

(Very) few notes on using Rstudio

Maria del Carmen Montero Calasanz
Lea A. I. Vaas
Benjamin Hofner
Markus Göker

Why use ?

- RStudio is an integrated development environment (IDE) for R.



An IDE that was built
just for R

- Syntax highlighting, code completion, and smart indentation
- Execute R code directly from the source editor
- Quickly jump to function definitions



Bring your workflow
together

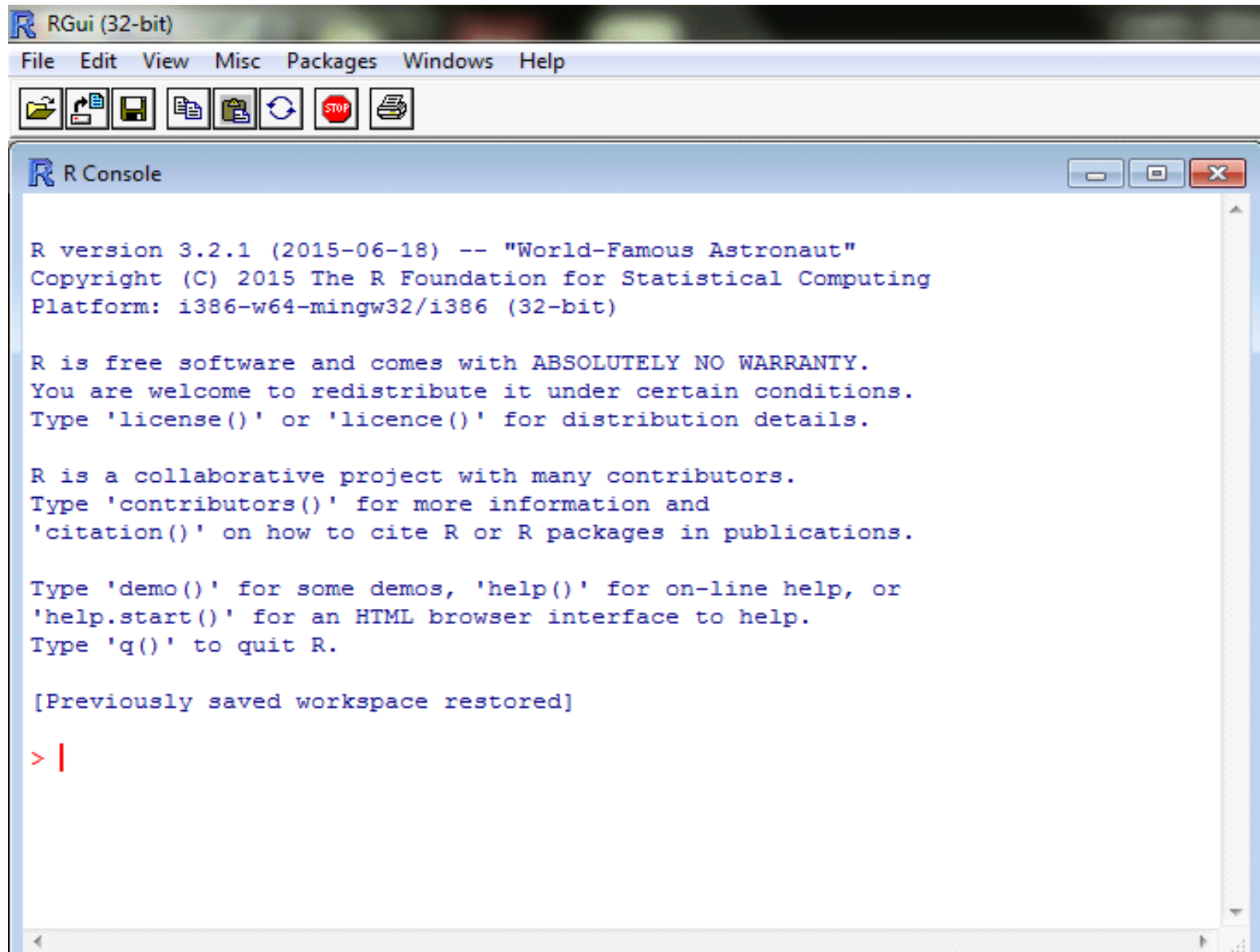
- Integrated R help and documentation
- Easily manage multiple working directories using projects
- Workspace browser and data viewer



