

Getting started with R and RStudio





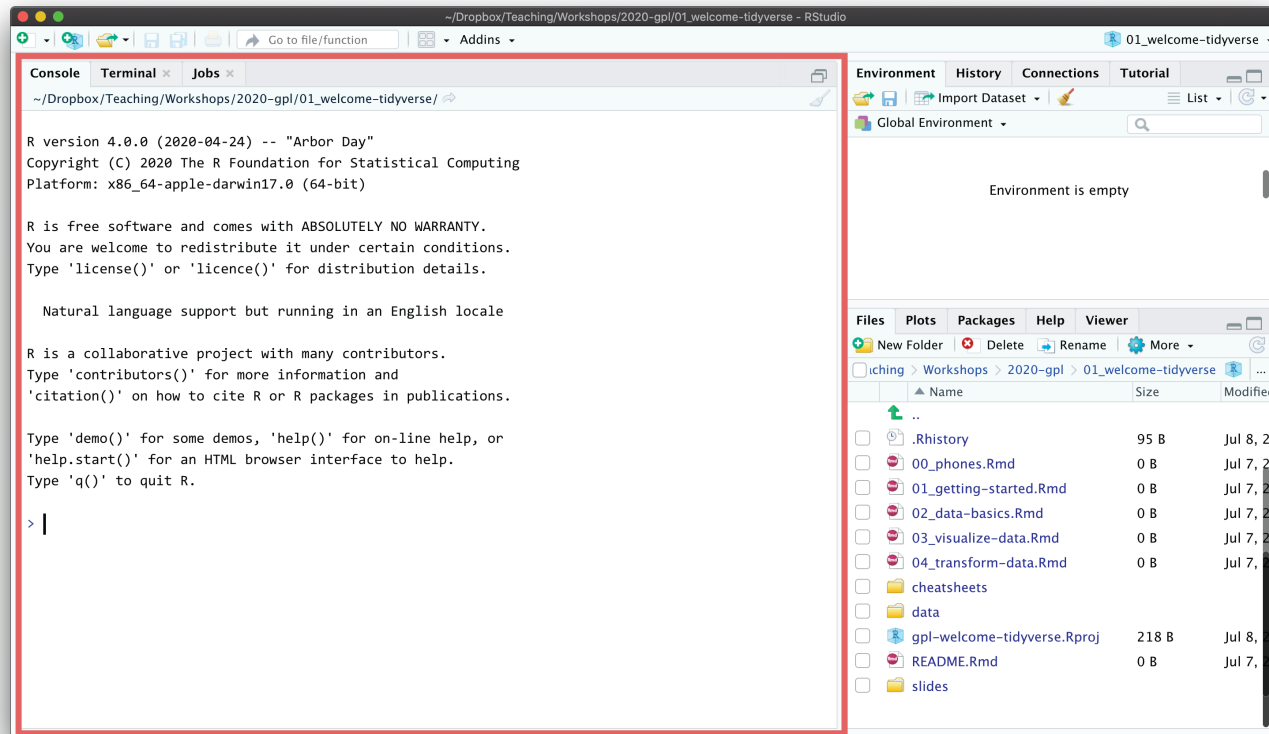
The engine



The dashboard

A tour of RStudio

Console



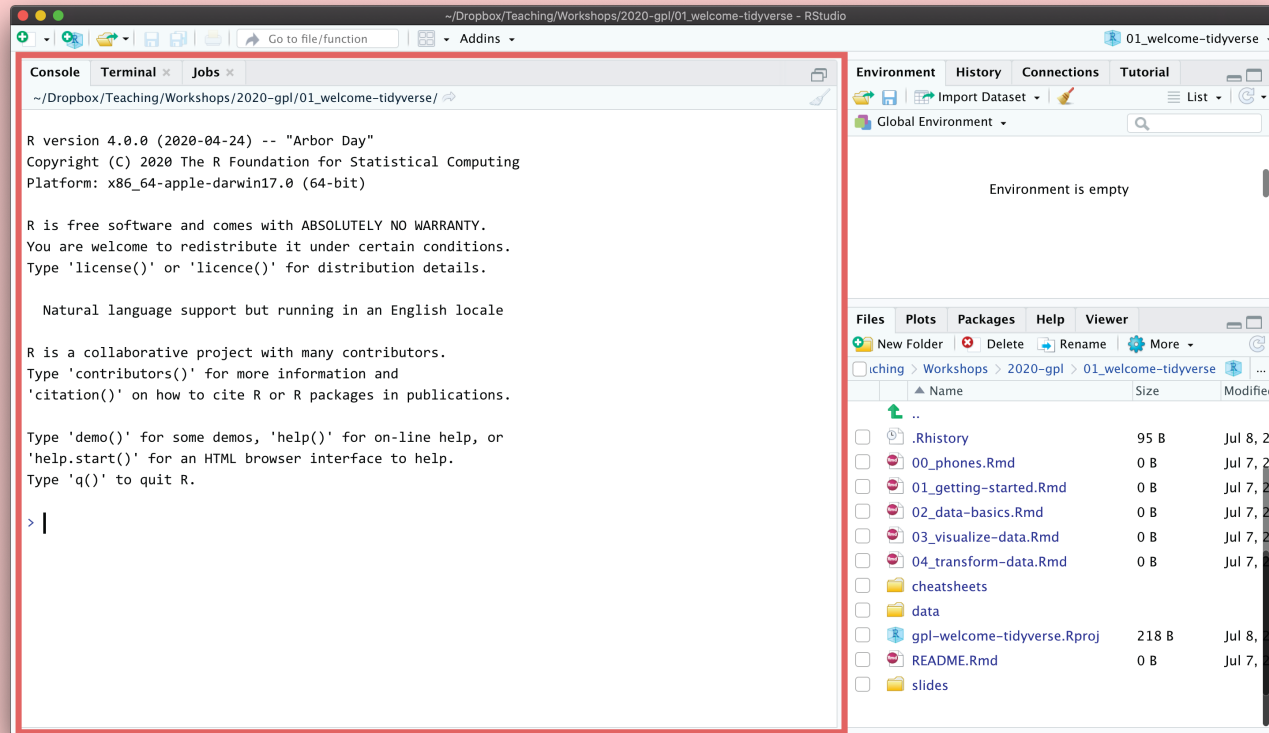
The screenshot shows the RStudio interface. The console pane on the left displays the R startup message, including the version (4.0.0), copyright information, and instructions for using the software. The environment pane on the right shows that the environment is currently empty. The files pane at the bottom lists the project files, including Rmd files for various topics and a project file.

```
~/Dropbox/Teaching/Workshops/2020-gpl/01_welcome-tidyverse/ RStudio  
01_welcome-tidyverse  
Environment History Connections Tutorial  
Global Environment  
Environment is empty  
Files Plots Packages Help Viewer  
New Folder Delete Rename More  
iChing > Workshops > 2020-gpl > 01_welcome-tidyverse  
Name Size Modified  
..  
[ ] .Rhistory 95 B Jul 8, 2  
[ ] 00_phones.Rmd 0 B Jul 7, 2  
[ ] 01_getting-started.Rmd 0 B Jul 7, 2  
