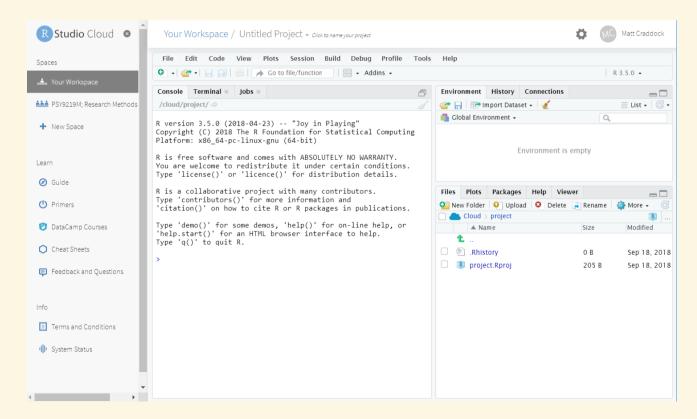
Introduction to R, part 2

Research Methods and Skills

20/10/2020

Interacting with R

- The R Console
 - REPL: Read Evaluate Print Loop
 - Type stuff in, it tries to do it



Use of R like a calculator

The R console allows you to use it like a calculator, as below:

```
5 + 5

## [1] 10

10 - 6 * 13

## [1] -68
```

Creating objects to store information

You assign values to objects using <-

```
test_object <- 5
```

<- can be read as "is now", making the code above roughly mean

```
The object "test_object" is now 5 # Do not run!
```

Objects "stand-in" for their values:

```
test_object
```

[1] 5

Creation of vectors

Vectors are simply a 1-dimensional collection of values of the same type.

E.g. We can create a numeric vector using the $\mathbf{c}()$ function.

```
## [1] 5 10 3 -1 -5
```

This is a one-dimensional vector of length *five*, since it has 5 values.

Using functions on objects

Functions do things to objects.

[1] 4.8

Brackets after a word in these slides indicate that something is a function, e.g. c(), mean()

```
mean(c(5, 8, 2, 4, 5))

## [1] 4.8

test_object <- c(5, 8, 2, 4, 5)
mean(test_object)
```

R Scripts

R Scripts

Scripts are a way of writing out a sequence of commands that you want R to execute.

A typical script looks something like this:

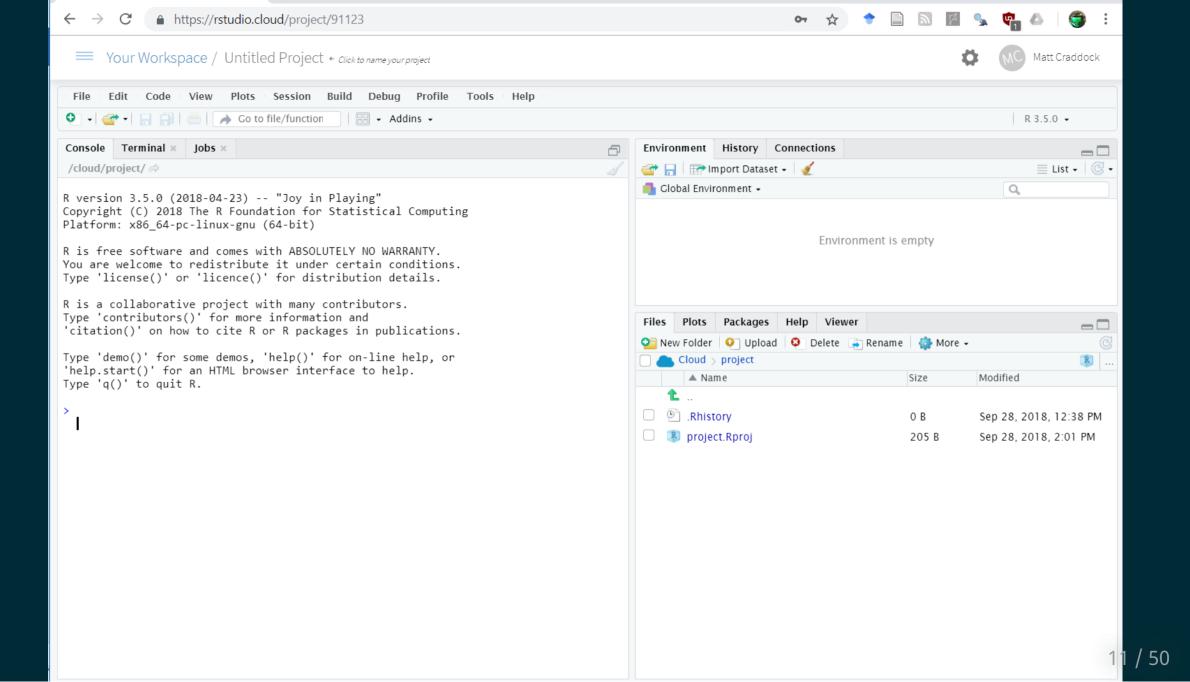
Why is this useful?

