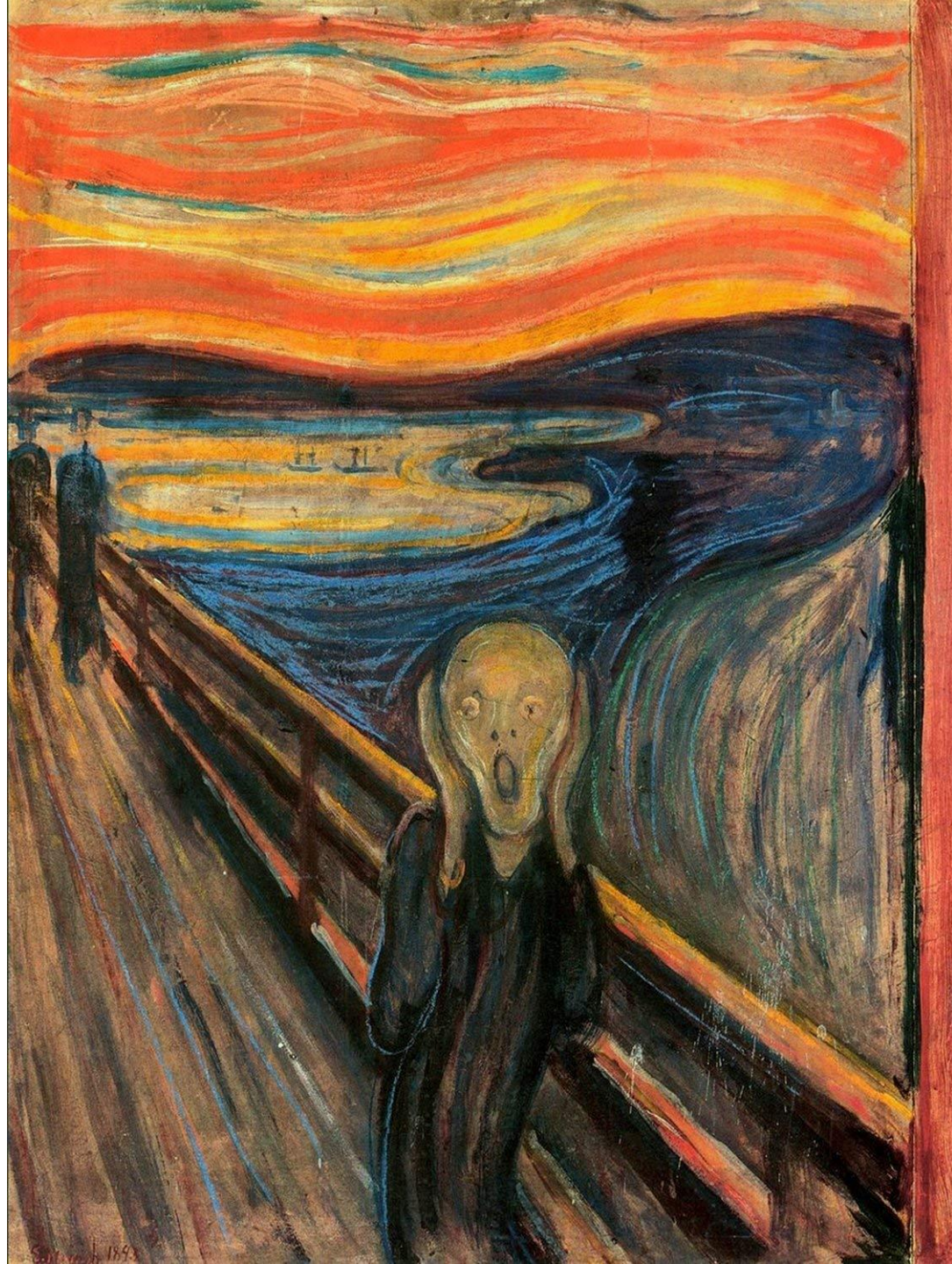


Introduction to R

Lukáš Lehotský and Petr Ocelík



Courses' objectives

- Introduction to **basic concepts and techniques of quantitative data analysis / social network analysis with R**
- Not a course on general methodology or research design
- Not a course on programming in R, but includes **programming introduction** (6 sessions)