Powerful authoring &
Debugging

- Interactive debugger to diagnose and fix errors quickly
- Extensive package development tools
- Authoring with Sweave and R Markdown

This is plain R



This is Rstudio

The screenshot displays the RStudio IDE interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Tools, and Help. Below the menu is a toolbar with icons for file operations and running code. The main editor window shows an R script with the following code:

```
1 library(pkgutils) # needed for mypdf()
2 library(opm) # version >= 1.2.28
3
4
5
6
7 # Extract and clean strain and replicate information from CSV data in OPMS
8 # object 'x'. Assume either 'Strain Number', 'Sample Number', 'Strain Name' or
9 # 'Strain Type' contain this information. Assume the part after the last space
10 # is the replicate ID, the part before it the strain ID.
11 #
12 find_strain_and_replicate <- function(x) {
13   clean <- function(x) {
14     x <- gsub("[^\\w\\s]", "", x, perl = TRUE)
15     x <- gsub("\\s+", " ", x, perl = TRUE)
16     sub("\\s+", "", sub("\\s$", "", x, perl = TRUE), perl = TRUE)
17   }
18   stopifnot(is(x, "OPMS"), is.matrix(x <- csv_data(x)))
19   stopifnot((sn <- "Strain Number") %in% colnames(x))
20   x <- as.data.frame(x, stringsAsFactors = FALSE)
21   for (name in c("Sample Number", "Strain Name", "Strain Type")) {
22     if (!any(empty <- !nzchar(x[, name])))
23       break()
24     if (pos <- match(name, colnames(x), nomatch = 0L))
25       x[, empty, sn] <- x[, empty, pos]
26   }
27   x[, sn] <- clean(x[, sn])
28   replicate <- sub("\\s$", "", x[, sn], perl = TRUE)
29 }
```

The Environment pane on the right shows the Global Environment with two values: 'x' (Large OPMS (2 elements, 527.2 Kb)) and 'x1' (Large OPMS (2 elements, 566.3 Kb)). It also lists a function 'find_strain_and_r...'.

The File Explorer pane on the right shows the directory structure: C:\Users\sam\Desktop>csv2>BMG133. It lists files: BMG133xA_PM01_allHours_kinetic.csv (226.9 KB, Aug 10, 2015, 9:08 PM), BMG133xB_PM01_allHours_kinetic.csv (227.6 KB, Aug 10, 2015, 9:09 PM), xyplot.pdf (269.3 KB, Aug 12, 2015, 1:08 PM), and data.yml (907 KB, Aug 12, 2015, 6:28 PM).

The Console at the bottom shows the following output:

```
> setwd("C:/Users/sam/Desktop/csv2/BMG133")
> source("C:/Users/sam/Desktop/agosto/opm session/doit.part1.R")
Strain Replicate
[1,] "BMG133" "A"
[2,] "BMG133" "B"
warning message:
In .local(object, ...) : running times are not uniform
>
```

Working with files

The screenshot displays the RStudio interface with a script editor on the left and a file explorer on the right. The script editor shows a function `find_strain_and_replicate` that processes CSV data. The file explorer shows a directory structure with a context menu open over the `csv2` folder.

