

Data import and export

Lukáš Lehotský and Petr Ocelík

Summary of the last lesson

- Logical conditions or index positions allow for object filtering, when inserted into square brackets
- Square brackets must include all object dimensions
- When filtering data frames, filtering condition usually applies to rows

```
condition <- df[ ,3] > 1000000
```

```
df[ condition , ]
```

	<i>cars</i>	<i>type</i>	<i>price</i>	<i>consumption</i>
1	<i>BMW</i>	<i>3</i>	<i>1200000</i>	<i>6.2</i>
2	<i>Audi</i>	<i>A4</i>	<i>1164000</i>	<i>5.9</i>

```
df.sub <- df[ condition , ]
```

Working directory

- Folder, where all imports and exports are taking place
- Makes data **import and export** easier
- Functions `setwd()` and `getwd()`

```
getwd()
```

```
[1] "C:/Users/Lukas/..."
```

```
setwd("C:\\Users\\Lukas\\Documents\\R intro")
```

```
setwd("C:/Users/Lukas/Documents/R intro")
```


File Edit Code View Plots Session Build Debug Profile Tools Help

New Session

Interrupt R

Terminate R...

Restart R

Ctrl+Shift+F10

Set Working Directory

To Source File Location

To Files Pane Location

Choose Directory...

Ctrl+Shift+H

Load Workspace...

Save Workspace As...

Clear Workspace...

Quit Session...

Ctrl+Q

Run

Source

Environment

History

Import Dataset

Global Environment

Environment is empty

Files

Plots

Packages

Help

Viewer

Install

Update

Name	Description	Version
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5
<input type="checkbox"/> acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
<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"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.0
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<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

1:1 (Top Level)

