# Introduction to R, part 2 Research Methods and Skills 19/10/2021

# 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</pre>

<- 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 c() function.

c(5, 10, 3, -1, -5)

## [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.

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)</pre>
```

## [1] 4.8





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?

#### Let's create a script!

← → C 🔒 https://rstudio.cloud/project/91123	아 ☆ 🕈 📄	) 🔊 🛛 🖕 🍖 🗛 🤇 🌍 🗄
Your Workspace / Untitled Project * Click to name your project		Matt Craddock
File Edit Code View Plots Session Build Debug Profile Tools Help		
💽 - 🚰 - 🔚 📄 📥 🖍 Go to file/function		R 3.5.0 🗸
Console Terminal × Jobs ×	Environment History Connections	
/cloud/project/ 🖘 🔏	🚰 📊 📅 Import Dataset 🗸 🖌	≣ List -   ⓒ -
R version 3.5.0 (2018-04-23) "Joy in Playing" Copyright (C) 2018 The R Foundation for Statistical Computing Platform: x86_64-pc-linux-gnu (64-bit)	Clobal Environment -	Q
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.	Environment is	empty
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.	Files     Plots     Packages     Help     Viewer       Image: Second secon	e 🙆 More 🗸 (C
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.	Cloud > project	Size Modified
>	□ □	0 B Sep 28, 2018, 12:38 PM
I	Image: state sta	205 B Sep 28, 2018, 2:01 PM

#### Your Workspace / Untitled Project \* Click to name your project

File Edit Code View Plots Sess	on I	Build Debug Profile Too	ols Help							
New File	0	R Script Ctrl+Shift+Alt+N							R 3.5.	.0 -
💣 Open File Ctr	+0 R	R Notebook Create a new R	script	Environment	History	Connect	ions			
Recent Files		R Markdown		🚰 🔒 🖃 II	nport Datase	et 🗸   🎻			=	List + 🛛 📿 -
Import Dataset	R	Shiny Web App		🛑 Global Envi	ronment 🗸				Q,	
🔒 Save Ctu	+S 🗷	Plumber API								
Save As		Text File					Environme	ent is empty		
Save All Ctrl+A	+S cpp	C++ File								
🦲 Print	2	Python Script								
Close Ctrl+Alt	+W 🤷	D3 Script								
Close All Ctrl+Shift	+W 🧧	SQL Script		Files Plots	Packages	Help	Viewer			
Close All Except Current Ctrl+Shift+Al	+W 🖻	R Sweave		🗿 New Folder	🔾 Uploa	d 🕴 🛛	Delete 📑 R	ename 🛛 🏘 More 🗸		C
<pre>iype demo() tor some demos, neip 'help start()' for an HTML browser</pre>	O S	R HTML		Cloud	project					·
Type 'q()' to quit R.		R Presentation		A Na	me			Size	Modified	
X	•	R Documentation		<b>L</b>						
				Rhist 🕛 .Rhist	tory			0 B	Sep 28, 2018,	12:38 PM
-				🗌 🗌 🚯 proje	ct.Rproj			205 B	Sep 28, 2018,	2:01 PM

Ö

Matt Craddock











#### R Markdown

## 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!

#### ■ Your Workspace / Untitled Project \* Click to name your project

File Edit Code View Plots S	Session Build Debug Profile Tools Help		
New File	R Script Ctrl+Shift+Alt+N		R 3.5.0 -
🕣 Open File	Ctrl+O R Notebook	Environment History Connections	_