Script Editor Content:

```
1 library(pkgutils) # needed for mypdf()
2 library(opm) # version >= 1.2.28
3
4
5
6
7 # Extract and clean strain and replicate information from CSV data in OPMS
8 # object 'x'. Assume either 'Strain Number', 'Sample Number', 'Strain Name' or
9 # 'Strain Type' contain this information. Assume the part after the last space
10 # is the replicate ID, the part before it the strain ID.
11 #
12 find_strain_and_replicate <- function(x) {
13   clean <- function(x) {
14     x <- gsub("[\\w\\s]", "", x, perl = TRUE)
15     x <- gsub("\\s+", " ", x, perl = TRUE)
16     sub("^\\s", "", sub("\\s$", "", x, perl = TRUE), perl = TRUE)
17   }
18   stopifnot(is(x, "OPMS"), is.matrix(x <- csv_data(x)))
19   stopifnot((sn <- "Strain Number") %in% colnames(x))
20   x <- as.data.frame(x, stringsAsFactors = FALSE)
21   for (name in c("Sample Number", "Strain Name", "Strain Type")) {
22     if (!any(empty <- !nzchar(x[, sn])))
23       break()
24     if (pos <- match(name, colnames(x), nomatch = 0L))
25       x[, empty, sn] <- x[, empty, pos]
26   }
27   x[, sn] <- clean(x[, sn])
28   replicate <- sub("^(\\s+)$", "", x[, sn], perl = TRUE)
29   data.frame(x[, sn], replicate, stringsAsFactors = FALSE)
30 }
```

Environment Panel:

- Global Environment
- Values
 - x: Large OPMS (2 elements, 527.2 Kb)
 - x1: Large OPMS (2 elements, 566.3 Kb)
- Functions
 - find_strain_and_r... function (x)

File Explorer:

- C: > Users > sam > Desktop > csv2 > E
- Name
- ..
- BMG133xA_PM01_allHours_kinetic
- BMG133xB_PM01_allHours_kinetic
- xyplot.pdf
- data.yml

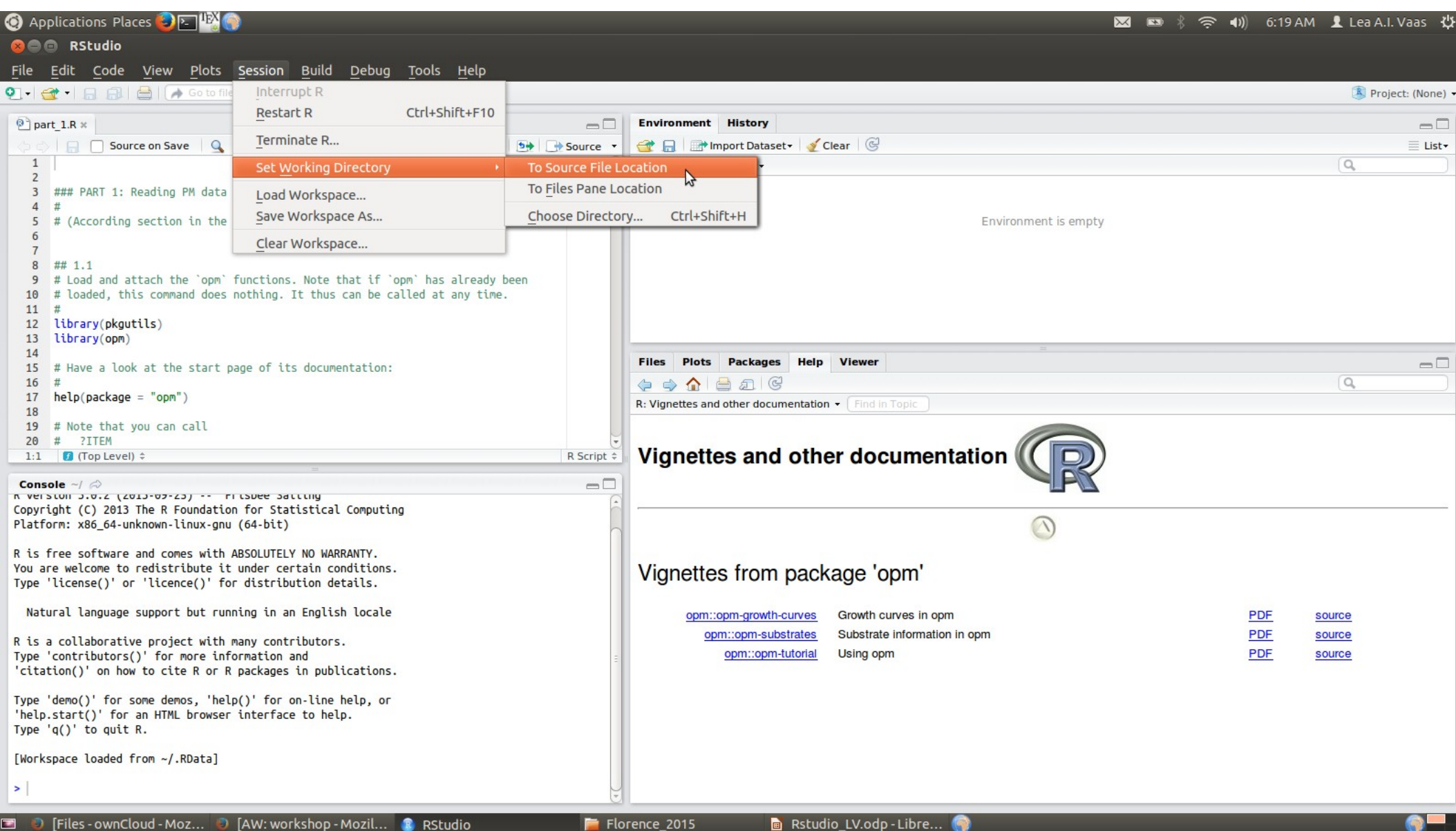
Context Menu:

