UNIVERSITY OF PORT HARCOURT

STATISTICS:
THE MAN IN YOUR NEIGHBOURHOOD

An Inaugural Lecture

By

PROFESSOR MARY PASCHAL IWUNDU
B.Sc. (Uyo); M.Sc., Ph. D (Nigeria); FNSA
Department of Mathematics and Statistics
Faculty of Science

INAUGURAL LECTURE SERIES

NO. 199

29TH MAY, 2025

University of Port Harcourt Printing Press Ltd.
University of Port Harcourt,
Port Harcourt,
Nigeria.
E-mail: uniport.press@uniport.edu.ng

ISBN: 1119-9849
INAUGURAL LECTURE SERIES NO. 199
DELIVERED: 29TH MAY, 2025

Designed, Printed and Bound By UPPL

ORDER OF PROCEEDINGS

2.45 pm. Guests are seated

3.00pm. Academic Procession begins

The Procession shall enter the CBN Centre of Excellence auditorium, University Park, and the Congregation shall stand as the Procession enters the hall in the following order:

Academic Officer
Professors
Deans of Faculties/School
Dean, School of Graduate Studies
Provost, College of Health Sciences
Lecturer
University Librarian
Registrar
Deputy Vice Chancellor Research and Development
Deputy Vice Chancellor Academic
Deputy Vice Chancellor Administration
Vice Chancellor

After the Vice Chancellor has ascended the dais, the Congregation shall remain standing for the University of Port Harcourt Anthem.

The Congregation shall thereafter resume their seats.

THE VICE CHANCELLOR'S OPENING REMARKS.

The Registrar shall rise, cap, invite the Vice Chancellor to make his opening remarks and introduce the Lecturer.

The Lecturer shall remain standing during the Introduction.

THE INAUGURAL LECTURE

The Lecturer shall step on the rostrum, cap and deliver her Inaugural Lecture. After the lecture, she shall step towards the Vice Chancellor, cap and deliver a copy of the Inaugural Lecture to the Vice Chancellor and resume her seat. The Vice Chancellor shall present the document to the Registrar.

CLOSING

The Registrar shall rise, cap and invite the Vice Chancellor to make his Closing Remarks.

The Vice Chancellor's Closing Remarks.

The Vice Chancellor shall then rise, cap and make his Closing Remarks. The Congregation shall rise for the University of Port Harcourt Anthem and remain standing as the Academic [Honour] Procession retreats in the following order:

Vice Chancellor
Deputy Vice Chancellor Administration
Deputy Vice Chancellor Academic
Deputy Vice Chancellor Research and Development
Registrar
University Librarian
Lecturer
Provost, College of Health Sciences
Dean, School of Graduate Studies
Deans of Faculties/School
Professors
Academic Officer

PROTOCOLS

- The Vice-Chancellor, University of Port Harcourt
- The Deputy Vice-Chancellor (Administration)
- The Deputy Vice-Chancellor (Academic)
- The Deputy Vice-Chancellor (Research & Development)
- The Registrar, University of Port Harcourt
- The University Bursar
- The University Librarian
- Past Principal Officers of the University of Port Harcourt
- Members of the Governing Council of the University
- The Provost, College of Health Sciences
- The Dean, School of Graduate Studies
- Deans of Faculties and of Student Affairs
- Directors of Institutes and Centres
- Heads of Departments
- Principal Officers from Other Universities and Institutions
- Chairman and Members of the Screening Committee on Inaugural Lecture Series
- Distinguished Professors and Academic Colleagues
- Senior Administrative, Technical and Non-Teaching Staff of the University of Port Harcourt
- Clergy and Religious Leaders
- Distinguished Invited Guests
- Unique Students of the University of Port Harcourt
- Esteemed Friends and Family
- Online Viewers
- Members of the Press
- Ladies and Gentlemen

DEDICATION

TO MY LATE MOTHER
She was flawless and priceless

ACKNOWLEDGMENTS

As I reflect on the path to this milestone, I am first and foremost thankful to the Almighty God, whose mercy and love have sustained me. Who am I that He is mindful of me and cares for me? He has made me a little lower than the angels and crowned me with glory and honour. To Him be all the glory. Amen.

I would like to express my profound gratitude to the Vice-Chancellor of our esteemed University of Port Harcourt, Prof. Owunari Abraham Georgewill, for approving the Inaugural Lecture Series No. 199 and providing us with a platform to showcase our talents. I also appreciate the enabling environment he has fostered, which allows us to contribute meaningfully to both the university and society.

My heartfelt thanks extend to the Deputy Vice-Chancellors: Prof. Clifford Ofurum (Administration), Prof. Kingsley Owete (Academic), and Prof. Iyeopu Siminialayi (Research and Development). I also thank the Registrar, Dr. Gloria Obiageri Chindah; the Bursar, Dr. Godpower Obah; and the University Librarian, Prof. Helen Emasealu, for their continued support.

I am deeply grateful to the past Vice-Chancellors and Principal Officers of the University of Port Harcourt, especially Prof. Joseph Ajienka, the 7th Vice-Chancellor, under whose leadership I was honoured to serve as Head of the Department of Mathematics and Statistics. My tenure as Head was marked by significant achievements, and I remain committed to the university's progress.

I would like to especially acknowledge the Inaugural Lecture Screening Committee for their unwavering dedication to

excellence in the presentation and delivery of Inaugural Lectures. A professor's greatest distinction lies in encouraging others, and I am proud to have had such role models in my journey.

Education is a lifelong pursuit, and I am indebted to the memory of my dear mother, the late Mrs. Cornelia Usoro, who nurtured my intellectual growth from a young age. She was not only my mother but also my home teacher. Mom, your memory is forever cherished. You were priceless, and I feel your love every day. I am assured of your care, and if I still needed a mother, I would still prefer you to be mine. May your beautiful soul continue to rest in peace.

Those early days of counting with sticks and stones were memorials of impactful learning. I would like to give respect to the foundational teachers who shaped my early education. Among them is the late Mr. Thompson, a renowned teacher at Christ the King Primary School in Uyo, Akwa Ibom State, who gave me the opportunity to complete Primary 5 and 6 in one academic year.

Additionally, I revere the late Mrs. R. M. Geo-Jaja, my White Principal at Immaculate Conception Secondary School in Itak-Ikono, Akwa Ibom State, who instilled in us a sense of self-worth. Under her mentorship, along with the guidance of numerous local, national, and international educators and affiliates, I learned to believe in myself and aspire to greatness. May their memories be forever blessed.

I am also grateful to the university lecturers and mentors who guided me along my academic path, particularly Prof. M. U. Umoren, Prof. E. E. Joshua, Prof. I. B. Onukogu, Late Prof. J. N. Adichie, Prof. P. E. Chigbu, Prof. E. C. Nduka, and my treasured colleague and friend, the Deputy Vice-Chancellor (Administration), University of Uyo, Prof. Aniekan Offiong.

To all my colleagues in the Department of Mathematics and Statistics at the University of Port Harcourt, the Department of Statistics at the University of Uyo, and the Department of Statistics at the University of Nigeria, Nsukka, I am eternally grateful for the support you have given me. You have all played a role in bringing me to this moment.

Let me also acknowledge a then-serving corps member of the NYSC who in my SS1 saw in me the capacity to become a Mathematician. He himself had the capacity to bring out the best in us. Sir, the script you endorsed 100% in the 1985/1986 academic session has not been discarded (at least not by me). Wherever you are, I wish you well.

My deep appreciation goes to my wonderful family. You are all truly special, and I am blessed to be a part of this incredible family. To my uncles and aunties, thank you for showing me the true meaning of love and care. I hold you all in the highest regard. To my siblings, cousins and extended family, your unwavering support has been my strength. A special thanks to my elder sister, Mrs. Inemesit David, who has been a pillar of support and love.

I sincerely appreciate my husband's family for their goodwill. I am grateful for the connections we share. I must especially thank my dear husband, Paschal, for being a beacon of support and affection. To our wonderful children—Dr. Chinwendu, Engr. Chukwudi, Miss. Chisomaga, and Master Noble—you are my greatest blessings and my greatest pride. Your achievements bring me fulfillment and joy. Watching you grow into remarkable individuals is one of my life's greatest rewards. I cherish you beyond words and will always be your unwavering support.

The Church holds a central place in my journey to success. As the Cyprian of Carthage wisely said, “No one can have God for his Father who has not the Church for his mother.” I am deeply grateful to the Church for her unwavering support, guidance, and spiritual nourishment. Through my engagements in and with the Church, I have gained treasures far greater than silver and gold. I have been strengthened, uplifted, and guided by divine wisdom.

God has made my feet like the hind’s feet—firm, steady, and prepared to navigate the paths of trials and triumphs. He has set me securely upon my high places, leading me with His unfailing grace (Psalms 18:33). Today, I stand strong and can boldly declare that God has been faithful. Remarkably, He remains the same yesterday, today, and forever.

I sincerely acknowledge all my friends, each of whom has added joy, encouragement, and support to my journey. Some are especially dear, standing by me through life's seasons with unwavering kindness and warmth. Your friendship is a treasure, and I deeply appreciate you all.

I also extend my heartfelt appreciation to all who have contributed to shaping this lecture, including those who provided invaluable guidance, insightful perspectives, and unwavering encouragement along the way. Their support and collaboration have been instrumental in bringing this work to fruition, and I am deeply grateful for their contributions.

To our great students, you are precious gifts to the academia. Without students, there are no professors, and without you, there is no me. Think differently today: You are braver than you believe, stronger than you seem, and smarter than you think. Embrace your potential, for the future of knowledge and progress rests in your hands.

To the Unique Uniport, I seek an “Alma Matership” or an “Alma Materism” with you. Permit me to be the first user of these words. As we accept them as part of language evolution, let us appreciate the need to form a lasting connection or identity with one’s Alma Mater.

I love you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF PLATES

PREAMBLE

MY JOURNEY

I am very delighted to be the inaugural lecturer today. I am a Professor of Statistics, in the Department of Mathematics and Statistics, Faculty of Science. With utmost humility, it is indeed an honour to be granted the approval of the Vice-Chancellor for the Inaugural Lecture Series No. 199. In the symbolic or spiritual language of numbers, "1" represents beginnings, leadership, and individuality. "99" is often associated with wisdom and humanity. Together, "199" signifies a culmination of a journey from individual vision to collective impact.

Mr. Vice-Chancellor, Sir, the decision to embark on my journey began like a coin experiment, requiring acceptance of one of two options. It was like a game of choice involving a coin with a Head (H) and a Tail (T). It is certain that when a coin is tossed once, the result will be either a Head or a Tail, but not both. For a balanced coin, the probability of getting a Head is the same as the probability of getting a Tail. When the coin is not balanced, these probabilities differ. However, the totality of the probabilities remains unity (1).

How well can I describe my journey?
In statistical terminologies, one would ask:
Was the journey a random process?
Was it a random walk?
Was it hill-climbing or valley descending?
Was it a type of steepest ascent or descent?
Were there Hypothetical Conjunctures?
Were the conjunctures tested?

Looking back, it has been like a Marathon. The type Dean Rarnazes posits “It doesn’t matter how you do it, just get out there and do it”. Because “Life in itself is a Marathon, not a sprint and requires taking a Bigger-Picture approach”-Iyanka Trump.

In 1989, I gained admission into the Bachelor of Science programme in Statistics of the University of Uyo and graduated with a second-class upper division after walking through the fundamentals and rudiments of the course requirements for the award of BSc. Statistics. I felt satisfied that I had run the first third of the race with my head. Thanks to my lecturers then and a special regard to Prof. M. U. Umoren.

In 1995, I accepted the offer of a Graduate Assistantship in the Department of Mathematics, Statistics and Computer Science of the University of Uyo and realized that I was in for a marathon; a long-lasting and possibly difficult task, having pits, valleys and hills. Like mountain-climbing you could fall yet when you fall, you should have the strength to get up and keep going, enduring the pains as you look up to the joy at the finish line.

In 1997, I proceeded to the University of Nigeria, Nsukka, for a Master of Science programme in Statistics with a programme goal to specialize in the field of Experimental Design. Truly, it was more than I thought. Oh, I remember Prof. I. B. Onukogu “Young lady, you need to come to class on time”. This got me smart and awakened me to the fact that I needed to run the second third of the race with my legs. The marathon continued and to me it was the start of the marathon in its continuity, searching for the global optimum.

It became clear that the marathon was a type of search algorithm for the global optimum without falling into the pit of several local optima. And so, in my race and decisions, the Best of Bests should be my goal. As one belonging to the Classical and Bayesian schools of thought, running the race with my legs that were conditioned on my head yielded me an MSc. Statistics degree in Optimal Design of Experiments with a CGPA of 5.0 on a 5-point scale.

