



HELLO
my name is

Rebecca

HELLO
my name is

Rob

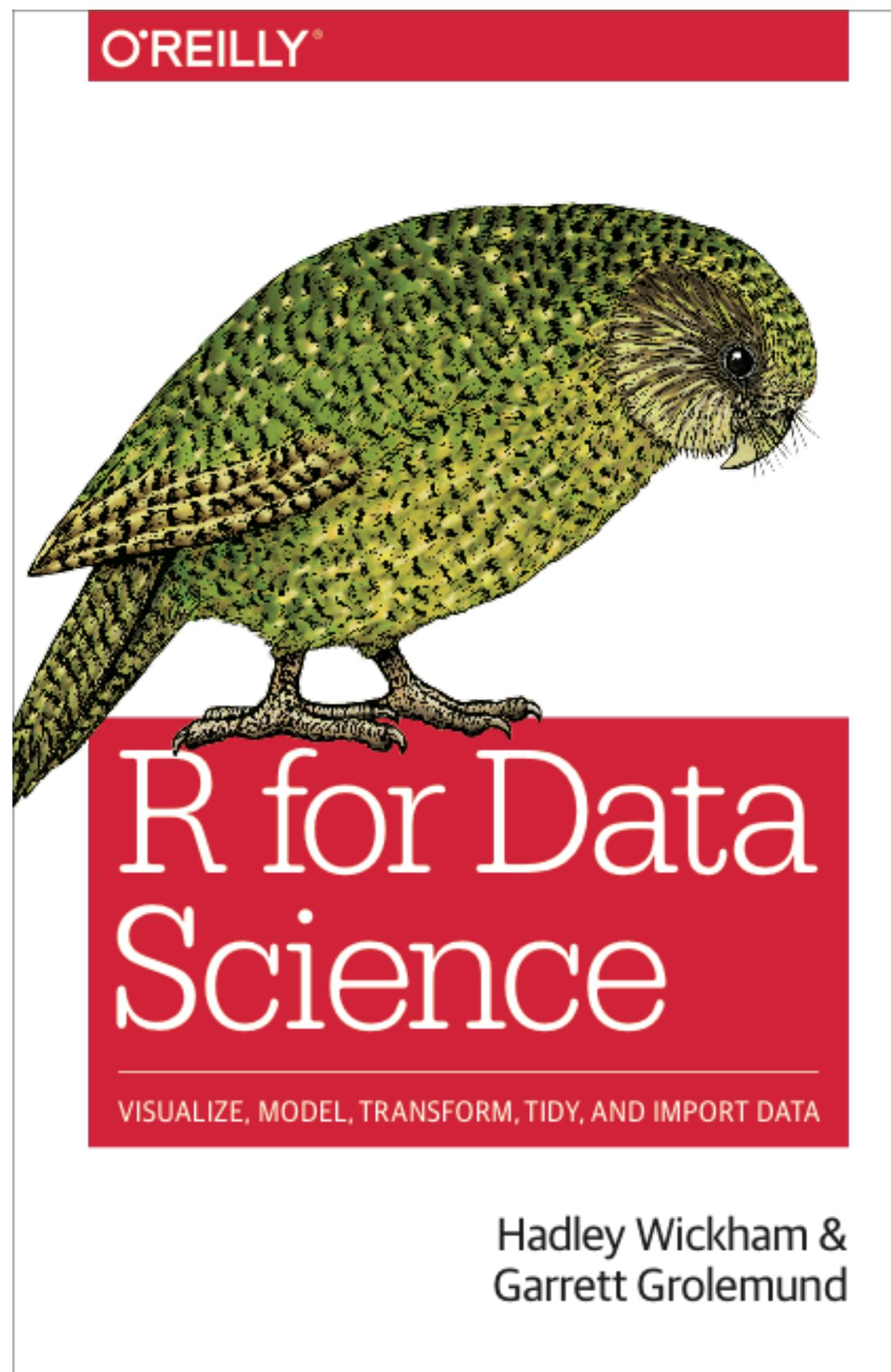
HELLO

my name is

TA

Day 2

09:00 am - 10:30 am	Introduction to R and Rstudio
10:30 am - 10:45 am	Morning Tea (provided)
10:45 am - 12:30 pm	Introduction to the R tidyverse - visualisation
12:30 pm - 1:30 pm	Lunch (BYO)
1:30 pm - 3:15 pm	Introduction to the R tidyverse - transform data
3:15 pm - 3:30 pm	Afternoon Tea (provided)
3:30 pm - 5:00 pm	Introduction to the R tidyverse - tidy data Wrap up



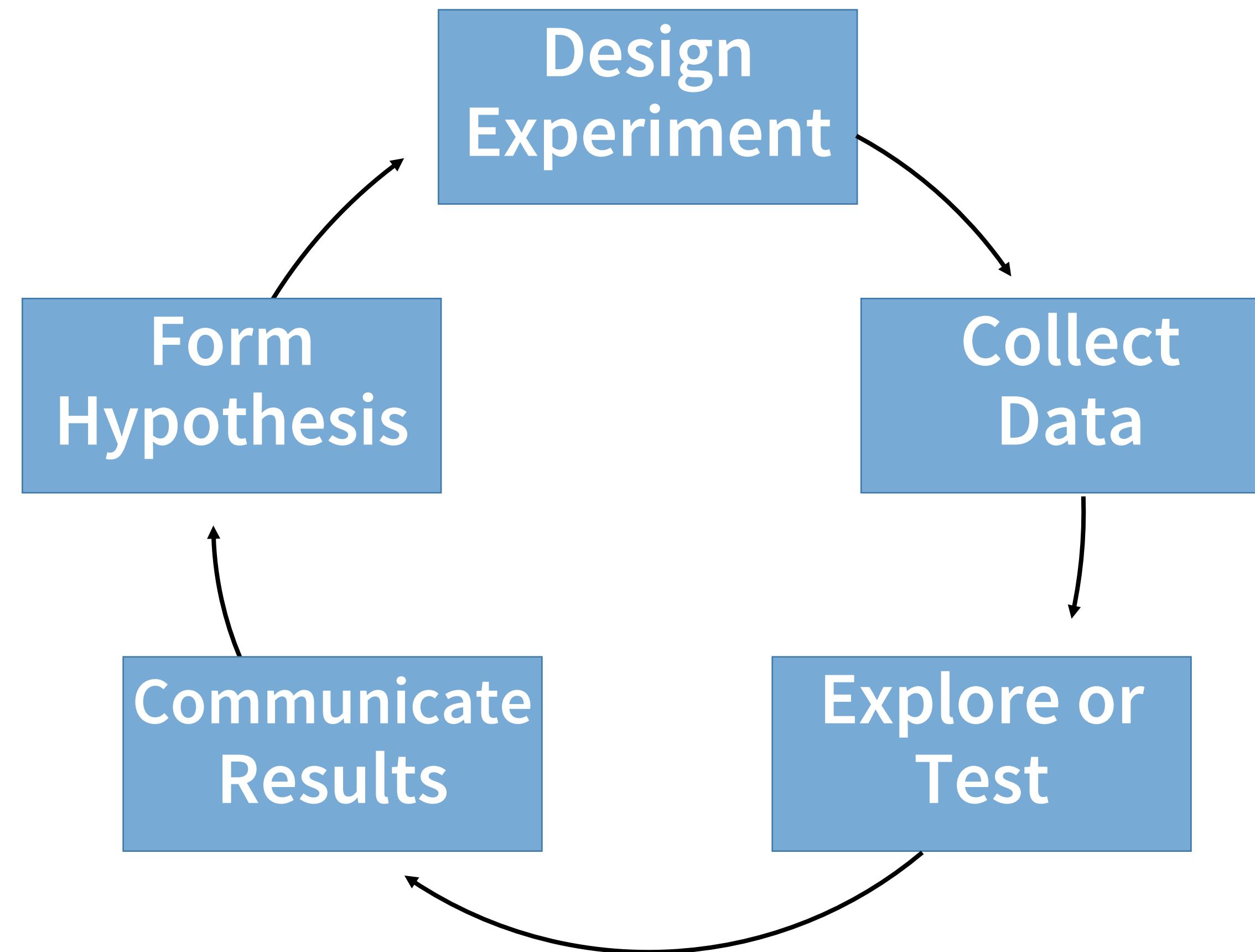
Your Turn

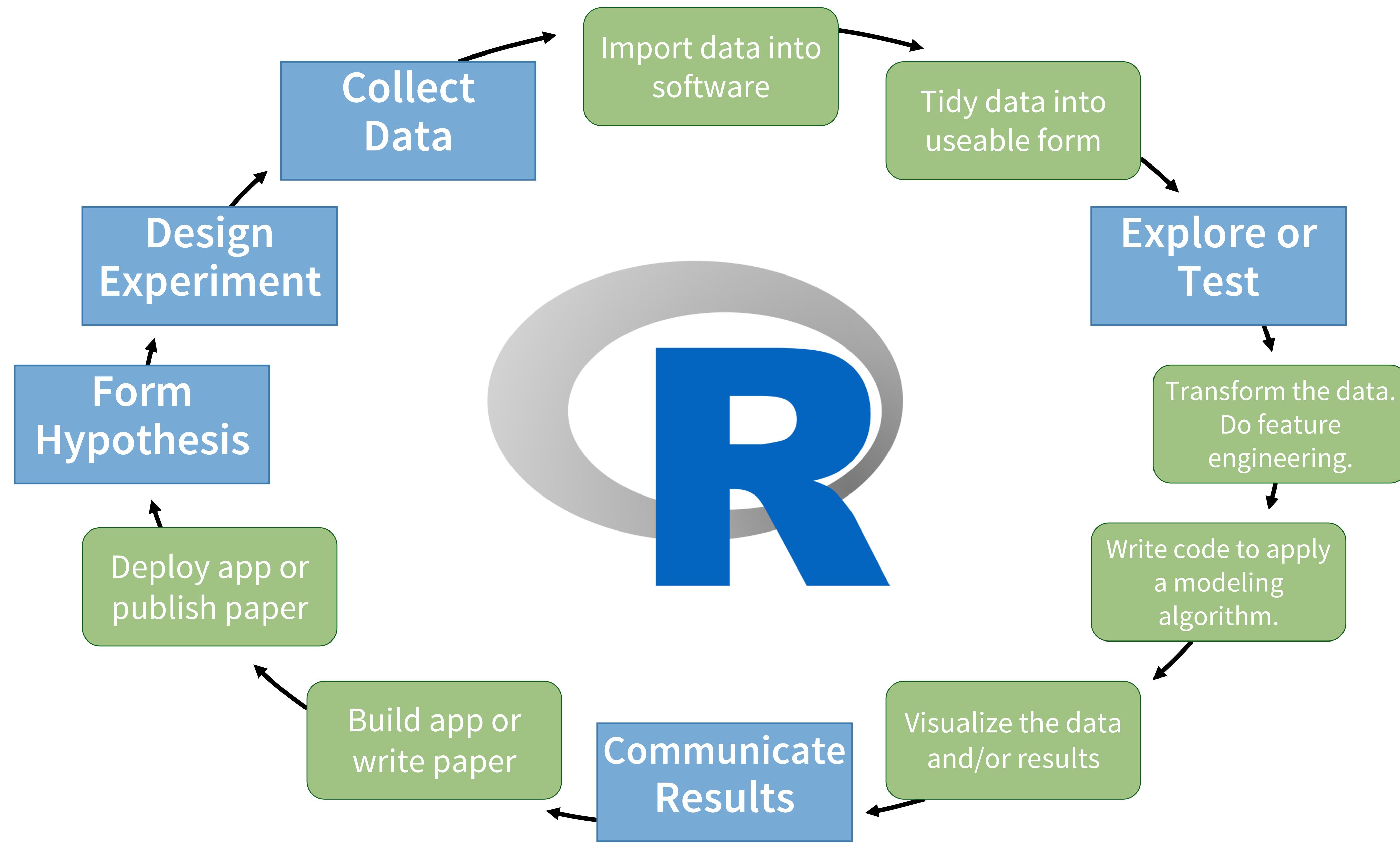
At your tables, re-introduce yourself to your group members. Tell them:

1. Who you are
2. Whether you have used R before
3. What you plan to do use R for

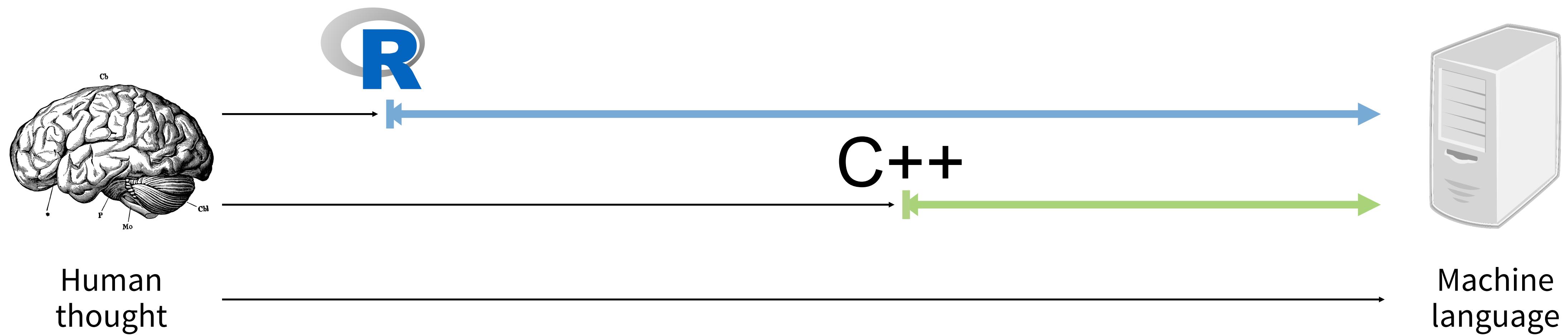


"Data Science"

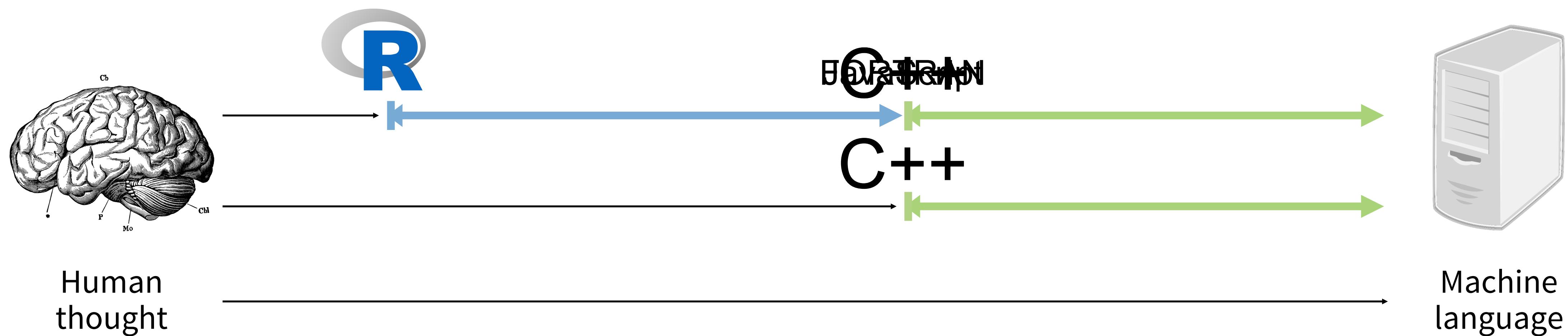




R - A computer language for scientists



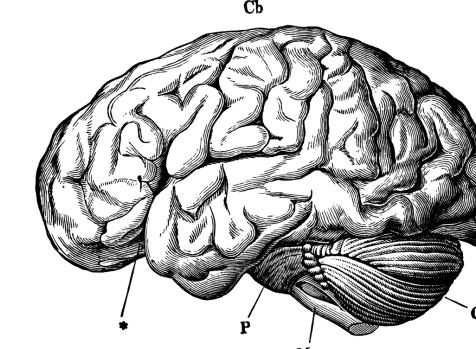
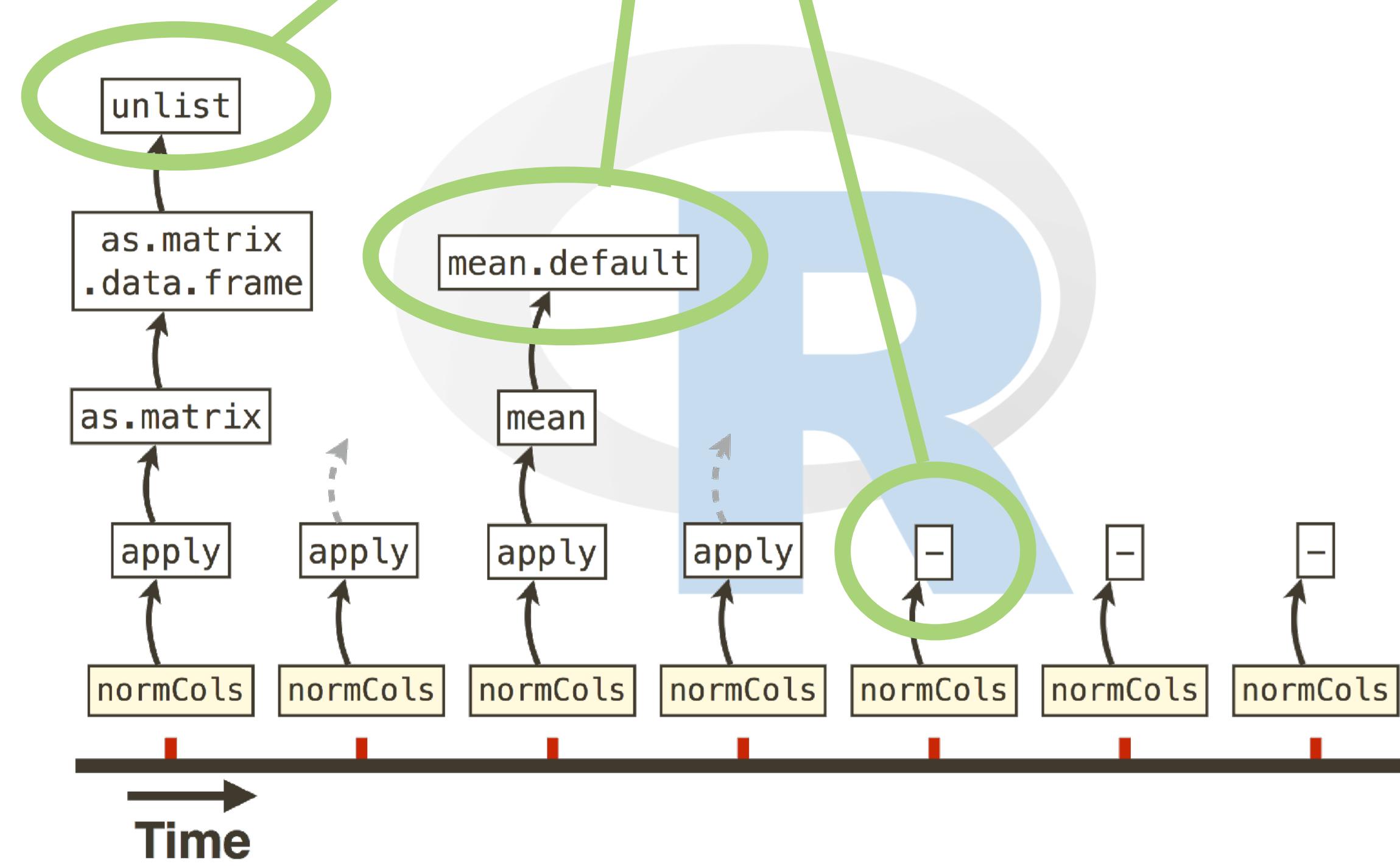
R - A computer language for scientists





