FACTORIAL EXPERIMENTS AND ITS APPLICATIONS IN INDUSTRIES

BY

MR OHAEGBULAM PRECIOUS OSINACHI

PSC1709447

DEPARTMENT OF STATISTICS,

FACULTY OF PHYSICAL SCIENCES,

UNIVERSITY OF BENIN,

EDO STATE

JANUARY 2023

UNDERTAKING

This project work was carried out by me **OHAEGBULAM PRECIOUS OSINACHII** with matriculation **PSC1709447**. I have not copied the work of any other author(s) all texts used have been duly cited and acknowledged.

______________________________ _______________________

Ohaegbulam Precious Osinachi Date

CERTIFICATION

This is to certify that the project work was carried out by **MR OHAEGBULAM PREIOUS OSINACHI** with Mat. No: **PSC1709447** of the Department of Statistics, faculty of Physical Sciences, University of Benin in partial fulfillment for the requirement for the award of the Bachelor of sciences (B.S.C) Degree in Department of Statistics.

___________________ ___________________

Prof. J.E OSEMWENKHAE Date

Project Supervisor

___________________ ___________________

Prof. C.C. ISHIEKWENE Date

Head of Department

DEDICATION

I dedicate this study to God Almighty who has always been with me and given me the grace to carry out this project work. I acknowledge that without His inspiration, directives, guidance and empowerment I would not be able to come this far. I must not fail to mention the emotional, financial and inspiring support I got from my family, friends and community. This is a huge win for all well-wishers.

Finally, this project is dedicated to everyone whose research and knowledge has contributed immensely to this project work. I am grateful to everyone that will go through this project work for any use or inference.

ACKNOWLODGEMENT

In my experience, I have indeed come to the fact that education is a light. It is important to acknowledge that I would not be able to see this light without the effort of my parents MR & MRS OHAEGBULAM. They set my feet on the right part and tried to ensure that I follow this path.

What is the essence of light without shining it? This acknowledgement goes to all light bearers which include teachers, mentors, friends and family who have supported this work and most importantly to all lectures and staffs of the Department of Statistics, University of Benin, Benin City. The peak of this light is to all Professors of Statistics. It was a great privilege to be supervised by Prof. Joseph E. OSEMWENKHAE, he contributed hugely and guided me to see the light of this project work.

Finally, I must not fail to acknowledge the strong panelists that approved this work, led by the Dean of Physical Science, Prof. Joseph E.OSEMWENKHAE.

Thank you.

TABLE OF CONTENT

ABSTRACT

This study focused on factorial experiments and its applications in industries. In this study, we examined the application of a two-factor factorial design to determine the significant difference in the mean yield of onion with respect to the effect of fertilizers and plant densities. Primary data (yield of onion) was collected from the Department of Crop science, Faculty of Agriculture, University of Benin. This research work covers only two factors which are fertilizers at three levels (PM, OM and NPK) and plant densities at two levels (LOW and HIGH). The analysis techniques employed was a 2*3 replicated factorial design with 6 replicates per cell. Data collected was analyzed using SPSS version 22. The hypothesis tests were carried out at a (5%) significance level and the decision rule was to reject the null hypothesis if the calculated significance value (p-value) is less than the α (5%). Results from the analyses revealed among others that there is significant difference in the fertilizer effect and plant densities effect on the yield of onion with a significance level of 0.0001 and 0.016 respectively. In addition, there is no significant interaction effect between fertilizers and plant densities with significance value of 0.840 on the yield of onion.

CHAPTER ONE

INTRODUCTION

1.0 BACKGROUND OF STUDY

The pioneer of modern experimental design and statistician Sir Ronald Fisher noted that experiments are "simply experience well organized in advance, and designed to provide a secure basis of new knowledge" (Fisher, 1935:8). The following three elements define an experiment: (1) the manipulation of one or more independent variables; (2) the application of controls, such as allocating participants or experimental units at random to one or more independent variables; and (3) the careful observation or measurement of one or more dependent variables. Experiments differ from other research methodologies due to the first and second characteristics—the manipulation of an independent variable and the application of controls like randomization.

An efficient procedure for planning experiments such that the data collected may be analyzed to provide conclusive results is the (statistical) design of experiment (DOE). In an experiment, we deliberately change one or more process variables (or factors) in order to observe the effect the changes would have on one or more response variables.

An experimental design is a plan for assigning experimental units to treatment levels and the statistical analysis associated with the plan (Kirk, 1995). An experimental design is the creation of a thorough experimental strategy before the experiment is conducted. Establishing an experiment's goals and choosing the process factors for the study are the first steps in the design of experiments. A thoughtfully selected experimental design increases the quantity of "knowledge" that can be gathered for a given amount of experimental effort. Certain effects known as treatment populations or treatment combinations are under the experimenter's control. In most cases, the experimenter decides whether the experimental unit will be divided into groups termed BLOCKS. Statistical techniques are used in the planning and execution of the experiment's design and analysis to make sure that the required data are gathered and processed to enable reliable findings.

A factorial design as one of the areas of design of experiment is often used by scientists wishing to understand the effect of two or more independent variables upon a single dependent variable. Factorial experimental design, or simply factorial design, is a systematic method for formulating the steps needed to successfully implement a factorial experiment. Estimating the effects of various factors on the output process with a minimal number of observations is crucial to being able to optimize the output of the process. In a factorial experiment, the effects of varying the levels of the various factors affecting the process output are investigated. Each complete trial or replication of the experiment takes into account all the possible combinations of the varying levels of these factors. Effective factorial design ensures that the least number of experiment runs are conducted to generate the maximum amount of information about how input variables affect the output of a process (Batra and Seema, 2012).

Traditionally research methods generally study the effect of one variable at a time, because it is statistically easier to manipulate. However, in many cases, two or more factors may be interdependent, and it is the impractical or false to attempt to analyze them in the traditional way. Factorial designs are frequently used by agricultural scientists to assess the effects of variables on crops in the field. in research of this size, it is challenging and impractical to isolate and examine each variable separately. Factorial experiment allows subtle manipulation of a large number of interdependent variables. Despite its drawbacks, the approach is a helpful one for simplifying research and letting strong statistical techniques show any correlations. Factorial designs are extremely useful to field scientists as a preliminary study, allowing them to judge whether there is a link between variables, whilst reducing the possibility of experimental error and confounding variables.

The factorial design highlights the relationship between variables and permits multiple levels of analysis. Also, it enables the isolation and individual analysis of the effect of changing a particular variable. A factorial design must be carefully prepared because a mistake in one of the levels or in the general operationalization may ruin a lot of work. Other than these slight distractions, a factorial design is a mainstay of many scientific disciplines, delivering great result in the field.

1.1 RATIONALE FOR THE USE OF FACTORIAL DESIGN IN INDUSTRIES

The employment of experimental procedures is common in both academic and professional contexts. In industrial contexts, the main objective is typically to derive the most accurate information about the variables influencing a manufacturing process from as few (costly) observations as possible. Interaction effects are frequently viewed as a "nuisance" in industrial settings (they are often of no interest; they only complicate the process of identifying important factors). Every machine employed in a manufacturing process generally allows its operators to change a variety of parameters, which affects the final quality of the product the machine produces. The production engineer can discover which variables have the most effects on the final product's quality through experimentation, which enables them to change the machine's parameters in a systematic manner. With the use of this data, the settings can be continuously enhanced to produce the best possible results.

1.2 AIM AND OBJECTIVES OF THE STUDY

The aim of this study is to examine the factorial experiments and its applications in industries.

