Getting Started with R

3 September 2021

Molly Lewis

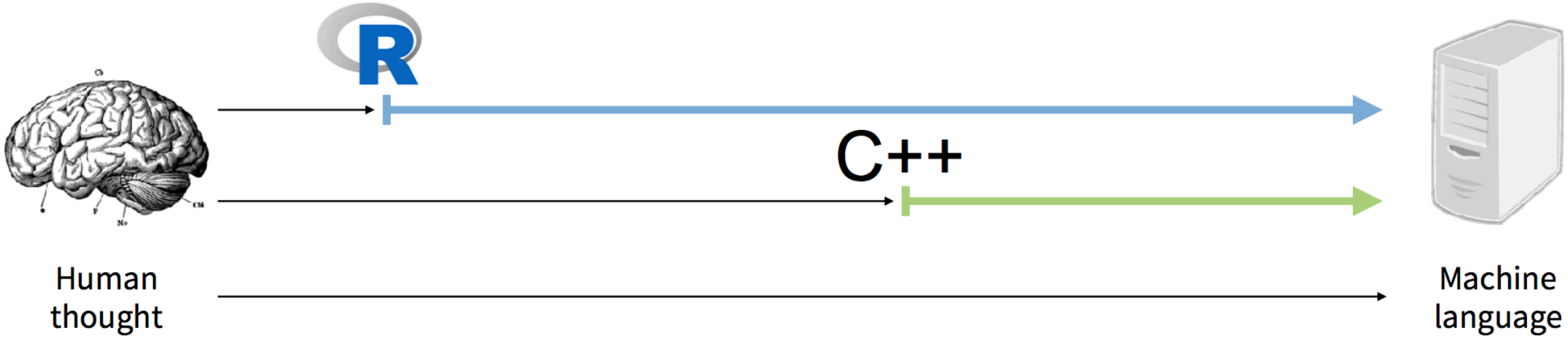
Lab

Artwork by @allison_horst

Why R for working with data

- Free and open-source
- Programming language (not point-and-click)
- Excellent graphics
- Easy to generate with reproducible reports
- Easy to integrate with other tools