Down the line, I enrolled to take a Ph.D. in Statistics under the mentorship and supervision of Prof. I. B. Onukogu. I learned “Pitching”; how best you can convince your best audience, possibly made up of the good, the bad, and the ugly. This time he said, “Mary, you know this subject better than any other person. It’s your work!”. It was time to defend the last third of the race in 2007 with my heart conditioned on my head and my legs. Thanks to Prof. P. E. Chigbu who co-supervised my Ph.D. Dissertation on the title “A Hill-Climbing Combinatorial Algorithm for Constructing D-Optimal Exact Design”. It was indeed a combinatorial procedure that took center stage in the Journal STATISTICA Vol. 67, No.4 (2007). STATISTICA is indexed in Clarivate Analytics Web of Science and Scopus.

My journey from being a Graduate Assistant to becoming a professor took twenty-two years of intense effort, with studentship, motherhood, and lectureship. To me, it is indeed notable progress. However, as Gary Gregory rightly says, “The finish line is not the finish line; it becomes the starting line of your next race”.

Mr. Vice-Chancellor Sir, I am in a race that is a non-stop marathon and that marathon for me is a statistical walk-through whose pathway contains several utensils and designs that are well packaged as statistical toolkits for research and innovation, with the drive that as I reach one goal, I find a new goal and keep going. Indeed, achieving any goal requires starting well, going in the right direction, and utilizing the right steps and modalities. That is the essence of Experimental Design.

STATISTICS: THE MAN IN YOUR NEIGHBOURHOOD

INTRODUCTION

There is a popular saying by the renowned statistician John Tukey: *“The best thing about being a statistician is that you get to play in everyone's backyard.”* This statement captures the essence of Statistics—its quiet but powerful presence in every field, every home, and indeed, every neighbourhood. Statistics is an integral part of living as it influences nearly everything around us and helps to shape important daily decisions. In today’s data-driven society, Statistics serves as the mathematical foundation upon which much of the information around us is built and interpreted. Unfortunately, information expressed in what is observed in our everyday lives and engagements would be meaningless if the right interpretation is not given for an informed decision. Statistics brings meaningfulness to data. It stands as the cornerstone of research, the power in our thoughts, the beauty of our minds, and the insights into evidence-based decisions.

ORIGIN OF STATISTICS

Mr. Vice-Chancellor, Sir, the birth of Statistics is often dated to 1662, when John Graunt, along with William Petty, developed early human statistical and census methods that provided a framework for modern demography. Specifically, he produced the first life table (also called a mortality table) that shows, for each age, what the probability is that a person of that age will die before their next birthday ("probability of death" or “probability of survival”). In other words, it represents the survivorship (population’s longevity) of people from a certain population.

John Graunt has been regarded as the founder of demography and perhaps the first epidemiologist, though by profession he was a haberdasher. He was the first to document the phenomenon of 'excess deaths' during epidemics. He provided a template for numerical analysis of demographic and health data and initiated the concepts of statistical association, statistical inference and population sampling. Sir William Petty, a 17th-century English economist, was known for his early statistical methods to analyze demographic data.

William Petty (26 May 1623 – 16 Dec 1687)

John Graunt (24 April 1620 – 18 April 1674)

Basic forms of Statistics have been used since the beginning of civilization. Today, Statistics has gained prominence and as rightly put ***"Statistical thinking will one day be as necessary for efficient citizenship as the ability to read or write."*** *H. G. Wells*.

THE MAN IN YOUR NEIGHBOURHOOD

Mr. Vice-Chancellor, Sir, permit me to state that as with the two founding fathers of Statistics there is a man in your neighbourhood. He has some adoring features and offers solutions to life challenges. He finds relevance in Machine

Learning, a branch of Artificial Intelligence, and is inherently a necessary component of Data Science. He has great analytical, summarization, predictive, and precision power. When he takes an action, he is unbiased, efficient, sufficient and consistent. That man is Statistics. In the face of uncertainty, he stands strong. Among other disciplines, Statistics is one of the most important disciplines for data and environmental Scientists. In fact, "Statistics is the sole of research and sharpest tool of Science". Statistics and statistical applications are so vital that we can say *YOU ARE STATISTICS AND STATISTICS IS YOU.*

- In Mathematics, he is called Mathematical Statistics.
- In Economics, he is called Econometrics.
- In Business/Management, he is called Business Analytics.
- In Psychology, he is called Psychometrics.
- In Biology and Public Health, he is called Biostatistics.
- In Sociology, he is called Quantitative Sociology.
- In Engineering, he is called Engineering Statistics.
- In Education, he is called Educational Statistics.
- In Computer Science, he is called Data Science.
- In Environmental Science, he is called Environmental Statistics.
- In Marketing, he is called Marketing Analytics.
- In Anthropology, he is called Quantitative Anthropology.
- In Linguistics, he is called Quantitative Linguistics.
- In Geography, he is called Spatial Statistics.
- In Medicine, he is called Medical Statistics.
- In Philosophy, he is called Philosophical Statistics.
- In Physics, he is called Statistical Physics.
- In Chemistry, he is called Chemometrics.
- In Sports Science, he is called Sports Analytics.
- In Demography, he is called Demographic Methods.

- In Music, he is called Music Analytics.
- In Dentistry, he is called Dental Statistics.
- In Theater and Performing Arts, he is called Performance Analytics.
- In Law, he is called Legal Statistics.
- The list is unending, and the significance is profound. These justify that Statistics is the universal language that speaks to every profession.

DEFINING WHAT STATISTICS IS

Many sources have tried to define what Statistics is and all seem to be unified in the concept of data collection and analysis. From Oxford Languages and Google, it is seen as a noun that expresses the practice or science of collecting and analyzing numerical data in large quantities, especially for the purpose of inferring proportions in a whole, from those in a representative sample. Statistics is considered a branch of science that involves the drawing of inferences from sample data to the whole population and includes effective ways to gather, review, analyze, and draw conclusions from the data.

According to the APA Dictionary of Psychology, Statistics is “the branch of Mathematics in which data are used descriptively or inferentially to find or support answers for scientific and other quantifiable questions”. We can agree that Statistics is a branch that deals with every aspect of the data and is a crucial process that helps to make decisions based on the data. Hence, we can summarily define Statistics as the discipline that concerns the collection, organization, presentation, analysis, and interpretation of data for informed decisions. For the modern researcher, Statistics is the “Heel of Achilles”. Its complexity requires careful interpretation, as misunderstanding or misinterpretation can lead to erroneous conclusions and flawed decisions.

While presenting “A Christian View of the Foundations of Statistics”, Geertsema (1987) considered Statistics to be connected to all empirical sciences, acting as a mathematical auxiliary to them. All sciences make use of data to some extent, and Statistics is just that ancillary science that has the task of handling the empirical data. Hence, this unique characteristic of Statistics implies a closely assumed relationship between Statistics and Philosophy of Science.

The importance of data literacy should be given great advocacy in our increasingly data-driven world. It is worthy of note that statistics provides a framework for making data-driven decisions, minimizing uncertainties, and reducing biases. In fact, among other disciplines, statistics is one of the most important disciplines for data scientists (University of Virginia, 2021). Taking my mind to the classroom between 1989 and 1994 where Statistics was taught in segmented but intertwined categories, each category directly or indirectly involved and still involves data generation and requires a crucial process that helps to make decisions based on the data.

Figure 1: Segmented Categories of Statistics

In research, the main purpose of using statistics is to plan the data collection process in terms of experimental designs and statistical surveys. By way of a simple definition, a statistical experiment is an ordered procedure that is performed to verify and determine the validity of a statistical hypothesis (a statement) postulated about the population from which the sample data are derived. The experimental design helps to choose the proper method of collecting the data and employs the sample data in the correct analysis process to effectively produce well-authenticated results.

BUT WHAT IS DATA?

Data is the heart of research and a necessary requirement in empirical study. It refers to the raw information or observations collected from various sources in the world. For example, there is a strong cross-cultural assumption that women talk a lot more than men. This is taking a realistic place in research as documentations show that most women speak on average 20,000 words per day, and approximately 13,000 more than the average male ("Women Speak 20,000 Words," n.d.; "Women Use More Words," 2025). Without jumping the gun, this is data. It is a type of discrete data.

The World Health Organization classifies Body Mass Index ≥ 30.0 as Obesity. That again is data and very insightful if we pay attention to the story they tell. A United Nations Children's Fund Report revealed that one in six (16.67%) Nigerians between the ages of 15 and 24 often feel depressed (UNICEF, 2021). This also is data and quite informative as this is the age group where many young Nigerians are pursuing tertiary education. In common usage, data is a collection of discrete and continuous values that convey information, describing the quantity, quality, facts, sequences of symbols, and other basic units of meaning that may be further interpreted formally.

It is of interest to note that data collected need not be numerical as even pictures are data, moments on the internet form data, the social media sites visited and/or explored constitute data, sleep orientations are data (called circular data), the way one smiles is data, and every action around humans or animal subjects is data. These unstructured data types are manipulated through the general data concept called Datafication, which offer technological trends that aim at transforming every aspect of human's engagement into quantifiable data that can be tracked, monitored, and analyzed. Regardless of the form of data, data speaks, and understanding what data says is Science.

DO YOU NEED DATA? DO YOU USE DATA?

Mr. Vice-Chancellor, Sir, to lend credibility to the importance of data and analysis in our present world, allow me to share the perspectives of renowned experts:

"We are surrounded by data but starved for insights." — *Jay Baer*

"Data will talk to you if you're willing to listen." — *Jim Bergeson*

"You can't improve what you don't measure." — *Cliff Lerner*

"Without data, you're just another person with an opinion." — *Edwards Deming*

“Those who rule data will rule the entire world.” — *Masayoshi Son*
“God loves Statistics.” — *Jan Geertsema*

These statements powerfully underscore the critical role of data in shaping decisions, policies, and progress across all spheres of life. Interestingly, Data is everywhere, and Data is life. Every day, a vast amount of data is generated in the business sector, administrative sector, medical sector, and in our homes as well. This proves that humans need data, businesses need data, everyone needs data, and without data, everything in life will pretty much be meaningless. Being surrounded by data is a clear indication that Statistics is in your neighbourhood, and where there is data, there is statistics. The better the data, the better the results.

Think of how life would be without Statistics. To me, life without Statistics is a bitter experience. Statistics is required to study human behaviour, social environments, health, business environments, drugs composition and risks. Statistical methods are applied to assess risk in the insurance and financial industries. Health statistics are used to understand risk factors for communities, track and monitor health events such as diseases, see the impact of health policy changes, and assess the quality and safety of health care. Statistics in Psychology is focused on the collection and analysis of data involving the science of mind and behaviour. In fact, Statistics is useful in Psychology because it helps a psychologist to determine what is typical or normal for a particular group.

Insurance statistics, on the other hand, give the frequency with which untoward events occur in a class of insured individuals, and with variance in the severity of the effects of those events.

Social statistics requires the use of statistics to study human behaviour and social environments and help enhance information or knowledge on an individual, object, or event. For government agencies to thrive, Government Statistics provide accurate, timely, and credible information, which informs current and future policies. Statistical physics helps in the study of special laws that govern the behaviour and properties of macroscopic bodies. In Chemistry, statistical methods are applied in chemical studies for an informed analysis of chemical compounds towards a more efficient management of flow of the information.

Statistics in engineering is used for designing experiments, analyzing data, summarising and presenting information, and drawing reliable conclusions. It guides risk management, quality control, reliability analysis, and making informed decisions in design and operational processes. In education, Statistics is the bedrock of research; no theorem is justified without statistics. In the Medical field, Statistics is referred to as “indispensable knowledge” that the modern researcher cannot refuse to know and use. An adequate knowledge of Statistics is necessary for the proper designing of an epidemiological study or a clinical trial, as improper statistical methods may result in erroneous conclusions which may lead to unethical practice. The necessity to apply statistics to any scientific problem leads the researcher to assume the burden of a deeper knowledge of science.

CHALLENGES AND MISUSE OF DATA

There are challenges associated with data usage. One such is the widespread misuse and misrepresentation of statistical information which oftentimes give the Statistics Profession a bad name. There is the wrong use of statistical designs, statistical tests, and falsifications of statistical data. It is often said that data do not lie. Unfortunately, some people who use data to support their claims in scientific and social experiments do lie. An investigation of a famous social psychologist (name withheld) showed he was guilty of data falsification on over 55 papers he authored and were published in top-rate journals, and in 10 PhD dissertations he supervised ("Psychologist Admits Faking Data," 2011). The social psychologist did not deny the falsification. He insisted that he loved social psychology “but had been frustrated by the messiness of experimental data (my thought: considered messiness), which rarely led to clear conclusions (my thought: because he wanted to force the data to speak his thoughts against the truth)”. This led him to concoct results that journals found attractive.

Mr. Vice-Chancellor Sir, it was a quest for aesthetics, a quest for beauty instead of a quest for the truth. Unfortunately, many statistical frauds are difficult to spot and only sometimes present themselves with some obvious inconsistencies. In modeling sometimes, a chosen model may not be a good fit for the data at hand, and the usual suggestion would be to find a more suitable model. What happens when data have been falsified? Do we keep searching for models that may fit the falsified data? No. The way to go is **“Allow data to speak and accept the story they tell”**. This will require collaborative efforts to actualize the UN World Data Forum's objectives of promoting data innovation, fostering collaboration, mobilizing high-level political and financial support for data, and creating a path for better data toward sustainable development.