The following are the objectives of the study:

1. To examine the importance of factorial designs in industries
2. To apply factorial designs on real-life data sets.

1.3 CHALLENGES OF USING FACTORIAL DESIGNS

The major essence of using factorial design in any given industry is to generate the maximum amount of information about how input variables affect the output of a process. This information is gotten by various statistical tools and equations. These tools and equations have helped to make the measurement and analysis of factorial design much easier to obtain and interpret. However, despite the availability of these measurement tools and equations, there are some problems still associated with the application of factorial design in a given industry. First, there is the challenge of using the appropriate tools for measurement and secondly the challenge of the accurate analysis of the information obtained in relations to the purpose of use of the analysis output.

1.4 SIGNIFICANCE OF STUDY

This study seeks to provide its readers/experimenters with information on the impact of the application of factorial design in industries and thus, hopefully, will help them understand the importance of using factorial design as well as how it can be applied in industries.

1.5 SCOPE OF STUDY

This study was carried out at Ugbowo campus, University of Benin, Benin City, Edo State; 2022

1.6 DEFINITION OF TERMS

We define the following key terms and concepts used in this study.

1.6.1 EXPLANATORY VARIABLE: Experimental variables are variables that explain (predict) variability in the response variable. E.g. soil type.

1.6.2 RESPONSE VARIABLE: a response variable is a variable whose distribution is of interest. This could be quantitative (size, height, weight etc.) or qualitative (pass or fail, male or female etc.).

1.6.3 EXPERIMENTS: An experiment is the test performed to investigate or get answer to a question which the researcher wants to know.

1.6.4 ANALYSIS: This is the statistical analysis of the data from the experiment.

1.6.5 FACTORS: Factors are the distinct types of condition that are manipulated on the experimental units. E.g. Age, gender fertilizer etc. A factor can be fixed or random.

1.6.6 TREATMENTS: Treatments are different procedures we want to compare. This could be different kinds or amount of fertilizers.

1.6.7 EXPERIMENTAL UNITS: These are the things to which treatments are applied. E.g. different plots of land receiving different fertilizers, different customers receiving different rate structures etc.

1.6.8 MEASUREMENT UNITS/SAMPLING/RESPONSE UNITS: This is the object or person or thing being measured in an experiment. This may be different from the experimental units.

1.6.9 EXPERIMENTAL ERROR: This is the random variation present in all responses among experimental units. Different experimental units will give different responses to the same treatment. It is often said that when the same treatment is applied to the same unit over and over again, it will result in different trials.

1.6.10 RANDOMIZATION: This is the use of a known, understood probabilistic mechanism for the assignment of treatments to units. It ensures that each treatment has an equal chance of being assigned to any experimental unit.

CHAPTER TWO

REVIEW OF RELATED LITERATURE

2.1 INTRODUCTION

There are several articles on the uses and application of factorial design in transportation, agriculture, finance, telecommunication, hospitals etc. We review some of them here.

Richard and Darkwah (2006) employed two-factorial design to investigate how effective National Road Safety Strategy (NRSS 11) was in the reduction of Road Traffic Accidents (RTA) based on the data for the period 2006-2009 and extending to 2010 to inform policy strategy for the next face for the implementation of NRSS 11 which was to take off in 2011. After five years of running of NRSS 11, the study assessed its effectiveness using factorial design. For that purpose, data was collected from the Ashanti and Northern regional offices of the Motor Traffic Unit (MTU) of Ghana Police Service. In the study, two factors were considered. The first is REGION which served as factor A and comprises of 2 random levels, Northern and Ashanti regions. The second is ACCIDENT NATURE which served as factor B with 3 fixed levels as the number of minor cases, major cases and fatal cases. The test statistic F-ratio was used for the hypothesis testing of no interaction between the accident nature and region (that is, to find out whether accident nature depends on region or not); and also for the hypothesis testing of no difference among the treatment combination means (that is, to find out whether there is a difference in variability among the cells means or not). The research concluded that there is statistical evidence that accident nature depends on region and also that the effect of NRSS 11 is not the same across the two regions, Northern and Ashanti. The research further revealed that the seriousness reduction rate of RTA in Ashanti region is higher than that in the Northern region, meaning that the effect of NRSS 11 in Ashanti region is greater than that in Northern region. All the same, the NRSS 11 generally helped in the reduction of the seriousness of RTA.

Davide and Waste (2018) in their work proposed the use of factorial experimental design as a standard experimental method in the application of froth flotation to plastic separation as opposed to the widely-used OVAT approach (manipulation of one variable at a time). Instead of recovering plastics from the separation products, the kinetic model's characteristics were used as process responses. A series of 32 experimental tests were conducted utilizing combinations of two polymers with roughly the same density, PVC and PS (with mineral charges), to clarify and demonstrate the suggested methodology. The three controlled variables were pH, air flow rate, and frother concentration. A three-level full factorial design was conducted. The models showing the relationships between the controlled variables and their interactions with the responses (first order kinetic model parameters) were built. In order to choose the model that suited the data the best, the Corrected Akaike Information Criterion was utilized, and an analysis of variance (ANOVA) was carried out to determine which model terms were statistically significant. The researched showed that froth flotation can be used to efficiently separate PVC from PS with mineral charges by reducing the floatability of PVC, which largely depends on the action of pH. This factor has the most effect on the flotation rate constants during the tested interval. The results obtained showed that the pure error may be of the same magnitude as the sum of squares of the errors, suggesting that there is significant variability within the same experimental conditions. Thus, we need to employ special care during evaluation and generalization of the process.

Barnes *et al.* (2010) conducted a research to compare the relative effectiveness of hospital infection control measures. A 2^k factorial design was designed on the output of a stochastic, agent-based simulation to compare the effects of the hand hygiene compliance of healthcare workers and the nurse-to-patient ratio on the transmission of methicillin-resistant staphylococcus aurenus (MRSA) in a 20-bed ICU. The survey assumed that both the number of physicians and their hand hygiene behavior was constant (that is, only parameters related to nurses were varied). Two physicians were responsible for 10 patients each in the ICU, and physician compliance was held constant at 65% throughout the experiment. The research showed that increasing the nurse-to-patient ratio is more effective at levels below approximately 60% compliance of nurses. However, with greater baseline compliance levels, encouraging nurses to wash their hands more frequently becomes the more advantageous course of action. Additionally, the marginal advantage of increasing both components to high levels is limited by interaction effects between the two infection control methods.

King and Napal (1998) in their study used a factorial design to investigate how factors, such as the degree of meaning one obtains from one's job, happiness with one's job and one’s income, affect how others rate the person's desirability and moral character. Each subject was given the task of reading a career satisfaction questionnaire and rating the person who filled out the questionnaire on several dimensions. The questionnaires were all fictitious and were prepared by the researcher to represent one of eight conditions, represented by a 2 (high vs. low happiness) X 2 (high vs. low meaning) X 2 (high vs. low income). They found that the ratings of desirability and moral goodness were influenced by the factors of happiness in one's job and the amount of meaning one derived from one's job, but that income did not have an effect. All three independent variables influence the quality of life ratings.

Mutiu *et al.* (2016) in their study examined the application of a three-factor factorial design to determine the significant difference in the mean yield of maize in Nigeria with respect to the effect of fertilizers, herbicides and water volumes. For the successful execution of the research work, primary data (yield of maize) were collected from farm land cultivated on half plot of land in the year 2016. There were 216 ridges manufactured in total, which were divided into segments (9), each of which contained 24 ridges. Additionally, the 24 ridges were divided into three segments for a total of 8 replicates per factor level. This research work covered only three factors which are fertilizers at three levels [N:P:K(20:10:10), N:P:K(15:15:15), and UREA], herbicides at three levels (Altraforce, Xtraforce and Metaforce) and water volumes at three levels (5litres, 7.5litres, and 10litres). The maize (Soar 1) was planted in June 2016, the herbicides (Altraforce, Xtraforce and Metaforce) were applied a day after planting, the water volumes (5Litres, 7.5Litres and 10Litres) were applied everyday according to how the ridges were segmented irrespective of rainfall. The fertilizers [N:P:K(20:10:10), N:P:K(15:15:15), and UREA] were applied in August and the maize were harvested in September on the farm land and weighed per ridge in kilogram (kg). A 33 replicated factorial design with 8 replicates in each cell was employed as the analytic method. Results from the analysis revealed among others that there is significant difference in the fertilizers effect on the yield of maize while there is no significant difference in the herbicides effect. Similar to this, there is no significant difference in the effect of water volume on maize yield. 'Herbicides and water volumes' as well as 'fertilizers and herbicides' have significant interaction effect on the yield of maize.

