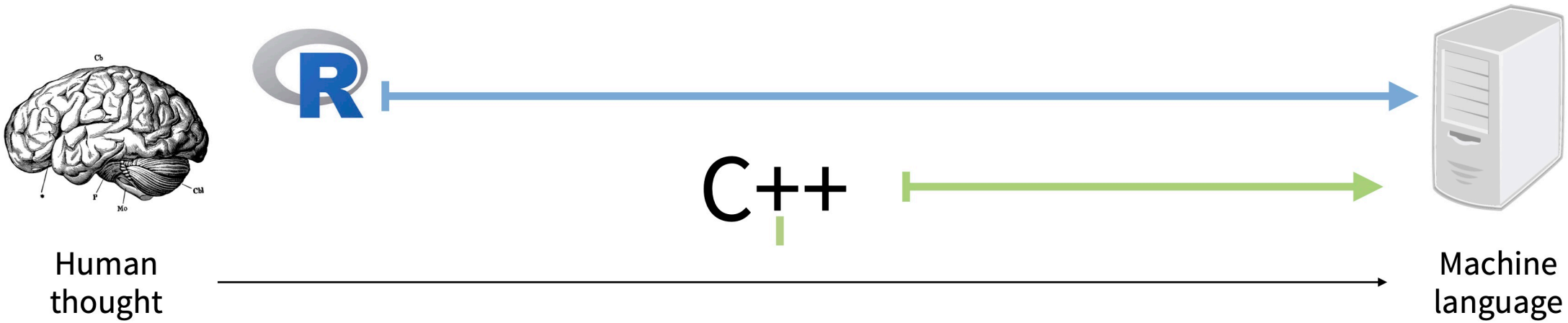
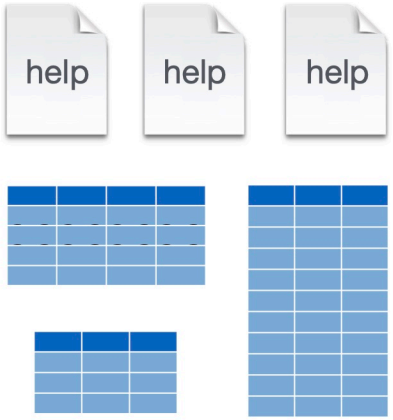
R, RStudio and RMarkdown

presented by: Jackson Hoffart

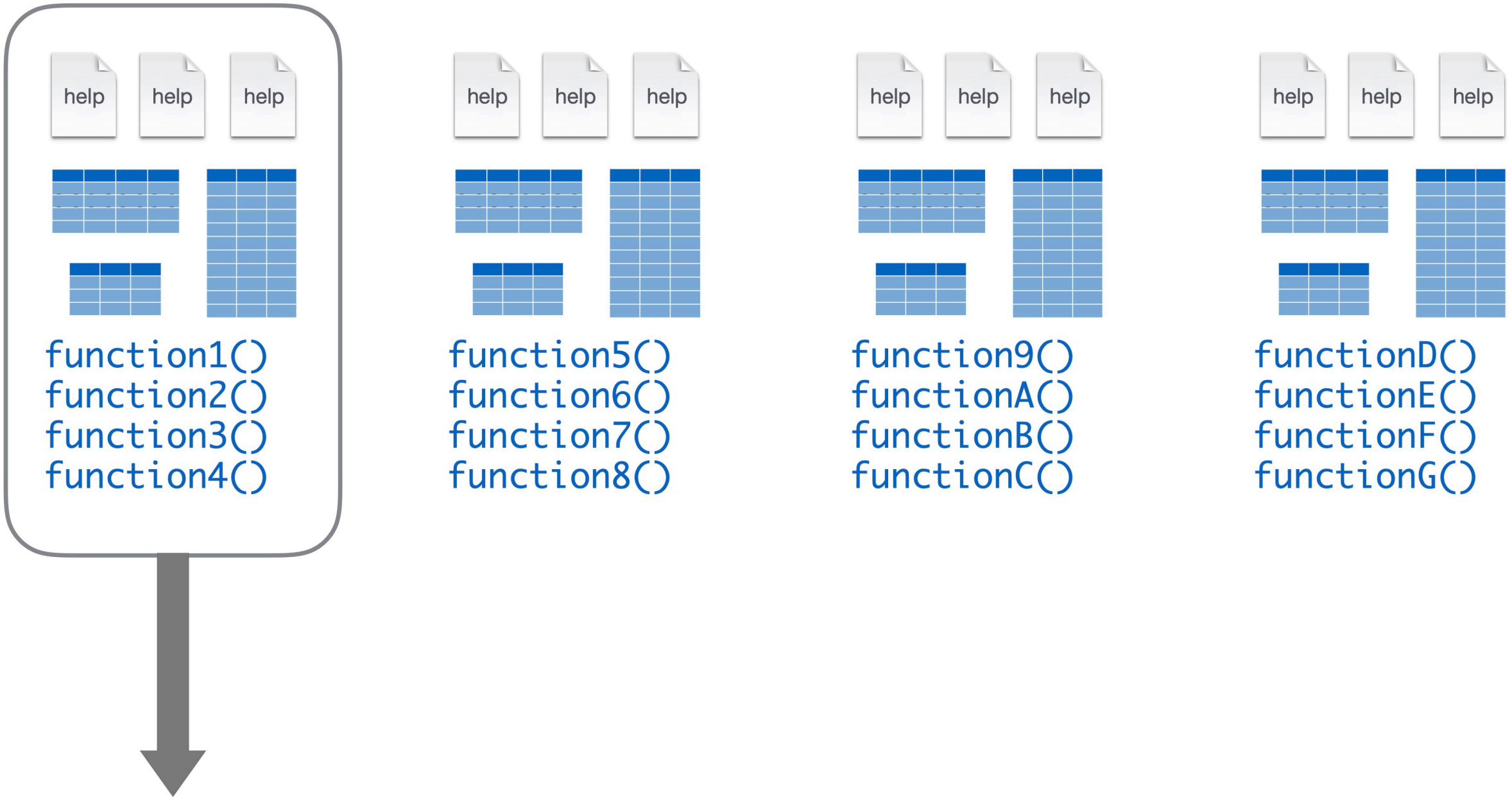
slides adapted from: https://github.com/rstudio-conf-2020/data-science-tidy