HOW RELIABLE ARE YOUR DATA?

Attention must be given to getting reliable data. Data reliability refers to the completeness and accuracy of data as a measure of how well it can be counted on to be consistent and free from errors across time and sources ("What Is Data

Reliability?" n.d.). The more reliable data are, the more trustworthy they become. Trust in data provides a solid foundation for drawing meaningful insights and well-informed decision-making. Inaccurate or unreliable data can lead to incorrect conclusions, flawed models, and poor decision-making. It is worth noting that something as “little” as mistakenly using O for 0 could cause a big issue in data analysis. This is a type of bad data in a dataset. As in data dispensation, bad data in a database refers to inaccurate, inconsistent, incomplete, or improperly formatted information that can compromise the quality and reliability of the data.

Some bad datasets include Missing Data, Inaccurate Data, Outdated data, Inconsistent Data, Outliers, Data Insecurity (Storing sensitive data like passwords in plaintext instead of securely hashing them), Inappropriate Data (violating privacy or legal regulations), Garbage Data (serving no legitimate purpose), and several other anomalies. To maintain data integrity and ensure the reliability of a database, it is important to implement data validation, data cleansing, data auditing, and several other data quality assurance processes. These measures enhance data-driven decision-making and improve the overall effectiveness of database systems.

VARIOUS STATISTICAL RESEARCH METHODS IN THE NEIGHBOURHOOD

The field of statistics is broad but generally classified as inferential or descriptive. Each of the two broad classifications requires data generation which is a function of an experimental design. Statistics offer various research methods that seem to cut across all areas of research in human and material environments. The research methods include Descriptive analysis, Inferential analysis, Predictive analysis, Prescriptive analysis, Exploratory data analysis, Causal analysis and Mechanistic analysis (Ali & Bhaskar, 2016).

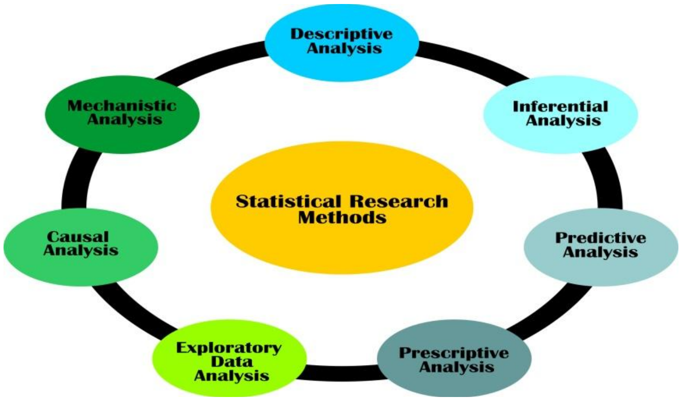

Figure 2: Various Statistical Research Methods

i. **Descriptive Analysis:**

Descriptive analysis is an important phase in data exploration that involves summarizing and describing the primary properties of a dataset. It provides vital insights into the data's frequency distribution, central tendency, dispersion, and identifying position. As an example, UN News reported that 783 million people faced hunger worldwide in 2022 as a result of the COVID-19 pandemic. Nigeria was ranked 103rd on the 2022 Global Hunger Index, with a score of 27.3**,** 109th on the 2023 Global Hunger Index, with a score of 28.3, and 110th on the 2024 Global Hunger Index, with a score of 28.8 (Akinola, 2024). These statistics identify the Nigeria’s position on Global Hunger ranking and are hence descriptive. The Global Hunger Index (GHI) is a tool for assessing whether nations are meeting SDG that relates to Hunger. The Global Hunger Index indicates that the hunger level in Nigeria is serious ("Nigeria's Hunger Level ‘Serious,’ Ranks 103 Out of 121 Countries," 2022).

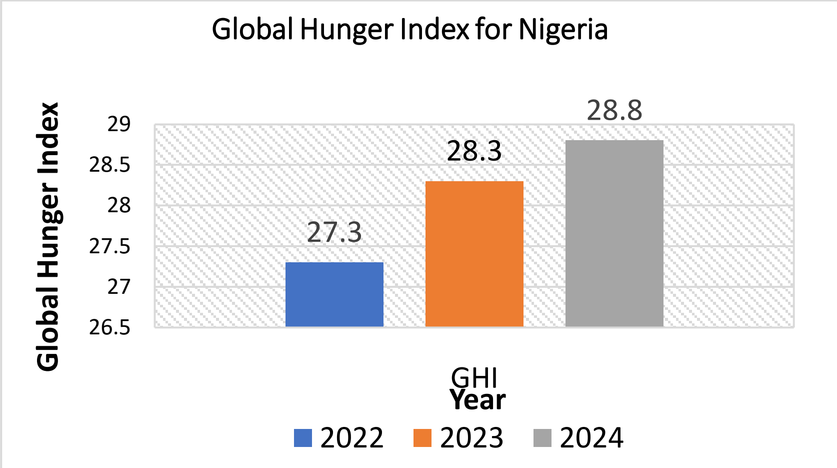

Figure 3: 2022-2024 Global Hunger Index

This is insightful and should inform evidence-based decisions and policy in view of achieving SDG 1. Hunger can be viewed as a dimension of extreme poverty, and poverty has been described as the worst form of violence. As rightly put by Nelson Mandela, "As long as poverty, injustice, and gross inequality persist in our world, none of us can truly rest."

ii. **Exploratory Data Analysis:**

Exploratory data analysis (EDA) helps to summarize the main characteristics of a dataset using data visualization methods. The technique identifies measures of central tendency and variability, gives data visualization plots, and establishes patterns, correlations, and data cleaning. Exploratory data analysis is usually the recommended first step in data analysis where the analyst utilizes visualization techniques to describe the dataset characterization. We took a tour of a laboratory within the University of Port Harcourt environment to study the demand for pregnancy tests among youths aged 15 to 25 years from 2019 to 2024 (Iwundu & Josiah, 2024). The visualization graphs present insights into the dataset. In particular, demand for pregnancy tests faced a decline from June 2019 and a sudden rise from May 2021and beyond. As with the demand for pregnancy tests, there is also a rise in the number of positive pregnancy tests.

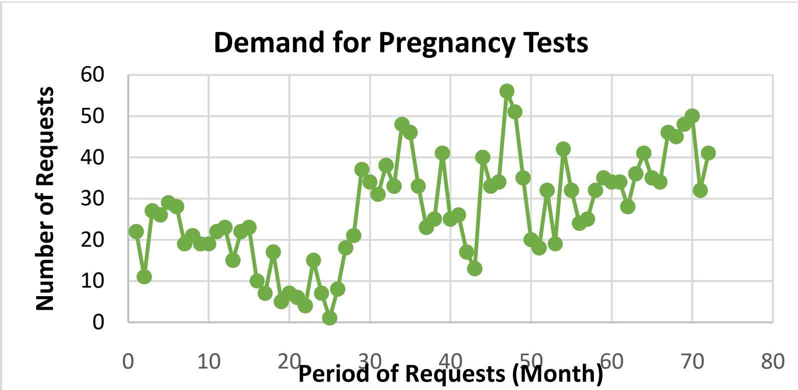

Figure 4: Series Plot of the Demand for Pregnancy Tests

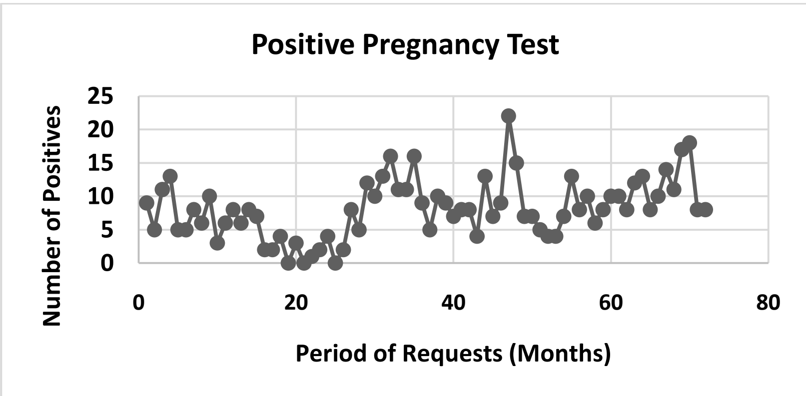

Figure 5: Series Plot of Positive Pregnancy Tests

For informed decision, it may be of interest to investigate the reason behind increase in the demand for pregnancy tests. Some questions to pay attention to include why the demand for pregnancy tests was highest in a particular month or season? What seasons are youths most vulnerable? Why is the demand for pregnancy tests increasing? These and many more help in decision-making towards achieving the Sustainable Development Goals (SDGs) of ensuring healthy lives and promoting well-being for all ages. Data-driven decisions could inform the need for reproductive health education programmes to increase knowledge, self-awareness, and self-assertiveness among youths. A related study is in Mekonen (2024).

iii. **Inferential Analysis:**

Every day, we face situations where we must make decisions, whether simple or complex, because life and progress depend on choices. When making decisions based on data, it is not practical to obtain all the measurements in a population of interest. Inferential analysis requires using sample data to draw conclusions or make predictions about the larger population. It is essential to note that data greatly affects the power of a statistical test and could cause two errors: Type I Error and Type II Error. A Type I error occurs when a test incorrectly indicates the presence of a condition when there is no such condition (False Positive). A Type II error occurs when a test fails to detect a condition that is actually present (False Negative). Both Type I and Type II errors greatly impact on statistical tests. In fact, Type II Error directly affects the power of the test as

$$Power = 1 - P(Type\ II\ Error)$$

For a good inference, choosing the right sample size and a representative sample is crucial. It is necessary to lay importance to the sample size that is a good representation of the population under study. It helps in ensuring reliable tests of hypotheses. A small sample may not represent realities well, leading to statistical errors and poor decisions, while an excessively large sample may be costly and unnecessary. For instance, a medical test that falsely indicates a patient has a disease when he is actually void of the disease indicates a Type I Error. In administration of criminal justice, a Type I Error occurs if a person is wrongly convicted of a crime he did not commit, based on false evidence or reasoning. This error is like blowing a false alarm!

Imagine in Business Research where tests suggest a new product would not move the market when it could have been a hit. This is a Type II Error and leads to lost opportunities. In Structural Engineering, a Type II Error occurs when a building or bridge has undetected structural weaknesses, leading to potential failure or collapse under stress. In Business, Type II Error is committed where a test suggests a new product would not move the market when it could have been a hit. This leads to lost opportunities. The power of a test lies in the test's effectiveness to correctly detect a real effect. Best practice demands that every test should seek to minimize these two errors, while aiming to increase the test's effectiveness to correctly detect a real effect.

We used birth-weight data to examine how different sample sizes affect the power of a statistical test (Iwundu & Onisokumen, 2016). The results of the sample size selection study indicated that as sample size increases, the probability of Type II Error decreases and the power of the test increases. Also, indicative was that an optimal sample size allows stronger evidence for correctly rejecting a false null hypothesis, thus, enhancing the test's effectiveness to correctly detect a real effect.