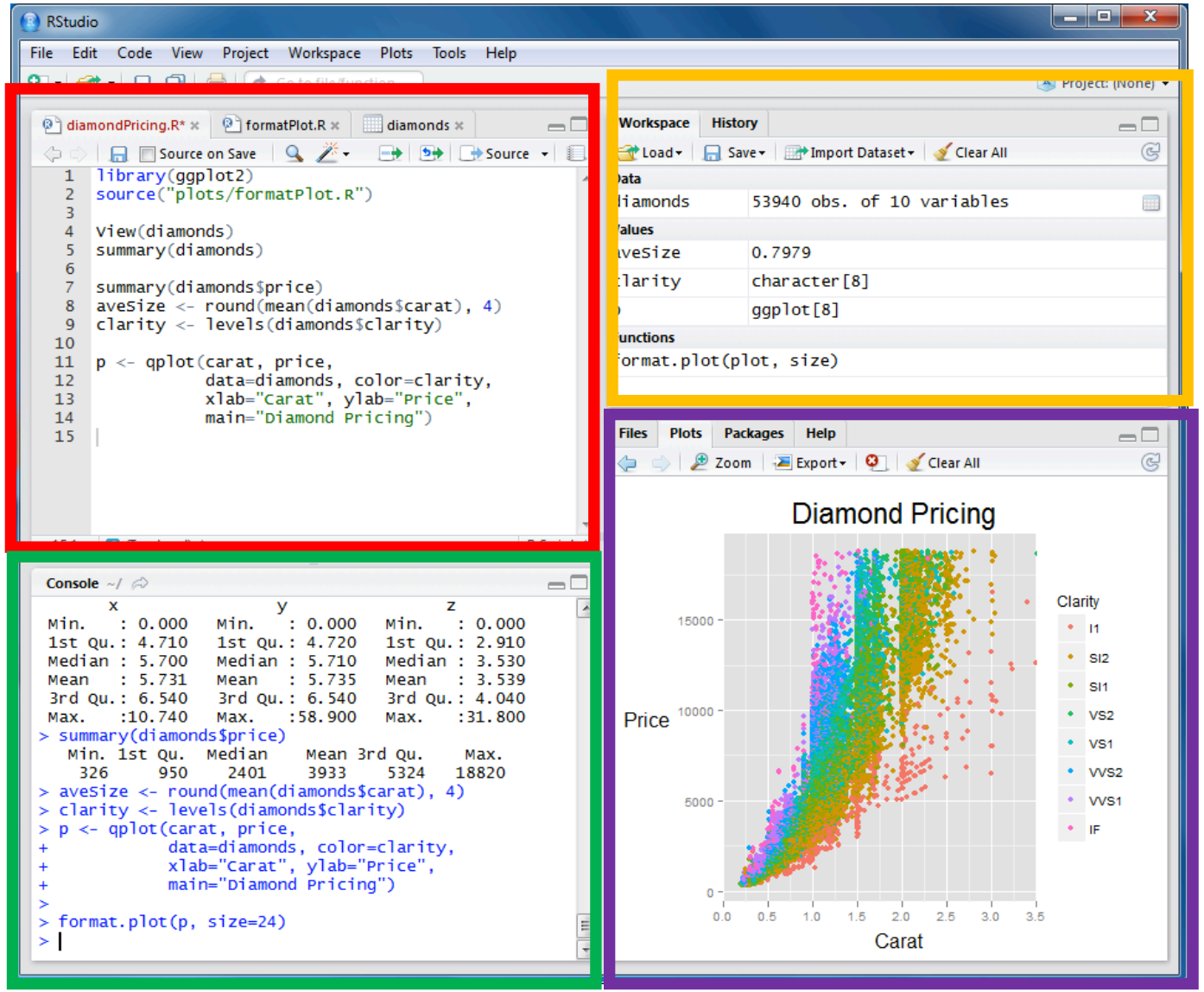
R - A computer language for scientists



Rstudio Integrated Development Environment (IDE)

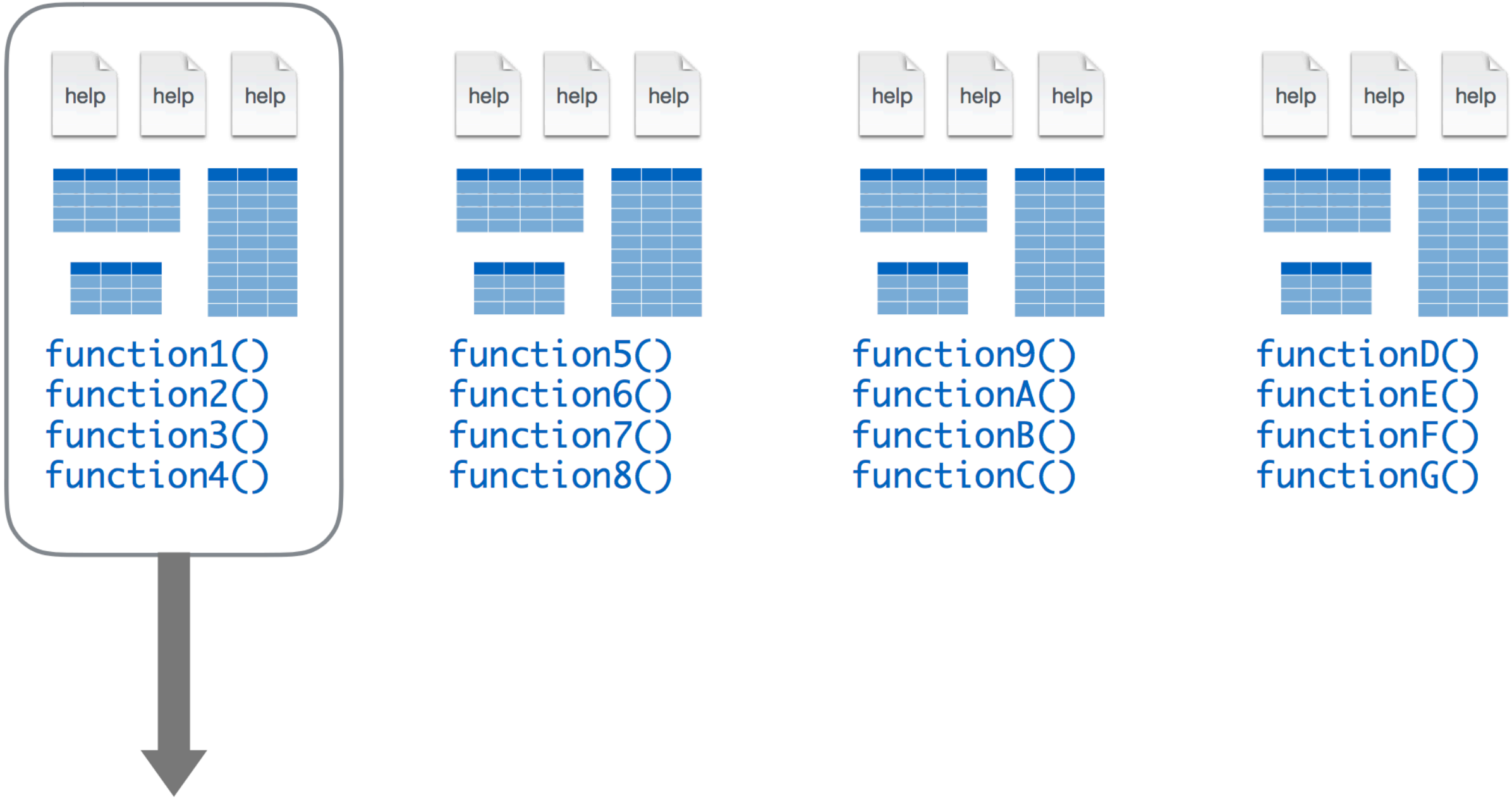
Source: edit file that you can run again later.