[ ] 02_data-basics.Rmd 0 B Jul 7, 2  
[ ] 03_visualize-data.Rmd 0 B Jul 7, 2  
[ ] 04_transform-data.Rmd 0 B Jul 7, 2  
[ ] cheatsheets  
[ ] data  
[ ] gpl-welcome-tidyverse.Rproj 218 B Jul 8,  
[ ] README.Rmd 0 B Jul 7,  
[ ] slides
```

R is awaiting your instructions

Type code here, press enter, and R will run it

Your turn



The screenshot shows the RStudio interface. The console window on the left contains the following text:

```
~/Dropbox/Teaching/Workshops/2020-gpl/01_welcome-tidyverse/

R version 4.0.0 (2020-04-24) -- "Arbor Day"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin17.0 (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.

> |
```

The file explorer on the right shows a list of files and folders:

Name	Size	Modified
..		
.Rhistory	95 B	Jul 8, 2
00_phones.Rmd	0 B	Jul 7, 2
01_getting-started.Rmd	0 B	Jul 7, 2
02_data-basics.Rmd	0 B	Jul 7, 2
03_visualize-data.Rmd	0 B	Jul 7, 2
04_transform-data.Rmd	0 B	Jul 7, 2
cheatsheets		
data		
gpl-welcome-tidyverse.Rproj	218 B	Jul 8, 2
README.Rmd	0 B	Jul 7, 2
slides		

