Design of experiments, Major Project requirements.

The role of the project is to put the content of design and analysis of experiments in the context of agricultural research.

**Measuring seedlings**

We will be measuring the stem of the seedling.

* Gentle stretch out the seedling and measure from below the leaf stem to the top of the root.
* Record the measurement in mm.
* This is indicated by the red arrows.

A suggested work flow is:

1. 1 person will pull the seedling.
2. 1 person will give it a wash.
3. 1 person will measure the seedling.
4. 1 person will record the results.

More than one person can measure. If you decide to do this, you will need to ensure that you conduct within and between assessor reliability.

Within assessor reliability

* Each measurer will do this.
* Randomly select 5 seedlings
* When you have finished measuring them the first time put them aside.
* Measure them for a total of three times through the process at random times.

Between assessor reliability

* You will measure the 5 seedlings your measuring partner has put aside, and they will measure yours.
* It does not matter which person is the expert for the analysis.

Remember….it is essential that you record **which** seedlings you measure so they can be traced back to the seed on the sticky paper.

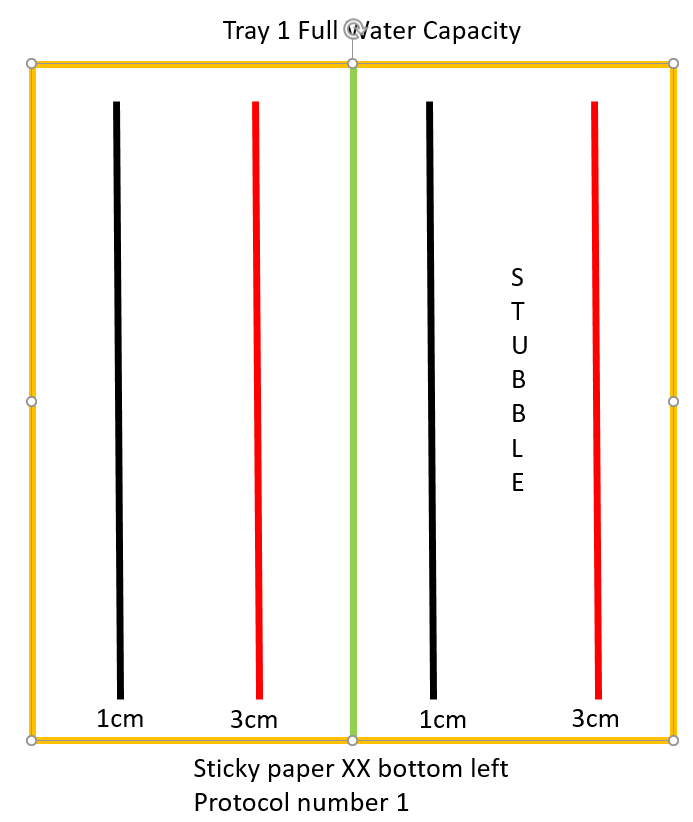
**Confirmation of the setup.**

This was a complex practical. To carry out multiple tasks with speed, efficiency, with a degree of complexity is always going to result in some errors. **This is not detrimental to your marks**. Your discussion of what was done is more important for your grade.

**This first part is required by the end of the practical 5pm Friday 04/06/2021.** For the group analysis, I will need to verify exactly what was done with your trays, as some mistakes may exclude your trays from the group analysis results. Hence, I require you to draw me a diagram clearly showing the following:

1. Each tray and the tray number.
2. The trays that received water saturation.
3. The trays that received no extra water.
4. The side of each tray that received stubble.
5. The placement of each planting row with its depth.
6. Your protocol number.
7. Before and after weights of weighed tray.
8. What seed sticky paper you used.
9. Clarification of the orientation of the sticky paper,
   1. i.e. which corner the written identifier was on.

For example….do this for each tray.



This will allow me to double check. **No marks will be taken if mistakes have been made.** It is important you let me know where you might have deviated from the plan.

**Important information for handing up**

* Your Genstat file must be present for your assignment to be marked.
  + If Genstat looks like it is having a fit, copy and paste the input log onto a word document and save that.
  + Save often.
* All files that you hand up must be appropriately labelled and start with your student number.
* There will be **5 marks** allocated to formatting, spelling, layout etc.
* English as a second language students will not be penalised for grammar, however you do have a spelling check in Microsoft Word.
* This is a question and answer format. Not a report format.