R - A computer language for scientists

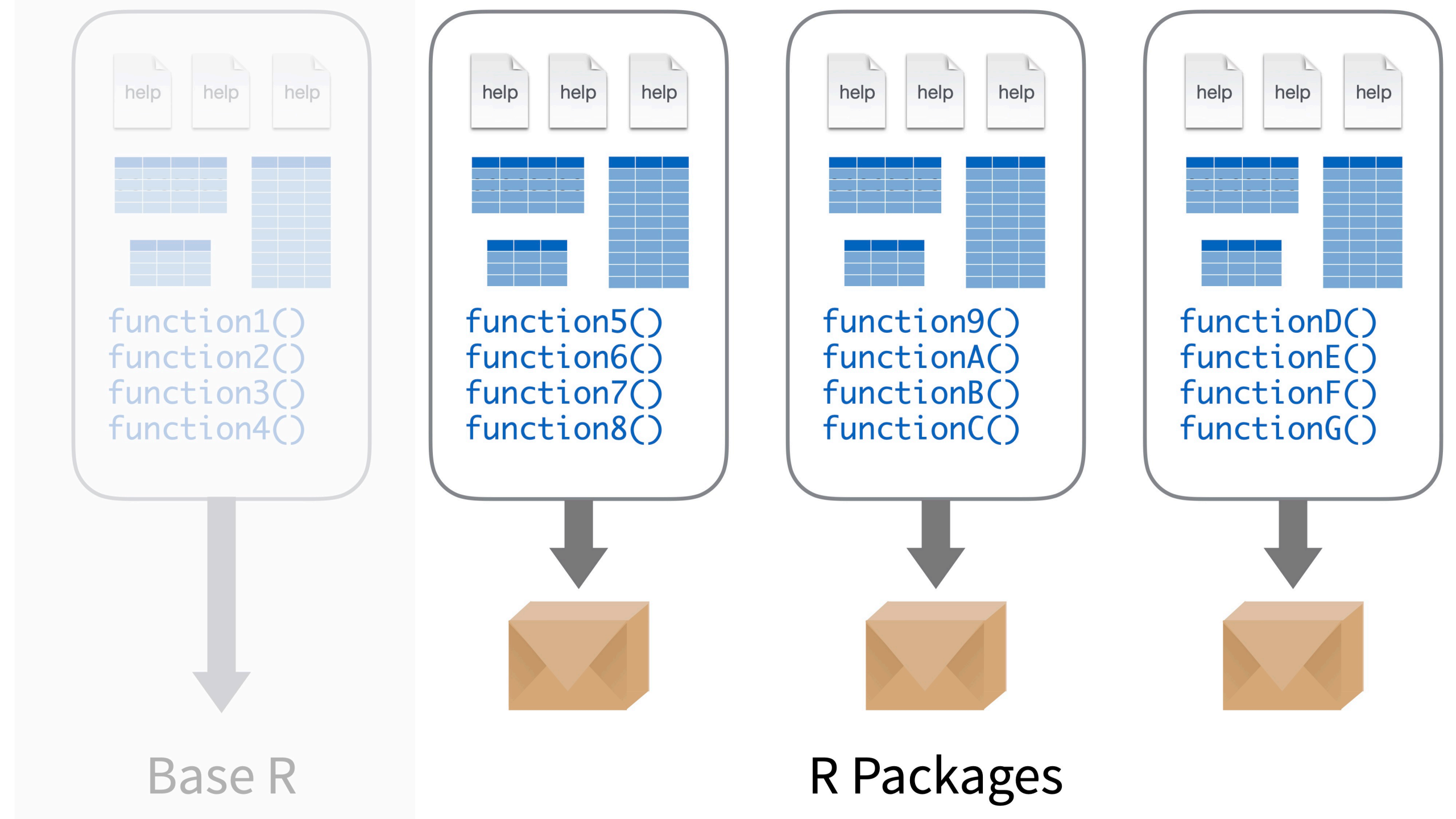


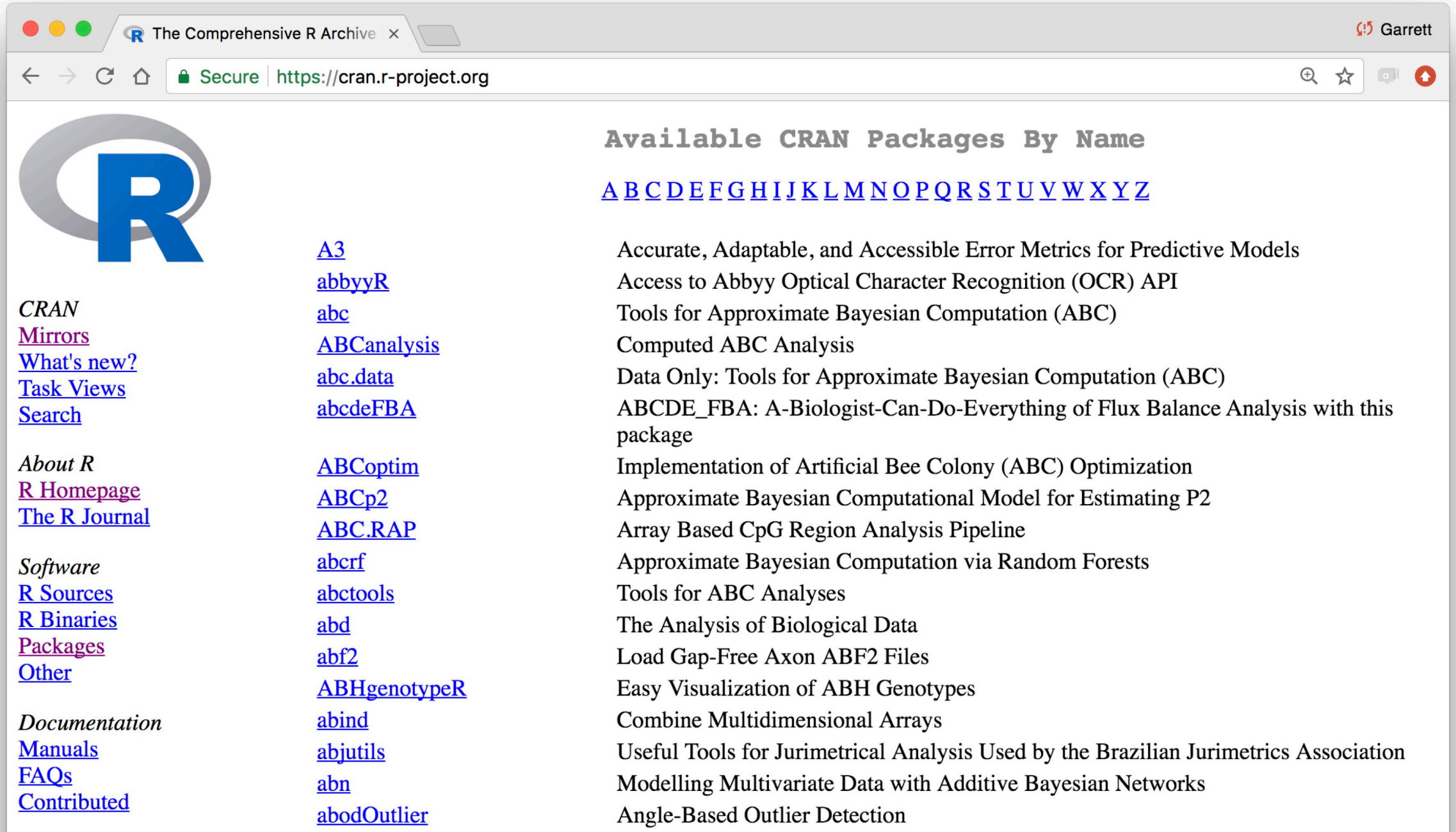



```
function1()
function2()
function3()
function4()
```