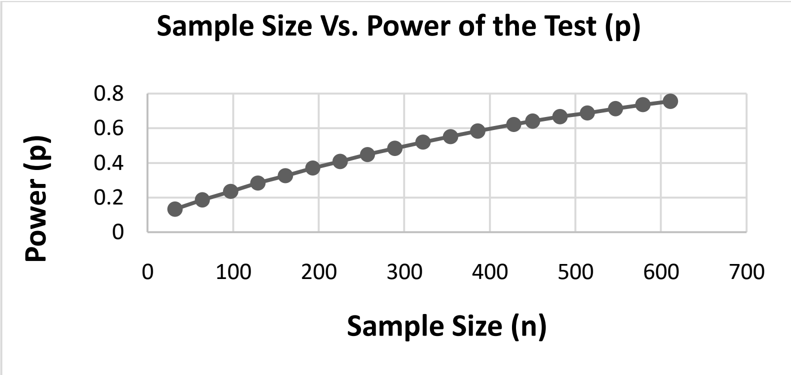

Figure 6: Sample Size and the Power of the Test (p)

Data information needs to be accurate as decisions based on wrong data could be very costly. A good example is in the controversies surrounding census figures and resource allocation, which have brought persistent disputes that highlight significant statistical and economic concerns, particularly regarding data integrity and its impact on policy formulation (Okonkwo, 2019; Adegboye & Olaniyi, 2021; Eze & Yusuf, 2020). Manipulation of census data leads to inaccurate population estimates, which in turn leads to wrong decisions in resource distribution, political representation, and social and economic planning. This misrepresentation undermines democratic principles and skews policymaking in favour of regions with exaggerated demographic figures. From econometric and statistical perspectives, the inaccuracies in census data exemplify both Type I and Type II errors in decision-making.

iv. **Predictive Analysis**:

Predictive analysis involves the use of current and historical data, modeling techniques, and machine learning to determine future performance, predict and plan for future events or opportunities, etc. An example of predictive analysis involves using money supply and exchange rates to predict inflation. It is indicative that excessive money supply leads to inflation, and depreciating currency leads to higher prices of imported goods and this contributes to inflation.

v. **Prescriptive Analysis:**

Prescriptive analysis requires the use of advanced processes and tools to analyze data to recommend the optimal course of action or strategy moving forward. Recalling that different offenders have different motives for the same behaviour, it seeks to answer the question, “What should we do?” This type of analysis requires using optimization algorithms to recommend optimal strategies for products and services. We have utilized various optimization algorithms for process control. Specifically, we proposed the variance-weighted gradient algorithm and the variance-weighted projection algorithm for optimizing responses. The algorithms make use of experimental design principles that seeks to minimize variance at every step of action (Otaru & Iwundu, 2017a; Otaru & Iwundu, 2017b; Iwundu & Otaru, 2019).

vi. **Causal Analysis:**

Casual analysis (also known as explanatory analysis) is a process for identifying and addressing the causes and effects of a problem. Causal analysis identifies the extent and nature of cause-and-effect relationships between two or more variables. It is often used by companies and organizations to determine the impact of changes in policies, products, features, or processes on critical company metrics. The era of artificial intelligence has ushered hype into the public consciousness. A large percentage of academic writing is AI aided and contains AI generated contents. This is a type of cause-and effect relationship and is examined using Causal Analysis. Specifically, Causal Analysis would help to identify whether the observed academic writing changes are indeed due to the hype in AI and no other factors.

vii. **Mechanistic Analysis**:

Mechanistic analysis is used to understand exact changes in variables that lead to other changes in other variables. It finds applications in Science and Engineering situations that require high precision and little room for error and deals with understanding complex systems. Mechanistic analysis seeks to explain underlying processes and the interactions that exist within a system. Interestingly, most mechanistic analyses rely on experimental design techniques from the design phase to the optimization phase. As an example, we modeled an experimental situation requiring the efficiency of Arsenite removal from groundwater using optimal design techniques (Iwundu & Cosmos, 2022).

The electrocoagulation process involving seven operating parameters, namely, initial pH, operating time, current, initial As (III) concentration, size of Al ball anode, air flow rate, and height of Al in the reactor including their interactions is a type of Mechanistic Analysis. Model Selection techniques arrived at reduced models containing only variables that optimized the process. The Reduced models offered good model fits and satisfactory optimality properties under D-optimality criterion. The totality of the Arsenite Removal process falls under Response Surface Methodology (RSM), the area of my major research focus. RSM offers mathematical and statistical techniques for studying the relationships between a response or multiple responses and several independent variables. The aim is usually to optimize the response or responses (in the presence of multi-objective functions).

OPTIMIZING DATA NEEDS THROUGH STATISTICAL EXPERIMENTAL DESIGN

Data information needs to be accurate as decisions based on wrong or insufficient data could be very costly. A good example is in the controversies surrounding census figures and resource allocation, which have brought persistent disputes that highlight significant statistical and economic concerns, particularly regarding data integrity and its impact on policy formulation. It is clearly proven that manipulations of census data lead to inaccurate population estimates, which in turn leads to wrong decisions in resource distribution, political representation, social and economic planning. This misrepresentation undermines democratic principles and skews policymaking in favour of regions with exaggerated demographic figures. From econometric and statistical perspectives, the inaccuracies in census data exemplify both Type I and Type II errors in decision-making.

Two basic ways of data collection are through surveys and experimental designs. Survey methodology collects information from a targeted group of people about their opinions, behaviours, or knowledge, through written questionnaires, face-to-face or telephone interviews, or electronic means. Experimental design methodology is a structured approach for collecting data and making discoveries. It projects a systematic, efficient method that enables researchers to study the relationship between multiple input variables (factors) and key output variables (responses). Ronald Fisher first introduced four enduring principles of Design of Experiment (DOE) in 1926: the factorial, randomization, replication, and blocking principles. As of today, generating and analyzing designs have grown from hand calculation to using computers for a more effective and efficient DOE.

Tracing back through generations, traditional farming systems handed down from generation to generation have always employed experimental design to arrive at conclusions pertaining to optimal spacing of crops, optimal soil depth for seeds, optimal planting season for different crops, before the advent of scientific agriculture. These optimal settings were products of experimental design even though the terminology was unknown to our ancestors. Although Sir Ronald Fisher (1890–1962) of blessed memory laid the foundation for modern experimental methodology, his principles were merely what our ancestors practised without formalization.

Maximizing the amount of data that can be gathered from an experiment, while minimizing the time, costs, and errors are some goals of an experimental design. A good experimental design entails determining the factors that must be investigated, selecting the right sample size, and setting up an experiment that will produce precise, consistent and reliable results. There are many designs in practice and many statistical methodologies. I have seen correctly used experimental procedures. I have also seen the misuse of experimental designs and statistical analyses, especially when data are put in tabular frames, having rows and columns. A well-structured experimental design facilitates the collection of good data. Good data promotes quality research. Statistical methods rely on the integrity of data and design's efficiency. All three are interdependent components of the research process. Their effective integration ensures that research findings are valid, reliable, and meaningful.

To obtain quality data that promote good research output, we have utilized standard experimental designs as well as computer-generated designs. Some standard designs we have utilized include Complete Factorial Designs, 2^k Factorial

Designs, Placket-Burman Designs, 3^k Factorial Designs, Fractional Factorial Designs, Central Composite Designs, and Box-Behnken Designs. These designs are known to be efficient and very robust under varying experimental conditions.

EXPERIMENTAL DESIGN PROCEDURE

What guides good experimental research is correctly adhering to the general experimental design procedure. The procedure requires using suitable designs in line with the research goals and phases. The experimental design procedure starts with defining the design objectives and variables, planning the data collection process via experimentation, employing a suitable data analysis/modeling procedure, interpreting the results, and drawing useful and insightful interpretations. There is actually no good data without a good design and there is no good model without good data.

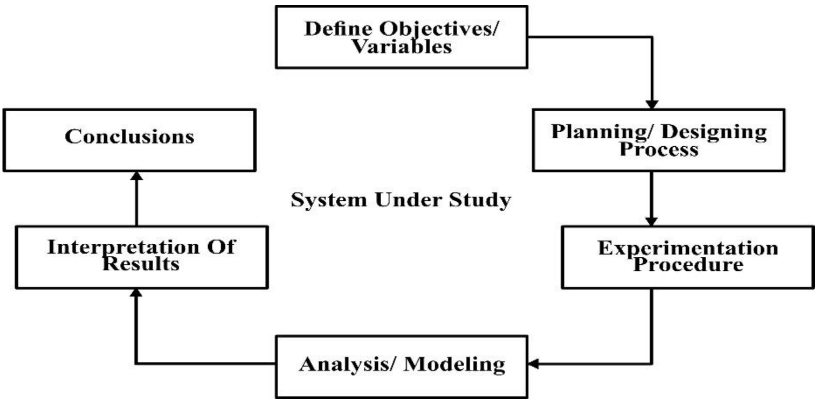

Figure 7: The General Experimental Design Procedure

SEQUENTIAL SEARCH FOR BEST EXPERIMENTAL CONDITIONS

The determination of the experimental design is crucial and requires a proper understanding of the research problem, careful selection of design and associated design points to satisfy optimality conditions. We undertook a project titled "An exchange algorithm for global optimization of response function". The idea was to provide a sequential search technique for locating the global optimizer of a response surface that contains several local optimizers (Usoro, 1998).

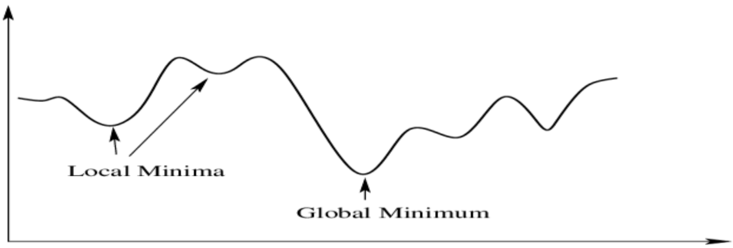

Figure 8: Optimization graph of Local and Global Minima

Whereas a local optimizer is a set of input values where the response function reaches its optimal value within a small neighbourhood, a global optimizer is a set of input values that yield the absolute best outcome across the entire input domain. The research utilized optimal design principles for arriving at the global optimum in few iterative moves, and without converging at a local optimum. The research offered a technique for arriving at the global optimum without falling into the pits of local optimizers. The search requires:

i. Getting the best (optimal) starting point.
ii. Getting the best (optimal) direction of search.
iii. Getting the best (optimal) step length.
iv. Making a move towards the optimum; and
v. Stopping when there is no further improvement.

The algorithm finds application in many industrial experiments, cutting across several scientific fields including food science, agriculture, and process engineering. It can be used in the Maximization of oil yield from palm fruit while ensuring high oil quality; where the Palm Oil Extraction depends on Extraction Temperature (°C), Pressing Time (minutes), Palm fruit peel or Flesh (mm), and Pressing Pressure (Pa). It can be used in the minimization of Food Waste Volume (kg) generated during food processing, where the Food waste Volume depends on raw material quality, processing equipment efficiency, production time, packaging technique, and production personnel.

DETERMINIATION OF AN EXPERIMENTAL DESIGN: THE A-G-E-D OPTIMALITY CRITERIA

Some questions to ask while choosing experimental designs include: How efficient are they? Are they optimal? Do they have computational advantages? Are they economical? Are they user-friendly? Do they meet the experimental objective and goal? Are they robust under varying experimental conditions? One way to think about good experimental designs is by incorporating the A-G-E-D optimality criteria and efficiency, which have been central to my research. The popularly used regression models have inherent in them the A-G-E-D alphabetic optimality and efficiency properties. When trying to choose one of several regression models, the choice may be a function of the estimates of model parameters.

Usually in Statistics, unbiasedness and minimum variance of estimators are fundamental, ensuring estimators provide reliable and efficient results. Think of an estimator as a measuring tool such as thermometer that is used to check room temperature. A good thermometer should give readings that are on average correct (unbiased) and should not fluctuate too much when used multiple times (low variance). Unbiasedness means that if you take many measurements, the average of those measurements should be close to the actual temperature. Hence, if a thermometer always shows a slightly higher or lower temperature than the real one, it is biased. Minimum variance property means that if you measure the temperature multiple times under the same conditions, the readings should be consistent. A thermometer that gives widely different readings each time (even if its average is correct) is not ideal. In statistics, we want estimators that, like a good thermometer, provide results that are both accurate (unbiased) and consistent (low variance), so we can make reliable conclusions from data.

In experimental design, A-optimality criterion primarily focuses on minimizing the average variance of an estimator, making estimates more stable by reducing overall variation. Suppose an agricultural scientist wants to estimate the average yield increase due to a new fertilizer. He conducts an experiment with different plots of land and measures crop yields. If the estimate has high variance, the results will fluctuate widely, making it difficult to draw reliable conclusions. To get a more precise and stable estimate, the scientist needs to minimize variance in the experiment.

Variance minimization may be achieved by: (i) Employing a well-designed experiment that ensures balanced soil types, similar weather conditions, and uniform irrigation across test plots. This reduces natural variation that could inflate variance. (ii) Testing the fertilizer on more plots (concept of increased sample size). This helps reduce variance because larger samples provide more stable estimates. (iii) Grouping the plots into blocks that are based on soil type, and applying the fertilizer within each block helps control external variations, leading to lower variance. (iv) Using statistical designs and models that help optimize data collection to achieve more precise estimates with less uncertainty. The 2^k factorial designs and its fractional factorial are good choices of designs that minimize variances.

In the agricultural experiment involving crop yield, if the scientist tests the fertilizer on 10 different soil types, G-optimality ensures that no soil type has an extremely unreliable yield estimate, even if some soils naturally vary more than others. This leads to fairer comparisons and more confident predictions about how well the fertilizer works in different conditions. G-optimality thus ensures that the variance is more evenly spread across all test conditions and that no single estimate has extremely high variance. Without G-optimality, some conditions might have highly uncertain predictions, making it difficult to generalize results. By minimizing the prediction variance, the amount of uncertainty in estimating the crop yield is minimized. This keeps the worst-case variance as low as possible.

In like manner, some test conditions might give very uncertain estimates (high variance). E-optimality ensures that even the least precise estimate is as accurate as possible, preventing any test condition from having extremely high uncertainty. Unlike A-optimality, which reduces average variance, and G-optimality, which minimizes the worst-case prediction variance, E-optimality ensures that all estimates have a reasonable level of precision, avoiding extremities in any condition. It helps improve the consistency of results across all test conditions, making it easier to generalize findings about the test's effectiveness.