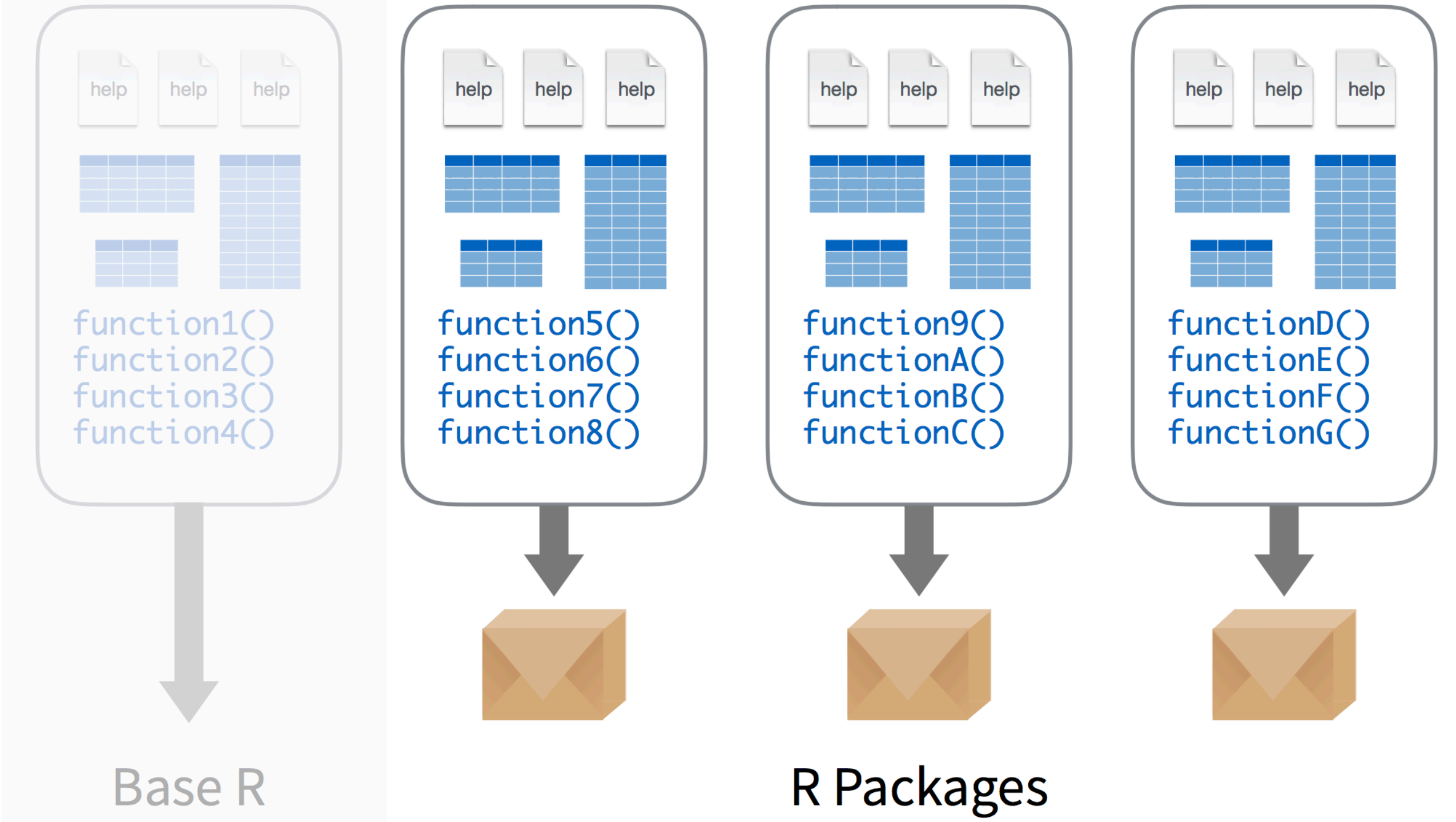
Console: type/paste commands to get output from R



