

FAUZIA CHEVONNE MOHAMMED

AIIP SUMMATIVE ASSESSMENT STATISTICS, 2019

1. This survey dealt with a lot of variables. Some of the variable types include dependent and independent variables. Examples of the independent variables are Agricultural Seasons e.g. summer, winter; and dependent variables are Crop type, Harvest quantity, Farming Yield etc.
2. The study design for this survey is a Multistage stratified sampling method. Specifically, the population was sampled first according to location coordinates in Africa (North, South, East, West), The first sample retrieved was at a Country level i.e. countries were selected based on institutions concerned with Climate change (those that expressed interest). Furthermore, stratified sampling was carried out at District levels. Districts were chosen based on Agro-climatic zones (Districts that represent key Agro climatic zones). Lastly, households were selected using simple random sampling. In summary, it can be deduced that multi-staged Stratified sampling was used for the population. This method is appropriate. Although the study catered for biases, some of the sources of bias are self-selection bias and non-response bias. Although the Institutions (at the country level) were not the direct respondents in the survey, them expressing interest could have opened the study to the risk of Self-selection bias. Also, with regard to non-response bias, some selected respondents were unable to participate as intended.
3. A) What's the effect of the amount spent on extension services on farming yield?

B) What adaptation strategies are commonly used by farmers to cope with climate changes
4. For question A) above, Statistical analysis should commence by measuring central tendency using the Median. Then in order to quantify whether the

outcome of the survey is not unusual from the population, further analysis should be conducted using a hypothesis test.

Null hypothesis: No effect on farming yield; Alternative Hypothesis: Has effect on farming yield.

Hypothesis Test: Pearson's Correlation coefficient should be used to test for correlation between the amount spent on extension services and the farming yield, considering that both are continuous variables. The significance level is set at 0.05, if the probability is higher than this level, then we reject the null hypothesis.

5. If a significant effect occurred, decision to reject the null hypothesis must be made. Nevertheless, in addition to the statistical significance evaluation, the practical significance must also be evaluated before a conclusion can be made.
6. Scatter Plots could be used to see the relationship (pattern) between the data set. Boxplots could also be used especially for those with outliers, e.g. one of the countries had an odd set of extension services. Histogram could be used whereby the cost of extension services is placed in bins/intervals and this is plotted to check the frequency of yield within each interval (cost).
7. Comparing datasets (FAO/World Bank data on national average values and survey data on fertilizer use, pesticide use and irrigated area) could allow for data validation and over time develop richer data sets for benchmarking.