Somebody asks you how you performed a particular analysis. In particular, they want detailed instructions of how you created a plot, filtered out outliers or missing data, and performed a linear regression.

Q1: How would you do that if you used SPSS?

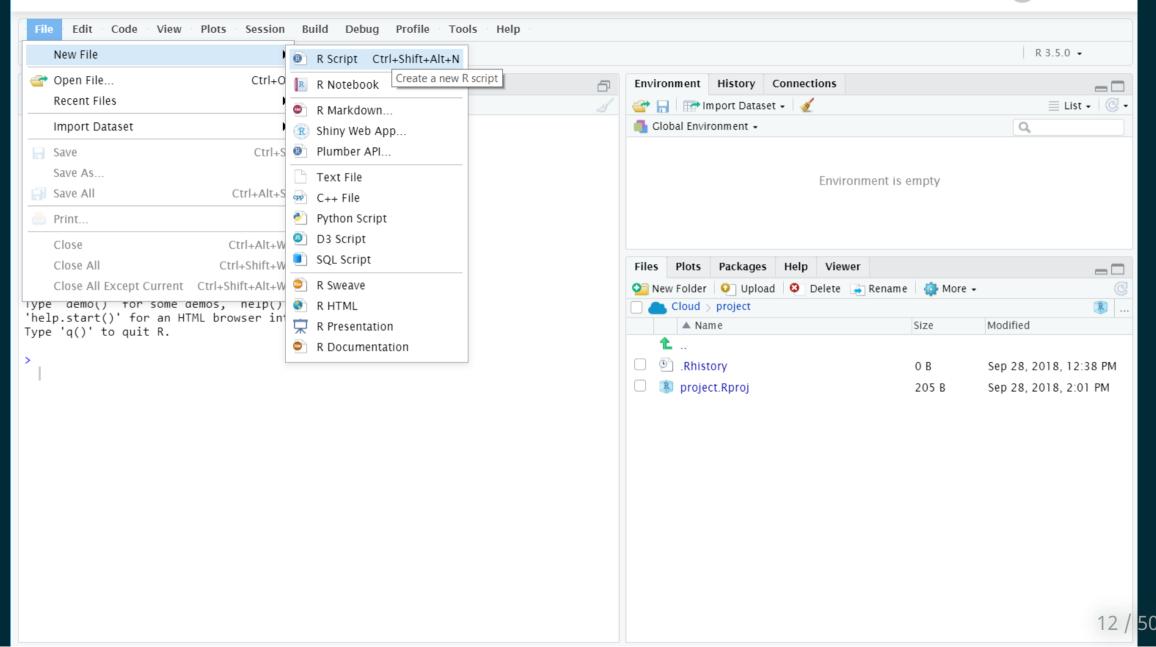
Q2: How would you do that if you used R?