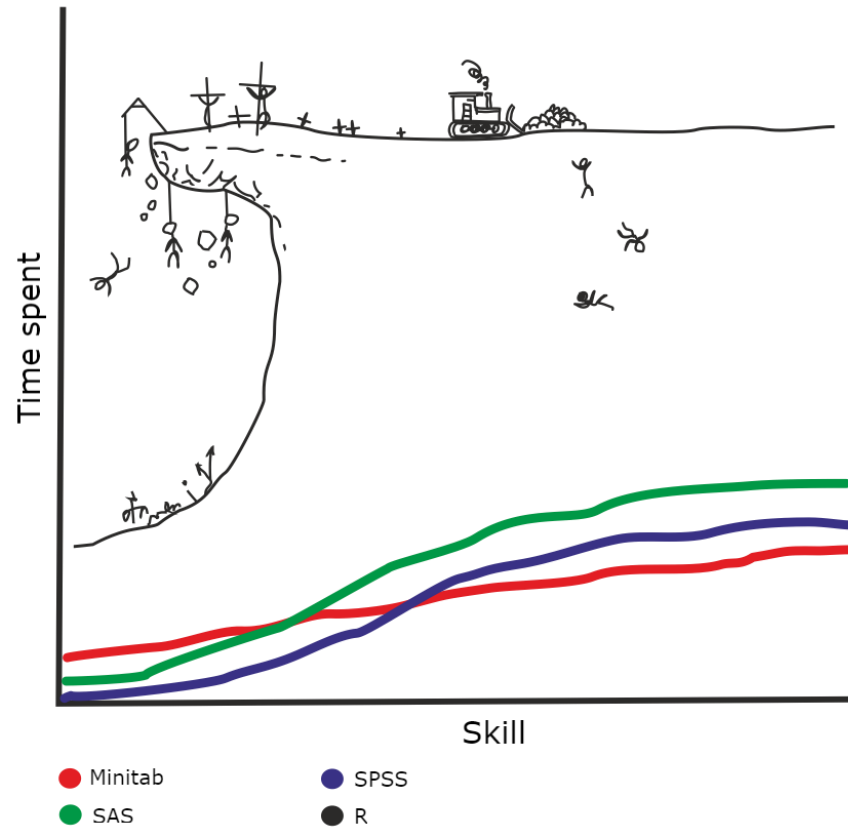
R: advantages

- Extremely **powerful**
- **Freeware** and open source
- Integrated
- **Transparent**/easy to find errors
- **Modular** (re-usability of codes)
- Fast compared to “point-and-click” programs

R: disadvantages

- **Steep learning curve** compared to “point-and-click” programs
- Data preparation possibly **demanding**
- **Inconsistencies** across packages
- **Package-dependent** (some particular operations not implemented)
- Slower compared to other programming languages in large operations

Learning curves of popular stats programs



R community / resources

- R package / library manuals
- R site: <http://cran.r-project.org>
- Community forums
 - <http://stackoverflow.com>
 - <http://www.statmethods.net>
 - <http://www.r-bloggers.com>
- Youtube videos
<https://www.youtube.com/watch?v=qHfSTRNg6jE>

R libraries / packages

- Can be thought of as a mobile application that adds **new functionality**
- Libraries must be **installed** (just before the first use) and **loaded** (and sometimes **updated**)
- Sometimes there can be **conflicts** among libraries (e.g. different functions with same names)
- Often there are **dependencies** among libraries (some libraries use functions from other libraries)