Workspace/History: see list of variables and previous commands

Files/Plots/Packages/Help


help
help
help
function1()
function2()
function3()
function4()
help
help
help
function5()
function6()
function7()
function8()
help
help
help
function9()
functionA()
functionB()
functionC()
help
help
help
functionD()
functionE()
functionF()
functionG()
Base R


help
help
help
function1()
function2()
function3()
function4()
Base R
help
help
help
function5()
function6()
function7()
function8()
help
help
help
function9()
functionA()
functionB()
functionC()
help
help
help
functionD()
functionE()
functionF()
functionG()
R Packages

The Comprehensive R Archive

Garrett

Secure https://cran.r-project.org

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed

Available CRAN Packages By Name

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
abbyyR	Access to Abbyy Optical Character Recognition (OCR) API
abc	Tools for Approximate Bayesian Computation (ABC)
ABCanalysis	Computed ABC Analysis
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
abcdeFBA	ABCDE_FBA: A-Biologist-Can-Do-Everything of Flux Balance Analysis with this package
ABCoptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2
ABC.RAP	Array Based CpG Region Analysis Pipeline
abcrf	Approximate Bayesian Computation via Random Forests
abctools	Tools for ABC Analyses
abd	The Analysis of Biological Data
abf2	Load Gap-Free Axon ABF2 Files
ABHgenotypeR	Easy Visualization of ABH Genotypes
abind	Combine Multidimensional Arrays
abjutils	Useful Tools for Jurimetrical Analysis Used by the Brazilian Jurimetrics Association
abn	Modelling Multivariate Data with Additive Bayesian Networks
abodOutlier	Angle-Based Outlier Detection

Using packages

1

```
install.packages("foo")
```

Downloads files to computer

1 x per computer

2

```
library("foo")
```

Loads package

1 x per R Session

Tidyverse

The [tidyverse](#) is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures.


tidyverse
data visualisation
ggplot2
dplyr
data wrangling
reading data
readr
tibble
modern data frames
string manipulation
stringr
forcats
dealing with factors
data tidying
tidyr
purrr
functional programming