- Copy...
- Move...
- Set As Working Directory
- Show Folder in New Window

Console:

```
C:/Users/sam/Desktop/csv2/BMG133/
> setwd("C:/Users/sam/Desktop/csv2/BMG133")
> source('C:/Users/sam/Desktop/agosto/opm session/doit.part1.R')
Strain Replicate
[1,] "BMG133" "A"
[2,] "BMG133" "B"
warning message:
In .local(object, ...) : running times are not uniform
>
```

Setting the working directory



The screenshot shows the RStudio interface with the 'Session' menu open. The 'Set Working Directory' option is highlighted, and its submenu is also open, showing 'To Source File Location' as the selected option. The 'Environment' pane is empty, and the 'Files' pane shows the R vignettes page.

Session Menu Options:

- Interrupt R
- Restart R (Ctrl+Shift+F10)
- Terminate R...
- Set Working Directory**
 - To Source File Location
 - To Files Pane Location
 - Choose Directory... (Ctrl+Shift+H)
- Load Workspace...
- Save Workspace As...
- Clear Workspace...

Environment Pane:

- Environment
- History
- Source
- Import Dataset
- Clear
- Environment is empty

Files Pane:

- Files
- Plots
- Packages
- Help
- Viewer
- R: Vignettes and other documentation
- Find in Topic
- Vignettes and other documentation
- Vignettes from package 'opm'
- [opm::opm-growth-curves](#) Growth curves in opm [PDF](#) [source](#)
- [opm::opm-substrates](#) Substrate information in opm [PDF](#) [source](#)
- [opm::opm-tutorial](#) Using opm [PDF](#) [source](#)

