

Statistics summative

Firstly, the method used to gather samples is assessed. Agricultural data was collected from several countries, so the population was firstly split into countries. Next districts were chosen, which have high agriculture activity. And households with varying farm sizes were selected. As such it looks as follows: countries (11) > districts (30-60) > farm types (2-5) > households (5 – 10). From this we can see that the sample was clearly chosen according to stratified sampling. Data is first divided into clusters after which the population is divided into smaller groups according to set criteria, households are then randomly selected and measured.

The method used allowed the study to cover a large area and diverse types of farming. There was however a bias towards places where fertile land occurred, as districts where farming is less common were not sampled at all, also only one country (Egypt) was measured in north Africa providing a dataset more skewed towards western Africa.

The two key questions that I chose are:

- Are meat cattle more likely to graze on own land than milk cattle?
- Are the daily wages paid for small, medium and large-scale farms the same?

For question 3 the daily wages paid for small, medium and large-scale farms will be compared.

The null hypothesis is that there is no difference in the average wages being paid for small, medium and large-scale farms.

The alternate hypothesis is that the size of the farm has a significant impact on daily wages.

To reduce the impact of exchange rates and governing type on the results only one country will be chosen for this analysis, this will reduce sample size but also eliminate important uncontrolled variables.

The test to be used for this experiment is the ANOVA test.

The ANOVA test is chosen because we have 3 **independent** groups. Our goal is to **Discriminate** between the groups to determine whether there is a significant difference between them.

Significance will be determined the F and p values are considered. If the p-value is lower than a pre-determined value (0.05) and practical significance suggests otherwise the null hypothesis is rejected.

If a significant effect resulted the following can be concluded: The daily wages paid on a farm is likely to depend on the scale of the farm being worked on. Because a significant result was noted from the ANOVA test, in other words the means of at least 2 of the data sets differed significantly we can discard the null hypothesis. Many other socio-economic factors can influence this test as this test only looks at 4 independent variables and many more uncontrolled variables can possibly have significant impacts on the dataset, like household income.

If results were not significant the following conclusion can be reached: It is unlikely that the scale of the farm has any significant impact on the wages being paid, as all three groups have a median that is within a tolerable margin of each other the null hypothesis is kept.