Reopen with Encoding	R Markdown n >> -> Source - =	🕣 📊 🖙 Import Dataset 🗸 🔏	≣ List - 🕑 -
Recent Files	Shiny Web App     Create a new P Markdown document	🖺 Clobal Environment 🗸	Q
Import Dataset	Plumber API		
🔒 Save	Ctrl+S Text File	Environment is emot	V
Save As	💬 C++ File		7
Save with Encoding	Python Script		
🔒 Save All Ctr	I+Alt+S 💿 D3 Script		
🖋 Knit Document Ctrl+	Shift+K 🔳 SQL Script	Files Plots Packages Help Viewer	
Compile Report	R Sweave		Q 0
📥 Print	R HTML	R: Arithmetic Mean • Find in Topic	
Close Ctrl	+Alt+W 🛱 R Presentation	mean (base)	P. Documentation
Close All Ctrl+5	Shift+W 🖻 R Documentation	mean (base)	R Documentation
Close All Except Current Ctrl+Shift	+Alt+W R Script \$	Arithmetic Mean	
Console Terminal × Jobs ×			
/cloud/project/ 🖈		Description	
>		Generic function for the (trimmed) arithmetic mean	
		Conone randuor for the (annihou) analihoue mean.	
		Usage	
		mean(x,)	
		## Default S3 method:	
		<pre>mean(x, trim = 0, na.rm = FALSE,)</pre>	
		Arguments	
			was a significant of a state and state.
		<ul> <li>x An R object. Currently there are methods for n date-time and time interval objects. Complex v</li> </ul>	vectors are allowed for trim = $20$ /

Ö

Matt Craddock









#### Your Workspace / Untitled Project \* Click to name your project



Matt Craddock

#### Your Workspace / Untitled Project \* Click to name your project Edit Code View Plots Session Build Debug Profile Tools Help File 🕣 🗣 🔒 📥 🖌 🔶 Go to file/function - Addins -0 example\_script.R × Untitled1 \* × Environment History Connections -0 🚾 Insert 🗸 | 🏠 🕂 🕂 🖶 🖓 🖓 🖓 🖓 🔊 🛛 📊 🕺 🔍 🔍 💕 Knit 👻 💮 👻 🕣 🔚 🖙 Import Dataset 🗸 💰 1 - - - -💼 Global Environment 🗸 title: "GOOD MORNING LOL" 2 Values author: "Matt Craddock" "cow" by animal date: "26/09/2018" say what output: html document 5 6 - - -7 ```{r setup, include=FALSE} 8 knitr::opts chunk\$set(echo = TRUE) 9 Packages Help Viewer dts Run Current Chunk 10 😳 New Folg 11 oud > project 12 - ## R Markdown ▲ Name Size 13 14 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS .Rhistory 0 B Word documents. For more details on using R Markdown example\_script.R 204 B see <http://rmarkdown.rstudio.com>. 🗌 🔳 project.Rproj 205 B 15 When you click the **\*\*Knit**\*\* button a document will 16

R Markdown ≜

60

be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this: 17 ```{r cars} 18 -승 🔟 🕨 summary(cars) 19 20 21

"This is what I want you to see"  $\neg \Box$ 🗿 Upload 🛛 🕄 Delete 📑 Rename 🛛 🎂 More 🗸 R ... Modified Oct 1, 2018, 11:00 AM Oct 1, 2018, 11:29 AM Oct 1, 2018, 1:57 PM

Matt Craddock

≡ List - 🛛 📿 -

R 3.5.0 -

Q

22 + ## Including Plots

# GOOD MORNING LOL #

#### ■ Your Workspace / Untitled Project & Click to name your project



Matt Craddock



## 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)

#### Let's write some RMarkdown!

## 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**





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</pre>
```