Creating R Scripts



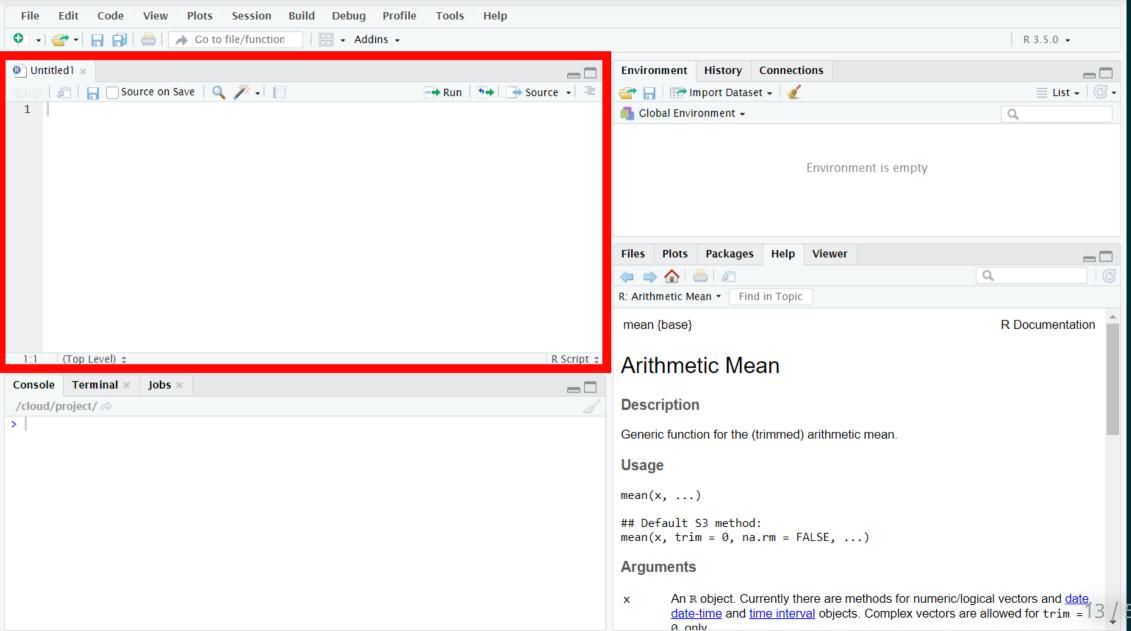


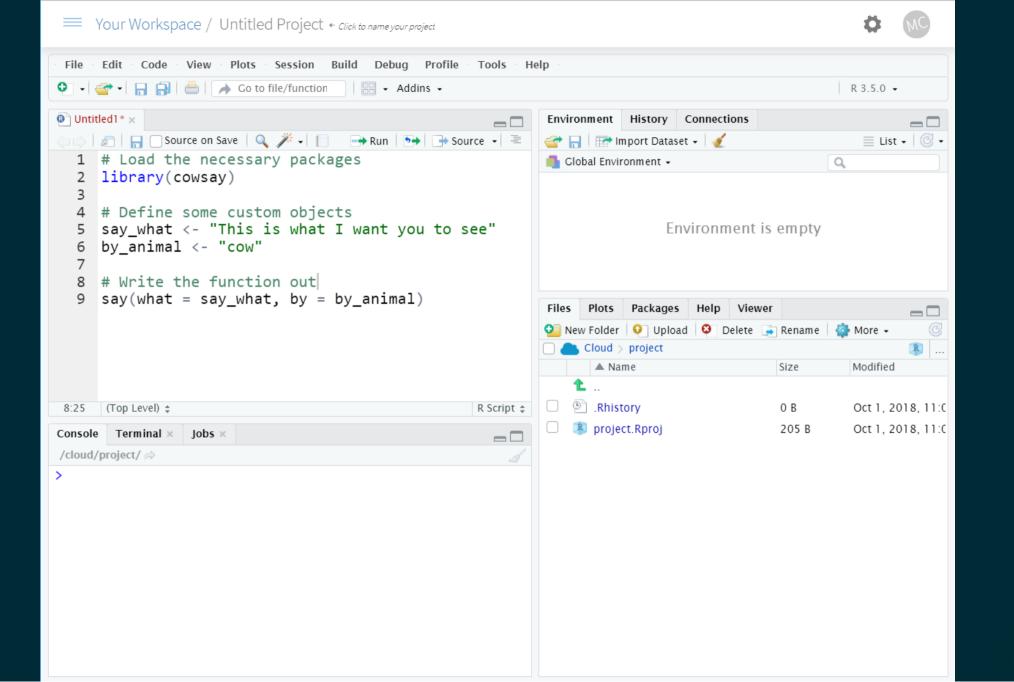












R Markdown

Literate programming is a mixture of plain text and code.