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



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

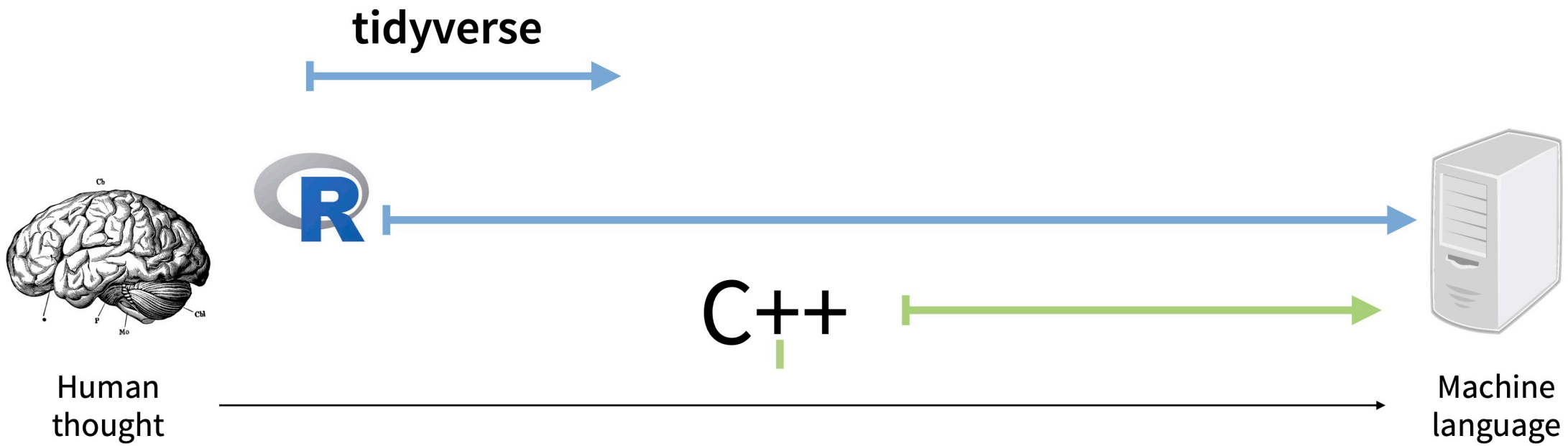
```
library("foo")
```

Loads package

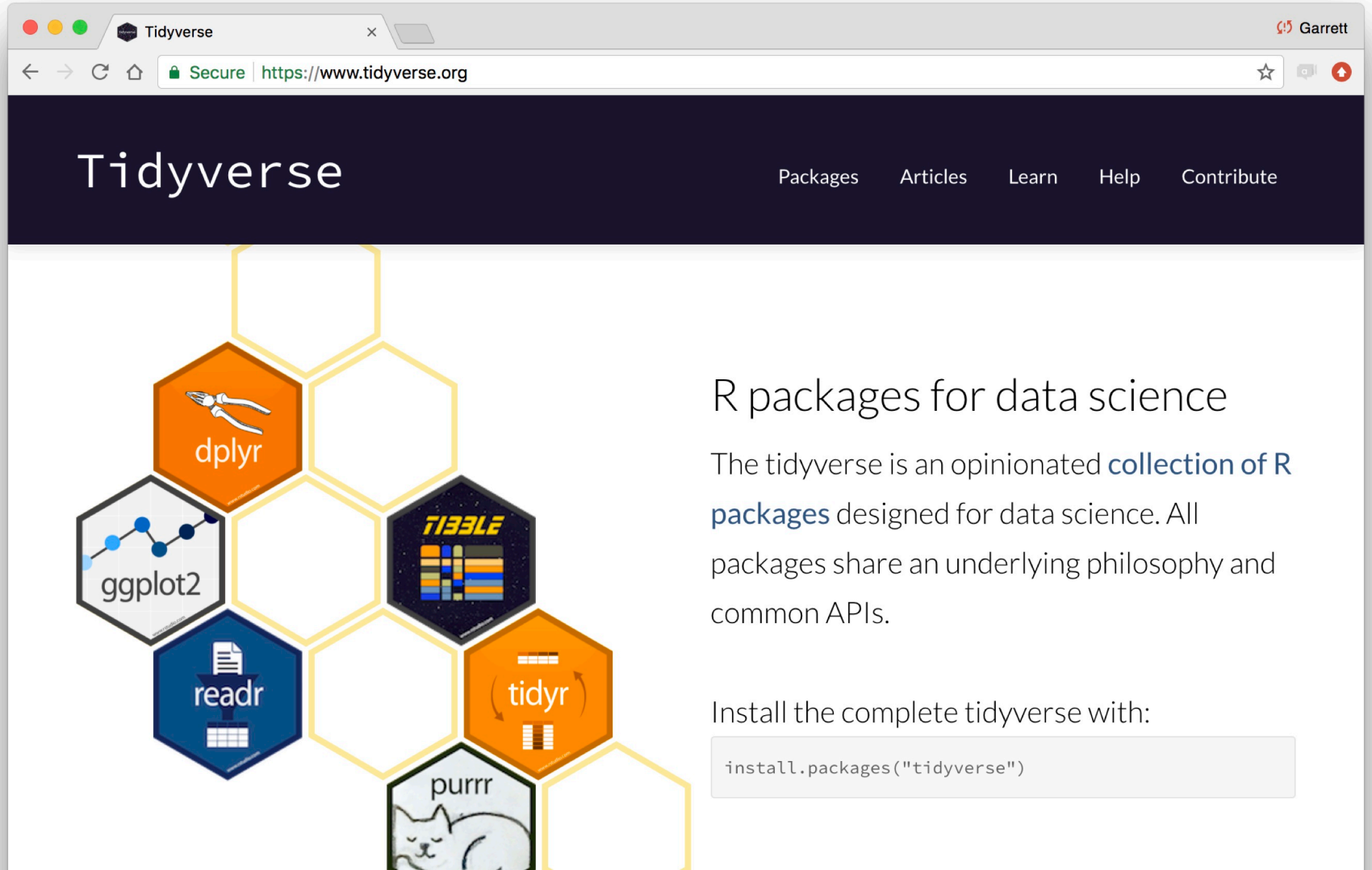
1 x per R Session

I've done this for you for this workshop

The tidyverse - A set of R packages to unify some data science tasks

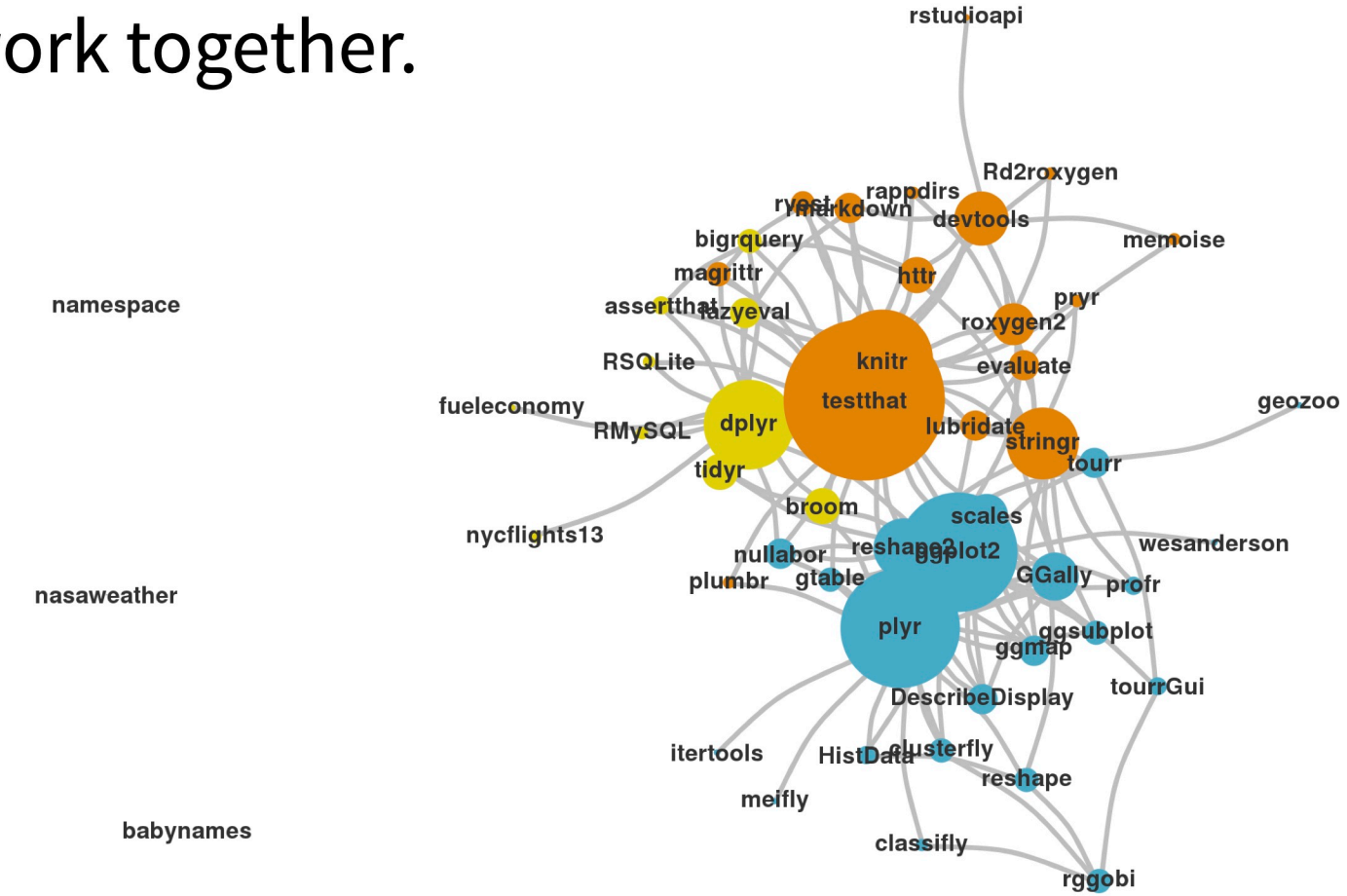


tidyverse.org



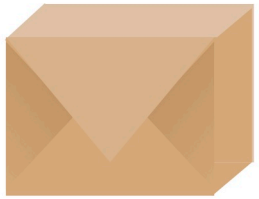
The Tidyverse

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



Display by Adolfo Álvarez

tidyverse



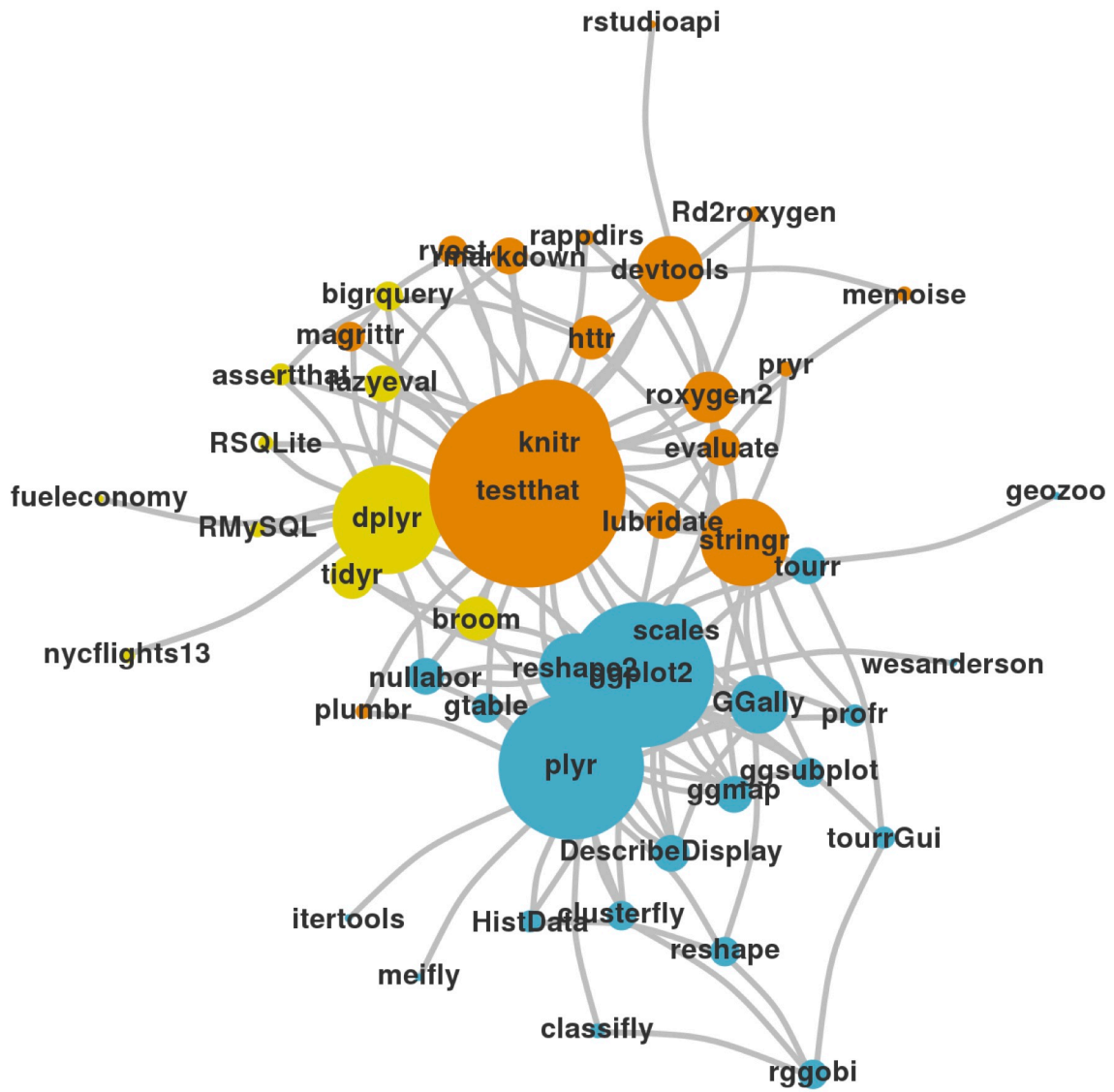
An R package that serves as a short cut for installing and loading the components of the tidyverse.

```
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

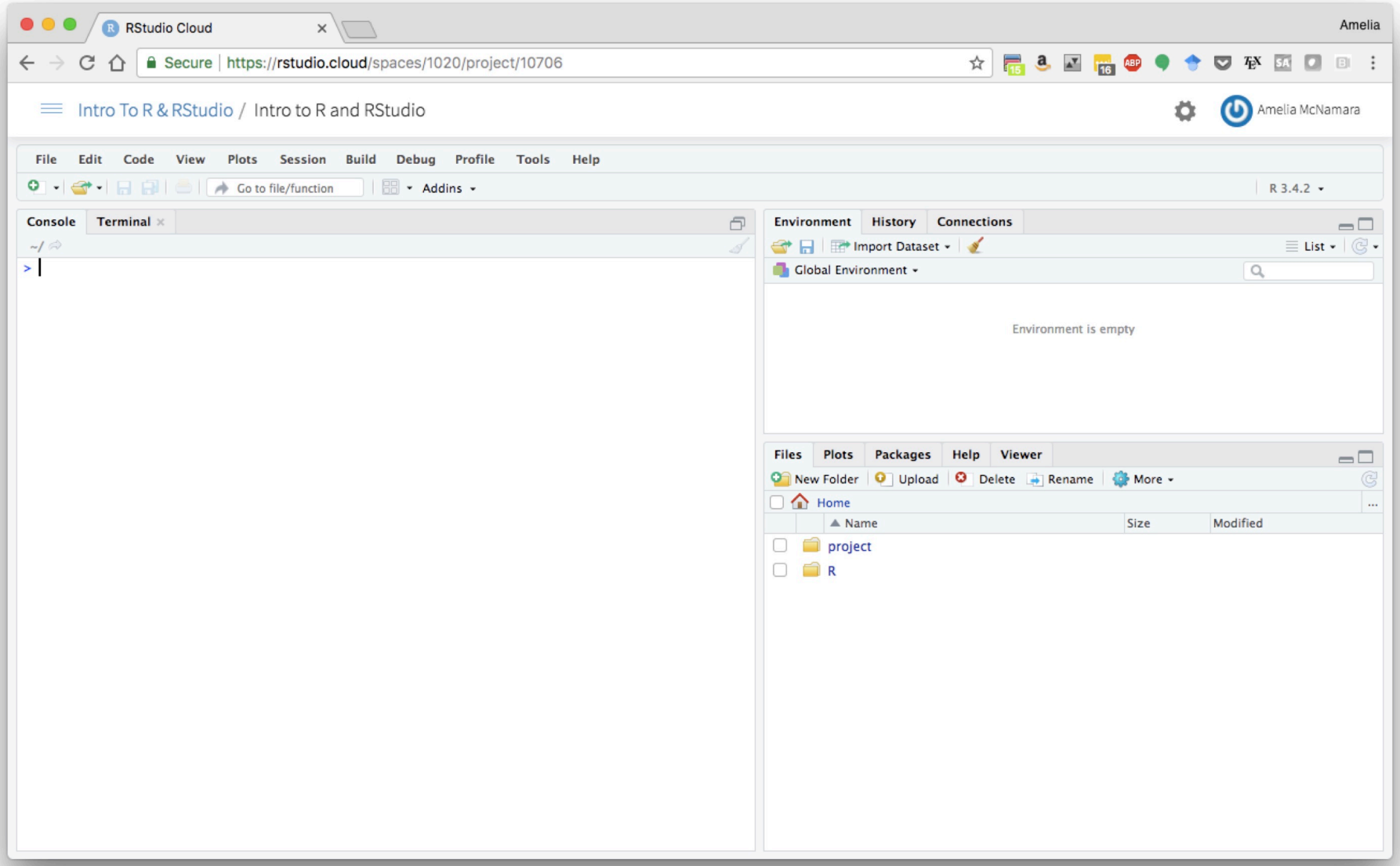
```
library("ggplot2")
library("dplyr")
library("tidyr")
library("readr")
library("purrr")
library("tibble")
```