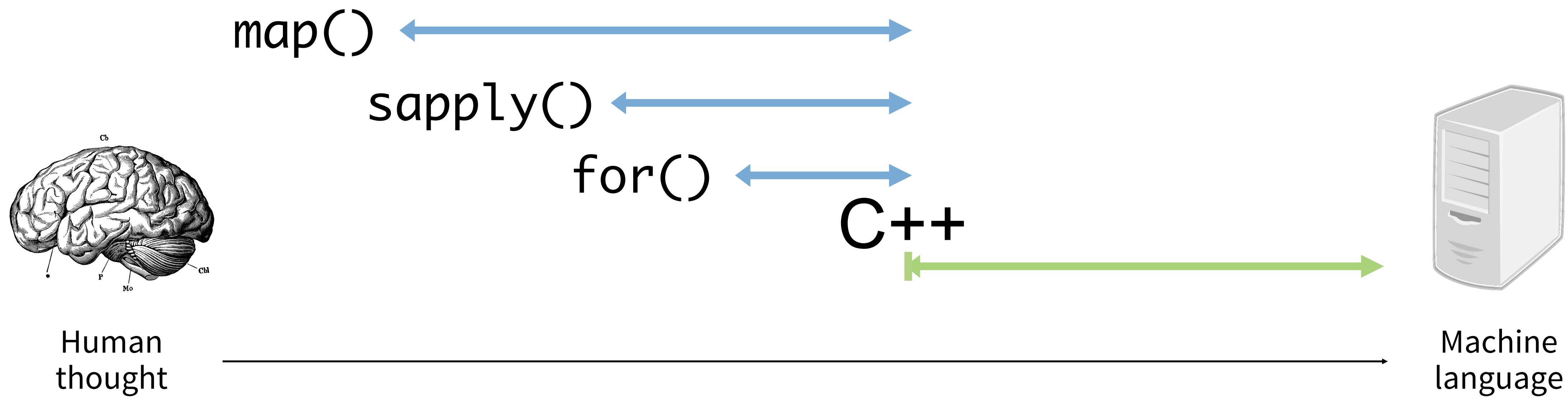
Machine language

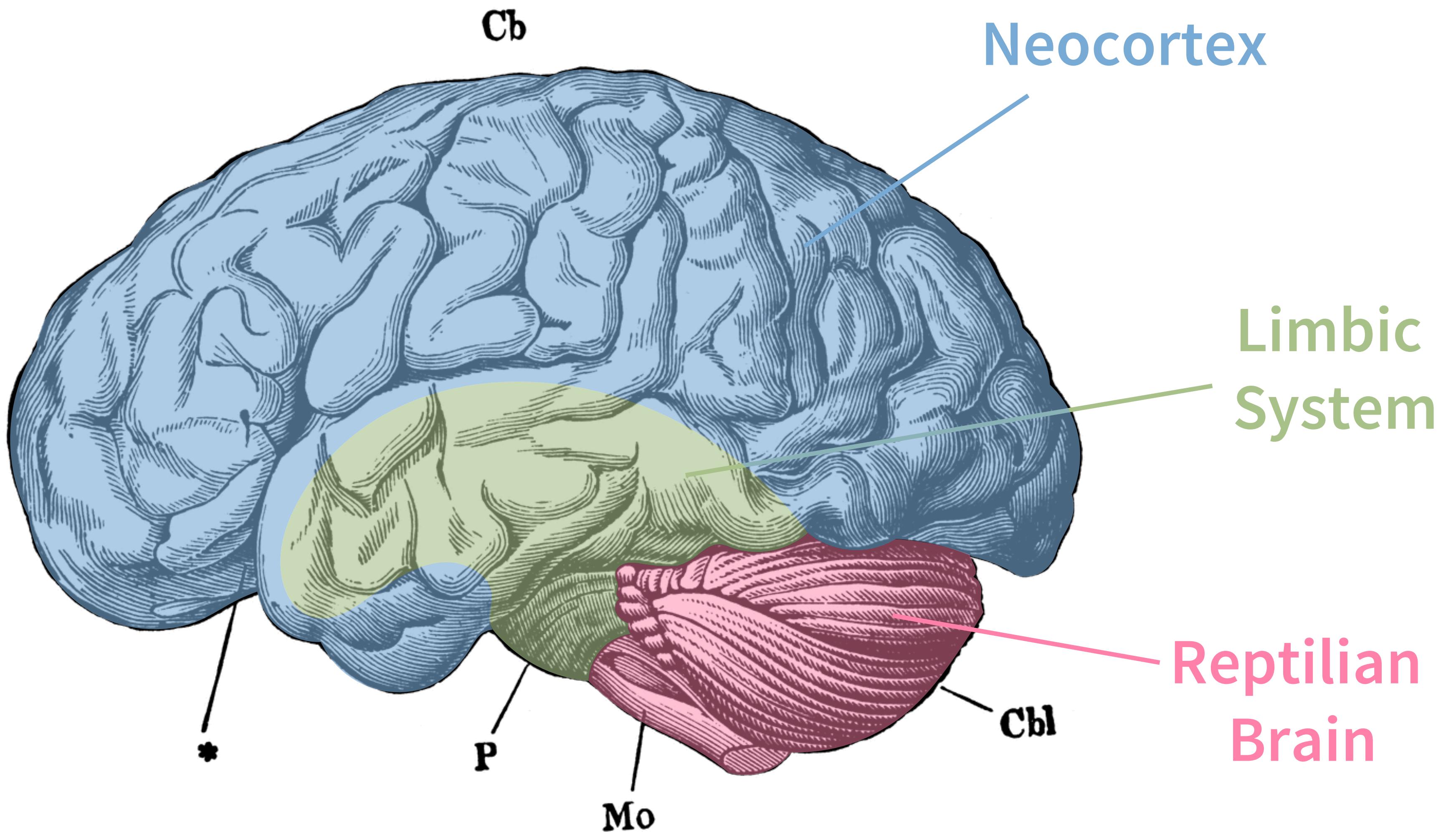
C++, FORTRAN, etc.

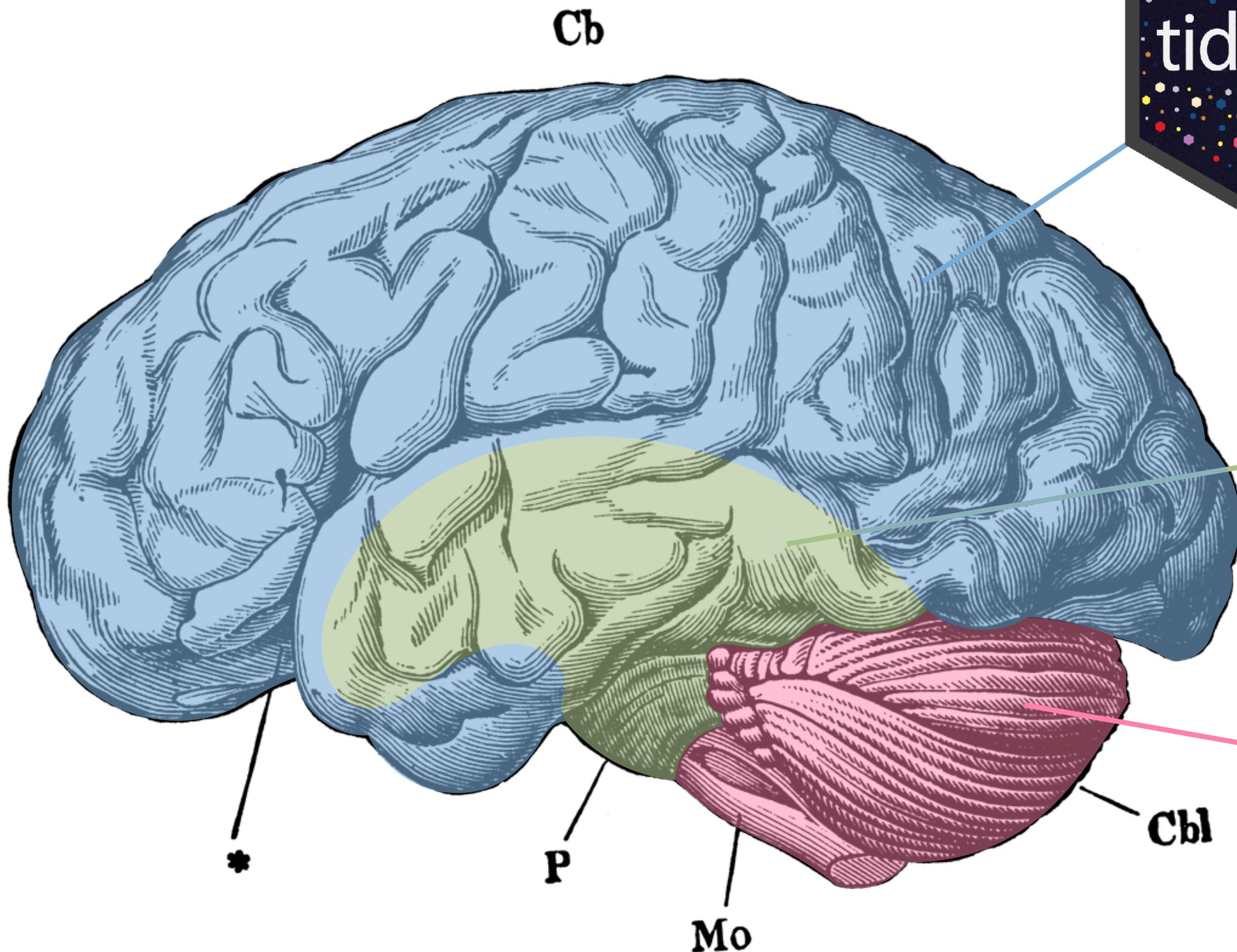


Human language

R - A computer language for scientists







Limbic
System

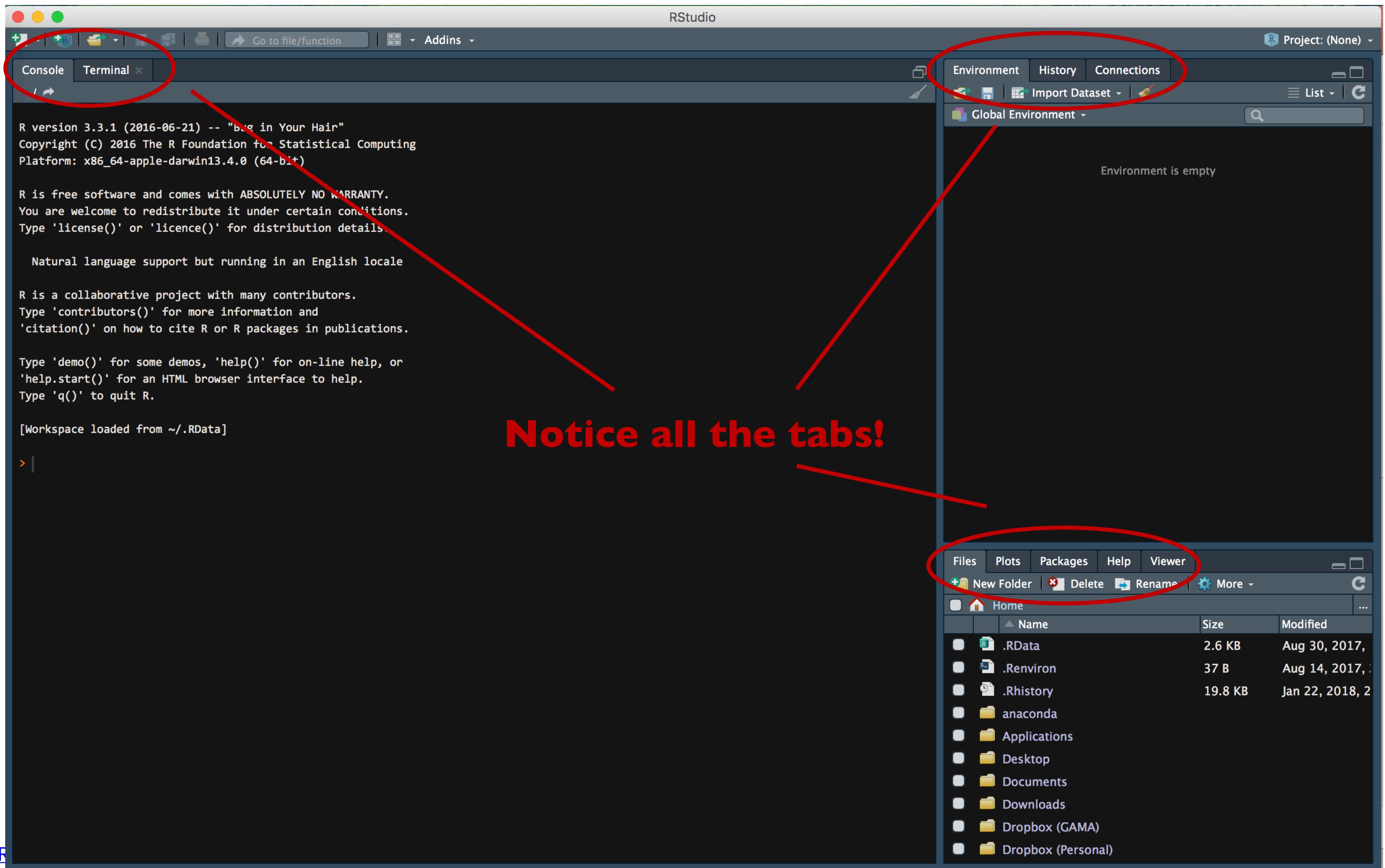
Reptilian
Brain

RStudio



The screenshot shows the RStudio interface with the following components:

- Console Pane:** Displays the R startup message and workspace loading information.
- Global Environment Pane:** Shows the "Global Environment" tab with a message stating "Environment is empty". A large red text overlay reads:
**All the data, variables
and functions you
created will show up
here**
- Files Pane:** Shows the file structure under "Home". A large red text overlay reads:
This is where all the coding happens
- Bottom Status Bar:** Shows "CC BY-SA R"



RStudio

Project: (None)

Console Terminal

R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (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]

> | From here you can access:

- **The console, your interface to R**
- **The terminal, your interface to bash**
- **A script editor will show up here when you create a new script; your platform to document your reproducible analysis**

Environment History Connections

Import Dataset

Global Environment

Environment is empty

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

Name	Size	Modified
.RData	2.6 KB	Aug 30, 2017,
.Renviron	37 B	Aug 14, 2017, 1
.Rhistory	19.8 KB	Jan 22, 2018, 2
anaconda		
Applications		
Desktop		
Documents		
Downloads		
Dropbox (GAMA)		
Dropbox (Personal)		

CC BY-SA

RStudio

Console Terminal

~/

```
R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (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]
```

> | **From here you can access:**

- **The environment where all your data for this session is stored**
- **The history of your commands**
- **Connections to databases (advanced)**

Project: (None)

Environment History Connections

Import Dataset

Global Environment

Environment is empty

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

	Name	Size	Modified
..	.RData	2.6 KB	Aug 30, 2017,
..	.Renviron	37 B	Aug 14, 2017, 1
..	.Rhistory	19.8 KB	Jan 22, 2018, 2
..	anaconda		
..	Applications		
..	Desktop		
..	Documents		
..	Downloads		
..	Dropbox (GAMA)		
..	Dropbox (Personal)		

RStudio

Console Terminal Addins Environment History Connections Project: (None)

From here you can access:

- The file browser
- Plots will be shown here once you create some
- The packages tab allows you to explore the installed packages, download new ones or update existing ones
- Help allows you to search the documentation. Very useful!!
- The viewer can be used to view local web content, e.g. Googlevis charts or animations created in R

R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32-win3.1.0 (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
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for online 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

New Folder Delete Rename More

Name Size Modified

Name	Size	Modified
.RData	2.6 KB	Aug 30, 2017,
.Renviron	37 B	Aug 14, 2017,
.Rhistory	19.8 KB	Jan 22, 2018, 2
anaconda		
Applications		
Desktop		
Documents		
Downloads		
Dropbox (GAMA)		
Dropbox (Personal)		

CC BY-SA F

RStudio

Project: (None)

Console Terminal

R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (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]

>

From here you can:

- Create a new file, e.g. a script
- Create a new R project
- Open an existing file

Environment History Connections

Import Dataset

Global Environment

Environment is empty

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

	Name	Size	Modified
	.RData	2.6 KB	Aug 30, 2017,
	.Renvironment	37 B	Aug 14, 2017, :
	.Rhistory	19.8 KB	Jan 22, 2018, 2
	anaconda		
	Applications		
	Desktop		
	Documents		
	Downloads		
	Dropbox (GAMA)		
	Dropbox (Personal)		

CC BY-SA

RStudio Addins Project: (None)

Untitled1 x Go to file/function Addins Run Source

Source on Save Import Dataset Global Environment

Environment History Connections List C

Global Environment

Environment is empty

1:1 (Top Level) R Script

Console Terminal

~/

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]

The console and terminal have now been moved to your bottom half of the screen.

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

	Name	Size	Modified
	.RData	2.6 KB	Aug 30, 2017,
	.Renviron	37 B	Aug 14, 2017, 1
	.Rhistory	19.8 KB	Jan 22, 2018, 2
	anaconda		
	Applications		
	Desktop		
	Documents		
	Downloads		
	Dropbox (GAMA)		
	Dropbox (Personal)		

CC BY-SA R >

RStudio

Addins

Project: (None)

Untitled1

Source on Save | Run | Source | Environment History Connections

Import Dataset | Global Environment

Environment is empty

1:1 (Top Level) R Script

Console Terminal ~/

R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (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]

>

All window sizes can be adjusted
by dragging the edges up/down
or left/right

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

Name	Size	Modified
.RData	2.6 KB	Aug 30, 2017, 10:17 AM
.Renviron	37 B	Aug 14, 2017, 2:51 PM
.Rhistory	19.8 KB	Jan 22, 2018, 2:43 PM
anaconda		
Applications		
Desktop		
Documents		
Downloads		
Dropbox (GAMA)		
Dropbox (Personal)		

CC BY-SA RStu

R Untitled1 x

Source on Save | Run

1 |

1:1 (Top Level) R Script

Console Terminal x

~/

```
R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (64-bit)

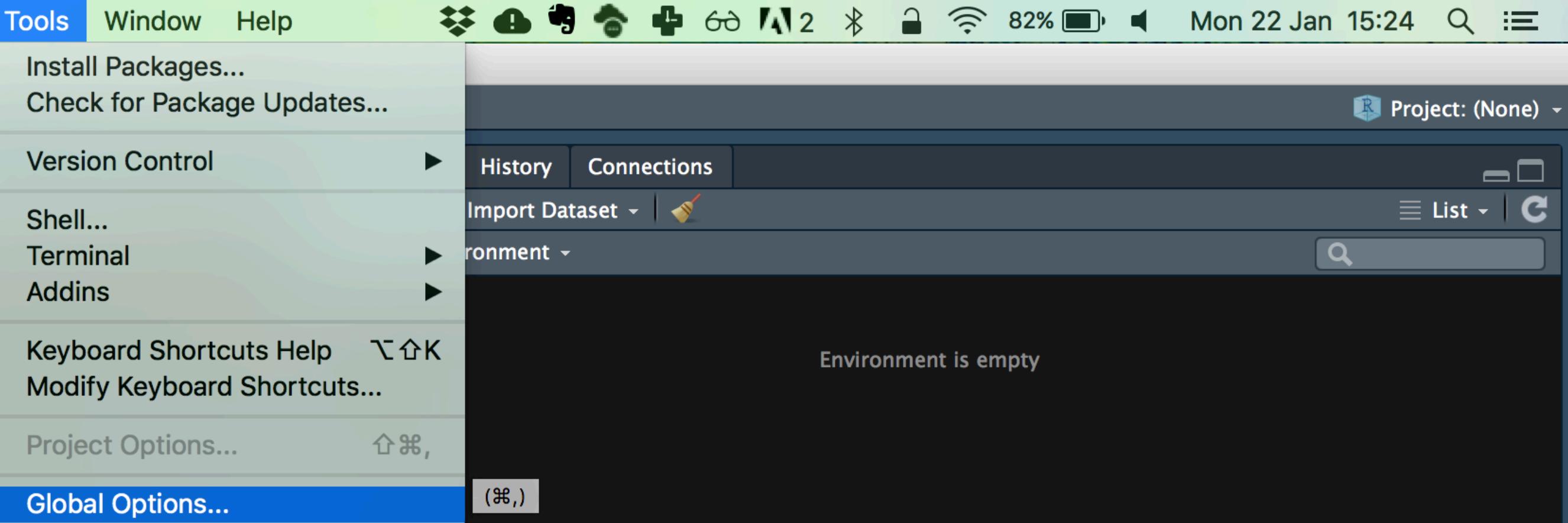
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

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]
```



You can change the layout and look of your RStudio and many other settings by going to

Tools > Global Options

Files	Plots	Packages	Help	Viewer	...
+ New Folder	- Delete	Rename	More	C	...
Home					
Name		Size	Modified		
.RData	2.6 KB	Aug 30, 2017, 10:17 AM			
.Renviron	37 B	Aug 14, 2017, 2:51 PM			
.Rhistory	19.8 KB	Jan 22, 2018, 2:43 PM			
anaconda					
Applications					
Desktop					
Documents					
Downloads					
Dropbox (GAMA)					
Dropbox (Personal)					

The appearance theme here for example is Merbivore

R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (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]

>

Untitled1 x

Source on Save | Run | Source | Environment | History | Connections | Global Environment | Import Dataset | List | C

Environment is empty

More ▾

Size Modified

2.6 KB Aug 30, 2017, 10:17 AM

37 B Aug 14, 2017, 2:51 PM

19.8 KB Jan 22, 2018, 2:43 PM

OK Cancel Apply

Applications

Desktop Documents Downloads Dropbox (GAMA) Dropbox (Personal)

Options

R Studio theme: Modern

Editor font: Consolas

Editor Font size: 10

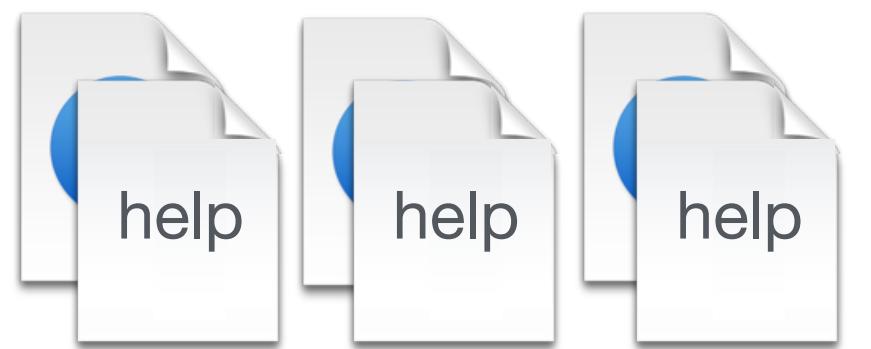
Editor theme: Merbivore (selected)

Crimson Editor Dawn Dracula Dreamweaver Eclipse Idle Fingers Katzenmilch Kr Theme Material Merbivore Soft Merbivore Mono Industrial Monokai Pastel On Dark Solarized Dark Solarized Light

R Packages

A large, semi-transparent watermark of the R logo is positioned in the bottom right corner. The logo consists of a dark green circle containing a white letter 'R' with a diagonal line through it.

