{290,000}{5} = 58,000.00k$$

That is, the mean monthly earning of the group of five workers is 58,000.

The above formula is the formula for simple arithmetic mean. It is so called because it is the formula used in calculating the mean when frequencies are not attached to the values (Nwabuokei, 1999).

When frequencies are attached to the values, we find the mean by using the formula for Weighted Arithmetic Mean (WAM) given as:

$$X = \frac{\sum xf}{\sum f}$$

Where $\sum xf$ = the sum of the product of the values and their corresponding frequencies.

Where $\sum f$= the sum of frequencies.

For example, find the average monthly earning using the data in the table below:

Distribution of the Monthly Earnings of a Group of Workers

Earnings (#)	No. of workers (F)
90	1
100	4
110	8
115	5
120	2
TOTAL	20

Solution

In this example, the X-values (the monthly earnings) have frequencies attached to them. Therefore, in finding the average (mean) monthly earning of the workers, it would be wrong to simply add the five X-values and divide the sum by 5. Rather, we shall weight the values with their respective frequencies and divide the result obtained by the sum of frequencies; that is to say, we are to make use of the formula for weighted arithmetic mean.

Hence the mean monthly earning of the 20 workers will be:

$$X = \frac{\sum xf}{\sum f} = \frac{(90x1)+(100x4)+(115x5)+(120x2)}{1+4+8+5+2}$$

$$X = \frac{90+400+880+575+240}{20}$$

$$X = 109.25k$$

The Arithmetic Mean for Grouped Data

When we have grouped data given in form of a frequency distribution, the individual X-values are unknown. In order to find the mean in this case, we assume that the Class Mark (Class Midpoint) of each class represents the mean value of that class. By so doing, we are able to find a good approximation of the sum of the values in each class if we multiply the midpoint of each class by its class frequency. The mean is then found by the formula for Weighted Arithmetic Mean or this:

$$X = \frac{\sum FXi}{\sum F}$$

Where X=mean, $\sum$ = summation
FXi= frequency multiplied by the midpoints.

Let's use the exam scores of the 100 students in the statistics examinations

Computation of the Arithmetic Mean for Grouped Data

Score	F	X (midpoint)	FX
0-10	2	5	-
10-20	10	15	150 (10x15)
20-30	8	25	200
30-40	8	35	280
40-50	13	45	585
50-60	23	55	1265
60-70	18	65	1170
70-80	10	75	750
80-90	4	85	340
90-100	4	95	380
Total	100		5130

Substituting $\frac{\sum xfi}{\sum f} = \frac{51300}{100} = 51.35$ Ans

Note: when we have open classes, the classes are closed accordingly, and then we proceed to find the mean as described in the steps above.

The Relationship between the Mean, the Median and the Mode
When we have a Unimodal frequency, the following relationship exists.

When a distribution is symmetrical (a distribution in which the items are uniformly dispersed with respect to the mean), the mean equals the median and the mode i.e. mean= median=mode.

When we have a positively skewed distribution (a skew distribution is one in which the dispersion of the items with respect to the mean is not uniform), i.e. one with a longer tail extending towards higher values of the variant, the mean will be greater than the median which in turn would be greater than the mode i.e. Mean>Me>mode.

When we have a negatively skewed distribution, that is, one with a longer tail extending towards lower values of the variant, then the mean is less than the median which in turn is less than the mode i.e.

X<Me<Mo.
>=Greater than
<=Less than

In summary, the results obtained in our earlier examples for the distribution of scores of 100 students in the statistics examination reads thus:

- Mode=56.7%
- Median=53.9%
- Mean=51.3%

We note that the mean is less than the median, which in turn is less than the mode. Hence, we conclude that the distribution of the scores of the 100 students is negatively skewed.

4.0 CONCLUSION

The mean, median and mode are three popular measures of central tendency that could help a researcher describe his data in terms of such related statistical calculations. To really understand your data, it is important to determine these measures of central tendency. This is not compulsory in most media research endeavours but a description of a particular study without calculating the mean, median and the mode may be missing in vital components of such description.

5.0 SUMMARY

This unit discussed three most popular measures of central tendency also called measures of location: mean, median, and the mode. Examples of how to calculate these measures were also discussed in this unit.

6.0 TUTOR-MARKED ASSIGNMENT

SCORES	FREQUENCY
0-100	4
100-200	12
200-300	10
300-400	10
400-500	15
500-600	25
600-700	12
700-800	12
800-900	8
900-1000	8

i. Calculate the mean, mode and the median for this frequency distribution table.

ii. When will it be appropriate to apply the descriptive statistics?

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Champion, D. J. (1981). *Basic Statistics for Social Research*. NY: Macmillan.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Nwabuokei, P. O. (1999). *Fundamentals of Statistics*. Enugu: Koruna Books.

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Okeke, M.C. (1997). *Statistics: A Business Approach*. Onitsha: Onwubiko Prints Ltd.

Otokiti, S. *et al*. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M.& McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Spiegel, M. R. (S.I. Edition). *Theory and Problems of Statistics*. New York: McGraw-Hill.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note Developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 2 MEASURES OF DISPERSION

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 Measures of Dispersion
        3.1.1 The Range
        3.1.2 Mean Deviation
        3.1.3 The Variance
        3.1.4 The Standard of Deviation
        3.1.5 The Co-efficient of Variation
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Reading

1.0 INTRODUCTION

Dispersion is defined as the result of spread of individual observations around any measure of central tendency, for example, the mean. This extent of spread may be small or large. If it is small, it shows that the individual observations cluster around this measure of central tendency. If it is large, it indicates this measure of individual observations are scattered over large distances from this measure of central tendency. Measures of dispersion are very useful in finding out how representative any measure of central tendency is, as description of any set of data (Nwabuokei, 1999).

If we know that the average weekly wage of skilled workers at NOUN is #25,000 a week, it is not unusual to find some workers earning #10,000, or others earning #50,000 or even somebody earning up to #80,000. In order to be sure that this average of #25,000 is a true representative of the average wage of these NOUN staff, we need to know about the spread (dispersion) of the individual wage around this wage. If this spread is small, we may be inclined to believe that this average wage is a true representative of the monthly wages of these workers (Nwabuokei, 1999).

A good measure of dispersion should be based on all the observations and should be easily calculated. The measures of dispersion that are frequently mentioned in statistics are:

- Range

- The mean deviation
- The variance and
- The standard deviation

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- identify, explain and correctly calculate the following measures of dispersion: Range, Mean Deviation, Variance, and Standard Deviation
- sum the entire unit with the Co-efficient of variation which is an equalising measure.