R

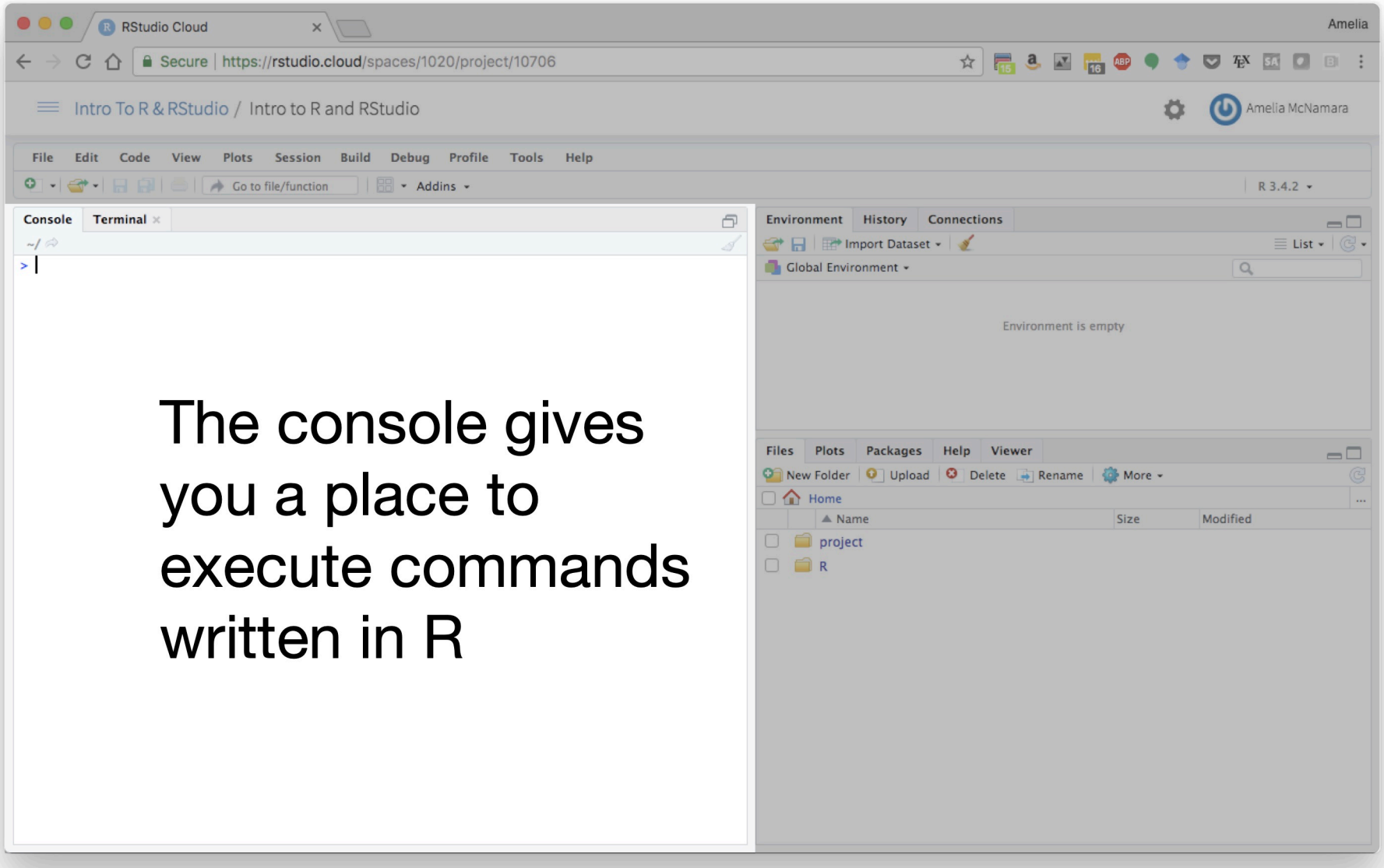
RStudio: a software program

1. like Microsoft Word, Excel, etc.
2. built to help you write R code, run R code, and analyze data with R
3. text editor, version control, keyboard shortcuts, debugging tools, and much more

RStudio

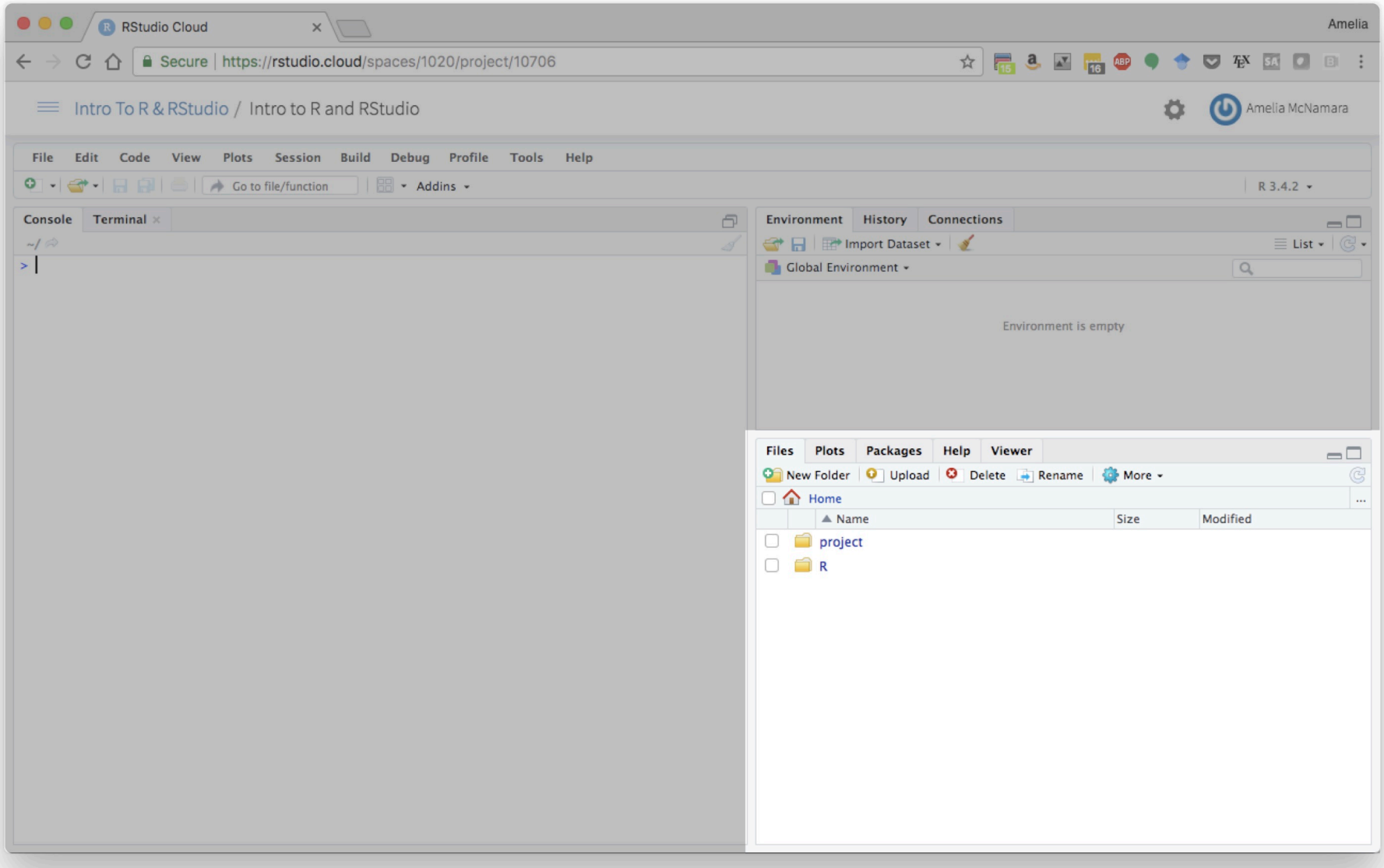

RStudio Cloud
Amelia
Secure | https://rstudio.cloud/spaces/1020/project/10706
Intro To R & RStudio / Intro to R and RStudio
Amelia McNamara
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 3.4.2
Console
Terminal
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Files
Plots
Packages
Help
Viewer
New Folder
Upload
Delete
Rename
More
Home
Name
Size
Modified
project
R

RStudio

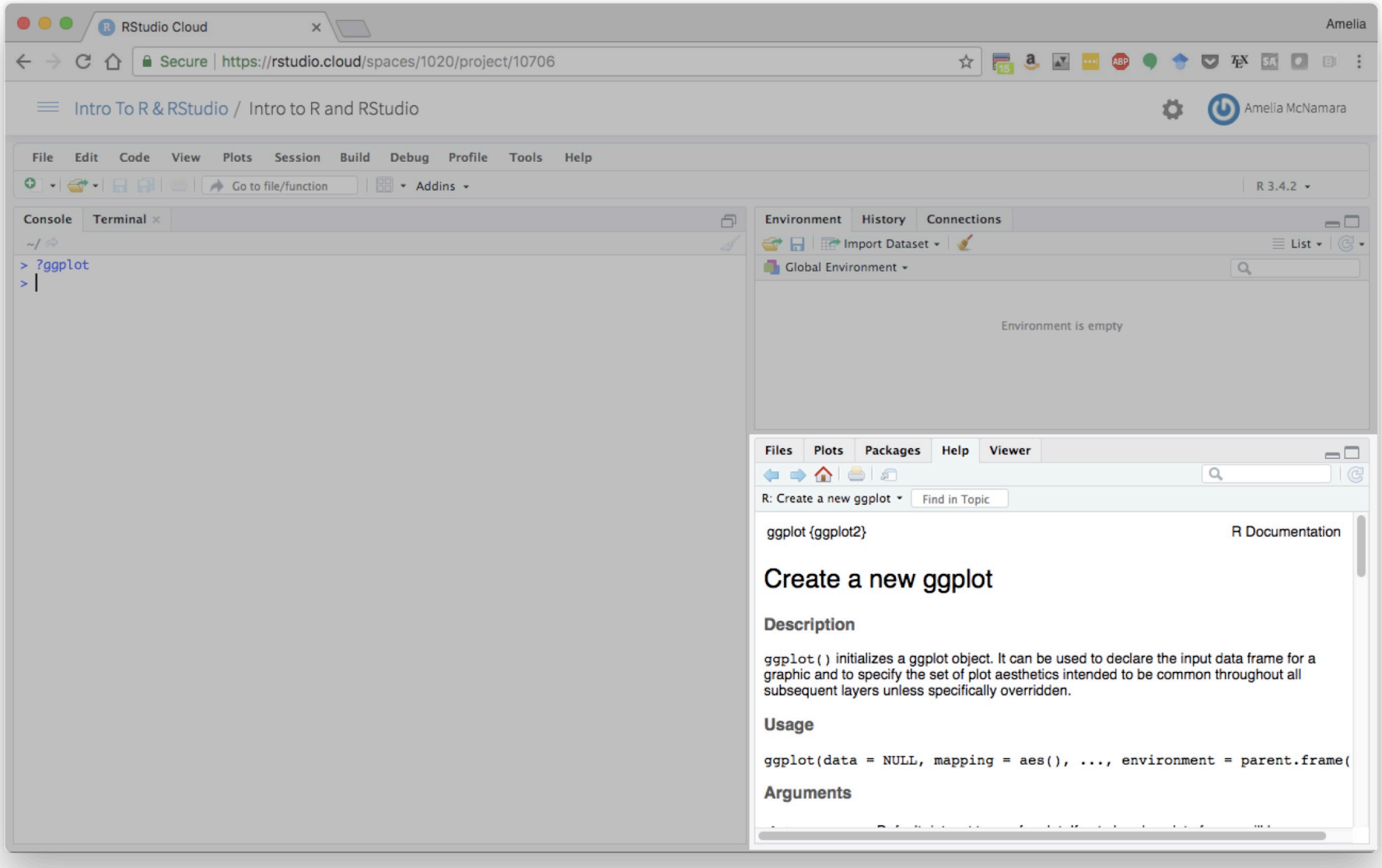


The console gives you a place to execute commands written in R

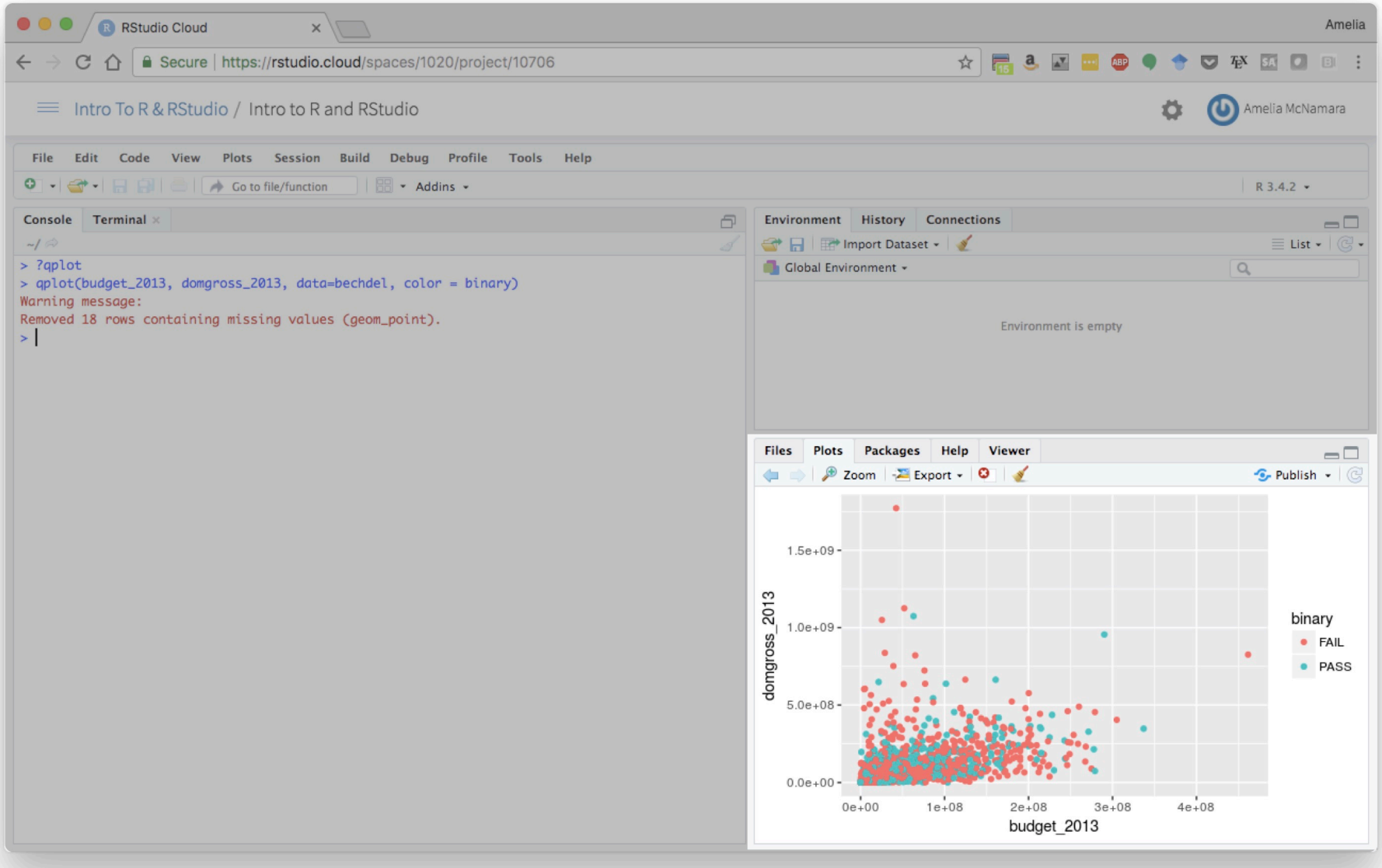
RStudio


RStudio Cloud
Amelia
Secure | https://rstudio.cloud/spaces/1020/project/10706
