An analysis that reopens a year later

RStudio projects, here(), renv and the two habits — setwd() and a saved workspace — that make an R analysis run on exactly one machine on exactly one day.

Pierre Robentz Cassion

Lesson 1 of 8

R and the Tidyverse for Programme Data — A project that reopens

What this lesson covers

- The failure this lesson prevents
- The three habits to drop
- The project
- `here()` rather than relative paths
- renv: the package versions, recorded
- Loading packages
- Verifying the environment before trusting it
- What to hand to a colleague
- What comes next

The failure this lesson prevents

- Someone asks you to explain last quarter's figure.

The three habits to drop

- `setwd()`

The three habits to drop — In R

```
setwd("/Users/marie/Desktop/nutrition")
```

The three habits to drop

- `rm(list = ls())` at the top
- Restart R instead — Ctrl/Cmd + Shift + F10 in RStudio
- A saved workspace

The three habits to drop — Example

```
Tools -> Global Options -> General
  Restore .RData into workspace at startup    [ ]
  Save workspace to .RData on exit            Never
```

The project — Example

```
muac-analysis/
  muac-analysis.Rproj
  data/
    raw/            the export exactly as it arrived, never edited
    interim/
  outputs/
    tables/
    figures/
  R/
    read_register.R
    indicator_table.R
  renv.lock
  README.md
  .gitignore
```

The project

- `data/raw` is read-only — The export as it arrived is the only thing you cannot reproduce; everything else is...

`here()` rather than relative paths — In R

```r
library(here)

muac <- readr::read_csv(here("data", "raw", "muac-screening-artibonite-2024.v1.csv"))
```

`here()` rather than relative paths — In R

```
here()
#> [1] "/home/marie/muac-analysis"
```

renv: the package versions, recorded — In R

```
install.packages("renv")

renv::init()       # once per project
```

renv: the package versions, recorded — In R

```
renv::snapshot()    # after installing or upgrading anything
renv::restore()     # on another machine, or a year later
```

renv: the package versions, recorded

- Commit `renv.lock` — It is the difference between "install the tidyverse" and "install the versions that produced this...
- Installing where there is no internet

renv: the package versions, recorded — In R

```
renv::init()
renv::snapshot()

# Fill a local cache with everything the lockfile names.
renv::install()
renv::isolate()
```

renv: the package versions, recorded — In R

```r
# On the connected machine
dir.create("pkgs")
download.packages(c("dplyr", "readr", "haven"), destdir = "pkgs", type = "source")

# On the offline machine
install.packages(
  c("dplyr", "readr", "haven"),
  repos = NULL,
  type = "source",
  contriburl = paste0("file://", normalizePath("pkgs"))
)
```

Loading packages — In R

```
library(readr)
library(dplyr)
```

Loading packages

- Do not use `require()` in a script — It returns `FALSE` and continues when the package is missing, so the script fails...
- Do not call `install.packages()` from a script — A script that installs software as a side effect of being run is a...

Loading packages — In R

```
muac |> dplyr::filter(muac_mm < 125)
```

Verifying the environment before trusting it — In R

```
sessionInfo()
```

Verifying the environment before trusting it — In R

```
stopifnot(getRversion() >= "4.2.0")
stopifnot(requireNamespace("dplyr", quietly = TRUE))
```

What to hand to a colleague

Commit	Do not commit
`R/`, `.Rproj`	`renv/library/`
`renv.lock`	`data/raw/` — see below
`README.md`	`outputs/`
`.gitignore`	`.Rhistory`, `.RData`

What to hand to a colleague — Example

```
.Rproj.user/
.Rhistory
.RData
renv/library/
data/raw/
outputs/
```

What to hand to a colleague

- Never commit raw beneficiary data — identified or pseudonymised

What to hand to a colleague — In R

```r
token <- Sys.getenv("DHIS2_TOKEN")
stopifnot(nzchar(token))
```

What comes next

- The project reopens and the packages are pinned.

Read the full lesson, with runnable code

https://data-analysis.cassion.dev/courses/r-for-programme-data/r-01-projects-and-renv