

## RUFORUM Annual General Meeting 2019

Scientific Data Management Skill Enhancement Training for Postgraduate Students  
at University of Cape Coast, Ghana.

29<sup>th</sup> November-03<sup>rd</sup> December, 2019

### CONCEPT NOTE

#### Background

Research in Africa involves generating agricultural output and other technologies that can improve food security but the impact is yet to be felt in sub-Saharan African (SSA) households. Proposal writing entails various stages such as reviews to identify researchable issues, building and testing theory under which the research problem emanates, integration and reflection of components that provide the map towards solving the research problem. Scientific data management enhances the capacity of postgraduate students to meaningful engage in conducting quality research by developing appropriate research proposals, design of studies, collection and analysis of data for meaningful reporting. Postgraduate students are heavily involved in large scale experiments or surveys that sometimes lead to complex designs and to subsequent messy data. Figuring out how to handle data resulting from such experiments/surveys takes time, and getting appropriate assistance is difficult. The students are also constrained on how to effectively analyze data using appropriate statistical software, interpret the results and communicate well to the target audience. In recognition of these shortcomings, the RUFORUM Network is mandated to strengthen the quality and relevance of postgraduate training and research in African universities—especially in agriculture, science, technology, and innovation—through activities designed to improve the capacity of African universities and research centres to generate knowledge relevant to Africa’s development challenges. RUFORUM is organising a Scientific Data Management Skill Enhancement Training for Postgraduate Students at University of Cape Coast, Ghana. The training is structured to encompass broad biometrical needs that will equip the postgraduate students

#### CO-ORGANISERS:



with skills required in conducting their research efficiently and effectively. The content incorporated in this course is drawn from broader topics ranging from research process, proposal development, planning of experiments/surveys, designing and implementing experiments/surveys, conducting data analysis for qualitative and quantitative data. The students will also be exposed to key statistical software (Genstat and SPSS) for data analysis and reporting.

**Aims and Objectives of the side Event**

The aim of this training is to improve efficient flow of agricultural information and research specifically to achieve the following; understand the various research processes essential in proposal writing, biometrical components pertaining to design and analysis of experiments/surveys; applying various statistical techniques correctly at all stages of research and report the results effectively. The training will equip postgraduate students with the skills and knowledge in software use (GenStat & SPSS) in data analysis and presentation of results in a format that would ensure their wide dissemination as peer reviewed publications and policy formulation. It is expected that this training will give them the hands-on skills they need to improve the quality and quantity of their research publications. The training will also be an opportunity for postgraduate students to prepare their proposal or draft thesis.

*Specific objectives*

To provide participants with:

- i) Appropriate research techniques in designing suitable scientific proposal in selected agricultural and related fields in science and innovation;
- ii) Statistical knowledge in design, analysis and result interpretation of experiments/surveys to agricultural, socio-economics and other biophysical studies;
- iii) Ability to manage complex data, analyse data using statistical software and interpret results objectively;
- iv) Skills to manage every component of their theses in order to complete studies within the specified timeframe; and
- v) Skills to produce manuscripts for publication that contains sound statistical facts.

**CO-ORGANISERS:**



**Approach  
and  
Methodology**

**Delivery Method and Requirements**

The training will cover aspects of research proposal development, and scientific data management. The module will have a balanced approach, including very little theoretical underpinning, analytical tools, and practical application of the learning to solve real-world problems. Delivery will be mixed mode, including interactive lectures and practicals designed to complement the lecture material. The approach will be participatory, with students expected to be active learners, and to commit themselves to intensive and critical self-study. Assignments will be designed to train and test critical thinking skills. Real life data sets brought by facilitators or drawn from students prior to the start of the course will be used throughout in examples and exercises. The mode of instruction is divided into two parts namely, limited theory/examples and computer exercises. All the exercises will use both SPSS and GenStat Software. Each participant will be expected to bring a laptop and a set of data. The daily programme will be divided into sections that will allow for an overview of the topics followed by computer-based practicals and discussion on the statistical results. Basic principles followed by computer examples will be introduced first. Participants will analyse their data using techniques already introduced on a daily basis. Discussions on interpretation and presentation of the results will be held every day during the plenary sections. The participants will evaluate the modules on a daily basis and shortcomings addressed immediately. An overall course evaluation will be done at the end of the module.