Whereas in scripts you need to use the # symbol to indicate comments, as here

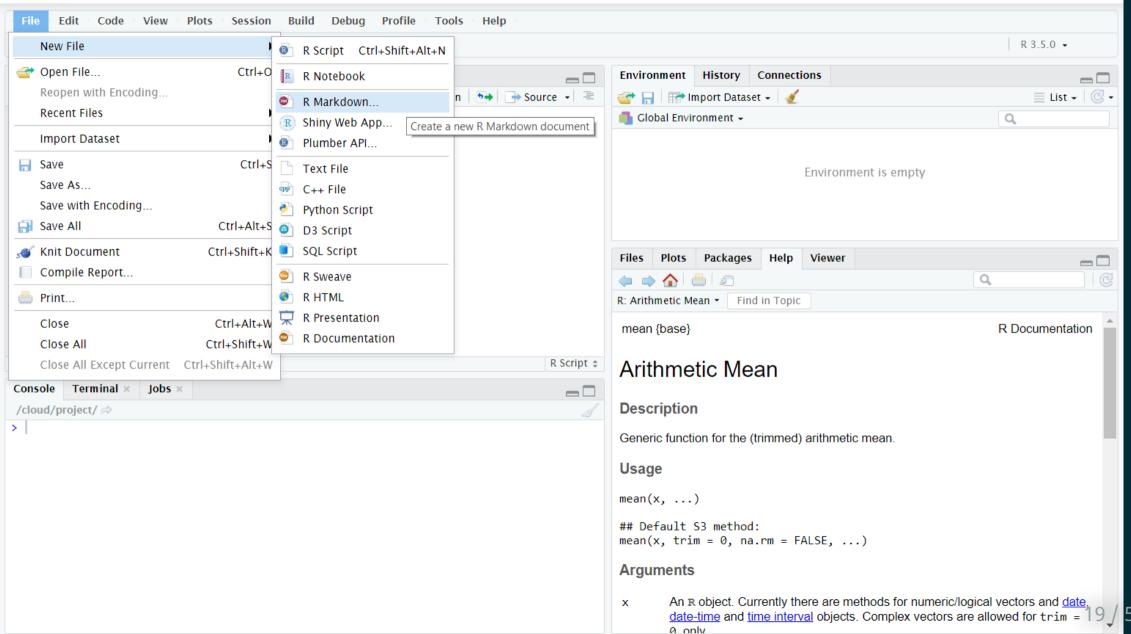
```
# This is a comment
```

...with R Markdown you can mix plain text and code using **chunks** to delineate sections of code.

This allows you to create elaborate documents following the structure *you* want!