R

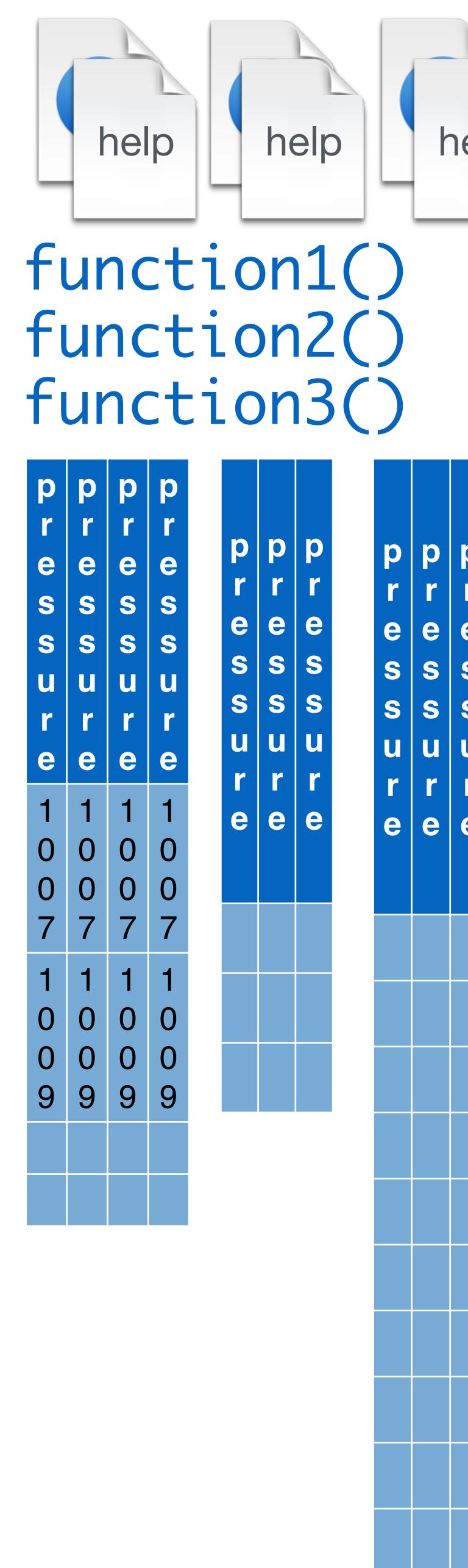
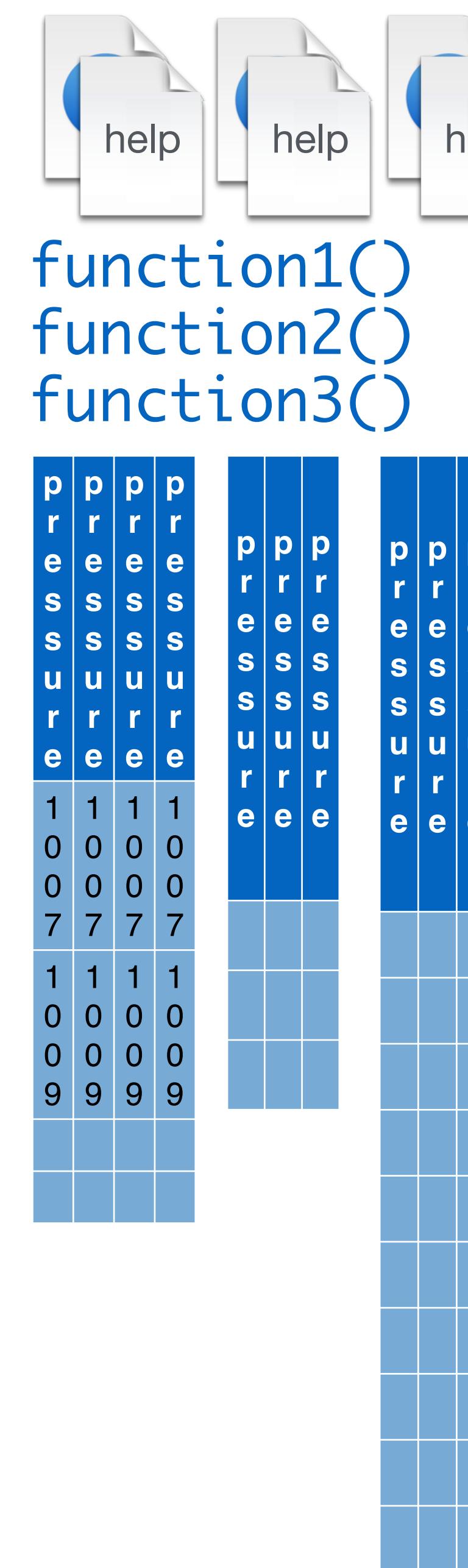
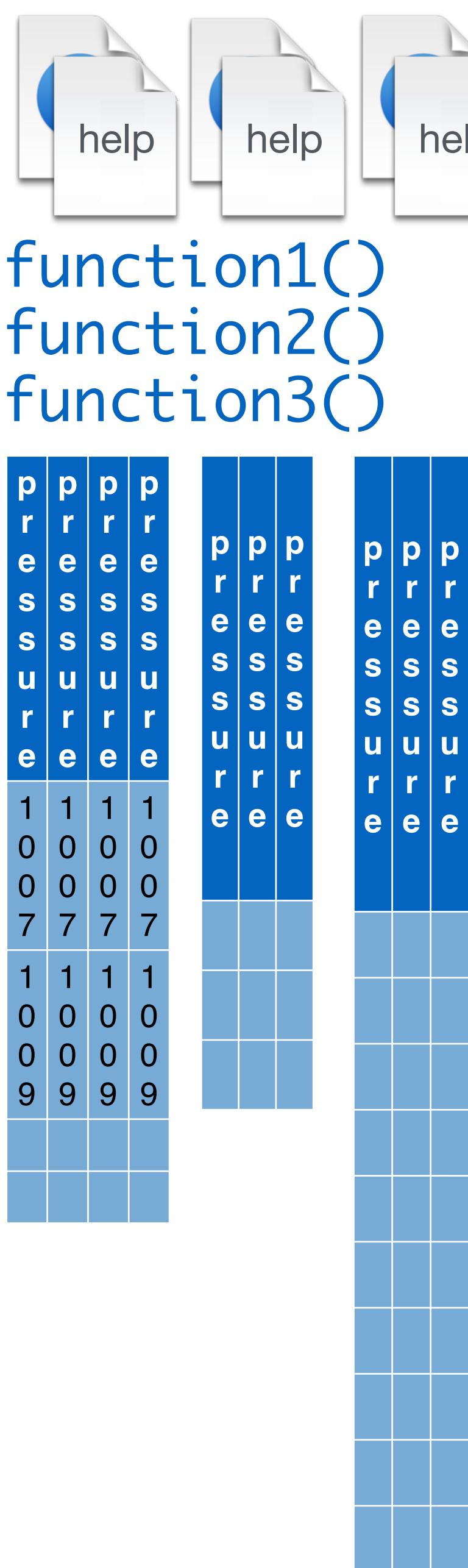
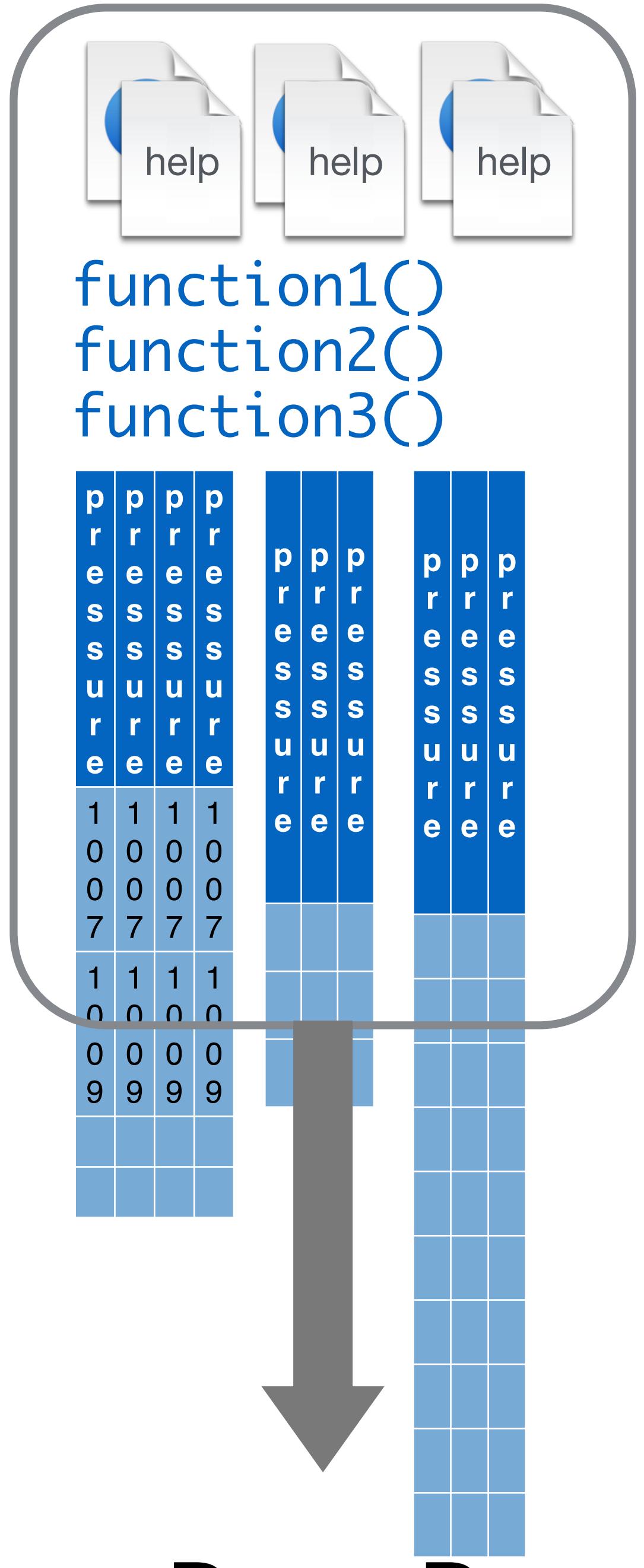