3.0 MAIN CONTENT

3.1 Measures of Dispersion

3.1.1 The Range

The range (R) is the simplest measure of dispersion. It is the difference between the greatest and the least values in the distribution. This is stated in symbols thus:

- R= Xmax – Xmin

Where R=Range
Xmax = the greatest value in the distribution
Xmin = the least value in the distribution

Example: Suppose we have data on the monthly earnings of six NOUN staff as follows: #20,000, #34,000, #19,000, #55,000, #94,000, and #33,000.

The Range = R = # 94,000 - #19,000=75,000

When we have a grouped frequency distribution, the range is considered to be the difference between the lowest class limit of the lowest class and the upper class limit of the highest class.

Example
Frequency Distribution of Age of Staff Members at NOUN

Group	Frequency
10-20	5
20-30	8
30-40	7
40-50	4

Find the Range of this distribution.
The Range R = 50 – 10 = 40 Answer

Uses of the Range
The range is usually employed in situations where the aim is to know the extent of extreme variations. In business for example, the performance of stocks in the stock market is usually reported in terms of their ranges, by quoting the low and high prices during each session. Another example of the use of the range is meteorology. Meteorologists usually report low and high temperatures rather than the daily hourly readings (Nwabuokei, 1999).

In modern business statistics, the range is widely used for the control of the average of manufactured products. For example, when a machine is turning out large numbers of a particular item, a small sample of four or five items is examined; say every 30 minutes, to see whether the process is remaining constant within the limits of error allowed. A quality control chart for ranges can be set up to ensure that variation within the samples is not too much (Nwabuokei, 1999).

Challenges/Limitations in the Use of the Range
One major limitation to the use of the range is that it does not make use of all the observations in the distribution; rather it uses only two of the observations - the largest and the smallest. As a result of this, the range does not show the clustering of the items relative to the centre. It tends to increase with the increase in the sample size. Consequently, the range is most appropriate when the total number of observations is small (Nwabuokei, 1999).

SELF-ASSESSMENT EXERCISE

Discuss two limitations associated with the range.

3.1.2 The Mean Deviation (MD)

The mean deviation, sometimes called the average deviation, is defined as the arithmetic mean of the absolute values of the deviations from the mean. The formula for the mean deviation is given below:

For Ungrouped Data

$$MD = \frac{\Sigma |Xi - \bar{x}|}{N}$$

Where di = Xi – x = the deviation of each value from the mean. The notation, / / means that the sign of deviations of the values are neglected; that is to say that negative deviations are treated as positive. E.g. in a statistical quiz for six students who missed the original examination the following scores were recorded over 20 points: 6, 4, 5, 8, 10 and 15. Calculate the mean deviation.

Answer

Formulae $\frac{\Sigma |di|}{N}$

Scores (Xi)	Mean ($\bar{x}$)	Absolute deviation for the mean (/di/ = / X – X/)
6	8	6 – 8 = 2
4	8	4 – 8 = 4
5	8	5 – 8 = 3
8	8	8 – 8 = 0
10	8	10 – 8 = 2
15	8	15 – 8 =7
TOTAL		18

$$\bar{x} = \frac{6 + 4+5+8+10+15}{6} = \frac{48}{6} = 8$$

Applying the formulae

$$\frac{\Sigma |di|}{N} = \frac{18}{6} = 3 \text{ Answer}$$

Note : Σdi = summation of the absolute deviation from the mean.

N = Total number of observations.

For Grouped Data

Formulae $=di = \frac{\Sigma |Xi - \bar{X}Fi|}{\Sigma f i}$

Example: Compute the MD for a daily pay of a group of 20 laborers at NOUN recorded thus:

- 3 persons earn between 3.70 – 4.60, 4 earn between 4.60 – 5.50, 6 earn between 5.50 – 6.40, another 3 earn between 6.40 – 7.30 and 4 earn 7.30 – 8.20

Daily pay naira	No of laborers (F)	Midpoint (Xi)	(Xi Fi)	Deviation in the mean (Xi–$\bar{x}$ =/di/)	/ X - $\bar{x}$ / (/di / Fi)
3.70 – 4.60	3	4.15	12.45	4.15 – 5.995 = -1.845	1.845 x 3 = 5.535
4.60 – 5.50	4	5.05	20.20	-0.945	3.780
5.50 – 6.40	6	5.95	35.70	-0.045	0.270
6.40 – 7.30	3	6.85	20.55	+ 0.855	2.565
7.30 - 8.20	4	7.75	31.00	+ 1.755	7.020
Total	20		119.90		19.170

Mean = $\Sigma xf = 119.90 = 5.995$
$\Sigma f \quad 20$

Applying the formulae $\Sigma /di / \frac{Fi}{\Sigma f} = 19.170 = \# 0.9585$
20

3.1.3 The Variance

The variance is the average of the squared deviations of the items from their mean. Denoted by S^2, the variance is calculated by the formulae:
For ungrouped data $S^2 = \frac{\Sigma (X- X)^2}{N} = \frac{\Sigma \bar{d}i^2}{N}$

Where $di = (Xi – \bar{X})$

For grouped data $S^2 = \frac{\Sigma (X- X)^2 f}{\Sigma f} = \frac{\Sigma \bar{d}i^2 Fi}{\Sigma f}$

Computation of the Variance for Ungrouped Data
The computation of the variance for ungrouped data is carried out as follows:

- determine the Mean
- determine the deviation, d, of each item from the mean, i.e for each item, Find $di = (Xi – \bar{X})$
- square these deviations, i.e determine $d^2 = (Xi – \bar{X})^2$
- find the sum of these squared deviations, i.e determine $\Sigma (X – \bar{X})^2$
- divided this sum by the total number of observations n, i.e compute

- $S^2 = \Sigma (X - \bar{X})^2 = \frac{\Sigma di^2}{N}$ $\quad N$

By this five operations, we obtain the variance (S^2)

Computation of the Variance for Grouped Data
In order to calculate the variance for grouped frequency table, we proceed as follows:

- determine the X values. These are the class midpoints
- find the Mean of the distribution
- find the deviations of the Midpoints from the Mean, $di = (Xi - \bar{X})$
- square these deviations, i.e. determine $di^2 = (Xi - \bar{X})^2$
- multiply the squared deviations by their corresponding frequencies, i.e.
- determine $(Xi - \bar{X})^2F$
- compute $\Sigma (Xi - \bar{X})^2F$
- divide the sum in (No. 5 above) by the sum of frequencies to obtain the variance
- $S^2 = \frac{\Sigma X^2F}{\Sigma f} - \frac{(\Sigma XF)^2}{\Sigma f}$