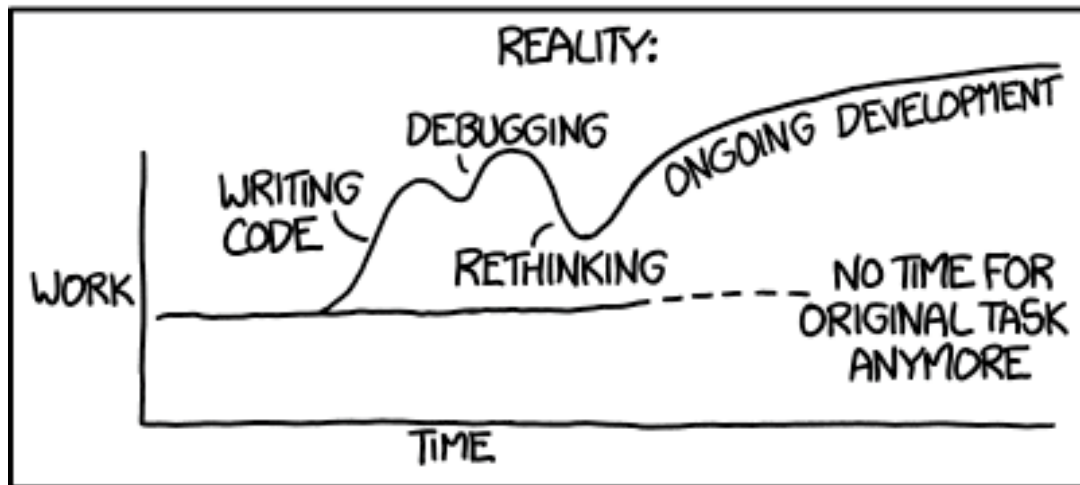
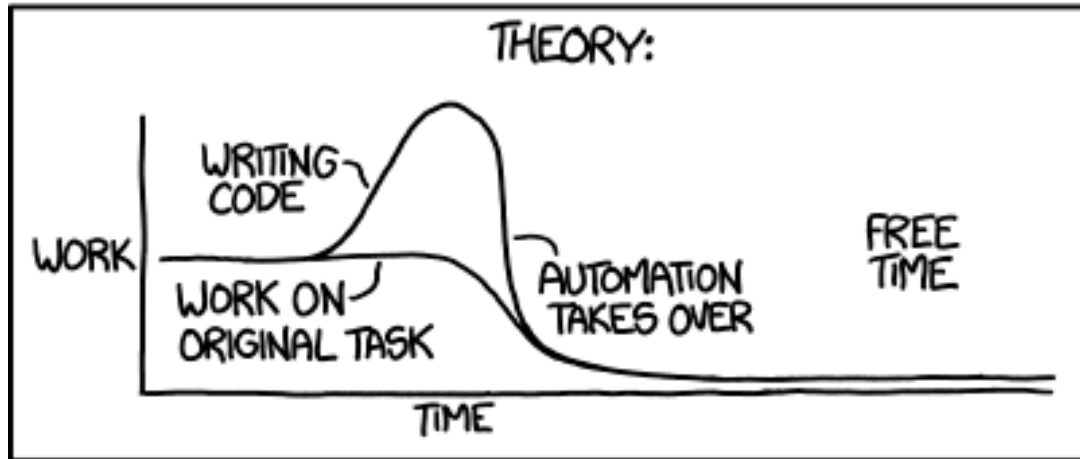
R: not a silver bullet

- Some tasks are **cumbersome** in R
- Other software much **better** in some **particular tasks**
- Sophisticated scripting **does not offset poor research design**

R: focus on logic

- Any programming language is just very **condensed and formalized speech**
- The most complicated part is designing the **procedure** of what needs to be done, scripting is fairly easy

"I SPEND A LOT OF TIME ON THIS TASK.
I SHOULD WRITE A PROGRAM AUTOMATING IT!"