**Course Pre-requisite**

This course builds on the knowledge acquired by participants during their postgraduate and undergraduate studies. It assumes understanding of basic biometrics applied to quantitative and qualitative data, and in addition, numeracy skills acquired overtime. The module provides a solid understanding of statistical techniques that relate to quantitative/qualitative aspects from application, and analytical perspective, thus balancing between theory and applied concepts.

**Duration**

The write-shop will take five working days each day starting at 8:30 am up to 5:45 pm, with tea and lunch breaks in between.

**CO-ORGANISERS:**



### **An Overview/Summary of Training Content**

The training will cover research process and skills in proposal writing, a whole range of applied biometrics, say, from basic concepts to computing and results presentation. The topics will concentrate on introducing basic concepts on proposal writing, statistical concepts in a non-theoretical way, with computer-based practicals being used to illustrate the different concepts. These are highly practical topics intended to increase the participants' awareness of biometrical techniques for data management and analysis in their own specialist areas. The students will be grouped according to their area of interested and given assignments with data related to their area followed by presentations at the end of the day. These modules have been modified using the results from the training needs assessment sent to the participants.

#### ***Module 1- Research proposal development/writing***

Research process Reviews; Theory building, Theory testing, Reflection and integration, Components of research process and Research implementation strategy. Problem analysis and conceptualization of research ideas Problem tree analysis: Process and stages of problem tree analysis, Practical on student's proposal problem tree analysis. Formulation of research objectives, research questions and hypotheses and Ethics of research.

#### ***Module 2- Data Collection***

Introduction to planning and design of experiments and surveys, Data -Types of scientific data records, Data collection procedures, data quality measures, integrity and consistency checks, Variables and measurements, an overview of types of experiments and their applications; types and use of surveys; sampling techniques; sample size determination; types of qualitative research and their uses; and tools for study designs

#### ***Module 3- Managing the Data***

Software and techniques for data entry and effective retrieval in GENSTAT/SAS/SPSS. Detection of outliers or influential observations and exploratory/Graphical Data Analysis using statistical software

#### ***Module 4-Inferential statistics and Interpretation***

Parametric statistics: Analysis of variance (ANOVA), separation of means,; Moving beyond ANOVA; Correlation and Regression analysis (simple linear regression, multiple regression, validation of assumptions under regression etc); model building:- modelling and analysis of experimental and survey data; Introduction to types of statistical models (Generalised linear models [log-linear models, logistic regression

### **CO-ORGANISERS:**



model) and their applications; analysis of repeated measurements; mixed models etc ; Analysis of categorical data; use of chi-square tests; analysis of score and rank data; applications of GLM in survey data; multivariate analysis [use of factor/principal component analysis, canonical correlation analysis, discriminant analysis];

**Module 5-Interpretation, presentation and discussion of results**

This will entail presentation; interpretation and discussion of results with reference the participants' study objectives.

**Module 6. Student's clinic and wrap up**

This will cover individual student's clinics, where the participants share their designed studies with the facilitators. the facilitator will then advise the student and this improves their proposals (design, data management, analysis and interpretation).

Training outputs

At the end of the training participants would be able to:

1. Develop/or refine titles of prospective research proposal;
2. Formulate/or refine research objectives, questions and hypotheses of prospective proposals;
3. Grasp the fundamental concepts behind experimental/survey designs and statistical data analyses within the context of developing countries;
4. Understand and apply key statistical concepts such as correlation& regression analysis; categorical data analysis techniques and generalised linear models, etc;
5. Use statistical software to describe, analyze and model the state of a biological or agricultural system in both a quantitative and qualitative manner; and

Training outcomes/expectations

At the end of the training participants would be able to:

1. Apply sound research methods in developing/or writing research proposal for academic work and development;  
Apply sound research methods in developing/or writing research proposal for academic work and development;
2. Grasp and apply fundamental concepts behind experimental/survey designs and statistical data analyses within the context of developing countries;
3. Understand and apply key statistical concepts in data analyses and reporting of results for effective decision making and policy formulation; and

**Outputs/  
outcomes**

**CO-ORGANISERS:**



4. Use suitable statistical software to describe, analyze and model the state of a biological or agricultural system in both a quantitative and qualitative manner.