Console:

```
R version 3.0.2 (2013-09-23) -- FriSep Setting
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-unknown-linux-gnu (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.

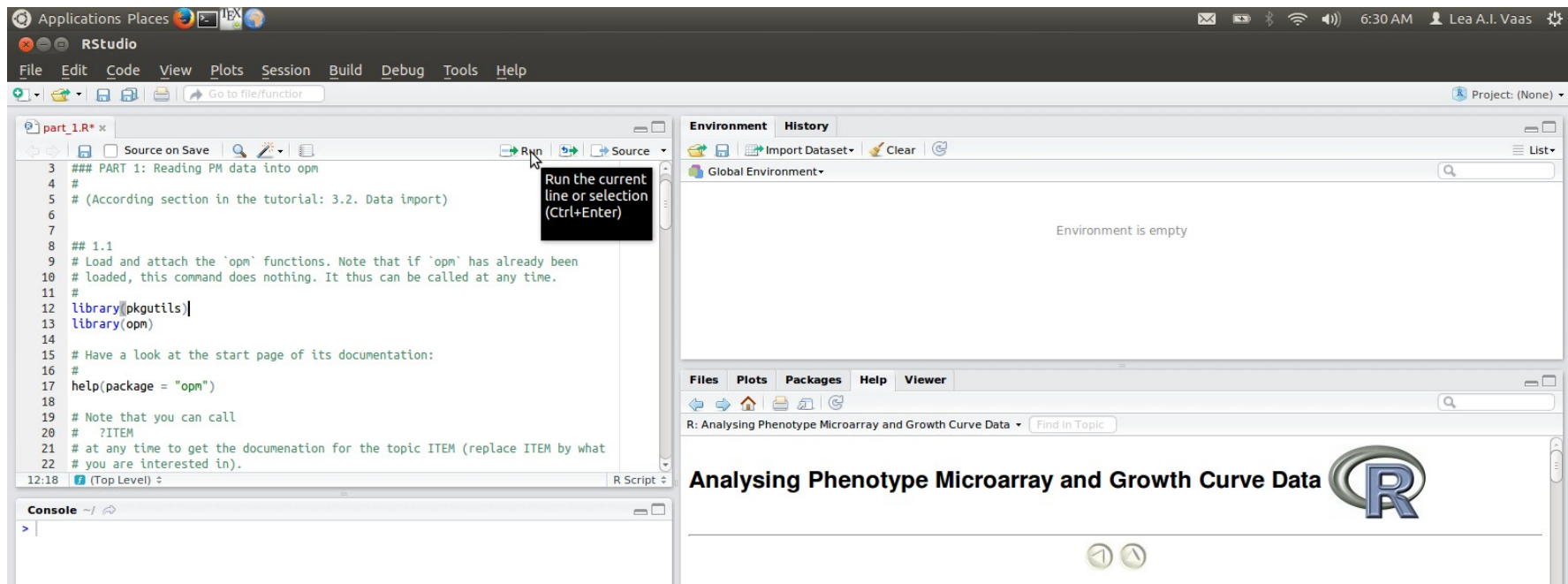
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

> |
```