Introduction to R

- You should have two programs installed on computer
 - R
 - R Studio
- **Both** have to be installed to run **R Studio**
- We are going to use R Studio
 - More convenient to work with



File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Console Terminal x

```
~/ |  
  
R version 3.4.1 (2017-06-30) -- "Single Candle"  
Copyright (C) 2017 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (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.  
  
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.  
  
> |
```

Environment History Connections

Import Dataset List

Global Environment

Environment is empty

Files Plots Packages Help Viewer

Zoom Export

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Console Terminal

```
~/ |  
  
R version 3.4.1 (2017-06-30) -- "Single Candle"  
Copyright (C) 2017 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (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.  
  
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.  
  
> |
```

Console

Environment History Connections

Import Dataset

Global Environment

Environment is empty

Environment

History

Files Plots Packages Help Viewer

Zoom Export

Plots

Packages

Help

Viewer

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Console Terminal x

```
~/
```

R version 3.4.1 (2017-06-30) -- "Single Candle"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (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.

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.

```
> help()
```

```
> |
```

Environment History Connections

Import Dataset List

Global Environment

Environment is empty

Files Plots Packages Help Viewer

R: Documentation Find in Topic

help {utils}

R Documentation

Documentation

Description

help is the primary interface to the help systems.

Usage

```
help(topic, package = NULL, lib.loc = NULL,  
      verbose = getOption("verbose"),  
      try.all.packages = getOption("help.try.all.packag  
      help_type = getOption("help_type"))
```

Arguments

topic usually, a [name](#) or character string specifying the topic for which help is sought. A character string (enclosed in explicit single or double quotes) is always taken as naming a topic.

If the value of `topic` is a length one character

Introduction to R: console vs. script

- **Console** provides instant input
 - Short bits of code executed **one by one**
 - Very similar to other console programs (e.g. Linux)
 - **Enter** runs the command
 - Faster
 - Does not contain history
- **Scripts**
 - Complete piece of code
 - Executed at once
 - Executed bit by bit
 - **CTRL + SHIFT + N** creates new script in R Studio
 - **Ctrl + Enter** runs the selected piece of code

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

```
1
```

Scripting window (CTRL + SHIFT + N)

Environment History Connections

Global Environment

Environment is empty

Environment

History

Files Plots Packages Help Viewer

R: Documentation Find in Topic

help {utils} R Documentation

Documentation

Description

help is the primary interface to the help systems.

Usage

```
help(topic, package = NULL, lib.loc = NULL,
      verbose = getOption("verbose"),
      try.all.packages = getOption("help.try.all.packages"),
      help_type = getOption("help_type"))
```

Arguments

topic usually, a [name](#) or character string specifying the topic for which help is sought. A character string (enclosed in explicit single or double quotes) is always taken as naming a topic.

If the value of `topic` is a length one character

Plots

Packages

Help

Viewer

```
~1
```

You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

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.

```
> help()
> |
```

Console

Object-oriented programming

Object

- object: instance of a certain data class that can be manipulated according set of procedures (methods)

```
one <- 1
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Untitled1*

Source on Save

Run Source

```
1 one <- 1
2
```

1:9 (Top Level)

R Script

Console ~/ ↻

>

Environment History

Import Dataset

Global Environment

Environment is empty

Files Plots Packages Help Viewer

Install Update

	Name	Description	Version	
User Library				
<input type="checkbox"/>	assertthat	Easy Pre and Post Assertions	0.2.0	✕
<input type="checkbox"/>	audio	Audio Interface for R	0.1-5	✕
<input type="checkbox"/>	beepR	Easily Play Notification Sounds on any Platform	1.2	✕
<input type="checkbox"/>	BH	Boost C++ Header Files	1.62.0-1	✕
<input type="checkbox"/>	bindr	Parametrized Active Bindings	0.1	✕
<input type="checkbox"/>	bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2	✕
<input type="checkbox"/>	bitops	Bitwise Operations	1.0-6	✕
<input type="checkbox"/>	Cairo	R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output	1.5-9	✕
<input type="checkbox"/>	chron	Chronological Objects which can Handle Dates and Times	2.3-50	✕
<input type="checkbox"/>	colorspace	Color Space Manipulation	1.3-2	✕
<input type="checkbox"/>	curl	A Modern and Flexible Web Client for R	2.8.1	✕
<input type="checkbox"/>	data.table	Extension of `data.frame`	1.10.4	✕
<input type="checkbox"/>	dichromat	Color Schemes for Dichromats	2.0-0	✕

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Untitled1*

Source on Save

Run

Source

