FEDERAL UNIVERSITY OTUOKE

STATISTICS, THE DISCIPLINE CREATED BY GOD AND APPLIED CONSCIOUSLY AND UNCONSCIOUSLY IN OUR EVERYDAY LIVES

By

JOHNSON OHAKWE

B.Sc., M.Sc., PhD

Professor of Applied Statistics

11th Inaugural Lecture

Wednesday, 29th November, 2023



JOHNSON OHAKWE
B.Sc, M.Sc, PhD**.**
Professor of Applied Statistics

Statistics, the Discipline Created by God and Applied Consciously and Unconsciously in our Everyday Lives

Copyright: Federal University Otuoke, Bayelsa State

ISBN: 978-978-57207-0-8

First published 2023
All Rights Reserved

No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the permission of the copyright owners.

Printed by:
Eagle Eye Printing Press
Port Harcourt, Rivers State
+2348036716938

DEDICATION

This lecture is solely dedicated to my beloved father, Mazi Johnson Okere Ohakwe, for all his efforts toward my education.

ACKNOWLEDGEMENTS

First and foremost, I thank the Almighty God, the "Unmoved Mover" and "Uncaused Causer" of all things without whom my journey through life would not have been possible. I also wish to take a moment to express my deep appreciation to HIM, the ultimate source of knowledge, wisdom, and inspiration.

Yahweh, your name is Yahweh
You are the miracle working God, your name is Yahweh
(one more time)

I would like to extend my heartfelt appreciation to the Vice Chancellor, Professor Teddy Charles Adias for providing me the opportunity of presenting this 11th Inaugural lecture.Your leadership has not only shaped the present but also illuminated the path to a brighter future for our institution. We are privileged to have a leader whose passion for education, commitment to excellence, and unwavering dedication to our academic community continue to inspire us all. Thank you, sir.

Let me also quickly express my profound gratitude to the Deputy Vice Chancellor, Prof. Chris Onyema who took me like his younger brother. Your kind words of advice and guidance are inspirational to me. You are truly a big brother.

Furthermore, I humbly commend other principal Officers of the University for providing support to the Vice Chancellor: Mr. Kenneth Sonrandein (Acting Registrar), Professor (Mrs.) Felicia E. Etim (University Librarian) and Mr.Preteowei Topman Suwari (Bursar).

To all my teachers, from primary school to university, I extend my immense gratitude. Some of them include Mr. Okoro Ajua and Miss Asi (now Mrs Mark Ogbonna) at Ahaba Imenyi Community

Primary School, Mr G. Anorzie, Mr I. Okorie, Mr. O. Ukoh, Mr. Okemiri and Mr. Iroh at Ahaba Imenyi Secondary School.Others include Prof. D.C. Chikezie (who took me like a son), Prof. A.C. Akpanta (who supervised my B.Sc project), Prof. Aloy Onyeka, Prof. J.C. Nwabueze (the first female Professor of Statistics in Nigeria and former Deputy Vice Chancllor, Academic, Michael Okpara University of Agriculture Umudike) and Prof. C. O. Omekara at Abia State University Uturu. Your dedication, passion for teaching, and unwavering belief in the potential of your students have had a profound impact on my life. I am forever grateful for the immeasurable gifts you have bestowed upon me.
I specially extend my heartfelt thanks to Prof. Iheanyi Sylvester Iwueze who supervised my M.Sc. dissertation and PhD. thesis. This is the man that I called an "Academic Bulldozer" and a tree that made an "academic forest". When he joined Abia State University, Uturu, the University became the leading center for postgraduate programmes in Statistics within the Southeast region. Thank you, Sir. You are an extra-ordinary supervisor. Once he gave you an academic task to crack, he also went home to crack the same task, so that the next meeting with him was to compare results (those from the supervisor and the supervisee). Sir,you opened our eyes to the beauty of Statistics. As I progress in my career, I am committed to upholding the principle of excellence, dedication, and mentorship that you exemplified.

I appreciate all my colleagues in the Faculty of Science: the Dean Professor (Mrs.) Iheoma M. Adekunle, Professor Collins C. Chiemeke (Sub-Dean), Professor Stephen O. Ibe (my HOD), Professor Okpo Unanaonwi (HOD Biology), Dr Nnamdi Omehe (HOD, Physics), Dr Olufunso O. Abosede (HOD, Chemistry), Dr. Ifeoma Vivian Anyiam (HOD, Microbiology), Dr. Dennis Kpomah (HOD, Biochemistry), Prof. Mbosowo Etukudoh (Director, SIWES), Dr Akens Hamilton Amachree (Ag. Director, Academic Planning), Prof. Eno Ekanem, Prof. Stephen A.

Uzoekwe, Dr Chibuzo Amaefula, Dr UzoAnekwe, Dr Nlerum A. Promise and Dr (Mrs) Chinonye Obasi and all others that could not be mentioned here. You are all appreciated.

I also appreciate specially my colleagues in the Department of Mathematics and Statistics and Computer Science and Informatics: Dr Egbuhuzor P. Udechukwu, Dr Akpanibah E. Edikan, Dr Eli Innocent, Dr (Mrs) Amadi Chinyere, Dr (Mrs.) Udoh Ndipmong, Mr Azor Promise, Mr.Samaila Kupongoh, Mr Okardi Bioble, Mrs. Moko Anasuodei, Dr (Mrs.) May Stow, Mrs. Maudlyn Victor-Ikoh, Dr Ikirigo Samuel Apigi, Eleje Best Chinwendu, and many others that could not be listed as a result of want of space.

Special thanks to the Dean, School of Postgraduate Studies, Prof. Benjamin Okaba; Dean, Faculty of Education, Prof. Pac Ordu; Dean, Faculty of Management Sciences; Prof. Cletus Akenbor; Dean, Faculty of Humanities, Prof. Akachi Odoemene; Dean, Faculty of Social Sciences, Prof. Francis Oluleye; Dean, Faculty of Engineering, Prof. Scott Akpila; Director, Research and Quality Assurance, Prof Emmanuel S. Akpan; Director, Centre for Continuing Education, Prof. Diri I. Teilanyo; Prof. Okechukwu Edward Okeke (a great uncle); Ag. Director, Entrepreneurship Development Centre, Dr Raimi Lasisi; and Ag. Director, Institution of Foundation Studies, Dr (Mrs.) Oyintonyo Micheal-Olomu. You all are great friends.

I also wish to appreciate my colleagues in the Faculty of Physical Sciences Abia State University, Uturu, where I have been an Adjunct Lecturer since 2010: the Dean, Prof. J.C. Igwe, Prof. J. N. Nnadi, Prof. Osu Bright, Dr Uche C. Onyemachi (my HoD), Dr Eric Otuonye, Professor I.C. Nnorom (HOD, Industrial Chemistry), Dr H. N. Odikwa (HOD, Computer Science),Dr E. Nwaokoro (HOD, Physics), Prof. I.U.Chiemeka, Prof. A. Onwuegbuche, Dr. A. Obike, Dr. Chidinma. Olunkwa, Mrs

Fidelia Lawrence-Kiwu, Miss Okereke-Jude Esther Ogechi, Mrs Adiele and Mr. Bright Offorha, among others.

I wish to extend my special appreciations to the entire staff of Federal University Otuoke for their support. You are all great and I love you all.

Specific mention must be made of the contributions of the following past Vice Chancellors to my growth in the University Community:

(i) Prof. S.O. Emejuaiwe offered me appointment of a Graduate Assistant in the Department of Statistics, Abia State University, Uturu in 2000, as the best graduating student.

(ii) Prof. Fidelis Uzochukwu Okafor, the ViceChancellor, Anambra State University, Uli, offered me appointment as Lecturer I in the Department of Statistics in 2010 and made me the pioneer HOD of the Department.

(iii) Prof. Mobolaji Aluko (the pioneer Vice Chancellor of Federal University, Otuoke) gave me appointment as a Senior Lecturer in 2012.

(iv) Prof. Seth Accra Jaja promoted me to the rank of Reader and Professor and also made me the HOD of the Department of Mathematics and Statistics for four years. Sir, your contribution to my academic career would long be stenciled on my mental sheath as long as I live.

I cannot easily forget the following personalities for imparting positively on my life: Mazi Johnson OkereOhakwe(my father) and Mrs Patience Iroaku Ohakwe (my mother, of blessed memory). Others include Hon. Samuel Ezelu Ogbonna, Hon. Prince Ohochukwu and his wife Dame Roseline Ohochukwu, Mrs. Nwanyinnaya Obioma, Mrs. Nkechi Ekeh, Dr Obrien Osuiwu (my

in-law), Mr. Gabriel Oyemike (my In-law), Mr. Barinyima Badon (my In-law), Dr (Bar) Chika Odoemena (Dean, Faculty of Law, Edwin Clark University), Mr.Onyekachi Onyeforo, Dr Friday Moses Onuobia, Engr. Ugochukwu O. Igwe, Ohakwe Abel Onyema, Rev. (Dr) Ejikeme Ibegbulam, and Mr Onwuka Akachukwu.

Special thanks to my siblings: Mrs Monica Osuiwu, Mrs. Lilian Ihuoma Gabriel Oyemike, Mr.Okapara Johnson, Mr. Ikechukwu Johnson, Mrs.Chinonyerem Barinyima and Mr. Chukwuemeka Johnson.

To my Committee of Friends, led by Dr Olufunso Abosede, I say a big thank you for galvanizing such a huge support for this project. You all are great. I love you all and may God bless you bountifully.

A super thanks to Prof I. S. Iwueze , Prof. A. C Akpanta and Prof. Okechukwu E. Okeke for their invaluable contributions to improve the quality of this lecture material.

I cannot extend my appreciation to my family without first thanking the parents of my wife, my in-laws – Mr. Mark Ejiofor and late Mrs. Chinenye Mark Ejiofor – for allowing me to marry their only daughter. The biblical saying in Proverbs 18: 22 ("He who finds a wife finds a good thing…") is very affirmative of my marriage with my beloved wife, Comfort Ogechi. Indeed, her coming into my life ushered in lots of blessings into my life. God bless you, my Baby. To our children—Chidiebube, Chinomso, Chinemere, and Chukwudike—you are all wonderful and may God continue to preserve you in Jesus Name. Amen.

Professor Johnson Ohakwe
Otuoke, November 2023

CONTENTS

Dedication ----------iv
Acknowledgements ---------- v
Table of Contents ---------- x
List of Tables ----------xii
List of Figures----------xiii
Protocols ---------- xiv
Preamble ---------- xvi

Chapter One: Statistics and How they are Used ---------- 1
1.0 Introduction ---------- 1
1.1 Definition of Statistics ---------- 1
1.2 Two Broad divisions of Statistics ---------- 1
1.2.1 Descriptive Statistics ---------- 2
1.2.2 Inferential Statistics ---------- 4
1.3 Data and Methods of Data Presentations ---------- 12
1.3.1 Data----------12
1.3.2 Data presentation ---------- 12
1.3.2.1 Tabular Technique----------13
1.3.2.2 Graphical Technique ---------- 16
1.4 Statistical Enquiry Cycle ----------22

Chapter Two: Variations and Uncertainties24
2.0 Introduction ----------24
2.1 Variations and Uncertainties: An Integral part of Human Life ----------24
2.1.1 Variations----------24
2.1.2 Uncertainties ----------25
2.2 Who created variations? ----------26
2.2.1 The Creator of variations/diversities ----------26
2.2.2 The Creator of Uncertainties ----------27
2.3 God, the Father of Statistics----------28
2.4 Statistics, a Super Discipline that has Applications in virtually Every Field of Study ----------30

2.5 Statistics, a Unique Field that is unconsciously Applied in our Everyday Lives---------- 33

Chapter Three: Power of Statistics in Decision Making
3.0 Introduction ---------- 35
3.1 The Power of Statistics in Decision-making under Potential Variations and Uncertainties---------- 35
3.1.1 The Power of Statistics in Decision-making under Potential Variations ---------- 35
3.1.2 The Power of Statistics in Decision-making under Uncertainties ---------- 37
3.2 Statistics, a tool that can lead to misguided decisions 40
3.3 Misuses and Abuses of Statistics ---------- 41
3.4 Outliers and its Impact on data-driven decision making 44

Chapter Four: Statistics, a tool for Good Governance and National Development---------- 47
4.0 Introduction ---------- 47
4.1 Statistics, a tool for Good Governance and National Development ---------- 47
4.2 Establishment of National Bureau of Statistics: A recognition of the power of statistics for good Governance and National Development ---------- 48