Case studies are often used to teach students in subjects like law, medicine, and business. Lee and Hutchison (1998) employed design of experiments to determine if the degree of elaboration and reflection of the examples helps students to learn the material. In one experiment, two factors were manipulated: the type of presentation of the material (case, augmented or strategy) and whether reflective questions were asked of the students. Main effects were found for both the type of presentation and whether reflective questions were asked. Students tended to learn more when they were given reflective questions that required them to actively analyze the material, and the more information provided to them, the more they learnt. No interaction was found.

Iwok and Akpan (2016) studied the success of telecommunication providers as a function of network popularity, period of the day and gender of agent using factorial design. Based on information gathered from two commercial telephone centers situated at Ibom Plaza, Akwa Ibom State, Nigeria, high and low levels were ascribed to the three criteria. The separate and combined effects of the three factors were examined using the Analysis of Variance (ANOVA) approach. The study revealed that the three factors – Network popularity, period of the day and gender of agent separately affect the success of MTN and AIRTEL and by extension, the success of other mobile telecommunication service providers. That is, the success of a mobile telecommunication operator in a given community at any given time depends on the extent of its network coverage and quality of service at that time, period of the day and gender of the agent. They further concluded that, the interaction of network popularity with gender of agent has effect on the success of the two mobile telecommunication operators. This implies that, by using an agent with good customer relation, a mobile operator with wide network coverage and quality service delivery is bound to be successful.

Jacobby (1999) in this study used factorial experiments to measure age-related differences in memory. It was a factorial study, in which one factor was the age of the participant (young college students, older community residents who averaged 75 years old). The second factor was how quickly the participants had to respond. It was hypothesized that making older participants respond to memory tasks more quickly might interfere with their ability to recall information. Specifically, older adults were expected to confuse what words were in the original list to be learned and what words were not in the list, but were asked about during the testing procedure. Surprisingly, the hypothesis was not accepted.

2.2 FACTORIAL EXPERIMENTAL PLANNING

Perhaps one of the most critical stages of a scientific work is the optimization methodology of experimental parameters of relevance. According to the methodology chosen it is not always possible to evaluate the interaction between the variables and may contribute to a result that does not match the true. Hence the use of factorial experimental designs that can analyze multiple variable at the same time to the same factor (CUNICO, 2008; PERALTA-ZAMORA, 2005).

Day-Mayer *et al.* (2017), points out that in the methodological systems planning of experimental projects, the methods must be carefully analyzed, being sophisticated techniques with statistical experiments. To better understand the processes that are being monitored in a particular research is of utmost importance to the observation of the effects of the variables and the interaction between them (NETO *et al*. 2007; PEREIRA

FILHO *et al.* 2002). In this sense, the optimization of multivariable systems have shown great momentum and is useful in various fields of knowledge and received attention in studies involving biotechnological processes (PERALTA-ZAMORA *et al.* 2005).

Silva *et al*. (2016) used multivariate analyzes together with exploratory study units known as clustering. A hierarchical method with main components to assess the association between the production components and product doses Ribumin; a conditioner of soil applied in corn cultivars. The purpose of the multivariate analysis technique for grouping is to gather the sample units and groups so that there is homogeneity within the group and heterogeneity between them. However, multivariate optimization systems,

based on factorial design of experiments, have proven a useful and simple alternative given the need to simultaneously evaluate the effect of a large number of variables and the interaction between them from a small number of trials (PEREIRA FILHO *et al* 2002; PERALTA-ZAMORA *et al* 2005;. CUNICO *et al* 2008).

Freund *et al.* (2010b) points out that in the factorial experiment it is expected to examine the effect of two or more causes of the same type of sampling unit. In a laboratory study, factorial experiment can be utilized to analyze the differences in the yields of several varieties, as well as the different levels of function of a microorganism, for example. A model in each combination of all levels of factors are applied to its experimental units. From this perspective, understanding the project and statistical control of all experimental units of a sample leads to excellence of the quality of processes of goods and services of the final product to be desired.

The Experimental Planning Factor is one statistical tool that can be used to check the logic of the real synthesis method and establish whether it is necessary to revalidate the tests used to access the data provided in a search. A factorial experimental design defines the imposition to make changes in the data production of synthesis in order to establish control of processes such as the monitoring and control of processes with the purpose of detecting possible upgrades in the development and processes of research tests. Additionally,

It assists in taking preventive decisions. (ANDERSON; WHITCOMB, 2010; FREUND; WILSON, MOHR, 2010a, 2010b).

Anderson and Whitcomb (2010) having a factorial experimental design is to have a tool that allows to investigate each sample tested during the research meeting the requirements of procedures and protocols used by the researcher prior to testing.

2.3 MERITS OF FACTORIAL EXPERIMENTS

Compared to one-factor-at-a-time (OFAT) experiments which examines the effect of only a single factor or variable, factorial experiments offer several advantages.

- Factorial designs examine additional factors at no additional cost.
- OFAT experiment design cannot be used to detect the effect when one factor is different for different levels of another factor. Factorial designs are used to detect such interactions. Using OFAT when interactions are present can lead to serious misunderstanding of how the response changes with the factors.
- OFAT experiments are less effective than factorial designs. They provide more information at similar or lower cost. They can identify optimal conditions more quickly than OFAT experiments.
- Factorial designs enable the estimation of a factor's effects at various levels of the other factors, producing results that are reliable under a variety of experimental conditions.

2.4 DEMERITS OF FACTORIAL EXPERIMENTS

The major disadvantage is the difficulty of experimenting with more than 2 factors or many levels. A factorial design has to be planned well, as an error in one of the levels can jeopardize the entire work.

.

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 INTRODUCTION

This chapter presents the methodology used in this study. Discussed here are the research design, the sources of data that includes the rationale of the study, data collection, and tools used for data analysis.

3.1 RESEARCH DESIGN

Factorial design will be used in this study to examine the application of factorial experiment in Agriculture.

3.2 DATA COLLECTION METHOD

This study utilizes secondary data extracted from the Department of Crop science, Faculty of Agriculture, University of Benin.

3.2.1 DATA DESCRIPTION

An experiment was carried out to assess the effects of fertilizer type (factor A, with a = 3 levels: Poultry manure(PO), Organic manure(OM) and NPK (Nitrogen, phosphorus and potassium)) and plant density (factor B , with 2 levels: low and high) on the yield of Onion(bulb).

Each of the 6 (a*b) treatments was randomly applied to n = 6 plots (experimental units) hence, 36 (a*b*n) total observations…

3.3 METHOD OF DATA ANALYSIS

A 2*3 factorial design is applied in this study. A 2*3 factorial design is a type of experimental design that allows researchers to understand the effect of two independent variables on a single dependent variable. In this type of design, one independent variable has two levels and the other independent variable has three levels. Table 1 below gives the general layout for a two-factor factorial design.