3.1.4 The Standard Deviation

Because of the operation of squaring, the variance is expressed in square units rather than in the original unit. It becomes necessary, therefore, to extract the square root to restore the original unit. The measure of dispersion obtained by taking the square root of the variance is called the standard deviation. Hence, the standard deviation could be defined as the positive square root of the variance. It is sometimes called the root-mean-squared deviation (Nwabuokei, 1999). Denoted by the letter S, the standard deviation is computed by the following formulae:

For Ungrouped Data

$$S = \sqrt{\frac{\Sigma (X - \bar{X})^2}{N}} = \sqrt{\frac{\Sigma di^2}{N}}$$

For grouped data

$$S^2 = \sqrt{\frac{\Sigma (X-X)^2F}{\Sigma Fi}} = \sqrt{\frac{\Sigma di^2Fi}{\Sigma Fi}}$$

While the mean deviation is computed in reference to either arithmetic mean or the median, the standard deviation is always calculated in

reference to the arithmetic mean. The standard deviation is by far the most important measure of dispersion and as a result it features very prominently in statistics. For example, it is one of the parameters that determine the mathematical equation for the probability distribution of the continuous normal variable. It also used in computing various measures of skewness. It is also used in testing the reliability of certain statistical measures (Nwabuokei, 1999).

Computation of Standard Deviation for Ungrouped Data

The computation of the variance and the standard for ungrouped data is carried out as follows:

- determine the Mean
- determine the deviation, d, of each item from the mean, i.e for each item, Find $di = (Xi - \bar{X})$
- square these deviations, i.e determine $d^2 = (Xi - \bar{X})^2$
- find the sum of these squared deviations, i.e determine $\Sigma (X - \bar{X})^2$
- divided this sum by the total number of observations n, i.e compute
- $S^2 = \frac{\Sigma (X - \bar{X})^2}{N} = \frac{\Sigma di^2}{N}$

By this five operations, we obtain the variance (S^2)

Take the square root of the variance to obtain the standard deviation, i.e find

$$S = \sqrt{\frac{\Sigma (X - \bar{X})^2}{N}} = \sqrt{\frac{\Sigma di^2}{N}}$$

Example: A random variable X assumes values 2, 3,5,6,4
Find the variance and the standard deviation.

Answer
Since there are no frequencies attached to the values, we have to use the formula for ungrouped data

Variable Xi	Deviations from the mean $di = (Xi - \bar{X})$	Square deviations $d^2=(Xi - \bar{X})^2$
2	2-4 = -2	$-2^2 = 4$
3	3-4 = -1	$-1^2 = 1$
5	5-4 = 1	$1^2 = 1$
6	6-4 = 2	$2^2 = 4$
4	4-4 = 0	$0^2 = 0$
20		10

$$\bar{X} = \frac{\Sigma X}{N} = \frac{20}{5} = 4$$

The variance is therefore $S^2 = \frac{\Sigma (Xi - \bar{X})^2}{N}$

$$S^2 = \frac{10}{5} = 2$$

Standard Deviation $S = \sqrt{S^2} = \sqrt{2} = 1.414$

Computation of the Variance and the Standard Deviation for Grouped Data

In order to calculate the variance and the standard deviation for grouped frequency table, we proceed as follows:

- determine the X values. These are the class midpoints
- find the mean of the distribution
- find the deviations of the Midpoints from the Mean, $di = (Xi - \bar{X})$
- square these deviations, i.e. determine $di^2 = (Xi - \bar{X})^2$
- multiply the squared deviations by their corresponding frequencies, i.e.
- determine $(Xi - \bar{X})^2 F$
- compute $\Sigma (Xi - \bar{X})^2 F$
- divide the sum in (No. 5 above) by the sum of frequencies to obtain the variance
- $S^2 = \frac{\Sigma X^2 F}{\Sigma f} - \frac{(\Sigma XF)^2}{\Sigma f}$
- take the square root of the variance to obtain the standard deviation:

$$S^2 = S^2 = \sqrt{\frac{\Sigma (Xi - \bar{X})^2 F}{\Sigma F}} \quad - \quad \sqrt{\frac{\Sigma X^2 F}{\Sigma F} - \frac{(\Sigma XF)^2}{\Sigma F}}$$

Example

Use the data in the table below above to compute the variance and the standard deviation for the grouped frequency table

Computation of the Variance and the Standard Deviation for Grouped Data

Daily pay of labourers in #	No of labourers (F)	Midpoints (Xi)	Xi x Fi (XiFi)	Xi - $\bar{X}$ = di	$(X - \bar{X})^2$ di^2	$(X - \bar{X})^2 Fi = d^2 Fi$
3.70 -4.60	3	4.15	12.45	-1,845	3.404	10.21
4.60 -5.50	4	5.05	20.20	-0.945	0.893	3.57
5.50 -6.40	6	5.95	35.70	-0.045	0.002	0.01
6.40 -7.30	3	6.85	20.55	0.855	0.731	2.19
7.30 -8.20	4	7.75	31.00	1.755	3.080	12.32
	20		119.90			28.30

Note: Mean = $\frac{\Sigma Xifi}{\Sigma fi} = \frac{119.90}{20} = \#5,995$

$S^2 = \frac{\Sigma (X- X)^2 fi}{\Sigma fi} = \frac{28.30}{20} = 1.415$ Answer

$S = \sqrt{S^2} = \sqrt{1.415} = \underline{\#1.19}$ Answer

It is pertinent to mention here that in finding the standard deviation from sample data, the sum of frequencies, Σf, that is, the total number of observation (n) is replaced by n-1. The reason for this is that the resulting value would be a better estimate of the standard deviation of the population from which the sample is drawn. Usually n is replaced by n-1 when the size of the sample is small. For large sample size, n or n-1 could be used. In either case, there would be very little difference in the answer. A sample of size $n<30$ is regarded as small. When $n \geq 30$, we say the sample is large.

3.1.5 Relative Dispersion: The Coefficient of Variation

The standard deviation is a measure of absolute dispersion in a set of observations. It is expressed in the unit in which the observations are made. Because of this limitation, it becomes impossible to compare variability in two or more series of values which are measured in different units or whose means are quite different. In such situations, comparison is possible only if we use a measure of variability that would relate the dispersion to some equalising factor. Such a relative measure is obtained by expressing the standard deviation as a percentage of the arithmetic mean. This relative measure of dispersion is called coefficient of variation (Nwabuokei, 1999). The coefficient of variation (CV) is defined as:

$CV = \frac{S}{\bar{x}} * \frac{100}{1}$

[The standard deviation divided by the mean, multiplied by 100 over 1]

Example
At NOUN, the average monthly earning of semi-skilled male workers is #200 with a standard deviation of #15; while for semi-skilled female workers, it is #180 with a deviation of #9. In which of these groups of workers are monthly earnings more uniform?