Type `2 + 2` in the console

Press enter

```
2 + 2
```

```
## [1] 4
```

**This is ephemeral though.
If you want to run this again, you'll have to type it again.**

Store R code in a document instead

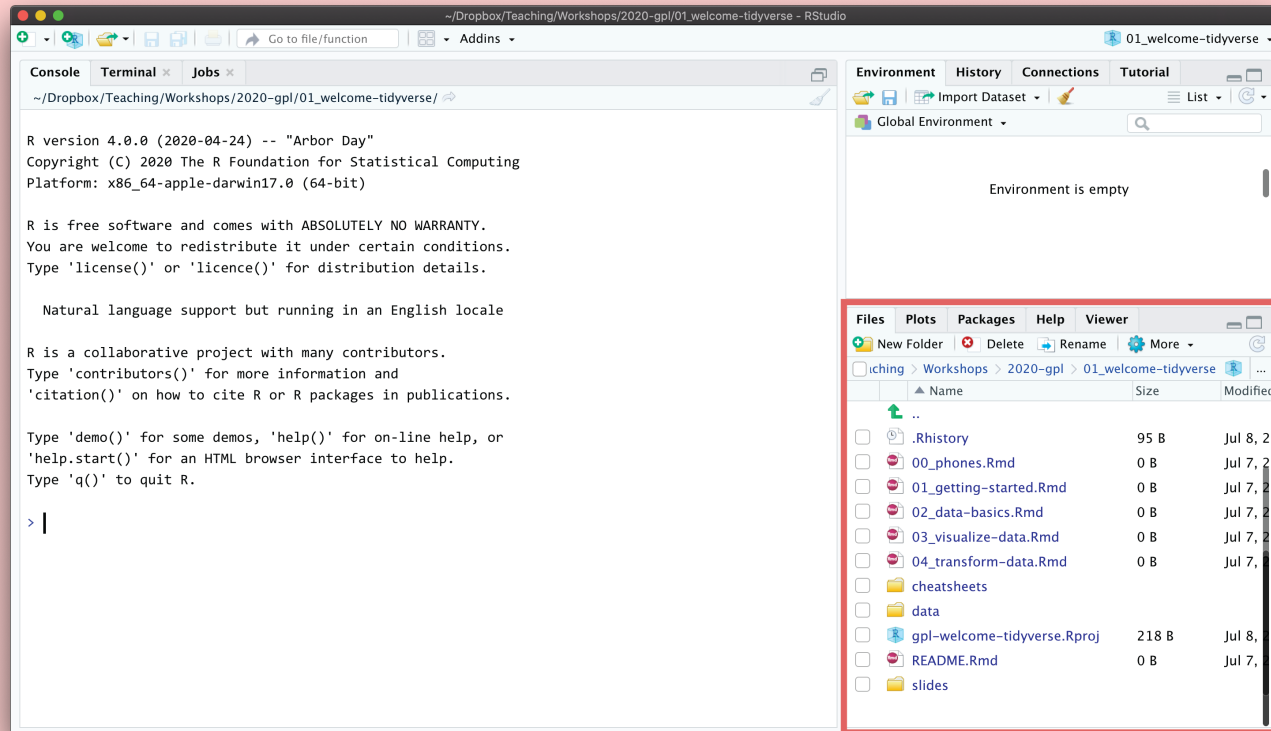
Files pane

The screenshot shows the RStudio interface with three main panes. The Console pane on the left displays the R version 4.0.0 (2020-04-24) -- "Arbor Day" and copyright information. The Environment pane in the top right shows "Global Environment" and "Environment is empty". The Files pane in the bottom right, highlighted with a red border, shows a file explorer view of the current working directory. The file list includes:

Name	Size	Modified
..		
.Rhistory	95 B	Jul 8, 2
00_phones.Rmd	0 B	Jul 7, 2
01_getting-started.Rmd	0 B	Jul 7, 2
02_data-basics.Rmd	0 B	Jul 7, 2
03_visualize-data.Rmd	0 B	Jul 7, 2
04_transform-data.Rmd	0 B	Jul 7, 2
cheatsheets		
data		
gpl-welcome-tidyverse.Rproj	218 B	Jul 8,
README.Rmd	0 B	Jul 7,
slides		