Table 1: General layout for two-factor factorial design

	Factor B				Totals	Means
Factor A	**1**	**2**	…	**B**	$y_{i..}$	$\bar{y}_{i.}$
1	y_{111}, y_{112},… y_{11n}	y_{121}, y_{122},… y_{12n}	…	y_{1b1}, y_{1b2},… y_{1bn}	$y_{1..}$	$\bar{y}_{1.}$

2	y_{211}, y_{212},…, y_{21n}	Y_{221}, y_{222},…, y_{22n}		y_{2b1}, y_{2b2},…, y_{2bn}	$y_{2..}$	$\bar{y}_{2.}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
A	y_{a11}, y_{a12},…, y_{a1n}	y_{a21}, y_{a22},…, y_{a2n}		y_{ab}, y_{ab2},…, y_{abn}	$y_{a..}$	$\bar{y}_{a.}$
Totals($y_{.j.}$)	$y_{.1.}$	$y_{.2.}$	…	$y_{.b.}$	$y_{...}$	$\bar{y}_{...}$
Means ($\bar{y}_{.j.}$)	$\bar{y}_{.1.}$	$\bar{y}_{.2.}$	…	$\bar{y}_{.b.}$		

In Table 1, the first column is headed by factor A and comprises random levels counted as (1,2,…,a); the second column is headed by factor B and has *b* fixed levels counted as (1,2…,*b*); the third column headed by Totals($y_{i..}$, i= 1,2,…a) shows the sum of the replications (observations) for the *i*th level of factor A and the last column, headed by Means ($\bar{y}_{i..}$) is the average of the observations for the *i*th level of factor A. The last but one row, headed by Totals ($y_{.j.}$, j= 1,2…,*b*) is the sum of the observations for the *j*th level of factor B and the last row, headed by Means ($\bar{y}_{.j.}$), is the average of the observations for the *j*th level of factor B.

As there are *a* levels of factor A and *b* levels of factor B, *ab* is the total number of cells for treatment combinations. A treatment combination (cell) is a level of factor A applied in conjunction with a level of factor B. Also, if there are *n* observations in each cell, the total number of observations (replicates) in the experiment is given by *abn*. In addition, the *k*th observation taken at the *i*th level of factor A and *j*th level of factor B is denoted as y_{ijk}, where i= 1,2…, *a*; j= 1,2…, *b* and k= 1,2…, *n*. For instance, y_{31n} is the *n*th observation taken at the third level of factor A and at the first level of factor B.

Below are some useful symbols,

$y_{i..} = \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}$ $\qquad \bar{y}_{i..} = \frac{y_{i..}}{bn}$

$y_{.j.} = \sum_{i=1}^{a} \sum_{k=1}^{n} y_{ijk}$ $\qquad \bar{y}_{.j.} = \frac{y_{.j.}}{an}$

$y_{ij.} = \sum_{k=1}^{n} y_{ijk}$ $\qquad \bar{y}_{ij.} = \frac{y_{ij.}}{n}$

$y_{...} = \sum_{i}^{a} \sum_{j}^{b} \sum_{k}^{n} y_{ijk}$ $\qquad \bar{y}_{...} = \frac{y_{...}}{abn}$

3.3.1 ASSUMPTIONS

- Observations in each cell makes up a random sample size *n* which is assumed to come from a population that is normally distributed with mean (μ_{ij}) and variance σ^2.
- It is assumed that the sample observations in each *ab* cell have the same variance σ^2.
- The error term is identically and independently normally distributed with mean equal to zero and variance σ^2.

3.4 ANALYTICAL METHOD (MODEL FORMULATION)

Statistical model formulation implies the development of prediction equations by mathematical or statistical method from experimental data. A two-factor design with fixed-effects is used in this study. This is because the levels of both factors are in the control of the researcher (that is, the treatments were specifically chosen by the researcher) and the goal is to test hypotheses about the treatment means. Conclusions drawn here will only apply to the treatments considered and cannot be extended to other treatments that wasn't in the study.

Let Y_{ijk} be the value of the response from the *i*th level of factor A, *j*th level of factor B and *k*th level of replication.

The model is given by:

$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$

where i = 1,2,...a; j=1,2…,*b*; k= 1,2,…*n*.

μ = the overall effect estimated by $\bar{y}_{...}$

α_i = main treatment effect of factor A at *i*th level and is estimated by **$(\bar{y}_{i..} - \bar{y}_{...})$**;

β_j = main treatment effect of factor b at *j*th level and is estimated by

$(\bar{y}_{.j.} - \bar{y}_{...})$;

$(\alpha\beta)_{ij}$ = the interaction effect of *a* and *b* at the *ij*th level and is estimated by **$(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})$.**

e_{ijk} = random error and is estimated by **$(y_{ijk} - \bar{y}_{ij.})$**

3.5 HYPOTHESIS TESTING

In a two-factor factorial, we are equally interested in both row and column factors, A and B. Explicitly, we test the hypotheses about the equality of row treatment effects, say

$$H_o = \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$$

$$H_1 = \text{at least one } \alpha_i \neq 0$$

and the equality of column treatment effects, say

$H_o = \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_b = 0$

$H_1 =$ at least one $\boldsymbol{\beta}_j \neq 0$

We are also interested in testing if there exist column and row interaction effect. Thus we test

$H_o = (\boldsymbol{\alpha\beta})_{ij} = 0$ for all (i,j)

$H_1 =$ at least one $(\boldsymbol{\alpha\beta})_{ij} \neq 0$

3.6 SEQUENCE OF TESTING

The first hypothesis you must test is the interaction effect before the main effects or simple effect.

Step 1: Test for interaction effect. If interaction exists (if the H_o is rejected) skip to step 3 but if interaction doesn't exist, (if H_o is not rejected) go to step 2.

Step 2: Test the hypotheses for the main effect of factor A and main effect of factor B.

Step 3: Analyze the simple effect of one of the factors on the other factor.

3.7 SUM OF SQUARES

Sum of squares is a statistical tool used to determine the dispersion of data values from their mean value. It is the sum of squared deviations from a mean score.

Two-factor factorial analysis makes use of five sum of squares:

- **SUM OF SQUARES FOR FACTOR A (SS_A):** this measures variation of the marginal means of factor A($\bar{y}_{i..}$) around the grand mean ($\bar{y}_{...}$). It is computed as

$$SS_A = bn \sum_i^a (\bar{y}_{i..} - \bar{y}_{...})^2$$

- **SUM OF SQUARES FOR FACTOR B (SS_B):** this measures variation of the marginal means of factor B ($\bar{y}_{.j.}$) around the grand mean ($\bar{y}_{...}$). it is computed as

$$SS_B = an \sum_i^b (\bar{y}_{.j.} - \bar{y}_{...})^2$$

- **INTERACTION SUM OF SQUARES (SS_{AB}):** this is the sum of squares for the interaction between factor A and factor B and can be computed as

$$SS_{AB} = n \sum_i^a \sum_j^b \sum_k^n (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{...})^2$$

- **SUM OF SQUARES FOR WITHIN GROUP (SS_W):** Also known as the error sum of squares(SSE). It measures variation of all observations (y_{ijk}) around their respective group means ($\bar{y}_{ij.}$). it is computed as:

$$SS_W = \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{ij.})^2$$

- **TOTAL SUM OF SQUARES (SS_T):** this measures variation of all observations(y_{ijk}) around the grand mean ($\bar{y}_{...}$). It is computed as:

$$SS_T = \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{...})^2$$

The total sum of squares is the sum of the component sum of squares that is:

$$SS_T = SS_A + SS_B + SS_{AB} + SS_W$$

It can also be expressed as

$$\sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{...})^2 = bn \sum_i^a (\bar{y}_{i..} - \bar{y}_{...})^2 + an \sum_i^b (\bar{y}_{.j.} - \bar{y}_{...})^2 + n \sum_i^a \sum_j^b \sum_k^n (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{...})^2 + \sum_i^a \sum_j^b \sum_k^n (y_{ijk} - \bar{y}_{ij.})^2$$

This is established by the Fisher-Cochran theorem stated below:

Let Z be a sum of squares with N degrees of freedom and suppose $z_1, z_2 \ldots z_k$ are k sum of squares with respective degrees of freedom $N_1, N_2 \ldots N_K$ such that

$Z = z_1 + z_2 \ldots + z_k$ and

$N = N_1 + N_2 \ldots + N_k$

Then if Z has a chi-square distribution with N degrees of freedom then $z_1, z_2 \ldots z_k$ have independent chi-square distribution with $N_1, N_2 \ldots N_K$ degrees of freedom.