Solution
For Males: S = #15, $\bar{X}$ = #200

$$CV = \frac{15}{200} \times \frac{100}{1} = 7.5\% \text{ Answer}$$

For Females: S = #9, $\bar{X}$ = 180

$$CV = \frac{9}{180} \times \frac{100}{1} = 5\% \text{ Answer}$$

Since the CV is greater for the male than for the female workers, we conclude that monthly earnings for the females were relatively more uniform, that is, displayed less variation than earnings of the male workers.

4.0 CONCLUSION

According to Obikeze (1990:104), "describing our data by summary averages alone does not give a complete picture of the distribution. It requires the use of other summary measures that describe the spread or dispersion of the data from the centrality values to complete the picture". This is the main value that measures of dispersion bring to statistical manipulations.

5.0 SUMMARY

This unit discussed measures of dispersion with particular attention to the range, mean deviation, variance, standard deviation, and coefficient of variation which is a relative dispersion. It is now my understanding that with the examples of each given, you can comfortably calculate each of the measures of dispersion discussed.

6.0 TUTOR-MARKED ASSIGNMENT

The distribution of capital investment of 100 companies is represented below:

Capital Investment (#m): **1-3, 3-5, 5-7, 7-9, 9-11, 11-13**
Number of Companies: **15, 30, 20, 15, 15, 5**
Calculate the following:

i. The mean of the distribution
ii. The mean deviation
iii. The variance
iv. The standard deviation
v. The Co-efficient of variation.

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Champion, D. J. (1981). *Basic Statistics for Social Research*. NY: Macmillan.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Nwabuokei, P. O. (1999). *Fundamentals of Statistics*. Enugu: Koruna Books.

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Okeke, M.C. (1997). *Statistics: A Business Approach*. Onitsha: Onwubiko Prints Ltd.

Otokiti, S.*et al*. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M.& McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Spiegel, M. R. (S.I. Edition). *Theory and Problems of Statistics*. New York: McGraw-Hill.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 3 RECOMMENDATIONS FOR APPROPRIATE USE OF STATISTICS

CONTENTS

1.0 Introduction
2.0 Objective
3.0 Main Content
    3.1 Recommendations for Appropriate Use of Statistics
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

Many statistical tools are available for researchers to use in the analysis of their data. This usually becomes very problematic for the young researcher who is unfamiliar with the various statistical tools and the best place to use them. This unit sums up our discussion so far and gives you a guideline on the appropriate statistical tool to apply to any statistical analysis.

2.0 OBJECTIVE

At the end of this unit, you should be able to:

- do summary view of the recommendation for the use of appropriate statistical tools in data analysis.

3.0 MAIN CONTENT

3.1 Recommendations for Appropriate Use of Statistics

Many statistical techniques are available to help the mass communication researcher correctly interpret data and unearth important relationships between variables. The availability of many statistical tools also makes it imperative that the researcher makes efforts to select only the appropriate statistic in his analysis assignment. This assignment is not usually easy for some beginning researchers and academics who attempt to publish in reputable international journals. Corroborating the above views, Poindexter and McCombs (2000:148) note thus:

- Although statistical programs that specialize in analysing survey data can easily and quickly produce statistics and report their

> significance, it is still your responsibility to select a statistic that is appropriate for your data. For example, it is incorrect to use statistics only suitable for interval or ratio levels of measurements on nominal level data. Your data must conform to the level of measurement that the statistics requires. Also, in a two-variable analysis, when one variable is at the nominal level and the other at the ordinal level of measurement, you should select a statistic that is appropriate for the variable at the lower level of measurement (Weaver, 1981:85).

Poindexter and McCombs (2000) further note that some statistical tools are more powerful than the others and advise that the most powerful of the statistical tools should be used per time for any type of data so that one can maximise the benefits and understand the survey results better.

Below is a summary guide to selecting appropriate statistics and analytical techniques as prepared by Poindexter and McCombs (2000:149) thus:

Guide to Selecting appropriate Statistics and Analytical Techniques

S/N	**Level of Measurement**	**Analysing One Variable**
1.	Nominal	Frequencies, Mode, Rank Order, Confidence Interval
2.	Ordinal	Frequencies, Mode, Median, Range
3.	Interval or Ratio	Frequencies, Mode, Median, Mean, Variance, Standard Deviation
	Level of Measurement	**Analysing Two Variables**
1.	Nominal	Cross-tabulation, chi square, phi, Cramer's V
2.	Ordinal	Kendall's $tau_{b,}$ Kendall's $tau_{c,}$ gamma, Spearman's rho
3.	Interval or Ratio	Pearson's product-moment correlation (r), t-test
	Level of Measurement	**Analysing Two Variables**
1.	Nominal	Cross-tabulation
2.	Ordinal	Cross-tabulation
3.	Interval or Ratio	Factor analysis, Analysis of Variance (ANOVA), Multiple regression

4.0 CONCLUSION

In conclusion, the views of Poindexter and McCombs (2000) suffice here as they argue that "although the appropriate use of statistics is very important in understanding the data, it is also important to evaluate the

quality of the overall sample. If the sample is poor, sophisticated statistics will not make the data better".

5.0 SUMMARY

This unit x-rayed the recommendations for the appropriate use of statistics for researchers in media related studies. It also discussed the suggested guide to selecting appropriate statistics and analytical techniques for social science researchers.

6.0 TUTOR-MARKED ASSIGNMENT

i. Discuss two analysis techniques for one variable concerning a nominal level data.
ii. Discuss the situations when cross-tabulation will be the appropriate type of statistical tool to use.

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics*. New York: Longman.

Champion, D. J. (1981). *Basic Statistics for Social Research*. NY: Macmillan.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill.

Nwabuokei, P. O. (1999). *Fundamentals of Statistics*. Enugu: Koruna Books.

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Okeke, M.C. (1997). *Statistics: A Business Approach*. Onitsha: Onwubiko Prints Ltd.

Otokiti, S. *et al*. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M.& McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Spiegel, M. R. (S.I. Edition). *Theory and Problems of Statistics*. New York: McGraw-Hill.

Wilson, D. *et al*. (2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.