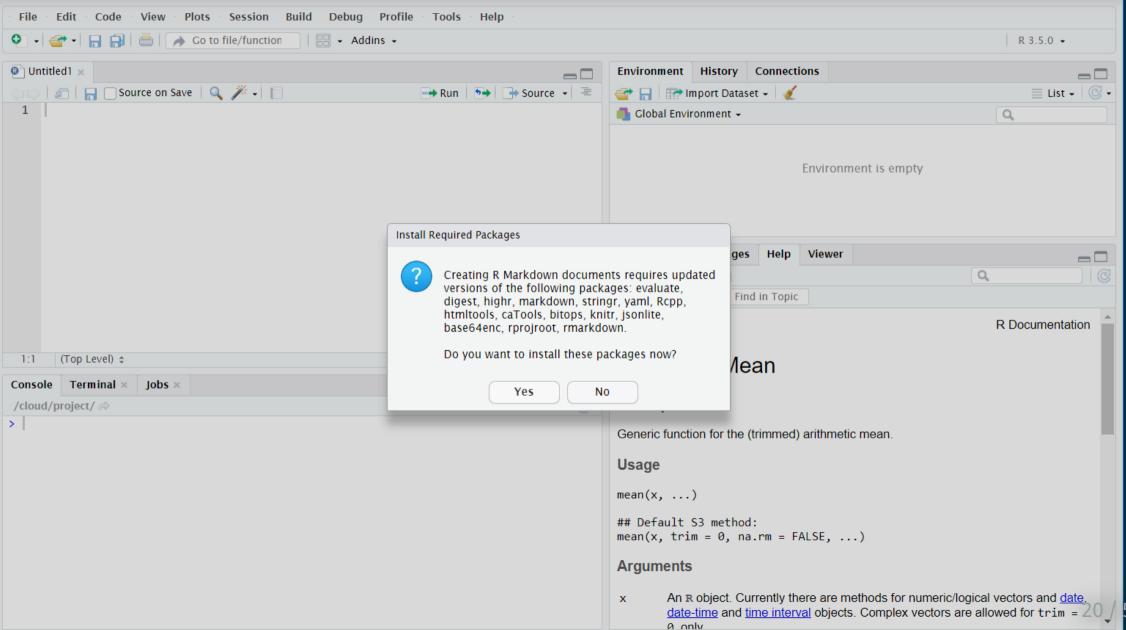






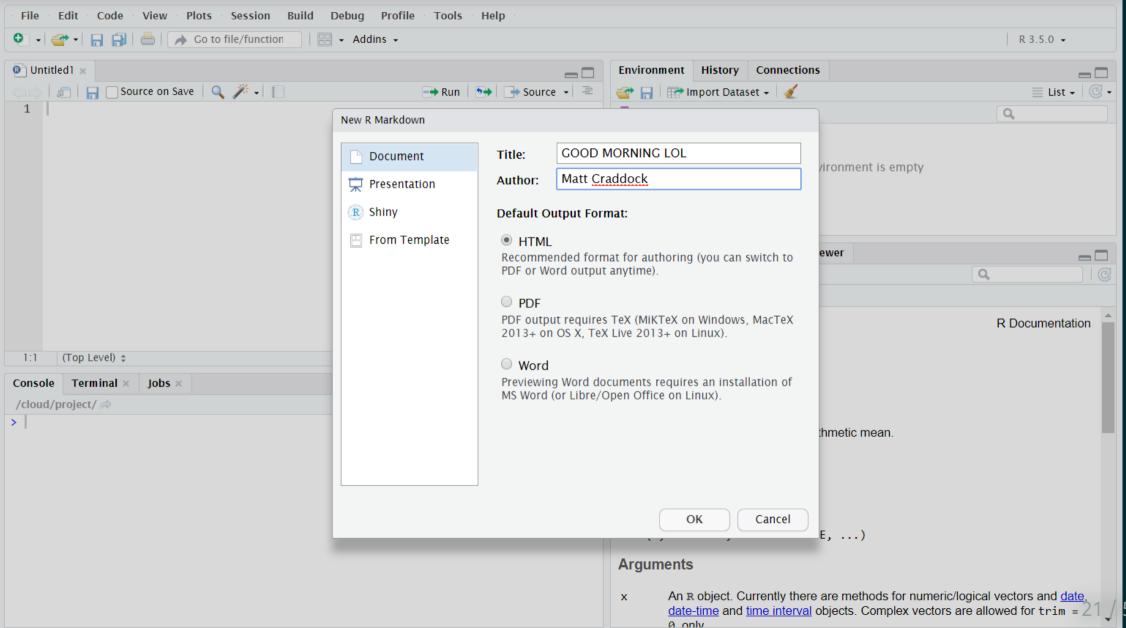






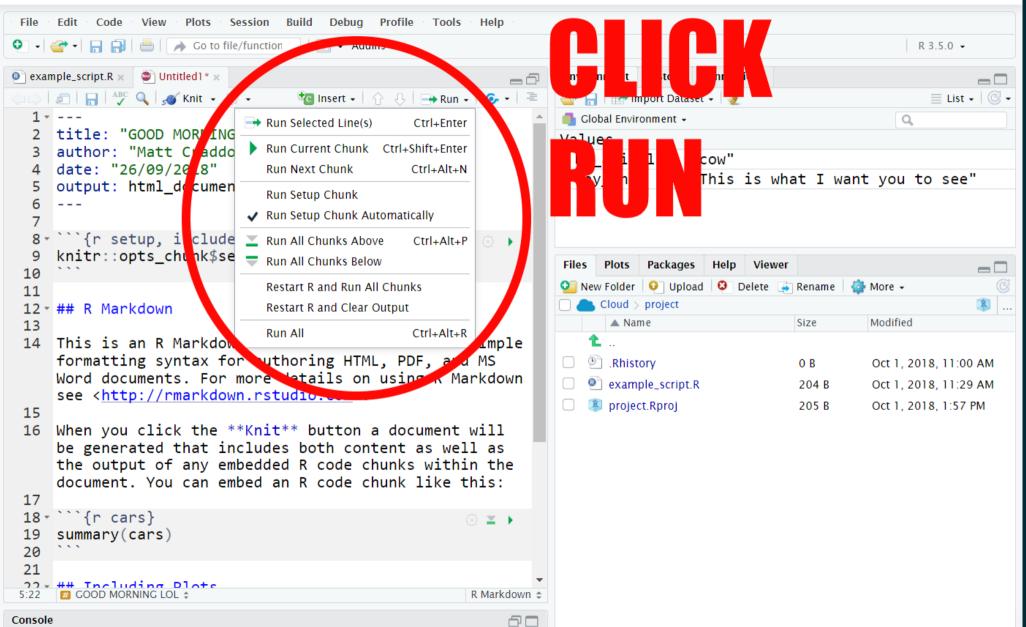






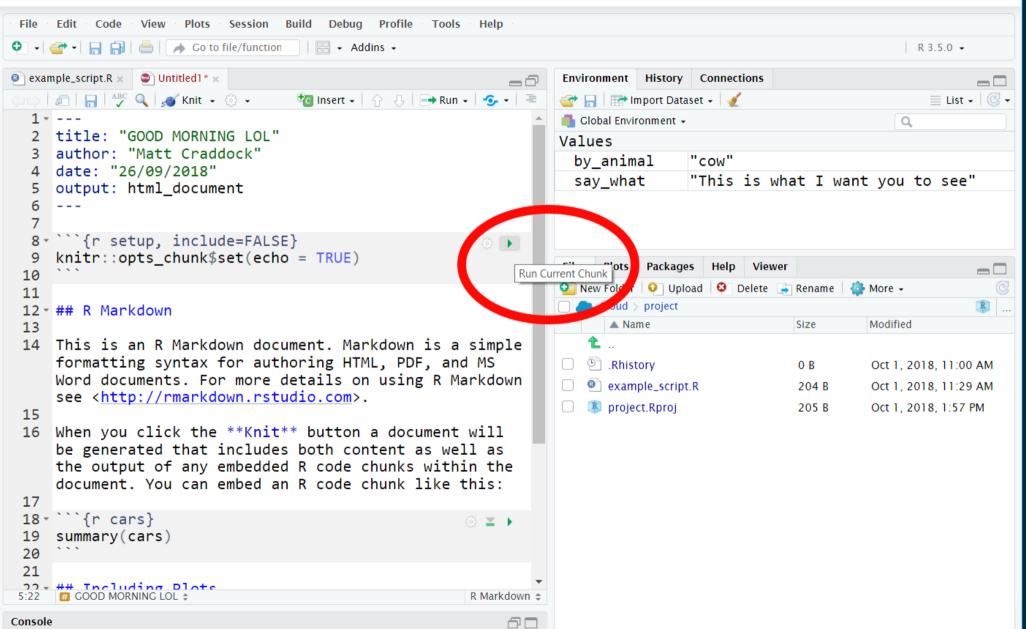