Of several alphabetic optimality criteria, the D-optimality criterion is the must applied in experimental design settings. It is a design criterion that helps in reducing the overall uncertainty in the estimates, thus making the results as precise as possible. D-optimality ensures that the experiment is designed to get the most accurate estimates while using the least amount of data. Instead of testing every possible condition which may be too expensive and time-consuming, D-optimality selects the best set of test conditions that provide the most useful information with least uncertainty, thus making the results as precise as possible and maximizing the reliability of results.

Researchers in the fields of Agriculture, Medicine, Food Science and Engineering can utilize D-optimality in designing experimental studies to get the most reliable results with minimal resources. Imagine a manufacturing company wishes to improve the strength of a new type of plastic. Several factors affect plastic strength, and include Material Composition, Production Temperature, Production Pressure. Instead of testing all possible combinations of these factors (which could be thousands of experiments), D-optimality enables the selection of the best set of test conditions that give the most precise results with fewer tests.

Although it requires expert planning and relies on statistical software and careful experimental design to select the best test conditions, it saves time and money, improves product quality and reduces waste. In medical research, it is a vital resource in drug testing and clinical trials where researchers do not test new drugs on every patient; they select a D-optimal set of patients that provides the best information with fewer trials. Bhattacharya (2021) emphasized the role of Central Composite Design for Response Surface Methodology and its application in Pharmacy and Engineering.

There is yet another optimality criterion, the integrated variance (IV or I) criterion, which seeks designs that minimize the average prediction variance over the experimental region. The concept of integrated variance can be viewed in clinical trial for a new drug, where interest is in knowing how effective the drug is across different patient groups classified by age, gender or some other health conditions. Each group might respond differently, and there is uncertainty (variance) in how well the drug works for each group. Integrated variance optimality is like ensuring that the clinical trial is designed in a way that minimizes this uncertainty across all groups. We employed I-Optimality principles in design characterization for second-order models (Iwundu & Israel, 2024).

RESPONSE SURFACE METHODOLOGY

The concept of Response Surface Methodology (RSM) has been studied as an effective mathematical and statistical technique for process and product optimization. It finds application in the optimization of industrial processes and products that depend on several experimental variables. Scientific fields of Process Engineering, Agricultural and Food Sciences, Chemical, and Pharmaceutical Sciences, and several many other study areas find the concept of RSM very essential.

It generally requires searching for the best value of the response variable and requires moving sequentially along the path that leads to that optimal value. The first phase of experimentation is called the screening phase and involves the screening of significant effects. The second phase of experimentation is called the optimization phase and requires exploring the vicinity of the optimal response to identify the optimal value of the Response Surface.

SCREENING DESIGNS (FIRST-ORDER DESIGNS)

In the context of the Design of Experiments (DOEs), the first stage of the experimentation focuses on identifying the most significant factors from a somewhat potentially large list of variables. Screening designs are suitably used at this stage of experimentation as they enable users to narrow down a long list of potentially important factors and the associated interactions to only a few significant ones. The goal could be to assess the effect of several independent or control factors on a response factor in an attempt to establish the significance of the main factor effects and the interaction effects. At the screening stage of experimentation, the two-level factorial designs and its fractional factorial designs are readily available screening designs that are known to be very informative, efficient, and quite simple to use. They enable one to study very conveniently the joint effects of several input variables on the response and assume that the input (independent) variables and the output (response or dependent) variables are linearly related.

As an example, an engineer supposes that temperature, pressure and time contribute to the Haber Process for ammonia synthesis. He studies the factors at {Temperature: 400-500°C}; {Pressure: 200-300 atm}; {Time: 5 to 10 minutes per pass}. These specifications define the region of the experimentation based on the factors, and he needs to take observations under the defined experimental conditions such that measurements do not vary greatly. He needs the optimal settings of temperature, pressure and time that result in the maximum yield of ammonia. Moreover, if there is any factor that does not significantly contribute to ammonia synthesis, he needs it screened out.

The recommended design is the 2^k factorial design or its fractions called the fractional factorial designs. These designs are efficient and minimize the variances of the parameter estimates, while giving the optimal performance index. Moreover, they are convenient designs useful when there are many factors that may be considered to affect a response and further help to screen out factors in an experiment that do not contribute to an optimal response.

In real-world situations, true relationships between response variables and a set of independent variables are usually unknown. Suggestions are to approximate them with low-order polynomials which are linear models of first or second order, using straight-line graphs or quadratic graphs, such as polynomials of first and second degrees, respectively. These polynomial are models which can be described using equations or graphs. The best model has optimal settings of input or independent variables. To obtain such a best model, one must aim to get the set of values of input variables, that best describe the process under study.

APPLYING OPTIMAL EXPERIMENTAL DESIGN TO SCIENTIFIC PROBLEMS

The necessity to apply optimal experimental design to any scientific problem leads the researcher to assume the burden of a deeper knowledge of science. This is the role of optimality.

FISH PRODUCTION EXPERIMENT

We studied the effect of feed types on the growth of catfish (CLARIAS-Gariepinus) using a two-level optimal factorial design in four factors (Omosioni & Iwundu, 2012). The two-level factorial design requires that for the experiment containing four factors, each factor should be studied at two levels. This makes it possible to assess the effect of the four factors and their interactions with some level of ease. The two-level factorial design being a first-order design finds application in many scientific experiments when it could be assumed that the relationship between a response variable and a set of independent variables can be modelled by a first-order polynomial (a linear equation).

The experimental factors were four feed types, namely Sharp (S), Coppens (C), Euro (E), and Dizengoff (D), and the experiment was over a period of four weeks. Each factor (feed type) was studied at a high and a low level and resulted in 16 treatment combinations. The idea was to obtain the feed combination best suited to produce healthy fish within a duration of four weeks. The complete experiment was replicated three times thus allowing three observations to be taken at each treatment combination. The optimal feed combination, comprising 5 grams of Sharp (S), 10 grams of Coppens (C), 10 grams of Euro (E), and 10 grams of Dizengoff (D), impacted on the feed conversion efficiency, giving rise to massive fish size that addresses food needs.

Plate 1: Harvest of Catfish (CLARIAS-Gariepinus)

The experimental design assumed an initial 11-parameter screening model that described the presence of the main factor effects and their interactions. After appropriate statistical tests were carried out, the model was reduced to 5-parameter model. It is worth noting that the reduced model does not follow a hierarchical structure where items are linked to each other in a type of parent-child relationship and requires that "if we include an interaction in a model, we should also include the main effect, even if the p-value associated with their coefficients are not significant". Hence the final model contained 7-parameters. Unfortunately, this model has no standard experimental design. Fortunately, too, we have proffered experimental design solutions to cases of reduced non-standard models (Iwundu & Israel, 2024).

FRESH MELON SEED EXPERIMENT

We have demonstrated experimentally that A-G-E-D optimality properties bear a relationship to fit statistics such as R^2 statistic, $R^2_{Adjusted}$ statistic, C_p statistic, PRESS, AIC, etc. We examined the Drying Characteristics of fresh melon seed in relation to Time, Temperature and Moisture Content (Iwundu & Omosioni, 2015). Specifically, in modeling the drying characteristics of fresh grains (melon seed), the stepwise procedures that require sequentially selecting the best models were utilized. Unbiased estimates for the best drying conditions were determined by using suitable models. The adequacy of the selected models was assessed using standard and known assessment criteria and fit statistics namely as R^2 statistic, $R^2_{Adjusted}$ statistic, C_p statistic, PRESS, AIC, and MSE.

D-optimality criterion was introduced as a new assessment criterion for establishing the adequacy of the regression models, and its performance was compared with those of the standard assessment criteria. D-optimality criterion was successfully utilized as a new criterion for assessing the adequacy of regression models. Its discriminating ability is as with $R^2_{Adjusted}$ statistic, C_p statistic, PRESS, and AIC. It is important to highlight that these statistics are model validation techniques and processes intended to verify that the models are performing as expected. They find application in Machine

Learning, a branch of Artificial Intelligence. For more details on A-G-E-D optimality properties and fit statistics, see the publications of Iwundu and Omosioni (2015) and Iwundu and Cosmos (2022), that are related to the modeling and optimization of industrial processes and products that depend on several experimental variables.

ADVANCES IN OPTIMAL DESIGN OF EXPERIMENTS: COMBINATORIAL PROCEDURE FOR CONSTRUCTING D-OPTIMAL DESIGN

Knowing the importance of experimental design strategies (like D-optimality) in minimizing parameter variances and covariance, we obtained combinatorial procedure useful in the construction of D-optimal designs for any design size (Onukogu & Iwundu, 2007). The procedure is iterative, flexible, adaptable to any design region, and can be easily implemented using most programming languages. Using the combinatorial technique, it is easy to determine all the designs that are concurrently D-optimal. The experimenter can therefore choose one of these designs based on convenience and minimality of cost. This helps to enhance design efficiency and accuracy.

The combinatorial procedure is adaptable for other optimality criteria and has been implemented on E-optimality criterion in an unpublished M.Sc. Dissertation (Ogbonna, 2010). An application area of the procedure is in investment, requiring an optimal guide on how much to invest in shares, treasury bill, money market and commercial papers, to get an optimal returns on investment. This is achievable with the aid of Statistics, as among the eloquent proofs that Statistics is in your neighbourhood.

ADVANCES IN OPTIMAL DESIGN OF EXPERIMENTS: MULTI-RESPONSE OPTIMIZATION

We have made contributions in Multi-Response Optimization (MRO), which focuses on simultaneously optimizing multiple dependent responses, that have different goals, by balancing trade-offs among them thereby making compromises between the goals in view of finding the best overall solution. For example, in manufacturing, increasing production speed may lead to a decrease in product quality. Multi-Response Optimization allows trade-offs among the two goals towards effective optimization. We proposed two optimization algorithms that depend on experimental design principles of minimal variances while optimizing multi-response functions. The algorithms are the Variance-Weighted Gradient Algorithm and Variance Weighted Gradient Projection (Otaru & Iwundu, 2017a; Otaru & Iwundu, 2017b; Iwundu & Otaru, 2019). These algorithms require finding the best settings of input variables that optimize multiple output variables.

Since it is usually impossible to optimize all responses, the trade-offs help in finding a compromise solution that provides the best possible balance among the different objectives. The multi-response algorithms are alternative techniques to known gradient and non-gradient optimization techniques, encountered in engineering, food and drugs processes. It is interesting to note that the Variance-Weighted Gradient Algorithm and the Variance Weighted Gradient Projection algorithm employ experimental design principles based on statistical variance, and simultaneously optimize several response surfaces, and serve as competitive alternative to popularly used one-at-a-time optimization methods.

The methods rely on the general line search equation whose components are the starting (initial) point of search, direction of search, step-length and the point arrived at the j^{th} iteration. At an iterative move, the optimizer(s) reached are successively added to the previous immediate design measure(s). The use of projection operator scheme that allows the projection of design points from one design space to another is employed. This ensures that the solutions stay within the feasible region of experimentation. An application is in the optimization of drug formulations to achieve maximum effectiveness while simultaneously minimizing side effects.

EXPLORATION OF THE VICINITY OR REGION OF OPTIMALITY

In searching for the optimal value of a response, the search for the optimum advances through paths of improvement with the first-order model being used at the initial stage of experimentation. As the search advances and exhibits a lack of fit of the first-order model, there is the to explore the region of experimentation further using the second-order model (that can measure curvature) around the vicinity of optimality. This helps to locate the optimum more precisely.

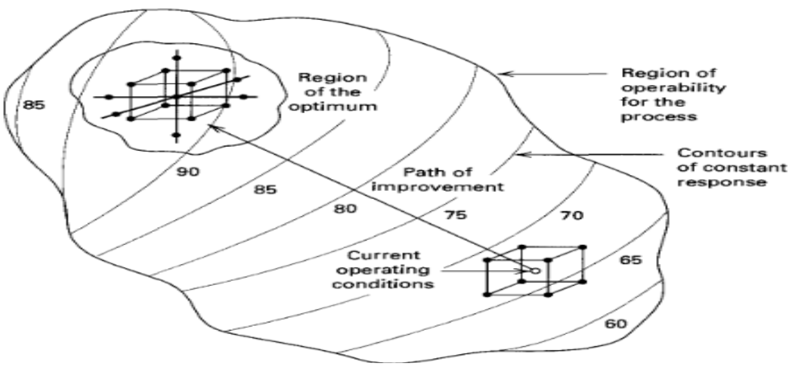

Figure 9: Region of Optimal Exploration

To save cost of experimentation yet maintain efficiency, Central Composite Designs (CCDs) and Box-Behnken Designs (BBDs) are some experimental designs that are may be used to explore the vicinity of optimality. These designs are useful in the field of Response Surface Methodology (RSM) and serve as great design choices in building a second order (quadratic) model for the response variable without needing to use a complete three-level factorial experiment. They require fewer experimental runs than would the complete three-level factorial design and help researchers study the joint effects of more factors conveniently. They are the most applied second-order designs for optimization of processes cutting across food science, engineering and numerous industrial processes.

