

# Getting started with R and RStudio





The engine

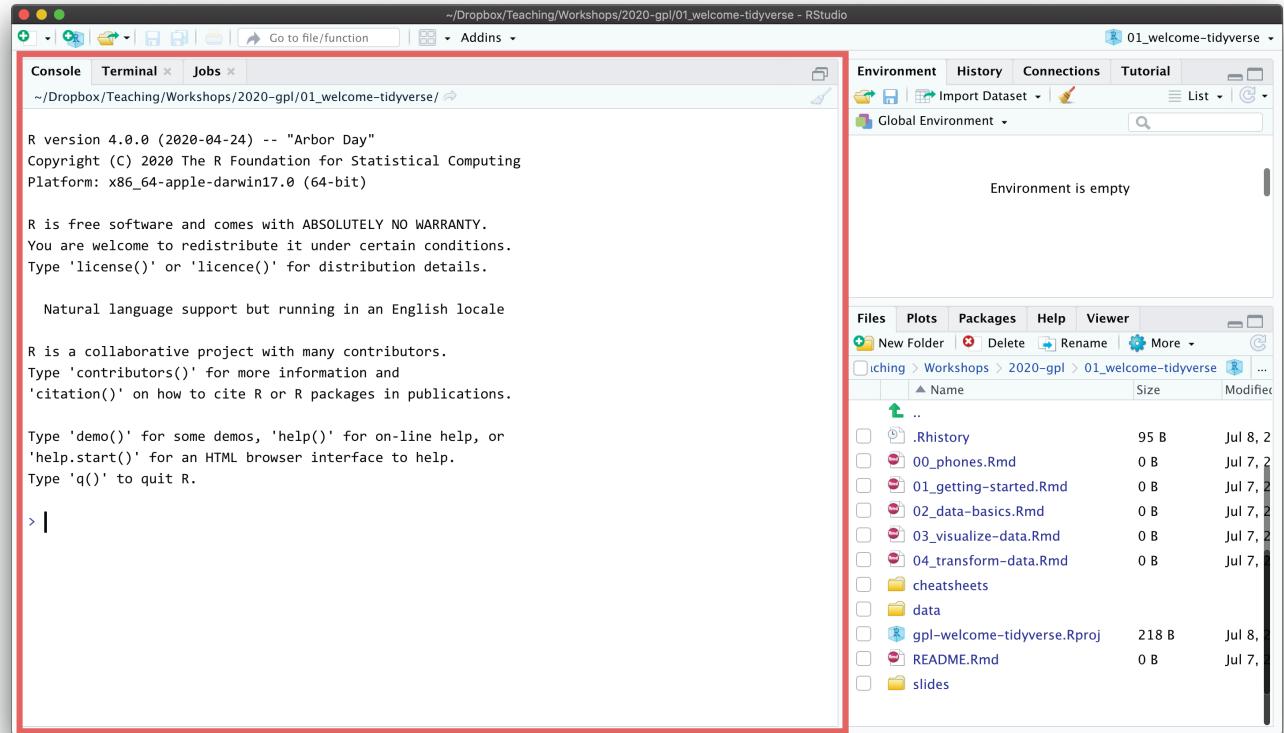


The dashboard

# A tour of RStudio



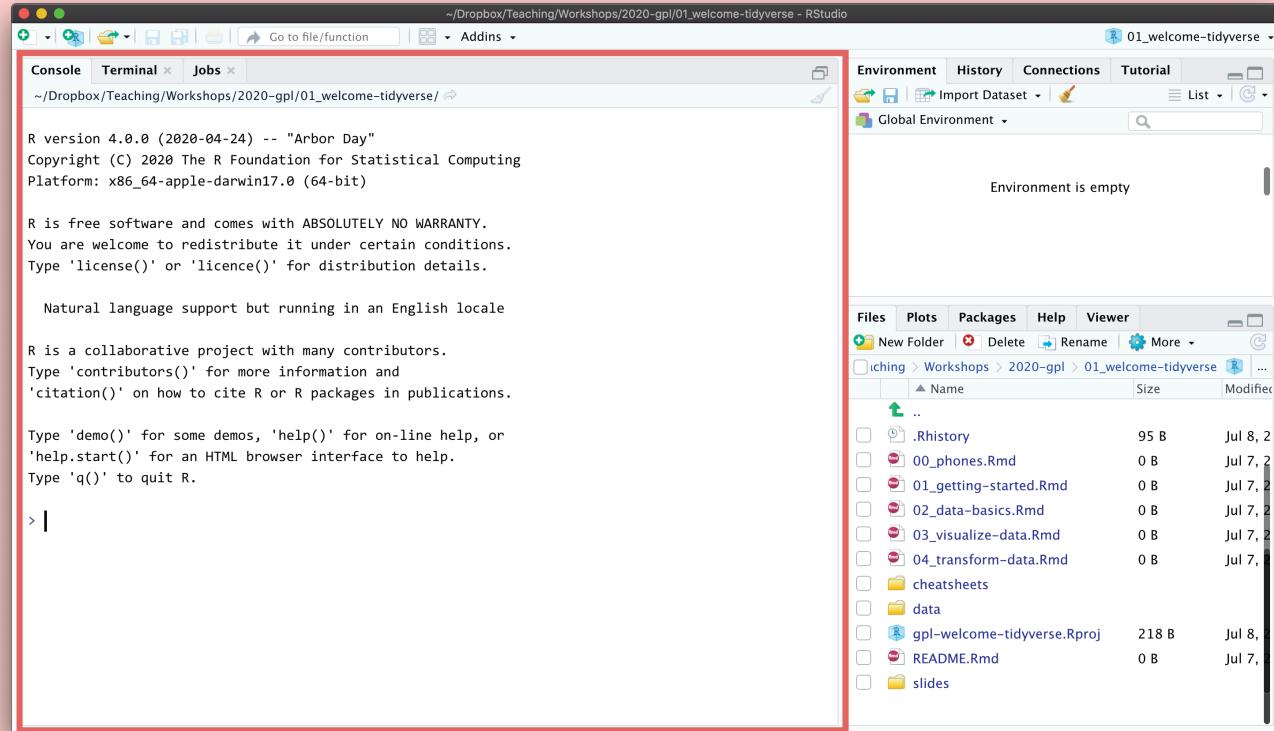
# Console



R is awaiting your  
instructions

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

# Your turn



A screenshot of the RStudio interface. The left pane shows the R console with the following text:

```
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 right pane shows the file browser with the following directory structure:

Name	Size	Modified
.Rhistory	95 B	Jul 8, 2020
