

Interuniversity Institute for Biostatistics
and statistical Bioinformatics

An introduction to R: Basic skills in R Studio and R markdown

Developed by
Rudradev Sengupta, Bernard Osangir and Ziv Shkedy
Hasselt University

>eR-BioStat development workshop and Machine learning for real world data

Eldoret, 06/10/24-12/10/24
Kenya



ER-BioStat

GitHub <https://github.com/eR-Biostat>

@erbiostat

Overview

- R Studio/R markdown.
 - How to produce output using R markdown ?
 - How to develop output using R markdown ?
-
- Main concept: we run the same analysis multiple times and show how to produce different type of output.
 - Analysis: simple plot + simple linear regression.

Rmd programs for the class today

- To run the analysis presented in the slides you need the following Rmd programs:

Dataset	Rmd file	Output
mtcars	R_course_UHasselt_demo_V0.Rmd	R code in a R markdown file
mtcars	R_course_UHasselt_demo_V1.Rmd	produce simple HTML/PDF output for a report
Airquality	R_course_UHasselt_2021(html)_V1.Rmd	Produce a HTML output (example how to work with R markdown)
Airquality	R_course_UHasselt_2021(pdf)_V1.Rmd	Produce HTML/PDF output
Old Faithful	R_course_UHasselt_2021(pdf)_V2.Rmd	A report for an example of a simple analysis (HTML/PDF)

See later in the slides how these files are used and which type of output can be produced.

R Studio

What is R Studio ?

- R Studio is an integrated development environment (IDE) for R.
- It includes a **console, syntax-highlighting editor** that supports direct code execution.
- Tools for plotting, history, debugging and workspace management.
- R Studio is available in **open Source** and **commercial** editions and runs on the desktop (Windows, Mac, and Linux).

What is R Studio ?

- More information:

<https://rstudio.com/products/rstudio/>

R Studio

The screenshot shows the R Studio interface. The left pane is the 'Console' window, displaying the R startup message and workspace details:

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

[workspace loaded from ~/.RData]

> |
```

The right pane contains three main sections: 'Environment', 'History', and 'Data'. The 'Data' section lists various objects in the global environment:

Object	Description
dat	16999 obs. of 12 variables
dat1	21 obs. of 2 variables
data1	12 obs. of 2 variables
datmat	num [1:100, 1:100] 17.4 17.4 17.4 17.4 ...
datatest	num [1:10, 1:10] 24.7 24.4 24.4 21.3 20.8 ...
idd	int [1:30, 1:30] 1 2 3 4 5 6 7 8 9 10 ...
out	40001 obs. of 4 variables
outp02	40001 obs. of 4 variables
outp04	40001 obs. of 4 variables
scandat1	16000 obs. of 12 variables

The bottom right corner shows the system tray with icons for network, battery, and date/time (EN 11:33 30/09/2020).

An overlaid red text box in the center-left area reads:

An initial screen for R Studio
(console)

The cars dataset in R

The data give the speed of cars and the distances taken to stop. Note that the data were recorded in the 1920s.

Two variables:

- Cars' speed.
- Cars' stopping distance.

```
> head(cars)
   speed dist
1     4    2
2     4   10
3     7    4
4     7   22
5     8   16
6     9   10
```

R Studio: example

R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

[workspace loaded from ~/.RData]

```
> plot(cars$speed, cars$dist)
Warning message:
R graphics engine version 12 is not supported by this version of RStudio. The Plots tab will b
> |
```

A plot in R Studio

- The **cars** dataset.
- Plot of the car's speed versus stopping distance.

plot(x,y)

A scatter plot showing the relationship between car speed (x-axis) and stopping distance (y-axis). The x-axis ranges from 5 to 25, and the y-axis ranges from 0 to 120. The data points show a positive correlation, with stopping distance increasing as speed increases. A red text annotation 'plot(x,y)' is overlaid on the top left of the plot area.

File History Resize

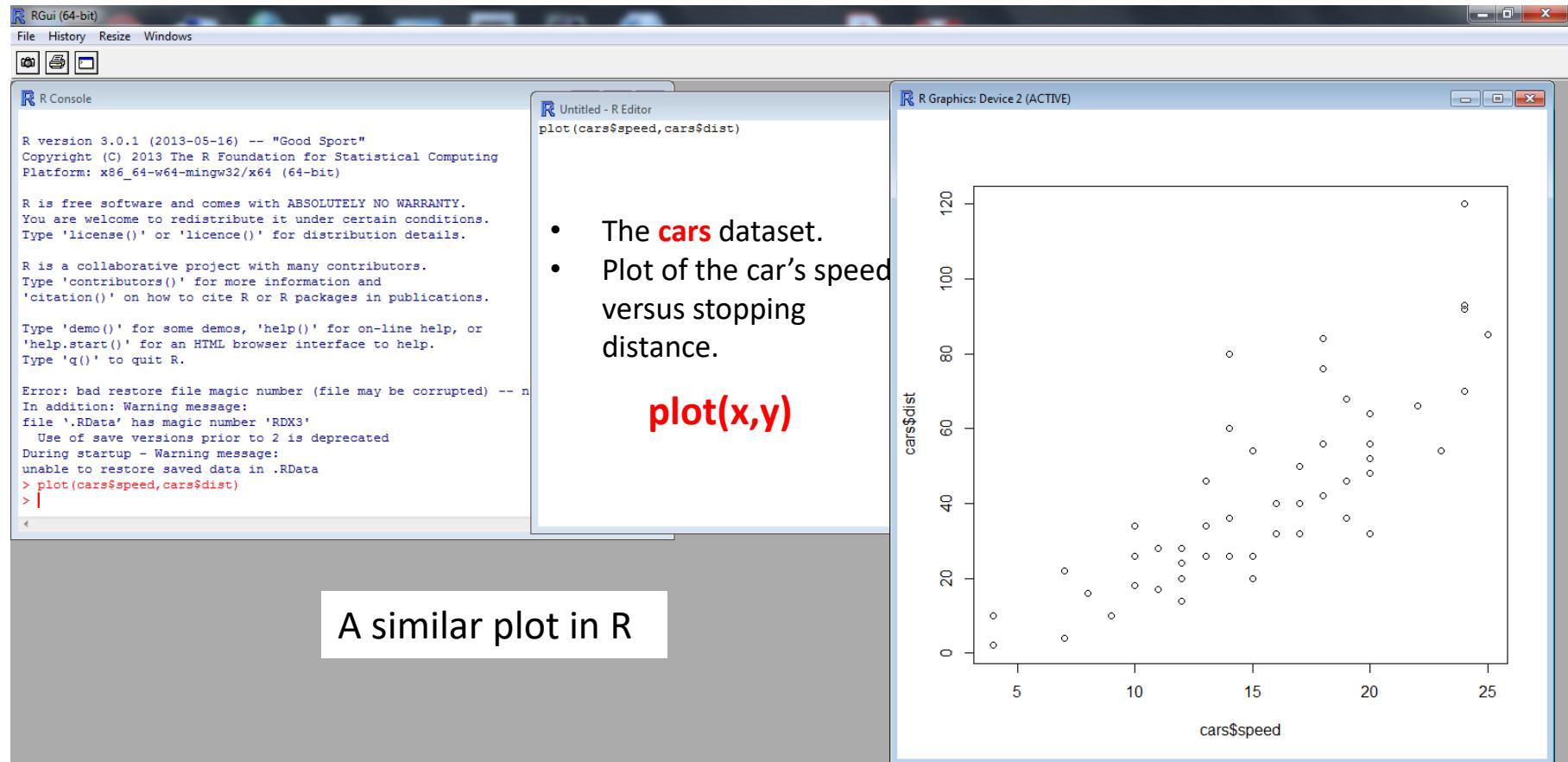
Project: (None)

Version

1.4-5
1.1-4
1.1
0.2.1
2015.6.25
1.1.4
0.1-3
1.69.0-1
2.0.1
1.30.4
2.12.0
3.9.0
1.0-6
0.16
bookdown
Authoring Books and Technical Documents with R Markdown
boot
Bootstrap Functions (Originally by Angelo Canty for S)

EN 11:34 30/09/2020

R: example



General structure of R

- R Functions.
- R packages (written by developers) as a part of the software:
 - `glm()`.
- R packages (written by developers) uploaded in CRAN, Bioconductor, Github....
- R packages (written by the user=you).

R functions

Example:

```
function(data)
```

A procedure that was programmed in R that uses data to produce output.

```
> var(x)
```

The r function

```
data
```

Calculate the sample variance.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

R packages

- A package in R: a **collection of functions** can be used to conduct an analysis or data management.
- Example: the **lme4** package for linear mixed models.
- Some packages are included in R and some can be install from CRAN/Bioconductor.
- **CRAN**: website with R packages ready to be installed and used.

<https://cran.r-project.org/>

R Studio: packages

R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

[workspace loaded from ~/.RData]

```
> plot(cars$speed,cars$dist)
Warning message:
R graphics engine version 12 is not supported by this version of RStudio.
> |
```

Upload new a package to R Studio

Install Packages

Install from: Repository (CRAN)

Packages (separate multiple with space or comma):

Install to Library: C:/Program Files/R/R-3.6.1/library [Default]

Install dependencies

Install Cancel

Environment History

Import Dataset

Global Environment

Data

Name	Description
dat	16999 obs. of 12 variables
dat1	21 obs. of 2 variables
data1	12 obs. of 2 variables
datamat	num [1:100, 1:100] 17.4 17.4 17.4 17.4 17.4 ...
dist	num [1:10, 1:10] 24.7 24.4 24.4 21.3 20.8 ...
int	int [1:30, 1:30] 1 2 3 4 5 6 7 8 9 10 ...
q1	40001 obs. of 4 variables
q2	40001 obs. of 4 variables
q3	40001 obs. of 4 variables
q4	40001 obs. of 4 variables
q5	16000 obs. of 12 variables

File Plots Packages Help Viewer

Install Update

Name Description Version

System Library

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-5
additivityTests	Additivity Tests in the Two Way Anova with Single Sub-class Numbers	1.1-4
askpass	Safe Password Entry for R, Git, and SSH	1.1
assertthat	Easy Pre and Post Assertions	0.2.1
asympow	Calculate Power Utilizing Asymptotic Likelihood Ratio Methods	2015.6.25
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.4
base64enc	Tools for base64 encoding	0.1-3
BH	Boost C++ Header Files	1.69.0-1
biclust	BiCluster Algorithms	2.0.1
BiocManager	Access the Bioconductor Project Package Repository	1.30.4
BiocStyle	Standard styles for vignettes and other Bioconductor documents	2.12.0
BiocVersion	Set the appropriate version of Bioconductor packages	3.9.0
bitops	Bitwise Operations	1.0-6
bookdown	Authoring Books and Technical Documents with R Markdown	0.16
boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-23

EN 11:35 30/09/2020

Example 1

The mtcars data

Part 1: R and R Studio

- Example of a data analysis: the `mtcars` data.
- Use `help(mtcars)` to get information about the data.
- Running the analysis in R.
- Running the analysis in R-Studio.

The mtcars data: analysis in R

- Use `help(mtcars)` in R.

The screenshot shows an R session with a browser window displaying the mtcars dataset documentation. The R console on the left shows the command `help(mtcars)` being run. The browser window title is "R: Motor Trend Car Road Tests" and the URL is "127.0.0.1:16512/library/datasets/html/mtcars.html". The page content includes the `mtcars` dataset description, usage information, and examples. Two red arrows point to the first two items in the variable list: "[, 1] mpg Miles/(US) gallon" and "[, 2] cyl Number of cylinders".

mtcars {datasets}

Description

The data was extracted from the 1974 *Motor Trend* US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

Usage

mtcars

Format

A data frame with 32 observations on 11 variables.

[, 1] mpg Miles/(US) gallon ←
[, 2] cyl Number of cylinders ←
[, 3] disp Displacement (cu.in.)
[, 4] hp Gross horsepower
[, 5] drat Rear axle ratio
[, 6] wt Weight (lb/1000)
[, 7] qsec 1/4 mile time
[, 8] vs V/S
[, 9] am Transmission (0 = automatic, 1 = manual)
[,10] gear Number of forward gears
[,11] carb Number of carburetors

Source

Henderson and Velleman (1981), Building multiple regression models interactively. *Biometrics*, 37, 391–411.

Examples

```
require(graphics)
pairs(mtcars, main = "mtcars data")
coplot(mpg ~ disp | as.factor(cyl), data = mtcars,
       panel = panel.smooth, rows = 1)
```

Description of the data in the help system.

The mtcars data in R

```
> head(mtcars)
```

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

- Variables in the data.
- For our example: **mpg** and **wt**.

Analysis

- Plot mpg Vs. weight.
- Calculate the mean weight.
- Fit a simple regression model for mpg on Weight.

To run the analysis in the example, use the program in R Studio:

[R_course_UHasselt_demo_V0.Rmd](#)

Analysis in basic R & output

R Gui (64-bit)

File History Resize Windows

R Console

```
> plot(mtcars$wt, mtcars$mpg)
> mean(mtcars$mpg)
[1] 20.0962
> fit.lm<-lm(mtcars$mpg~mtcars$wt)
> summary(fit.lm)

Call:
lm(formula = mtcars$mpg ~ mtcars$wt)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.5432 -2.3647 -0.1252  1.4096  6.8727 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 37.2851    1.8776 19.858 < 2e-16 ***
mtcars$wt   -5.3445    0.5591 -9.559 1.29e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 

Residual standard error: 3.046 on 30 degrees of freedom
Multiple R-squared:  0.7528, Adjusted R-squared:  0.7528 
F-statistic: 91.38 on 1 and 30 DF, p-value: 1.29
```

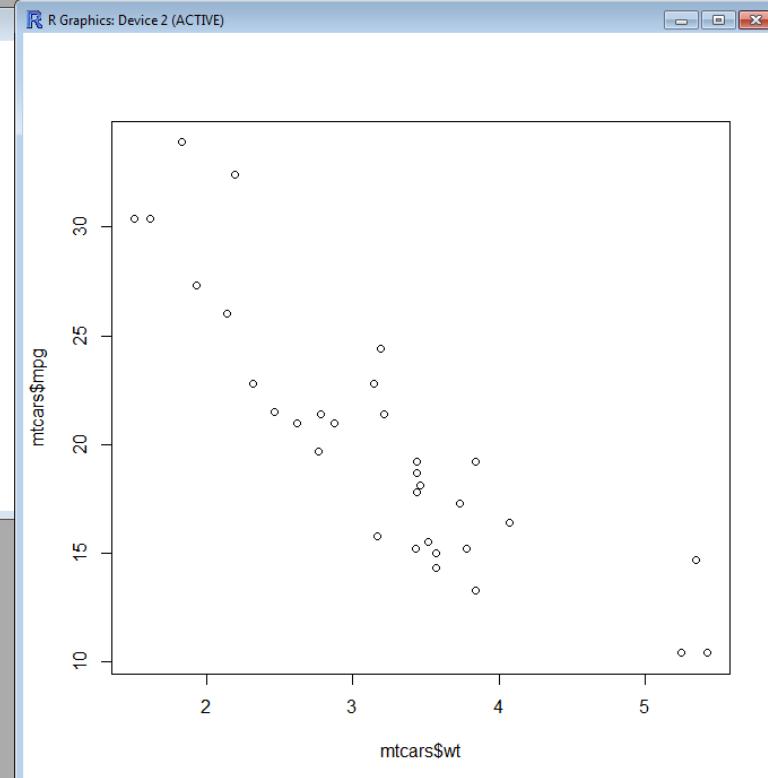
see next slide

The estimated model

Untitled - R Editor

