Statistical Analysis using R

Course overview and Requirements

IITA Biometrics Unit

01-02 & 07-09 December 2021

Overview

- This course is intended to give participants a foundational understanding of statistical analysis in the R language.

- General topics we will focus on include:

    - Creating and organizing projects

    - Wrangling and preparing data for analyses

    - Exploring data with graphics and summaries

    - Analyzing data

Overview

- By the end of this course, you should feel confident approaching data from the agricultural sector

- We will cover virtually no statistics

- However it is the intention that this course will leave you comfortable to

  - understand/solve statistical problems using **R**

  - interpret results generated by **R**

We assume:

- you installed **R**

- you installed **RStudio**

- you installed **Rtools** (Windows)

- you installed the necessary **libraries**

All is well!

Slide Formatting

Remember these:

Bold usually indicates an important vocabulary term.

`This` indicate emphasis but also are used to point out things you must click with a mouse, for example: `File > Save as...`

`Code` represents R code you type into the editor or console used to perform actions, for example

Code chunks that span the page represent *actual R code embedded in the slides*.

```r
# Sometimes important stuff is highlighted!
5 * 4
```

```
## [1] 20
```

The lines preceded by ## represent the output, or result, of running the code in the code chunk. We'll talk about this more later!

Logistics

Morning Sessions: Lectures

- From 9:00 am to 1:00 pm with a break of 20 min

- All participants mic to be muted. But questions in the chat

- A troubleshooting room to be opened. Raised the issue in the chat and we move you there

- After each *important* topic, **Your turn** sessions: short time to allow participants practice the concepts just showed

Materials

All course materials will be made available. This includes:

- These slides (**html**)

- PDFs of slides if you like those

- Files (**xlsx** and **csv**) for the practical

- References and Useful links to other resources

A dropbox link will be shared

Course overview

1. Short Introduction to R and RStudio

2. Preparation of Data for Statistical Analysis

3. Data wrangling

4. Experimental Designs for Plant Breeding

5. ANOVA and MET analysis

6. Multivariate analysis

7. Graphics in R with ggplot2