Chapter Five: My Humble Contribution to Knowledge-- 50

Chapter Six: Suggestion for Future Research, Recommendations and Conclusion ---------- 61
6.1 Suggestion for future research ---------- 61
6.2 Recommendations ---------- 61
6.3 Conclusion---------- 62

References ---------- 64

LIST OF TABLES

Table 1.1: Scores of students on seven subjects ---------- 10
Table 1.2: Confidence interval for the mean score of the seven students in Table 1.1 ---------- 10
Table 1.3: Survey results on the favourite soup of a group of Ten (10) students---------- 13
Table 1.4: Frequency table of students' test scores -------- 14
Table 1.5: Survey Results of a group of 100 Respondents on Preferred Car Make---------- 14
Table 1.6: Monthly Average Income of Two Groups of Workers (Group A and Group B)---------- 15
Table 1.7: Data Representing Monthly Sales of Fruits by a Market Woman---------- 16
Table 1.8: Average Monthly Temperature of a City in Abia State over a period of one year ---------- 18
Table 1.9: Height and Weight of a Group of Ten (10) Statistics Students ---------- 19
Table 3.1: Mean and Standard deviation of Daily Revenue from GWBS and GWAS ---------- 36
Table 3.2: Annual Stock Returns for Stocks A and B for a period of Six years. ---------- 38
Table 3.3: Mean Return and Standard Deviations of Stocks A and B in Table 3.2.---------- 38
Table 3.4: Investment options---------- 39
Table 3.5: Weekly Returns of two Pastors (A and B) in a Local Church ---------- 44
Table 3.6: Summary Statistics of the data in Table 3.5----- 45

LIST OF FIGURES

Figure 1.1: A Simple Bar Chart Representing the Monthly Sales of Fruits by Market Woman ---------- 17
Figure 1.2: A line graph of the Average Monthly Temperature (^{0}C) Versus Month ---------- 18
Figure 1.3: Scatter plots of Weight and Height for a group of Ten (10) Statistics Students ---------- 20
Figure 1.4: Pie Chart of the Values of Fruits: Mango, Banana and Carrot sold by a Market Woman - 21

PROTOCOLS

The Vice Chancellor and Chairman of this occasion, Distinguished Professor Teddy Charles Adias

Members of FUO governing Council

The Deputy Vice Chancellor, Distinguished Professor Chris Onyema

Other Principal Officers of the University:

- The Registrar, Mr Keneth Suonrandein
- The Librarian, Prof. Felicia Etim;
- The Bursar, Mr. Preteowei TopmaSuwari

Deans of School of Postgraduate Studies, Faculties and Student Affairs

Professors

Directors of Institutes, Directorates, Centres and Units

Heads of Departments,

Distinguished Academics and Professional Colleagues

Staff of Federal University Otuoke here present

Royal Fathers here present

My Lords, Spiritual and Temporal

Great FUO Students

Friends of the University

Distinguished Invited Guests

My wife and other family members
Gentlemen of the press

Distinguished Ladies and Gentlemen

PREAMBLE

I need not pause to say how greatly delighted I am to stand before you to present the 11th Inaugural lecture of this great University titled "Statistics, the Discipline created by God and applied consciously and unconsciously in our everyday lives". Mr Vice Chancellor, I cannot thank you enough for this great privilege given to me but can only say God bless you. As we already know, an inaugural lecture is one of the major celebrations in academia. It is an opportunity for a professor to share the highlights of his academic endeavours with the academic community, family and friends and with the general public. To put it nicely, an inaugural lecture is an opportunity for the Professor to convince the academic community and the general public the relevance of his discipline and his own contribution to its development and this I will try to do in the subsequent pages.

Mr. Vice Chancellor, before I proceed with the calling of this lecture, I will give a brief history of my journeys into the world of statistics and into academia.

FROM PETROCHEMICAL ENGINEERING DREAMS TO THE WORLD OF STATISTICS

Admission to Study Statistics

Life often has a way of leading us down unexpected paths, changing our plans and aspirations. In the unpredictable journey of education, dreams sometimes take a detour, and it is in those moments of adjustment that we discover new perspectives, opportunities, and passions. Such was my case—a student who once harbored fervent dreams of studying Petrochemical Engineering in his secondary school days but offered admission to study Statistics instead.

I took my first JAMB Examination when I was in SS2 and did not bother to check the result because I just wanted to have experience of what the exam was all about. In my second JAMB Examination, I applied to study Petrochemical Engineering in University of Port Harcourt (UNIPORT), as first choice University and also Petrochemical Engineering in Federal University of Technology (FUTO), Owerri as second choice University. When the admission list came out in UNIPORT my name was not on the list. However, to my greatest surprise as a novice, I saw names of candidates with much lower scores than mine on the admission list. Bewildered as I was about the development, I simply concluded that it was because I did not know anybody that my name did not appear on the list.

Having felt disappointed by the university of my first choice, I decided to go to FUTO and also check the admission list. The day I was going to FUTO through Okigwe, I overhead people inside the bus discussing the sale of supplementary forms at Abia State University, Uturu (ABSU). I then stopped over at ABSU, purchased the form and never continued the journey to FUTO.

When the supplementary admission list was published my name was on Physics programme. I immediately applied for change of discipline to study Industrial Chemistry because I did not want the A-1 grade I had in Chemistry to be wasted. As providence would have it, the Co-ordinator of the Physical Science Programme, Late Professor Nkemdirim Onuoha approved the intended change of discipline, but the Dean, Late Professor Mike Oleka disapproved it and approved Statistics for me. That was how I got to study Statistics in Abia State University, Uturu, and this I can confidently say was providential.

My Journey into Academia

My journey into academia is also providential. Before my graduation as an undergraduate student, my Academic Adviser, then Mr. Anthony Chidi Akpanta(now Prof. Anthony Chidi Akpanta) called me to his office and asked me, "Would be interested in joining the University as a Lecturer?" My answer was "No". He further asked: "Why, because you have the potentials to be one". I simply replied that I have no interest in teaching but rather interested to work in banks, multinational companies, and federal government ministries and agencies.

Immediately, after my youth service in February 2000, I travelled to Abuja on an invitation by one of my uncles who was the Deputy Chairman of Bwari Area Council in Abuja. My uncle, Hon. Samuel Ezelu Ogbonna, did not want me to be employed in the Area Council but by any of the federal government ministries, departments and agencies, and he was making frantic efforts in this regard. After a couple of months, precisely five months of waiting, my patience suddenly cracked and I sent a letter to Prof. A.C. Akpanta asking if the opportunity of joining the University as a Graduate Assistant (GA) was still available and he responded by asking me to come back immediately and apply, and that I did.

As a matter of coincidence, the week I enquired from Prof A.C Akpanta on the availability of the position he earlier encouraged me to apply for was the week the Vice Chancellor, Prof S. O. Emejuaiwe announced that Departments across the University should encourage their best graduating students to apply for the position of Graduate Assistant. Who knows whether I would have had the opportunity to join academia if I did not make the right enquiries at the right time! That is the reason why I always say that my journey into academia is also providential.

CHAPTER ONE:

STATISTICS AND HOW THEY ARE USED

1.0 INTRODUCTION

This chapter contains the definition and the two broad divides of statistics, data and methods of data presentations and statistical enquiry cycle (SEC).

1.1 DEFINITION OF STATISTICS

Statistics has been defined in many ways by different scholars and academic bodies. One scholar defines it as the science of not only collecting, classifying, presenting, summarizing and tabulating data including the drawing of inferences using observed data and the whole problem of making decision in the face of uncertainty, but also the designing of theoretical basis for analyzing data (Akpanta, 2002).The American Statistical Society defined statistics as the science of learning from data, and of measuring, controlling and communicating uncertainty.

On the basis on my experience as a Statistics scholar, I define statistics elaborately as the science of variations and uncertainties that systematically examines, quantifies, and analyzes the inherent variability and unpredictability in data generated from observations, experiments, physical and natural phenomena and so on. It provides tools, methods, and techniques to describe variations, infer from data, quantify uncertainties, make informed decisions, identify patterns and predict future outcomes.

1.2 TWO BROAD DIVISIONS OF STATISTICS

There are two broad divisions of Statistics, namely Descriptive Statistics and Inferential Statistics.

1.2.1 DESCRIPTIVE STATISTICS

Descriptive statistics provide a summary of the main characteristics of a data set, such as measures of central tendency (e.g., mean, median, mode) and measures of dispersion (e.g., range, and standard deviation).

(i) Measures of Central Tendency

Mean: The mean, also known as the arithmetic mean or average, is calculated by summing all the values in the data set and dividing by the total number of values. It represents the balancing point of the data and is sensitive to extreme values. Given a set of data, $x_1, x_2, \ldots, x_n$ of size, n, the formula for the sample mean is:

$$\text{Sample mean} = \bar{\mathbf{x}} = \frac{\left(\textbf{Sum of all values}\right)}{\textbf{Number of values}} = \frac{\sum_{i=1}^{n} \mathbf{x_i}}{\mathbf{n}} \qquad (1.1)$$

Median: The median is the middle value of a data set when it is sorted in ascending or descending order. If the data set has an odd number of values, the median is the value at the centre. If the data set has an even number of values, the median is the average of the two middle values. The median is less affected by extreme values and provides a measure of the central value.

Mode: The mode represents the most frequently occurring value or values in a data set. It is useful for identifying the most common value or values in a distribution. In some cases, there may be no mode if all values occur with the same frequency, or there may be multiple modes if multiple values occur with the highest frequency.

(ii) Measures of Dispersion

Measures of dispersion, also called measures of variability, are measures that help us understand the spread or dispersion of a data

set. Some common measures of variability include the range, variance, standard deviation and coefficient of variation.

Range: The range is the simplest measure of variability and represents the difference between the maximum and minimum values in a data set. It gives an idea of the total spread of the data but is sensitive to outliers.

Variance: Variance is a measure of how much the values in a data set deviate from the mean. The unbiased estimate of the variance is calculated by taking the squared differences between each value and the mean, and dividing with the total number of observations minus one. Variance provides an understanding of the overall variability but is influenced by extreme values. Mathematically, the sample variance of the n values $x_1, x_2, \ldots, x_n$ is:

$$\textbf{Sample Variance (SV)} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \qquad (1.2)$$

To calculate the sample variance (SV), we follow these steps:

1. Calculate the sample mean (average) of the data set by summing all the values and dividing by the total number of values.
2. For each data point, subtract the mean and square the difference.
3. Sum up all the squared differences obtained in step 2.
4. Divide the sum of squared differences by the total number of data points less by one (n – 1) to obtain the sample variance.

Sample Standard Deviation (SD): The sample standard deviation (SD) is the square root of the sample variance and provides a more easily interpretable measure of variability. It represents the average amount by which values in the data set differ from the mean. A higher standard deviation indicates a

greater dispersion of values. The SD is calculated as the square root of (1.2), that is SD is calculated as

$$\mathbf{SD} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}} \qquad (1.3)$$

Coefficient of Variation (CV): It is a statistical measure that quantifies the relative variability of data points in a data set, expressed as a percentage. The coefficient of variation is particularly useful when comparing the variability of data sets with different units or scales. Mathematically,

$$CV = \left(\frac{S\tan dard\ \ deviation(\sigma)}{Mean(\mu)}\right) * 100 \qquad (1.4)$$

where:
σ = Population Standard deviation
μ = Population Mean or Average
Given a set of data, $x_1, x_2, \ldots, x_n$ of size, n, the sample coefficient of variation, r is given as

$$r = \frac{\bar{x}}{SD} * 100 \qquad (1.5)$$

1.2.2. INFERENTIAL STATISTICS

Inferential statistics involves using information from a sample to make inferences, predictions, and generalizations about a larger population. It allows scientists or researchers to draw conclusions and make statements about the population based on the analysis from a representative sample. Inferential statistics includes techniques such as hypothesis testing, confidence intervals, and regression analysis.

(i) Hypothesis Testing

Z-Test: This is used to test for equality of two population means when the standard deviations are known or can be estimated or when the distributions of samples are either normally distributed or the sample sizes are greater than 30.

T-Test: This is used also to test for equality of two population means when the standard deviations are unknown but can be estimated and the sample sizes are less than 30.

Chi-Square Test: This is used for testing relationships or independence between categorical variables.

F-Test: This is used to compare the variances of two or more groups to test for equality of the variances.

(ii) Confidence Intervals

Confidence intervals provide a range within which a population parameter (e.g., mean, variance, and proportion) is likely to fall with some level of confidence.