```
summary(mtcars)
plot(mtcars$wt, mtcars$mpg)
mean(mtcars$mpg)
fit.lm<-lm(mtcars$mpg~mtcars$wt)
summary(fit.lm)
```

R Graphics: Device 2 (ACTIVE)



A scatter plot showing the relationship between car weight (mtcars\$wt) on the x-axis and fuel economy (mtcars\$mpg) on the y-axis. The x-axis ranges from approximately 1.6 to 5.4, and the y-axis ranges from 10 to 35. The data points show a clear negative correlation, indicating that heavier cars tend to have lower fuel economy.

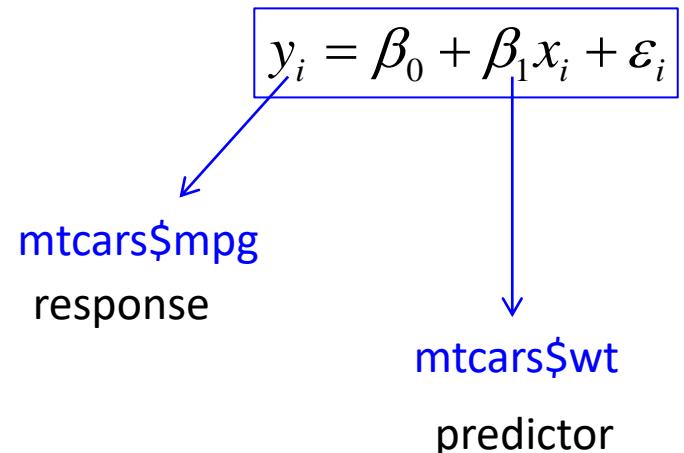
EN 10:06 29/09/2020

The R code for the analysis

```
summary(mtcars)
plot(mtcars$wt,mtcars$mpg)
mean(mtcars$mpg)
fit.lm<-lm(mtcars$mpg~mtcars$wt)
summary(fit.lm)
```

fit a simple linear regression model

The regression output



The mtcars data: an analysis in R Studio

The R code

```
1 ---  
2 title: "Demo 1"  
3 author: "Rudradev Sengupta & ziv shkedy"  
4 date: "29 September"  
5 output:  
6   pdf_document: default  
7   html_document: default  
8 params:  
9   snapshot: lubridate::ymd_hms("2015-01-01 12:30:00")  
10  start: lubridate::ymd("2015-01-01")  
11 ---  
12   
13 ```{r setup, include=FALSE}  
14 knitr::opts_chunk$set(echo = TRUE)  
15   
16   
17  
18 summary(mtcars)  
19 plot(mtcars$wt,mtcars$mpg)  
20 mean(mtcars$mpg)  
21 fit.lm<-lm(mtcars$mpg~mtcars$wt)  
22 summary(fit.lm)|
```

22:16 (Top Level) R Markdown

Console R Markdown x

```
>/\r  
> fit.lm<-lm(mtcars$mpg~mtcars$wt)  
> summary(fit.lm)  
  
Call:  
lm(formula = mtcars$mpg ~ mtcars$wt)  
  
Residuals:  
    Min      1Q      Median      3Q      Max  
-4.5432 -2.3647 -0.1252  1.4096  6.8727  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 37.2851    1.8776  19.858 < 2e-16 ***  
mtcars$wt    -5.3445    0.5591  -9.559 1.29e-10 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 3.046 on 30 degrees of freedom  
Multiple R-squared:  0.7528    Adjusted R-squared:  0.7416
```

Environment History

Global Environment

Name	Type	Value
b1	Named num	9773
b2	Named num	0.282
b3	Named num	28.1
bi	Named num	17.9668497569183
cases_d	num [1:60]	1.34 1.78 2.36 3.12 4.14 ...
cases.1	num [1:21]	1 2 3 7 13 17 24 24 51 62 ...
cases.21	num [1:21]	1 2 3 7 13 17 24 24 51 62 ...
dose	num [1:12]	1 1 1 2 2 3 3 3 4 ...
f.val	num [1:16998]	11.769 1.52 12.755 0.132 3 ...
fdose	Factor w/ 4 levels	"1","2","3","4": 1 1 ...
fit.lm	List of 12	

Files Plots Packages Help Viewer

System Library

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-5
additivityTests	Additivity Tests in the Two Way Anova with Single Sub-class Numbers	11-4
askpass	Safe Password Entry for R, Git, and SSH	1.1
assertthat	Easy Pre and Post Assertions	0.2.1
asympow	Calculate Power Utilizing Asymptotic Likelihood Ratio Methods	2015.6.25
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.4
base64enc	Tools for base64 encoding	0.1-3
BH	Boost C++ Header Files	1.69.0-1
bioclust	BiCluster Algorithms	2.0.1
BiocManager	Access the Bioconductor Project Package Repository	1.30.4
BiocStyle	Standard styles for vignettes and other Bioconductor documents	2.12.0
BiocVersion	Set the appropriate version of Bioconductor packages	3.9.0
bitops	Bitwise Operations	1.0-6
bookdown	Authoring Books and Technical Documents with R Markdown	0.16

EN 29/09/2020 10:10

The mtcars data : an analysis in R Studio

The screenshot shows the RStudio interface with several panels:

- Code Editor:** Shows R code for setting up a new R Markdown document and performing a linear regression analysis on the mtcars dataset.
- Console:** Displays the results of the R code execution, including the regression output.
- Global Environment:** Shows the objects available in the current R session, such as `b1`, `b2`, `b3`, `bi`, `cases_d`, `cases.1`, `cases.21`, `dose`, `f.val`, `fdose`, and `fit.lm`.
- Run Menu:** A context menu is open over the code editor, specifically over the line `fit.lm<-lm(mtcars$mpg~mtcars$wt)`. The menu options include `Run Selected Line(s) Ctrl+Enter`, `Run Current Chunk`, `Run Next Chunk`, `Run Setup Chunk`, `Run Setup Chunk Automatically`, `Run All Chunks Above Ctrl+Alt+P`, `Run All Chunks Below`, `Restart R and Run All Chunks`, `Restart R and Clear Output`, and `Run All Ctrl+Alt+R`. The option `Run Current Chunk` is highlighted with a red arrow.
- Text Labels:** Two red arrows point to specific areas: one to the "Run" menu in the top right and another to the "Run Current Chunk" option in the context menu.
- Bottom Taskbar:** Shows icons for various applications like Word, Excel, and R.

The R code

```
1 ---  
2 title: "Demo 1"  
3 author: "Rudradeep Sengupta & ziv shkedy"  
4 date: "29 September"  
5 output:  
6   pdf_document: default  
7   html_document: default  
8 params:  
9   snapshot: lubridate::ymd("2015-01-01 12:30:00")  
10  start: lubridate::ymd("2015-01-01")  
11 ---  
13 ``{r setup, include=FALSE}  
14 knitr::opts_chunk$set(echo = TRUE)  
15  
16  
18 summary(mtcars)  
19 plot(mtcars$wt,mtcars$mpg)  
20 mean(mtcars$mpg)  
21 fit.lm<-lm(mtcars$mpg~mtcars$wt)  
22 summary(fit.lm)
```

The output

```
Call:  
lm(formula = mtcars$mpg ~ mtcars$wt)  
Residuals:  
    Min      1Q  Median      3Q     Max  
-4.5432 -2.3647 -0.1252  1.4096  6.8727  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 37.2851   1.8776 19.858 < 2e-16 ***  
mtcars$wt   -5.3445   0.5591 -9.559 1.29e-10 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
Residual standard error: 3.046 on 30 degrees of freedom  
Multiple R-squared:  0.7528, Adjusted R-squared:  0.7446  
F-statistic: 91.38 on 1 and 30 DF,  p-value: 1.294e-10
```

Analysis in R Studio: the output

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Addins

R_Course_UHasselt_demo_V0.Rmd x R_Course_UHasselt_demo_V1.Rmd x

code

```
1 ---  
2 title: "demo 1"  
3 author: "Rudradev Sengupta & ziv shkedy"  
4 date: "29 September"  
5 output:  
6   pdf_document: default  
7   html_document: default  
8 params:  
9   snapshot: lubridate::ymd_hms("2015-01-01 12:30:00")  
10  start: lubridate::ymd("2015-01-01")  
11 ---  
12   
13 {r setup, include=FALSE}  
14 knitr::opts_chunk$set(echo = TRUE)  
15   
16  
17  
18 summary(mtcars)  
19 plot(mtcars$wt, mtcars$mpg)  
20 mean(mtcars$mpg)  
21 fit.lm<-lm(mtcars$mpg~mtcars$wt)  
22 summary(fit.lm)
```

18:1 (Top Level) ↓

Console R Markdown x

~/

```
> fit.lm<-lm(mtcars$mpg~mtcars$wt)
> summary(fit.lm)

Call:
lm(formula = mtcars$mpg ~ mtcars$wt)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.5432 -2.3647 -0.1252  1.4096  6.8727 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 37.2851    1.8776 19.858 < 2e-16 ***
mtcars$wt    -5.3445    0.5591 -9.559 1.29e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.046 on 30 degrees of freedom
Multiple R-squared:  0.7528  Adjusted R-squared:  0.7446 
```

The estimated model

R Graphics: Device 2 (ACTIVE)

File History Resize

Graphical window

Scatter plot showing the relationship between mtcars\$mpg (Y-axis, ranging from 10 to 35) and mtcars\$wt (X-axis, ranging from 2 to 5). The data points show a negative correlation, indicating that as weight increases, fuel efficiency (MPG) tends to decrease.

mtcars\$mpg

mtcars\$wt

mtcars

bioconductor

biocstyle

Standard styles for vignettes and other Bioconductor documents

2.12.0

BiocVersion

Set the appropriate version of Bioconductor packages

3.9.0

bitops

Bitwise Operations

1.0-6

bookdown

Authoring Books and Technical Documents with R Markdown

0.16

EN 24 10:05 29/09/2020

R Studio

- So far, R Studio (in this setting): similar to R BUT....
- R Studio + R markdown: a different level of output.

R markdown

R markdown: what?

- Markdown allows you to write a file format independent document using an easy-to-read and easy-to-write plain text format.
- Instead of marking up text so that is easy for a computer to read
 - e.g. HTML: <html><body>Name</body></html>
- The goal is to mark down text so that it is easy and human readable (instead of machine readable):
 - e.g. **Name**
- The most easy way to understand it: R markdown allows us to write free text and run R together in the same file.

R markdown: what?

- Markdown is a specific Markup language which is structured very loosely => any file format can be generated using pandoc.
- Pandoc: R function to convert documents To other formats.
- From one Markdown document you can generate different file formats: html, PDF, docx, slideshows, rtf, etc.
 - The downside is that there is slightly less control over formatting.

Example of slideshows: next class.

R markdown: what?

- Extension of Markdown via R:
 - Allowing *R code* and its *results* to be merged with Markdown.
 - Ensuring that *R Markdown documents are fully reproducible*.
 - Enabling extra modifications to original markdown specification.
- Provides an *unified authoring framework for data science*, combining your code, its results, etc.
- In practice: just by changing the dataset then entire analysis can be rerun and the new report can be produced.
- Integrates a number of R packages and external tools.

R markdown: what?

- A lot of online materials.
- RMarkdown Cheat Sheet: *Help > Cheatsheets > R Markdown Cheat Sheet* (<https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>)
- RMarkdown Reference Guide: *Help > Cheatsheets > R Markdown Reference Guide*
- Both cheatsheets are also available at <http://rstudio.com/cheatsheets>
- *Help > Markdown Quick Reference*

The R markdown program: components

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

```
1 ---  
2 title: "Demo"  
3 author: "Rudradev Sengupta"  
4 date: "16 July 2019"  
5 output: html_document  
6 ---  
7  
8 ...{r setup, include=FALSE}  
9 knitr::opts_chunk$set(echo = TRUE)  
10  
11  
12 ## R Markdown  
13  
14 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS  
15 word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.  
16 when you click the **Knit** button a document will 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  
18 ...{r cars}  
19 summary(cars)  
20  
21 ## Including Plots  
22  
23 you can also embed plots, for example:  
24  
25 ...{r pressure, echo=FALSE}  
26 plot(pressure)  
27  
28 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code  
that generated the plot.  
29  
30
```

Annotations on the left side of the code editor highlight sections of the document:

- A brace labeled "YAML Header" covers the first 6 lines.
- A brace labeled "Formatted Text" covers the free text and the "Note that..." section.
- A brace labeled "Code Chunks" covers the two code blocks (lines 18-19 and 25-27).

Two red arrows point from the annotations to the corresponding code blocks in the editor.

In the bottom panel, the "Console" tab shows the command used to run the R Markdown file and the resulting output file:

```
C:/programs_Rudra/RStudio/bin/pandoc/pandoc" +RTS -K512m -RTS demo.utf8.md --to html4 --from markdown+auto  
ink_bare_uris+ascii_identifiers+tex_math_single_backslash --output demo.html --smart --email-obfuscation non  
e-self-contained --standalone --section-divs --template "C:/programs_Rudra/R-3.5.1/library/rmarkdown/rmd/h  
\default.html" --no-highlight --variable highlightjs=1 --variable "theme:bootstrap" --include-in-header "C:/  
Users\rsengup4\AppData\Local\Temp\1\Rtmpg9j6Tx\rmardown-str37f8437e3e4d.html" --mathjax --variable "mathjax  
-url:https://mathjax.rstudio.com/latest/MathJax.js?config=TeX-AMS-MML_HTMLorMML"  
output created: demo.html
```

There are principally three sections to an R Markdown document:

- YAML header surrounded by `---`
- Code chunks surrounded by `````
- Free text mixed with simple text formatting like `#heading` and `_italics_`

Programname.Rmd

Starting point

- Available materials to try out:
 - <https://teams.microsoft.com/#/files/Reference%20Material?threadId=19:05ec12c79df7460ca9cdfbd8b620f16a@thread.skype&ctx=channel&context=Rmarkdown%2520Help> – templates by Stefan to create pdf/html/.. documents
 - <https://teams.microsoft.com/#/files/Reference%20Material?threadId=19:05ec12c79df7460ca9cdfbd8b620f16a@thread.skype&ctx=channel&context=Code%2520Repository%252FAssignment%25206> – solutions for Assignment 6, by different groups in US

Starting point

- Available materials to try out:
 - https://rmarkdown.rstudio.com/articles_intro.html – materials from RStudio

Example 1 (continue)

The mtcars data

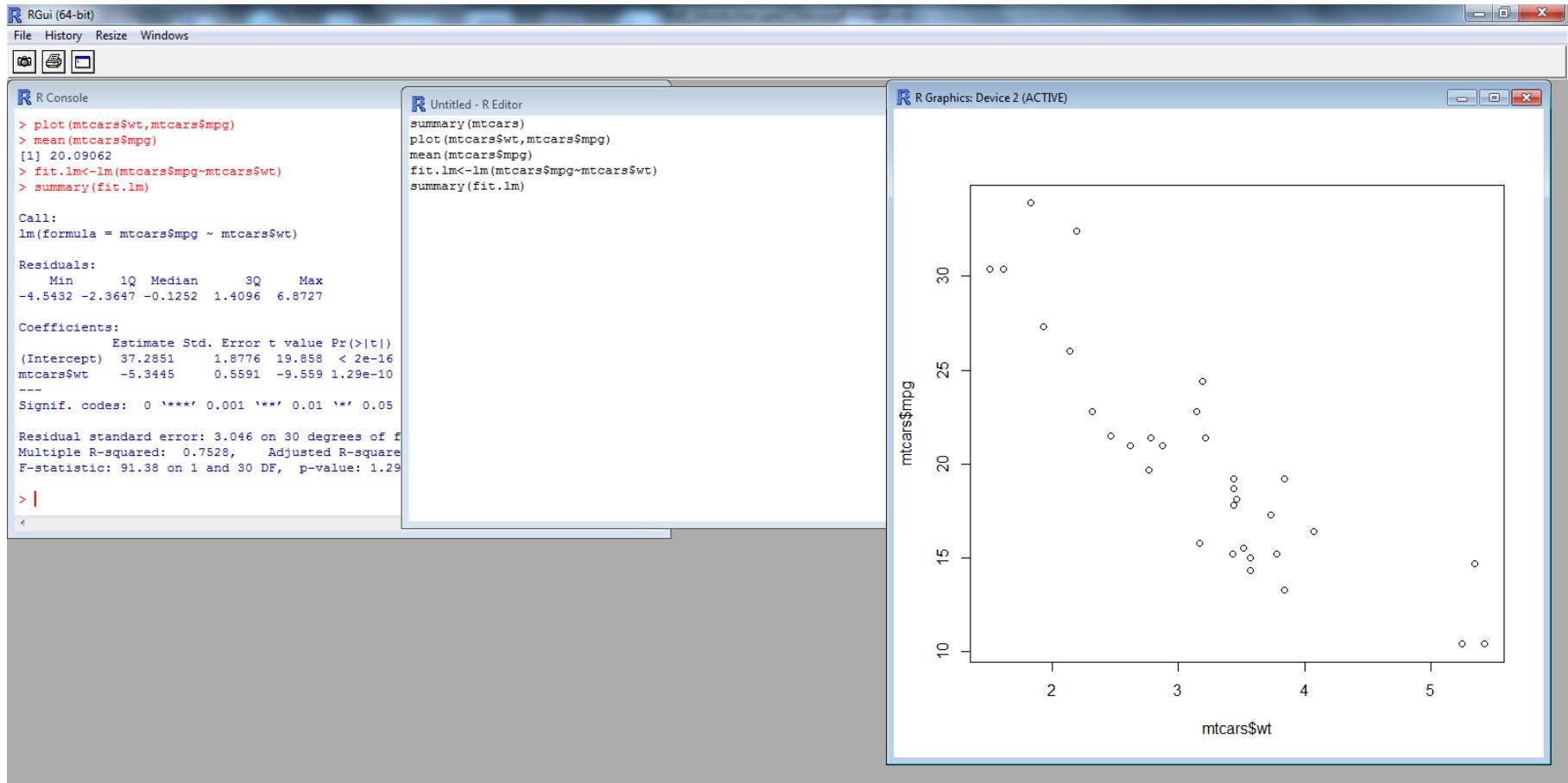
Part 1: Analysis with R and R Studio

Analysis

- Plot mpg Vs. weight.
- Calculate the mean weight.
- Fit a regression model for Mpg on Weight.
- R code: see next page.

Analysis in R

- The same slides as 18-22.



Analysis in R Studio

The screenshot shows the RStudio interface with the following components:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Includes icons for file operations, search, and R-related functions.
- Code Editor:** Displays an R Markdown script with code and comments. The code includes setting parameters for a report, loading the mtcars dataset, creating a linear model (fit.1m), and summarizing it.
- Console:** Shows the R session output, including the creation of fit.1m, its summary, and the results of the lm() function.
- Environment:** A pane showing the global environment with various objects and their types and values.
- Packages:** A pane showing the system library with installed packages and their details.
- Bottom Bar:** Icons for various applications and the R logo.
- System Tray:** Shows the date and time (10:10, 29/09/2020).

Analysis in R Studio

The screenshot shows the R Studio interface with a context menu open over a code editor window. The menu options include:

- Run Selected Line(s) Ctrl+Enter
- Run Current Chunk
- Run Next Chunk
- Run Setup Chunk
- Run Setup Chunk Automatically
- Run All Chunks Above Ctrl+Alt+P
- Run All Chunks Below
- Restart R and Run All Chunks
- Restart R and Clear Output
- Run All Ctrl+Alt+R

The 'Run Current Chunk' option is highlighted with a red box. The code editor contains R Markdown code, including a chunk setup and a linear regression analysis of the mtcars dataset.

Code Editor Content:

```
1 ---  
2 title: "Demo 1"  
3 author: "Rudradev Sengupta & ziv shkedy"  
4 date: "29 September"  
5 output:  
6   pdf_document: default  
7   html_document: default  
8 params:  
9   snapshot: lubridate::ymd_hms("2015-01-01 12:30:00")  
10  start: lubridate::ymd("2015-01-01")  
11 ---  
13 for setup, include=FALSE}  
14 knitr::opts_chunk$set(echo = TRUE)  
15  
16  
18 summary(mtcars)  
19 plot(mtcars$wt,mtcars$mpg)  
20 mean(mtcars$mpg)  
21 fit.lm<-lm(mtcars$mpg~mtcars$wt)  
22 summary(fit.lm)
```

Console Output:

```
18:1 [Top Level]   
R Markdown   
Console | R Markdown x  
~/   
fit.lm<-lm(mtcars$mpg~mtcars$wt)  
Residuals:  
    Min      1Q  Median      3Q     Max  
-4.5432 -2.3647 -0.1252  1.4096  6.8727  
Coefficients:  
Estimate Std. Error t value Pr(>|t|)  
(Intercept) 37.2851  1.8776 19.858 < 2e-16 ***  
mtcars$wt    -5.3445  0.5591 -9.559 1.29e-10 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
Residual standard error: 3.046 on 30 degrees of freedom  
Multiple R-squared:  0.7528, Adjusted R-squared:  0.7446  
F-statistic: 91.38 on 1 and 30 DF, p-value: 1.294e-10  
> |
```

Global Environment:

Name	Type	Value
b1	Named num	9773
b2	Named num	0.282
b3	Named num	28.1
b1	num	[1:60] 1.34 1.78 2.36 3.12 4.14 ...
cases_d	num	[1:21] 1 2 3 7 13 17 24 24 51 62 ...
cases.1	num	[1:21] 1 2 3 7 13 17 24 24 51 62 ...
cases.21	num	[1:12] 1 1 1 2 2 2 3 3 3 4 ...
dose	num	[1:12] 1 1 1 2 2 2 3 3 3 4 ...
f.val	num	[1:16998] 11.769 1.52 12.755 0.132 3 ...
fdose	Factor w/ 4 levels "1","2","3","4": 1 1 ...	
fit.lm	List of 12	

System Library:

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-5
additivityTests	Additivity Tests in the Two Way Anova with Single Sub-class Numbers	1.1-4
askpass	Safe Password Entry for R, Git, and SSH	1.1
assertthat	Easy Pre and Post Assertions	0.2.1
asympow	Calculate Power Utilizing Asymptotic Likelihood Ratio Methods	2015.6.25
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.4
base64enc	Tools for base64 encoding	0.1-3
BH	Boost C++ Header Files	1.69.0-1
biclust	BiCluster Algorithms	2.0.1
BioManager	Access the Bioconductor Project Package Repository	1.30.4
BioStyle	Standard styles for vignettes and other Bioconductor documents	2.12.0
BiocVersion	Set the appropriate version of Bioconductor packages	3.9.0
bitops	Bitwise Operations	1.0-6
bookdown	Authoring Books and Technical Documents with R Markdown	0.16

Analysis in R Studio

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

R_Course_UHasselt_demo_V0.Rmd x R_Course_UHasselt_demo_V1.Rmd x

```
1 ---  
2 title: "demo 1"  
3 author: "Rudradev Sengupta & ziv shkedy"  
4 date: "29 September"  
5 output:  
6   pdf_document: default  
7   html_document: default  
8 params:  
9   snapshot: lubridate::ymd_hms("2015-01-01 12:30:00")  
10  start: lubridate::ymd("2015-01-01")  
11 ---  
12   
13 {r setup, include=FALSE}  
14 knitr::opts_chunk$set(echo = TRUE)  
15   
16  
17  
18 summary(mtcars)  
19 plot(mtcars$wt, mtcars$mpg)  
20 mean(mtcars$mpg)  
21 fit.lm<-lm(mtcars$mpg~mtcars$wt)  
22 summary(fit.lm)
```

18:1 (Top Level) `~ /`

Console R Markdown x

```
> fit.lm<-lm(mtcars$mpg~mtcars$wt)
> summary(fit.lm)

Call:
lm(formula = mtcars$mpg ~ mtcars$wt)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.5432 -2.3647 -0.1252  1.4096  6.8727 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 37.2851    1.8776 19.858 < 2e-16 ***
mtcars$wt    -5.3445    0.5591 -9.559 1.29e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.046 on 30 degrees of freedom
Multiple R-squared:  0.7528  Adjusted R-squared:  0.7446 
```

R Graphics: Device 2 (ACTIVE)

File History Resize

Scatter plot showing the relationship between car weight (mtcars\$wt) on the x-axis and fuel economy (mpg) on the y-axis. The x-axis ranges from approximately 1.5 to 5.5, and the y-axis ranges from 10 to 35. The plot shows a clear negative linear trend, with data points scattered around a downward-sloping line of best fit.

mtcars\$mpg

mtcars\$wt

Standard styles for vignettes and other Bioconductor documents 2.12.0

BiocStyle

Set the appropriate version of Bioconductor packages 3.9.0

BiocVersion

Bitwise Operations 1.0-6

bitops

Authoring Books and Technical Documents with R Markdown 0.16

bookdown

EN 40 10:05 29/09/2020

Part 2: Analysis with R markdown

To run the analysis in the example, use the program in R Studio:

`R_course_UHasselt_demo_V1.Rmd`

Different types of Output

- Two types of output:
 - HTML.
 - PDF.
- Text and R code in the same document.

Running R in Markdown

- General form:

```{r}  
R code  
```



Code chunks for
the analysis.

- Example of a scatterplot:

```{r}  
plot(x,y)  
```

Use the function `plot()` to
produce a scatterplot.

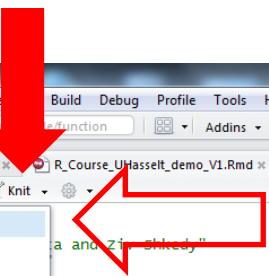
The Rmd program

Title and info.

The screenshot shows the RStudio interface with the following components:

- Top Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project Bar:** Project: (None).
- Code Editor:** Displays the R Markdown file `R_Course_UHasselt_demo_V0.Rmd` and `R_Course_UHasselt_demo_V1.Rmd`.
 - YAML Header:** Lines 1-15 of `V1.Rmd` show the YAML header, which is highlighted with a red bracket and arrow.
 - R code:** Lines 23-36 of `V1.Rmd` show an R code chunk, indicated by a red bracket and arrow.
 - Example of free text:** Lines 17-36 of `V1.Rmd` show an example of free text, enclosed in a blue box.
- Environment Tab:** Shows the Global Environment with various data frames and objects.
- Packages Tab:** Shows the System Library with a list of installed packages.
- Console:** Displays the R command-line interface with help messages and workspace information.
- Bottom Bar:** Icons for Windows, Word, Excel, PDF, Video, Image, Google Chrome, Internet Explorer, R, and RStudio.
- Bottom Right:** Language (EN), Date (30/09/2020), and Time (12:07).

Knit to HTML (how to run the program)



Running the program: knit to HTML

The screenshot shows the RStudio interface. On the left, the code editor displays an R Markdown file named "R_Course_UHasselt_demo_V1.Rmd". A red arrow points to the "Knit" button in the toolbar above the editor. The main pane shows the rendered HTML output with the title "Running the program: knit to HTML". The right pane contains the "Environment" and "Global Environment" panes, showing various R objects and their values. The bottom navigation bar includes icons for Windows, Word, Excel, and PDF.