3.8 DEGREES OF FREEDOM

Degrees of freedom refers to the number of independent sample points used to compute a statistic minus the number of parameters estimated from the sample point. The number of degrees of freedom associated with the sum of squares are shown in the table below:

Table 2

SUM OF SQUARES	DEGREES OF FREEDOM
Factor A	a - 1
Factor B	b - 1
AB interaction	(a – 1)(b -1)
Within groups	ab(n – 1)
Total	abn - 1

The degrees of freedom for total sum of squares is equal to the sum of the degrees of freedom of component sum of squares. This is also established by Fisher-Cochran theorem stated in section 3.7 above.

3.9 MEAN SQUARES

The population variance is estimated by a mean square. The mean square is obtained by dividing the sum of squares (SS) by its corresponding degrees of freedom (df).

For a two-factor factorial experiment, the four mean squares employed to conduct analysis of variance are shown below:

- **Mean square for factor A(MS_A):** this measures variation due to the main effect of factor A. It is computed as:

$$MS_A = SS_a / df_a$$

- **Mean square for factor B(MS_B):** measures variation due to the main effect of factor B. It is computed as:

$$MS_B = SS_b / df_b$$

- **Mean square for interaction AB(MS_{AB}):** this measures variation due to the AB interaction effect. It can be computed as:

$$MS_{AB} = SS_{AB} / df_{AB}$$

- **Mean square for error(MS_E):** this measures variation due to the differences among experimental units within the same treatment group. It is computed as:

$$MS_E = SS_e / df_e$$

3.10 THE F-TEST

The F-test is used to determine whether we accept or do not reject the null hypothesis (H_0). The F-value is the mean square for each main effect and the interaction effect divided by the within variance. When F-value is small, we do not reject the null hypothesis and when F-value is larger than F-table value, we reject the null hypothesis (H_o). Appealing to the Fisher-Cochran's theorem; for a two factor factorial experiment the F-values are computed as:

$$F_A = MS_A/MS_E \sim f_{\alpha\,(a-1,\, ne)}$$

$$F_B = MS_B/MS_E \sim f_{\alpha\,(b-1,\, ne)} \text{ and}$$

$$F_{ab} = MS_{AB}/MS_E \sim f_{\alpha\,((a-1)(b-1),\, ne)}$$

where f_α are read from the f-distribution table and n_e is the error degree of freedom; ab(n-1).

3.11 ANOVA

Table 3 **A Typical two-way ANOVA table**

Source of variation	Degrees of Freedom	Sum of squares	Mean squares	f-value
Factor A	a -1	SS_A	MS_A	$MS_A/MS_E=F_A$
Factor B	b -1	SS_B	MS_B	$MS_B/MS_E=F_B$
Interaction AB	(a-1)(b-1)	SS_{AB}	MS_{AB}	$MS_{AB}/MS_E=F_{AB}$
Within group error	ab(n-1)	SS_E	MS_E	
Total	abn – 1	SS_T		

DECISION

- Reject H_0 if $F_{ab} > F_\alpha$ or P-value < α and conclude that there is interaction

- Reject H_0 if $F_a > F_\alpha$ or P-value $< \alpha$ and conclude that factor A has a significant main effect
- Reject H_0 if $F_b > F_\alpha$ or P-value $< \alpha$ and conclude that factor B has a significant main effect

3.12 MODEL ADEQUACY

The adequacy of the model needs to be checked before the conclusions from the analysis of variance is adopted. The primary diagnostic tool for model adequacy checking is residual plot. Basically, a residual plot is an error in the fit of a model. When testing for interaction, and for main effects of factor A and B, we assumed that the errors are a sample from a normal distribution with mean equal to zero and variance constant. If this assumption holds, then the p-values for the tests of main effects and interaction effects are valid. The residual plot tells if this assumption holds or not.

The residuals for a two-factor factorial model is given by:

$E_{ijk} = y_{ijk} - \bar{y}_{ij.}$

One of the residual plots is the normal probability plot. For the normal probability plot of residuals, the plot exhibits some kind of linearity if the underlying error distribution is normal, hence the adequacy of the model. However, the occurrence of an outlier is a common defect that shows up on the plot which can seriously distort the analysis of variance. Mostly, the cause of the outlier could be human error such as calculation error, data coding error or copying error. However, a suspected outlier could be checked by examining the standardized residual value (d_{ijk}) given by:

$$d_{ijk} = E_{ijk} / \sqrt{MSe}$$

A standardized residual value (d_{ijk}) bigger than 3 or 4 in absolute is a potential outlier which can cause a distortion to the conclusion drawn from the ANOVA (Montgomery, Douglas C. (1984). Design and analysis of experiments. New York:Wiley).

The coefficient of determination (R^2) could be used to further verify the results. That is, $R^2 = r\%$, then $r\%$ of the variability of the response variable in the model is explained by factor A, factor B and their interaction.

CHAPTER 4

DATA PRESENTATION, ANALYSIS AND INTERPRETION

This chapter presents and analyses the data from the experiment. The test is also carried out and findings reported.

4.1 DATA PRESENTATION

We would be looking at an experiment carried out at the Department of Crop Science, Faculty of Agriculture, University of Benin to compare the effects of fertilizer type (factor A, with a = 3 levels: PO, OM, NPK) and plant density (factor B , with 2 levels: low and high) on the yield of Onion(bulb).

Each of the 6 (ab) treatments was randomly applied to n = 6 plots (experimental units) that is, 36(abn) total observations. The experiments were performed in random order. For more details of this experiment see Department of Crop Science, Faculty of Agriculture, UNIBEN. The result of the experiment is given below:

Table 4: Observed Data For 3 Varieties Of Fertilizers At 2 Densities

	Plant Density (B)	
Fertilizer Type (A)	LOW	HIGH
PM	17.8, 19.1, 20.6, 21.2, 22.7, 23.3	22.1, 22.5, 14.1,19.1, 20.7, 22.6
OM	19, 18.7, 25, 21.30, 22.90, 23.80	23.80,24.30,25.40, 21.30, 22.70, 24.30
NPK	25.30, 26, 27.6, 26.80, 28.30, 29.20	27.90,31,30.70, 27.20, 28.30, 29.10

4.2 DATA ANALYSIS

In this case study, the researcher wants to answer the following questions:

1. Does the effect of fertilizer type on the yield of onion depend on plant density or vice-versa? If the effects do not depend on the other factor, we could ask:
2. What effects do fertilizer type and plant density have on the yield of onions?
3. What is the effect of changing the plant density from low to high on the expected yield of onions?
4. What is the effect of changing the fertilizer type from, e.g. PM to OM on the expected yield of onion?

It is always important to look at the sample average yields for each treatment of each level of factor A and at each level of factor B. This is shown in table 5 below:

Table 5. Summary Table

	Plant Density (B)		Sample average yield for each level of factor A
Fertilizer Type (A)	LOW	HIGH	
PM	20.7833	21.85	21.3167
OM	21.7833	23.6333	22.7083

NPK	27.200	29.0333	28.1167
Sample average yield for each level of factor B	23.2556	24.8389	24.0472

For example, 21.3167 is the average yield for PM over both levels of planting densities. The value 23.2556 is the average yield for plots planted in low density across all fertilizer types. The grand mean is 24.0472.

All data obtained were subjected to analysis of variance (ANOVA) using statistical package for social science (SPSS). Differences between the treatment means were separated using Turkey test.