**Venue and Participants**

The side event will take place at the University of Cape Coast in Ghana. The training targets postgraduate students whose research is likely to contribute to agri-entrepreneurship, innovations, policy making and leadership management from different training programmes such as agricultural economics, plant protection, food science, natural resources management, aquaculture and fisheries sciences.

**Side Event Organizers and Contact**

The organiser is organized by RUFORUM and University of Cape Coast.

Facilitators: DR. VINCENT O. OEBA ([vongusoeba@gmail.com](mailto:vongusoeba@gmail.com))

DR SUSAN B. TUMWEBAZE ([balaba2@yahoo.com](mailto:balaba2@yahoo.com))

**CO-ORGANISERS:**





### 3.0. Training Programme

#### Scientific Data Management Skill Enhancement Training for Postgraduate Students at University of Cape Coast, Ghana- December 2019

<i>Time</i>	<b>Friday 29 November 2019</b>	<b>Saturday 30 November 2019</b>	<b>Sunday 1<sup>st</sup> December 2019</b>	<b>Monday 2 December 2019</b>	<b>Tuesday 03 December 2019</b>
<b>0830-1015</b>	<ul style="list-style-type: none"> <li>Registration –UCC</li> <li><b>Opening Ceremony</b></li> <li>Introduction of participants-UCC and RUFORUM</li> <li>Welcome remarks-Gulu University</li> <li>Objectives and overview of TNA report-Facilitators-Susan Balaba</li> <li>Training expectations, Facilitators -Vincent Oeba</li> <li>Remarks from Executive Secretary, RUFORUM</li> <li>Opening speech by Guest of honor-UCC</li> </ul>	<b>Qualitative research approaches and design of data collection tools</b> <ul style="list-style-type: none"> <li>Introduction to qualitative research</li> <li>Data collection tools for experimental designs</li> <li>Data collection tools for surveys</li> <li>Review of student’s draft data collection tools</li> </ul>	<b>Data analysis and presentation of results: Experimental data</b> <ul style="list-style-type: none"> <li>Introduction to descriptive statistics of experimental data in Genstat</li> <li>Analysis of variance (ANOVA) and separation of means</li> <li>Relationships &amp; Association</li> <li>Demonstration and practicals</li> </ul>	<b>Data analysis and interpretation of results</b> <ul style="list-style-type: none"> <li>Introduction to SPSS and exploratory data analysis for survey data</li> <li>Introduction to descriptive statistics of survey data using SPSS</li> <li>Frequencies, means, contingency tables</li> <li>Chi-square test of association</li> <li>Demonstrations and practicals</li> </ul>	<b>Qualitative data analysis</b> <ul style="list-style-type: none"> <li>Introduction to qualitative data analysis approaches/ methods</li> <li>Demonstration and practicals</li> </ul>
		VO/SB	SB/VO	VO/SB	VO/SB
<b>1015-1030</b>	<b>Health Break</b>	<b>Health Break</b>	<b>Health Break</b>	<b>Health Break</b>	<b>Health Break</b>
<b>1030-1300</b>	<b>Research Design:</b> <ul style="list-style-type: none"> <li>Overview of research process</li> <li>Problem analysis and conceptualisation of research ideas</li> </ul>	<b>Data management</b> <ul style="list-style-type: none"> <li>Introduction to data management</li> <li>Data checking techniques</li> </ul>	<ul style="list-style-type: none"> <li>Performing analysis of variance with examples</li> <li>Introduction to inferential</li> </ul>	<b>Introduction to parametric statistics of survey data</b> <ul style="list-style-type: none"> <li>Regression analysis</li> </ul>	<ul style="list-style-type: none"> <li>Presentation of results</li> <li>Students clinics with facilitators</li> </ul>