UNIT 4 SIGNIFICANCE TESTING

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
- 3.1 Hypothesis Testing
- 3.2 Nature of Research Questions
- 3.3 Research Hypothesis: A Definition
  - 3.3.1 Advantages of Using Hypothesis in Communication Research
  - 3.3.2 Ingredients of a Good Hypothesis
- 3.4 Testing Hypothesis for Statistical Significance
- 3.5 The Null and the Alternative Hypotheses
- 3.6 Two-Tailed Test and One-Tailed Test
- 3.7 Two Types of Errors
- 3.8 Level of Significance
- 3.9 The Critical Region and The Acceptance Region
- 3.10 Test of Hypothesis Using the X^2 and the F-Distributions
- 3.11 Hypothesis Testing Using X^2
- 3.12 Contingency Tables
- 3.13 Types of Chi-Square Tests

4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 Reference/Further Reading

1.0 INTRODUCTION

To have an insight into the reasoning involved in statistical hypothesis testing, let us examine the following procedure with which we are familiar. Consider the process by which an accused person is tried in a law court under our legal system. The accused, behind the bar, is considered innocent until proven guilty. The proof of guilt rests on the prosecution. In the language of hypothesis testing, we say that we want to test the null hypothesis Ho, that the person behind the bar is innocent. An alternative hypothesis H1: that the person behind the bar is guilty therefore exists. Available evidences before the jury are now examined. If the jurors find that the evidence is inconsistence with the null hypothesis, they will reject the Ho and therefore accept the alternative, that the defendant is guilty (Nwabuokei, 1999).

Analogous to the above illustration and on the assumption that the null hypothesis is true, Nwabuokei (1999) argue that “a random sample is

taken and necessary computations are made. If the results obtained from the sample differ significantly from those expected under the assumption that the null hypothesis is true, we say that they are significant and they would be inclined to reject Ho. The above procedure which enables us to decide whether or not to reject any null hypothesis Ho is called 'Test of Hypothesis', 'Test of Significance or Rule of Decision'.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define a hypothesis
- distinguish it from a research question
- compute a typical hypothesis and make a rational decision based on the analysis.

3.0 MAIN CONTENT

3.1 Hypothesis Testing

This is an important aspect of data analysis that demands special attention. Before one begins any form of research that will involve two or more variables, a research question (RQ) or Statistical Hypothesis must be formulated to guide the research. These shall also form the basis of analysis. The RQ and hypothesis are identical except that RQ's do not predict outcomes whereas hypothesis do.

3.2 Nature of Research Questions

a. Researchers who use research questions are not basically interested in testing the statistical significance of their findings.
b. Researchers who use research questions are more interested with discovering general indications than statistical significance.
c. Research questions also help to gather preliminary data that could lead to the formation of hypothesis.
d. Research questions are also used mostly in explanatory studies which aim to research for data indications rather than an attempt to find causality (Fukey, 1986 cited in Wimmer & Dominick, 2011).

3.3 Research Hypothesis: A Definition

Hypothesis is a tentative and testable proposition about research variables or phenomena. A statistical hypothesis is a statement or assumption about an unknown population parameter. A test performed

in order to verify whether a hypothesis is true or false is called a test of hypothesis (Nwabuokei, 1999).

The hypothesis about any population parameter is tested by using information obtained from a sample drawn from the population in question. If the result obtained from the sample, (in form of calculated sample quantities), is inconsistent with the hypothesis being tested, we have to reject the hypothesis; on the other hand, sample evidence supporting the hypothesis leads to its acceptance. Consequently, a statistical hypothesis is formulated for the purpose of rejecting it or accepting it (Nwabuokei, 1999).

3.3.1 Advantages of Using Hypothesis in Communication Research

a. They provide direction and focus for the study.
b. They help researchers to eliminate trial and error research.
c. They also help out intervening and confounding variables.
d. They also allow for the qualification of variables.

3.3.2 Ingredients of a Good Hypothesis

a. A good Hypothesis should be compatible with current knowledge in the area **[compatibility].**
b. A good Hypothesis should be stated briefly **[brevity].**
c. A good Hypothesis should be testable **[testability].**
d. A good Hypothesis should be tentative **[tentative].**
e. A good Hypothesis should be logically consistent **[consistency].**

SELF-ASSESSMENT EXERCISE

i. Discuss four ingredients of a good hypothesis.
ii. Distinguish between Hypothesis and Research Questions

3.4 Testing Hypothesis for Statistical Significance

Hypothesis testing could also be regarded as significance testing. At the end of this process, the researchers either reject the Ho or accept it. Usually a probability level or significance level is stated as a starting point of test. The two most common probability levels are .01 and .05 or 1% and 5% levels of significance respectively. A 5% level of significance indicates that the researcher has a 5% chance of making a wrong decision about negating the null hypothesis or accepting the H_1 (research hypothesis)

However, we must note that researchers usually state the H_1 rather than Ho. In reality therefore acceptance or rejection applies mostly to H_1. Furthermore, if the outcome of the study indicates a probability lower than the significance level, the researcher should reject the Ho and if it has a high probability, Ho is accepted.

3.5 The Null and the Alternative Hypotheses

The hypothesis that is being tested is called the Null Hypothesis and is denoted by Ho. The hypothesis that we are willing to accept if we reject the null hypothesis is called the Alternative Hypothesis denoted by H_1. We shall now illustrate these two hypotheses with the following examples.

Example
Suppose that during the Alpha Semester, a class of 100 students was taught statistics by the use of a certain method. At the end of the semester, the class recorded an average score of 65% in this subject with a standard deviation of 5%. During the Omega Semester, a new statistics lecturer was employed to handle the course. The new lecturer claims that he has developed a new method for teaching the course, which according to him, is more effective than the first one. In order to test his claim, a sample of 30 students taught by the new method was examined at the end of the Omega Semester. The average score was found to be 68%. The question arises: Is this new method really more effective than the old one? In other words, should a decision be taken to adopt the new method of teaching statistics? In order to answer this question, we need to carry out a test of hypothesis (Nwabuokei, 1999). Such a test will enable us verify the following possibilities:

- whether the observed higher sample average of 68% was actually due to the fact that the second method is better than the first one, that is, whether there is a significant difference between the two methods, or
- whether the observed sample average occurred by chance. For example, it might have happened that the sample was made up of the very brilliant students. If that is the case, it would be wrong to attribute the higher average score observed for the 30 students to the superiority of the second method (Nwabuokei, 1999).