All the files in your current working directory

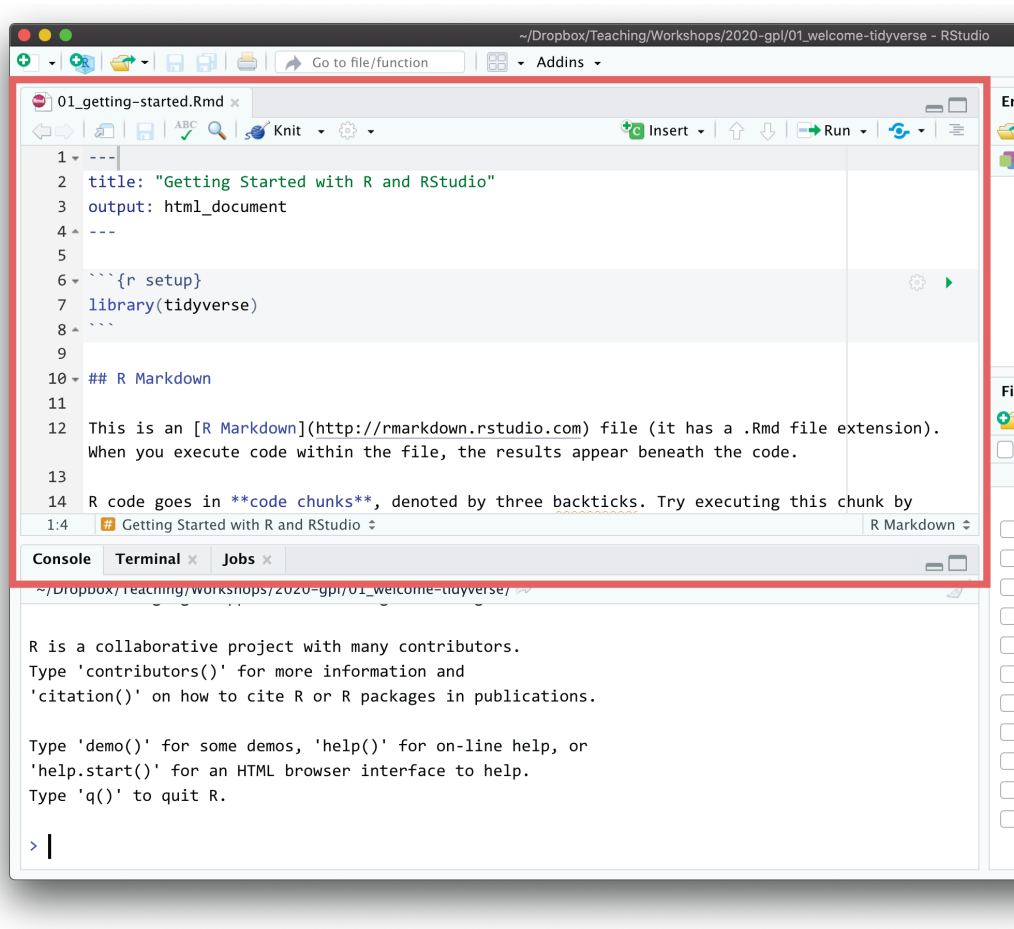
Your turn



Find 01_getting-started.Rmd

Click on its name to open the file

Source pane

A screenshot of the RStudio Source Pane. The pane is titled "01_getting-started.Rmd" and contains R Markdown code. The code is as follows:

```
1 ---
2 title: "Getting Started with R and RStudio"
3 output: html_document
4 ---
5
6 ```{r setup}
7 library(tidyverse)
8 ```
9
10 ## R Markdown
11
12 This is an [R Markdown](http://rmarkdown.rstudio.com) file (it has a .Rmd file extension).
13 When you execute code within the file, the results appear beneath the code.
14
15 R code goes in code chunks, denoted by three backticks. Try executing this chunk by
16 # Getting Started with R and RStudio
```

The Source Pane is highlighted with a red border. Below the Source Pane is the Console pane, which displays the R help text for the 'tidyverse' package:

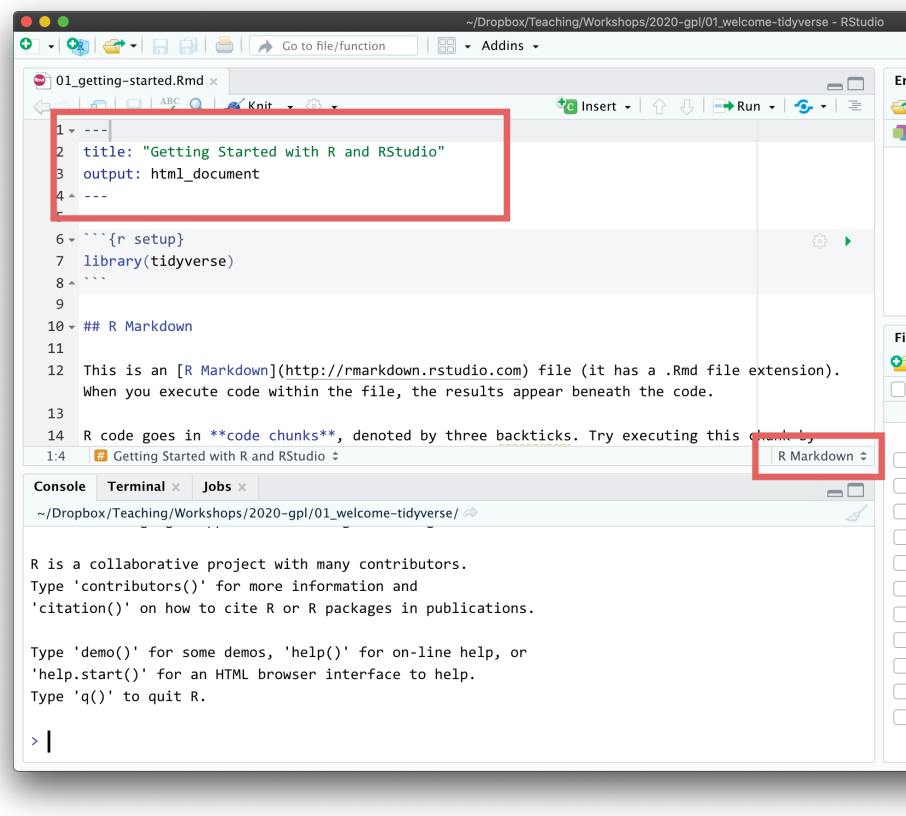
```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

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.

> |
```