Intro To R & RStudio / Intro to R and RStudio
Amelia McNamara
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 3.4.2
Console
Terminal
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Files
Plots
Packages
Help
Viewer
New Folder
Upload
Delete
Rename
More
Home
Name
Size
Modified
project
R

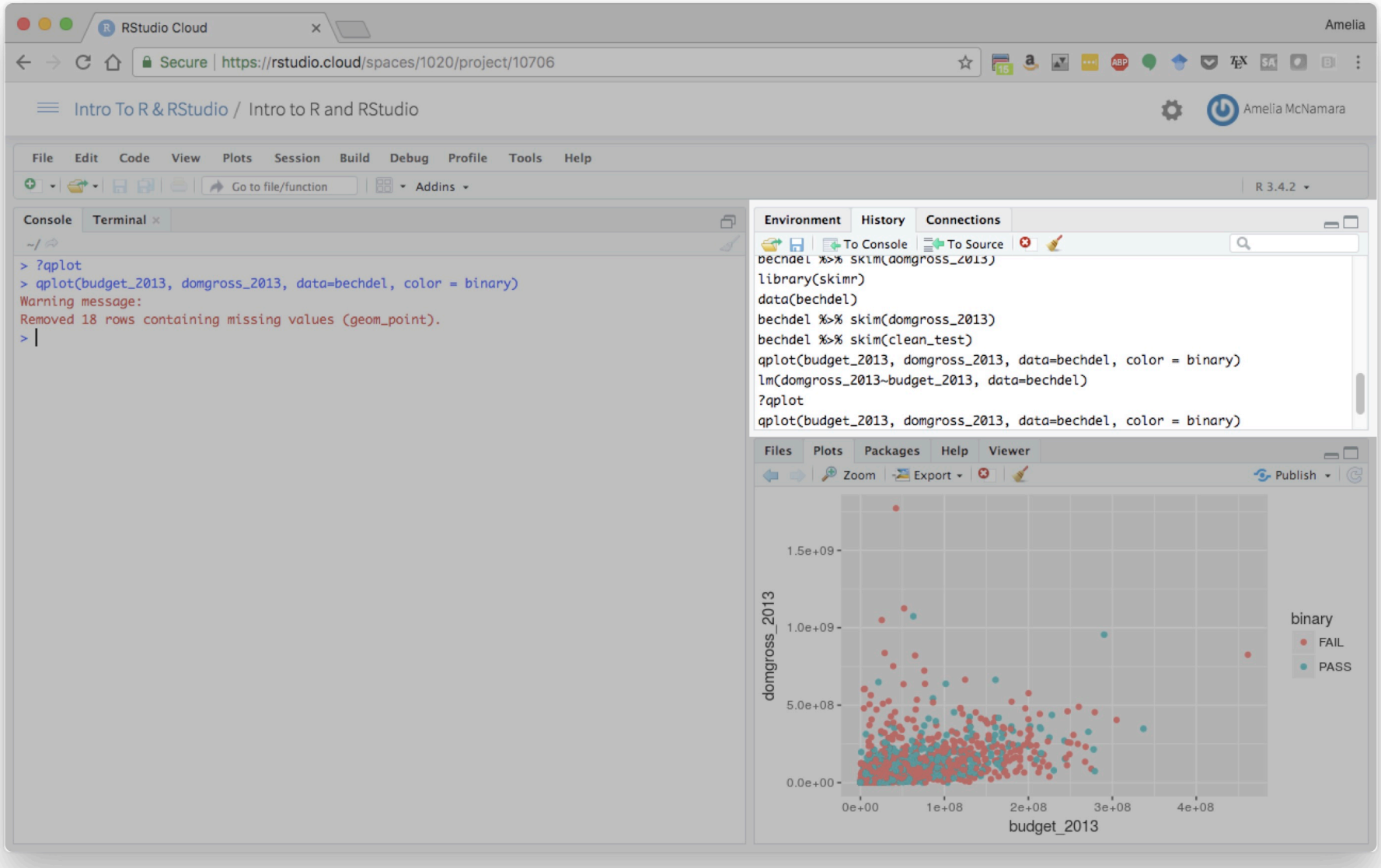
RStudio


RStudio Cloud
Amelia
Secure | https://rstudio.cloud/spaces/1020/project/10706
Intro To R & RStudio / Intro to R and RStudio
Amelia McNamara
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 3.4.2
Console Terminal
~/
> ?ggplot
>
Environment History Connections
Import Dataset
List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
R: Create a new ggplot
Find in Topic
ggplot {ggplot2}
R Documentation
Create a new ggplot
Description
ggplot() initializes a ggplot object. It can be used to declare the input data frame for a graphic and to specify the set of plot aesthetics intended to be common throughout all subsequent layers unless specifically overridden.