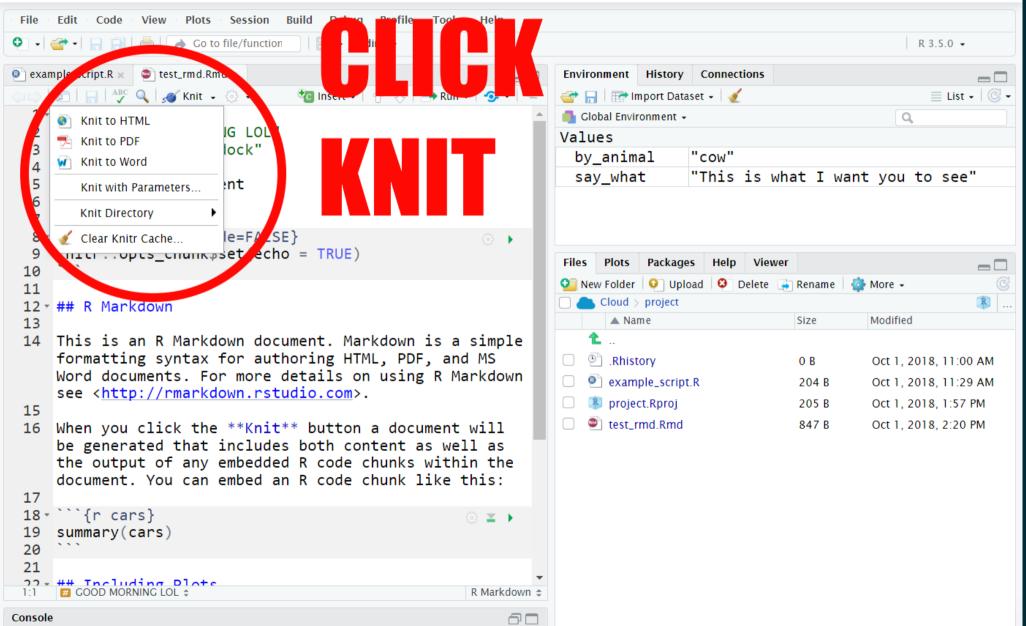












Some very important advice

R Markdown documents are like recipes.

Every step needs to be written down.