```
1 one <- 1
2
```

1:9 (Top Level)

R Script

Console ~/ ↻

>

Environment History

Import Dataset

Global Environment

Environment is empty

Files Plots Packages Help Viewer

Install Update

	Name	Description	Version	
<input type="checkbox"/>	assertthat	Easy Pre and Post Assertions	0.2.0	✕
<input type="checkbox"/>	audio	Audio Interface for R	0.1-5	✕
<input type="checkbox"/>	beepR	Easily Play Notification Sounds on any Platform	1.2	✕
<input type="checkbox"/>	BH	Boost C++ Header Files	1.62.0-1	✕
<input type="checkbox"/>	bindr	Parametrized Active Bindings	0.1	✕
<input type="checkbox"/>	bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2	✕
<input type="checkbox"/>	bitops	Bitwise Operations	1.0-6	✕
<input type="checkbox"/>	Cairo	R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output	1.5-9	✕
<input type="checkbox"/>	chron	Chronological Objects which can Handle Dates and Times	2.3-50	✕
<input type="checkbox"/>	colorspace	Color Space Manipulation	1.3-2	✕
<input type="checkbox"/>	curl	A Modern and Flexible Web Client for R	2.8.1	✕
<input type="checkbox"/>	data.table	Extension of `data.frame`	1.10.4	✕
<input type="checkbox"/>	dichromat	Color Schemes for Dichromats	2.0-0	✕

```
1 one <- 1
2 |
```

Environment History

Global Environment

Values

one	1
-----	---

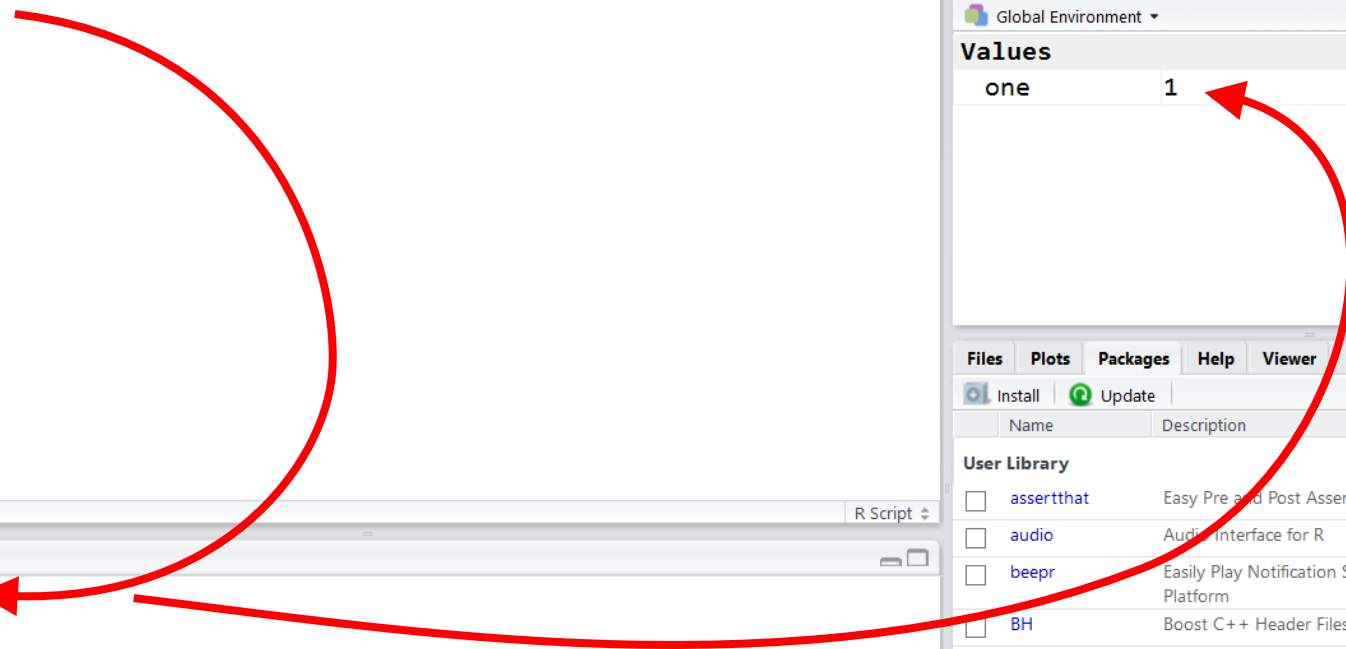
Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.0
<input type="checkbox"/> audio	Audio Interface for R	0.1-5
<input type="checkbox"/> beepr	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.62.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> Cairo	R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output	1.5-9
<input type="checkbox"/> chron	Chronological Objects which can Handle Dates and Times	2.3-50
<input type="checkbox"/> colorspace	Color Space Manipulation	1.3-2
<input type="checkbox"/> curl	A Modern and Flexible Web Client for R	2.8.1
<input type="checkbox"/> data.table	Extension of 'data.frame'	1.10.4
<input type="checkbox"/> dichromat	Color Schemes for Dichromats	2.0-0

Console