```
1 #> 
2 #> tit Knit to HTML
3 #> aut Knit to PDF
4 #> dat Knit to Word
5 #> out Knit with Parameters...
6 #> h Clear Knitr Cache...
7 #> par id_hms("2015-01-01 12:30:00")
8 #> s start: lubridate::ymd("2015-01-01")
9 #> ---
10 #> 
11 #> 
12 #> 
13 #> ``{r setup, include=FALSE}
14 knitr::opts_chunk$set(echo = TRUE)
15 
16 
17 ## R Markdown
18 
19 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
20 
21 When you click the **Knit** button a document will 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:
22 
23 ``{r cars}
24 R Markdown
25 
```

Residuals:

Min	1Q	Median	3Q	Max
-4.5432	-2.3647	-0.1252	1.4096	6.8727

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	37.2851	1.8776	19.858	< 2e-16 ***
mtcars\$wt	-5.3445	0.5591	-9.559	1.29e-10 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.046 on 30 degrees of freedom
Multiple R-squared: 0.7528, Adjusted R-squared: 0.7446
F-statistic: 91.38 on 1 and 30 DF, p-value: 1.294e-10

> |

EN ▾ 10:20 29/09/2020

Output: HTML

See page 43

The screenshot shows the RStudio interface with two panes. The left pane displays an R Markdown file (R_Course_UHasselt_demo_V1.Rmd) containing code and text. A red arrow points from the line "# R Markdown" to the right pane. The right pane shows the generated HTML document (R_Course_UHasselt_demo_V1.html). The HTML page has a title "Demo 2" and author information "Rudradev Sengupta and Ziv Shkedy, 29 September 2020". Below the title is a section titled "R Markdown" with a descriptive paragraph and an R code chunk. A red box highlights the "R Markdown" section, and another red arrow points from the "R Markdown" text in the left pane to this highlighted area. The bottom of the right pane shows the generated HTML output for the R code chunk.

Demo 2

Rudradev Sengupta and Ziv Shkedy
29 September 2020

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both the embedded R code chunks within the document. You can embed an R code chunk like this:

```
project_start <- params$start
summary(mtcars)
```

mpg	cyl	disp	hp
Min. :10.40	Min. :4.000	Min. :71.1	Min. :52.0
1st Qu.:15.43	1st Qu.:4.000	1st Qu.:120.8	1st Qu.:96.5
Median :19.20	Median :6.000	Median :196.3	Median :123.0
Mean :20.09	Mean :6.188	Mean :230.7	Mean :146.7
3rd Qu.:22.80	3rd Qu.:8.000	3rd Qu.:326.0	3rd Qu.:180.0
Max. :33.90	Max. :8.000	Max. :472.0	Max. :335.0
drat	wt	qsec	vs
Min. :2.760	Min. :1.513	Min. :14.50	Min. :0.0000
1st Qu.:3.080	1st Qu.:2.581	1st Qu.:16.89	1st Qu.:0.0000
Median :3.695	Median :3.325	Median :17.71	Median :0.0000
Mean :3.597	Mean :3.217	Mean :17.85	Mean :0.4375
3rd Qu.:3.920	3rd Qu.:3.610	3rd Qu.:18.90	3rd Qu.:1.0000
Max. :4.930	Max. :5.424	Max. :22.90	Max. :1.0000
am	gear	carb	
Min. :0.0000	Min. :3.000	Min. :1.000	
1st Qu.:0.0000	1st Qu.:3.000	1st Qu.:2.000	
Median :0.0000	Median :4.000	Median :2.000	
Mean :0.4062	Mean :3.688	Mean :2.812	
3rd Qu.:1.0000	3rd Qu.:4.000	3rd Qu.:4.000	
Max. :1.0000	Max. :5.000	Max. :8.000	

Including project metadata:

See page 43

The output

EN 10:20 29/09/2020 40

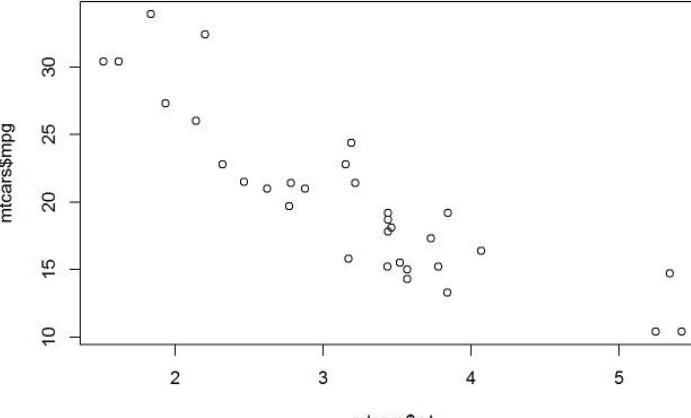
The HTML output

RMD Introduction.pptx - Microsoft PowerPoint
C:/projects/eR-Biostat/courses/introduction to R/R markdown/R_Course_UHasselt_demo_V1.html | Open in Browser | Find

Including Plots

You can also embed plots, for example:

```
plot(mtcars$wt, mtcars$mpg)
```



A scatter plot showing the relationship between car weight (wt) on the x-axis and fuel economy (mpg) on the y-axis. The x-axis ranges from approximately 2.0 to 5.0, and the y-axis ranges from 10 to 35. The data points show a negative correlation, with fuel economy generally decreasing as weight increases.

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

41 The Rmd program
42 Knit to HTML
43 Output
44 Output

Slide 44 of 55 | "Office Theme" | Dut

EN 12:09 30/09/2020 83%

The HTML output

RMD_Introduction.pptx - Microsoft PowerPoint

C:/projects/eR-Biostat/courses/introduction to R/R markdown/R_Course_UHasselt_demo_V1.html

Slides Outline

42 Knit to HTML

43 Output

44 The HTML output

45 The HTML output

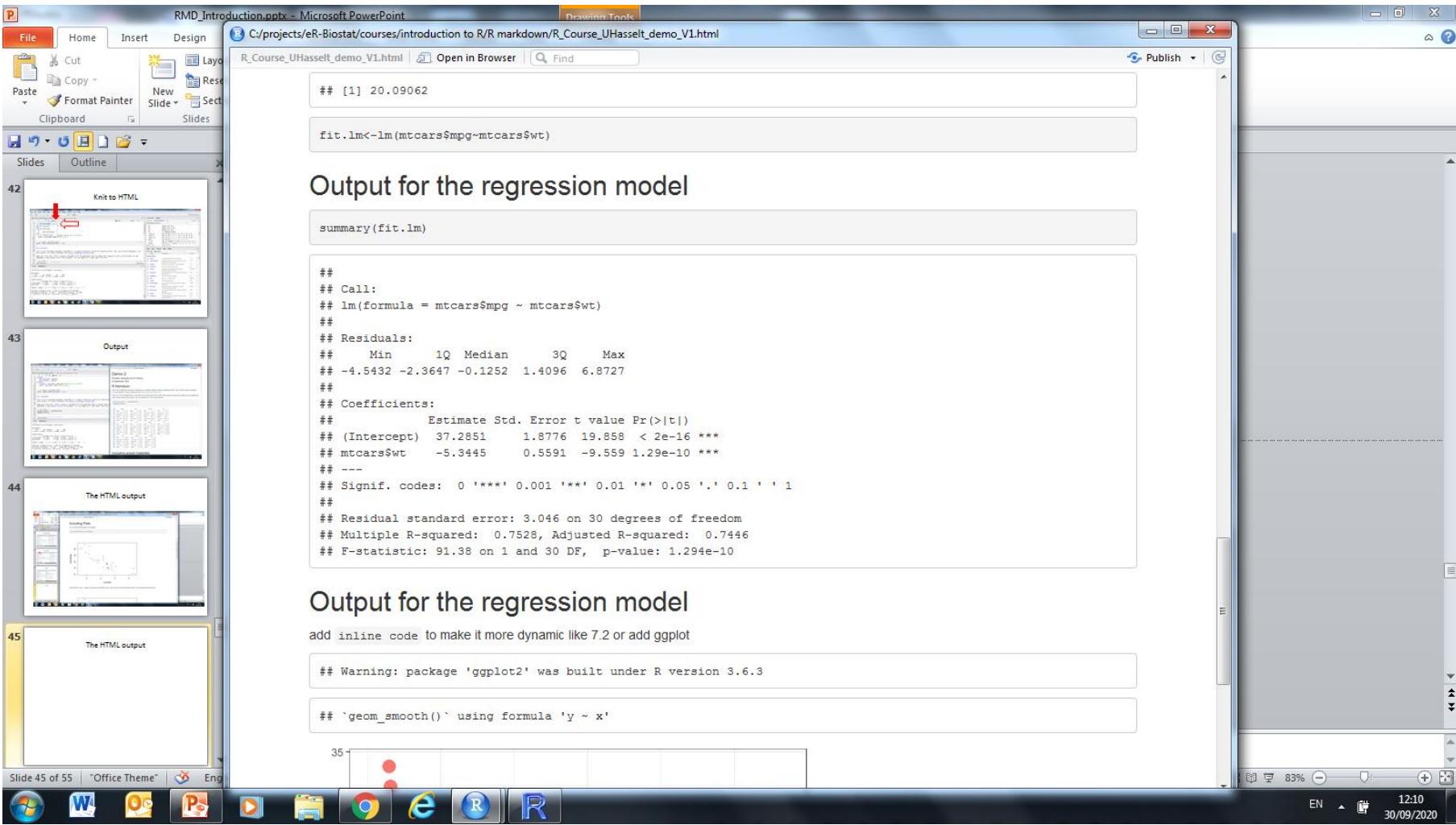
Output for the regression model

```
## [1] 20.09062  
fit.lm<-lm(mtcars$mpg~mtcars$wt)  
  
summary(fit.lm)  
  
##  
## Call:  
## lm(formula = mtcars$mpg ~ mtcars$wt)  
##  
## Residuals:  
## Min 1Q Median 3Q Max  
## -4.5432 -2.3647 -0.1252 1.4096 6.8727  
##  
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 37.2851 1.8776 19.858 < 2e-16 ***  
## mtcars$wt -5.3445 0.5591 -9.559 1.29e-10 ***  
## ---  
## Signif. codes: 0 '****' 0.001 '** 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 3.046 on 30 degrees of freedom  
## Multiple R-squared: 0.7528, Adjusted R-squared: 0.7446  
## F-statistic: 91.38 on 1 and 30 DF, p-value: 1.294e-10
```

Output for the regression model

add inline code to make it more dynamic like 7.2 or add ggplot

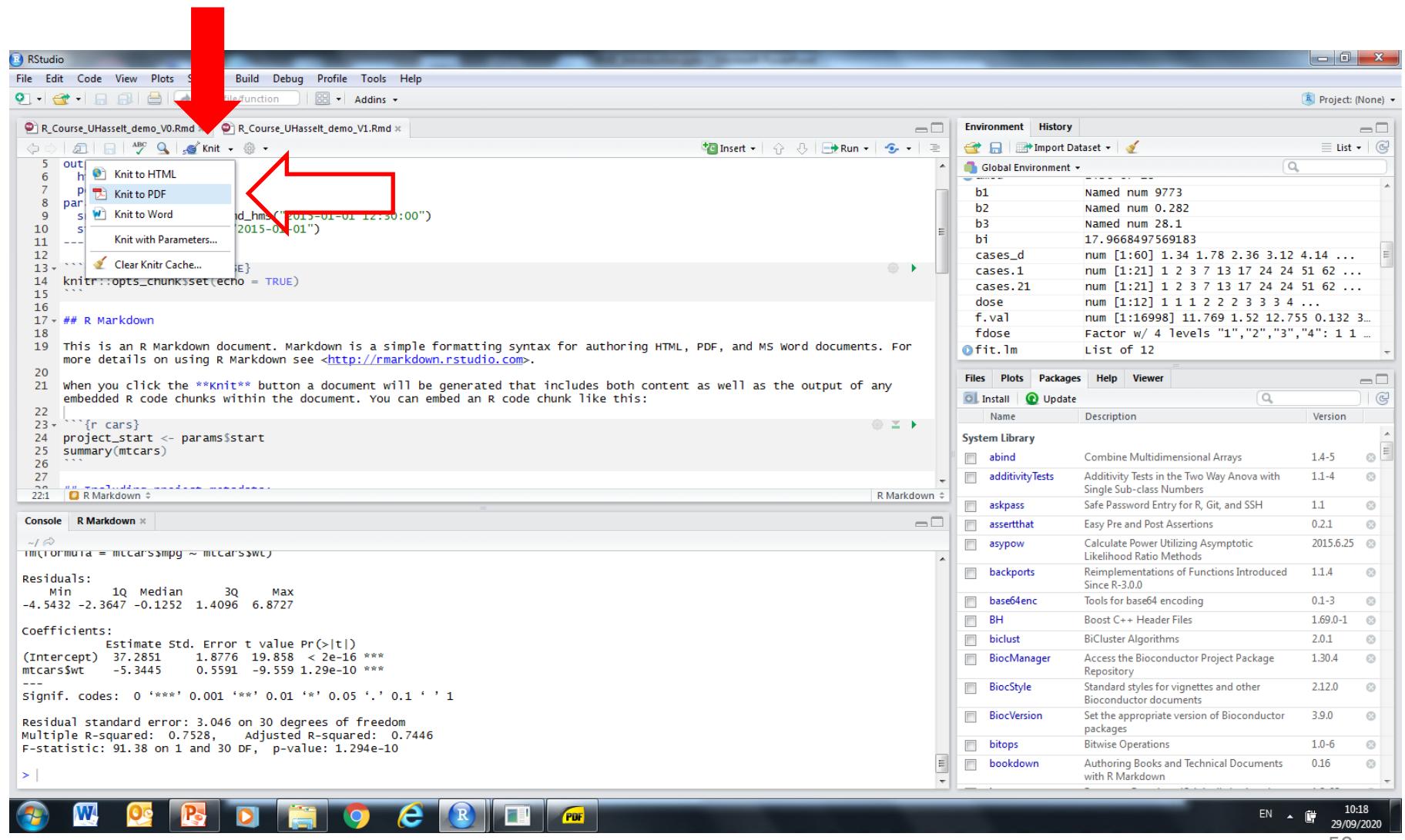
```
## Warning: package 'ggplot2' was built under R version 3.6.3  
  
## `geom_smooth()` using formula 'y ~ x'
```



The Rmd output

- Output in ONE document.
- Text and R output.
- R code can be included.
- Type of document: HTML/PDF.
- The data analyst design the output.

Knit to PDF



RStudio

File Edit Code View Plots Shiny Build Debug Profile Tools Help

Knit to function Addins

R_Course_UHasselt_demo_V0.Rmd R_Course_UHasselt_demo_V1.Rmd

Knit to HTML
Knit to PDF
Knit to Word
Knit with Parameters...
Clear Knitr Cache...
knitr::opts_chunk\$set(echo = TRUE)

```
5 out
6 h
7 par
8 s
9 id_hms("2015-01-01 12:30:00")
10 s
11 id_ymd("2015-01-01")
12 i
13 e}
14 knitr::opts_chunk$set(echo = TRUE)
15
16
17 ## R Markdown
18
19 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
20
21 when you click the **knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:
22
23 ``{r cars}
24 project_start <- params$start
25 summary(mtcars)
26
27
28 R Markdown
```

Console R Markdown