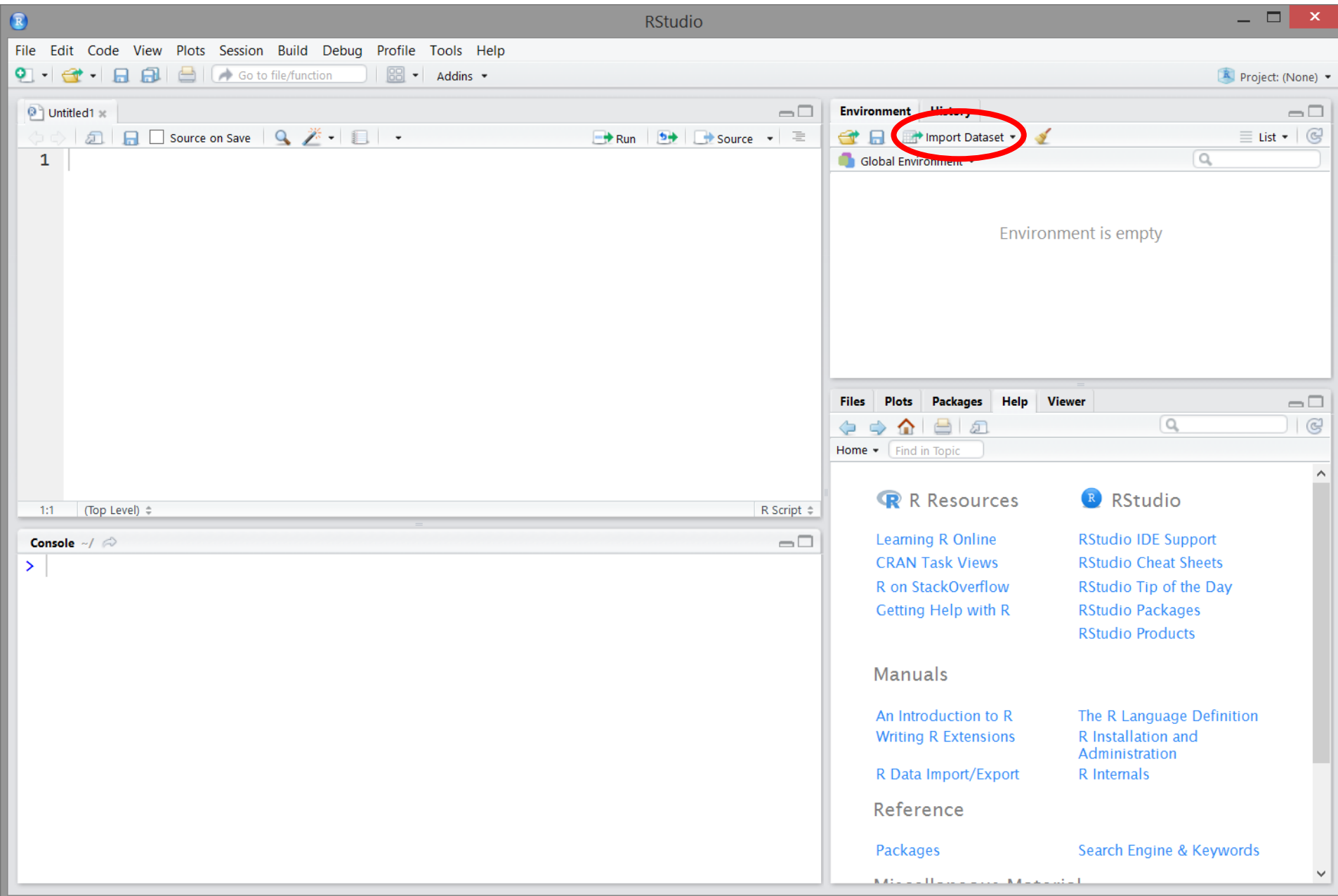
R Script

Console ~/

>

Data import

- R Studio visual import
- Functions
 - Structured data (tables)
 - Unstructured data
 - R native files



Data import: structured data

- Most common format of **tabular data**
 - CSV – comma-separated value
 - TSV – tab-separated value
 - TXT – plain text file
 - Other formats (XLSX, ...)
- CSV/TSV are most desired
- Essentially, CSV and TSV files are text files

Data import: structured data

	151	153	155	160	162
151	0	2	2	2	2
153	2	0	2	2	2
155	2	2	0	2	3
160	2	2	2	0	2

Data import: structured data

- BALIBOMBING2002_2002.csv
 - Matrix
 - Stored as a comma-separated value
 - First row – column names (note first cell is empty)
 - First column – row names
 - No missing values/texts

Data import: structured data

- Most common function `read.table()`
- Import **any** structured (tabular) data
- Highly **customizable** – **many arguments**
- **Always** imports data as **data frame**

```
df <- read.table( file = "BALIBOMBING2002_2002.csv",  
                 header = TRUE,  
                 sep = ",",  
                 row.names = 1,  
                 na.strings = "NA",  
                 stringsAsFactors = FALSE,  
                 fileEncoding = "UTF-8")
```

df x

Filter

	X151	X153	X155	X160	X162	X163	X175	X177	X180	X183	X189	X650
151	0	2	2	2	2	2	2	2	2	2	2	2
153	2	0	2	2	2	2	2	2	2	2	2	2
155	2	2	0	2	3	3	2	2	2	2	2	2
160	2	2	2	0	2	2	2	2	2	2	2	2
162	2	2	3	2	0	3	2	2	2	2	2	0
163	2	2	3	2	3	0	2	2	2	2	2	0
175	2	2	2	2	2	2	0	2	2	2	2	0
177	2	2	2	2	2	2	2	0	2	2	2	2
180	2	2	2	2	2	2	2	2	0	2	2	2
183	2	2	2	2	2	2	2	2	2	0	0	0
189	2	2	2	2	0	0	0	2	2	0	0	0
650	2	2	0	0	0	0	0	0	0	0	0	0
1579	2	2	2	2	2	2	2	2	2	2	2	2