```
> one <- 1
> |
```

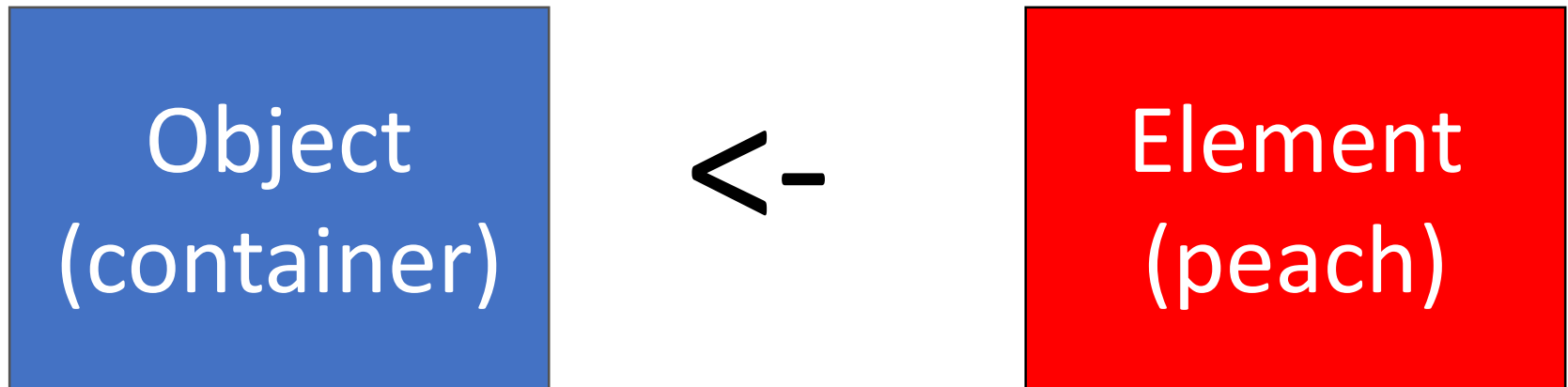


Object: what is it?

- **Object** is container
- **Element** is anything in container – a peach
- To reuse elements, they must be **stored as objects**
 - Any name **defined** by user
 - **Remain the same** unless overwritten
 - Must be **removed** by user as well



Object: creating/storing objects



Object: use

- Once objects exist, you may use it as whole or its elements for various operations
- It may be reused **again and again**
- Functions may be applied