(**iii) Regression Analysis:** Linear and non-linear regression models are used to explore relationships between variables and make predictions.

Hypothesis Testing

It is not an overstatement to say that every scientific research involves a form of hypothesis testing. It helps researchers make evidence-based conclusions and provides a structured framework for evaluating the validity of claims using empirical data. The ultimate goal of hypothesis testing is to assess whether the patterns observed in sample data are likely due to real effects or simply random variation.

The process of hypothesis testing typically involves the following steps:

1. **Formulate Hypotheses:** Start by formulating two hypotheses: the null hypothesis (H_0) and the alternative hypothesis (H_a or H_1). The null hypothesis represents the default assumption or no effect, while the alternative hypothesis represents the claim you want to find evidence for.
2. **Collect Data:** Obtain a sample of data relevant to the hypothesis being tested. This could involve conducting experiments, surveys, observations, or any other data collection methods.
3. **Choose Significance Level:** Select a significance level (often denoted as α) that represents the threshold for considering evidence strong enough to reject the null hypothesis. Common significance levels are 0.05 (5%) and 0.01 (1%).
4. **Perform Statistical Test:** Choose an appropriate statistical test based on the nature of the data and the hypotheses being tested. The test will generate a test statistic that quantifies how far the sample data deviates from what is expected under the null hypothesis.
5. **Calculate the test statistic or the P-value:** The test statistic is a numerical value calculated from sample data that is used to make an informed decision about whether to reject or fail to reject the null hypothesis while the p-value is the probability of observing a test statistic as extreme as the one calculated, assuming that the null hypothesis is true. A small p-value indicates strong evidence against the null hypothesis.
6. **Compare P-value with Significance Level:** Compare the calculated test statistic with the critical value (determined from the chosen level of significance, α, and can easily be obtained from a statistical table) or compare the calculated p-value with the chosen significance level, α. The decision rule is to reject the null hypothesis if the calculated test statistic is greater than the critical value or equivalently the calculated p-value is less than or equal to the significance level, α. Otherwise, you do not have enough evidence to reject the null hypothesis

7. **Draw Conclusion:** Based on the comparison of the calculated test statistic with the critical value or the calculated p-value and the significance level, we make a decision about whether to reject or fail to reject the null hypothesis. This decision provides insight into the validity of the claim made in the alternative hypothesis.

THE TWO POSSIBLE ERRORS COMMITTED IN HYPOTHESIS TESTING

In hypothesis testing, there are two types of errors that can be committed namely: Type I and Type II errors.

Type 1 Error (False Positive): This is the error committed by rejecting the null hypothesis when it is true and by implication concluding that there is an effect when there is none.

Type 2 Error (False Negative): This is the error committed by accepting the null hypothesis when it is false. In other words, it happens when you fail to reject a false null hypothesis or miss detecting a true effect.

Ultimately, which error is considered more serious will depend on the potential impact of the error in a specific context. In some cases, the consequences of a Type 1 error might outweigh those of a Type 2 error, while in other cases, the opposite could be true. This is why careful consideration of the costs, benefits, and potential risks associated with each type of error is crucial when making decisions about hypothesis testing and other critical applications. In what follows, let us examine some specific scenarios where one type of error is more serious than the other.

CASE 1: SOME CASES WHERE TYPE I ERROR IS CONSIDERED TO BE MORE SERIOUS THAN TYPE II ERROR

1. **Criminal Justice System:** In criminal justice, a Type 1 error occurs when an innocent person is wrongly convicted of a crime he did not commit and possibly face a death penalty as the case may be, while Type 2 errors might result in some guilty individuals not being convicted. The social and personal repercussions of unjustly punishing or even executing an innocent person make Type 1 errors more severe.
2. **Product Safety Testing:** Imagine a safety test for a product used in critical applications, such as medical devices or aviation equipment. A Type 1 error could lead to approving a product that does not meet safety standards, potentially endangering lives. Although Type 2 errors might result in delaying the release of a safe product, the potential consequences of allowing unsafe products to reach the market make Type 1 errors more critical.

CASE 2: SOME CASES WHERE TYPE II ERROR IS CONSIDERED TO BE MORE SERIOUS THAN TYPE I ERROR

1. **Fire Detection Systems:** In fire detection systems, a Type 2 error would occur if the system fails to detect a real fire. This can be extremely dangerous in situations where early detection is crucial for saving lives and minimizing property damage. Even if false alarms (Type 1 errors) occur occasionally, failing to detect a fire can have catastrophic consequences.
2. **Airport Security for Dangerous Items:** In airport security, a Type 2 error occurs when a potentially dangerous item such as an explosive device goes undetected and makes it onto an aircraft. This could lead to serious security breaches, endangering the lives of passengers and crew. While Type 1 errors might inconvenience passengers, the potential risk

posed by failing to detect actual threats makes Type 2 errors more critical.

3. **Quality Control in Manufacturing:** In manufacturing, a Type 2 error would involve allowing defective products to reach consumers. This can lead to safety hazards, product recalls, and damage to a company's reputation. Although rejecting good quality products (Type 1 errors) might result in waste, the potential harm caused by distributing faulty products makes Type 2 errors more severe.

(ii) Confidence Intervals and Confidence Level

Researchers, analysts, and decision-makers often use confidence intervals to draw conclusions based on sample data while acknowledging the potential variability in population parameters.
Confidence interval is typically expressed as a range between two values, an upper limit and a lower limit, within which the population parameter is expected to lie.

Confidence Level

The confidence level is a measure of the degree of confidence you have that the true population parameter (e.g., the population mean or proportion) falls within a calculated confidence interval. It is typically expressed as a percentage, such as 90%, 95%, or 99%.

Interpretation

For example, suppose a 95% confidence interval for the mean height of students in a lecture hall is 4.5feet - 6.0feet, it implies that we are 95% confident that the true students' mean height falls within this interval.

It is important to mention that the higher the confidence, the wider the width of the confidence interval. For example, suppose we have the scores of a student on seven subjects as given in Table 1.1, while the confidence interval for the mean score for various confidence levels are given in Table1.2.

Table 1.1: Scores of a student on seven subjects

	Subject						
	English	Maths	Igbo	CRK	Economics	Geography	Literature
Score	56	81	52	65	69	45	51

Table 1.2: Confidence Interval for the Mean Score of the student for varying Confidence Levels

	90% Confidence Interval (CI)	95% Confidence Interval (CI)	99% Confidence Interval (CI)
CI for the Mean Score	50.699 – 69.015	48.325 – 71.389	42.325 – 77.330

The following can be deducted from Table 1.2.
The width of 90% CI < the width of 95% CI < the width of 99% CI

That is, higher confidence levels, such as 99%, will result in wider intervals but provide greater confidence in the parameter's inclusion within the interval.

(iii) Regression Analysis

Regression analysis is a statistical technique used to examine the relationship between one dependent variable and one or more independent variables. The goal is to model and understand the association between variables, predict the value of the dependent variable based on the values of the independent variables, and identify the strength and nature of the relationships. Suppose a dependent variable Y, is to be regressed on an independent variable X, a simple linear regression of Y on X can be given as follows;

$$Y_i = \alpha + \beta X_i + e_i, \text{ i} = 1, 2, \ldots, \text{n} \qquad (1.6)$$

where
Y_i is the dependent variable, for example, blood pressure

X_i is the independent variable, for example body mass index (BMI)
e_i is the error component that is assumed to be normally distributed with zero mean and constant variance, and
n is the sample size or the number of observations in a data set.
The estimates of α and β are given as follows;

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} Y_i}{n} - \hat{\beta}\frac{\sum_{i=1}^{n} X_i}{n} = \bar{Y} - \hat{\beta}\,\bar{X} \qquad (1.7)$$

$$\hat{\beta} = \frac{n\sum_{i=1}^{n} X_i Y_i - \left(\sum_{i=1}^{n} Y_i\right)\left(\sum_{i=1}^{n} X_i\right)}{n\sum_{i=1}^{n} X_i^{\,2} - \left(\sum_{i=1}^{n} X_i\right)^2} \qquad (1.8)$$

For a non-linear relationship between Y and X, we can have a model of the form

$$Y_i = \alpha + \beta X_i + \gamma X_i^{\,2} + e_i \qquad (1.9)$$

where the coefficients , α, β, and γ can be estimated from solving the following normal equations:

$$\sum_{i=1}^{n} Y_i = n\alpha + \beta\sum_{i=1}^{n} X_i + \gamma\sum_{i=1}^{n} X_i^{\,2} \qquad (1.10)$$

$$\sum_{i=1}^{n} Y_i X_i = \alpha\sum_{i=1}^{n} X_i + \beta\sum_{i=1}^{n} X_i^{\,2} + \gamma\sum_{i=1}^{n} X_i^{\,3} \qquad (1.11)$$

$$\sum_{i=1}^{n} Y_i X_i^{\,2} = \alpha\sum_{i=1}^{n} X_i^{\,2} + \beta\sum_{i=1}^{n} X_i^{\,3} + \gamma\sum_{i=1}^{n} X_i^{\,4} \qquad (1.12)$$

1.2 DATA AND METHODS OF DATA PRESENTATION

1.3.1 DATA

Data is the raw material of knowledge and is everywhere. It refers to a collection of facts, statistics, or information that can be in various forms, such as numbers, text, images, or multimedia. It is the raw material from which information and knowledge can be derived through analysis and interpretation.

Data can be classified to be primary or secondary, depending on the source. The difference between **primary** and **secondary data** in Statistics is that p**rimary data** is collected firsthand by a researcher (organization, person or group of persons, authority, agency or party etc.) through experiments, observations, surveys, questionnaires, focus groups, conducting interviews and taking (required) measurements, while the **secondary data** is readily available (collected by someone else, agencies, organization, etc.) and is available to the public through publications, journals and newspapers.

1.3.2 DATA PRESENTATION

In our data-driven world, the ability to effectively present and communicate data is a crucial skill. Whether you are a scientist, a business analyst, or a policymaker, the way data is presented plays a pivotal role in conveying information and, when properly done, has the power to transform numbers and facts into actionable insights. Effective data presentation simplifies and distills complex data set into clear and concise visual representations that facilitate understanding, enhance decision-making and even aid in storytelling.

There are two primary methods of data presentation namely; tabular and graphical techniques. Each has its advantages and is chosen based on the nature of the data, the audience, and the objectives of the presentation.

1.3.2.1 TABULAR TECHNIQUES

Tables are organized rows and columns of data. They are excellent tools for presenting data due to their structure, readability, and ability to facilitate data comparisons. Some commonly used tables are data tables, frequency tables, contingency tables and comparison tables.

(i)Data tables: They are the most often used for presenting raw data. They display data in rows and columns, where each row represents an observation or data point, and each column represents a variable or attribute. For illustrative purpose, suppose we conducted a survey on the favorite soup of a group of ten (10) FUO, students and assuming that we have the following survey data as given in Table 1.3

Table 1.3: Survey Results on Favorite Soup of a Group of Ten (10) FUO Students

S/N	Name of Student	State of Origin	Age	Gender	Favorite Soup
1	Blessing	Delta	23	Female	Okro
2	Chioma	Imo	21	Female	Ofe Owerri
3	Agnes	Bayelsa	25	Female	Banga
4	John	Bayelsa	19	Male	Banga
5	William	Akwa Ibom	24	Male	Afang
6	Jackson	Edo	22	Male	Egusi
7	Charles	Rivers	20	Male	White Soup
8	Robinson	Bayelsa	19	Male	Banga
9	Christian	Bayelsa	23	Male	Native soup
10	Ekaete	Akwa Ibom	21	Female	EdikangIkong

(ii)Frequency Tables: They are used to display the frequency or count of different values or categories in a data set. For illustration, given a data set of students' test scores:85, 90, 75, 80, 90, 85, 70, 80, 90, 85, we can create a frequency table for this data set, as follows, by counting the number of times each score appears:

Table 1.4: Frequency Table of Students' Test Scores

Score	Frequency
70	1
75	1
80	2
85	3
90	3

Table 1.4 indicates that the score of 70 appeared once, the score of 75 appeared once, the score of 80 appeared twice, the score of 85 appeared three times, and the score of 90 appeared three times.

(iii) Contingency Tables. They are also known as cross-tabulation tables and are used to show the relationships between two or more categorical variables. They are used for exploring associations and dependencies between variables. For illustration, suppose we have surveyed 100 individuals and asked them about their gender and preferred car maker, and we have recorded the following data:
40 respondents are Male, and they prefer Toyota.
25 respondents are Male, and they prefer Honda.
20 respondents are Female, and they prefer Toyota.
15 respondents are Female, and they prefer Honda.

