United States International University-Africa

Education to take you places

SCHOOL OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF DATA SCIENCE AND ANALYTICS

SUMMER 2024 – ASSIGNMENT 2

COURSE CODE: STA 4020A

UNIT NAME: DESIGN AND ANALYSIS OF EXPERIMENTS

DATE: 23RD MAY 2024 **TOTAL MARKS**: 30 MARKS

INSTRUCTIONS:

For this exercise:

1. ANSWER ALL QUESTIONS
2. Do all your working in the Rmarkdown (.rmd).
3. Submissions should be in either a **`.ipynb`** or **`.rmd`** **file**
4. NO SUBMISSIONS SHOULD BE DONE VIA EMAIL

QUESTION 1

A researcher is conducting an experiment to test the effectiveness of four different fertilizers (A, B, C, and D) on the yield of a specific crop. The experiment is conducted in four different fields (blocks) to account for variability in soil quality and other environmental factors. Each field is divided into four plots, and each fertilizer is randomly assigned to one plot within each field.

The yield data (in kilograms) from the experiment are as follows:

Field	Fertilizer A	Fertilizer B	Fertilizer C	Fertilizer D
1	20	18	15	22
2	25	24	20	23
3	18	17	14	20
4	22	21	19	24

a) State the null and alternative hypotheses for this experiment.
b) Construct an ANOVA table for the RCBD analysis and determine whether there is a significant difference among the fertilizers and (or) replicates at the 0.05 significance level.
c) Interpret the results of the ANOVA table.
d) Conduct a post-hoc analysis if necessary

QUESTION 2

A researcher conducted an experiment to evaluate the growth performance of four different plant varieties (A, B, C, and D) in five different soil types (blocks). Each soil type represents a different field condition. The experiment was set up in a randomized complete block design, with each combination of plant variety and soil type replicated four times. However, due to unforeseen circumstances, two observations are missing from the data:

Block (Soil Type)	Variety A	Variety B	Variety C	Variety D
1	15	17	20	?
2	?	18	22	21
3	16	19	23	20
4	14	16	?	19
5	18	20	24	22

a) Estimate the missing values in this dataset.
b) Conduct an analysis of variance (ANOVA) to test whether there are significant differences in the mean growth performance among the four plant varieties and(or) the soil types at the 0.05 significance level, using the completed dataset.
c) Interpret the results of the ANOVA and discuss any implications for the study.
d) Conduct a post-hoc analysis if necessary