00_phones.Rmd	0 B	Jul 7, 2020
01_getting-started.Rmd	0 B	Jul 7, 2020
02_data-basics.Rmd	0 B	Jul 7, 2020
03_visualize-data.Rmd	0 B	Jul 7, 2020
04_transform-data.Rmd	0 B	Jul 7, 2020
cheatsheets		
data		
gpl-welcome-tidyverse.Rproj	218 B	Jul 8, 2020
README.Rmd	0 B	Jul 7, 2020
slides		

Type 2 + 2 in the  
console

Press enter

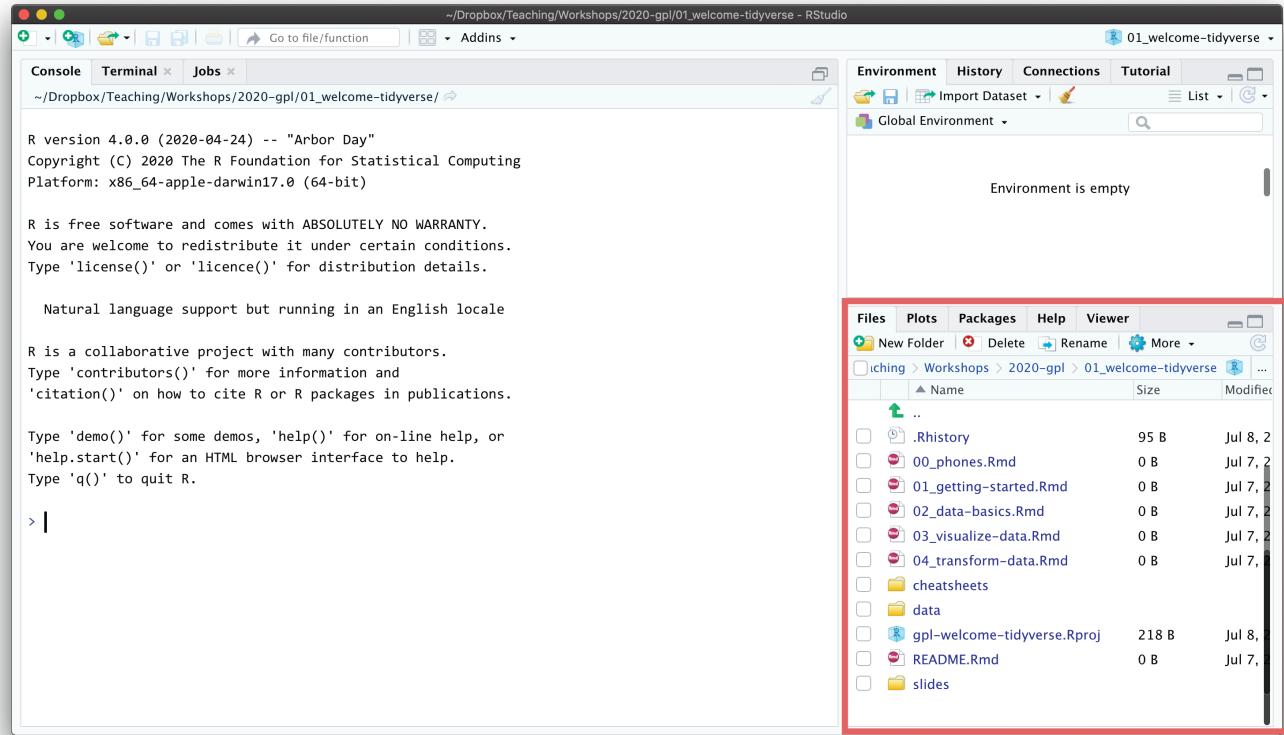
2 + 2

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

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

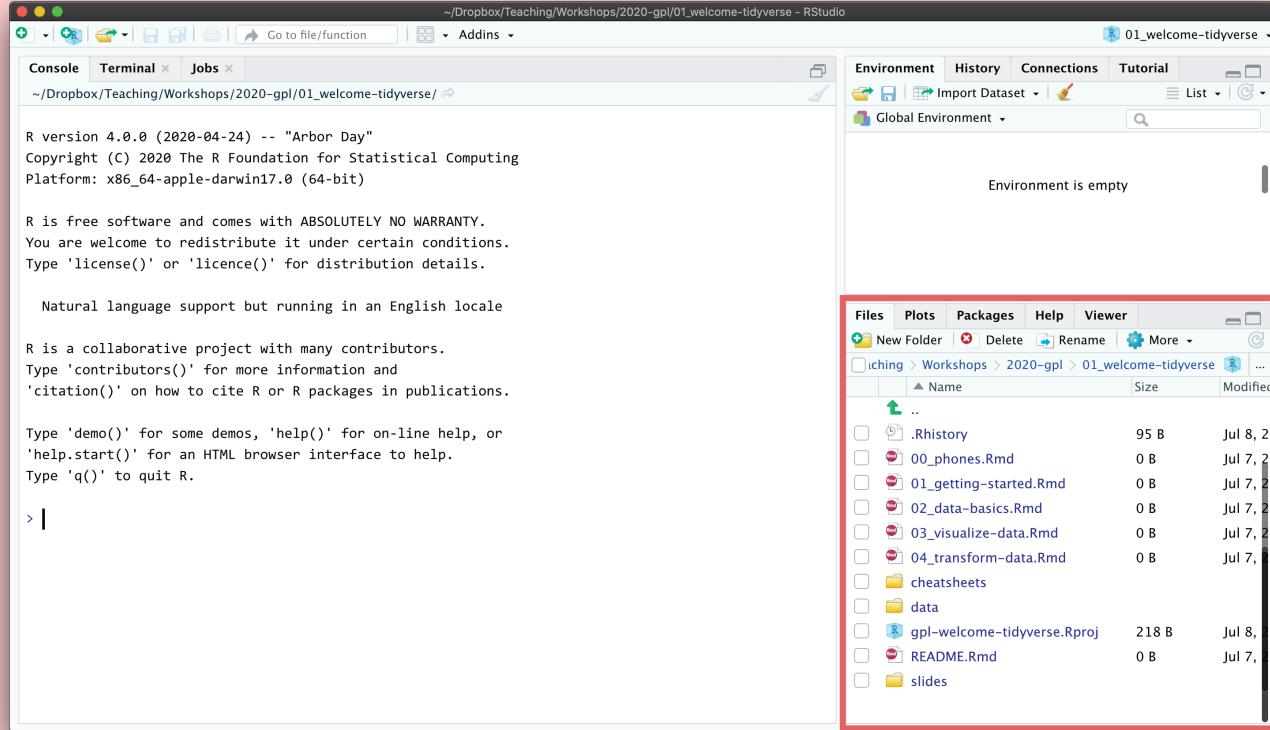
Store R code in a document instead

# Files pane



All the files in your  
current working  
directory

# Your turn



The screenshot shows the RStudio interface. The top bar indicates the current project is "01\_welcome-tidyverse". The left pane, titled "Console", displays the standard R startup message. The right pane, titled "Environment", shows that the environment is currently empty. Below the console, the "File Browser" pane is open, showing the directory structure of the project:

Name	Size	Modified
.Rhistory	95 B	Jul 8, 2020
00_phones.Rmd	0 B	Jul 7, 2020
01_getting-started.Rmd	0 B	Jul 7, 2020
02_data-basics.Rmd	0 B	Jul 7, 2020
03_visualize-data.Rmd	0 B	Jul 7, 2020
04_transform-data.Rmd	0 B	Jul 7, 2020
cheatsheets		
data		
gpl-welcome-tidyverse.Rproj	218 B	Jul 8, 2020
README.Rmd	0 B	Jul 7, 2020
slides		

Find `01_getting-started.Rmd`

Click on its name to open the file

# Source pane

The screenshot shows the RStudio interface. The main window displays an R Markdown file named "01\_getting-started.Rmd". The code in the file includes a YAML header, an R chunk for setup, and a section about R Markdown. The R console at the bottom shows the standard R welcome message.