Using a contingency table, we can present the results of the survey as given in Table 1.5.

Table 1.5: Survey Results of a Group of 100 Respondents on Preferred Car Maker

	Preferred Car Maker		
Gender	Toyota	Honda	Row Total
Male	40	25	65
Female	20	15	35
Column Total	60	40	100

(iv) Comparison Tables: They are used to compare data from different sources, groups, or time periods. They allow for side-by-side comparisons, making it easy to identify similarities and differences. For numerical illustration, suppose we want to compare the average monthly incomes of two groups, Oil workers denoted as Group A and Lecturers denoted as Group B, over the period of 5-months. Suppose, we collect the following data:

Group A:
January: ₦900,000
February: ₦1,000,000
March: ₦1,200,000
April: ₦1,350,000
May: ₦1,500,000

Group B:
January: ₦350,000
February: ₦389,000
March: ₦402,000
April: ₦450,000
May: ₦480,000

To compare the average monthly incomes of the two groups, we can create a comparison table as shown in Table 1.6.

Table 1.6: Monthly Average Income of Two Groups of Workers (Groups A and B)

Month	Group A Monthly Income	Group A Monthly Income
January	₦900,000	₦ 350,000
February	₦1000,000	₦389,000
March	₦1,200,000	₦402,000
April	₦1,350,000	₦450,000
May	₦1,500,000	₦480,000

1.3.2.2 GRAPHICAL TECHNIQUES

Some commonly used graphical tools are bar charts, line charts, scatter plots and pie charts.

(i) Bar Charts

Bar charts display data in rectangular bars of varying lengths, where the length represents the quantity being measured. They are effective for comparing discrete categories or groups. Common variations include vertical bar charts (column charts) and horizontal bar charts.

For example, suppose the data in Table 1.7 represent the sales of different fruits by a Market woman in a particular month. We can create a bar chart to visualize the sales figures.

Table 1.7: Data Representing Monthly Sales of Fruits by a Market Woman

Fruits	Sales
Apple	120
Banana	80
Orange	150
Grapes	100
Mango	70

In this simple bar chart example, we would have a vertical bar chart with the x-axis representing the fruit names and the y-axis representing the sales figures, as shown in Figure 1.1.

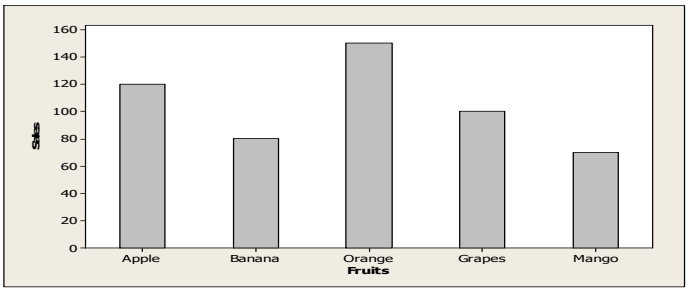


Figure 1.1: A Simple Bar Chart Representing Monthly Sales of Fruits by a Market Woman

(ii) Line Charts

A line graph, also known as a line chart, is a graphical type of data presentation that uses lines to represent the relationship between two variables over a continuous period or interval. It is commonly used to show trends, patterns, and changes in data over time.

For illustrative purpose, suppose the data in Table 1.8 represent the average monthly temperature of a city in Abia State over a period of one year.

To create a line graph for this data, we would plot the average monthly temperature on the y-axis and the corresponding months on the x-axis as shown in Figure 1.2.

Table 1.8: The Average Monthly Temperature of a City in Abia State over a Period of one year

Month	Temperature (°C)
Jan	30
Feb	32
Mar	25
Apr	28
May	22
Jun	26
Jul	30
Aug	28
Sep	25
Oct	30
Nov	35
Dec	32

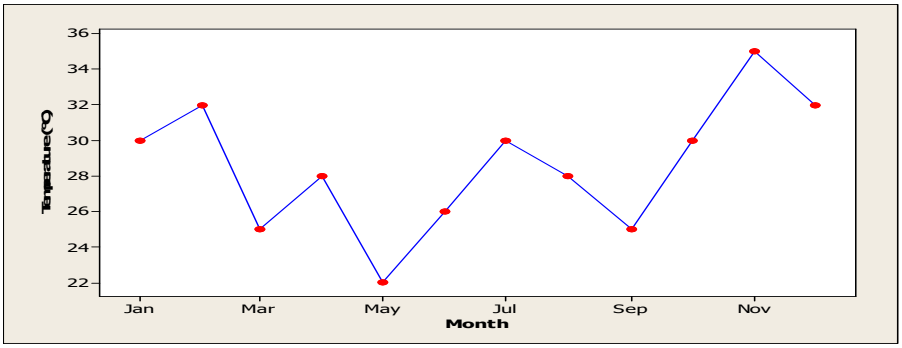


Figure 1.2: A Line Graph of Average Monthly Temperature (^{0}C) versus Month

The line graph consists of twelve data points, with each point representing the temperature for a specific month. These points are connected by straight lines, creating a continuous line that traverses the graph from left to right.

Visually, the line rises and falls, reflecting the changes in temperature over the course of the year. We can observe trends such as the fall in temperature from January to May, followed by a rise/decrease from May to December.

(iii) Scatter Plots

A scatter plot is a type of graphical data presentation method that displays the relationship between two variables. It uses a set of data points, each representing a combination of values for the two variables, and plots them on a Cartesian coordinate system. In a scatter plot, one variable is plotted on the x-axis, and the other variable is plotted on the y-axis.

Scatter plots are particularly useful for examining the relationship between variables, identifying patterns, trends, or correlations, and detecting outliers. For illustration, suppose the data in Table 1.9 represent the Height (in inches) and Weight (in pounds) of a group of 10 Statistics students and we want to examine the relationship between the height and weight of a group of the students.

Table 1.9: Height and Weight of a group of Ten (10) Statistics Students

Height (inches)	Weight (pounds)
62	120
65	140
68	160
70	165
72	180
66	150
62	120
67	156
69	170
66	168

To create a scatter plot for this data, we would plot each individual's height on the x-axis and their weight on the y-axis. Each data point represents an individual's height and weight pair. The scatter plot given in Figure 1.3 consists of ten data points, each plotted at the corresponding height and weight coordinates.

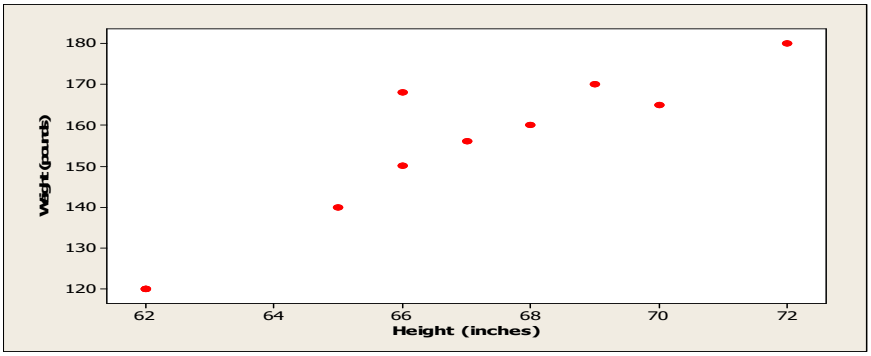


Figure 1.3: Scatter Plots of Weight and Height for a Group Ten Statistics Students

By examining the scatter plot, we can observe any patterns or trends in the data. If the data points cluster around a straight line that slopes upwards, it suggests a positive correlation between height and weight, indicating that as height increases, weight tends to increase as well. On the other hand, if the data points are scattered randomly without any clear pattern, it suggests no or weak correlation between the variables.

(iv) Pie Charts

A pie chart is a method of data presentation in a circle that is divided into sectors, representing proportions or percentages of a whole. It is used to display categorical data and show the distribution or composition of different categories within a data set.

In a pie chart, each category is represented by a slice of the pie, and the size of each slice is proportional to the value or degree or percentage it represents. The entire pie represents 360^0 or 100% or the total of the data being represented.

Suppose a Market woman has 60 mangoes, 30 oranges, 40 bananas and 30 carrots for sale. To calculate the angle in degrees for each sector corresponding to the fruits, we obtain the value (60 + 30 + 40 + 20 = 150) and calculate the degrees as follows:

First sector (Mango): (60/150) * 360 = 144 degrees
Second sector (Orange): (30/150) * 360 = 72 degrees
Third sector (Banana): (40/150) * 360 = 96 degrees
Fourth sector (Carrot): (20/150) * 360 = 48 degrees

The pie chart describing the sales of the fruits is given in Figure 1.4

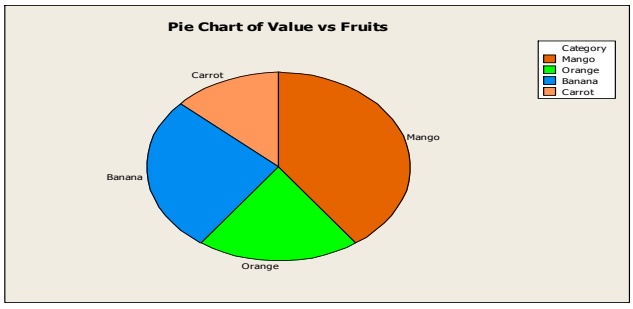


Figure 1.4: Pie Chart of the values of Fruits; Mango, Orange, Banana and Carrot

Choosing Between Tabular and Graphical Techniques

Tabular techniques are typically better for presenting precise numerical details and facilitating data comparison. However, graphical techniques are ideal for conveying patterns, trends, and relationships within the data. The choice depends on the nature of the data, the audience, and the objectives of the presentation.

Often, a combination of both techniques is used to provide a comprehensive view of the data.

1.4 STATISTICAL ENQUIRY CYCLE (SEC)

The Statistical Enquiry Cycle (SEC) is a structured approach used to guide the process of conducting statistical investigations and analyses. It helps ensure that appropriate data are collected, analyzed, and interpreted effectively to draw meaningful conclusions. Whether you are a statistician or non-statistician, we advise you to adopt the statistical enquiry cycle for any data-driven decision making. The step-by-step approach with the acronym PPDAC is given as follows;
P = Problem; P = Plan; D = Data; A = Analysis; C = Conclusion

1. Problem: This initial "P" spotlights the problem (or question) crystallization phase. This is the starting point, where you define the research question or problem you want to investigate using statistics. Formulating a clear and focused question helps guide the entire process.
2. Plan: The plan step focuses on deciding your approach, methods, tools, and resources that will be needed in solving the problem. It also involves outlining the people/objects/entities to obtain data on, what things we should measure and how we are to go about the investigation.
3. Data: The data stage is all about collecting the relevant data that can help answer the research question or solve the problem. The data collection process needs to be well-designed to ensure accuracy and representativeness.
4. Analysis: Here is where the real crunching happens. At this stage, the data is analyzed using the appropriate techniques. Prior to the formal data analysis, data exploration methods such as creating graphs, calculating summary statistics, and identifying patterns or trends are used to gain insights and understand the nature of the data.

5. Conclusion: After the analysis of the data, meaningful conclusions and insights are drawn. Based on the findings, recommendations or decisions that address the initial conceptualized problem are made.

It is important to highlight that the PPDAC cycle is not just a one-time linear process. It is often iterative, in the sense that we might need to cycle through the steps a couple of times as we refine our understanding and insights. This iterative approach allows us to continuously improve our analysis and make more informed decisions.

CHAPTER TWO:

VARIATIONS AND UNCERTAINTIES

2.0 INTRODUCTION

In this chapter, we will discuss the concepts of variations and uncertainties as integral parts of human existence. Also in this chapter, we will answer the questions on who created the variations and uncertainties and who is the father of statistics. Furthermore, we will discuss the applications of statistics in virtually every field of study hereinafter called the conscious applications of statistics and discuss some of the unconscious applications of statistics.

2.1 VARIATIONS AND UNCERTAINTIES; AN INTEGRAL PART OF HUMAN LIFE

Variations and uncertainties are fundamental aspects of the human nature and the natural world. It is within this inevitable dynamic interplay of uncertainties and variations in every aspect of human life that statistics finds its purpose and power and that makes it an indispensable discipline in virtually every field of study.

2.1.1 VARIATIONS

Variations encompass the differences, fluctuations, or deviations from an expected or average value in a system or data set and they can be categorized into two types, namely random and systematic variations.