Central Composite Designs are classified into three basic forms, namely, Circumscribed, Face-centered and Inscribed Central Composite Designs, depending on the experimental region.

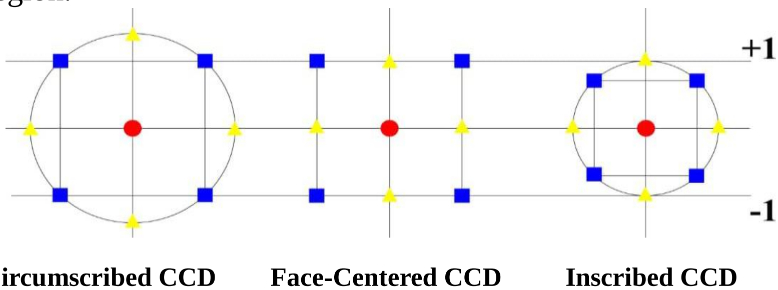

Figure 10: Circumscribed (Rotatable), Face-Centered and Inscribed CCDs in two variables

The effect of D-optimality criterion in the construction of optimal designs that are defined on the design regions of the face-centered central composite design, rotatable (circumscribed) central composite design and inscribed central composite design, respectively, was investigated using a second-order response surface model. For the six-parameter second-order polynomial model used, the D-optimal design defined over the rotatable (circumscribed) Central Composite Design (CCD) region had better determinant values than those obtained over the face-centered central composite design region and the inscribed central composite design region. These results are indicative that D-optimal designs defined over the rotatable (circumscribed) CCD region give better and more precise parameter estimates as the variances and covariances of the parameters are minimized (Iwundu & Otaru, 2014). The Rotatable Central Composite Designs offered better prediction across the design space and hence are better designs for modeling quadratic effects. Useful Numerical Statistics for some response surface designs have been documented in Iwundu (2016a).

Box-Behnken designs are some forms of composite designs. Unlike the central composite designs that utilize the vertex (corner) points, Box-Behnken designs do not utilize the vertex (corner) points of the design region. They are a class of efficient designs with a reduced number of experimental runs and are desirable when avoiding extreme conditions, making it useful when extreme values might be infeasible or unsafe. Figures 11 and 12 are geometric views of a Central Composite Design and a Box-Behnken design, respectively. Experimental designs like Box-Behnken and Central Composite Designs help researchers optimize processes with minimal experiments. An application is in the problem requiring the efficiency of Arsenite removal from groundwater (Iwundu & Cosmos, 2022).

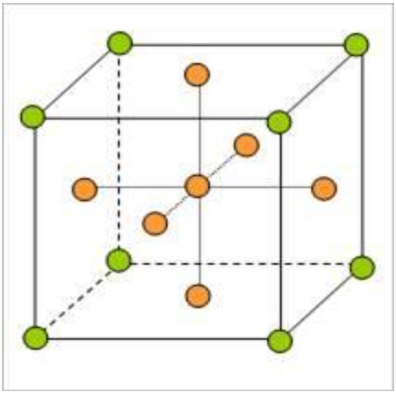

Figure 11: Face-Centered CCD in three Variables

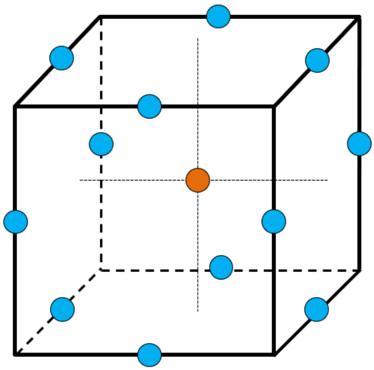

Figure 12: Box-Behnken Design in three Variables

SPECIAL ISSUES IN OPTIMAL DESIGN OF EXPERIMENTS: REDUCED MODELS

In most practical modeling situations involving several control variables, the CCDs and BBDs have been satisfactorily used in the optimization phase of experimentation. However, the second-order model which justifies the use of these designs is assumed prior to experimentation. After data have been collected, model fitting may reveal that not all model parameters are significant and hence the need to remove insignificant parameters from the model thus resulting in a reduced posterior model. We proposed the Modified Central Composite Designs (MCCDs) for non-standard reduced models (Iwundu, 2018). The MCCDs utilize the Hat Matrix principle, a concept in Statistics, Linear Algebra and Regression Analysis that helps us understand how observed data points influence the predicted values in a model. The Hat matrix served as an effective scientific tool for the construction of designs that led to increase in design efficiency, and offer alternative designs for non-standard models, that are similar to the structure of CCDs, yet of reduced experimental runs (Iwundu, 2018).

It was interesting to employ the “Hat Matrix” concept as a means of constructing designs as this was never in the properties and capacities of its usage. Although CCDs and BBDs are reasonably robust to model misspecifications, the MCCDs require fewer design points, are economically efficient, have proven to be better design choices than the CCDs in minimizing the maximum variance of the predicted values, and have a competitive advantage in minimizing the variance and co-variances of parameter estimates. Furthermore, CCDs and BBDs assume that the model is of the standard type and that the design region is regular. The use of CCDs and BBDs for non-standard models and irregular experimental regions is a violation of the basic assumptions.

SPECIAL ISSUES IN OPTIMAL DESIGN OF EXPERIMENTS: NON-OPTIMAL DESIGN POINTS

The presence of non-optimal design points in an experimental design measure greatly affects the convergence of a search algorithm to a desired optimum. It is seen as a misfit or "odd one out”. In the world of Statistics, this describes an outlier, a value in a dataset that is very different from the other values.

We proposed Filtering and Re-construction as a viable procedure for sequentially locating D-optimal design measures. The method effectively improves experimental design in the search for an optimal design measure. The filtering and re-construction technique, a back-to-back procedure addresses situations where outlying non-optimal design points had been admitted into the design possibly either by the creation of a poor initial design or by its influence on the next design point(s). By the method, outlying non-optimal design points are removed and the design reconstructed, thus resulting in a significant improvement on the performance of the design as measured by the determinant value of design’s information matrix (Iwundu, 2016). The algorithm for the technique is certain to converged to the desired optimum.

SPECIAL ISSUES IN OPTIMAL DESIGN OF EXPERIMENTS: CONVERGENCE AT NON-GLOBAL OPTIMUM

When searching for optimizers of a response function, some algorithms fail to proceed to the optimal solutions, but converge only locally, within a small neighbourhood of the input variables. Sometimes, the algorithms fail to converge and rather cycles around the same non-optimal solutions or converge very slowly, with very slow progress. We proffered a “Unique Solution” to the problems associated with convergence of exchange algorithms (Iwundu, 2010). By this approach, it is possible to obtain designs with the best characteristics.

SPECIAL ISSUES IN OPTIMAL DESIGN OF EXPERIMENTS: MISSING OBSERVATIONS

Cases of missing observations are some unavoidable practical issues in many experiments as it is not certain that all responses of experimental trials would be realized during experimentation. Also, encountering missing data points is not far-fetched as an instrument malfunction could lead to a data point that is not consistent with the other data points thus resulting in an outlying observation. During the data validation and cleaning, the outlying observation may be discarded from further consideration in the analysis of the experiment, thus resulting in some missing observations. Sometimes also, a treatment combination may not be feasible due to certain constraints in the system under study. Unfortunately, and regardless of the form, missing data points could grossly affect the statistical power of a test, offer biased estimates of parameters, and give invalid conclusions.

Due to the frequent use of the Central Composite Design in experimental studies involving second-order optimization, we examined the effect of missing observations in the compartments of the Central Composite Design and established that higher losses in Relative A- and D-efficiencies are attributed to missing vertex (factorial) points. Relative G-efficiency is mildly affected by the missing vertex or axial point or both. Relative A-, D- and G-efficiency were unaffected by missing center points (Iwundu, 2017). The practical importance of these results is that there is a need to access recovery techniques for the treatment of missing observations that grossly affect the design's efficiency. Such recovery techniques include imputation and estimation methods.

SPECIAL ISSUES IN OPTIMAL DESIGN OF EXPERIMENTS: IRREGULAR EXPERIMENTAL REGION

Much scientific and industrial research is applicable to experimental regions with regular geometry that are scaled to cubes or spheres. Also, standard statistical designs such as factorial and fractional designs, as well as the popularly used Central Composite designs and Box-Behnken designs in many scientific and industrial research, are only applicable to experimental regions that can be scaled to cubes or spheres. Unfortunately, irregular regions occur in real world situations. It could be a result of a sudden voltage surge that might cause a system to react unpredictably, thus altering the experimental conditions.

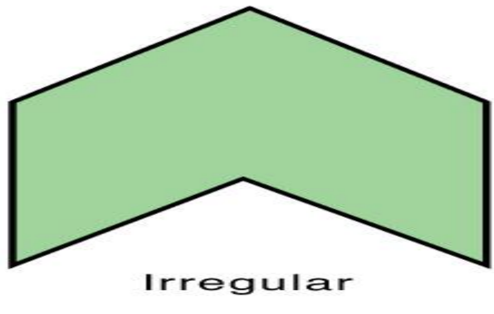

Figure 13: An Irregular Region

For geometric experimental regions that may not be scaled to cubes and spheres and whose lengths and interior angles may vary, we have provided techniques that allow the construction of optimal designs for the constrained or irregular experimental regions (Iwundu & Ebong, 2014). We have further established that rather than force a standard design into an irregular region, Custom and Computer-generated designs should be utilized (Iwundu & Israel, 2024). These designs are efficient and tailored to meet specific research objectives, constraints, and available resources.

Although the use of Central Composite Design (CCD) and Box-Behnken Design (BBD) is common practice in experimental sciences requiring the design of experiments, sometimes the designs are still cost-demanding and the need for economical designs becomes ideal. For varying problems, we recommend using computer-generated and custom designs or some existing economical designs including Small Composite Design (SCD), Hoke design, Koshal design, and Hybrid design. In a recent study, we have established that Computer-Generated Designs, Custom Designs and Hybrid designs outperform the CCDs and BBDs (Iwundu & Nwuche, 2023 and Iwundu & Israel, 2024).

Advanced experimental design software provides algorithms for construction of custom designs, namely, Custom A-, D-, and I-optimal designs. We examined the performances of Custom A-, D-, and I-optimal designs on non-standard second-order models using the alphabetic A-, D-, and G-optimality efficiencies, as well as the Average Variance of Prediction. The results reveal that Custom-A optimal designs perform generally better in terms of G-efficiency. They show high superiority to A-efficiency as the worst G-efficiency value of the created Custom-A optimal designs exceeds the best A-efficiency value of the designs and does well in terms of D-efficiency. Custom-D optimal designs perform generally best in terms of G-efficiency, as the worst G-efficiency value exceeds all A- and D-efficiency values. Custom-I optimal designs perform generally best in terms of G-efficiency as the worst G-efficiency value is better than the best A-efficiency value and performs generally better than the corresponding D-efficiency values. For the Average Variance of Prediction, Custom A- and I-optimal designs perform competitively well, with relatively low Average Variances of Prediction. On the contrary, the Average Variance of Prediction is generally larger for Custom-D optimal designs. Hence when seeking designs that minimize the variance of the predicted response, it suffices to construct Custom A-, D-, or I-optimal designs, with a preference for Custom-D optimal designs (Iwundu & Israel, 2024).

SPECIAL ISSUES IN OPTIMAL DESIGN OF EXPERIMENTS: LACK-OF-FIT

Modeling is used to establish relationships between variables and starts by assuming a linear relationship between variables. An issue in modelling is the Lack-of-fit of a chosen model (Seshu-Babu et al. 2015). Lack-of-fit occurs when the chosen model does not accurately describe the real pattern in the data, leading to errors or incorrect predictions. We have presented efficient sequential and non-sequential third-order designs that are useful when there is lack of fit of second-order model in establishing relationships between variables (Iwundu & Agadaga, 2021, Iwundu & Agadaga, 2022). The construction of the sequential and non-sequential third-order response surface designs utilized the principles of the Hat Matrix. The designs yielded high efficiency values using the D- and G-optimality criteria.

HAT-MATRIX AIDED COMPOSITE DESIGNS FOR SECONDS-ORDER MODELS

Hat-Matrix aided composite designs, comparable with standard response surface methodology designs and Computer-Generated designs for Seconds-Order models were constructed and presented alongside their optimality and efficiency properties. The designs depend on the principles of the loss function. Through the Hat-matrix, design points that enhance efficiency of second-order designs were selected. Unlike computer-generated designs which may not be unique for a specific model and which may present some less efficient designs, the Hat-Matrix aided designs are unique and require at least two categories of discrete design runs formed from the complete 3^k factorial design runs, only on the basis of the diagonal elements of the hat matrix that promote maximizing determinant of the information matrix thereby minimizing the loss function (Iwundu & Otaru, 2019).