<i>Time</i>	<b>Friday</b> 29 November 2019	<b>Saturday</b> 30 November 2019	<b>Sunday</b> 1 <sup>st</sup> December 2019	<b>Monday</b> 2 December 2019	<b>Tuesday</b> 03 December 2019
	<ul style="list-style-type: none"> <li>Formulating research objectives, research questions and hypothesis</li> <li>Ethics of research</li> <li>Overview of student's draft research ideas/proposal and objectives</li> </ul>	VO/SB	<ul style="list-style-type: none"> <li>statistics in Genstat and Interpretation of results</li> <li>Moving beyond ANOVA</li> </ul>	<ul style="list-style-type: none"> <li>Generalized linear models</li> <li>Demonstration and practicals</li> </ul>	
<b>1300-1400</b>	<i>Lunch Break</i>	<i>Lunch Break</i>	<i>Lunch Break</i>	<i>Lunch Break</i>	<i>Lunch Break</i>
<b>1400-1515</b>	<b>Study designs</b> Overview of experimental design s <ul style="list-style-type: none"> <li>Highlights of study design &amp; selection of variable</li> <li>Design aspect</li> <li>Approach to analysis</li> <li>Hypotheses testing</li> <li>Case studies on designs from student's research ideas/proposal</li> </ul>	<ul style="list-style-type: none"> <li>Introduction to Genstat Exploration of experiments and qualitative data</li> <li>Data manipulation and exploration - Quick diagnostic</li> </ul>	<ul style="list-style-type: none"> <li>Simple &amp; multiple linear regression using Genstat</li> <li>Analysis of repeated measurements</li> </ul>	<ul style="list-style-type: none"> <li>Introduction to principal component analysis</li> <li>Logistic regression analysis</li> <li>Log linear modelling</li> <li>Practicals and demonstrations</li> </ul>	Students clinics
<b>1515-1530</b>	<i>Health Break</i>	<i>Health Break</i>	<i>Health Break</i>	<i>Health Break</i>	<i>Health Break</i>
<b>1530-1645</b>	<b>Study designs</b> <ul style="list-style-type: none"> <li>Overview of sampling designs</li> <li>Key elements of sampling</li> <li>Sampling techniques</li> <li>Sample size determination</li> </ul>	Data manipulation and exploration- quick diagnostic in Genstat	<ul style="list-style-type: none"> <li>Using general analysis REML approach in Genstat</li> </ul>	<ul style="list-style-type: none"> <li>Introduction to Non-parametric tests</li> <li>Demonstration and practicals</li> </ul>	<ul style="list-style-type: none"> <li>Student clinics</li> <li>Wrap up</li> </ul>



<i>Time</i>	<b>Friday 29 November 2019</b>	<b>Saturday 30 November 2019</b>	<b>Sunday 1<sup>st</sup> December 2019</b>	<b>Monday 2 December 2019</b>	<b>Tuesday 03 December 2019</b>
	<ul style="list-style-type: none"> <li>Case studies on designs from student's research ideas/proposal</li> </ul> VO/SB				
1645-1745	<ul style="list-style-type: none"> <li><i>Installation of statistical software</i></li> <li>Students clinics on developing research ideas and study designs</li> <li>Installation of statistical Software (Genstat and SPSS)</li> </ul> SB/VO	<ul style="list-style-type: none"> <li>Computer practical &amp; discussion of results</li> <li>Students clinics</li> </ul> VO/SB	<ul style="list-style-type: none"> <li>Computer practical &amp; discussion of results</li> <li>Students clinics</li> </ul> VO/SB	Students clinics VO/SB	

**SB: Susan Balaba Tumwebaze; VO: Vincent Oeba**

### 3.1 Annex 2: Training needs assessment tool

#### **Training Needs Assessment questionnaire**

#### **Pre-Course Questionnaire for Scientific Data Management Skill Enhancement Training for Postgraduate Students at University of Cape Coast, Ghana- 29<sup>th</sup> November-03<sup>rd</sup> December, 2019**