Random Variation: These are variations resulting from inherent randomness and unpredictability in a system. They are chance variations and are uncontrollable, for example, natural variations in weather patterns represent random variations.

Systematic Variation: These are variations resulting from specific, identifiable factors and often follow a discernible pattern. These types of variations are identifiable and rectifiable, for example, variations in a company's product due to use of a bad raw material**.**

2.1.2 UNCERTAINTIES

Uncertainty refers to the lack of exactness in our measurements or estimations or predicting the outcome/occurrence of an event or a physical phenomenon. Two categories of uncertainty are aleatory and epistemic uncertainties.

Aleatory Uncertainty: This type of uncertainty is inherent to the system and arises from natural variability. For example, in the financial market, stock prices fluctuate due to random market events, making them inherently uncertain.

Epistemic Uncertainty: Epistemic uncertainty, on the other hand, stems from a lack of knowledge or incomplete information. In healthcare, a patient's prognosis may be uncertain due to limited medical data.

Variations and uncertainties abound in every aspect of life, as listed below:

1. Weather patterns: Temperature, humidity, precipitation and wind speed can vary significantly from day to day and predicting these variables with exactness is impossible due to the complex interactions of different atmospheric factors.
2. Performance in sports: The performance of athletes from one match or game to another are not exactly the same due to some factors like player's health, team strategies and state of mind that can influence the outcome, introducing both variation and uncertainty.

3. Natural disasters: Events like earthquakes, hurricanes and volcanic eruptions showcase significant variations in their magnitudes and impacts. Predicting the exact timing and intensity of such events can be uncertain due to the intricate interactions of geological and meteorological processes.
4. Medical research: In clinical trials and medical research, variations in patient responses are common due to differences in genetics, lifestyle and other factors. Uncertainty arises when extrapolating the study results to a broader population.
5. Agricultural yields: Crop yield can vary from year to year due to changing weather patterns, pests, diseases and soil conditions and, as a result, farmers are confronted the uncertainties of predicting harvest yield.

The list of examples to elucidate the concept of variation and uncertainty in every aspect of human life is endless. However, the big question is: "Who created these variations and uncertainties in human existence?"

2.2 WHO CREATED VARIATIONS AND UNCERTAINTIES?

We would answer the pertinent question of this lecture under three sub-sections, as follows: the creator of variations/diversity; creator of uncertainty; and God, the Father of Statistics.

2.2.1: THE CREATOR OF VARIATIONS/DIVERSITIES

Unsurprisingly, God designed variety into humanity when God chose to make Adam and Eve in His own image. The first diversity among humans was the dichotomy of male and female. But the diversity God designed for humanity goes far beyond gender. People belong to different tribes, nations, and genetic-ethnic groups. These varied backgrounds are defined and

recognized by myriads of details, including language and appearance.

God did not limit His provision of diversity to Adam and Eve's extended family. God programmed variety into animals, plants and even tiny microorganisms. Also, non-living components of God's creation display geophysical diversity (mountains, valleys, and sand dunes) and even astrophysical diversity (planets and pulsars, supernovas and spiral galaxies). God's vast display of diversity in creation demonstrates how much He prizes variety.

The following passages from the Bible lend credence to this assertion:
Genesis 1:11-12: "Then God said, 'Let the earth sprout vegetation, plants yielding seed and fruit trees on earth bearing fruit after their kind with seed in them' and it was so"
Genesis 1: 24-25: "Then God said, 'Let the earth bring forth living creatures after their kind: cattle and creeping things and beasts of the earth after their kind' and it was so".
There is no question that Nature displays differences in details of diverse animals, plants, microorganisms, earth's geophysical environment, and even the innumerable galaxies of outer space.

2.2.2 THE CREATOR OF UNCERTAINTIES

Let me refer to the Bible once again:
Proverbs 16:9: "The heart of man plans his way, but the Lord establishes his steps." This verse highlights the tension between human planning and God's ultimate control. It suggests that while we may make plans, there is an element of uncertainty in whether those plans will come to fruition as expected.

James 4:13-15: "Come now, you who say, 'Today or tomorrow we will go into such and such a town and spend a year there and trade and make a profit' - yet you do not know what tomorrow will

bring." This passage underscores the uncertainty of the future and our limited understanding of what lies ahead, despite our intentions and plans.

Matthew 6:34: "Therefore do not be anxious about tomorrow, for tomorrow will be anxious for itself". This verse encourages us to focus on the present moment rather than worrying about the uncertainties of the future, implying that there will always be an element of the unknown.

From these passages, one could argue that God's creation of uncertainty serves several purposes:
Testing and Growth: Uncertainty challenges us to trust in God's plan and develop qualities like faith, patience, and resilience.
Reliance on God: When we face uncertainty, we are reminded of our need for God's guidance and wisdom.
Humility: Recognizing that we don't have all the answers fosters humility and a deeper connection to God.
Appreciation of the Present: Uncertainty encourages us to live in the moment and appreciate the blessings we have today.

2.3.3 GOD, THE FATHER OF STATISTICS

God's proclamation in Genesis (1:14), "Let there be lights in the firmament of the heaven to divide the day from the night; and let them be for signs and for seasons, and for days, and years", signified that God is the composer of time and seasons. Therefore, the concept of decision making based on seasonal analysis, an aspect of statistics under time series analysis and forecasting, is divine as evidenced by the following two examples:

(i) The decision by birds of the air to migrate at a particular time of the year is based on seasonal analysis (an aspect of statistical decision-making). Here the birds did not attend any school of statistics, nor did they do any programme in

statistics, to learn the act of decision-making based on seasonal analysis but by the act of creation.

(ii) Similarly, the shedding of leaves or reduction of leaves' sizes by tress during dry seasons or period of water scarcity to minimize water loss through transpiration is a form of decision making by the trees based on seasonal analysis that was not learnt in any school of Statistics but by the act of creation.

God, the Father of Statistics, instructed Noah in the Bible to carry out an aspect of statistics called Sampling when he wanted to populate his ark (Akpanta, 2023). As reported in the Bible:

"1.The Lord said to Noah, 'Go into the boat with your whole family; I have found that you are the only one in all the world who does what is right. 2. Take with you seven pairs of each kind of ritually clean animal, but only one pair of each kind of unclean animal. 3. Take also seven pairs of each kind of bird. Do this so that every kind of animal and bird will be kept alive to reproduce again on the earth'". (Genesis 7:1-3)

Furthermore, God, the Father of Statistics, instructed Moses to conduct two censuses (an aspect of statistics) for the Israelites' in the 13th century BCE, before their entry into the Promised Land.

Deductions

1. God, the Creator, or the Divine, has ordained the universe's order and complexity. Variations in life, including the diversity of species, cultures, human experiences and variations in biodiversity, are manifestations of divine creativity.

2. Uncertainties and variations are part of God's design for a purpose, which may include personal growth, moral development, reliance on God, fostering deeper connection to God or appreciation of our present blessings.

3. The concept of God giving humanity charge or dominion over the Earth is well established in Bible. For example, in Genesis 1:26-28 (New International Version):

"Then God said 'Let us make mankind in our image, in our likeness, so that they may rule over the fish in the sea and the birds in the sky, over the livestock and all the wild animals, and over all the creatures that move along the ground.'"

4. God is the father of Statistics

Based on the above deductions (1-4) that attest to the fact that God is the creator of variations and uncertainties and that He gave humanity stewardship over the earth which involves grappling with the variations and uncertainties of creation, the emergence of statistics as a discipline is a divine gift, an intellectual and practical tool provided to humanity to navigate the variations and uncertainties inherent in creation. Thus, we can logically with confidence say that God is the creator of the Discipline of "Statistics".

2.4 STATISTICS, A SUPER DISCIPLINE THAT HAS APPLICATIONS IN VIRTUALLY EVERY FIELD OF STUDY

Statistics is a super discipline that provides tools and techniques that can be applied across various domains of study. From science to business, sports to medicine, and beyond, statistics plays a crucial role in shaping our understanding of the world.

Think about it like this: we are constantly bombarded with information and data from all directions. Statistics gives us the tools to sift through this sea of data and extract meaningful insights. It is like having a superpower that helps us distinguish between what is relevant and what is noise.

Here are some examples of applications of statistics across various frontiers of knowledge which I describe as conscious applications of statistics:

1. Actuarial science is the discipline that applies mathematical and statistical methods to assess risk in the insurance and finance industries.
2. Astrostatistics is the discipline that applies statistical analysis to the understanding of astronomical data.
3. Biostatistics is a branch of biology that studies biological phenomena and observations by means of statistical analysis.
4. Business analytics is a rapidly developing business process that applies statistical methods to data sets (often very large) for the purpose of understanding business performances and opportunities.
5. Chemometrics is the science of relating measurements made on a chemical system or process to the state of the system via application of mathematical or statistical methods.
6. Demography is the statistical study of all populations. It can be a very general science that can be applied to any kind of dynamic population, that is, one that changes over time or space.
7. Econometrics is a branch of economics that applies statistical methods to the empirical study of economic theories and relationships.
8. Environmental statistics is the application of statistical methods to environmental science. Weather, climate, air and water quality are included, as are studies of plant and animal populations.
9. Forensic statistics is the application of probability models and statistical techniques to scientific evidence, such as DNA evidence, and the law. In contrast to "everyday" statistics, to not engender bias or unduly draw conclusions, forensic statisticians report likelihoods as likelihood ratios (LR).

10. Geostatistics is a branch of geography that deals with the analysis of data from disciplines such as petroleum geology, hydrogeology, hydrology, meteorology, oceanography, geochemistry, and geography.
11. Jurimetrics is the application of probability and statistics to law.
12. Operations research (or operational research) is an interdisciplinary branch of applied mathematics and formal science that uses methods such as mathematical modeling, statistics, and algorithms to arrive at optimal or near optimal solutions to complex problems. Management science focuses on problems in the business world.
13. Quality control reviews the factors involved in manufacturing and production. It can make use of statistical sampling of product items to aid decisions in process control or in accepting deliveries.
14. Quantitative psychology is the science of statistically explaining and changing mental processes and behaviors in humans.
16. Statistical mechanics is the application of probability theory, which ranges from mathematical tools for dealing with large populations, to the field of mechanics, which is concerned with the motion of particles or objects when subjected to a force.
16. Statistical physics is one of the fundamental theories of physics, and uses methods of probability theory in solving physical problems.
17. Statistical signal processing utilizes the statistical properties of signals to perform signal processing tasks.
18. Statistical thermodynamics is the study of the microscopic behaviors of thermodynamic systems using probability theory and provides a molecular level interpretation of thermodynamic quantities such as work, heat, free energy, and entropy.

2.5 STATISTICS, A UNIQUE FIELD OF STUDY THAT IS UNCONSCIOUSLY APPLIED IN OUR EVERYDAY LIVES

It is incontrovertible that there are situations in our day-to-day activities where the tools of statistics are unconsciously or unknowingly applied. The list of scenarios where we unconsciously or unknowingly apply the concepts of statistics abound in our everyday lives but for the limitations of this lecture, we give the following list:

- The security man at a bank premises will always shout to customers at an ATM point: "Please be on queue and you would be served on the basis of first-come-first –served" (FCFS). Here the security man unknowingly applying the concept of queuing theory where the service discipline is FCFS—a statistical method.
- A bus driver from Lagos to Port Harcourt would want to know those passengers that would disembark on the way before the last bust stop and thus load their luggage last in the bus's luggage compartment because they will access their luggage sooner than others. This is also another practice of statistical concept in queuing theory where the service discipline is Last-in-First-out (LIFO).
- A Doctor does not need to drain the entire blood of a patient or take blood samples from every part of the patient's body to examine him or her for malaria but only takes a sample from a particular part of the body for this purpose. This is an act of inferring for the whole based on sample result—a practice of statistics.
- The housewife who wants to buy some cups of rice at a village market examines few grains of rice on her palm from different bags of rice in order to decide from which bag of rice to buy from. Here she practices acceptance sampling which is a tool in statistical quality control.

- Courtship in a marriage setting is a statistical study where both people who plan to spend their lives together have the opportunity to study themselves and assess whether they would be able to manage their inherent variations as different individuals and make informed decisions about marriage.
- A cook does not need to drain the whole pot of soup to determine the salt content of the soup. After adding salt, she only stirs and tastes a drop on her palm.
- The rural farmer's decision to either engage in early or late planting is based on seasonal analysis of the climatic conditions. Here the farmer applies the concept of decision making based on seasonal time series analysis, a statistical tool under time series analysis and forecasting.