On the other hand, according to Nwabuokei (1999), suppose the sample of 30 students so selected recorded an average score of say 63%, which is lower than the average scored by the entire class taught by the first method during the Alpha Semester. Here, again, a test of hypothesis will enable us verify:

- whether the observed low sample score of 63% has occurred because the new method is by no means more effective than the old one
- whether the observed low sample average score occurred because the sample happened to be made up of less brilliant students; and not because the new method is less effective than the old one

The null hypothesis is usually specified in terms of the population parameter of interest. It is the hypothesis of no difference, and consequently, it is stated with the equality sign. Throughout the process of analysis, the Ho is assumed true. Evidence provided from the sample that is inconsistent with the null hypothesis leads to its rejection. On the other hand, evidence supporting the hypothesis leads to its acceptance (Nwabuokei, 1999).

3.6 Two-Tailed Test and One-Tailed Test

A two-tailed test is a test in which the alternative hypothesis is non-directional (i.e. it is two-sided). For a two-tailed test, when the null hypothesis is rejected, the alternative hypothesis does not indicate whether the true mean is greater or less than the specified in the null hypothesis. A two-tailed test has two critical values (Nwabuokei, 1999).

A one-tailed test, sometimes called a one-sided test, is one where the alternative hypothesis is one-sided (directional). In such a case, when the null hypothesis (Ho) is rejected, we conclude that the true mean is as specified by the alternative hypothesis. A one-tailed test has one critical value- to the right of the distribution for case (1) and to the left of the distribution for case (2) (Nwabuokei, 1999).

We can also test hypothesis about a population proportion. For example, we may wish to test that a population proportion P – is equal to a specified value Po, against the alternative that it is not. We state our hypothesis as follows:

1. Ho: P = Po
   H1: P ≠ Po (a two-tailed test)
2. Ho: P = Po
   H1: P > Po (a one-tailed test to the right)
3. Ho: P = Po
   H1: P < Po (a one-tailed test to the left)

3.7 Two Types of Errors

In testing statistical hypothesis we can commit two types of errors, - the TYPE 1 ERROR and the TYPE 2 ERROR

1. **Type 1 Error**
   A Type 1 Error has been committed if we reject the null hypothesis, Ho when in fact it is true. This error is sometimes called the α-error (Alpha error).
2. **Type 2 Error**
   A Type 2 Error has been committed if we accept the null hypothesis, Ho when in fact it is false. This error is sometimes called the β-error (Beta error). In other words, a type 2 error is made when Ho is erroneously accepted.

It is pertinent to mention here that these two types of errors arise because the truth or falsity of Ho is unknown; even after it is accepted or rejected. Consequently, that we accept Ho does not necessarily mean it is true. In the same manner, that we reject Ho does not mean it is false (Nwabuokei, 1999).

3.8 Level of Significance

The level of significance of a test is the probability of committing the type 1 error; that is, the probability of rejecting Ho when in fact it should be accepted. It represents the highest probability with which we are willing to risk a type 1 error. The level of significance is denoted by alpha sign α, and its magnitude is usually specified before samples are drawn for the test (Nwabuokei, 1999).

The most frequently used levels of significance in hypothesis testing are $\alpha = 0.05$ (5% level of significance), and $\alpha = 0.01$ (i.e. 1% level of significance). Other values of α, say $\alpha = 0.001$, 0.10 etc could also be used. Suppose a test is performed at a 5% level of significance (i.e. $\alpha = 0.05$), it means that there are 5 chances in a hundred that true null hypothesis would be rejected (Nwabuokei, 1999). A test is said to be **significant** if the Ho is rejected at the 5% level. A test at 1% level of significance means that there is only one chance in a hundred that a true Ho will be rejected. A test is said to be highly significant if the null hypothesis is rejected at the 1% level (Nwabuokei, 1999).

3.9 The Critical Region and the Acceptance Region

Having selected the level of significance of a test, the CRITICAL REGION sometimes called the REJECTION REGION is determined. This rejection region (critical region) is defined as a set of the possible values of the test statistics that leads to the acceptance of the null hypothesis. The size of the rejection region is identical with the magnitude of the level of significance, while the size of the acceptance

region coincides with the confidence coefficient, 1 – α (Nwabuokei, 1999).

A test statistics serves as a decision maker in hypothesis testing, because a decision to reject or accept the null hypothesis depends on its magnitude. The values of the test statistic that separate the acceptance region from the rejection region are called CRITICAL VALUES.

3.10 Test of Hypothesis Using the X^2 and the F-Distributions

Very often, we are interested in testing whether more than two population proportions, or more than two population means may be considered equal. The x^2(chi-square) tests are employed for testing more than two population proportions. Tests concerning more than two population means are based on the F distribution.

Expected and Observed Frequencies
Sometimes in practice, we observe that the results obtained in random samples do not actually correspond to the expected results, in accordance with the rules of probability. So in hypothesis testing both the expected and the observed frequencies must be determined before the hypothesis is computed correctly.

3.11 Hypothesis Testing Using X^2

Usually, when testing any hypothesis with X^2 statistic, the null hypothesis (Ho) is that observed and expected frequencies agree with one another, while the alternative hypothesis (H_1) is that they do not agree. Under this assumption, the expected frequencies are computed from a random sample and the value of X^2 calculated. If this value of X^2 is greater than $X^2(\alpha,V)$ – the critical value for the X^2 at the alpha level of significance, with V degrees of freedom, then Ho is rejected and we conclude that there is a significant difference between observed and expected frequencies. In essence, the Ho is accepted if the two sets of frequencies are identical or almost so. If, on the other hand, the discrepancies between the expected and the observed frequencies are large, the Ho is rejected.

3.12 Contingency Tables

A table in which observed frequencies are arranged in r-rows and c-columns (r,c>1) is called contingency table. For instance:

Opinion of Students on OPEC

OPINION	Covenant University	UNILAG	LASU	TOTAL
In favor of remaining in OPEC	74	128	57	259
Against remaining in OPEC	24	70	23	117
TOTAL	98	198	80	376

The contingency table is suitable for testing the independence of two attributes, when it is possible to classify each observed frequency into one and only one cell.