##### *Section A : Personnel Information*

Sex:         Male         Female

2. Name of degree being under taken

---

3. College/School/Department

---

4. Title of your research

---

---

5. Stage of research process

---

6. *Email*

---

##### ***SECTION B: Competence in proposal writing***

7. Knowledge in identifying a research gap:  None  Slight  Moderate  Good  Excellent  
Knowledge in stating the aim and objectives:  None  Slight  Moderate  Good  Excellent
8. Knowledge in stating research questions and hypothesis:  None  Slight  Moderate  Good  Excellent
9. Knowledge in writing a research design:  None  Slight  Moderate  Good  Excellent
10. Knowledge in identifying and performing the appropriate data analysis for the stated objectives:  None  Slight  Moderate  Good  Excellent
11. Knowledge in developing a conceptual framework :  None  Slight  Moderate  Good  Excellent
12. Knowledge in developing a timeframe and budget :  None  Slight  Moderate  Good  Excellent
13. Knowledge in manuscript and thesis writing :  None  Slight  Moderate  Good  Excellent

##### ***SECTION C: Competence in Data Management, Study designs and Analysis***

##### **Data Management**

14. Knowledge in design of data collection tools:  None  Slight  Moderate  Good  Excellent

15. Knowledge in data management techniques:  None  Slight  Moderate  Good  Excellent
16. Knowledge in design of spreadsheet for data entry:  None  Slight  Moderate  Good  Excellent
17. Knowledge in data checking:  None  Slight  Moderate  Good  Excellent
18. Knowledge in importation to statistical Software:  None  Slight  Moderate  Good  Excellent

### Basic Statistics and Interpretation of Results

19. Knowledge in some basic descriptive statistics (*measures of central tendency and measures of dispersion*):  None  Slight  Moderate  Good  Excellent
20. Knowledge in some basic inferential statistics (*hypotheses testing, t-test, ANOVA, confidence intervals*):  None  Slight  Moderate  Good  Excellent
21. Knowledge in data interpretation and reporting:  None  Slight  Moderate  Good  Excellent
22. Knowledge in design of data collection tools:  None  Slight  Moderate  Good  Excellent
23. Knowledge in presentation of results:  None  Slight  Moderate  Good  Excellent

### Statistical Software and Computer Skills

24. Which of the following statistical software you have access to and use/used?

<i>Statistical software aware</i>	<i>Aware of</i>	<i>Have access to</i>	<i>Use or used</i>
<i>GENSTAT</i>			
<i>SPSS</i>			
<i>SAS</i>			
<i>R</i>			
<i>MINITAB</i>			
<i>Any other specify:</i>			

25. What is your level of knowledge in GENSTAT?  
 None  Slight  Moderate  Good  Excellent
26. What is your level of knowledge in other statistics software (e.g. *SPSS, SAS, etc.*)?  
 None  Slight  Moderate  Good  Excellent
27. What is your level of knowledge in Spreadsheet/Microsoft Excel?  
 None  Slight  Moderate  Good  Excellent

### SPECIFIC TOPICS

Please tick in the appropriate cell in the grid below, your assessment of your needs and capabilities in the topics listed (*VD= very deficient, ND= Not deficient, MR= Major enhancement required, LR= little enhancement required*)

Component	Specific Topics	Level of competence on this topic					Level of enhancement required						
		VD		ND			MR		LR				
		1	2	3	4	5	1	2	3	4	5		
Proposal development and writing													
Data entry and management	General survey procedures												
	<b>Agricultural &amp; market survey procedures</b>												
	Large scale experimental designs												
	Techniques of data checking												
	Data storage and retrieval												
	Data management strategy												
Data analysis and Interpretations	Exploratory data analysis												
	Analysis of variance (ANOVA)												
	Chi-square and t-test												
	Non-parametric methods for survey data												
	Use mixed models												
	Use of generalised linear models (e.g. Logistic regression, Log-Linear)												
	Regression analysis (simple and multiple)												
	Quality management – principles and applications in agricultural research and data												
	<b>Interpretation of statistical results</b>												

Please email the form to the Course Facilitators: **Dr. Susan Balaba Tumwebaze** ([balaba2@yahoo.com](mailto:balaba2@yahoo.com)) and Dr. Vincent Oeba ([vongusoeba@gmail.com](mailto:vongusoeba@gmail.com))