We will now analyze the data using ANOVA with level of significance $\alpha=0.05$ = 5%. This implies that the null hypothesis will have 5 in 100 chances of being erroneously rejected when it is actually true. The ANOVA table is presented next:

Table 6: Two-way ANOVA Table

Source	Df	SS	MSS	F-value	P-value
Fertilizer type	2	309.707	154.854	45.165	<0.0001
Plant Density	1	22.562	22.562	6.581	0.016
Interaction	2	1.202	0.601	0.175	0.840
Error	30	102.858	3.429		
Total	35				

R Squared = 0.764 (Adjusted R Squared = 0.725)

4.3 HYPOTHESIS TESTING

We recall that the first hypothesis we must test is the interaction effect. Hence;

H_O: There is no interaction between factor A and factor B

H_1: There is a significant interaction between both factors

This hypothesis asks if the effect of low versus high plant density is the same for units in ordinary manure(OM), as it is for units in poultry manure(PM) and NPK.

The F-statistic:

$$F_{AB} = \frac{MSAB}{MSE} = \frac{0.601}{3.429} = \mathbf{0.175}$$

Conclusion: From Table 6, the p-value for the test for significance interaction factors is 0.482, this p-value is greater than the $\alpha = 0.05$, therefore we fail to reject the null hypothesis and conclude that there is no evidence of a significant interaction between fertilizer type and plant density. So it is appropriate to carry out further tests concerning the presence of the main effects.

4.3.1 MAIN EFFECT OF FACTOR A (FERTILIZER TYPE)

H_O: There is no effect of factor A on the response variable (onion yield)

H_1: There is a significant effect of factor A on the response variable

This hypothesis asks if the average onion yield is different for ordinary manure(OM), poultry manure(PM) and NPK.

The F-statistic:

$$\mathbf{F_A = \frac{MSA}{MSE} = \frac{154.854}{3.429} = 6.581}$$

Conclusion: from Table 6, the p-value (<0.001) is less than the α = 0.05 therefore we reject the null hypothesis and conclude that there is a significant difference in the yield between the three fertilizer types.

4.3.2 MAIN EFFECT OF FACTOR B (PLANT DENSITY)

H_O: There is no effect of factor B on the response variable (onion yield)

H_1: There is a significant effect of factor B on the response variable

This hypothesis asks if the average onion yield is different for seeds planted in low density and seeds planted in high density.

The F-statistic:

$$\mathbf{F_A = \frac{MSA}{MSE} = \frac{22.562}{3.429} = 45.165}$$

Conclusion: from Table 6, the p-value is 0.003. This is less than the α = 0.05, therefore we reject the null hypothesis and conclude that there is a significant difference in the yield between the two plant densities.

4.4 MULTIPLE COMPARISONS

The next step is to examine the multiple comparisons for each main effect to determine the differences. We will perform a Tukey-Kramer (Tukey's W) multiple comparison analysis to determine which means are similar and which means are different for only factor A since it has three levels. The result is given in table 7 and 8 below:

Table 7: Multiple Comparisons Of Fertilizer Type

Dependent Variable: BULB YIELD

Tukey HSD

(I) fertilizer type	(J) fertilizer type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
PM	OM	-1.3917	.75593	.174	-3.2552	.4719
	NPK	-6.8000*	.75593	.000	-8.6636	-4.9364
OM	PM	1.3917	.75593	.174	-.4719	3.2552
	NPK	-5.4083*	.75593	.000	-7.2719	-3.5448
NPK	PM	6.8000*	.75593	.000	4.9364	8.6636
	OM	5.4083*	.75593	.000	3.5448	7.2719

The mean difference is significant at the .05 level.

Intervals that does not contain zero implies that the means are different while intervals that contain zero implies that the means are similar. for example, the confidence interval for PM-OM is (-3.2552, 0.4719) which contains zero hence their means are similar while the confidence interval for PM-NPK and OM-NPK does not contain zero hence their means are different. Thus, we are 95% confident that applying NPK fertilizers to plants seem to produce higher average onion yield than plants applied with OM and PM fertilizers (which have similar population mean yield).

Table 8: BULB YIELD FOR FERTILIZER TYPE

Tukey HSD[a,b]

fertilizer type	N	Subset	
		1	2
PM	12	21.3167	
OM	12	22.7083	
NPK	12		28.1167
Sig.		.174	1.000

Means for groups in homogeneous subsets are displayed.

Alpha = .05.

This table corresponds to Table 7. Note that the PM and OM sample means (21.3167 & 22.7083) are grouped together (separately from the differing NPK sample mean (28.1167)). This shows that we are 95% confident that the PM and OM population means are similar, yet both differ from the NPK population mean (which agrees with the conclusion based on the simultaneous confidence intervals).

4.5 Model Adequacy

Before the conclusions of the ANOVA are accepted, we'll check the adequacy of the underlying model. As stated before, the primary diagnostic tool for model adequacy checking is residual plot, of which one of them is the normal probability plot. The residuals for a two-factor factorial model is given by:

$E_{ijk} = y_{ijk} - \bar{y}_{ij.}$

The residuals from the onion yield data are shown in the SPSS output below:

Table 9: RESIDUALS FOR ONION YIELD

	BULB YIELD (y_{ijk})	Predicted Value for yield ($\bar{y}_{ij.}$)	Residual for yield (E_{ijk})
1	17.80	20.78	-2.98
2	19.10	20.78	-1.68
3	20.60	20.78	-.18
4	21.20	20.78	.42
5	22.70	20.78	1.92
6	23.30	20.78	2.52
7	22.10	21.85	.25
8	22.50	21.85	.65
9	24.10	21.85	2.25
10	19.10	21.85	-2.75
11	20.70	21.85	-1.15
12	22.60	21.85	.75
13	19.00	21.78	-2.78
14	18.70	21.78	-3.08
15	25.00	21.78	3.22
16	21.30	21.78	-.48
17	22.90	21.78	1.12
18	23.80	21.78	2.02

19	23.80	23.63	.17
20	24.30	23.63	.67
21	25.40	23.63	1.77
22	21.30	23.63	-2.33
23	22.70	23.63	-.93
24	24.30	23.63	.67
25	25.30	27.20	-1.90

Table 9: RESIDUALS FOR ONION YIELD

26	26.00	27.20	-1.20
27	27.60	27.20	.40
28	26.80	27.20	-.40
29	28.30	27.20	1.10
30	29.20	27.20	2.00
31	27.90	29.03	-1.13
32	31.00	29.03	1.97
33	30.70	29.03	1.67
34	27.20	29.03	-1.83
35	28.30	29.03	-.73
36	29.10	29.03	.07

The normal probability plot of these residuals is displayed in the figure below, it plots the residuals versus the fitted values ($\hat{y}_{ijk.}$). A diagonal straight line means that error terms are normally distributed. If the line is skewed to the left or right, it means that you do not have a normally distributed data.