```
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).  
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  
1:4 # Getting Started with R and RStudio
```

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 an R Markdown file open. The code editor window displays the following content:

```
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).  
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  
15 Getting Started with R and RStudio
```

The first two code blocks (lines 1-4) and the third one (line 14) are highlighted with red boxes. The RStudio interface includes a toolbar, a sidebar with environment, file, and help tabs, and a console tab at the bottom.

Document format that combines text and code

Acts like a notebook for your analysis

# R Markdown

The screenshot shows the RStudio interface with an R Markdown file open. The code editor contains the following content:

```
## R Markdown
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.
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).
````{r}
ggplot(data = mpg) +
  geom_point(mapping = aes(x = cty, y = hwy), alpha = 0.2)
````
```

Below the code editor, a plot is displayed showing the relationship between city fuel economy (cty) on the x-axis and highway fuel economy (hwy) on the y-axis. The plot uses a semi-transparent dot aesthetic where the alpha value is set to 0.2.

## Text

# R Markdown

The screenshot shows the RStudio interface with the following details:

- File:** 01\_getting-started.Rmd
- Code Editor:** Contains R code. Lines 16-19 are highlighted with a red box:

```
16 `~{r}
17 ggplot(data = mpg) +
18   geom_point(mapping = aes(x = cty, y = hwy), alpha = 0.2)
19 ~~`
```
- Plot:** A scatter plot generated by the code. The x-axis is labeled "cty" and the y-axis is labeled "hwy". The plot shows a positive correlation between fuel economy in the city (cty) and on the highway (hwy). Most data points are clustered between 30 and 40 on both axes, with a few outliers at higher values.

Text

Code

# R Markdown

The screenshot shows the RStudio interface with an R Markdown file open. The code editor contains the following content:

```
## R Markdown
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.
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).
```{r}
ggplot(data = mpg) +
  geom_point(mapping = aes(x = cty, y = hwy), alpha = 0.2)
```

```

Below the code editor is a scatter plot generated by the R code. The x-axis is labeled "cty" and the y-axis is labeled "hwy". The plot shows a positive correlation between fuel economy in the city (cty) and on the highway (hwy). A red box highlights the plot area.

Text

Code

Output

# Your turn

The screenshot shows the RStudio interface with an R Markdown file open. The code editor displays the following content:

```
## R Markdown
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.
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).
````{r}
ggplot(data = mpg) +
  geom_point(mapping = aes(x = cty, y = hwy), alpha = 0.2)
````
```

Below the code editor is a scatter plot generated by the R code. The x-axis is labeled "cty" and the y-axis is labeled "hwy". The plot shows a positive correlation between fuel economy in the city (cty) and on the highway (hwy). Most data points are clustered between 30 and 45 on both axes, with a few outliers at higher values.

Read the instructions

Run the code chunk by clicking the play button

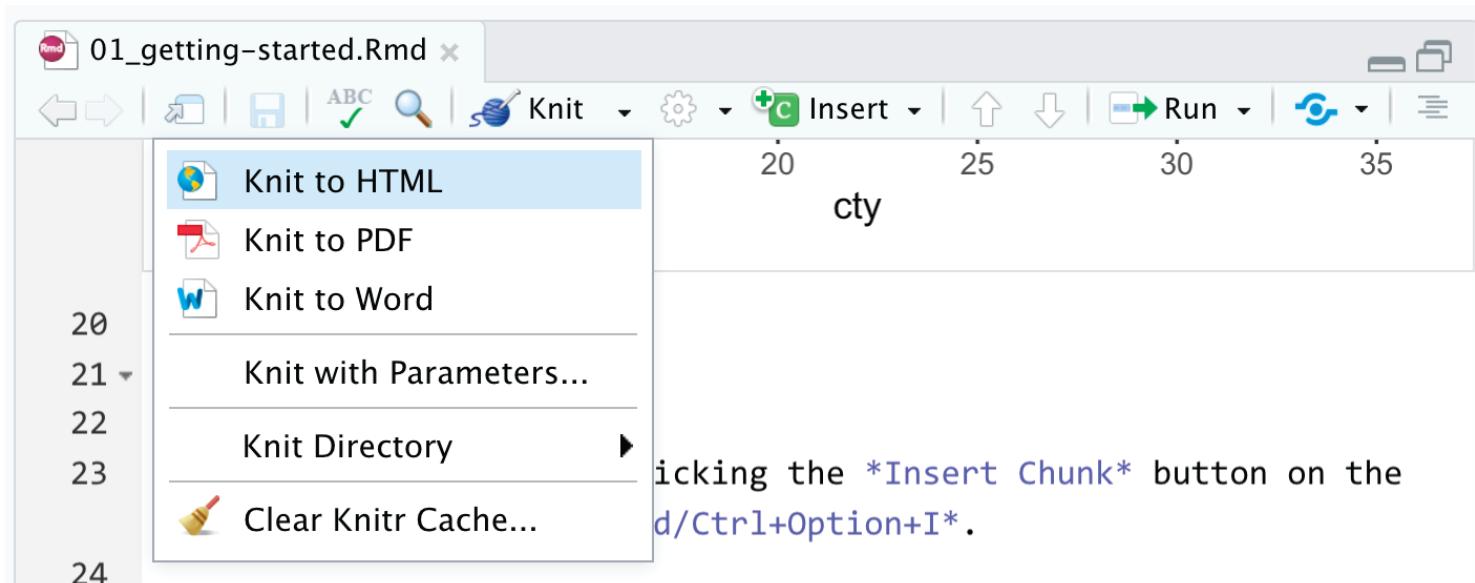
# Your turn

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)
```