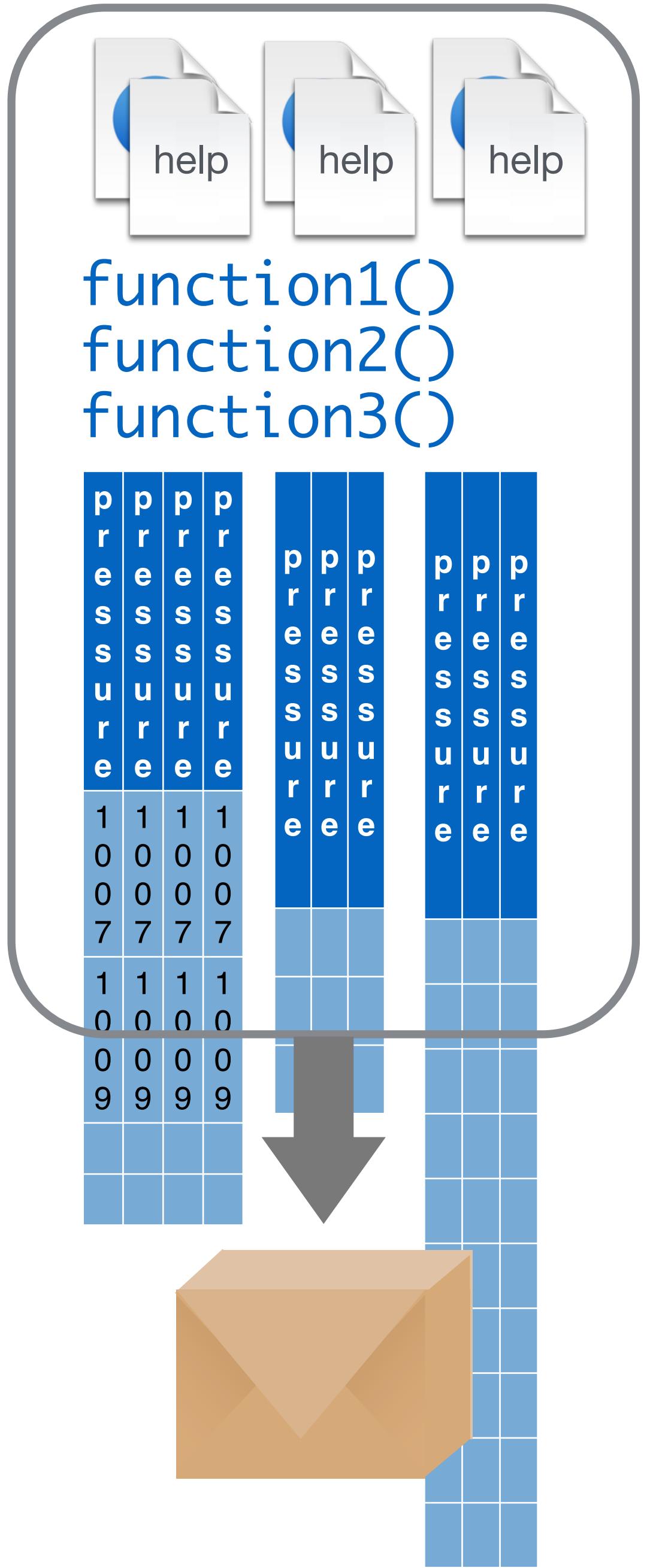
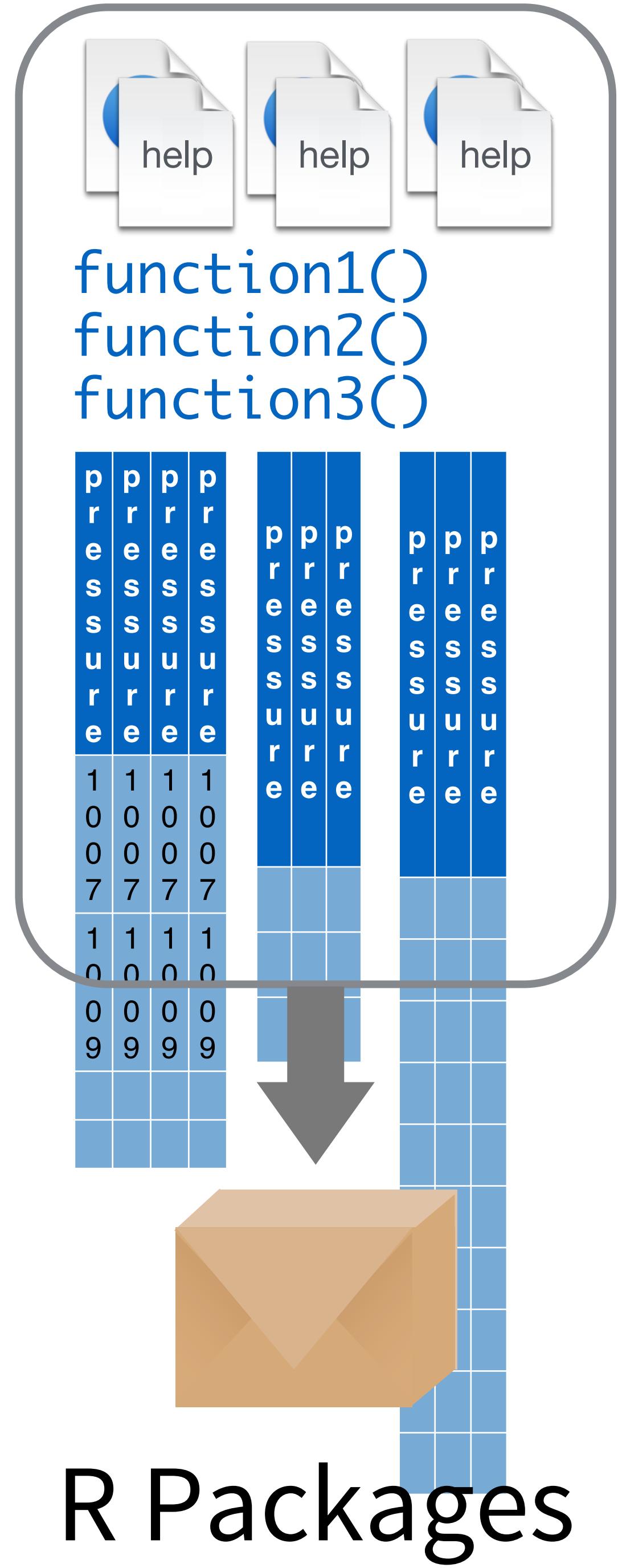
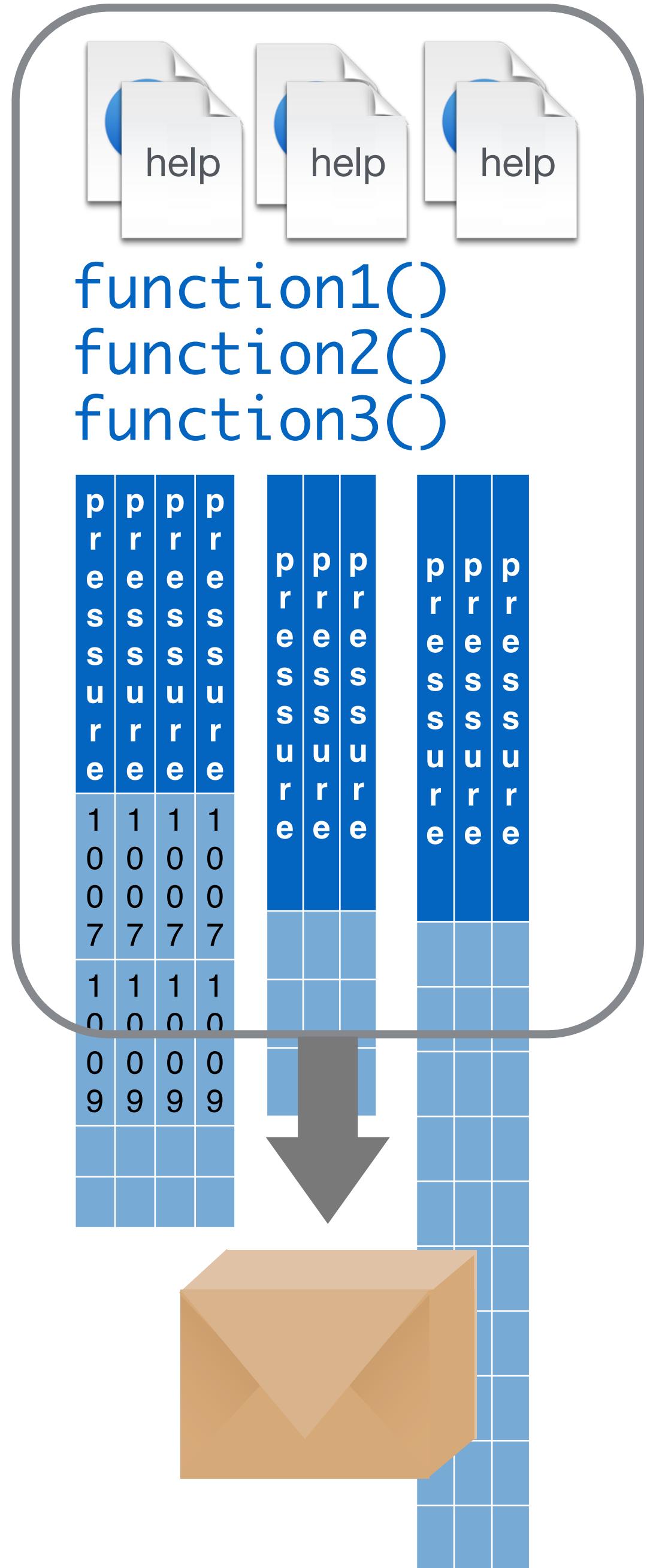
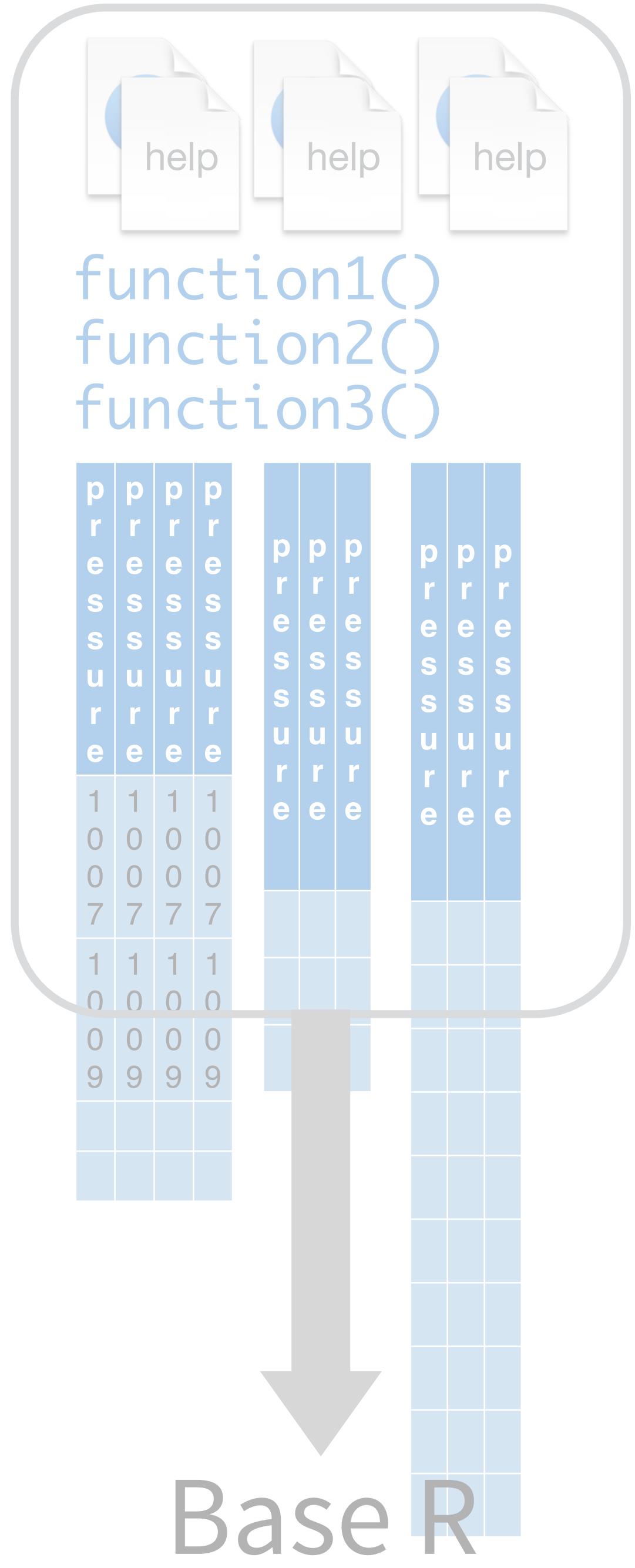
function1()
function2()
function3()