It can be easy to see with a histogram how data fits the norm or skews from the norm but with normal probability plot, it can be easier to see individual data items that don’t quite fit a normal distribution. The normal probability plot of these residuals is displayed in the figure below:

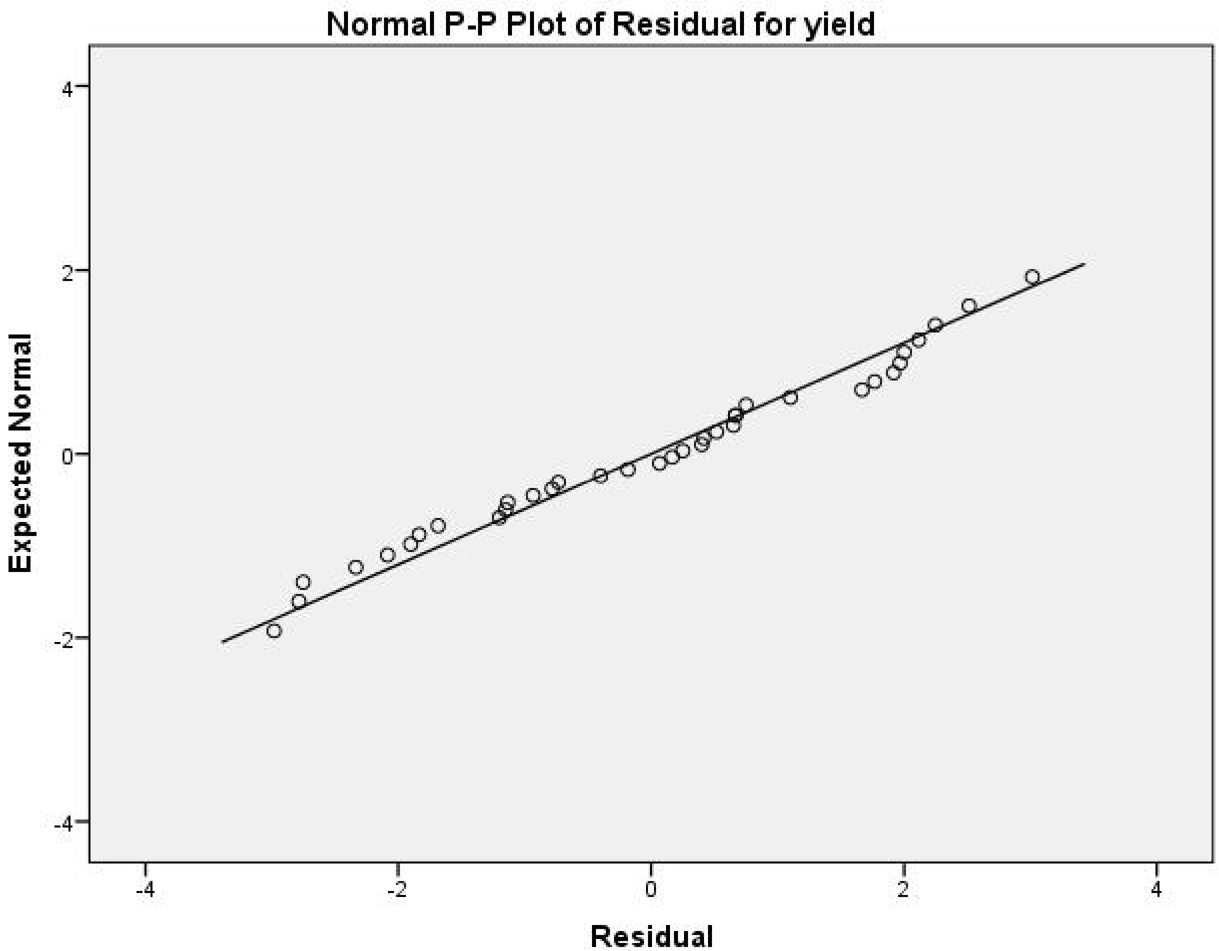

Figure 1: Normal probability plot of residuals for onion yield

The normal probability plot of the residuals above is approximately linear hence, supporting the condition that the error terms are normally distributed.

4.6 FINDINGS

To assist in interpreting the results of this experiment, it is helpful to construct a graph of the average responses at each treatment combination. For every combination of the levels plant density (low and high) and fertilizer type (OM, PM and NPK), we calculate the average value of the response. In figure 2.1 we use plant density on x-axis. In addition, settings corresponding to the same level of fertilizer type are connected with lines. In figure 2.2 we use fertilizer type on the x-axis and settings corresponding to the same level of plant density are connected with lines. In figures 2.1 and 2.2 The absence of significant interaction is indicated by the parallelism of the lines. In general, the lowest average onion yield is attained at low density for plants applied with PM fertilizers.

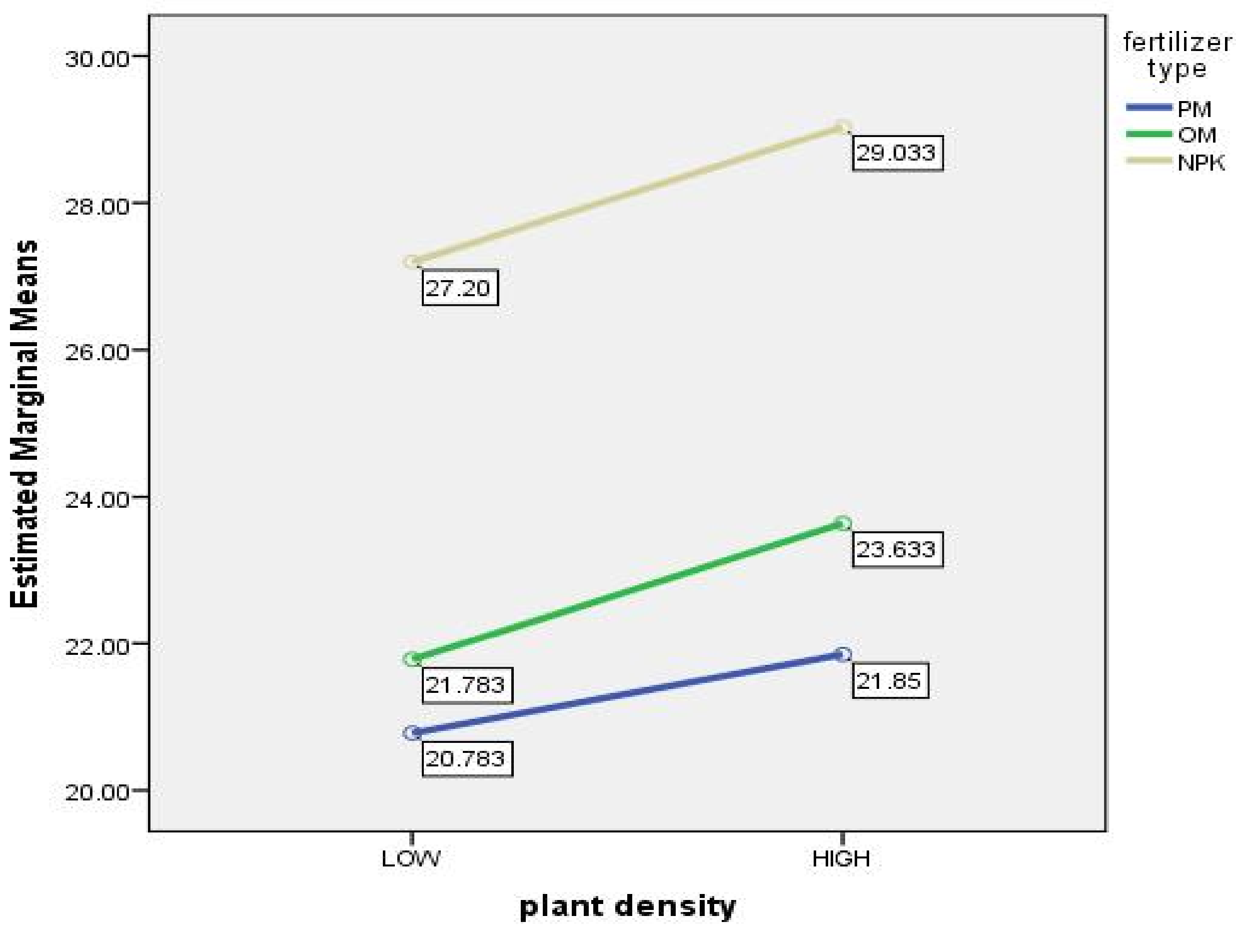

Figure 2.1: Average responses at each treatment combination

In figure 2.1 we can observe that the average yield of onion increases when changing from low to high density but the highest average onion yield is attained when plant density is high and NPK fertilizer type is applied.