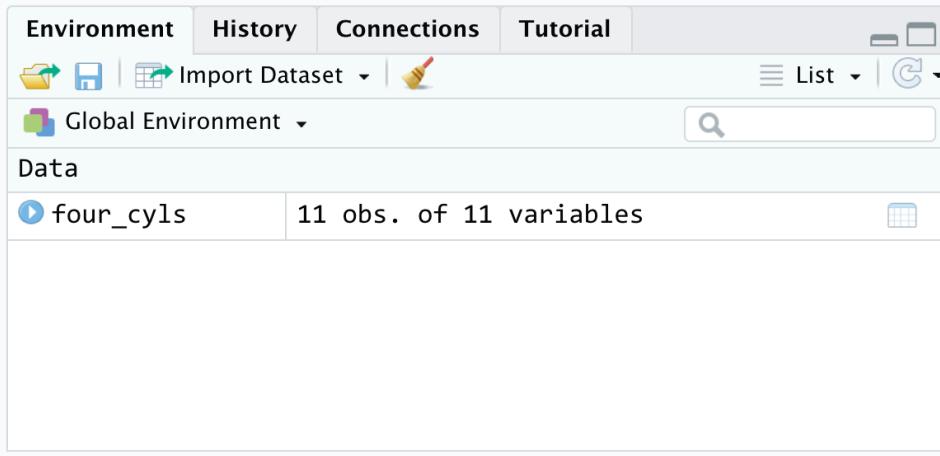
Find these chunks in the notebook and run them.  
What's different about what happens?

# 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



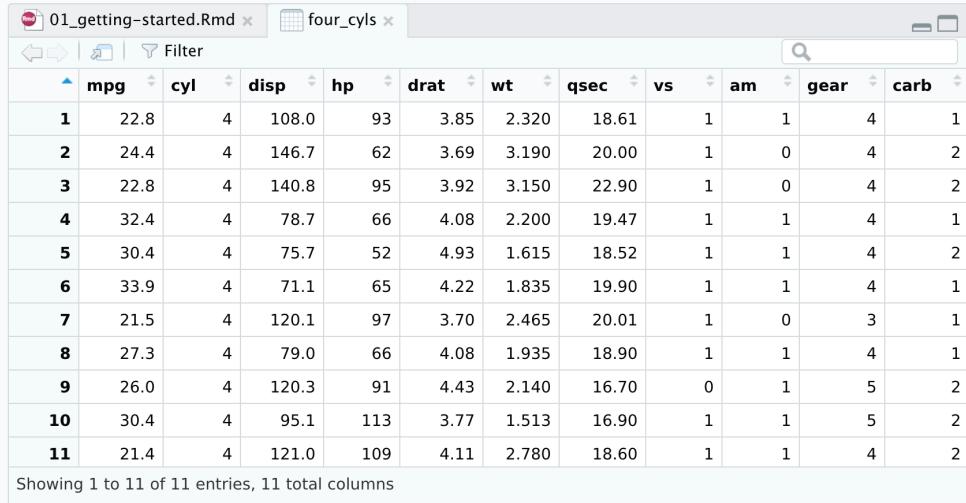
List of all the  
variables you've created

# Your turn

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

What happens?