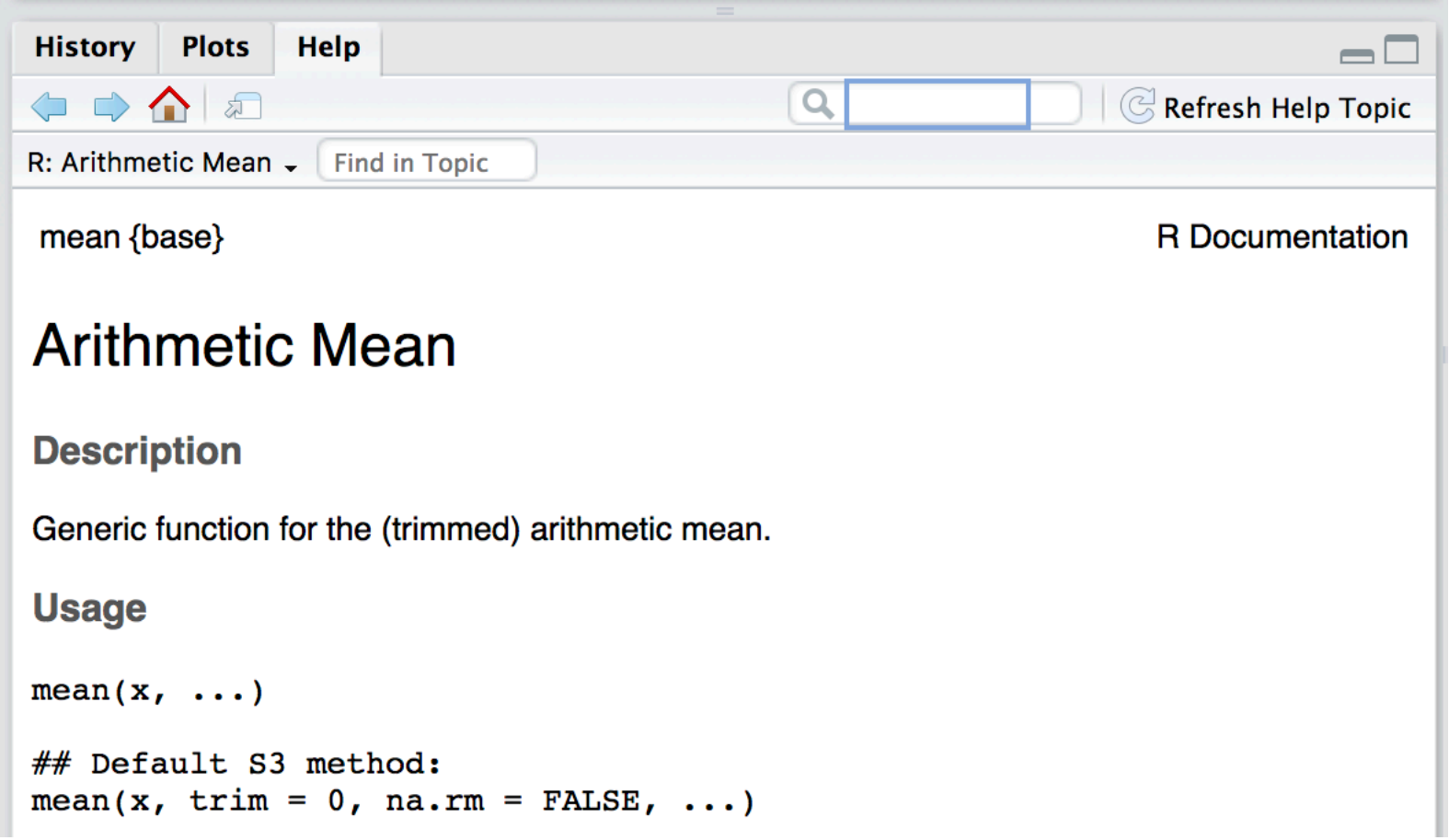
tidyverse

An R package that serves as a short cut for installing and loading the components of the tidyverse.

```
library("tidyverse")
```

Getting help

1. R help files



Anatomy of an R help file

Two ways to access:

1. Peruse in Help pane
2. ?<name> in console

The name of the function, and the library it is in.

mean {base} R Documentation

Arithmetic Mean

Description

What it does.

Generic function for the (trimmed) arithmetic mean.

Usage

```
mean(x, ...)

## Default S3 method:
mean(x, trim = 0, na.rm = FALSE, ...)
```

The function's name, and in the parentheses the named arguments it expects, in the order it expects them. If an argument has a default value, it is shown. Arguments without default values (e.g. x) must be provided by you.

Arguments

More details on each named argument. This will tell you what class of thing each argument has to be—an object, a number, a data frame, a logical value, etc.

`x`	An R object. Currently there are methods for numeric/logical vectors and date, date-time and time interval objects. Complex vectors are allowed for `trim = 0`, only.
`trim`	the fraction (0 to 0.5) of observations to be trimmed from each end of `x` before the mean is computed. Values of trim outside that range are taken as the nearest endpoint.
`na.rm`	a logical value indicating whether `NA` values should be stripped before the computation proceeds.
`...`	further arguments passed to or from other methods.

The ellipsis allows other arguments to be passed to and from the function.

Value

What the function returns—i.e., the result of whatever operation or calculation it performs. This can be a single number, as here, or a multi-part object such as a list, a data frame, a plot, or a model.

If `trim` is zero (the default), the arithmetic mean of the values in `x` is computed, as a numeric or complex vector of length one. If x is not logical (coerced to numeric), numeric (including integer) or complex, `NA_real_` is returned, with a warning.
If `trim` is non-zero, a symmetrically trimmed mean is computed with a fraction of `trim` observations deleted from each end before the mean is computed.