Showing 1 to 15 of 27 entries

Console Terminal x

C:/Users/Lukas/Downloads/

```
> df <- read.table(
+   file = "BALIBOMBING2002_2002.csv",
+   header = TRUE,
+   sep = ",",
+   row.names = 1,
+   na.strings = "NA",
+   stringsAsFactors = FALSE,
+   fileEncoding = "UTF-8")
>
```

Environment History Connections

Import Dataset

List

Global Environment

Data

df 27 obs. of 27 variables

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5
<input type="checkbox"/> acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
<input type="checkbox"/> antiword	Extract Text from Microsoft Word Documents	1.1
<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"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.2
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<input type="checkbox"/> beepr	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.66.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2.2
<input type="checkbox"/> bit	A Class for Vectors of 1-Bit Booleans	1.1-13
<input type="checkbox"/> bit64	A S3 Class for Vectors of 64bit Integers	0.9-7
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> blob	A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS')	1.1.1
<input type="checkbox"/> brew	Templating Framework for Report Generation	1.0-6
<input type="checkbox"/> ca	Simple, Multiple and Joint Correspondence Analysis	0.70

Data import: unstructured data

- Most common to use function `readLines()`
- Can import any unstructured **data as a character vector**
- Data conversions/mining must follow

```
text <- readLines(con = "BALIBOMBING2002_2002.csv",  
                  warn = TRUE,  
                  encoding = "UTF-8")
```

Data import: RData files

- R has **native data format**
- Allows to save whole environment (all objects) or particular object for later use
- Easiest way to **access/share previous R work**
- Function `load()`

```
load(file = "balibombing.rdata")
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

df x

Filter

	X151	X153	X155	X160	X162	X163	X175	X177	X180	X183	X189	X650
151	0	2	2	2	2	2	2	2	2	2	2	2
153	2	0	2	2	2	2	2	2	2	2	2	2
155	2	2	0	2	3	3	2	2	2	2	2	2
160	2	2	2	0	2	2	2	2	2	2	2	2
162	2	2	3	2	0	3	2	2	2	2	2	0
163	2	2	3	2	3	0	2	2	2	2	2	0
175	2	2	2	2	2	2	0	2	2	2	2	0
177	2	2	2	2	2	2	2	0	2	2	2	2
180	2	2	2	2	2	2	2	2	0	2	2	2
183	2	2	2	2	2	2	2	2	2	0	0	0
189	2	2	2	2	0	0	0	2	2	0	0	0
650	2	2	0	0	0	0	0	0	0	0	0	0
1579	2	2	2	2	2	2	2	2	2	2	2	2

Showing 1 to 15 of 27 entries

Console Terminal x

C:/Users/Lukas/Downloads/

```
> str(df)
'data.frame': 27 obs. of 27 variables:
 $ X151 : int 0 2 2 2 2 2 2 2 2 2 ...
 $ X153 : int 2 0 2 2 2 2 2 2 2 2 ...
 $ X155 : int 2 2 0 2 3 3 2 2 2 2 ...
 $ X160 : int 2 2 2 0 2 2 2 2 2 2 ...
 $ X162 : int 2 2 3 2 0 3 2 2 2 2 ...
 $ X163 : int 2 2 3 2 3 0 2 2 2 2 ...
 $ X175 : int 2 2 2 2 2 2 0 2 2 2 ...
 $ X177 : int 2 2 2 2 2 2 2 0 2 2 ...
 $ X180 : int 2 2 2 2 2 2 2 2 0 2 ...
 $ X183 : int 2 2 2 2 2 2 2 2 2 0 ...
 $ X189 : int 2 2 2 2 0 0 0 2 2 0 ...
```

Environment History Connections

Import Dataset

Global Environment

Data

df 27 obs. of 27 variables

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5
<input type="checkbox"/> acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
<input type="checkbox"/> antiword	Extract Text from Microsoft Word Documents	1.1
<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"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.2
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<input type="checkbox"/> beeper	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.66.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2.2
<input type="checkbox"/> bit	A Class for Vectors of 1-Bit Booleans	1.1-13
<input type="checkbox"/> bit64	A S3 Class for Vectors of 64bit Integers	0.9-7
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> blob	A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS')	1.1.1
<input type="checkbox"/> brew	Templating Framework for Report Generation	1.0-6
<input type="checkbox"/> ca	Simple, Multiple and Joint Correspondence Analysis	0.70

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

	X151	X153	X155	X160	X162	X163	X175	X177	X180	X183	X189	X650
151	0	2	2	2	2	2	2	2	2	2	2	2
153	2	0	2	2	2	2	2	2	2	2	2	2
155	2	2	0	2	3	3	2	2	2	2	2	2
160	2	2	2	0	2	2	2	2	2	2	2	2
162	2	2	3	2	0	3	2	2	2	2	2	0
163	2	2	3	2	3	0	2	2	2	2	2	0
175	2	2	2	2	2	2	0	2	2	2	2	0
177	2	2	2	2	2	2	2	0	2	2	2	2
180	2	2	2	2	2	2	2	2	0	2	2	2
183	2	2	2	2	2	2	2	2	2	0	0	0
189	2	2	2	2	0	0	0	2	2	0	0	0
650	2	2	0	0	0	0	0	0	0	0	0	0
1579	2	2	2	2	2	2	2	2	2	2	2	2

Showing 1 to 15 of 27 entries

Console Terminal

C:/Users/Lukas/Downloads/