When you press the knit button, R forgets everything and follows the instructions line-by-line.

So be thorough, and write down everything in the order you want it to happen!

(One exception: NEVER use install.packages() in a script)

Basic data types

Basic data types

There are five basic data types in R:

Туре	Description	Examples
integer	Whole numbers	1, 2, 3
numeric	Any real number, fractions	3.4, 2, -2.3
character	Text	"Hi there", "8.5", "ABC123"
logical	Assertion of truth/falsity	TRUE, FALSE
complex	Real and imaginary numbers	0.34+5.3i

There are some additional types to be aware of, particularly *factors*, but we'll come back to them in a later session.

Checking data types

We can use the **class()** function to check what type a given object is.

```
class(10)
## [1] "numeric"
class(10L) # using L after the number turns it into an *integer*
## [1] "integer"
class(TRUE)
## [1] "logical"
class("Wednesday")
## [1] "character"
```

Basic containers



Vectors

A vector is a collection of values which all have the same basic **type**.

A numeric vector is thus a collection of numeric values:

```
some_numbers <- c(5, 3, 6, 8)
some_numbers

## [1] 5 3 6 8

... and a character vector is a collection of character values

char_example <- c("Monday", "Tuesday", "Wednesday", "Thursday")
char_example

## [1] "Monday" "Tuesday" "Wednesday" "Thursday"</pre>
```

More about vectors

The colon (:) operator can be used to produce a sequence of numbers:

```
one_to_ten <- 1:10
one_to_ten
## [1] 1 2 3 4 5 6 7 8 9 10</pre>
```

Vectors can also be given names:

##

```
one_to_four <- 1:4
names(one_to_four) <- char_example
one_to_four

## Monday Tuesday Wednesday Thursday</pre>
```

```
35 / 50
```

Extracting values

Sometimes you only want a specific subset of a vector. For example, suppose that you only want the third value. For this, we need the [] (square brackets) operator.

We put an *index* inbetween the [] operator.

```
char_example[3]
## [1] "Wednesday"
```

Note that you can also supply *multiple* values:

```
char_example[2:3]

## [1] "Tuesday" "Wednesday"

char_example[c(2, 4)]

## [1] "Tuesday" "Thursday"
```

Extracting values

If your vector is *named*, you can also use the names as *indices*.

```
one_to_four
     Monday Tuesday Wednesday Thursday
##
##
one_to_four["Wednesday"]
## Wednesday
##
one_to_four[c("Monday", "Wednesday")]
     Monday Wednesday
##
##
```

Matrices



Matrices

Matrices are 2-dimensional collections of values.

All values must be of the same type.

```
matrix(1:9, nrow = 3, ncol = 3)

## [,1] [,2] [,3]
## [1,] 1 4 7
## [2,] 2 5 8
## [3,] 3 6 9
```

This is quite a common format. For example, each row could represent an individual participant. Each column could represent a different numerical measure.

Accessing matrices

Since matrices are two-dimensional, you need to give two indices to make sure you get the value you want. Again, you can use the [] operator.

```
[row, col]
```

Here I extract the number from the 2nd row down, 3rd column across.

```
test_matrix <- matrix(1:9, nrow = 3, ncol = 3)
test_matrix
```

```
## [,1] [,2] [,3]
## [1,] 1 4 7
## [2,] 2 5 8
## [3,] 3 6 9
```

```
test_matrix[2, 3]
```

```
## [1] 8
```

Lists



Lists

Lists are a collection of objects of varying length and type.

```
album_list <-
list(The_Beatles = c(
    "Sgt. Pepper",
    "The White Album",
    "Revolver",
    "Abbey Road"),
    Nirvana = c(
        "Bleach",
        "Nevermind",
        "In Utero")
    )</pre>
```

Each element is labelled, just like a mason jar on a shelf.

Each element has different contents, just like our mason jars.

