



Session 1: Introduction to R and R studio

Turkey Ouma, Meklit Chernet



Background

- This material aims to provide a manual that will be used by the ACAI NARS partners to train R users. The questions we asked ourselves before coming up with this content were:
 - What would be the best way to reach a high number of users with the training content?
 - And what would be the best way to make sure the lessons learnt at the training remain active?

The training team came up with 2 solutions:

1. Stepping the training down to researchers who are running field experiments this year and who will be facing problems to be solved that are part of the training package
2. The participants of the ToT implement the lessons learnt to a hypothetical problem as part 2 of the training.

Who would benefit from this training?

Those who have attended the previous trainings and have a clear grasp of the basics of R. This is a linear training in the sense that it solves some of the most common issues that arise from data management and visualization of field trial data. This training uses data collected using ODK and hosted on ona.io.

Previous trainings

- May – July 2020: three virtual trainings “Introduction to R statistical programming language” --- Meklit Chernet (IITA) and Turry Ouma (IITA)
- November 2020: In-person training with focus on intermediate and advance levels --- Ibnou Dieng (IITA), Meklit Chernet (IITA), Turry Ouma (IITA) and Sam Ofodile (IITA)

Feedback

- There will be pre and post course online questionnaires hosted on ona.io to enable us get feedback to improve future training sessions.

Is R and Rstudio same?

- **R** is a programming language used for statistical computing and graphics. **Rstudio** is an integrated development environment (IDE) that provides an interface which allows users to write and edit programs in **R** by offering statistical packages, many convenient features and tools. You can use **R** on its own to write a program and run the code. It is however not the case with **Rstudio** since it can only be used together with **R**.
-
- In order to use **Rstudio** your **R** version should be 3.0.1 or higher. The most recent version of **R** can be downloaded from [CRAN](https://cran.r-project.org/), the Comprehensive R Archive Network.
-
- If you already have **R** and/or **RStudio** installed, let's get started!

Source

1

Environment

Console

R version 4.1.3 (2022-03-10) -- "One Push-Up"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

Files, plots, packages, help, viewer

Name	Size	Modified
gitignore	13 B	Apr 27, 2021, 8:50 AM
http-oauth	0 B	Apr 27, 2021, 8:50 AM
.R		
RData	1.7 MB	May 2, 2022, 8:17 PM
Renviron	41 B	Jul 8, 2021, 10:11 AM
Rhistory	23.6 KB	May 11, 2022, 9:46 AM
--SatLog ODK III Training 2020_07_03 13_01.rtf	162 B	Jun 15, 2021, 1:13 PM
--SEA_RewardPayments.xlsx	165 B	Apr 15, 2019, 9:00 AM
--Sgit.docx	162 B	Jun 15, 2021, 1:14 PM
--Sren.xlsx	165 B	Jul 26, 2021, 8:23 AM
442019		
542019		
06052019	27.2 MB	May 6, 2019, 9:37 AM
06052019.tar	18.4 MB	May 6, 2019, 9:37 AM
ACAI		
ACAI.rar	1.2 GB	May 10, 2019, 2:44 PM
Akilimo EA dashboard2_MC [Autosaved].pptx	1.7 MB	Jul 5, 2019, 1:34 PM
Apowersoft		
app.R	3.5 KB	Sep 16, 2020, 10:01 AM
BAND BUDGET AND SIGNATURES.docx	16.3 KB	Apr 16, 2021, 10:44 AM



Before we start...



Set the Working directory

Create a folder called ACAI_R_training

Navigate to the file path and copy the path

Use `setwd()` to set the directory: Paste the path in between the parentheses: `setwd("D:\Training Materials\ACAI_R_training")`

In R, a single backslash is an escape character, and using it for directory paths will always produce an error.

Specify directory paths correctly in R, using a forward slash:
`setwd("D:/Training Materials/ACAI_R_training")`



Install relevant packages



Let's use "tidyverse" as an example

1. Install it from CRAN with: `install.packages("tidyverse")` or from GitHub: `devtools::install_github("tidyverse/tidyverse")`. OR in the Files pane of RStudio: Click on the "Packages" tab, Click on "Install", Type the name of the package under "Packages (separate multiple with space or comma):" Click "Install"

For GitHub you may need to install the {devtools} package. You can do this from CRAN with `install.packages("devtools")` then load the devtools package using `library(devtools)`.

In most cases just use `install_github("author/package_name")` as shown in the example above.

1. Load it for use in R with: `library("tidyverse")`.
2. Get help on it with: `package? tidyverse` and `help(package = " tidyverse")`.

Prepare data

Read the data in R

- We will use {readr} which is a part of the core {tidyverse} and supports seven file formats with seven read_ functions. For our case these two apply: read_csv() and write_csv().

```
read_csv("data/dataVAL_FR_PO.csv")
```

- An alternative is to use read.csv which is inbuilt in base R and reads the data as a dataframe.

```
read.csv ("data/dataVAL_FR_PO.csv", header=T, na.strings=c("", "NA"))
```

- Now let's assign this data frame to a variable so that it is stored in R's memory:

```
dataVAL_FR_PO <- read_csv("data/dataVAL_FR_PO.csv")
```