```
one <- 1
```

```
one + one
```

```
[1] 2
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Untitled1*

Source on Save

Run

Source

```
1 one <- 1
2
3 one + one
4 |
```

4:1 (Top Level)

R Script

Console

```
> one <- 1
> one + one
[1] 2
> |
```

Environment History

Import Dataset

Global Environment

Values

one	1
-----	---

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.0
<input type="checkbox"/> audio	Audio Interface for R	0.1-5
<input type="checkbox"/> beepR	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.62.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> Cairo	R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output	1.5-9
<input type="checkbox"/> chron	Chronological Objects which can Handle Dates and Times	2.3-50
<input type="checkbox"/> colorspace	Color Space Manipulation	1.3-2
<input type="checkbox"/> curl	A Modern and Flexible Web Client for R	2.8.1
<input type="checkbox"/> data.table	Extension of 'data.frame'	1.10.4
<input type="checkbox"/> dichromat	Color Schemes for Dichromats	2.0-0

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Untitled1*

Source on Save

Run

Source

```
1 one <- 1
2
3 one + one
4 |
```

4:1 (Top Level)

R Script

Console

```
> one <- 1
> one + one
[1] 2
> |
```

Environment History

Import Dataset

Global Environment

Values

one	1
-----	---

?

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.0
<input type="checkbox"/> audio	Audio Interface for R	0.1-5
<input type="checkbox"/> beepR	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.62.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> Cairo	R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output	1.5-9
<input type="checkbox"/> chron	Chronological Objects which can Handle Dates and Times	2.3-50
<input type="checkbox"/> colorspace	Color Space Manipulation	1.3-2
<input type="checkbox"/> curl	A Modern and Flexible Web Client for R	2.8.1
<input type="checkbox"/> data.table	Extension of `data.frame`	1.10.4
<input type="checkbox"/> dichromat	Color Schemes for Dichromats	2.0-0

Object: use

- Any output of any operation needs to be stored
 - **new object** may be created
 - **existing object** may be **rewritten**

```
one <- 1  
one + one  
[1] 2
```

```
two <- one + one  
two  
[1] 2
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Untitled1*

Source on Save

Run

Source

```
1 one <- 1
2
3 one + one
4
5 two <- one + one
6 |
```

6:1 (Top Level)

R Script

Console

```
> one <- 1
> one + one
[1] 2
> two <- one + one
> |
```

Environment History

Import Dataset

Global Environment

Values

one	1
two	2

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.0
<input type="checkbox"/> audio	Audio Interface for R	0.1-5
<input type="checkbox"/> beeper	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.62.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> Cairo	R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output	1.5-9
<input type="checkbox"/> chron	Chronological Objects which can Handle Dates and Times	2.3-50
<input type="checkbox"/> colorspace	Color Space Manipulation	1.3-2
<input type="checkbox"/> curl	A Modern and Flexible Web Client for R	2.8.1
<input type="checkbox"/> data.table	Extension of `data.frame`	1.10.4
<input type="checkbox"/> dichromat	Color Schemes for Dichromats	2.0-0

Untitled1*

Source on Save

Run

Source

```
1 one <- 1
2
3 one + one
4
5 two <- one + one
6
7 two
8 |
```

8:1 (Top Level)

R Script

Console ~/ ↻