Usage
ggplot(data = NULL, mapping = aes(), ..., environment = parent.frame(
Arguments

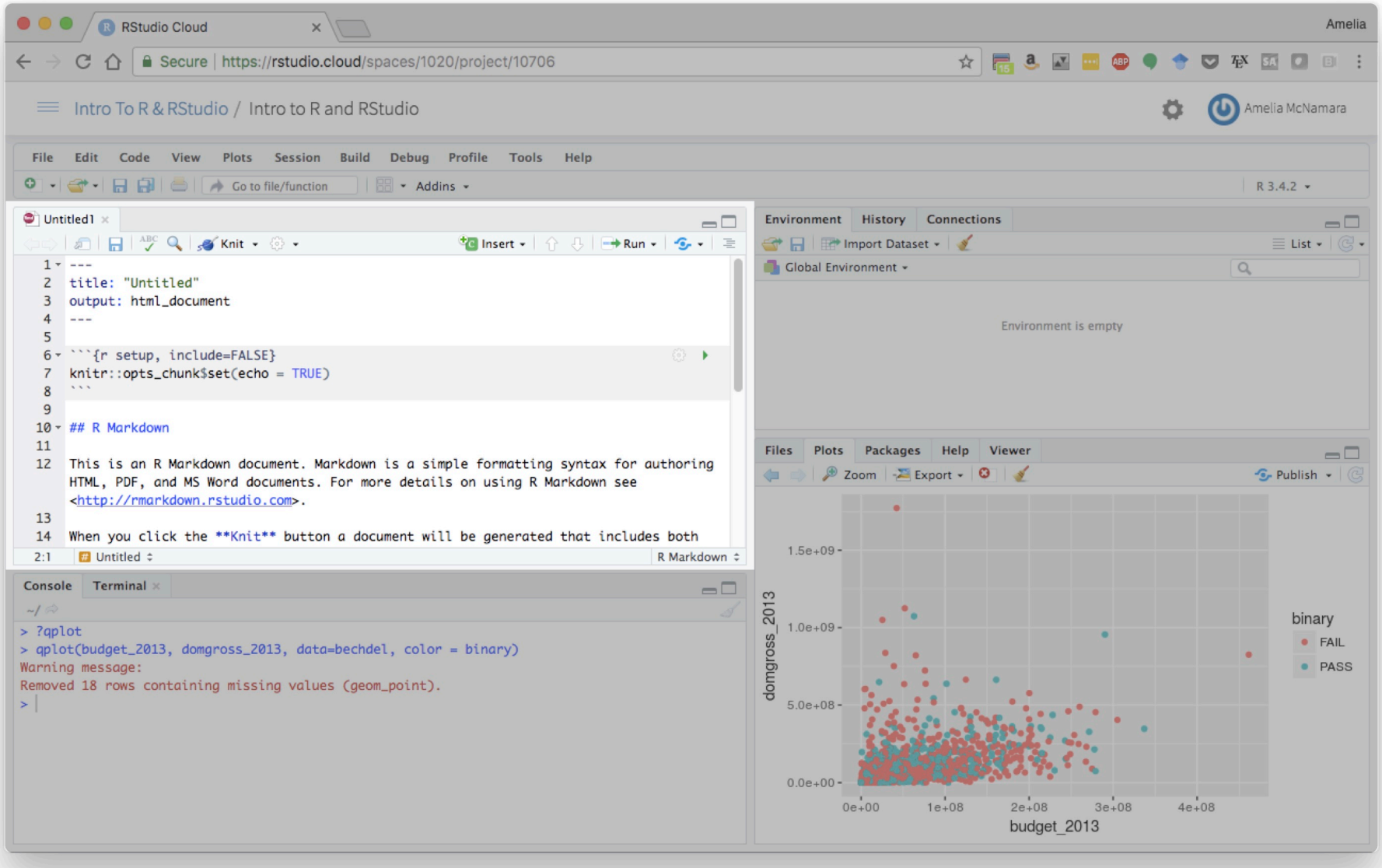
RStudio


RStudio Cloud
Amelia
Secure | https://rstudio.cloud/spaces/1020/project/10706
Intro To R & RStudio / Intro to R and RStudio
Amelia McNamara
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 3.4.2
Console Terminal
~/
> ?qplot
> qplot(budget_2013, domgross_2013, data=bechdel, color = binary)
Warning message:
Removed 18 rows containing missing values (geom_point).
>
Environment History Connections
Import Dataset
List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
Zoom
Export
Publish
1.5e+09
1.0e+09
5.0e+08
0.0e+00
domgross_2013
0e+00
1e+08
2e+08
3e+08
4e+08
budget_2013
binary
FAIL
PASS

RStudio

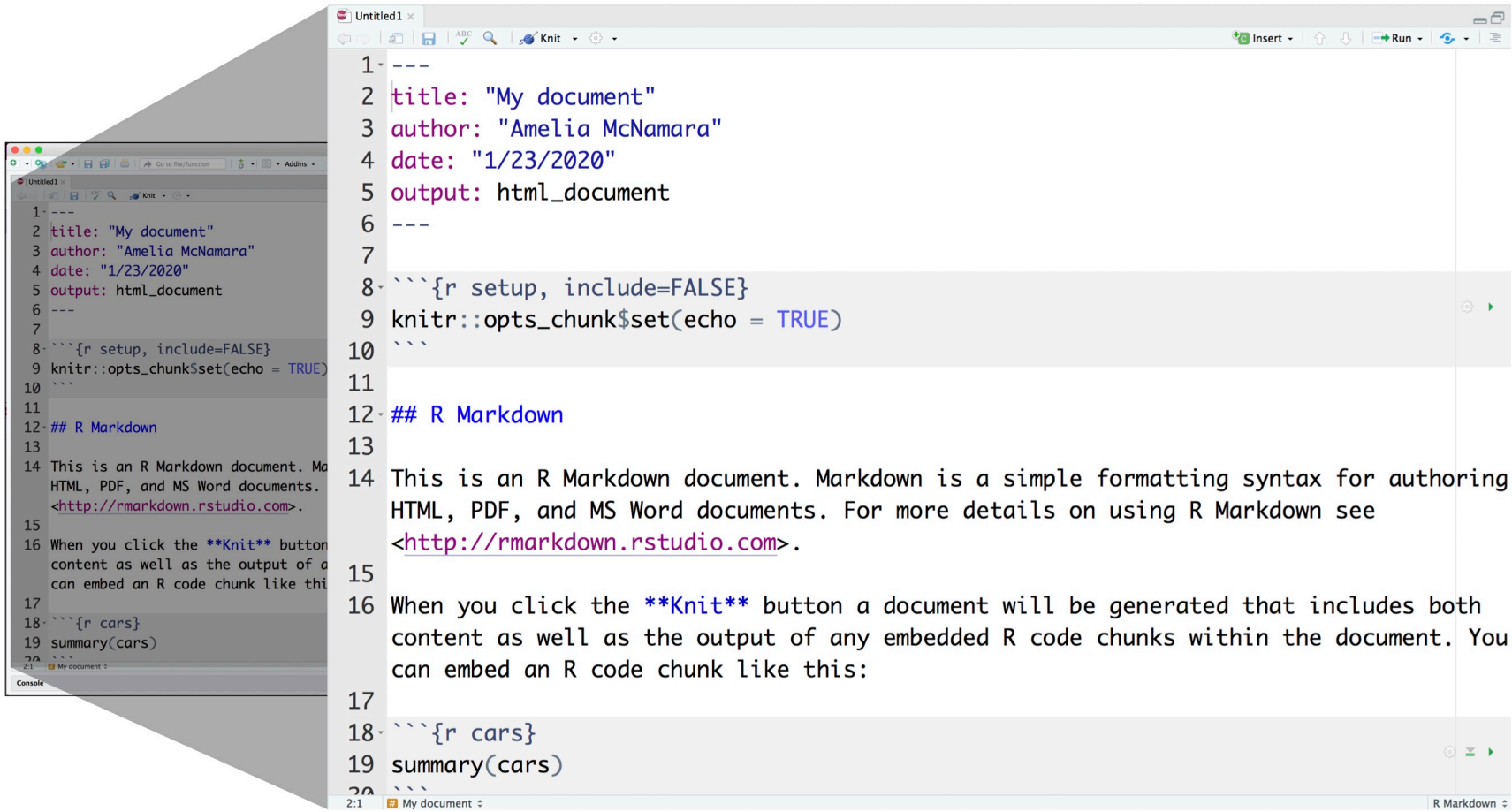

RStudio Cloud
Amelia
Secure | https://rstudio.cloud/spaces/1020/project/10706