THE MEETING OF EXPERIMENTAL DESIGN WITH OPERATIONS RESEARCH: ARE THE TWO MARRIED?

Marriage is imagined to be a contract between people. It is interesting how marriage can be made a central concept to Statistics. Experimental Design, as an integral part of Statistics, seeks to ensure experiments are well-planned and data are reliable. Operations Research, on the other hand, seeks to optimize everything for efficiency. The question to ask is: Are the two married?

The answer to this question is in the affirmative: Experimental design principles have demonstrated applicability in solving optimization problems commonly encountered in the field of Operations Research. We established a relationship between the feasible region of a Linear Programming (LP) problem and the design region of an optimal design problem, and we obtained solutions to the LP problems using experimental design principles associated with minimum variance and minimum mean squared error. For a given LP problem, we carefully select design points (coordinates) from the feasible region (called design region in the context of Design of Experiments), obtain an optimal starting point of search, an optimal step length and an optimal direction of search using variance minimization techniques. Through these engagements, we have successfully married the concepts of design planning and optimization (Odiakosa & Iwundu, 2013; Iwundu & Hezekiah, 2014; Iwundu & Ebong, 2014; Iwundu & Ndiyo, 2015; Iwundu & Ekpo, 2015). By this marriage, we establish that good research is a perfect balance of planning and optimization!

CONCLUSION

Mr. Vice-Chancellor Sir, Statistics is more than numbers. It stands as the unseen force in every good decision. It is a dependable tool, but its reliability depends on the quality of data. Today's world is the world of data and statistical data are the foundations of accountability and the lifeblood of decision-making. Statistics has been successfully utilized in human and environmental studies and is a very core part of Artificial Intelligence and an inherently necessary component of Data Science. Statistics is essential for all disciplines. Where there is no data, there is no Statistics. Where there is no Statistics, the world becomes stagnated. It is indeed the Man in your neighbourhood.

While thinking about the place of Statistics in the Bible, I came across a blog that says GOD LOVES STATISTICS (Published February 9, 2024 by Neil Elliot). It emphasized the importance of hundreds and thousands, serving as significant words in the Bible. As documented, the number *thousand* appears over 550 times, and the number *hundred* appears almost 700 times and are more than any other numbers and more than most words. They appear in the Old Testament mainly in the books of Moses and the history books (Joshua to Esther). In the New Testament, references to hundreds and thousands though occasional, appear in the book of Revelation as the majority user of *thousands*, both for the numbers of people and years. Truths are being established with these statistics, particularly the ability of God to bless his people. These also are proofs that Statistics is everywhere.

2 Peter 3:8 says, “One day is with the Lord as a thousand years, and a thousand years as one day”. This shows that God operates the Equivalence Theorem, a concept in Optimal Experimental Design that states that a design is G-optimal if it is D-optimal. It is interesting how statistics and experimental design play out in human and material environments, as well as the Divinity. God loves Statistics and every lover of God should love Statistics too. Since Statistics is in your neighbourhood, make Statistics your friend.

Mr. Vice-Chancellor Sir, Principal officers, Professors, students and esteemed audience, permit me to end this very long lecture with a repeat of the saying of John Tukey: “The best thing about being a statistician is that you get to play in everyone’s backyard". I hope I played in your backyard. Let Statistics be your friend.

FUTURE RESEARCH FOCUS

There are many of our designs that have been tested only theoretically. The utilization of these designs in field experiments with engineers, agronomists, food scientists, and several other industrial partners will give empirical support, confirming their effectiveness and efficiency.

I recently sought windows of collaboration with Biochemists in an experimental study of Adaptive Laboratory Evolution (ALE) of yeast_*Spathaspora passalidarum* towards the optimization of Bioethanol yield. We are discussing the application of optimal experimental design techniques for the optimization of Bioethanol yield. I am also collaborating with engineers in the field of Renewable Energy on the utilization of optimal design principles as effective tools for the production of Biodiesel, a renewable and biodegradable fuel derived from organic sources.

We also aim to establish an intersection of experimental design and machine learning, which will help in applying optimal design techniques to emerging areas of Science, Engineering, and Bioinformatics. Robustness of designs shall have a central focus in establishing designs that remain effective under changing assumptions and conditions.

Integration of the Bayesian methodology into experimental design is of interest as it will help to incorporate prior knowledge of the designs while updating the designs and models. We also hope to explore advanced methodologies in mixture experiments to optimize component interactions in complex systems. All these require collaboration, professional networks, and funding opportunities.

RECOMMENDATIONS

Mr. Vice-Chancellor Sir, the man in your neighbourhood is not just a statistician, he is a catalyst for progress. We are in the information age, where knowledge is power, and everyone should strive to stay informed, analytical, and future-ready. Today, big data analytics has become a standard tool in predictive analysis, widely applied across management, marketing, advertising, and various economic sectors. For effective data-driven decisions, we should encourage and develop core capabilities in data proficiency and analytics agility.

There is a need to promote data use for well-articulated data-driven decisions. For effective data articulation, we need framework for data ethics and governance, and a comprehensive data management system that outlines the moral principles that guide how data are collected, stored, analyzed, and used, and provides standards that direct collective action around the safe use of data, avoiding privacy invasion and human rights abuse. Research institutes and professional organizations should endeavour to clearly define expectations for researchers, and policies on the use of research data, as this will help ensure the safety of all human subjects and uphold data integrity.

In promoting data use, there is need to come up with methodologies that can handle the vast amount of evolving data for effective data-driven decisions bearing in mind that data is useless in and of itself unless it significantly and widely contributes to informing important public and private decisions. We should incorporate Statisticians and Data Scientists in every data sector to advance the proper utilization of data for evidence-based planning, policy formulation, and innovation. Their expertise ensures that data is not only collected, but also meaningfully utilized to drive impactful decisions and sustainable development.

Seeing the central part Statistics plays in scientific and non-scientific research, and in the totality of livelihood, and understanding its essence to Data Science, Machine Learning, Managerial Science, Artificial Intelligence, Business Intelligence, and other computational fields, I wish to implore that Statistics be taught as a university-wide course. This will curb the abuse in the use or misuse of Statistics and will amount to “THE STATISTICS OF THE PEOPLE BY THE PEOPLE AND FOR THE PEOPLE”.

POSTHUMOUS DEDICATION

Borrowing the saying of George Washington “All I am I owe to my mother. I attribute my success in life to the moral, intellectual and physical education I received from her. I am indebted to the memory of my dear mother Late Mrs. Cornelia O. Usoro, a woman that gave her all at ensuring I read and write and took it to the level of being my personal home teacher. I dedicate this lecture to her loving memory. She was flawless and priceless. **REST ON, MOM!**

APPRECIATION

Today would not have been successful without everyone seated here. My heart is filled with gratitude. I give all glory to God. I thank the Vice-Chancellor and the Management Team of the University. I belong to a faculty; I thank my Dean, Professor Chukwuemeka Ehirim and all faculty members. My department is home for me, I thank my Head of Department, Prof. Bazuaye Frank Etin-Osa and all staff and students of the Department of Mathematics and Statistics.

A very loud appreciation goes to my Committee of Friends—my backbone. You are dependable allies. Nothing is more precious than having true friends. Thank you for your unceasing commitment and unfailing love. I appreciate my dedicated staff at the Information and Communication Technology Centre (ICTC). You are a true family. My sincere appreciation also goes to the staff at the Intellectual Property and Technology Transfer Office (IPTTO); You never failed me throughout my service there. Thank you for helping us uphold work ethics and integrity.

I appreciate the AKWACOM, Igbo Progressive Group, Umu Mbieri, UPWA, OWSD, friends, colleagues and the entire university community who considered it necessary to attend this Inaugural Lecture Series. I appreciate my pastors and members of the church community. You’re such a shoulder to lean on. I appreciate Statisticians and all users of Statistics. I lose count; kindly permit me to thank everyone. You all are amazing.

To my family, moments with you are deeply cherished. Your love, support, and understanding have been my anchor through every phase of this journey. Thank you for being my source of strength and joy. To my in-laws, thank you for your unwavering love, encouragement, and prayers. Your support means the world to me.

To our esteemed online viewers, thank you for joining this special moment from afar. Your virtual presence is truly appreciated. Let us all embrace statistical thinking and recognize statistics as that man in your neighbourhood, working quietly yet powerfully to make sense of our world.

Thank you all, and may God bless you.

REFERENCES

Adegboye, O. O., & Olaniyi, T. A. (2021). Demographic data accuracy and economic development in Nigeria: Challenges and prospects. African Journal of Economics and Policy Studies, 15(2), 45-62.

Akinola, I. (2024, February 3). *Global Hunger Index 2023 indicates serious hunger level in Nigeria*. Tribune Online. https://tribuneonlineng.com/global-hunger-index-2023-indicates-serious-hunger-level-in-nigeria/

Ali, Z., & Bhaskar, S. (2016). Basic statistical tools in research and data analysis. *Indian Journal of Anaesthesia*, *60*(9), 662. https://doi.org/10.4103/0019-5049.190623

Bhattacharya, S. (2021). Central composite design for response surface methodology and its application in pharmacy. In P. Kayaroganam (Ed.), *Response Surface Methodology in Engineering Science*. IntechOpen.

Elliot, N. (n.d.). *God loves statistics – NumbersMatters*. Numbersmatters.Ca. Retrieved July 28, 2024, from https://numbersmatters.ca/2024/02/09/god-loves-statistics/

Eze, C. N., & Yusuf, M. B. (2020). Population statistics and electoral representation in Nigeria: A case study of census controversies. Nigerian Journal of Political Economy, 28(1), 33-50.

Geertsema, J. C. (1987). A Christian View of the Foundations of Statistics. *Perspectives on Science and Christian Faith* 39.3, 158-164.

Global hunger index (GHI). (n.d.). Who.int. Retrieved October 17, 2022, from https://www.who.int/data/nutrition/nlis/info/global-hunger-index-(ghi)

How much do data scientists need to know about statistics? (n.d.). School of Data Science. Retrieved

July 28, 2024, from https://datascience.virginia.edu/news/how-much-do-data-scientists-need-know-about-statistics

Iwundu, M. P. (2016a). Useful Numerical Statistics of Some Response Surface Methodology Designs. *Journal of Mathematics Research*, 8(4), 40-67. doi:10.5539/jmr.v8n4p40 URL: http://dx.doi.org/10.5539/jmr.v8n4p40.

Iwundu, M. P. (2016b). Design filtering and reconstruction: A procedure for sequentially locating D-optimal design measures. *International Journal of Statistics and Applications*, *6*(5), 300–308. http://article.sapub.org/10.5923.j.statistics.20160605.04.html

Iwundu, M. P. (2017). Missing observations: The loss in relative A-, D- and G-efficiency. *International Journal of Advanced Mathematical Sciences*, *5*(2), 43-49. https://doi.org/10.14419/ijams.v5i2.7786

Iwundu, M. P. (2018). Construction of modified central composite designs for non-standard models. *International Journal of Statistics and Probability*, *7*(5), 95. https://doi.org/10.5539/ijsp.v7n5p95

Iwundu, M. P., & Agadaga, G. O. (2021). Sequential third-order response surface designs. *Journal of the Nigerian Statistical Association*, *33*, 62–81.

Iwundu, M. P., & Agadaga, G. O. (2022). Five-level Non-Sequential third-order response surface designs and their efficiencies. *African Journal of Mathematics and Statistics Studies*, *5*(1), 14–32. https://doi.org/10.52589/AJMSS-IJYZJB2C

Iwundu, M. P., & Cosmos, J. (2022). The efficiency of seven-variable box-Behnken experimental design with varying center runs on full and reduced model types. *Journal of Mathematics and Statistics*, *18*(1), 196–207. https://doi.org/10.3844/jmssp.2022.196.207

Iwundu, M. P., & Ebong, D. W. (2014). Modified Quick Convergent Inflow Algorithm for solving Linear Programming Problems. *International Journal of Statistics and Probability,* 3(4), 54-66.

Iwundu, M. P., & Ekpo, S. U. (2016). Rules for identifying the initial design points for use in the Quick Convergent Inflow Algorithm. *International Journal of Statistics and Probability*, 5(1), 98-110.

Iwundu, M. P., & Hezekiah, J. E. (2014). Algorithmic approah to solving Linear Programming problems on segmented regions. *Asian Journal of Mathematics and Statistics,* 7(2): 40-59.

Iwundu, M. P, & Israel, C. (2024). *Performance Evaluation of Custom A, D, and I optimal designs for non-standard models.* American Journal of Theoretical and Applied Statistics.

Iwundu, M. P., & Ndiyo, I. E. (2015). Comparative study of the Quick Convergent Inflow Algorithm (QCIA) and the Modified Quick Convergent Inflow Algorithm (MQCIA). *International Journal of Statistics and Probability*, 4(2), 95-112.