CHAPTER THREE:

POWER OF STATISTICS IN DECISION MAKING

3.0 INTRODUCTION

In this chapter, we will use numerical examples to illustrate the power of statistics in decision-making in the midst of variations and uncertainties. We would also discuss the fact that statistics can lead to incorrect and misguided decisions when the data are flawed or compromised or biased. Furthermore, we would also discuss abuses and misuses of statistics and finally discuss the concept of outliers.

3.1 THE POWER OF STATISTICS IN DECISION MAKING UNDER POTENTIAL VARIATIONS AND UNCERTAINTIES

Here, we would just use two numerical examples to illustrate the power of statistics in decision making under potential variations and uncertainties in sections 3.1.1 and 3.1.2, respectively.

3.1.1 THE POWER OF STATISTICS IN DECISION-MAKING UNDER VARIATIONS

Let us use the following hypothetical example to illustrate the power of statistics in decision-making in the midst of variations. Suppose a restaurant owner in Otuoke named John is trying to optimize his daily special menu to maximize profit while dealing with fluctuations in customers' demand.

Step 1: Data Collection

Suppose John has the following data on the daily revenue from offering two different specials: Garri with Banga Soup (GWBS) and Garri with Afang Soup (GWAS) over the course of twenty (20) days:

GWBS
Daily Revenue (₦): 15000, 18000, 16000, 14000, 17000, 14500, 15500, 16500, 15500, 17500, 140000, 18500, 16000, 17000, 15500, 17500, 14500, 16000, 16500, 15000

GWAS
Daily Revenue (₦): 13000, 14000, 12000, 12500, 13500, 14000, 15000, 12000, 15500, 125000, 13000, 14000, 13500, 13000, 12000, 14500, 14000, 13000, 15500, 15000

Step 2: Descriptive Statistics
John starts by calculating basic descriptive statistics for both GWBS and GWAS, namely, the mean (average) and standard deviation to understand the central tendency and variability/dispersion in daily revenues. The results are given in Table 3.1.

Table 3.1: Means and Standard deviations of Daily Revenue of GWBS and GWAS

Daily Special Menu	Mean	Standard Deviation	Coefficient of Variation
GWBS	22300	27731	124.35
GWAS	13575	1127	8.30

Step 3: Decision Analysis
Now, John must decide which special to offer regularly. GWBS has a higher mean revenue, but it also has a higher standard deviation, indicating more variations in daily revenues. GWAS has a lower mean revenue but with less variation.

To make an informed decision, John can use further analysis. One approach is to calculate the coefficient of variation (CV), which measures the relative variability of each special by dividing the standard deviation by the mean:

For GWBS:
CV = (27731 / 22300) *100 = 124.35

For GWAS:
CV = (1127 / 13575) *100 = 8.30
The coefficient of variation helps John assess the risk associated with each special. A lower CV indicates lower relative variability, making GWAS a less risky choice in terms of daily revenue.

Step 4: Decision
Based on the statistical analysis, John can now make an informed decision. If he prioritizes consistency and wants to minimize the risk of low revenue days, GWAS is the better choice. While it has a lower mean revenue, it offers more predictable outcomes, making it a safer bet for steady daily operations.

3.1.2 THE POWER OF STATISTICS IN DECISION-MAKING UNDER UNCERTAINTY

Let us illustrate the power of statistics in decision-making under uncertainty with a numerical example involving a financial investment decision. Suppose Mr. Frank is contemplating whether to invest in two different stocks, Stock A and Stock B. Both stocks have shown varying historical returns, and Frank needs to make an informed decision considering the uncertainty in their future returns.

First step: Historical Data Analysis
Suppose that Frank gathers historical data on the annual returns of Stock A and Stock B over the last six years as given below in Table 3.2.

Table 3.2: Annual Stock Returns for Stock A and Stock B for a period of Six(6) Years

Year	Stock A	Stock B
Year 1	12%	9%
Year 2	15%	10%
Year 3	10%	13%
Year 4	8%	7%
Year 5	14%	12%
Year 6	8%	11%

Second Step: Calculating Measures of Central Tendency

Here Frank calculates the mean and standard deviations of the stocks. The mean represents the expected return, while the standard deviation measures the volatility or risk associated with each stock.

Table 3.3: Mean Return (μ) and Standard Deviation (σ) of the Annual Stock Returns

Stock	Mean Return (μ)	Standard Deviation (σ)
Stock A	11.17%	10.33%
Stock B	2.99%	2.16%

Third Step: Decision Analysis

Now Frank wants to make an investment decision considering the uncertainty associated with each stock. He has a certain amount of money to invest, and he needs to decide whether to put it all in one stock or diversify his investment. Suppose the investment options are given in Table 3.4

Table 3.4: Investment Options

	Investment option	
	Invest all the money in stock A	Invest all the money in stock A
Expected Return	11.170%	10.333%
Risk (σ)	2.99%	2.16%
Sharpe Ration (SR)	3.07	3.86

Fourth Step: Decision

Now, Frank can make an informed decision based on his risk tolerance and return expectations by calculating the Sharpe Ratio (Andrew(2003)) given as:

$$\textbf{Sharpe - Ratio (SR)} = \frac{\mu_r - R_f}{\sigma_r} \qquad (3.1)$$

where: μ_r is the mean return, R_f is the risk-free rate and σ is the standard deviation of the return data set.

The **SR** values, assuming a risk-free rate of 2%, are also given in Table 3.4 above.

The higher Sharpe Ratio indicates a better risk-adjusted return. In this case, Stock B has a higher Sharpe Ratio (3.86) compared to Stock A (3.07). This means that, for the given level of risk (as measured by standard deviation), Stock B has provided better offer of investment after adjusting for the risk-free rate. Therefore, based on the statistical analysis and the Sharpe ratio, Frank may choose to invest in Stock B, as it appears to offer a better trade-off between risk and return compared to Stock A.

This numerical example illustrates how statistics, through measures like mean, standard deviation and Sharpe-Ratio, enabled Frank to quantify uncertainty and make a rational investment decision. It allows him to weigh the potential returns against the

associated risks, aligning his investment strategy with his preferences and goals.

3.2 STATISTICS, A TOOL THAT CAN LEAD TO MISGUIDED DECISIONS

Evidence-based decision making is a logical and systematic approach to decision-making by relying on reliable and verifiable facts, data and information rather than personal opinions, assumptions or intuition. By leveraging statistics in decision-making, the decision maker would move beyond intuition, personal biases and subjective judgments to decisions that are grounded in evidence.

In as much as Statistics provides us with a tool for data-driven or evidence-based decisions, there is no question that inaccurate, incomplete, misleading, flawed or compromised data would lead to incorrect conclusions and misguided decisions. It is therefore crucial to validate and critically assess the data or evidence used for making decisions to avoid such pitfalls. The case of the stolen Esau's birthright by Jacob as contained in Genesis 27:15-16 is a case where compromised evidence led to a misguided decision. There, Isaac the father of Esau and Jacob, was presented with flawed or compromised evidence where Rebekah, his wife, dressed Jacob with Esau's clothes and covered his hands and neck with goat's skins to mimic Esau's hairy body and that led Isaac to taking a wrong decision of blessing Jacob as the first son instead of Esau. What follows are some areas of human endeavour where flawed data or evidences are tantamount to making dangerous or harmful decisions.

Financial Investments: Flawed data in the financial sector can lead to poor investment decisions. Imagine a scenario where the report on the financial health of a company is inaccurate or manipulated and the report is used by investors to make decisions.

Since the report does not reflect the true financial health of a company, investors might pour money into a failing company resulting in significant losses.

Healthcare Treatment Decisions: In healthcare, using flawed data for medical treatment decisions can have serious consequences. Suppose a doctor relies on outdated or inaccurate patient records when prescribing medication. This could lead to inappropriate dosages or the wrong treatment, potentially worsening the patient's condition or may even lead to death.

Educational Policy: Flawed data in educational assessments can impact policy decisions. For example, it could lead to misguided policies, such as allocating resources to the wrong areas or implementing ineffective teaching methods.

Criminal Justice: Flawed data in criminal justice systems can lead to wrongful convictions. If unreliable forensic evidence or biased data analysis is used, innocent individuals could be wrongly accused and imprisoned or may even face a death penalty as the case may be.

3.3 MISUSES AND ABUSES OF STATISTICS

There is no question that statistics is a tool for informed decision-making. However, there are cases where they are misused or abused. Misuse of statistics refers to the incorrect, inappropriate, or misleading application of statistical methods and principles, often resulting in flawed conclusions or misinterpretations. This can occur due to misunderstanding, negligence, deliberate manipulation, or a lack of statistical expertise. On the other hand, abuse of the use of statistics refers to the unethical or improper manipulation, misrepresentation, or misinterpretation of statistical methods and data for personal, ideological, or deceptive purposes. It involves distorting statistical information to support a particular interest, promote false narratives, or deceive others. Common misuses and abuses of statistics include:

1. **Misleading with averages:** Representing data with averages (e.g., mean, median) without considering the underlying distribution can give a false impression of what the data truly represents. For example, presenting the average income in a wealthy neighborhood as if it applies to all residents is misleading as it would not reflect the average income of the poor residents.
2. **Cherry-picking data:** This means selectively choosing data points or time periods that support a specific argument while ignoring or omitting data that contradicts it.
3. **Correlation vs. causation:** Correlation is a measure of the statistical relationship between two variables, while causation implies a direct cause-and-effect relationship. Therefore, inferring causation from correlation without sufficient evidence may be misleading
4. **Sample bias:** This means drawing conclusions from a non-representative sample that does not accurately reflect the population of interest, which often leads to incorrect generalizations.
5. **Data manipulation:** This is done by falsifying or altering data to achieve a desired outcome. This includes fabricating data or selectively removing outliers to fit a specific narrative.
6. **Overgeneralization:** Overgeneralization occurs through applying findings from a small or specific group to a larger and more diverse population without proper validation.
7. **Inadequate sample size:** This is done by drawing conclusions from a sample that is too small to provide statistically significant results, leading to unreliable findings.
8. **P-value -hacking:** P-value hacking, also known as p-hacking or data dredging, is often an unethical practice in statistics and data analysis that undermines the integrity of the scientific method. It involves the manipulation of data or the analytical process to achieve a statistically significant result when conducting statistical hypothesis tests. The primary motivation

behind p-value hacking is to present results that are more favorable or interesting, typically for publication or presentation, but it can lead to erroneous and misleading conclusions. It can be done through the following ways

(i) Testing multiple hypotheses without a specific hypothesis in mind.

(ii) Selective reporting where only reports of the results that are statistically significant are given while other results that are not statistically significant may be ignored or downplayed.

(iii) Data manipulation where researchers may engage in practices such as excluding outliers, transforming variables, or applying various statistical techniques, until they obtain the desired statistically significant result.

9. **Data snooping:** This means analyzing data exhaustively until a pattern or correlation is found, even if it is purely coincidental, and then claiming it as a significant discovery.
10. **Hasty generalizations:** This is the drawing of sweeping conclusions based on insufficient evidence or anecdotal experiences.
11. **Biased surveys.** This is the use of sampling techniques that are biased towards a particular response.
12. **Confirmation bias:** This is the seeking or interpreting of data in a way that confirms preexisting beliefs or biases.
13. **Over-fitting:** This is done by creating a model that fits the training data extremely closely, but fails to generalize well to new, unseen data. This can result in overly complex models that perform poorly on new observations.
14. **Lying with Statistics:** This is intentionally using statistics to deceive or mislead the audience by presenting false or distorted information, often through carefully crafted wording or framing.

Misusing and abusing statistics is an unhealthy practice that undermines the credibility of scientific research. It is important to promote ethical and responsible use of statistical methods, transparent reporting of findings, and critical evaluation of statistical claims to combat such abuses.

3.4 OUTLIERS AND THEIR IMPACT ON DATA DRIVEN DECISION MAKING

Outliers are data points that significantly differ from the majority of the data in a data set. They are observations that are either unusually high or low in value compared to the rest of the data. The presence of an outlier can cast doubt on the integrity of the data and significantly reduce its quality by distorting statistical measures and potentially leading to incorrect decisions.

Let us use the following hypothetical numerical example to illustrate the effect of an outlier on statistical analysis. Suppose there are two pastors (Pastor A and Pastor B) in a local church and suppose that Pastor A handled the church for the first ten weeks while Pastor B handled the church in the second ten weeks. Furthermore, suppose their weekly returns for the ten-week period are given in Table 3.5.