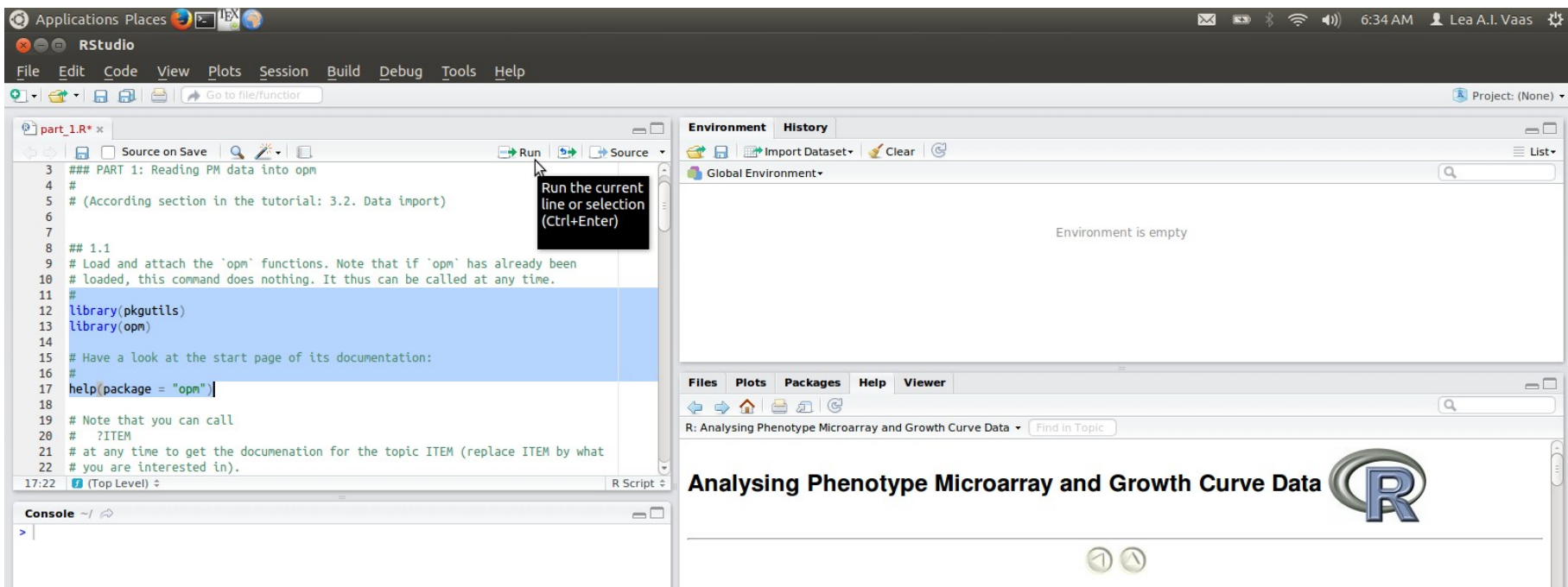
Executing R commands

- Execute only the line the cursor currently is in:
Ctrl + Enter
or press “run” button
(cursor then automatically jumps to the next line)



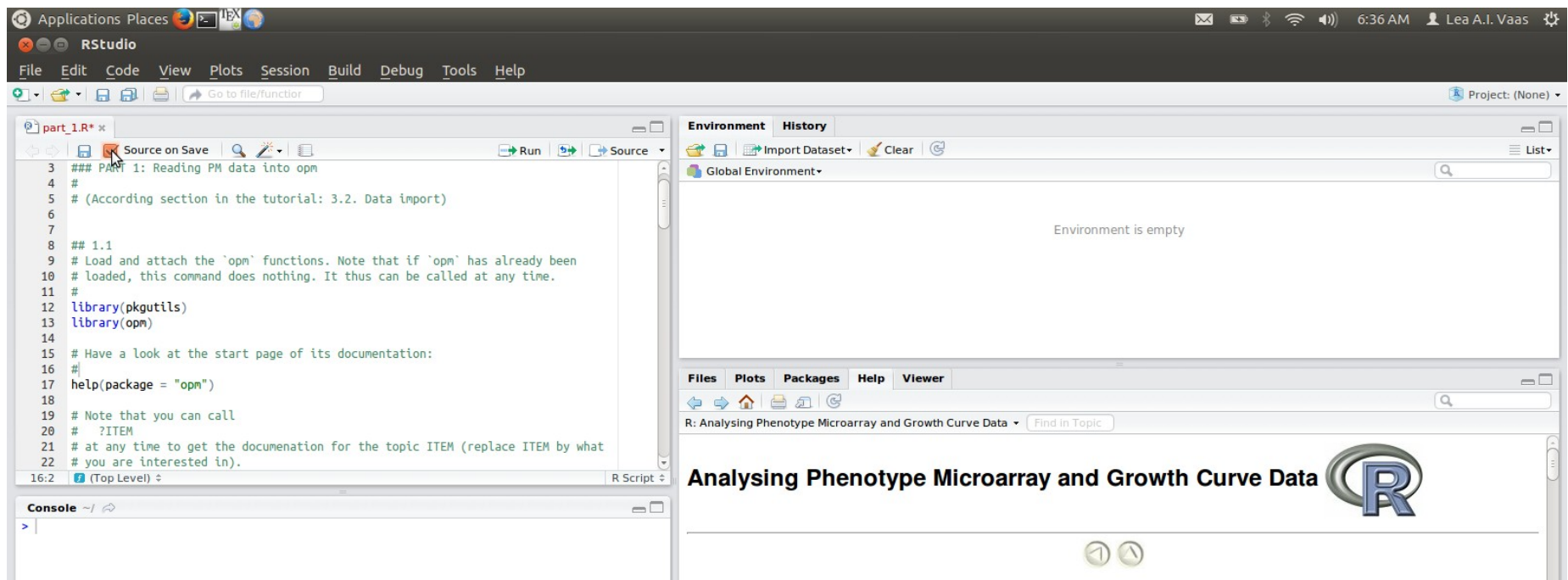
Executing R commands

- Execute several lines:
mark the lines and push Ctrl + Enter or the “run” button
(cursor does not move)



Executing R commands

- Execute a complete script:
push the “source” button OR tick box “Source on Save” on the Script tab, then click the save button



(Un-)comment lines

- Anything after “#” is not executed

press Ctrl+Shift+C to enter a comment or deactivate lines of code (or to activate them again)