Documents
open here

R Markdown



The screenshot shows the RStudio interface with a file named '01_getting-started.Rmd'. The editor contains the following code:

```
1 ---
2 title: "Getting Started with R and RStudio"
3 output: html_document
4 ---
5
6 ```{r setup}
7 library(tidyverse)
8 ```
9
10 ## R Markdown
11
12 This is an [R Markdown](http://rmarkdown.rstudio.com) file (it has a .Rmd file extension).
13 When you execute code within the file, the results appear beneath the code.
14
15 R code goes in code chunks, denoted by three backticks. Try executing this chunk by
16 clicking the Run button in the top right corner.
```

The code is executed, and the console shows the following output:

```
~/Dropbox/Teaching/Workshops/2020-gpl/01_welcome-tidyverse/

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.

> |
```

Two red boxes highlight the YAML header and the 'R Markdown' label in the console.

Document format that
combines text and code

Acts like a notebook
for your analysis

R Markdown



The screenshot shows an R Markdown document titled "01_getting-started.Rmd" in RStudio. The document content is as follows:

```
10 ## R Markdown
11
12 This is an [R Markdown](http://rmarkdown.rstudio.com) file (it has a .Rmd file extension).
13 When you execute code within the file, the results appear beneath the code.
14 R code goes in code chunks, denoted by three backticks. Try executing this chunk by
15 clicking the Run button (a small green triangle) within the chunk or by placing your cursor
16 inside it and pressing Ctrl+Shift+Enter (or Cmd+Shift+Enter on Mac).
17
18 ```{r}
19 ggplot(data = mpg) +
20   geom_point(mapping = aes(x = cty, y = hwy), alpha = 0.2)
21 ```
```

Below the code, a scatter plot is displayed. The y-axis is labeled "hwy" and ranges from 30 to 40. The x-axis is labeled "cty" and ranges from 10 to 20. The plot shows a positive correlation between city mileage (cty) and highway mileage (hwy). The points are semi-transparent (alpha = 0.2) and are colored in shades of gray.

Text

R Markdown



The screenshot shows an R Markdown document titled "01_getting-started.Rmd" in RStudio. The document contains the following content:

```
10 - ## R Markdown
11
12 This is an [R Markdown](http://rmarkdown.rstudio.com) file (it has a .Rmd file extension).
    When you execute code within the file, the results appear beneath the code.
13
14 R code goes in code chunks, denoted by three backticks. Try executing this chunk by
    clicking the Run button (a small green triangle) within the chunk or by placing your cursor
    inside it and pressing Ctrl+Shift+Enter (or Cmd+Shift+Enter on Mac).
15
16 ```{r}
17 ggplot(data = mpg) +
18   geom_point(mapping = aes(x = cty, y = hwy), alpha = 0.2)
19 ```
```

Below the code chunk, a scatter plot is displayed. The y-axis is labeled "hwy" and ranges from 30 to 40. The x-axis is labeled "cty" and ranges from 10 to 20. The plot shows a positive correlation between city mileage (cty) and highway mileage (hwy). The points are semi-transparent (alpha = 0.2) and are colored in shades of gray.