Table 3.5: Weekly Returns of Pastor A and Pastor B for the first and second ten-week period

First ten weeks	Returns in Thousands of Naira by Pastor A	Second Ten weeks	Returns in Thousands of Naira by Pastor B
Week 1	14466	Week 1	14466
Week 2	15058	Week 2	15058
Week 3	16452	Week 3	16452
Week 4	16242	Week 4	16242
Week 5	15771	Week 5	15771
Week 6	14400	Week 6	14400

Week 7	16398	Week 7	16398
Week 8	14066	Week 8	14066
Week 9	15058	Week 9	15058
Week 10	14784	Week 10	1950

It is important to note in Table 3.5 that the returns for both pastors are the same except in week 10 for Pastor B where there was the presence of an outlier. In what follows, we analyze the effect of the outlier by given the summary statistics of Table 3.5 in Table 3.6.

Table 3.6: Summary Statistics of the weekly Returns by Pastor A and Pastor B in a Local Church

Name	Average Weekly Return	Standard Deviation (A Measure of spread)	Minimum Return	Median
Pastor A	₦15270	±₦885	₦4066	₦15058
Pastor B	₦13986	±₦4317	₦1950	₦15058

The introduction of the outlier significantly skews the mean, making the Average/Mean Return by Pastor B to be lower than the typical weekly return for the ten-week period of overseeing the Church. The mean is very sensitive to outliers.

Standard deviation (A measure of spread)
The standard deviation measures the spread of the data about the central tendency.

Based on the summary statistics in Table 3.6, it is clear that the presence of the outlier made the standard deviation to increase drastically from ±₦885 to ±₦4317, reflecting the increased variability in the weekly return data set from Pastor B.

Median (Middle Value)
The middle value in the ordered weekly return data set for both Pastors remain the same at ₦15058 because the Median is relatively robust to outliers and is not always directly affected by outliers.

In summary, not only that the presence of the outlier has subjected the integrity of Pastor B to scrutiny, it significantly skewed the mean and standard deviation. In conclusion, outliers have the potentials to distort statistical analysis, affecting the accuracy and reliability of summary statistics like the mean and variance and thereby leading to misleading conclusions about the data. This is why statisticians will always warn you to identify and handle outliers appropriately, depending on the context of the analysis.

CHAPTER FOUR:

STATISTICS, A TOOL FOR GOOD GOVERNANCE AND NATIONAL DEVELOPMENT

4.0 INTRODUCTION

This chapter contains two sections: Section 4.1 describes statistics as tool for good governance and national development while section 4.2 describes the motive behind the establishment of the Federal Office of Statistics/National Bureau of Statistics.

4.1 STATISTICS, A TOOL FOR GOOD GOVERNANCE AND NATIONAL DEVELOPMENT

In the context of governance and national development, the role of statistics cannot be overstated. Accurate and timely data collection, analysis, and interpretation provide the necessary insights to drive policies, strategies, planning and resource allocations. The tools of statistics can be applied for good governance and national development in the following areas:

Data-Driven Policy Formulation: Statistical tools can be used by governments to analyze demographic, economic, and social data to formulate policies that address specific challenges. For instance, analyzing unemployment rates, income distribution, and education levels can lead to targeted policies to alleviate poverty, reduce illiteracy level and promote job growth.

Budget Allocation and Resource Management: Statistical analysis helps governments allocate budgets effectively by identifying areas of high demand and priority. For instance, analyzing healthcare data can guide resource allocation to improve medical facilities and services where they are most needed.

Environmental Management: Statistical tools are vital for monitoring environmental indicators like air and water quality, deforestation rates, and greenhouse gas emissions. Governments can use this data to develop and implement sustainable environmental policies.

Public Health Planning: Analyzing health-related data can help governments identify disease trends, vaccination coverage, and healthcare needs. This informs public health campaigns, resource allocation, and the development of healthcare infrastructure.

Urban Planning and Development: Statistical analysis of population growth, housing demand, and transportation patterns can assist governments in planning sustainable urban development, managing traffic congestion, and ensuring access to essential services.

In conclusion, statistics are not just a collection of numbers; they are the foundation of good governance and national development and therefore are indispensable tools for leaders striving to build better and more prosperous nations. To lead a nation without the guidance of statistics is an endeavor fraught with uncertainty and risk, where success becomes elusive, and the path to failure is all too clear. In the modern age, a nation not relying on statistics for planning is indeed planning to fail.

4.2 ESTABLISHMENT OF NATIONAL BUREAU OF STATISTICS: A RECOGNITION OF THE POWER OF STATISTICS FOR GOVERNANCE AND NATIONAL DEVELOPMENT

It is in recognition of the indispensability of Statistics in National development that the Federal Office of Statistics (FOS) that was later transformed to the National Bureau of Statistics (NBS) was established. NBS emerged from the merger of the Federal Office

of Statistics (FOS) and the National Data Bank (NDB) as part of the implementation of the Nigerian Statistical Master Plan (SMP) which was originally designed to span five years, 2005-2009 (Olubusoye et., al, (2015)). The FOS was transformed into the Nigerian Bureau of Statistics (NBS) under the Statistics Act, No. 18 of 2007. According to Akinyosoye (2006), the goal of the entire reform programme for NBS was to transform the Agency into a world class National Statistical Office (NSO), capable of producing adequate, highly quality and timely data relevant to demands of users in government, the universities and research institutes, private sector organizations and international agencies. Specifically, key responsibilities of NBS include:

(i) Data Collection: The NBS is responsible for data collection on various economic, social, and demographic aspects of the country, such as population, inflation, employment, and trade.
(ii) Data Analysis: It analyzes the collected data to generate statistics and reports that provide valuable insights into the economic and social conditions of Nigeria.
(iii) Economic Research: It conducts research and economic studies to support government policy making and decision-making.
(iv) National Accounts: It compiles and publishes national accounts, including Gross Domestic Product (GDP) figures, to measure the economic performance of Nigeria.
(v) Monitoring and Reporting: NBS monitors key economic indicators and reports on them regularly, which is essential for tracking progress and making informed policy decisions.
(vi) Providing Data to Stakeholders: It provides statistical data to various stakeholders, including government agencies, policymakers, researchers, businesses, and the public.

CHAPTER FIVE:

MY HUMBLE CONTRIBUTION TO KNOWLEDGE

Having published over 55 journal articles in the field of statistics, with over 97% of them in international journals, and with over 700 citations in Google scholar, I believe that I have made a reasonable contribution to knowledge in various areas of statistics, namely: time series analysis and forecasting, statistical modeling, data transformations, and distribution theory. Below is a list of selected research articles under selected fields of statistics:

1. Time Series Analysis and Forecasting

Time series analysis is a statistical and mathematical technique used to analyze and interpret data that is collected sequentially in time, usually at equally spaced time intervals

s/n	Authors	Title	Contribution to knowledge	Journal
(a)	Iwueze I. S and **Ohakwe J (2004)**	Buys – Ballot Estimates when Stochastic Trend is Quadratic.	We extended the Buys-Ballot estimation procedure developed by Iwueze and Nwogu (2004) for time series decomposition when trend-cyle component is linear to cases where the trend-cycle component is quadratic. Estimates were derived for the additive and multiplicative models. Two estimation methods namely chain-based estimation (CBE) and fixed based estimation (FBE) methods were established estimates. Our methods of estimation proved to be very effective when compared to least squares estimation method using simulated example.	Journal of the Nigerian Association of Mathematical Physics, Vol. 8: pp 311 – 318, 2004.

(b)	**Ohakwe J** and Iheanyi S. Iwueze (2009)	Characterizations of the Moments of the Purely Diagonal Bilinear Time Series Model of Order One,	Considering the similarity in the covariance structure of the purely diagonal bilinear model of order one (PDB(1)) and the linear moving average of order one (MA(1)) which can lead to misclassificationof one to the other, we therefore studied the dissimilarities between the two competing models. We established that the well-known similarity between the competing models is true only when the autocorrelation function at lag one(the tool of model identification)lies between zero and 0.16. At the region of similarity, we proposed use of fourth moments for distinction where the PDB(1) identifies as simple Autoregressive moving average model (ARMA(1,1)) while that of the MA(1) still identifies as an MA(1) model.	Interstat Journal of Statistics: No. 2, June 2009

(c)	Ohakwe J and Iheanyi S. Iwueze (2009)	Penalties For Misclassification of Bilinear And Linear Time Series Processes.	As a result of the possible misclassification in (paper No.b), we constructed a test of significance for testing bi-linearity. The test was named after the Authors (Ohakwe and Iwueze) and is called OI Statistics.	Interstat Journal of Statistics: No. 3, June 2009
(d)	I. S. Iwueze and **Ohakwe J. (2009)**	Covariance Analysis of the Squares of the Diagonal Completely Bilinear Time Series Models	Here, we Generalized the results of Paper No. b above to general non-linear purely diagonal bilinear the general linear moving average process.	Brazilian Journal of Probability and Statistics, Vol. 25, No. 1, 90 – 98, 2009

2. Statistical modeling **Statistical modeling provides a framework for making data driven decisions, identifying patterns and relationships, and quantifying uncertainty.**				
(e)	Innocent.C. Nnorom, **Ohakwe Johnson** and Oladele Osibanjo (2009)	Survey of Willingness of residents to participate in electronic waste recycling in Nigeria – A case study of mobile phone recycling	In this survey, we assessed the willingness of Consumers to pay for environmental friendlier phone (the so- called green phone). Here we used principal component analysis with varimax rotation to condense the survey items into smaller set of factors. The analysis showed that significant predictors of willingness to pay (WTP) extra for green electronics include awareness and concern about the deteriorating environment, age and general attitude towards to environment. **Award: No 15 on the list of Top 25 Hottest Articles, October – December, 2009 (Science Direct TOP 25**	Journal of Cleaner Production, Elsevier: 17 (2009), pp.1629 – 1637.

			Hottest Articles)	
(f)	Ogbonnaya C. Ogbonnaya, Diljit Singh and **Ohakwe Johnson (2011)**	Significant Predictors of Library Use by New International Students: the case study of University of Malaya, Kualla Lumpur, Malaysia.	In this study, we investigated the significant factors affecting the use of library by international students in the University of Malaya, Malaysia. The test items extracted were extracted from the literature reviews. Using ordered probit model, significant predictors were found to be programme of study, English language proficiency, computer-use skills and previous library experience. **Collaborators:** Ogbonnaya C. Ogbonnaya and Professor Diljit Singh, Faculty of Library and Computer Sciences, Institute of postgraduate studies, University of Malaya, Kualla Lumpur, Malaysia	Asian Journal of Applied Sciences 4 (1): 53 – 62, 2011.
3. Data Transformations				