```
~ / ~
lm(formula = mtcars$mpg ~ mtcars$wt)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.5432 -2.3647 -0.1252  1.4096  6.8727 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 37.2851    1.8776 19.858 < 2e-16 ***
mtcars$wt   -5.3445    0.5591 -9.559 1.29e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.046 on 30 degrees of freedom
Multiple R-squared:  0.7528, Adjusted R-squared:  0.7446 
F-statistic: 91.38 on 1 and 30 DF, p-value: 1.294e-10
```

Environment History

Global Environment

Name	Description	Version
b1	Named num 9773	
b2	Named num 0.282	
b3	Named num 28.1	
bi	17.9668497569183	
cases_d	num [1:60] 1.34 1.78 2.36 3.12 4.14 ...	
cases_1	num [1:21] 1 2 3 7 13 17 24 24 51 62 ...	
cases_21	num [1:21] 1 2 3 7 13 17 24 24 51 62 ...	
dose	num [1:12] 1 1 1 2 2 2 3 3 3 4 ...	
f.val	num [1:16998] 11.769 1.52 12.755 0.132 3 ...	
fdose	Factor w/ 4 levels "1","2","3","4": 1 1 ...	
fit.lm	List of 12	

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-5
additivityTests	Additivity Tests in the Two Way Anova with Single Sub-class Numbers	1.1-4
askpass	Safe Password Entry for R, Git, and SSH	1.1
assertthat	Easy Pre and Post Assertions	0.2.1
asypow	Calculate Power Utilizing Asymptotic Likelihood Ratio Methods	2015.6.25
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.4
base64enc	Tools for base64 encoding	0.1-3
BH	Boost C++ Header Files	1.69.0-1
biclust	BiCluster Algorithms	2.0.1
BioManager	Access the Bioconductor Project Package Repository	1.30.4
BiocStyle	Standard styles for vignettes and other Bioconductor documents	2.12.0
BiocVersion	Set the appropriate version of Bioconductor packages	3.9.0
bitops	Bitwise Operations	1.0-6
bookdown	Authoring Books and Technical Documents with R Markdown	0.16

EN 10:18 29/09/2020 50

The PDF output

See page 43 for the title

The screenshot shows the RStudio interface with two files open: `R_Course_UHasselt_demo_V0.Rmd` and `R_Course_UHasselt_demo_V1.Rmd`. A red arrow points from the `R_Course_UHasselt_demo_V0.Rmd` file to the `R_Course_UHasselt_demo_V1.pdf` window. Another red arrow points from the `R_Course_UHasselt_demo_V1.Rmd` file to the `R_Course_UHasselt_demo_V1.pdf` window.

R_Course_UHasselt_demo_V0.Rmd:

```
5 output:
6   html_document: default
7   pdf_document: default
8 params:
9   snapshot: lubridate::ymd_hms("2015-01-01 12:30:00")
10  start: lubridate::ymd("2015-01-01")
11 ---
12
13 `r setup, include=FALSE}
14 knitr::opts_chunk$set(echo = TRUE)
15
16
17 ## R Markdown
18
19 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, more details on using R Markdown see <http://rmarkdown.rstudio.com>.
20
21 when you click the **Knit** button a document will be generated that includes both content embedded R code chunks within the document. You can embed an R code chunk like this:
22
23 `r cars`
24 project_start <- params$start
25 summary(mtcars)
26
27
22:1 R Markdown
```

R_Course_UHasselt_demo_V1.Rmd:

```
5 output:
6   html_document: default
7   pdf_document: default
8 params:
9   snapshot: lubridate::ymd_hms("2015-01-01 12:30:00")
10  start: lubridate::ymd("2015-01-01")
11 ---
12
13 `r setup, include=FALSE}
14 knitr::opts_chunk$set(echo = TRUE)
15
16
17 ## R Markdown
18
19 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, more details on using R Markdown see <http://rmarkdown.rstudio.com>.
20
21 when you click the **Knit** button a document will be generated that includes both content embedded R code chunks within the document. You can embed an R code chunk like this:
22
23 `r cars`
24 project_start <- params$start
25 summary(mtcars)
26
27
22:1 R Markdown
```

R_Course_UHasselt_demo_V1.pdf:

Demo 2
Rudradev Sengupta and Ziv Shkedy
29 September 2020

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will 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:

```
project_start <- params$start
summary(mtcars)
```

mpg	cyl	disp	hp
Min. :10.40	Min. :4.000	Min. :71.1	Min. :52.0
1st Qu.:15.43	1st Qu.:4.000	1st Qu.:120.8	1st Qu.:96.5
Median :19.20	Median :6.000	Median :196.3	Median :123.0
Mean :20.09	Mean :6.188	Mean :230.7	Mean :146.7
3rd Qu.:22.80	3rd Qu.:8.000	3rd Qu.:326.0	3rd Qu.:180.0
Max. :33.90	Max. :8.000	Max. :432.0	Max. :335.0

drat	wt	qsec	vs
Min. :-2.760	Min. :1.513	Min. :14.50	Min. :0.0000
1st Qu.: -0.800	1st Qu.: 2.581	1st Qu.: 16.89	1st Qu.: 0.0000
Median : -3.695	Median : 3.325	Median : 17.71	Median : 0.0000
Mean : -3.597	Mean : -3.217	Mean : 17.85	Mean : 0.4375
3rd Qu.: -3.920	3rd Qu.: 3.610	3rd Qu.: 18.90	3rd Qu.: 1.0000
Max. : -4.930	Max. : 5.424	Max. : 22.90	Max. : 1.0000

am	gear	carb
Min. :0.0000	Min. :3.000	Min. :1.000
1st Qu.:0.0000	1st Qu.:3.000	1st Qu.:2.000
Median :0.0000	Median :4.000	Median :2.000
Mean :0.4062	Mean :3.688	Mean :2.812
3rd Qu.:1.0000	3rd Qu.:4.000	3rd Qu.:4.000
Max. :1.0000	Max. :5.000	Max. :8.000

Including project metadata:

```
## [1] *Project Start Date: lubridate::ymd("2015-01-01")*
```

Including Plots

You can also embed plots, for example:

```
plot(mtcars$wt, mtcars$mpg)
```

The PDF output

RStudio

R_Course_UHasselt_demo_V0.Rmd

File Edit Code View Plots Session

R_Course_UHasselt_demo_V1.pdf

Page: 1 / 4

Find:

Bookmarks

- Including project metadata:
- Including Plots
- Short Analysis
- Output for the regression model
- Output for the regression model

Median :0.0000 Median :4.000 Median :2.000
Mean :0.4062 Mean :3.688 Mean :2.812
3rd Qu.:1.0000 3rd Qu.:4.000 3rd Qu.:4.000
Max. :1.0000 Max. :5.000 Max. :8.000

Including project metadata:

```
## [1] "Project Start Date: lubridate::ymd(\"2015-01-01\")"
```

Including Plots

You can also embed plots, for example:

```
plot(mtcars$wt, mtcars$mpg)
```

mtcars\$mpg

mtcars\$wt

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.

Project: (None)

bles

4 17.4 17.4 17.4 17.4 ...
24.4 24.4 21.3 20.8 ...
4 5 6 7 8 9 10 ...

les

les

les

bles

Version

1.4-5
anova with Single Sub- 1.1-4
d SSH 1.1
0.21
otic Likelihood Ratio 2015.6.25
introduced Since R- 1.1.4
0.1-3
1.69.0-1
2.0.1
Package Repository 1.30.4
ther Bioconductor 2.12.0
conductor packages 3.9.0
1.0-6
documents with R 0.16
Angelo Canty for S 1.3-23

EN 12:12 30/09/2020

The PDF output

The screenshot shows the RStudio interface. On the left, the 'Console' tab is active, displaying R version 3.6.1 (2019-07-05) and workspace details. In the center, a PDF document titled 'R_Course_UHasselt_demo_V1.pdf' is open, showing R Markdown code and its output. The output includes a histogram of 'mtcars\$wt', a 'Short Analysis' section with mean and lm results, and a 'Output for the regression model' section with a summary of the lm fit. A red text overlay 'The regression output' is placed over the regression results. On the right, the 'Project' pane shows a list of files and packages. The bottom of the screen shows the Windows taskbar with various icons.

R_Course_UHasselt_demo_V1.pdf

Bookmarks

- R Markdown
- Including project metadata:
- Including Plots
- Short Analysis
- Output for the regression model
- Output for the regression model

2 3 4 5

mtcars\$wt

Short Analysis

```
mean(mtcars$mpg)
## [1] 20.09062
fit.lm<-lm(mtcars$mpg~mtcars$wt)
```

Output for the regression model

```
summary(fit.lm)

##
## Call:
## lm(formula = mtcars$mpg ~ mtcars$wt)
##
## Residuals:
##    Min     1Q   Median     3Q    Max 
## -4.5432 -2.3647 -0.1252  1.4096  6.8727 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 37.2851    1.8776  19.858 < 2e-16 ***
## mtcars$wt   -5.3445    0.5591 - 9.559 1.29e-10 ***
## 
## ---
```

The regression output

3

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

bles

4 17.4 17.4 17.4 17.4 ...

24.4 24.4 21.3 20.8 ...

4 5 6 7 8 9 10 ...

les

les

les

les

Version

1.4-5

anova with Single Sub-

1.1-4

1.1

0.21

Optic Likelihood Ratio

2015.6.25

Introduced Since R-

1.1.4

0.1-3

1.69.0-1

2.0.1

Package Repository

1.30.4

other Bioconductor

2.12.0

conductor packages

3.9.0

1.0-6

documents with R

0.16

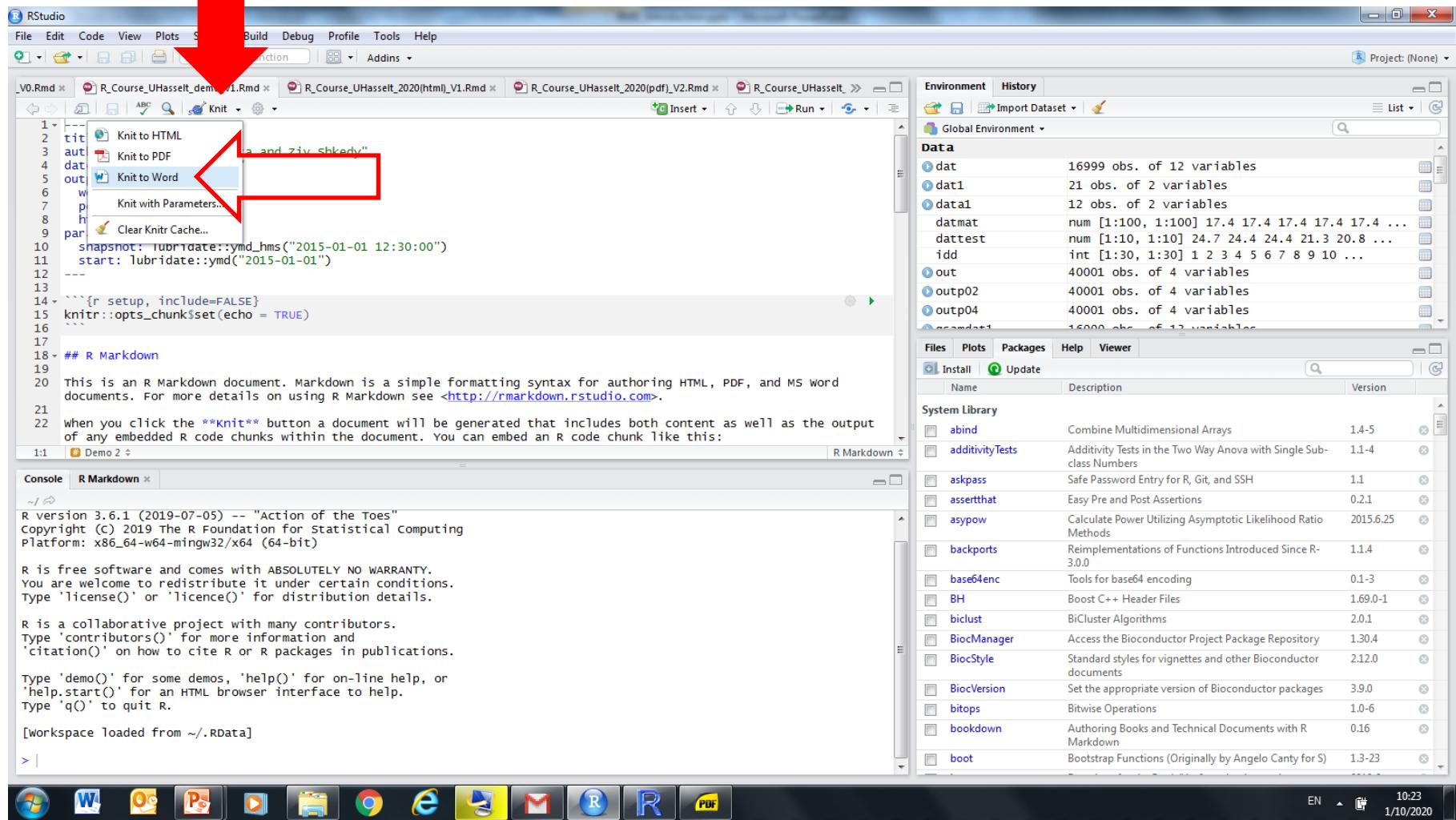
Angelo Canty for S

1.3-23

EN 12:12 30/09/2020

Important: To produce the PDF file, you will need to install **LaTex** or **Tex**

Knit to word



The screenshot shows the RStudio interface. In the top-left corner, there is a red arrow pointing down at the 'Knit' button in the toolbar. A dropdown menu is open from this button, showing several options: 'Knit to HTML', 'Knit to PDF', 'Knit to Word' (which is highlighted with a blue selection bar), and 'Knit with Parameters...'. Below the dropdown, the R Markdown code editor shows a snippet of R code and its corresponding Markdown output. The code includes a 'knitr::opts_chunk\$set(echo = TRUE)' line. The R Markdown pane also contains a note about generating documents with both content and R code chunks. The bottom-left pane is the 'Console', displaying the standard R startup message and license information. On the right side of the interface, there is a 'Data' viewer pane showing various R objects like 'dat', 'out', and 'idd' with their respective data structures. At the very bottom, a taskbar displays icons for various applications including Windows, Word, Excel, and Google Chrome.

The word doc output

A screenshot of Microsoft Word showing an R Markdown document. The document title is "Demo 2" and it includes authors "Rudradev Sengupta and Ziv Shkedy" and date "29 September 2020". It features an "R Markdown" section with a code chunk that prints the summary of the mtcars dataset. To the right, there is a scatter plot titled "plot(mtcars\$wt, mtcars\$mpg)" showing the relationship between weight (wt) and miles per gallon (mpg). A note below the plot states: "Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot." The Word ribbon at the top includes tabs like File, Home, Insert, Page Layout, References, Mailings, Review, and View. The ribbon also shows various font and style options, including a "AaBbCcDd" button which is highlighted with a yellow box.

Example 2

Output development

Focus

- How to develop an output document:
 - HTML.
 - PDF/HTML.
- Sections/Subsections.
- R code as a part of the text or not ?

The program

- A simple user guide for a document that includes:
 - Free text.
 - R code.
 - Graphical displays.
- To produce the document: run the programs in R Studio.

Development of a **HTML** document

- The output:
 - HTML document.

To run the analysis in the example, use the program in R Studio:

[R_course_UHasselt_2021\(html\)_V1.Rmd](#)

The program

R_course_UHasselt_2020(html)_V1.Rmd

A screenshot of the RStudio interface. On the left, the code editor shows an R Markdown file named "R_course_UHasselt_2020(html)_V1.Rmd". The first few lines of the file are the YAML header:

```
1 ---  
2 title: "Basic Skills in R Markdown"  
3 author:  
4 - name: ziv shkedy  
5   affiliation: Hasselt University, Belgium  
6 subtitle: The HTML file (2020)  
7  
8 #institute: UHasselt  
9 date: September, 29, 2020  
10 output:  
11   rmdformats::readthedown:  
12     highlight: kate  
13     use_bookdown: TRUE  
14 vignette: >  
15   %\VignetteEngine{knitr::rmarkdown}  
16 editor_options:  
17   chunk_output_type: console  
18 categories: ["R"]  
19 bibliography: bibliography.bib  
20 ---
```

A red arrow points from the top of the slide down to the "Knit" button in the toolbar. A red curly brace on the right side of the screen groups the first few lines of the file under the heading "YAML Header".

The RStudio interface includes a toolbar, a menu bar (File, Edit, Code, View, Plots, etc.), and several panes on the right: Environment, Data, Files, Plots, Packages, Help, and Viewer. The "Data" pane lists various R objects like dat, dat1, data1, datmat, dattest, idd, out, outp02, outp04, and scandat1. The "Packages" pane shows the System Library with numerous packages listed.

At the bottom, the R console displays the standard R startup message and workspace information.

System tray icons are visible at the very bottom of the screen.

The HTML output

The screenshot shows a Microsoft PowerPoint slide titled "Basic Skills in R Markdown". The slide content includes:

- Section 1: Basic Skills in R Markdown**
- Section 2: The HTML file (2020)**
 - ## Warning: package 'ggplot2' was built under R version 3.6.3
 - ## Warning: package 'mvtnorm' was built under R version 3.6.2
- Section 3: 1 Introduction**

This document provides basic tools to produce a html file using R markdown. The best way to use this document is to run the file in R studio and then read the .Rmd file to see how the output was created. The file can be used to produce a very basic html document an you can add later more components to you document.
- Section 4: 2 Sections and subsections**

This is a an example of a R markdown file that produces htnl output. This is a section in the document.
- Section 5: 2.1 Subsection**

This text appears in a subsection
- Section 6: 2.1.1 Subsubsection**

This text is a part of a subsection.
- Section 7: 3 Including R code**

On the left, the slide navigation pane shows slides 57 through 60. Slides 57 and 58 show screenshots of RStudio. Slide 59 is titled "The HTML output". Slide 60 is titled "The program". The bottom of the slide shows the author's information: Ziv Shkedy, Hasselt University, Belgium, and the date: September, 29, 2020.

**Very basic guide
how to produce a
markdown
document**

The program

An example how to create section/subsection/subsubsection...

Free text

R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

[workspace loaded from ~/RData]

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)

Global Environment

Environment History

Import Dataset

Insert Run

49 50 This document provides basic tools to produce a html file using R markdown. The best way to use this document is to run the file in R studio and then read the <tt>.Rmd</tt> file to see how the output was created. The file can be used to produce a very basic html document an you can add later more components to you document.

51 52 # Sections and subsections

53 54 This is a an example of a R markdown file that produces html output. This is a section in the document.

55 56 ## Subsection

57 58 This text appears in a subsection

59 60 ### Subsubsection

61 62 This text is a part of a subsection.

63 64

65 66 # Including R code

67 68 ## Print R code and output

69 70 This is an example how to include R code and output in the document. We use the <tt>airquality</tt> data as an

16:17 Basic Skills in R Markdown R Markdown

Console R Markdown

~ /

W P S E R PDF

12:30
30/09/2020

Package	Description	Version
additivityTests	Additivity Tests in the Two Way Anova with Single Sub-class Numbers	1.1-4
askpass	Safe Password Entry for R, Git, and SSH	1.1
assertthat	Easy Pre and Post Assertions	0.2.1
asypow	Calculate Power Utilizing Asymptotic Likelihood Ratio Methods	2015.6.25
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.4
base64enc	Tools for base64 encoding	0.1-3
BH	Boost C++ Header Files	1.69.0-1
bclust	BiCluster Algorithms	2.0.1
BiocManager	Access the Bioconductor Project Package Repository	1.30.4
BiocStyle	Standard styles for vignettes and other Bioconductor documents	2.12.0
BiocVersion	Set the appropriate version of Bioconductor packages	3.9.0
bitops	Bitwise Operations	1.0-6
bookdown	Authoring Books and Technical Documents with R Markdown	0.16
boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-23

The output

The screenshot shows the RStudio interface with an R Markdown file open in the left panel and its generated HTML output in the right panel.

Left Panel (R Markdown Editor):

- File: R_Course_UHasselt_demo_V0.Rmd
- Code:

```
1 ---  
2 title: "Basic skills"  
3 author:  
4 - name: Ziv Shkedy  
5 affiliation: Hasselt  
6 subtitle: The HTML file  
7  
8 #institute: UHasselt  
9 #date: September, 29, 2020  
10 output:  
11   rmdformats::readthedocs()  
12   highlight: kate  
13   use_bookdown: TRUE  
14 vignette: >  
15   %\VignetteEngine{knitr}  
16 editor_options:  
17   chunk_output_type: inline  
18 categories: ['R']  
19 bibliography: bibliography  
20 ---  
21  
22  
23 <script type="text/javascript">  
24   src="http://cdn.mathjax.org/mathjax/latest/MathJax.js?config=TeX-AMS-MML_HTMLorMML"</script>
```

- Console:

```
R version 3.6.1 (2019-07-05)  
Copyright (C) 2019 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for more information.  
  
R is a collaborative project.  
Type 'contributors()' for more information.  
Type 'citation()' on how to cite R or R packages.  
  
Type 'demo()' for some demos,  
'help.start()' for an HTML browser-based help system,  
Type 'q()' to quit R.
```

- Workspace:

```
[workspace loaded from ~/.RData]
```

- Bottom icons: Windows, Word, Excel, Powerpoint, YouTube, File Explorer, Google Chrome, Internet Explorer, R logo, PDF.

Right Panel (HTML Preview):

Basic Skills in R Markdown

The HTML file (2020)

```
## Warning: package 'ggplot2' was built under R version 3.6.3  
  
## Warning: package 'mvtnorm' was built under R version 3.6.2
```

1 Introduction

This document provides basic tools to produce a html file using R markdown. The best way to use this document is to run the file in R studio and then read the .Rmd file to see how the output was created. The file can be used to produce a very basic html document and you can add later more components to your document.

2 Sections and subsections

This is an example of a R markdown file that produces html output. This is a section in the document.

2.1 Subsection

This text appears in a subsection

2.1.1 Subsubsection

This text is a part of a subsection.

3 Including R code

R code in the output

The screenshot shows the RStudio interface. In the top-left pane, there are three tabs: 'R_Course_UHasselt_demo_V0.Rmd' (active), 'R_Course_UHasselt_demo_V1.Rmd', and 'R_Course_UHasselt_2020(html)_V1.Rmd'. The main workspace displays R code and its output. A red arrow points from the text 'How to include R code and output' to the code block starting at line 63.

```
63  
64  
65  
66 # Including R code  
67  
68 ## Print R code and output  
69  
70 This is an example how to include R code and output in the document. We use the <tt>airquality</tt> data as an example.  
71  
72  
73 ````{r}  
74 x<-na.omit(airquality$ozone)  
75 print(x)  
76  
77  
78 ## Do not print the R code but print the output  
79  
80 If we do not want to print the R code, but we want to see the output use the option <tt>echo=FALSE</tt>:  
81  
82 ````{r,echo=FALSE}  
83 x<-na.omit(airquality$ozone)  
84 print(x)  
85
```

In the bottom-left pane, the 'Console' tab is active, showing the R startup message and help information. The bottom-right pane shows the 'System Library' with a list of packages and their versions.

How to include R code and output

System Library	Description	Version
abind	Combine Multidimensional Arrays	1.4-5
additivityTests	Additivity Tests in the Two Way Anova with Single Sub-class Numbers	1.1-4
askpass	Safe Password Entry for R, Git, and SSH	1.1
assertthat	Easy Pre and Post Assertions	0.2.1
asypow	Calculate Power Utilizing Asymptotic Likelihood Ratio Methods	2015.6.25
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.4
base64enc	Tools for base64 encoding	0.1-3
BH	Boost C++ Header Files	1.69.0-1
biclust	BiCluster Algorithms	2.0.1
BiocManager	Access the Bioconductor Project Package Repository	1.30.4
BiocStyle	Standard styles for vignettes and other Bioconductor documents	2.12.0
BiocVersion	Set the appropriate version of Bioconductor packages	3.9.0
bitops	Bitwise Operations	1.0-6
bookdown	Authoring Books and Technical Documents with R Markdown	0.16
boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-23

The output

The screenshot shows the RStudio interface with an R Markdown file open. The left pane displays the R Markdown code, and the right pane shows the generated HTML output. A red box highlights the section where R code is included.

R Markdown Code:

```
59
60  ## Subsubsection
61
62 This text is a part of
63
64
65
66 # Including R code
67
68 ## Print R code and ou
69
70 This is an example how
example.
71
72
73 ``{r}
74 x<-na.omit(airquality)
75 print(x)
76
77
78 ## Do not print the R
79
80 If we do not want to p
81
```

HTML Output:

Basic Skills in R Markdown

1 Introduction

2 Sections and subsections

2.1 Subsection

2.1.1 Subsubsection

This text appears in a subsection

This text is a part of a subsection.

3 Including R code

3.1 Print R code and output

This is an example how to include R code and output in the document. We use the `airquality` data as an example.

```
x<-na.omit(airquality$Ozone)
print(x)
```

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81																																																																																							
##	[1]	41	36	12	18	28	23	19	8	7	16	11	14	18	14	34	6	30	11	1	11	4	32	##	[24]	45	115	37	29	71	39	23	21	37	20	12	13	135	49	32	64	40	77	97	97	85	10	##	[47]	7	48	35	61	79	63	16	80	108	20	52	82	50	64	59	39	9	16	78	35	66	122	##	[70]	110	44	28	65	22	59	23	31	44	21	9	45	168	73	76	118	84	85	96	78	73	91	##	[93]	32	20	23	21	24	44	21	28	9	13	46	18	13	24	16	13	23	36	7	14	30	14	##	[116]	20	##	attr(,"na.action")	##	[1]	5	10	25	26	27	32	33	34	35	36	37	39	42	43	45	46	52	53	54	55	56	57	##	[25]	60	61	65	72	75	83	84	102	103	107	115	119	150	##	attr(,"class")	##	[1]	"omit"

3.2 Do not print the R code but print the output

Development of a PDF/HTML document

- The output:
 - PDF/HTML document (the user can choose).

To run the analysis in the example, use the program in R Studio:

[R_course_UHasselt_2021\(pdf\)_V1.Rmd](#)

The program

R_course_UHasselt_2021(pdf)_V1.Rmd

Produce a PDF output

RStudio
File Edit Code View Plots Search Build Debug Profile Tools Help
Project: (None)

_Course_UHasselt_demo_V0.Rmd × _Course_UHasselt_demo_V1.Rmd × R_Course_UHasselt_2020(html)_V1.Rmd × R_Course_UHasselt_2020(pdf)_V1.Rmd ×

Knit ▾

- Knit to HTML
- Knit to PDF
- Knit to Word
- Knit with Parameters...
- Clear Knitr Cache...

R Markdown: the pdf file

```
1 --  
2 ti  
3 ou  
4 Knit to PDF  
5 Knit with Parameters...  
6 su  
7 la  
8 to  
9 to  
10 ---  
11 ---  
12 output: html_document  
13 use_bookdown: TRUE  
14 ---  
15  
16 ``{r,echo=FALSE}  
17 library(e1071)  
18 library(lattice)  
19 library(ggplot2)  
20 library(mvtnorm)  
21 ````  
22  
23 \newpage  
24  
16:18 [Chunk 1]
```

Console R Markdown ×

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"  
Copyright (C) 2019 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.

[workspace loaded from ~/.RData]
```

Environment History

Global Environment

Name	Description	Version
dat	16999 obs. of 12 variables	
dat1	21 obs. of 2 variables	
data1	12 obs. of 2 variables	
datmat	num [1:100, 1:100] 17.4 17.4 17.4 17.4 17.4 ...	
dattest	num [1:10, 1:10] 24.7 24.4 24.4 21.3 20.8 ...	
idd	int [1:30, 1:30] 1 2 3 4 5 6 7 8 9 10 ...	
out	40001 obs. of 4 variables	
outp02	40001 obs. of 4 variables	
outp04	40001 obs. of 4 variables	
scandata1	16000 obs. of 12 variables	

Files Plots Packages Help Viewer

System Library

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-5
additivityTests	Additivity Tests in the Two Way Anova with Single Sub-class Numbers	1.1-4
askpass	Safe Password Entry for R, Git, and SSH	1.1
assertthat	Easy Pre and Post Assertions	0.2.1
asympow	Calculate Power Utilizing Asymptotic Likelihood Ratio Methods	2015.6.25
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.4
base64enc	Tools for base64 encoding	0.1-3
BH	Boost C++ Header Files	1.69.0-1
bicluster	BiCluster Algorithms	2.0.1
BiocManager	Access the Bioconductor Project Package Repository	1.30.4
BiocStyle	Standard styles for vignettes and other Bioconductor documents	2.12.0
BiocVersion	Set the appropriate version of Bioconductor packages	3.9.0
bitops	Bitwise Operations	1.0-6
bookdown	Authoring Books and Technical Documents with R Markdown	0.16
boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-23

EN 12:39 30/09/2020

The PDF output

The screenshot shows the RStudio environment with three main panes:

- Left pane (Code Editor):** Displays the R Markdown code for "Basic Skills in R Markdown: the pdf file". The code includes YAML front matter, R code chunks, and a new page command.
- Middle pane (PDF Preview):** Shows the generated PDF titled "Basic Skills in R Markdown: the pdf file" by Ziv Shkedy (2020). The PDF contains a table of contents and several sections of text and code examples.
- Right pane (Dataset View):** Displays a dataset with 16999 observations and 12 variables, including numerical and integer columns.