Text

Code

R Markdown



The screenshot shows an R Markdown document titled "01_getting-started.Rmd" in RStudio. The document contains the following text and code:

```
10 - ## R Markdown
11
12 This is an [R Markdown](http://rmarkdown.rstudio.com) file (it has a .Rmd file extension).
    When you execute code within the file, the results appear beneath the code.
13
14 R code goes in code chunks, denoted by three backticks. Try executing this chunk by
    clicking the Run button (a small green triangle) within the chunk or by placing your cursor
    inside it and pressing Ctrl+Shift+Enter (or Cmd+Shift+Enter on Mac).
15
16 ```{r}
17 ggplot(data = mpg) +
18   geom_point(mapping = aes(x = cty, y = hwy), alpha = 0.2)
19 ```
```

The output of the code is a scatter plot showing the relationship between city miles per gallon (cty) on the x-axis and highway miles per gallon (hwy) on the y-axis. The plot is enclosed in a red box. The y-axis is labeled "hwy" and has tick marks at 30 and 40. The x-axis is labeled "cty" and has tick marks at 10, 20, 30, 40, and 50. The plot shows a positive correlation between cty and hwy, with data points scattered across the grid.

Text


Code

Output

Your turn

Read the instructions

Run the code chunk by clicking the play button



The screenshot shows the RStudio interface with a file named "01_getting-started.Rmd". The editor contains the following text:

```
10 - ## R Markdown
11
12 This is an [R Markdown](http://rmarkdown.rstudio.com) file (it has a .Rmd file extension).
    When you execute code within the file, the results appear beneath the code.
13
14 R code goes in code chunks, denoted by three backticks. Try executing this chunk by
    clicking the Run button (a small green triangle) within the chunk or by placing your cursor
    inside it and pressing Ctrl+Shift+Enter (or Cmd+Shift+Enter on Mac).
15
16 ```{r}
17 ggplot(data = mpg) +
18   geom_point(mapping = aes(x = cty, y = hwy), alpha = 0.2)
19 ```
```

The code chunk is highlighted with a red box. Below the code, a scatter plot is displayed with "hwy" on the y-axis and "cty" on the x-axis. The plot shows a positive correlation between city miles per gallon and highway miles per gallon for various car models.

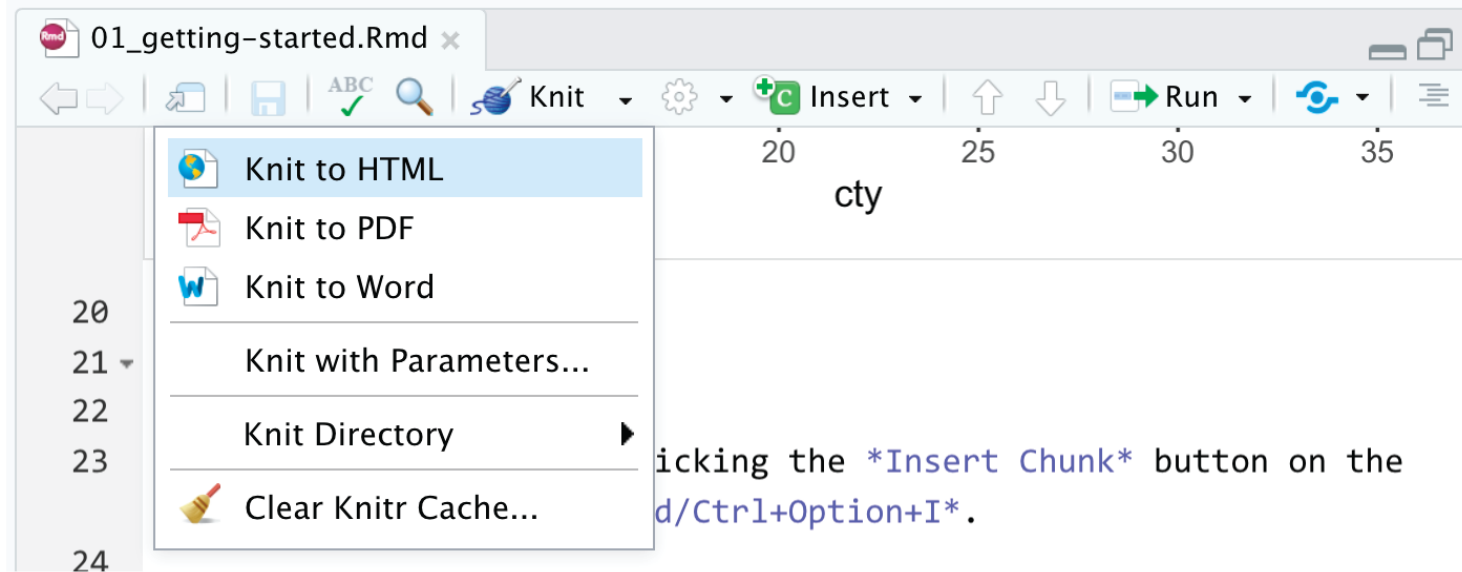
Your turn

Add a new chunk

Put $2 + 2$ in the chunk and run it

Knitting

"Knit" an R Markdown document into a standalone sharable file



R Markdown

The best way to combine R code and narrative

We'll use it throughout the workshop:

I'll provide starter code

You'll complete "Your turns"

In the end, you'll have an annotated record for yourself

Your turn

Spot the difference:

```
filter(mtcars, cyl == 4)
```

```
four_cyls <- filter(mtcars, cyl == 4)
```

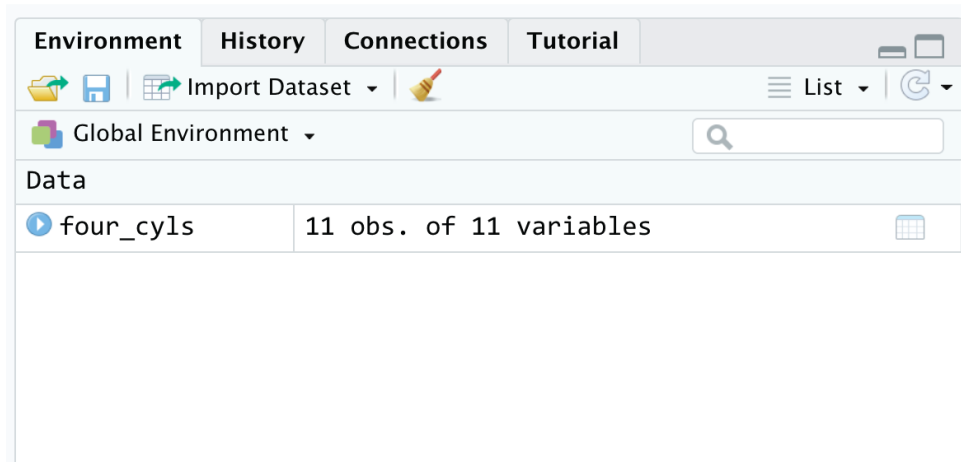
**Find these chunks in the notebook and run them.
What's different about what happens?**

Assignment

<- assigns the output from the righthand side to a variable with the name on the lefthand side

```
four_cyls <- filter(mtcars, cyl == 4)
```

Environment pane



List of all the
variables you've created

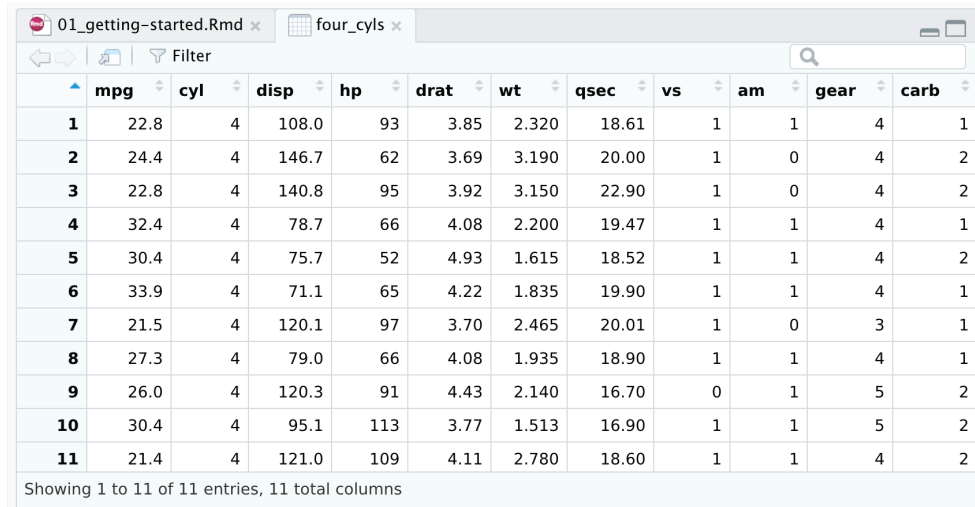
Your turn

Find `four_cyls` in the environment pane.
Click on the name `four_cyls`

What happens?

Viewer

Clicking on an object in the environment panel opens it an interactive viewer tab



The screenshot shows an R Studio Viewer window with a tab titled 'four_cyls'. The window displays a data table with 11 rows and 11 columns. The columns are labeled mpg, cyl, disp, hp, drat, wt, qsec, vs, am, gear, and carb. The rows are numbered 1 through 11. The table is displayed in a light blue theme. Below the table, it says 'Showing 1 to 11 of 11 entries, 11 total columns'.