Data transformation is a fundamental data preprocessing technique used in data analysis, statistics, and mathematical sciences. It is an application of a mathematical function aimed at altering the format, structure, or values of the data set to make it more suitable for analysis, modeling, or visualization. Data transformation can help improve the quality of the data, remove noise or anomalies, and make it more compatible with the assumptions of certain analytical methods. Reasons for data transformation include stabilizing variance, normalization, reducing the effect of outliers, making a measurement scale more meaningful and to linearize a relationship (J. Ohakwe et., al (2013). A successful transformation is achieved when the desirable properties of a data set remain unchanged after the transformation(Ohakwe et., al (2012). Examples are the logarithm, inverse, square –root and so on.				
(g)	Otuony E. L, I.S. Iwueze and **OhakweJ.(2011)**	The Effect of Square Root Transformation on the Error Component of the Multiplicative Time Series Model.	In time series modeling using multiplicative model whose error component (EC) is the Left-Truncated Normal Distribution (LTND), we established the distribution and properties of the EC under inverse transformation Furthermore, we established that the mean of inverse-transformed EC = 1.0 while its variance is 4times the variance of the untransformed EC for $0 < SD \leq 0.59$, where SD is the standard deviation of the untransformed EC.	International Journal of Statistics and Systems. Vol. 6, No. 4, 2011, pp. 461–476.ISSN: 0973 – 2675
(h)	Iheanyi S. Iwueze	Square Root	Here we determined the	Asian Journal of Mathematics

	and **Johnson Ohakwe(2011)**	Transformation of the Quadratic Equation.	necessary and sufficient conditions on the coefficients of quadratic equation such that its square root transformation remains approximately a quadratic equation or exactly a linear equation. Furthermore, numerical examples were used to illustrate the results obtained.	and Statistics.4 (4): 186 – 199, 2011. ISSN 1194 – 5418. DOI: 10.3923/ajms.2011.186.199
(i)	**Ohakwe J**, Akpanta A. C. and Dike O. A.	Unit Mean and Constant Variance of the Generalized Gamma Distributionafter Square Root Transformation in Statistical Modeling	Here we studied the effect of square root transformation on the error component of the multiplicative error model whose distribution belongs to the generalized gamma (GG) family. The purpose of the study was to determine the effect of the said transformation on the basic assumptions: unit mean and constant variance assumption required for modeling. The special cases of the GG considered are the three-parameter Gamma, Chi-	Mathematical Theory and Modeling, Vol. 2, No. 12, 2012, pp. 22 – 34

			square, Exponential, Weibull, Rayleigh and Maxwell distributed error component. From the results, the unit-mean assumption is maintained for all the distributions under study. There were reductions in the variances of the transformed distributions except for those of Gamm(a, b, 1) when a > 1, Rayleigh and Maxwell distributions where the variances increased. Furthermore, we established that square root transformations where applicable for the studied distributions is successful if variance of the transformed data is < 0.5	
4. Probability Distributions: **A probability distribution is a mathematical function that describes the likelihood of various outcomes or events in a random experiment. It assigns probabilities to each possible outcome of a random experiment and the sum of all these probabilities is equal to one.** **Under this area of study, we have introduced several new flexible probability density functions (pdf) into the**				

statistical literature. Here, we derived the probability density functions, their properties and illustrate its flexibility of applications using real-life data.				
(j)	I.E. Okorie, A.C. Akpanta and **Ohakwe J. (2016)**	The Exponentiated Gumbel Type 2 Distribution: Properties and Applications	Here We introduced into the statistical literature a generalized version of Standard Gumbel type-2 probability distribution called Exponentiated Gumbel Type 2 Distribution. Properties of the new distribution were established and the flexibility of the newly introduced distribution were illustrated using real data	International Journal of Mathematics and Mathematical Sciences, Volume 2016, 10 pages. Hindawi Publishing Corporation.
(k)	Idika E. Okorie, Anthony C. Akpanta and **Johnson Ohakwe(2016)**	Transmuted Erlang-Truncated Exponential distribution	Here we introduced a new life-time distribution into the statistics literature, gave its properties and illustrated its flexibility using a real- life data set. **Remark:** Sponsored by the University of Manchester, UK	Econ. Qual. Control, 31 (2), 71 – 84, 2016 DOI:10.1515/eqc-2016-0008 (Published by De Gruyter)
(l)	Idika E. Okorie, Anthony C. Akpanta and **J.**	The modified power function distribution	Here we proposed a new life time distribution that belongs to Marshal-Olkin G- family of	Cogent mathematics, 4(1), 1319592, 2017.

	Ohakwe and D.C. Chikezie (2017)		distributions called the Modified Power Function (MPF) distribution. **Remark:** Sponsored by the University of Manchester, UK	
(m)	I.E. Okorie, A.C. Akpanta, **J. Ohakwe,** D.C. Chikezie and E.O. Obi (2017)	The Kumaraswamy-G-Exponentiated Gumbel Type -2 Distribution	Here we extended the The Exponentiated Gumbel Type 2 Distribution to a wider family of distributions called the Kumaraswamy Generalized Exponentiated Gumbel Type-2 Distribution that has twenty-two (22) sub-class of distributions.	Afrika Statistika, Vol. 12 (3), pp. 1367-1396, 2017.

CHAPTER SIX:
SUGGESTION FOR FUTURE RESEARCH, RECOMMENDATIONS AND CONCLUSION

6.0 INTRODUCTION

This chapter contains suggestion for future research, recommendations and conclusion in sections 6.1, 6.2 and 6.3 respectively.

6.1 SUGGESTION FOR FUTURE RESEARCH

In addition to my future innovative researches in the areas of time series modeling and distribution theory, from my participation in StEP Initiative programmes in Philips Electronics Hitech Campus, Eindhoven, Netherlands and Umicore Materials Company, Brussell, Belgium, I want to engage in a multi-disciplinary research that will involve chemists and engineers in developing an E-waste materials Recovery Plant. With such a plant, we can recover Copper selenium and other precious metals contained in E-waste.

6.2 RECOMMENDATIONS

(i) I recommend the establishment of a University Statistical Consulting Unit (USCU) to be headed by a statistician where both postgraduate and undergraduate students consult for the proper design and analysis of their projects. This is as a result of the fact that most researchers end up collecting a lot of irrelevant data as a result of poor design.

(ii) Teaching of the Service Courses: The various statistics courses taught in other faculties such as Education, Social Sciences, Engineering and Management Sciences of the University must be taught by lecturers from the Department of Statistics so that students are properly taught and well equipped with the knowledge of statistical concepts and practices.

(iii) Considering the application of statistics in every field of study, I recommend the establishment of a University Statistical Laboratory where students across the University would be taught the practical applications of statistical tools as it relates to their field of study. The Laboratory must be equipped with computers installed with the relevant statistical and mathematical soft wares.

(iv) I recommend the deployment of statisticians to the Establishment and Examinations and Records Units of the University who will from time to time mine and interrogate the huge available data for updates and analyses for useful purposes.

6.3 CONCLUSION

Having defined Statistics as the science of variations and uncertainties that systematically examines, quantifies, and analyzes the inherent variability and unpredictability in data generated from observations, experiments, physical and natural phenomena, etc., and having established that variations and uncertainties which are integral aspects of our lives are act of creation by God, the Father of statistics, I profess that Statistics is indeed a Super Discipline created by God and applied consciously and unconsciously in our everyday lives. As we have seen,

(i) With statistics we can transform the jangling discord of variations in a data set across every field of study into meaningful insights.

(ii) With statistics we would be empowered to navigate the turbulent sea of variations and uncertainties that we encounter in everyday lives.

(iii) With statistics, we can uncover the hidden information in a complex data set for useful purposes.

Finally, as I encourage you to not to be afraid of Statistics, considering that we Statisticians have made it simple, I also thank you for your presence and attention.

Mr. Vice Chancellor, I hereby rest my case.

REFERENCES

Akinyosoye, V. O. (2006), National Bureau of Statistics. http://www.nigerianstat.gov.ng

Akpanta A. C. (2002). Statistics: The Steps Approach 1, Rabboni Publishers International, Enugu.

Akpanta A.C. (2023): "Statistics and the Nigerian Experience: Rich in Data, Poor in Application", 66th Inaugural Lecture of Abia State University, Uturu.

Andrew W. Lo (2003): Statistics of Sharpe Ratios, Financial Analyst Journal, 58(4), February 2023
DOI: 10.2469/faj.v58.n4.2453

Iwueze I. S. and **Ohakwe J** (2004) Buys – Ballot Estimates when Stochastic Trend is Quadratic.

Journal of the Nigerian Association of Mathematical Physics, Vol. 8: pp 311 – 318

Iwueze I. S. and **Ohakwe J** (2009) Penalties For Misclassification Of Bilinear And Linear Time Series Processes, Interstat Journal of Statistics: No. 3

http://interstat.statjournals.net/YEAR/2009/abstracts/0906003.php

Iwueze I. S. and **Ohakwe J** (2009) Covariance Analysis of the Squares of the Diagonal Completely Bilinear Time Series Models, Brazilian Journal of Probability and Statistics, Vol. 25, No. 1, 90 – 98, 2009. DOI: 10.1214/09-BJPS111
www.imstat.org/bjps/papers/BJPS111.pdf

Iwueze, Iheanyi S. and **Ohakwe, Johnson**(2011):Square Root Transformation of the Quadratic Equation. Asian Journal of Mathematics and Statistics.4 (4): 186 – 199, 2011. ISSN 1194 – 5418. DOI: 10.3923/ajms.2011.186.199.

Iwueze I.S. and Nwogu E. C. (2004): Buys-Ballot Estimates for Time Series Decomposition, Global Journal of Mathematics, 83 – 98.

Nnorom, Innocent C., Ohakwe Johnson and Osibanjo, Oladele (2009)Survey of Willingness of residents to participate in

electronic waste recycling in Nigeria – A case study of mobile phone recycling, Journal of Cleaner Production, Elsevier: 17 (2009) 1629 – 1637. Doi:10.1016/j.jclepro.2009.08.009
Available online at www.sciencedirect.com
Award: No 15 on the list of Top 25 Hottest Articles, October – December, 2009(Science Direct TOP 25 Hottest Articles)
http://top25.sciencedirect.com/subject/energy/11/journal/journal-of-cleaner-production/09596526/archive/24/
Ogbonnaya C. Ogbonnaya, Diljit Singh and **OhakweJohnson(2011)**: Significant Predictors of Library Use by New International Students: the case study of University of Malaya, Kualla Lumpur, Malaysia. Asian Journal of Applied Sciences 4 (1): 53 – 62, 2011.
ISSN: 1996 – 3343. DOI: 10.3923/ajaps.2011.53.62
http://scialert.net/fulltext/?doi=ajaps.2011.53.62&org=12
Ohakwe J.and I. S. Iwueze(2009) Characterizations of the Moments of the Purely Diagonal Bilinear Time Series Model of Order One, Interstat Journal of Statistics: No. 2, June 2009. http://interstat.statjournals.net/YEAR/2009/abstracts/0906002.php
Ohakwe J., Dike O.A. and Akpanta A.C. (2012).: The Implication of Square Root Transformation on a Gamma Distributed Error Component of a Multiplicative Time Series model. Proceedings of African Regional Conference on Sustainable Development, Volume 6, Number 4, pp. 65-78, June 11 – 14, 2012, University of Calabar, Nigeria.
Ohakwe J., Akpanta A.C.and Dike O.A. (2012): Unit Mean and Constant Variance of the Generalized Gamma Distribution after Square Root Transformation in Statistical Modeling. Mathematical Theory and Modeling, Vol. 2, No. 12, 2012, pp. 22 – 34.
Ohakwe,J., Iwuoha, O and E. L. Otuonye(2013)Condition for

successful Square Transformation in Time series modeling. Applied Mathematics, Vol. 4, pp. 680 - 687, 2013. DOI: 10.4236/am.2013.44093.

Okorie, I. E., Akpanta A. C. and **Ohakwe J** (2016): The Exponentiated Gumbel Type-2 Distribution: Properties and applications. International Journal of Mathematics and Mathematical Sciences, Volume 2016, 10 pages. Hindawi Publishing Corporation.

http://dx.doi.org/10.1155/2016/5898356

Okorie, I. E., Akpanta A. C. and **Ohakwe J**. (2016): Transmutted Erlang Truncated Exponential Distribution, Econ. Qual. Control, 31 (2), 71 – 84, 2016.

DOI:10.1515/eqc-2016-0008.

Olubusoye, O. E.,Keshinro, O. and Korter, G. O.(2015): Nigerian Statistical System: The Evolution, Progress and Challenges. Accessed on ResearchGate:

https://www.researchgate.net/publication/283715250.

FEDERAL UNIVERSITY OTUOKE
INAUGURAL LECTURE SERIES

S/N	LECTURER	TITLE	DATE
1	Prof. Teddy Charles Adias	Blood: Convergence of Science and Religion	24/07/2019
2	Prof. (Mrs.) Iheoma M. Adekunle	Chemistry in Analytical Rahapsody for Environmental Stability	21/08/2019
3	Prof. Okechukwu E. Okeke	Of Restructuring and Related Matters: The Historian as Moderator	04/09/2019
4	Prof. Charles A. Negbenebor	The Chemical and Microbial of Meat Storage Stability	02/10/2019
5	Prof. Eno O. Ekanem	The Eagle Eye of Nature – Facts, Figures and Cocepts: An Analyst Perspective	16/10/2019
6	Prof. Chris C. Onyema	Writing Next Door to (S)Hell: Oil and Environmental Rights Discourse in the Niger Delta	02/09/2021
7	Prof. Cletus O. Akenbor	Paradigm Shift in Accounting: Epistemological Phenomenological Perspectives	04/11/2021
8	Prof. Dirikon I. Teilanyo	'Permission Me, *Biko*, To Make *Dis Spokage*.'The Beatification of A Beautiful Beast?	10/02/2022
9	Prof. Akachi C. Odoemene	'Settled, Usettled and Resettled': Dialecctics of History and (Dis)Order in the Nigerian Society	26/07/2023
10	Prof. Francis A. Oluleye	'Quest for Economic Development in Africa': How far have we gone? How far can we go?	27/09/2023