Table of contents from the PDF:

Contents	Page
Introduction	2
Sections and subsections	2
Subsection	2
Subsubsection	2
Including R code	2
Print R code and output	2
Do not print the R code but print the output	2
Items	3
Use R as a part of your text	3
Example: summary statistics	3
Graphical displays in the document	3
How to add a link to your document	4
How to create a math formula	4
Just do it	7
Analysis of the cars data	7
Expected output	7

Text from the PDF:

```
Basic Skills in R Markdown: the pdf file  
Ziv Shkedy (2020)
```

Text from the R Console:

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"  
Copyright (C) 2019 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,  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]
```

Bottom navigation bar:

EN 12:39 30/09/2020

Text overlay:

Table of contents: see next slide

The program

A screenshot of the RStudio interface. The left pane shows an R Markdown file with the following code:

```
1 ---  
2 title: 'Basic skills in R Markdown: the pdf file'  
3 output:  
4   pdf_document: default  
5   html_document: default  
6   word_document: default  
7 subtitle: zivshkedy (2020)  
8 layout: page  
9 toc: yes  
10 ---  
11 ---  
12 output: html_document  
13 use_bookdown: TRUE  
14 ---  
15  
16 ```{r,echo=FALSE}  
17 library(e1071)  
18 library(lattice)  
19 library(ggplot2)  
20 library(mvtnorm)  
21 ````  
22  
23 \newpage  
24
```

A large red arrow points to the line `toc: yes`. To the right of the arrow, the text "toc: yes Produce the table of contents" is displayed in red.

The right pane shows the Environment and Global Environment panes. The Global Environment pane lists various objects:

Object	Description
dat	16999 obs. of 12 variables
dat1	21 obs. of 2 variables
data1	12 obs. of 2 variables
datmat	num [1:100, 1:100] 17.4 17.4 17.4 17.4 17.4 ...
dattest	num [1:10, 1:10] 24.7 24.4 24.4 21.3 20.8 ...
idd	int [1:30, 1:30] 1 2 3 4 5 6 7 8 9 10 ...
out	40001 obs. of 4 variables
outp02	40001 obs. of 4 variables
outp04	40001 obs. of 4 variables
scandata	16000 obs. of 12 variables

The bottom of the screen shows the R console output:

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"  
Copyright (C) 2019 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.

[workspace loaded from ~/.RData]
```

The status bar at the bottom right shows the date and time: 30/09/2020 12:40 EN

The program

A section that explains how to include a figure in the document.

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project (None) is selected.
- Code Editor:** Three files are listed: _Course_UHasselt_demo_V0.Rmd, R_Course_UHasselt_demo_V1.Rmd, and R_Course_UHasselt_2020(html)_V1.Rmd. The R_Course_UHasselt_2020(html)_V1.Rmd file is open, showing R code for generating a histogram and adding a caption.
- Code Content:**

```
90 var(x)
91 ...
92
93 ## Graphical displays in the document
94
95 A histogram for the ozone level can be produced using the function <tt>qplot</tt>:
96 "histogram"</tt>;
97 ````{r}
98 Ozone.R<-data.frame(x)
99 qplot(x, data = Ozone.R, geom = "histogram", binwidth = 0.1)
100
101
102 To add a caption to the figure we use <tt>(r figchp1,fig.cap="Sepal length. .... . . . .
103
104 ````{r figchp1,fig.cap="Sepal length (III)"}
105 Ozone.R<-data.frame(x)
106 qplot(x, data = Ozone.R, geom = "histogram", binwidth = 0.1)
107
108
109 We can refer to the figure from the text in the document. For example, Figure \@ref(fig:figchp2) presents a
110 histogram that was produced using the function <tt>qplot()</tt> function.
```
- Console:** Displays the R startup message and workspace information.
- Environment:** Shows the 'bioconductor' package version 1.5000.0, which contains 12 variables.
- Packages:** A table showing available packages in the System Library, including abind, additivityTests, askpass, assertthat, asympow, backports, base64enc, BH, biclust, BiocManager, BiocStyle, BiocVersion, bitops, bookdown, and boot.
- Bottom Bar:** Icons for various applications like Word, Excel, and R.
- Status Bar:** Shows the language (EN), date (30/09/2020), and time (12:43).

The PDF output

A section that explains how to include a figure in the document.

The screenshot shows the RStudio interface. On the left, the code editor displays R code for generating a histogram. The code includes a section titled "## Graphical displays in the document" and uses the qplot function with geom = "histogram". The R console at the bottom shows the R version information and a workspace message. In the center, a PDF viewer window titled "R_Course_UHasselt_2020-pdf_V1.pdf" shows the generated histogram. The PDF viewer has a toolbar with icons for file operations, page navigation, and search. The right side of the screen shows the RStudio project browser and file tree.

```
90 var(x)
91
92
93 ## Graphical displays in the document
94
95 A histogram for the Ozone level can be produced us
96 "histogram"</tt>:
97
98 ``{r}
99 Ozone.R<-data.frame(x)
100 qplot(x, data = Ozone.R, geom = "histogram", binwidth = 0.1)
101
102 To add a caption to the figure we use <tt>{r figch
103
104 ``{r figchp1,fig.cap="Sepal length (III)"}
105 Ozone.R<-data.frame(x)
106 qplot(x, data = Ozone.R, geom = "histogram", binwi
107
108
109
110 We can refer to the figure from the text in the do
111 histogram that was produced using the function <tt>
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
```

R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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,
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/RData]

EN 12:42 30/09/2020

Example 3

Very simple analysis

To run the analysis in the example, use the program in R Studio:

[R_course_UHasselt_2021\(pdf\)_V2.Rmd](#)

Produce a report for an analysis

- How to use Rmd to produce a report about an analysis ?
- Why this is important ?
- Example: the old faithful dataset.
 - Part 1: the analysis.
 - Part 2: the analysis + report using Rmd (you need to run the program [R_course_UHasselt_2020\(pdf\)_V2.Rmd](#) after the class to see the report.)

Exploratory analysis of the Old Faithful dataset

- Old Faithful is a geyser that is found in Yellowstone National Park, in Wyoming.
- It is one of the most famous attractions at Yellowstone National Park.
- Old Faithful is unique because of how long and how often it erupts.



Exploratory analysis of the Old Faithful dataset

- Our aim is to explore the association between the time between eruptions (waiting time) and the duration of the eruption.
- For the analysis we use **basic graphical R functions** and **functions to calculate descriptive statistics** for the data.
- Output: PDF format.

The data in R

head(faithful)

```
## eruptions waiting  
## 1 3.600 79  
## 2 1.800 54  
## 3 3.333 74  
## 4 2.283 62  
## 5 4.533 85  
## 6 2.883 55
```



The name of the data in R

A data frame with two variables:
eruption and waiting time.

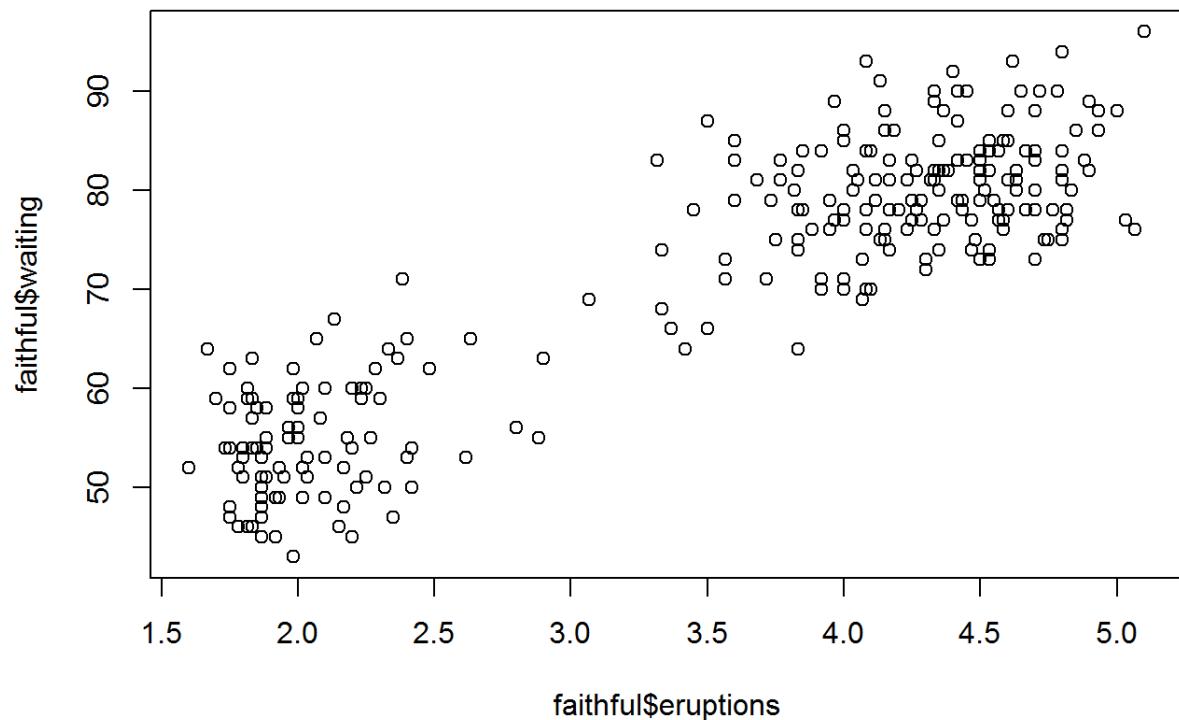
Eruption time and waiting time

```
plot(faithful$eruptions,faithful$waiting)
```



Basic graphical
function in R

```
plot(x,y)
```



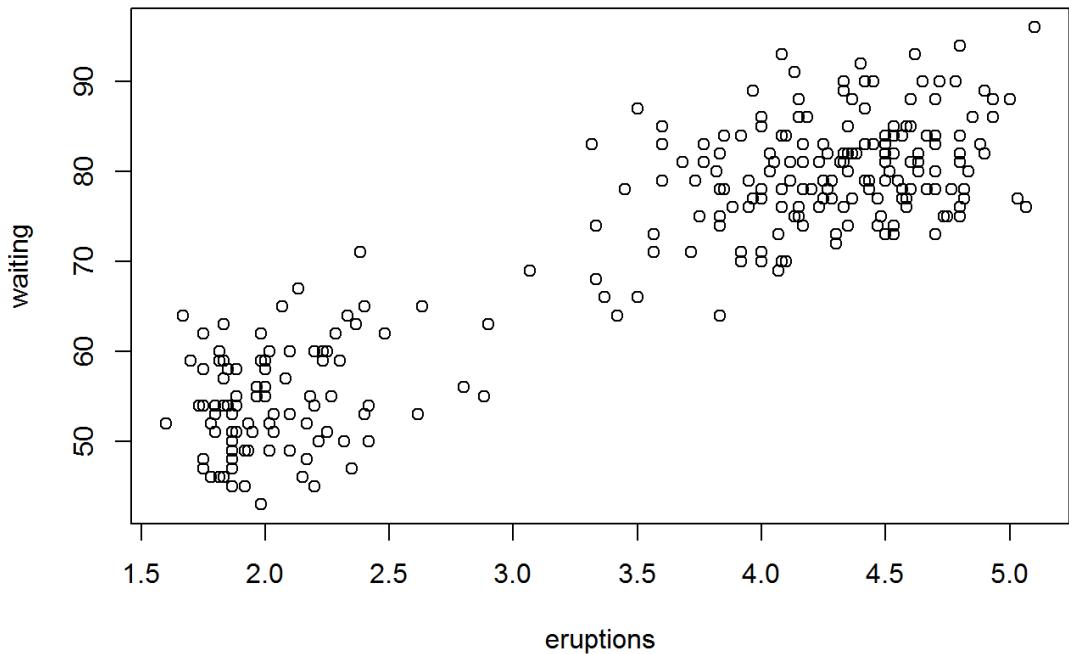
Eruption time and waiting time

```
plot(faithful$eruptions,faithful$waiting, xlab="eruptions",ylab="waiting")
```

```
title("The old faithful data")
```

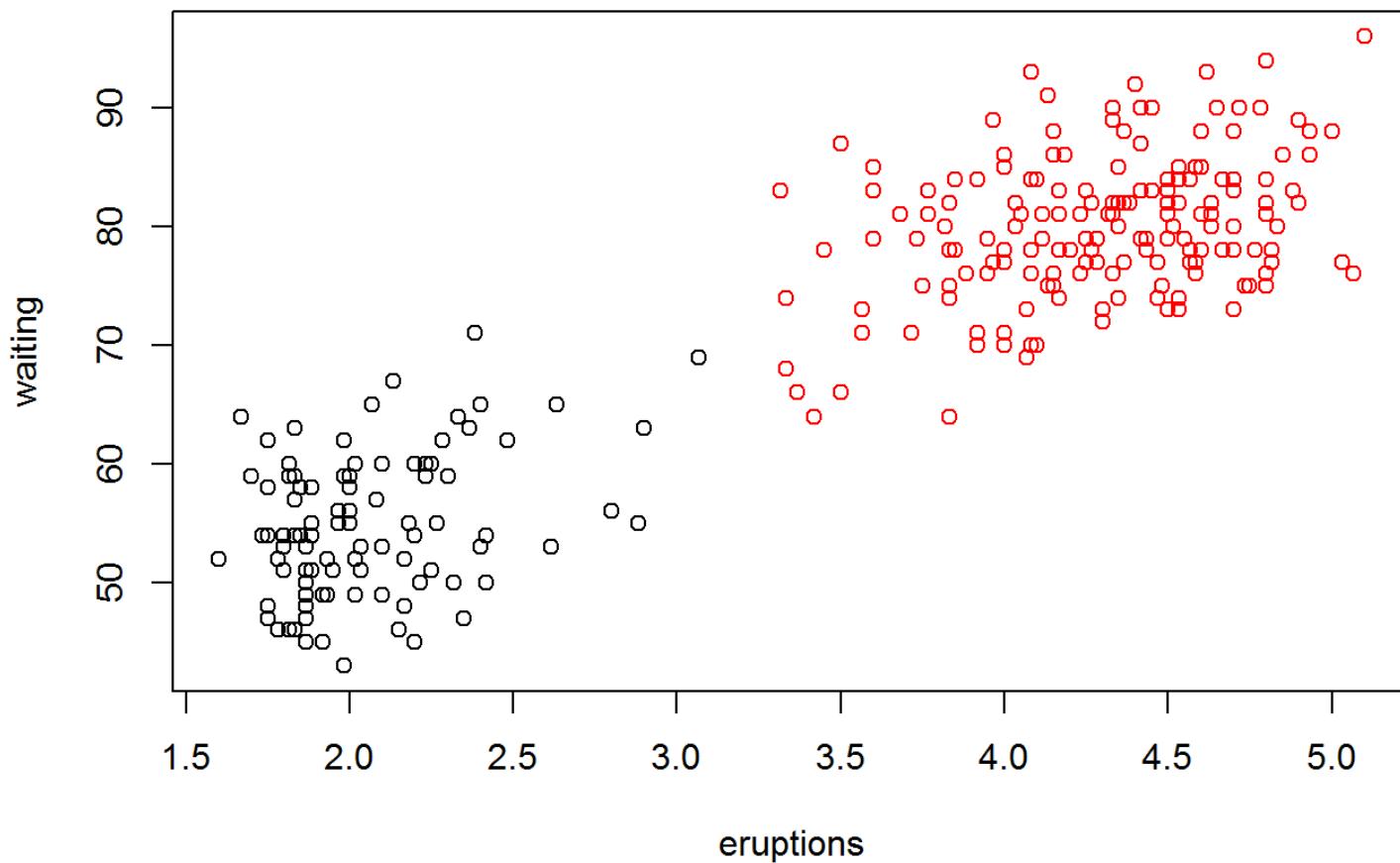
← The old faithful data →

Adding title and text
for the labels.