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
1	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
2	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
3	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
4	32.4	4	78.7	66	4.08	2.200	19.47	1	1	4	1
5	30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
6	33.9	4	71.1	65	4.22	1.835	19.90	1	1	4	1
7	21.5	4	120.1	97	3.70	2.465	20.01	1	0	3	1
8	27.3	4	79.0	66	4.08	1.935	18.90	1	1	4	1
9	26.0	4	120.3	91	4.43	2.140	16.70	0	1	5	2
10	30.4	4	95.1	113	3.77	1.513	16.90	1	1	5	2
11	21.4	4	121.0	109	4.11	2.780	18.60	1	1	4	2

Functions

```
four_cyls <- filter(mtcars, cyl == 4)
```

Functions do things

Functions take arguments, output results

If you want to keep the output, assign it to a variable

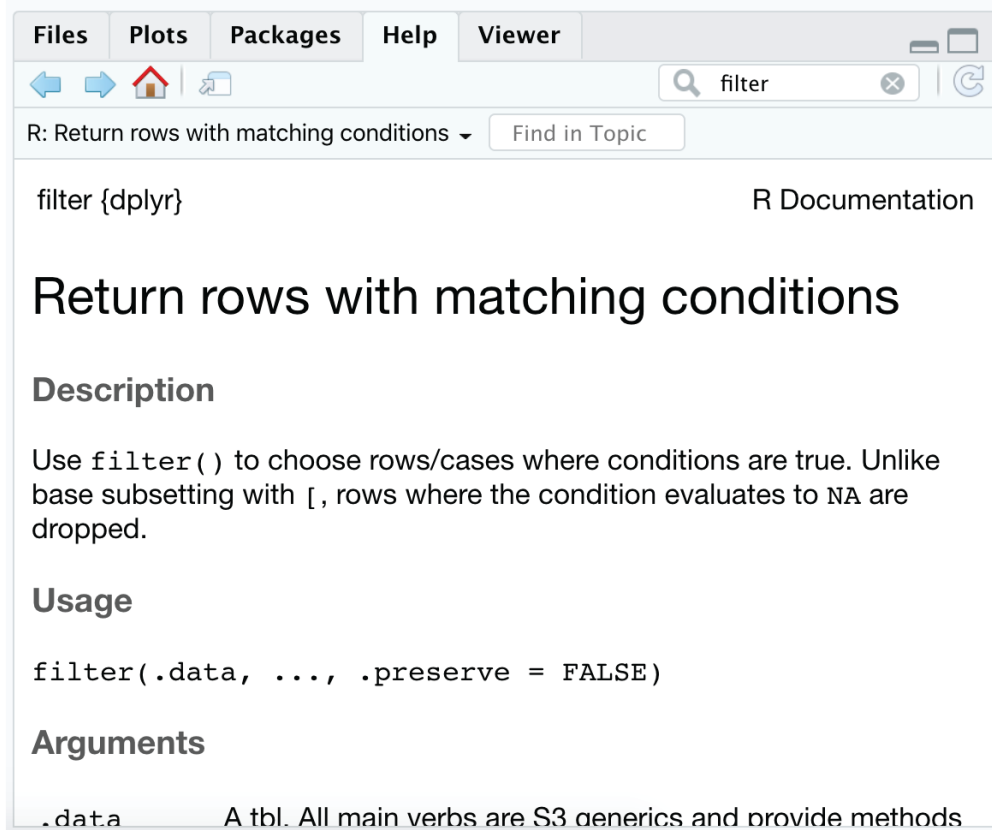
Help

To look up the help page for an R function,
type this in the console:

```
?function_name
```

(Or google it!)

Help pane



The screenshot shows the R Help pane with the following content:

Files Plots Packages Help Viewer

filter

R: Return rows with matching conditions Find in Topic

filter {dplyr} R Documentation

Return rows with matching conditions

Description

Use `filter()` to choose rows/cases where conditions are true. Unlike base subsetting with `[]`, rows where the condition evaluates to `NA` are dropped.

Usage

```
filter(.data, ..., .preserve = FALSE)
```

Arguments

`.data` A tbl. All main verbs are S3 generics and provide methods

These help pages provide details about the arguments you can supply a function

Often full of examples at the bottom

Your turn

Look at the help page for `seq`

Add a chunk that uses `seq()` to create a list of numbers from 3 to 45, spaced by 1.5 (so 3, 4.5, 6, 7.5, ...)

02:00

```
seq(from = 3, to = 45, by = 1.5)
```

```
## [1] 3.0 4.5 6.0 7.5 9.0 10.5 12.0 13.5 15.0 16.5 18.0 19.5  
## [16] 25.5 27.0 28.5 30.0 31.5 33.0 34.5 36.0 37.5 39.0 40.5 42.0
```