**Questions**

1. While there are many contrived examples given in university to teach students, this is not one of them. This trial/experiment is highly relevant to the agricultural industry. In a paragraph of no more than 300 words describe: (10 marks)
   1. The current understanding of the affect of each factor explored in this experiment on canola seed emergence, vigor, and final yield.
      1. Planting depth, stubble presence, and soil moisture at sowing are the factors.
   2. How you expect the response variables to react to the factors chosen for this experiment (emergence, seedling stem length)?
   3. Explain why seed size is a reasonable covariate choice.
      1. References are required for all of Q1.
2. State the null and alternative hypothesis, for seedling shoot length only, either in words or symbols such that the aim of the experiment would be obvious to someone who is unaware of what you are doing (2 marks).
   1. Explain and state the null value in your answer.
      1. Remember there are multiple comparisons.
      2. If you use symbols they must be clearly defined.
3. Briefly describe (5 marks).
   1. What blocking is?
   2. Why it is important in agricultural trials?
   3. What the blocks are in this experiment?
   4. What other variables could be used as blocks in other trials an agriculturalist may run in their field? 3 examples are required with a brief explanation as to why they would make good blocking choices.
4. Internal and external validity are extremely important to understanding the results of any trial/experiment (5 marks).
   1. Describe what internal and external validity is.
   2. Describe any potential violations of internal or external validity.
   3. Describe the measures that this trial took to ensure internal and external validity.
   4. If completely generalizable means all canola plants in Australia, and completely non-generalizable means only to the specific plants in your trays, what level of population can your results be generalised to, and why?
   5. Give your results for the within and between assessor validity with comments and cut off points.
5. We made sure that each person was to do the same job for each tray, rather than one person doing all jobs for a tray each to avoid confounding of treatments with people (3 marks).
   1. What is confounding?
   2. Specifically detail, with respect to the mean squares of the treatments and residuals, how confounding can affect the results.
6. Recreate the protocol in Genstat used to create this trial/experiment (7 marks)
   1. In the same manner as I provide you in the workbooks, create a step by step protocol to create a similar design as the one you used, such that students next year could follow it and replicate what you have done.
      1. Note that it will not be *exactly* the same design due to randomisation.
      2. You will need to manually create some of the columns.
      3. This is potentially the hardest bit. I will give hints and help for this one.
      4. Make sure the columns are re-coded such that the factors show the actual values rather than just 1 or 2. I.e want “1cm” and “3cm” for depth, not 1 and 2.
      5. Make sure your seed number for the design randomisation is included.
   2. Give the skeleton design with the degrees of freedom and include them in your document.
   3. Hand up a \*.gsh file with your design on it.
7. Analyse your results, for both seed emergence and seedling shoot length, and include the following (15 marks)
   1. If you made mistakes in your set up, this is where you would detail what mistakes were made and how they would affect the outcome.
   2. Run the power analysis by using the results from your experiment for comparing tray level factor effects. Use the residual error estimate from your analysis for your variance estimate.
      1. Briefly describe what power is and why it is important.
      2. What would occur to the power if you want to detect a larger, and a smaller, treatment mean difference.
   3. What the Treatment structure is (this will be the same for both response variables).
   4. What the Block Structure is (this will be the same for both response variables).
      1. Include a brief description of what nesting (big hint here for blocking structure) is and why it applies here.
   5. What level of interaction do you think is reasonable to ask for (this will be the same for both response variables).
      1. Include a brief description of what interactions are and when it is appropriate to not ask for them in an analysis, and why that might be desirable.
      2. You will need to talk about the SS, df here.
      3. Give the interaction plot from Genstat with proper axis labelling.
   6. Explain your results in context for both response variables.
      1. You will need to comment on more than just the p-values here.
      2. Discuss the percentage of variance different portions.
      3. Is it a good fitting model?
      4. Explain why there are two residual lines (within groups is the hint).
      5. Comment on each **testing** line.
      6. Present the means either in table form or graphically.
      7. Give the standard errors and comment on their size.
      8. Ask for the LSD results
      9. Give and comment on assumptions and residual plots.
8. Re-analyse your results including the covariate for both responses (5 marks).
   1. You will need to create a covariate column in Genstat and correctly allocate each seed to the appropriate result.
      1. You will be given a CSV file with the seed size and the photo of the seed.
      2. The algorithm randomised the seed number (not intentionally).
      3. You will need to sort this out.
   2. Explain what covariates are and why they are useful in the analysis.
   3. Describe and explain the covariate results.
      1. Discuss the standard errors with respect to the means.
      2. Discuss the percentage of variance different portions.
      3. Is this model a good fit?
   4. Compare the covariate results to the regular ANOVA results for seedling stem length only.
9. Describe a trial that you would do in your **future place of work** (7 marks).
   1. Describe:
      1. The response variable(s).
      2. Factors and levels (remember the levels do not need to be the same).
      3. Random effects.
      4. Blocking.
   2. Why would this trial would benefit you in your future place of work?