p	p	p	p
r	r	r	r
e	e	e	e
s	s	s	s
s	s	s	s
u	u	u	u
r	r	r	r
e	e	e	e
1	1	1	1
0	0	0	0
0	0	0	0
7	7	7	7
1	1	1	1
0	0	0	0
0	0	0	0
9	9	9	9

p	p	p	p
r	r	r	r
e	e	e	e
s	s	s	s
s	s	s	s
u	u	u	u
r	r	r	r
e	e	e	e
1	1	1	1
0	0	0	0
0	0	0	0
7	7	7	7
1	1	1	1
0	0	0	0
0	0	0	0
9	9	9	9

p	p	p	p
r	r	r	r
e	e	e	e
s	s	s	s
s	s	s	s
u	u	u	u
r	r	r	r
e	e	e	e
1	1	1	1
0	0	0	0
0	0	0	0
7	7	7	7
1	1	1	1
0	0	0	0
0	0	0	0
9	9	9	9





The screenshot shows a web browser window with the title "The Comprehensive R Archive". The address bar indicates a secure connection to "https://cran.r-project.org". The page content includes the CRAN logo, navigation links for mirrors, what's new, task views, search, about R, software, documentation, and contributed packages. A large section on the right lists available CRAN packages by name, with a link to the full list of packages from A to Z.



[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](#)

CC BY-SA

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

[abyyR](#) 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

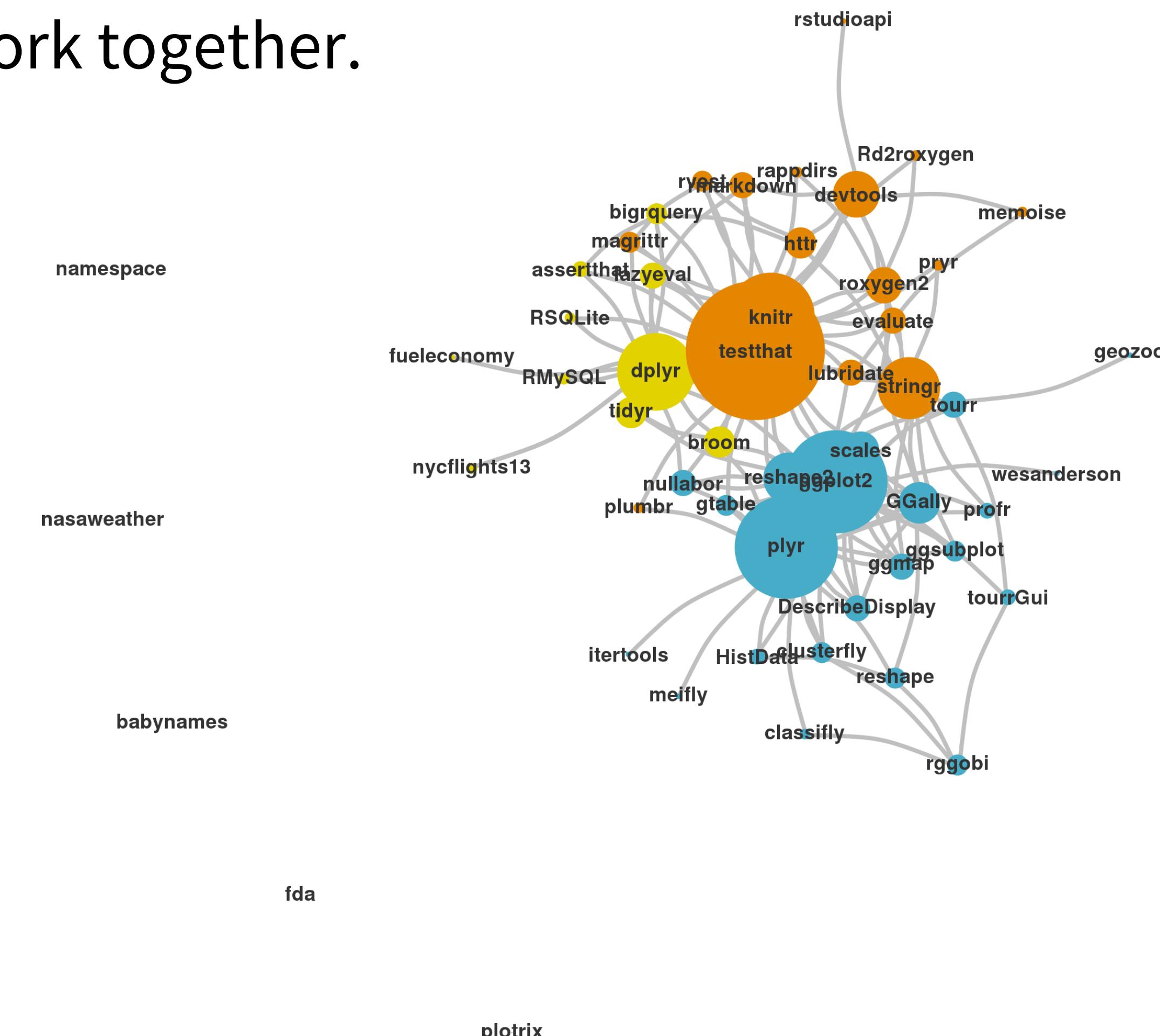
1 x per R Session

The Tidyverse



The Tidyverse

A collection of modern R packages that share common philosophies, embed best practices, and are designed to work together.



tidyverse.org

The screenshot shows the tidyverse.org homepage. At the top, there's a navigation bar with links for Packages, Articles, Learn, Help, and Contribute. Below the navigation, there's a large graphic featuring hexagonal icons for various R packages: dplyr (orange, with pliers icon), ggplot2 (grey, with line plot icon), readr (blue, with document icon), purrr (grey, with cat icon), tibble (dark blue, with grid icon), and tidyr (orange, with arrows icon). To the right of the graphic, there's descriptive text about the tidyverse and a code snippet for installing it.

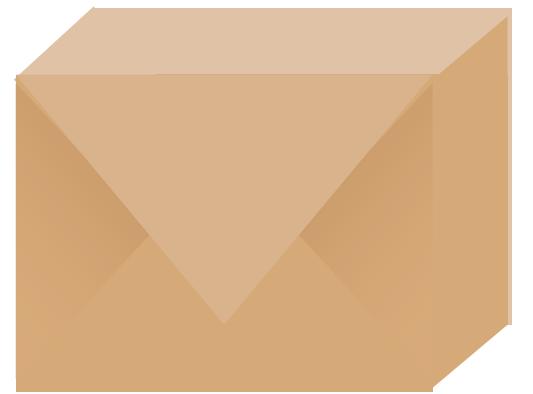
R packages for data science

The tidyverse is an opinionated **collection of R packages** designed for data science. All packages share an underlying philosophy and common APIs.

Install the complete tidyverse with:

```
install.packages("tidyverse")
```

tidyverse



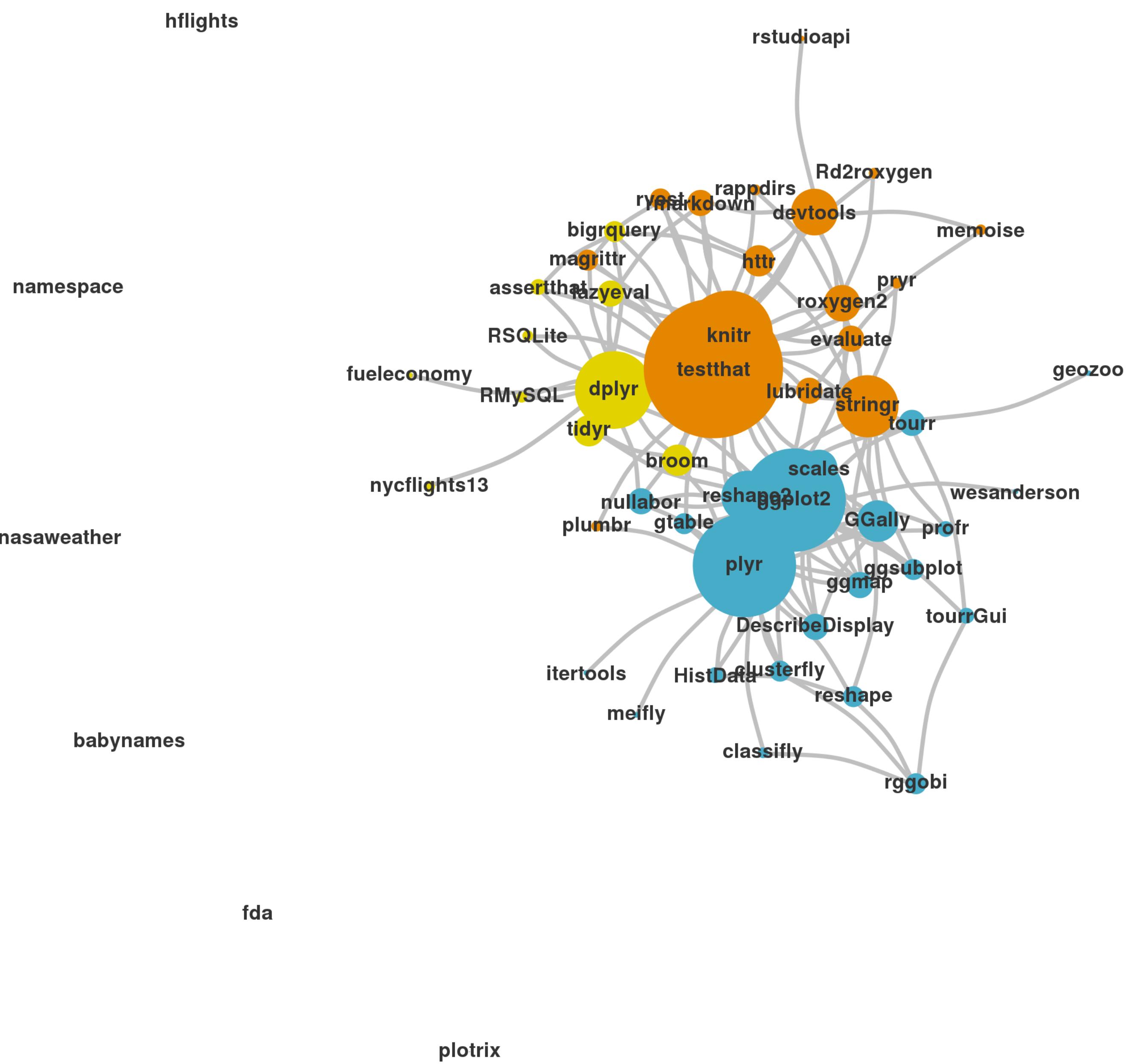
An R package that serves as a short cut for installing and loading the components of the tidyverse.