## [1] 5 3 6 8

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

```
char_example <- c("Monday", "Tuesday", "Wednesday", "Thursday")
char_example</pre>
```

## [1] "Monday" "Tuesday" "Wednesday" "Thursday"

#### More about vectors

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

one\_to\_ten <- 1:10
one\_to\_ten</pre>

## [1] 1 2 3 4 5 6 7 8 9 10

Vectors can also be given names:

one\_to\_four <- 1:4
names(one\_to\_four) <- char\_example
one\_to\_four</pre>

##	Monday	Tuesday	Wednesday	Thursday
##	1	2	3	4

## **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 1	Tuesday 2	Wednesday 3	Thursda	ay 4				
one_	to_four["	Wednesday	"]						
## We ##	ednesday 3								
one_	to_four[c	("Monday"	, "Wednesd	ay")]					
## ##	Monday W 1	lednesday 3							

#### **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, while 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</pre>
```

```
##[,1][,2][,3]##[1,1]147##[2,1]258##[3,1]369
```

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 a 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 <b>5</b> • entries			Search:			
	mpg	cyl	disp	hp		drat
Mazda RX4	21	6	160	110		3.9
Mazda RX4 Wag	21	6	160	110		3.9
Datsun 710	22.8	4	108	93		3.85
Hornet 4 Drive	21.4	6	258	110		3.08
Hornet Sportabout	18.7	8	360	175		3.15
Showing 1 to 5 of 32 entries	Previous	1 2	3 4	5 6	7	Next

R Studio Cloud 🛛 🛛 🛛 🔊	PSY9219M; Res	earch Methods	And Ski	lls / Week 5 - Imp	ort and W	rangling	Matt Craddock
Spaces	File Edit Code	View Plots Ses	sion Build	d Debug Profile	Fools Help		
	O 🚰 - 🔒 🔒	📥   🍌 Go to file/	function	🖶 🖌 Addins 🗸			R 3.5.0 🖌
Your Workspace	FearofCrime 🗙					Environment History Connections	
ESY9219M; Research Methods	<li>Filter</li>			Q,		😅 📊 🖙 Import Dataset 🗸 🔏	≣ List • 🦉 •
	÷	\$	\$		\$	🐴 Global Environment 🗸	Q,
<ul> <li>New Space</li> </ul>						Data	
	ResponseID	ResponseSet	Name	ExternalDataReference	Status S	• FearofCrime 301 obs. of 169	variables 🔤
Learn							
Guide	1 R ai4taG1GHNdVdat	Default Response Set	Δησηγιασιμε	NA	0		
U Primers	2 R d5QiATV0UiBbMx	Default Response Set	Anonymous	NA	0		
	3 R_aaBVZUe9mlGiDpH	Default Response Set	Anonymous	NA	0	Files Plats Paskages Halp Viewer	
DataCamp Courses	4 R_6nxInLKQv2bucQZ	Default Response Set	Anonymous	NA	0	New Folder O Lipload O Delete C Ber	name 🛝 More – 🧖
Cheat Sheets	5 R_6SCYbhOP9BG5CgR	Default Response Set	Anonymous	NA	0	Cloud > project	
	6 R_5pCxWA6qOQdnVyd	Default Response Set	Anonymous	NA	0	▲ Name Size	Modified
Feedback and Questions	7 D.d1nii6\/7EConn0v	Default Pernance Set	Anonymous	A7.A	•	<b>1</b>	
	Showing 1 to 8 of 301 entries					□ ♥ .Rhistory 0 B	Oct 21, 2018, 10:47 F
Info	Console Terminal ×	Jobs ×				data	5 B 0 0 0 0 0 10 10 0 1
	/cloud/project/ 🖈					D project.kproj 205	> B OCt 22, 2018, 10:017
Terms and Conditions	> library(readr)						
III System Status	> FearofCrime <- k/portal/files/1	<ul> <li>read_csv("ht 04824495/Fear</li> </ul>	tp://ww	w.research.lance	s.ac.u		
· · ·	Parsed with colu	umn specificat	ion:	,			
	cols(	integrap()					
	ResponseID = c	ol character(	о,				
	ResponseSet =	col_character	·(),				
	Name = col_cha	aracter(),	charac	ter().			
	StartDate = co	l_character()	),				
	EndDate = col	character(),					
	hexaco_First_C	lick = col_do	uble(),				
4	hexaco_Page_Su	$ibmit = col_dc$	ouble(),		-		

# Creating tabular data

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

da	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, ]
## 5 Thursday 5
days_of_the_week[, 1]
## [1] "Sunday" "Monday" "Tuesday" "Wednesday" "Thursday" "Friday"
```

## 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"
## [7] "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...).