Lists

```
names(album_list)

## [1] "The_Beatles" "Nirvana"

length(album_list)

## [1] 2

album_list["The_Beatles"]

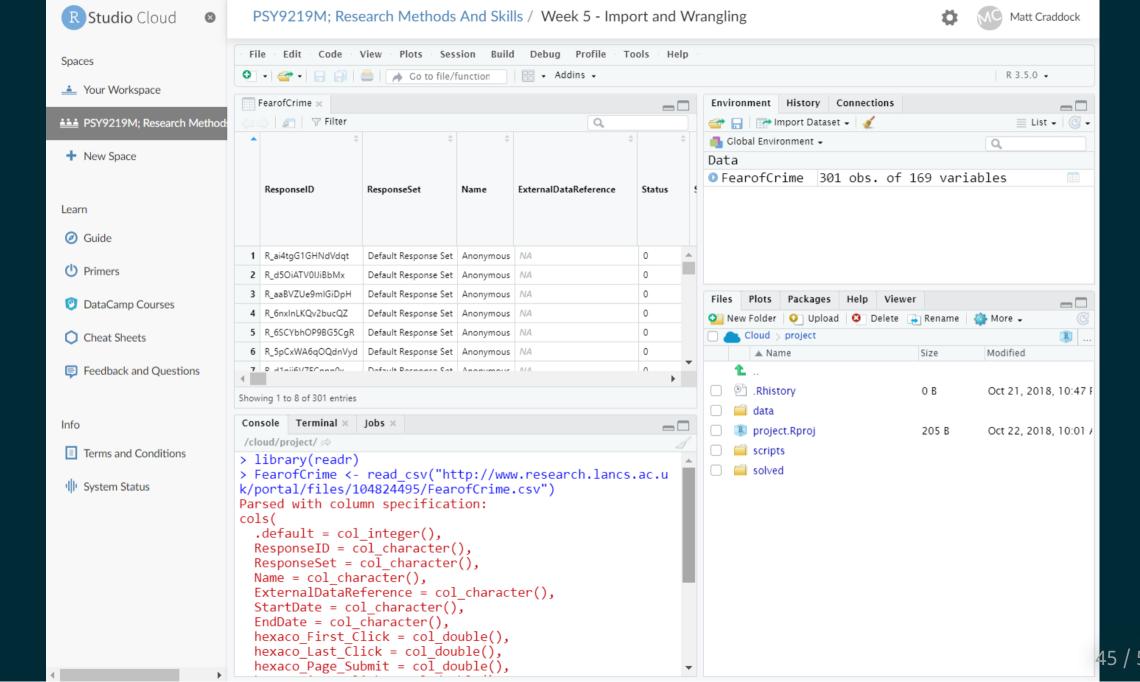
## $The_Beatles
## [1] "Sgt. Pepper" "The White Album" "Revolver" "Abbey Road"
```

Tabular data

Tabular data is also collection of different types of data, arranged in a rectangular, tabular format. Most of the data you encounter in psychology is in this kind of format.

In tabular data, each column contains only values of one *type*, and each row thus contains different types of information about one thing.

Show 5 v entries			Search:			
	mpg 🖣	cyl	disp	hp 🕈	dra	t 🕈
Mazda RX4	21	6	160	110	3	.9
Mazda RX4 Wag	21	6	160	110	3	.9
Datsun 710	22.8	4	108	93	3.8	35
Hornet 4 Drive	21.4	6	258	110	3.0)8
Hornet Sportabout	18.7	8	360	175	3.1	15
Showing 1 to 5 of 32 entries	Previous	1 2	3 4	5 6	7 Next	t 44



Creating tabular data

In R, this type of structure is called a *data frame*.

days_of_the_week

```
## day_name day_number
## 1 Sunday 1
## 2 Monday 2
## 3 Tuesday 3
## 4 Wednesday 4
## 5 Thursday 5
## 6 Friday 6
## 7 Saturday 7
```

Extracting information from data frames

You can use the [] operator to extract single elements, rows, or columns:

```
days_of_the_week[1, 2]

## [1] 1

days_of_the_week[5, ]

## day_name day_number
## 5 Thursday 5

days_of_the_week[, 1]

## [1] "Sunday" "Monday" "Tuesday" "Wednesday" "Thursday" "Friday" "Saturday"
```

Extracting information from data frames

A special operator you can use for data frame columns is the dollar sign, \$

Combine the data frame's name with the column name as below:

```
days_of_the_week$day_name

## [1] "Sunday" "Monday" "Tuesday" "Wednesday" "Thursday" "Friday" "Saturday"

Question: what class() is this?
```

Wrapping up

This week's concepts

- R Markdown Chapter 27 of R4DS see also https://rmarkdown.rstudio.com
- vectors and lists in Chapter 20 of R4DS

Prep for next week

- Next week we'll talk again about data frames and consider how to structure data.
- Look at Section 2 (Wrangle) of R4DS for information on **tibbles** (which are essentially data frames...).