Intro To R & RStudio / Intro to R and RStudio
Amelia McNamara
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 3.4.2
Console
Terminal
~/
> ?qplot
> qplot(budget_2013, domgross_2013, data=bechdel, color = binary)
Warning message:
Removed 18 rows containing missing values (geom_point).
>
Environment
History
Connections
To Console
To Source
bechdel %>% skim(domgross_2013)
library(skimr)
data(bechdel)
bechdel %>% skim(domgross_2013)
bechdel %>% skim(clean_test)
qplot(budget_2013, domgross_2013, data=bechdel, color = binary)
lm(domgross_2013~budget_2013, data=bechdel)
?qplot
qplot(budget_2013, domgross_2013, data=bechdel, color = binary)
Files
Plots
Packages
Help
Viewer
Zoom
Export
Publish
1.5e+09
1.0e+09
5.0e+08
0.0e+00
domgross_2013
0e+00
1e+08
2e+08
3e+08
4e+08
budget_2013
binary
FAIL
PASS

RStudio


RStudio Cloud
Amelia
Secure | https://rstudio.cloud/spaces/1020/project/10706
Intro To R & RStudio / Intro to R and RStudio
Amelia McNamara
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 3.4.2
Untitled1
Knit
Insert
Run
1 ---
2 title: "Untitled"
3 output: html_document
4 ---
5
6 ```{r setup, include=FALSE}
7 knitr::opts_chunk$set(echo = TRUE)
8 ```
9
10 ## R Markdown
11
12 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
13
14 When you click the **Knit** button a document will be generated that includes both
2:1 Untitled R Markdown
Console Terminal
~/
> ?qplot
> qplot(budget_2013, domgross_2013, data=bechdel, color = binary)
Warning message:
Removed 18 rows containing missing values (geom_point).
>
Environment History Connections
Import Dataset
List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
Zoom
Export
Publish
1.5e+09
1.0e+09
5.0e+08
0.0e+00
domgross_2013
0e+00
1e+08
2e+08
3e+08
4e+08
budget_2013
binary
FAIL
PASS

RMarkdown

An authoring format for Data Science.



RMarkdown (let's start!)