```
> one <- 1
> one + one
[1] 2
> two <- one + one
> two
[1] 2
> |
```

Environment History

Import Dataset

Global Environment

Values

one	1
two	2

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
User Library		
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.0
<input type="checkbox"/> audio	Audio Interface for R	0.1-5
<input type="checkbox"/> beepR	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.62.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> Cairo	R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output	1.5-9
<input type="checkbox"/> chron	Chronological Objects which can Handle Dates and Times	2.3-50
<input type="checkbox"/> colorspace	Color Space Manipulation	1.3-2
<input type="checkbox"/> curl	A Modern and Flexible Web Client for R	2.8.1
<input type="checkbox"/> data.table	Extension of `data.frame`	1.10.4
<input type="checkbox"/> dichromat	Color Schemes for Dichromats	2.0-0

Function



Function

- Pre-defined **methods** which allow operations over objects
- To **create** an object with more than one element, function `c ()` is used

```
onetofive <- c(1, 3, 5, 4, 2)
```

- Any object may be **manipulated** with a function

```
sort (onetofive)  
[1] 1 2 3 4 5
```


Function: arguments

- To extend functionality, functions have pre-defined **arguments**
 - **Arguments extend** the functionality of function
 - Some functions have many arguments, some none
- Results of the function must always be stored in the environment

```
sort(onetofive)
```

```
[1] 1 2 3 4 5
```

```
sort(onetofive, decreasing = TRUE)
```

```
[1] 5 4 3 2 1
```

```
onetofive <- sort(onetofive, decreasing = TRUE)
```

Function: syntax

- `functionname()` indicates function

```
sqrt(9)
[1] 3
```

- **Structure is** `function(arg1, arg2, ...)`

```
sample(0:100, 10, rep = FALSE)
[1] 48 50 37 94 42 39 21 19 63 95
```

- If **help** is necessary, just **add question mark or use function**
`help()` in front of the function name

```
?sample()
```

```
help(sample)
```

10.27

Data class

- **Properties of elements inside the object**
- **Numeric** – continuous numeric data
 - -1, 0.5, 10.27
- **Integer** – discrete numeric data
 - -1, 0, 1
- **Character** – string values
 - "anythingWithinQuotes"
- **Logical** – boolean output of logical operation
 - TRUE/FALSE, NA
- **Factor**
 - “agree”, “disagree”, “neutral”

Data class



Data class

- Sometimes, data elements are **ambiguous**
- Distinctions **important for computer**
- It is **useful to explore data class** of any object
 - This is not a necessary step
 - R comes with **pre-defined behavior** on classes – sometimes **counterintuitive** (e.g. in data sets, character variables are by default converted to factors)
 - Useful in **debugging** the code
- **If necessary, class can be converted/changed**

```
class(10.27)
[1] "numeric"
```

Data class: conversion of classes

```
as.numeric(10.27)
[1] 10.27
```

```
as.integer(10.27)
[1] 10
```

```
as.character(-1)
[1] "-1"
```

```
as.numeric("anything")
[1] NA
```

Warning message: NAs introduced by coercion

```
5 > 10
[1] FALSE
```

```
as.character(5 > 10)
[1] "FALSE"
```

Practice

- Create a new script
- Define 5 objects of your own liking (including object with more than one element)
- Get a class of each of the following items: 5, "5", NA, TRUE, "true", NULL
- Calculate the equation $\frac{(5.5*4)+(7.5*2.12)}{\frac{12}{7}}$ using R objects
- Try to get a square root of number 254

I JUST WROTE THE MOST BEAUTIFUL CODE OF MY LIFE.



THEY CASUALLY HANDED ME AN IMPOSSIBLE PROBLEM. IN 48 HOURS AND 200 LINES, I SOLVED IT.



ACADEMIA:

MY GOD ... THIS WILL MEAN A HALF-DOZEN PAPERS, A THESIS OR TWO, AND A PARAGRAPH IN EVERY TEXTBOOK ON QUEUEING THEORY!



BUSINESS:

YOU GOT THE PROGRAM TO STOP JAMMING UP? GREAT. WHILE YOU'RE FIXING STUFF, CAN YOU GET OUTLOOK TO SYNC WITH OUR NEW PHONES?