Iwundu, M. P., & Nwoshopo, D. E. (2022). Efficiency of modified central composite designs with fractional factorial replicates for five-variable nonstandard models. *Scientia Africana*, *21*(1), 195–206. https://doi.org/10.4314/sa.v21i1.17

Iwundu, M. P., & Nwuche, O. J. (2023). *Behaviour of some economical designs on second-order effects model and their relative efficiency* (Unpublished Masters Dissertation). University of Port Harcourt.

Iwundu, M. P., & Omosioni, P. E. (2015). On the adequacy of variable selection techniques on model building. *Asian*

Journal of Mathematics & Statistics, *8*(1), 19–34. https://doi.org/10.3923/ajms.2015.19.34

Iwundu, M. P., & Onisokumen, D. (2016). The effect of Sample Size Selection on Type I Error, Type II Error, Power of the test and P-value. *International Journal of Advancement in Physical Science, 8(*1), 174-178.

Iwundu, M. P., Otaru, O. P. & Onukogu, I. B. (2014). Simultaneous Optimization of Response Surfaces: Using Variance Weighted Gradient. *Journal of Knowledge Management, Economics and Information Technology***,** 4(1), 168-174.

Iwundu, M. P., & Otaru O. A. P. (2019). Construction of hat-matrix aided composite designs for Seconds-order models. *American Journal of Computational and Applied Mathematics*, *9*(3), 62–84. http://article.sapub.org/10.5923.j.ajcam.20190903.03.html

Iwundu. M. P., & Josiah, N. (2024). Exploratory Data Analysis: Demand for pregnancy test among youths aged 15 – 25 [Unpublished survey]. University of Port Harcourt.

Mekonen, E. G. (2024). Pooled prevalence and associated factors of teenage pregnancy among women aged 15 to 19 years in sub-Saharan Africa: evidence from 2019 to 2022 demographic and health survey data. *Contraception and Reproductive Medicine*, *9*(1). https://doi.org/10.1186/s40834-024-00289-5

Mental health disorders and their impact on Nigerian youth. (n.d.). *Slni.org. Retrieved August 2, 2024, from https://slni.org/mental-health-disorders-and-their-impact-on-nigerian-youth/*

More Nigerians to fall into extreme poverty, hunger, says world bank report. (n.d.).

Saharareporters.com. Retrieved October 4, 2022, from https://saharareporters.com/2022/06/12/more-nigerians-fall-extreme-poverty-hunger-says-world-bank-report

Nigeria's hunger level ‘serious,’ ranks 103 out of 121 countries. *(2022, October 15). Premium Times. Retrieved May 16, 2025, from https://www.premiumtimesng.com/news/headlines/559880-nigerias-hunger-level-serious-ranks-103-out-of-121-countries.html*

Odiakosa, O., & Iwundu, M. P. (2013). A Quick Convergent Inflow Algorithm for solving Linear Programming Problems. *Journal of Statistics Applications and Probability*, 2(2), 103-114.

Ogbonna, O. J. (2010). *Construction of E-optimal designs using a combinatorial method* (Unpublished master’s dissertation). Department of Mathematics and Statistics, University of Port Harcourt, Nigeria.

Okonkwo, J. C. (2019). Economic consequences of inaccurate population data in Nigeria. Journal of Development Economics, 12(4), 78-94.

Olawoyin, O. (2022, March 30). *Number of poor people in Nigeria to reach 95 million in 2022 – World Bank*. Premium Times Nigeria. https://www.premiumtimesng.com/business/business-news/520849-number-of-poor-people-in-nigeria-to-reach-95-million-in-2022-world-bank.html

Omosioni, P. E., & Iwundu, M. P. (2012). The Effect of Feed Type on the Growth of Catfish (Claria-Gariepinus) Using a 2^k Factorial Design. *SCIENTIA AFRICANA*, 11(2), 65-73.

Onukogu, I. B., & Iwundu, M. P. (2007). A combinatorial procedure for constructing D-optimal exact designs. *Statistica*, *67*(4), 415–523.

Otaru, O. A. P., & Iwundu, M. P. (2017a). The Variance Weighted Gradient Projection (VWGP): An alternative optimization approach of Response Surfaces. *International Journal of Statistics and Applications*, *7*(2), 57-72. DOI: 10.5923/j.statistics.20170702.02.

Otaru, O. A. P., & Iwundu, M. P. (2017b). Competitive assessment of two Variance Weighted Gradient (VWG) methods with some Standard Gradient and Non-gradient Optimization Methods. *International Journal of Statistics and Applications*, 7(2), 73-92. DOI:10.5923/j.statistics.20170702.02.

Psychologist admits faking data in dozens of studies. (2011, November 1). *New Scientist*. Retrieved May 16, 2025, from https://www.newscientist.com/article/dn21118-psychologist-admits-faking-data-in-dozens-of-studies/

Rapheal. (2021, August 1). *Women speak 20,000 words a day; 13,000 more than men*. The Sun Nigeria. https://thesun.ng/women-speak-20000-words-a-day-13000-more-than-men/

Romano, R., & Gambale, E. (2013). Statistics and Medicine: The indispensable know-how of the researcher. *Translational Medicine @ UniSa*, *5*, 28.

UNICEF. (2021, October 5). *Impact of COVID-19 on poor mental health of children and young people 'tip of the iceberg' – UNICEF*. UNICEF Nigeria. https://www.unicef.org/nigeria/press-releases/impact-covid-19-poor-mental-health-children-and-young-people-tip-iceberg-unicef

United Nations, Department of Economic, Social Affairs, & Statistics Division. (n.d.). *United nations world data forum*. Unstats.un.org. Retrieved July 28, 2024, from https://unstats.un.org/unsd/undataforum/about/

University of Virginia. (2021, May 4). *How much do data scientists need to know about statistics? https://datascience.virginia.edu/news/how-much-do-data-scientists-need-know-about-statistics*

What is data reliability? (2023, November 15). Ibm.com. https://www.ibm.com/topics/data-reliability

What is statistics from APA Dictionary of Psychology - *Google Search*. (n.d.). Google.com. Retrieved July 28, 2024, from,https://www.google.com/search?q=what+is+statistics+from+APA+Dictionary+of+Psychology&oq=what+is+statistics+from+APA+Dictionary+of+Psychology&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTIHCAEQIRigAdIBCTU4NjY4ajBqNKgCALACAA&sourceid=chrome&ie=UTF-8

Wikipedia contributors. (2024, July 24). *History of statistics*. Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=History_of_statistics&oldid=1236457100

Women speak 20,000 words a day, 13,000 more than men. (n.d.). *The Sun Nigeria*. Retrieved May 16, 2025, from https://thesun.ng/women-speak-20000-words-a-day-13000-more-than-men/

Women use more words daily than men in midlife, study finds. (2025, February). *Phys.org*. Retrieved May 16, 2025, from https://phys.org/news/2025-02-women-words-daily-men-midlife.html

(N.d.-a). Researchgate.net. Retrieved July 28, 2024, from https://www.researchgate.net/publication/322147021_Statistics_is_Essential_for_all_Fields_of_Scientific_Researches

(N.d.-b). Phys.org. Retrieved April 1, 2025, from https://phys.org/news/2025-02-women-words-daily-men-midlife.html

(N.d.-c). Researchgate.net. Retrieved July 28, 2024, from https://www.researchgate.net/figure/Types-of-statistical-research-methods-https-wwwenago_fig1_369895095

(N.d.-e). Nigeriainfo.Fm. Retrieved July 28, 2024, from https://www.nigeriainfo.fm/news/homepagelagos/rivers-state-school-in-deplorable-state/

(N.d.). Unicef.org. Retrieved April 1, 2025, from https://www.unicef.org/nigeria/press-releases/impact-covid-19-poor-mental-health-children-and-young-people-tip-iceberg-unicef

A BRIEF CITATION OF PROFESSOR MARY PASCHAL IWUNDU

Prof. Mary Paschal Iwundu is a distinguished Professor of Statistics, specializing in Experimental Design. She is a native of Akwa Ibom State and is deeply associated with Imo State by marriage. Prof. Iwundu is the beloved daughter of the late Mr. O. E. U. Usoro, a renowned teacher, and the late Mrs. Cornelia O. Usoro, a respected nurse. She spent much of her early years in Uyo, the capital city of Akwa Ibom State.

Her educational journey began at Christ the King Primary School in Uyo (1977–1982), followed by Immaculate Conception Secondary School in Itak-Ikono, Akwa Ibom State, where she graduated as a member of the Exquisite Class of 1988. Prof. Iwundu then pursued a Bachelor of Science in Statistics at the University of Uyo, graduating with a Second-Class Upper division in 1994.

After completing the mandatory National Youth Service Corps (NYSC) program at the Nigerian Ports Authority, Marina, Lagos (1994–1995), she began her academic career at the University of Uyo as a Graduate Assistant in November 1995, at the age of 24.

In 1997, she enrolled in the Master of Science program in Statistics at the University of Nigeria, Nsukka, where she earned a perfect 5.0 CGPA on 5-point scale in 1999, the highest ever recorded in the Department of Statistics and the established highest in the entire University of Nigeria.

Prof. Iwundu transferred to the University of Port Harcourt in 2003, where she joined the Department of Mathematics, Statistics and Computer Science. That same year, she began her Ph.D. Programme in Statistics at the University of Nigeria, Nsukka, graduating in 2008 with an A grade.

Prof. Iwundu holds B.Sc., M.Sc., and Ph.D. in Statistics, demonstrating expertise in both theoretical and computational aspects of the field. Her research interests focus on the Optimal Design of Experiments and their applications in Response Surface Methodology, and other areas of Applied Statistics.

Over the years, she has supervised many graduate students, including doctoral candidates in both Statistics and Business Management, the latter at Regenesys Business School, Sandton, Johannesburg, South Africa.

In 2020, Prof. Iwundu was formally announced as Professor of Statistics, although the promotion to that rank was effective from 2017. She has held numerous academic and administrative positions, including two terms as Head of the Department of Mathematics and Statistics (2012–2016), Post-

Graduate Programme Coordinator and Chairman, and several roles on various committees. She has served as Academic Consultant/Adviser as well as Professional Examiner and Assessor to several academic institutes.

Her leadership extends beyond the Department and Faculty, where she has contributed to the university's governance in positions such as Board Chairman of the Information and Communication Technology Centre, Director of the Intellectual Property and Technology Transfer Office, Director of the Information and Communication Technology Centre, Chairman of the Akwa Ibom Community, University of Port Harcourt.

Prof. Iwundu's professional development includes numerous certifications in Intellectual Property Law and Leadership from renowned institutions such as the University of Pennsylvania, USA, and the University of Illinois, USA. She is affiliated with Osiri University, headquartered in Lincoln, Nebraska, USA, where she has explored the rudiments of Data Science, and the philosophy of Ubuntu— "I am because we are"—in relation to servant leadership.

She has participated in various trainings and was actively involved in a training program on industry-academia collaboration, organized by the Japan Patent Office (JPO), where she earned the title "IP Friend of the Japan Patent Office" and was the only Nigerian participant.

As a visionary leader, Prof. Iwundu was recognized as the best-performing Intellectual Property and Technology Transfer Office Coordinator in Nigeria by the National Office for Technology Acquisition and Promotion (NOTAP) in 2023. This award placed the University of Port Harcourt among the top institutions in the country for innovation and intellectual property management.

In 2024, Prof. Iwundu organized the University of Port Harcourt's 2024 World Intellectual Property Day celebration, emphasizing the importance of innovation in achieving the Sustainable Development Goals (SDGs). She is a strong advocate for gender equality in accelerating innovation and creativity, yet very unbiased.

Throughout her 29-year academic career, Prof. Iwundu has shown much dedication to academic and administrative excellence, with commitment to academic and professional mentorship, serving as a role model to many. She is the first female Professor of Statistics at the University of Port Harcourt. She is recognized as the second officially announced female Professor of Statistics in South-South, even though by date of promotion, she ties as first female Professor of Statistics in the South-South. By extension, she is the second female Professor of Statistics in Nigeria.

Prof. Iwundu is the first female Director of the Information and Communication Technology Centre, University of Port Harcourt, and had served as the first female Director of the Intellectual Property and Technology Transfer Office of the University of Port Harcourt. She believes that having a woman on the team means there is a unique voice that brings about innovative solutions.

Beyond her academic and professional excellence, Prof. Iwundu is known for her strong ethical principles, and dedication to service and family values. She is happily married to Engr. Paschal Chuka Iwundu, and together, they are blessed with four children: Dr. Chinwendu Paschal Iwundu, Engr.

Chukwudi Paschal Iwundu, Miss Chisomaga Paschal Iwundu (LLB in view), and Master Noble Paschal Iwundu.

Today, as we gather to mark a significant milestone in Prof. Mary Paschal Iwundu's career, let us remember that learning is a lifelong journey, and the knowledge we acquire is a gift no one can take from us.

Ladies and Gentlemen, please join me in welcoming the Inaugural Lecturer, Prof. Mary Paschal Iwundu, FNSA.

Prof. Owunari Abraham Georgewill
Vice Chancellor