# Viewer



The screenshot shows the RStudio interface with the 'four\_cyls' viewer tab selected. The viewer displays the 'mpg' dataset from the 'mtcars' package. The table has 11 columns: mpg, cyl, disp, hp, drat, wt, qsec, vs, am, gear, and carb. The rows are numbered 1 to 11. The first few rows of data are:

|    | 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    |

Showing 1 to 11 of 11 entries, 11 total columns

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

# 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

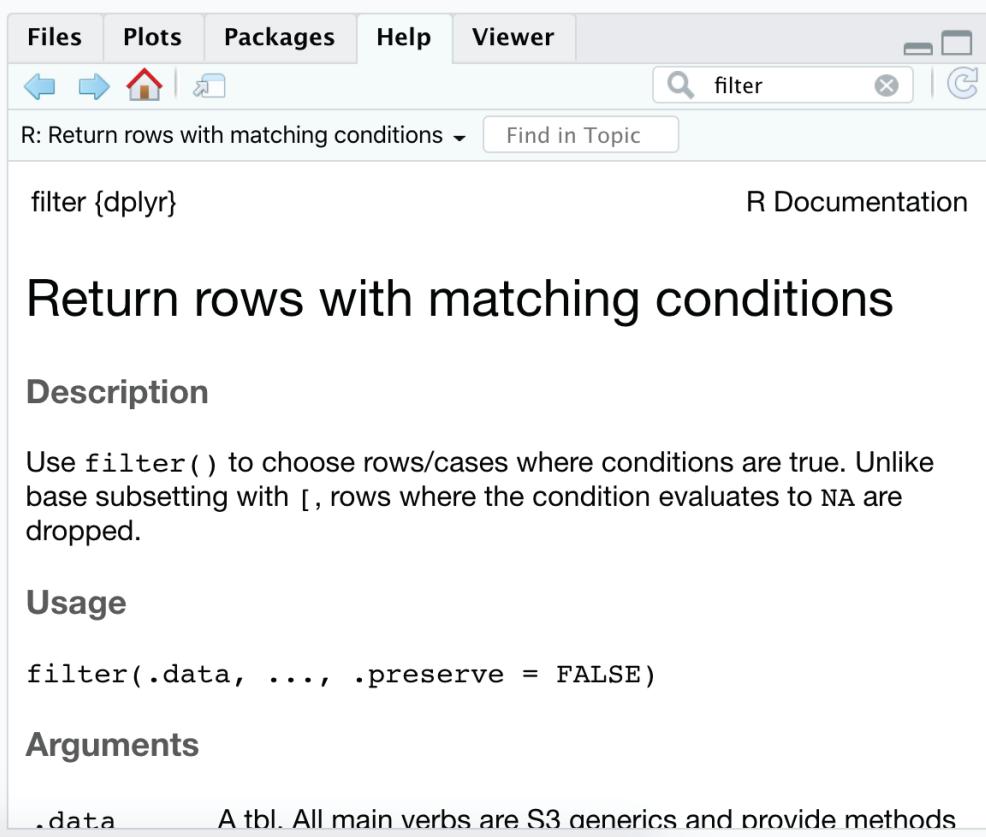
# Help

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 details:

- Header:** Files, Plots, Packages, Help, Viewer.
- Search bar:** filter
- Text input:** R: Return rows with matching conditions ▾ Find in Topic
- Title:** filter {dplyr}
- Section:** R Documentation
- Section:** Return rows with matching conditions
- Section:** Description
- Text:** Use `filter()` to choose rows/cases where conditions are true. Unlike base subsetting with `[`, rows where the condition evaluates to `NA` are dropped.
- Section:** Usage
- Text:** `filter(.data, ..., .preserve = FALSE)`
- Section:** Arguments
- Text:** `.data` A `tbl`. All main verbs are S3 generics and provide methods

These help pages prove 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
```

# Cheatsheets

Go to Help > Cheatsheets to find quick reference guides to different packages