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

```
install.packages("tidyverse")
```

does the equivalent of

```
install.packages("ggplot2")
install.packages("dplyr")
install.packages("tidyr")
install.packages("readr")
install.packages("purrr")
install.packages("tibble")
install.packages("hms")
install.packages("stringr")
install.packages("lubridate")
install.packages("forcats")
install.packages("DBI")
install.packages("haven")
install.packages("httr")
install.packages("jsonlite")
install.packages("readxl")
install.packages("rvest")
install.packages("xml2")
install.packages("modelr")
install.packages("broom")
```



```
install.packages("tidyverse")
```

does the equivalent of

```
install.packages("ggplot2")
install.packages("dplyr")
install.packages("tidyr")
install.packages("readr")
install.packages("purrr")
install.packages("tibble")
install.packages("hms")
install.packages("stringr")
install.packages("lubridate")
install.packages("forcats")
install.packages("DBI")
install.packages("haven")
install.packages("httr")
install.packages("jsonlite")
install.packages("readxl")
install.packages("rvest")
install.packages("xml2")
install.packages("modelr")
install.packages("broom")
```

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

does the equivalent of

```
library("ggplot2")
library("dplyr")
library("tidyr")
library("readr")
library("purrr")
library("tibble")
```

R Notebooks

(Let's start!)



Your Turn

Open **00-Introduction.Rmd**. Read through the notebook and do everything it tells you to do.



R Notebooks

An authoring format for Data Science.

The screenshot shows the RStudio interface with an R Notebook open. The notebook file is titled "R-Notebook.Rmd". The code editor pane contains the following R Markdown code:

```
1 ---  
2 title: "R Notebook"  
3 output: html_notebook  
4 ---  
5  
6 Text written in **markdown**  
7  
8 ```{r}  
9 # code written in R  
10 (x <- rnorm(7))  
11 ````  
12  
13 Text written in _markdown_  
14  
15 ```{r}  
16 # code written in R  
17 hist(x)  
18 ````  
19  
20 (Top Level) ◊
```

The R console pane below shows the output of the R code:

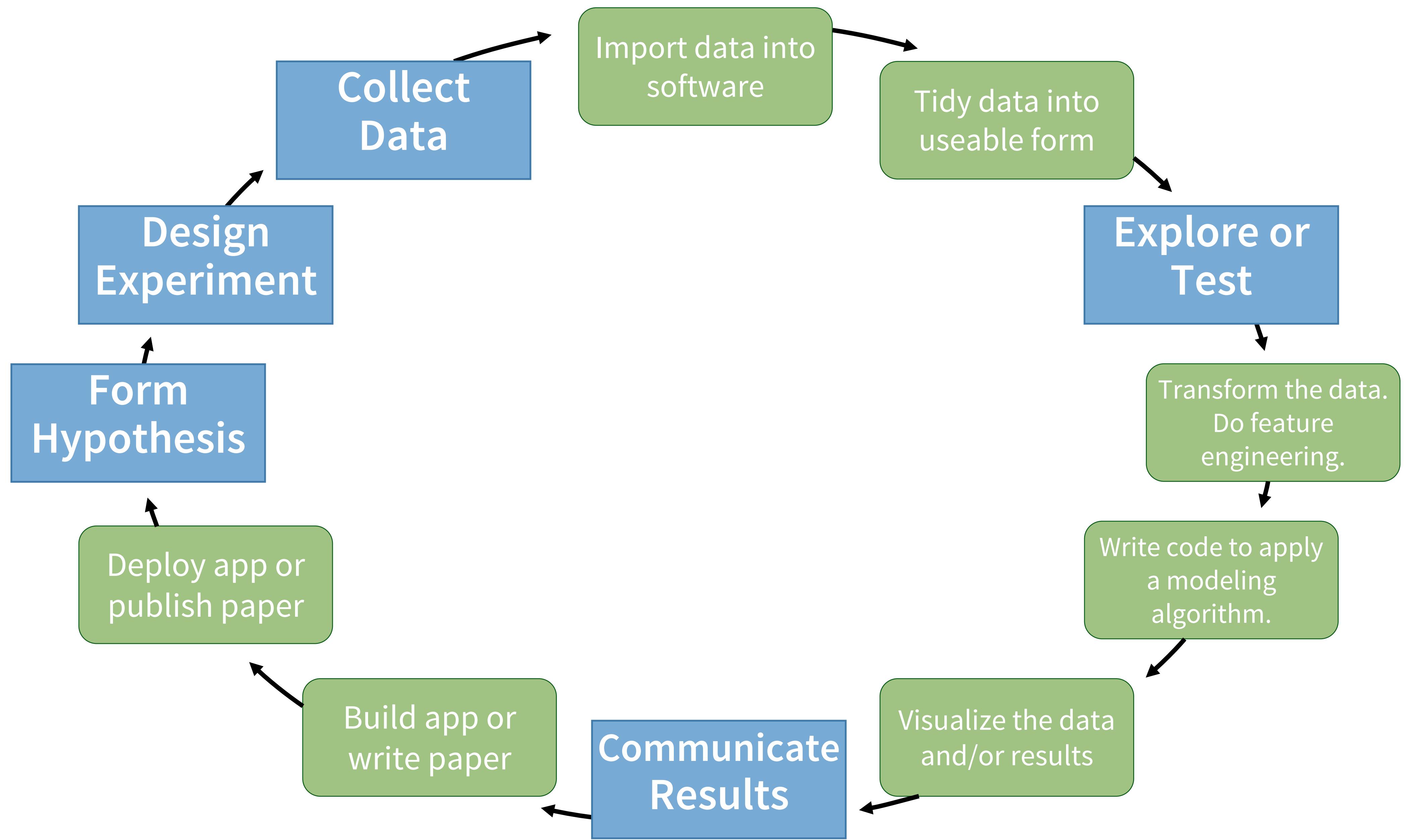
```
[1] -1.2  1.0 -0.5  0.9 -0.6 -1.1 -1.5
```

Three callout boxes highlight specific features:

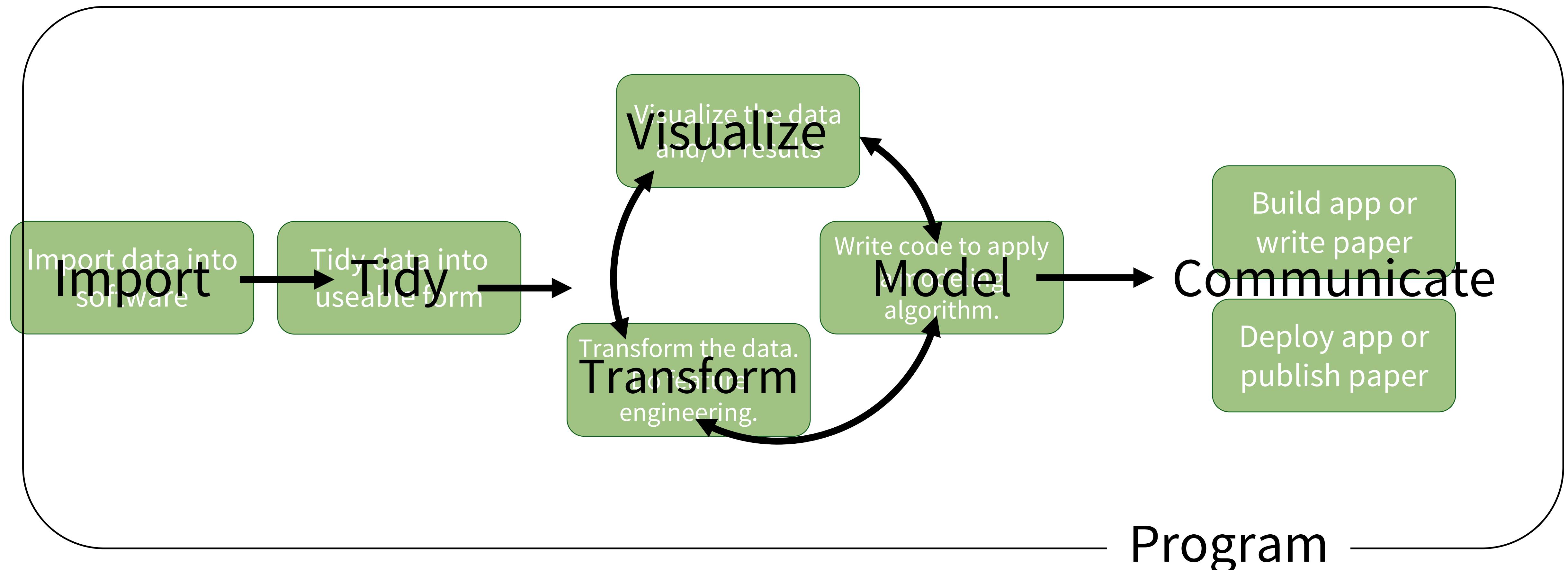
- A light gray callout box points to the green "Run All" button in the toolbar, containing the text: "Click to run all code chunks above".
- A medium gray callout box points to the green "Run Current Chunk" button in the toolbar, containing the text: "Click to run code in chunk".
- A dark gray callout box points to the R console output, containing the text: "Code result".

Outro





(Applied) Data Science



(Applied) Data Science