References

Becker, R. A., Chambers, J. M. and Wilks, A. R. (1988) *The New S Language*. Wadsworth & Brooks/Cole.

See Also

`weighted.mean`, `mean.POSIXct`, `colMeans` for row and column means.

Other related functions

Examples

```
x <- c(0:10, 50)
xm <- mean(x)
c(xm, mean(x, trim = 0.10))
```

Self-contained examples that you can run at the console. These may use built-in datasets or other R functions.

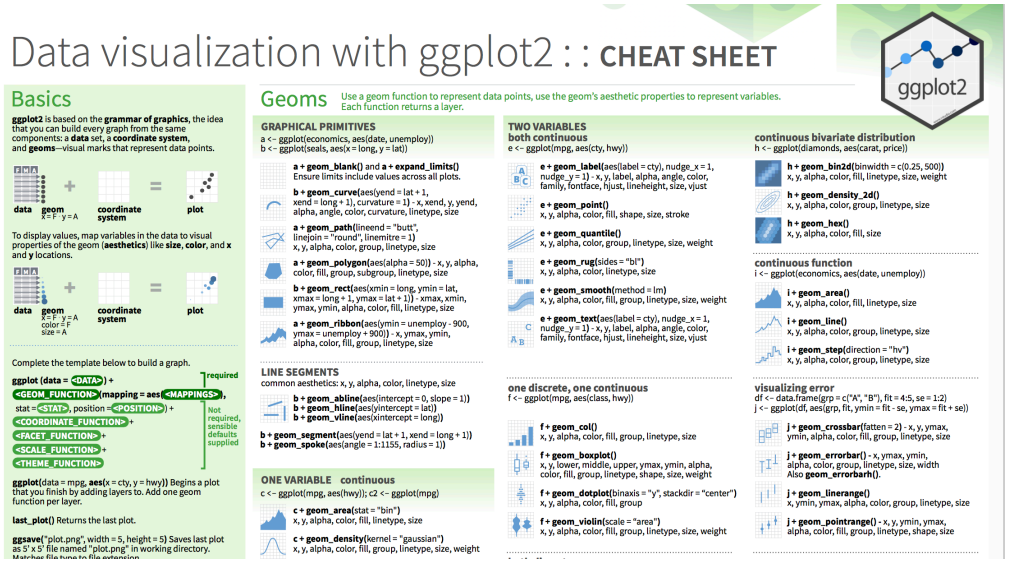
[Package *base* version 3.4.3 Index]

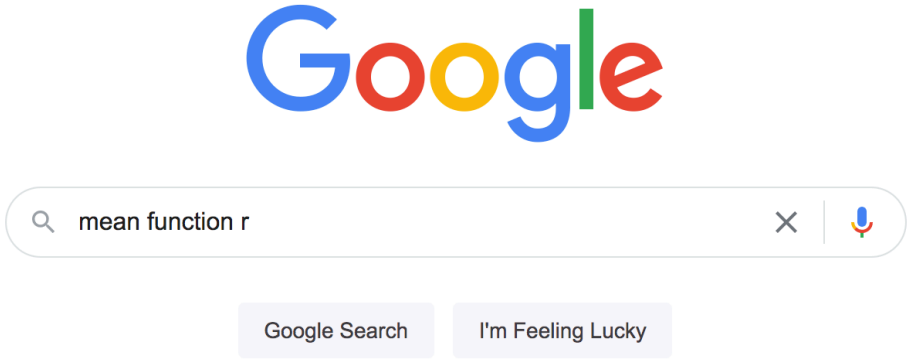
Visit the package's Index page to look for Demos and Vignettes detailing how it works.

Adapted from: https://socviz.co/appendix.html#a-little-more-about-r

Getting help

1. R help files
2. Cheatsheets (https://rstudio.cloud/learn/cheat-sheets)
3. Google!





Today

1. Make sure you have R, Rstudio and tidyverse installed on your laptop

2. Work through online tutorial on R programming basics: https://rstudio.cloud/learn/primers/1.2

3. (complete readings for today)

Acknowledgements

Slide 2 adapted from
https://www.andrew.cmu.edu/user/achoulde/94842/
by CC license

Slides 3-11, adapted from
https://github.com/rstudio-education/remaster-the-tidyverse/
By CC license

Slide 10 adapted from
https://education.rstudio.com/blog/2020/07/teaching-the-tidyverse-in-2020-part-1-getting-started/ by CC license