The formula for chi square is $\mathbf{X^2 = \Sigma\ (0^2/E) - N}$ **OR**

$$\mathbf{X^2 = \frac{\sum (O_i - E_i)}{E_i}}$$

Where:
N = Total number of frequencies
O = Observed data
E = Expected data
$\sum$ = Summation

Note also:
Degree of freedom = **(C-1)(R-1)** where
C = Number of columns
R = Number of Rows

Expected is usually gotten from **CM x RM** where
N

CM = Column marginal
RM = Row marginal
N = Total number of frequencies

Example
In a statistical investigation to determine if consumption of cosmetics depends on the material status of the consumers, in a random sample of 280 consumers (both sexes) was taken. It revealed that 12 married people, 86 unmarried people and 17 “others” consume large quantities of cosmetics while 20 married people, 125 unmarried people and 20 “others” consume very little quantities of cosmetics. On the basis of the

above information can one conclude at a significant level of 5% that consumption of cosmetics depends on marital status?

Solution
We first formulate the Ho & Hi as
Ho: Consumption of cosmetics and marital status are independent
H1: Consumption of cosmetics and marital status are not independent

We now display the information in a contingency with 2 Rows and 3 Columns thus:

	Married	**Not married**	**Others**	**TOTAL**
Large Cosmetics	12 (13.14)	86 (86.66)	17 (15.20)	115
Little Quantity of cosmetics	20 (18.86)	125 (124.34)	20 (21.80)	165
TOTAL	**32**	**211**	**37**	**280**

Using the formula X^2 = $\mathbf{X^2 = \Sigma (0^2/E)-N}$

We can construct another table to reflect this formula thus: Note – You can give letters of the alphabet to the sections of the contingency table (Columns & Rows).

	Observed (O)	Expected (E)	O^2	O^2/E
A	12	13.14	144	10.96
B	86	86.66	7396	85.36
C	17	15.20	289	19.01
D	20	18.86	400	21.21
E	125	124.34	15625	125.67
F	20	21.80	400	18.35
Total				280.56

Applying the formula

X^2 = $\mathbf{X^2 = \Sigma (0^2/E)-N}$
$X^2 = 280.56 – 280$
X^2 = 0.56 Answer

For the degree of freedom (DF)
Formula = (C-1)(R-1)
DF = (3-1)(2-1)
DF = 2x1
= 2

For the 5% significance level = 0.05

X^2 tabulated value (the critical value of X^2)
= 5.991 (this was found from the X^2 - table in statistics text books)

Decision
Since our X^2 computed is less than X^2 tabulated i.e. lies in the acceptance region, we accept Ho at $\alpha = 0.05$ and conclude that consumption of cosmetics and marital status are independent. That is, however, in consonance with the expected conclusion, contrary to the speculation that application of cosmetics is a habit for young spinsters and bachelors.

Example 2:
A researcher is interested in finding out if the perception of Nigerian media experts on the potential influence of globalisation and ICT on local cultural values is directly related to the knowledge of such experts in the application/use of ICT. Table 1 below shows the observed frequency for the hypothesis.

Table 1: Observed Frequencies for the Hypothesis

Perception on Cultural influence	Knowledge in the Use of ICT		
	Positive	**Negative**	**Not Sure**
Strongly Disagree	04	02	01
Disagree	-	-	-
Undecided	23	04	02
Agree	190	08	02
Strongly Agree	29	04	01

Use the formula $X^2 = \sum\left(\frac{O^2}{E}\right) - N$

Solution
H_0: The perception of Nigerian media experts on the potential influence of globalisation and ICT on local cultural values is not directly related to the knowledge of such experts in the application/use of ICT.

H_1: The perception of Nigerian media experts on the potential influence of globalisation and ICT on local cultural values is directly related to the knowledge of such experts in the application/use of ICT.

Calculated Value:
X^2 Computations

0	**E**	**0^2**	**$0^2/E$**
04	6.38	16	2.51
02	0.47	04	8.51
01	0.16	01	6.25
23	26.42	529	20.02
04	1.93	16	8.29
02	0.64	04	6.25
190	182.22	36100	198.11
08	13.33	64	4.80
02	4.44	04	0.90
29	30.98	841	27.15
04	2.27	16	7.05
01	0.76	01	1.32
			291.16

Applying the formula $\mathbf{X^2 = \sum \frac{(O^2)}{(E)} - N}$

$X^2 = 291.16 - 270$
$= 21.16$

Table Value
D.F. = (C-1) (R-1) = (3-1) (5-1)
= 2 x 4 = 8
:. X^2 (alpha= 0.05, 8) = 15.507

Decision
Since X^2 computed (21.16) is greater than X^2 tabulated (15.507), i.e. lies in the rejection region, we reject the H_0 and accept H_1 (There is a significant difference in observed and expected data or frequencies).

3.13 Types of Chi-Square Tests

a. Test concerning more than two population proportions
b. Test of independence
c. Test of goodness-of-fit

In each of these tests, the X^2 provides a means of comparing a set of observed frequency with a set of expected frequencies. Test concerning more than two population proportions is sometimes called the test of homogeneity. The word homogeneity in statistics is used to indicate "same" or "equal" (Wimmer & Dominick, 2011).

Test of independence is another form of contingency test in which we apply the X^2 distribution of two variables based on sample data. The example we did above is a test of independence.

Test of Goodness-Of-Fit

Test of Goodness-of-fit, according to Wimmer and Dominick (2011) "seeks to determine if a particular population has a specified theoretical distribution such as the binomial, poisson and the normal distributions. It is based on how good a fit we have between actually observed frequencies of sample data and the theoretical frequencies obtained from a hypothesised distribution. Mass media researchers often compare the observed frequencies of a phenomenon with the frequencies that might be expected or hypothesised." For example, assume that a researcher wants to determine if the sales of TV sets by four manufacturers in the current year are the same as the sales during the previous year. A logical hypothesis might be: "TV set sales of four major manufacturers is significantly different this year than those of the previous year."

Wimmer and Dominick (2003) solved this particular statistical problem thus:

Suppose the previous year's TV set sales were distributed as follows:

Manufacturer	Percent of sales
RCA	22
Sony	36
JVC	19
Mitsubishi	23

From these previous year's sales, the researcher can calculate the expected frequencies (using a sample of 1000) for each manufacturer's sale by multiplying the percentage of each company's sales by 1000. These are the expected frequencies:

- RCA – 220
- Sony – 360
- JVC – 190
- Mitsubishi – 230

Next, the researcher surveys a random sample of 1000 households known to have purchased one of the four manufacturers' TV sets during the current year. The data from this survey provides the following information:

Manufacturer	Expected frequency	Observed frequency
RCA	220	180
Sony	360	330
JVC	190	220
Mitsubishi	230	270

The researcher now must interpret this data to determine whether the change in frequency is actually significant. This can be done by reducing the data to a chi-square statistic performing a test known as chi-square "goodness-of-fit" test.