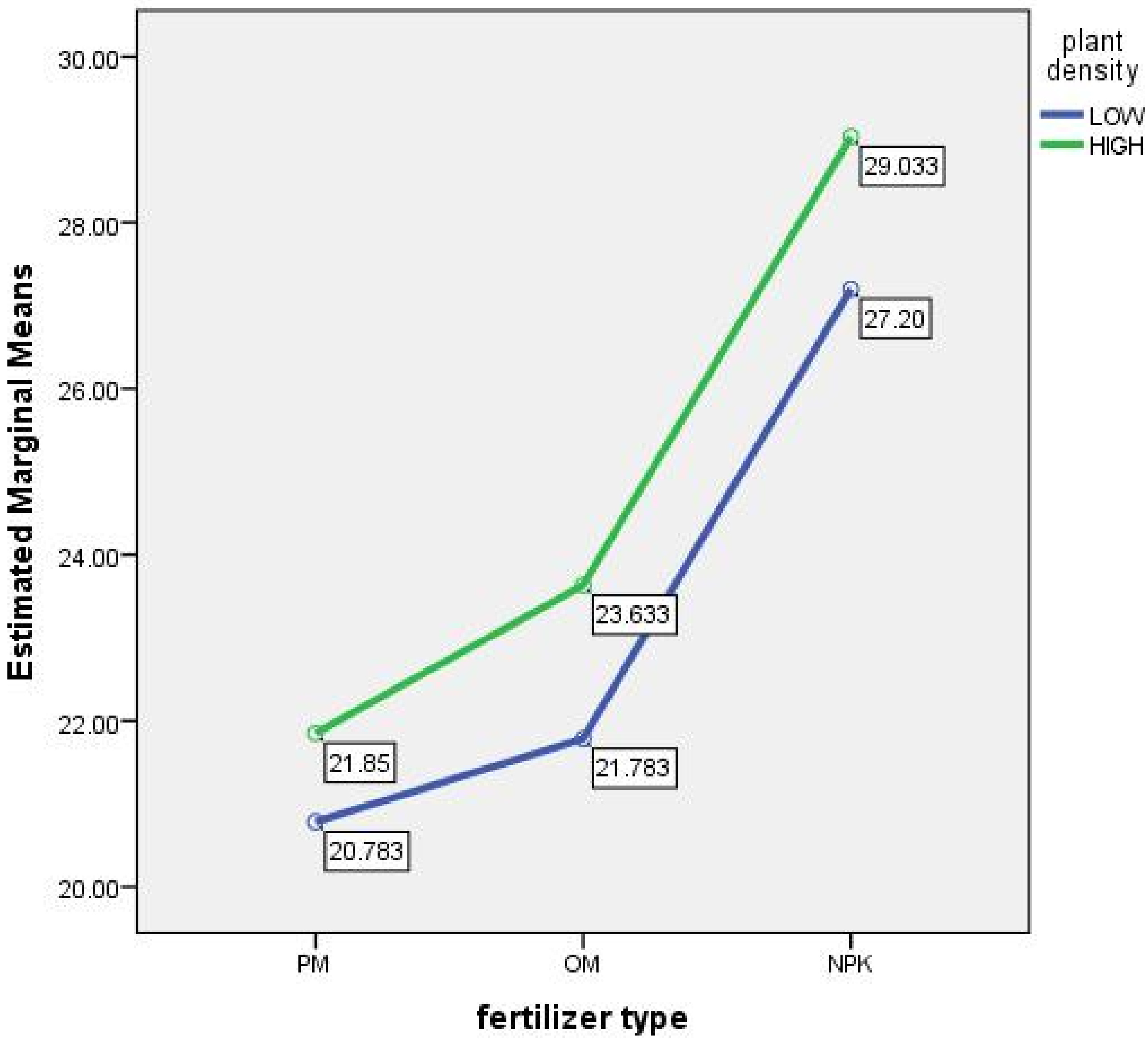

Figure 2.2: Average responses at each treatment combination

In figure 2.2 we can observe that the average yield of onion increases when changing the fertilizer type from PM to OM and this holds for both plant densities. We also observe that the average yield of onion increases when changing the fertilizer type from OM to NPK. This holds for both plant densities. However, the increase in high plant density is more pronounced compared to low plant density. Thus, applying NPK fertilizer give uniformly greater yield of onions regardless of plant density.

Furthermore, Table 7 and 8 revealed that plants applied with poultry manure and organic manure produce similar average onion yield whereas applying NPK to plants produced a higher average onion yield than the other fertilizer types and this corresponds to the figure 2.1and 2.2.

Additionally, An R-squared value of 0.764 implies that 76.4% of the variability of the response variable (onion yield) in the model is explained by factor A, factor B and their interaction.

CHAPTER 5

SUMMARY AND CONCLUSION

It is without doubt that the applied factorial design of experiments approach used in this study offers many advantages over conventional "one variable at a time" experiments by allowing researchers the ability to determine interactions between factors, more efficient use of data and statistical optimization of results.

5.1 CONCLUSIONS

From the findings of the study, we conclude that at 5% significance level:

1. There is no significant interaction effect between fertilizers and plant density on the yield of onion.
2. There is significant difference in the fertilizers effect on the yield of onion.
3. There is significant difference in the plant density effect on the onion yield.
4. Changing fertilizer type from OM or PM to NPK increases the average yield of onion.
5. Planting at low density with PM fertilizers produces the lowest average onion yield.

6. Planting at high density with NPK fertilizer produces the highest average onion yield.

5.2 RECOMMENDATIONS

In the light of the findings of this study, the following recommendations are made for adequate yield of onion.

1. The significant difference in the fertilizers effect implies that the three fertilizers do not perform equally on the yield. The marginal means for fertilizer type (PM--21.3167, OM--22.7083 and NPK—28.1167) suggests that the poultry manure and ordinary manure have similar effect on onion yield but NPK performs better with a mean of 28.1167 therefore NPK fertilizer is recommended for planting for optimal yield.
2. The significant difference in the plant density effect implies that the two plant densities do not have equal effect on the yield. The marginal means for plant density (LOW—23.2556 and HIGH—24.8389) suggests that high plant density performs better with mean of 24.8389 hence, it is recommended for planting for optimal yield.
3. Combining these two treatments-- NPK fertilizer and HIGH plant density produces the highest average onion yield. Hence, for optimal onion yield it is recommended to apply NPK and plant at a high density.

REFERENCES

ANDERSON; WHITCOMB, 2010; FREUND; WILSON, MOHR, 2010a, 2010b

Adeola, A.O. (2013). Experimental design and analysis of the action of different brands of soap and water. International Journal of Physical Sciences, Vol 3, pp 35-39. New Delhi. ISBN: 287-65-5423-093-4.

Fisher, R. A. (1960). The Design of Experiments, New York: Hafner Publishing Company.

Fisher, R.A. (1935). The Design of Experiments. Edinburgh and London: Oliver and Boyd.

International Journal of Advanced Engineering Research and Science (IJAERS) [Vol-5, Issue-6, Jun- 2018]

Kirk, R.E. (1995) Experimental Design: Procedures for the Behavioral Sciences (3rd edn.). Pacific Grove, CA: Brooks/Cole.

Montgomery, D.C. (2001). Design and analysis of experiment, 5th Edition. Arizona State University. John Willey & Sons, Inc. New York.

McGraw-Hill. New York, pp. 514-539. 169 Conference on Applied Statistics in Agriculture Kansas State University New Prairie.

Montgomery, Douglas C. (1984). Design and analysis of experiments. New York:Wiley.

PEREIRA FILHO *et al* 2002; PERALTA-ZAMORA *et al* 2005;. CUNICO *et al* 2008.

Winer, B. J. 1971. *Statistical principles in experimental design.* Second edition.

Yates, F. (1937) 'The design and analysis of factorial experiments', Imperial Bureau of Soil Science Technical Communication No. 35, Harpenden, UK.