Two clusters ??

The old faithful data



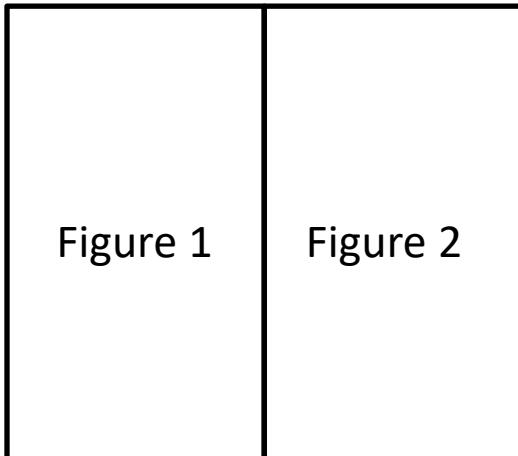
Distribution of eruption time

```
par(mfrow=c(1,2))
```

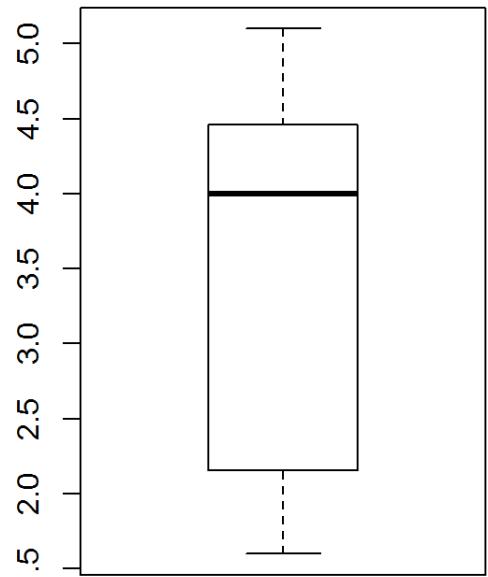
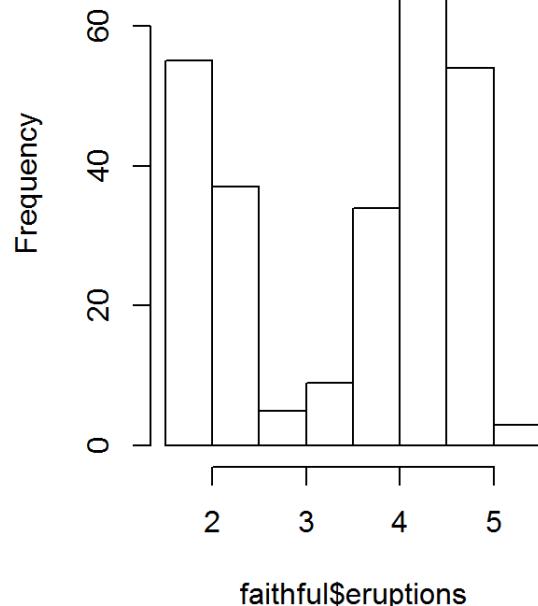


```
hist(faithful$eruptions)  
boxplot(faithful$eruptions)
```

Split the graphical window



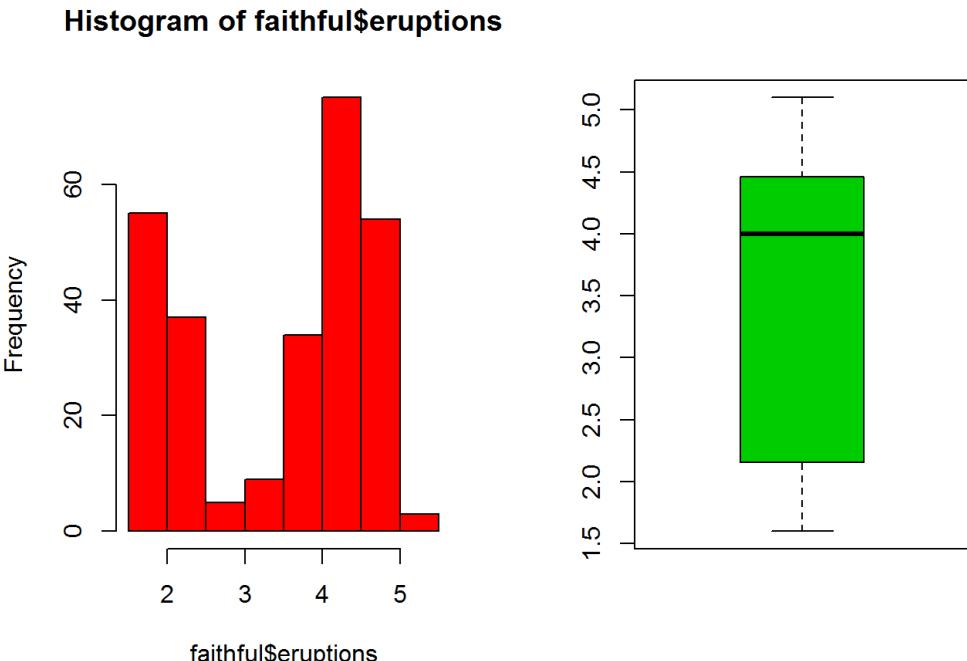
Histogram of faithful\$eruptions



Distribution of eruption time

```
par(mfrow=c(1,2))  
hist(faithful$eruptions, col=2)  
boxplot(faithful$eruptions,col=3)
```

Change plot settings.

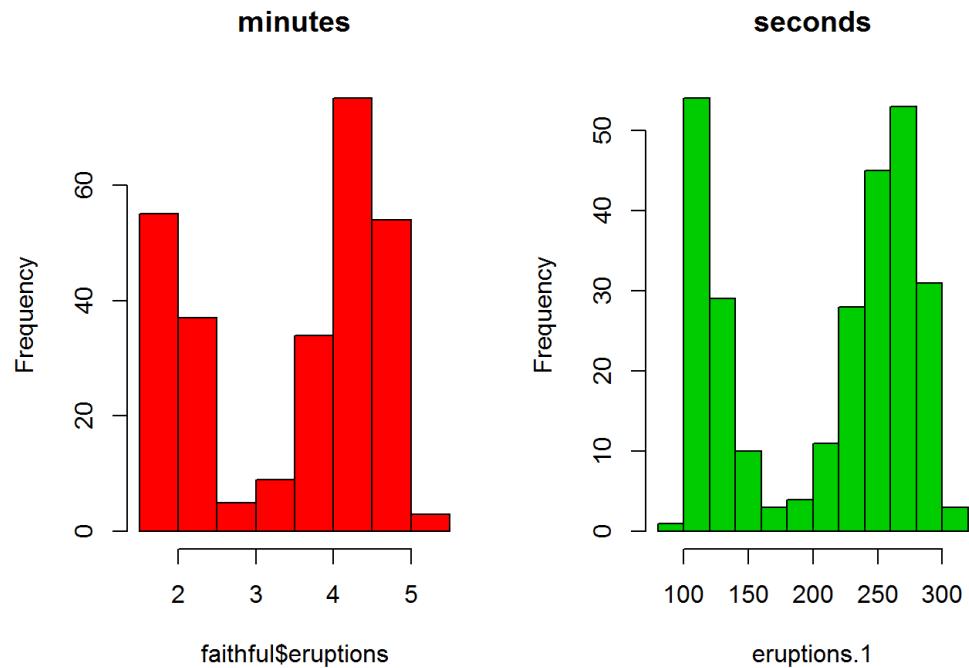


Distribution of eruption time in seconds

```
eruptions.1<-faithful$eruptions*60  
par(mfrow=c(1,2))  
hist(faithful$eruptions, col=2,main="minutes")  
hist(eruptions.1, col=3,main="seconds")
```

Create a new object:

```
eruptions.1
```



The Rmd program

- Repeat the analysis above.
- Produce a report for the analysis.
- Output: PDF format.
- Run the analysis in your own laptop to see the report.
- Rmd program:

R_course_UHasselt_2021(pdf)_V2.Rmd

The R markdown program & PDF output

R_course_UHasselt_2021(pdf)_V2.Rmd

The screenshot shows the RStudio interface with two main panes. The left pane displays an R Markdown file named R_course_UHasselt_2021(pdf)_V2.Rmd. The right pane shows the generated PDF output.

Code (Left Pane):

```
25  
26 # Introduction  
27  
28 Old Faithful is a geyser that is found in Yellowstone National Park, in wyom  
attractions at Yellowstone National Park. old Faithful is unique because of  
aim is to explore the time between eruptions and the duration of the eruption  
graphical R functions and functions to calculate descriptive statistics for  
29  
30  
31 # The data|  
32  
33 Data recorded from 1990 measured the time between eruptions and the duration  
minutes. This data set was based on 272 observations of old Faithful's erupt  
eruption time in mins and waiting time to next eruption (in mins). Both vari  
34  
35 ``{r}  
36 head(faithful)  
37 ```  
38  
39  
40  
41 # Analysis of the old faithful data  
42  
43 ## Eruption time and waiting time  
31:11 The data ↴
```

PDF Output (Right Pane):

Introduction

Old Faithful is a geyser that is found in Yellowstone National Park, in Wyoming. It is one of the most famous attractions at Yellowstone National Park. Old Faithful is unique because of how long and how often it erupts. Our aim is to explore the time between eruptions and the duration of the eruption. For the analysis we use basic graphical R functions and functions to calculate descriptive statistics for the data.

The data

Data recorded from 1990 measured the time between eruptions and the duration of the eruption, both taken in minutes. This data set was based on 272 observations of Old Faithful's eruptions and consists of 2 variables: eruption time in mins and waiting time to next eruption (in mins). Both variables are numeric.

	eruptions	waiting
## 1	3.600	79
## 2	1.800	54
## 3	3.333	74
## 4	2.283	62
## 5	4.533	85
## 6	2.883	55

Analysis of the old faithful data

Eruption time and waiting time

By looking at the scatter plot, we can see a visual representation of the data.

```
plot(faithful$eruptions,faithful$waiting)
```

System Tray (Bottom):

Windows icon, File Explorer icon, Print icon, Task View icon, Internet Explorer icon, Taskbar icon, PDF icon, R icon.

Bottom Right:

EN, 14:57, 30/09/2020

The R markdown program & PDF output

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Addins

Go to/file/function

_Course_UHasselt_demo_V0.Rmd x R_Course_UHasselt_demo_V1.Rmd x R_Course_UHasselt_2020(html)_V1.Rmd x

```
76  
77  
78 + `## {r}  
79 min(faithful$eruptions)  
80  
81  
82 Let us look once again in the scatterplot. we can add labels name and title  
83  
84  
85 + `## {r}  
86 plot(faithful$eruptions,faithful$waiting,  
87 xlab="eruptions",ylab="waiting")  
88 title("The old faithful data")  
89  
90  
91 In the next figure we use different colors for the two clusters.  
92  
93 + `## {r}  
94 plot(faithful$eruptions,faithful$waiting,  
95 xlab="eruptions",ylab="waiting")  
96 points(faithful$eruptions[faithful$eruptions>3.2],faithful$waiting[faithful$eruptions>3.2],  
97 title("The old faithful data")  
98  
99
```

31:11 The data ↴

Console R Markdown x

ordinary text without R code

```
label: unnamed-chunk-15 | 91%  
label: unnamed-chunk-16 | 94%  
ordinary text without R code  
  
label: unnamed-chunk-16 | 97%  
label: unnamed-chunk-16 | 100%  
ordinary text without R code
```

"C:/Program Files/RStudio/bin/pandoc/pandoc" +RTS -K512m -RTS R_Course_UHasselt_2020.Rmd+autoLink_bare_uris+ascii_identifiers+tex_math_single_backslash--output R_Course_UHasselt_2020.pdf--template "C:\PROGRA~1\R\R-36-1.1\library\RMARD-1\rmrd\latex\DEFAUL~3.TEX" --highlighting-css --variable graphics=yes --variable "geometry:margin=1in" --variable "compact-title"
output file: R_Course_UHasselt_2020-pdf-_V2.knit.md

pandoc.exe: Could not parse YAML header: mapping values are not allowed in this context
output created: R_Course_UHasselt_2020-pdf-_V2.pdf

R_Course_UHasselt_2020-pdf-_V2.pdf

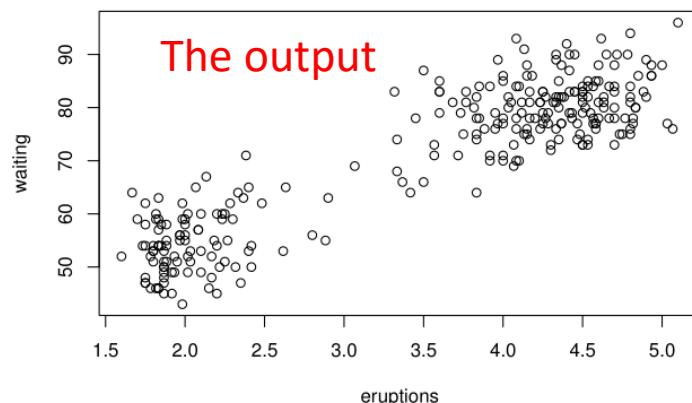
Page: 4 / 13

Find:

R code as a part of your text in the output

```
## [1] 96  
and the minimum eruption time is  
min(faithful$eruptions)  
## [1] 1.6  
Let us look once again in the scatterplot. We can add labels name and title  
plot(faithful$eruptions,faithful$waiting,  
xlab="eruptions",ylab="waiting")  
title("The old faithful data")
```

The old faithful data



The output

In the next figure we use different colors for the two clusters.

```
plot(faithful$eruptions,faithful$waiting,  
xlab="eruptions",ylab="waiting")
```

EN 14:59 30/09/2020

What did we saw today ?

- R Studio.
- R markdown + output.
- Basic graphical functions in R and how to control the figure (title, colors etc).
- Do the analysis in the file: [R_course_UHasselt_2021\(pdf\)_V2.Rmd](#)