A chi-square, as we noted earlier, is simply a value that shows the relationship between expected frequencies and observed frequencies this is computed by this formula:

$$X^2 = \sum \frac{(O_1-E_1)^2}{E1}$$

Where
O_1 is the observed frequencies
E_1 is the expected frequencies.

This means that the difference between each expected and observed frequencies must be squared and the divided by the expected frequency. The sum of the quotients is the chi-square for those frequencies. For the frequency distribution above, chi-square is calculated as follows:

$$X^2 = \frac{(O_1-E_1)^2}{E_1} + \frac{(O_2-E_2)^2}{E_2} + \frac{(O_3-E_3)^2}{E_3} + \frac{(O_4-E_4)^2}{E_4}$$

$$X^2 = \frac{(180-220)C}{220} + \frac{(330-360)^2}{360} + \frac{(220-190)^2}{190} + \frac{(270-230)^2}{230}$$

$$= \frac{(-40)^2}{220} + \frac{(-30)^2}{360} + \frac{(30)^2}{190} + \frac{(40)^2}{230}$$

$$= \frac{1600}{220} + \frac{900}{360} + \frac{900}{190} + \frac{1600}{230}$$

$$= 7.27+2.50+4.75+6.95$$

$X^2 = 21.45$

Note (You can consult www.georgetown.edu/cball/webtools/web_chi.html for an internet calculator for the chi-square statistic)

Once the X^2 value is known, the goodness-of-fit test determines whether this value represents a significant difference in the frequencies. To do this, two values are necessary. The first is probability level, which is pre-determined by the researcher; the second, called degrees of freedom

(DF), is the number of scores in any particular test that are free to vary in value. For example, if one has three unknown values (X,Y, and Z)such that X+Y+Z = 10, there are two df: any two of the three variables may be assigned any value without affecting the total, but the value of the third will then be predetermined. Thus, if X=2 and Y=5, then Z must be 3 because 2+5+3 = 10

In the goodness-of-fit test, DF is expressed in terms of K-1, where K is the number of categories. In the TV set sales study, K=4 and DF = 4-1 = 3. Next, an X^2 significance table is consulted. These tables are arranged by probability level and DF.

If the calculated X^2 value equals or exceeds the value found in the table, the differences in observed frequencies are considered to be statistically significant at the predetermined alpha level; if the calculated value is smaller, the results are non significant. Therefore, if the X^2 computed value is less than (<) the X^2 tabulated, the Ho is accepted. This is for hypothesis that both Ho and H1 are stated.

In our example above, the X^2 computed is 21.45 while the X^2 tabulated is 7.815. since 21.45 is greater than 7.815, the difference is significant, and the hypothesis is supported – that is, TV set sales of the four manufacturers are significantly different in the current year than sales in the previous year.

Uses of Chi-Square Goodness-Of-Fit Test

a. Wimmer and Dominick (2003) identify the following as uses to which Chi Square Goodness-of-fit test could be put to.
b. It could be used to measure changes in studying audience perceptions of advertising messages over time
c. It could be used in planning changes in TV programming
d. It could be used in analysing the results of public relations campaigns.

Limitations to the Use of Chi-Square Goodness-Of-Fit Test

Notwithstanding the popularity of the Chi Square Goodness-of-Fit test, it has three limitations as identified by Wimmer and Dominick (2003):

- since it is a nonparametric statistical procedure, the variables must be measured at the nominal and ordinal level
- the categories must be mutually exclusive, and each observation in each category must be independent from all others
- because the chi-square distribution is sharply skewed for small samples, Type 2 error may occur. Small samples may not produce significant results in cases that could have yielded significant results if a larger sample had been used. To avoid this

problem, most researchers suggest that each category contain at least 5 observations.

As an alternative to the chi-square goodness-of-fit test, some researchers prefer the KOLOMOGORV – SMIRNOV TEST, which is considered to be more powerful than the chi-square approach. In addition, minimum number of expected frequency in each cell is not required, as in the chi-square test (Wimmer & Dominick, 2003).

4.0 CONCLUSION

To conclude this unit, the views of Wilson, Esiri and Onwubere becomes pertinent here: "After a research topic has been thoroughly researched, you should have some prediction about what you think will happen in your study. This educated guess concerning the outcome of a research is called hypothesis. A hypothesis, therefore, can be defined as a set of assumptions, a speculation, which is tentatively accepted as the basis for an investigation. The hypothesis is worded so that it can be tested in your research work. That is why it has been stated that hypotheses are measurable, testable statements about the relationship(s) between variables."

5.0 SUMMARY

This unit focused on hypothesis testing otherwise known as significance testing. It juxtaposed hypothesis and research questions to identify the core differences in both. We now have the understanding of when and how to use both research questions and hypothesis. Above all, I am now confident you can calculate hypothesis using standardized formulas.

6.0 TUTOR-MARKED ASSIGNMENT

i. A media researcher is interested in finding out if there is a relationship between violent behaviour displayed by children and the number of hours exposed to televised violence. The Table below summarises the data generated from the survey.

	Number of hours exposed to tv violence		
Behaviour of Children	**Between 1 – 5hours**	**Between 5 – 10hours**	**Between 10 and above**
Violent	55	40	55
Non violent	25	15	10

ii. Test the null hypothesis that violent behaviour is independent of number of hours exposed to televised violence at the 1% alpha level of significance. Do not forget to suggest the hypothesis for this enquiry.

7.0 REFERENCES/FURTHER READING

Bruning, J. L. & Kintz, B. L. (1997). *Computational Handbook of Statistics.* New York: Longman.

Lucey, T. (1998). *Quantitative Techniques-An Instructional Manual*. London: Dp Publications Ltd.

Maner, M. (2000). *The Research Process: A Complete Guide and Reference for Writers.* Boston: McGraw Hill

Obikeze, D. O. (1990). *Methods of Data Analysis in the Social and Behavioural Sciences*. Enugu: Auto-Century Publishing Company Ltd.

Otokiti, S. *et al*. (2007). *Contemporary Statistical Methods.* Lagos: Vantage Publication Company.

Poindexter, P. M. & McCombs, M. E. (2000). *Research in Mass Communication (A Practical Guide).* Boston, USA: Bedford/St. Martins.

Sobowale, I. A. (2009). *Scientific Journalism* (2nd ed.). Lagos: Idosa Konsult.

Wilson, D. *et al*.(2008). *Communication Research*. Unpublished Lecture Note developed for the National Open University of Nigeria (NOUN).

Wimmer, R. D. & Dominick, J. R. (2011, 2003). *Mass Media Research. An Introduction*. Belmont, CA: Thomson/Wadsworth.