```
> str(df)
'data.frame': 27 obs. of 27 variables:
```

```
$ X151 : int  0 2 2 2 2 2 2 2 2 ...
$ X153 : int  2 0 2 2 2 2 2 2 2 ...
$ X155 : int  2 2 0 2 3 3 2 2 2 ...
$ X160 : int  2 2 2 0 2 2 2 2 2 ...
$ X162 : int  2 2 3 2 0 3 2 2 2 ...
$ X163 : int  2 2 3 2 3 0 2 2 2 ...
$ X175 : int  2 2 2 2 2 2 0 2 2 ...
$ X177 : int  2 2 2 2 2 2 2 0 2 ...
$ X180 : int  2 2 2 2 2 2 2 2 0 ...
$ X183 : int  2 2 2 2 2 2 2 2 0 ...
$ X189 : int  2 2 2 2 0 0 0 2 2 0 ...
```

Environment History Connections

Import Dataset

Global Environment

Data

df 27 obs. of 27 variables

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
User Library		
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5
<input type="checkbox"/> acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
<input type="checkbox"/> antiword	Extract Text from Microsoft Word Documents	1.1
<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"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.2
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<input type="checkbox"/> beeper	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.66.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2.2
<input type="checkbox"/> bit	A Class for Vectors of 1-Bit Booleans	1.1-13
<input type="checkbox"/> bit64	A S3 Class for Vectors of 64bit Integers	0.9-7
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> blob	A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS')	1.1.1
<input type="checkbox"/> brew	Templating Framework for Report Generation	1.0-6
<input type="checkbox"/> ca	Simple, Multiple and Joint Correspondence Analysis	0.70

Data conversions

- Data frame may be **converted to other object types**
 - Function `as.matrix()` turns DF to matrix
 - Function `as.list()` will turn DF into list of vectors
- Other object types may be converted to a data frame as well
 - Function `as.data.frame()`



Data conversions: DF to matrix

- Data are loaded as **data frames by default** – for networks, conversion is necessary
- **Class** of data needs to be considered beforehand
 - If **text included**, the whole matrix will be **coerced to character**
 - Conversion of data classes after the DF conversion is **painful** – needs to be done **column by column**

```
mat <- as.matrix(df)
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

df x

Filter

	X151	X153	X155	X160	X162	X163	X175	X177	X180	X183	X189	X650
151	0	2	2	2	2	2	2	2	2	2	2	2
153	2	0	2	2	2	2	2	2	2	2	2	2
155	2	2	0	2	3	3	2	2	2	2	2	2
160	2	2	2	0	2	2	2	2	2	2	2	2
162	2	2	3	2	0	3	2	2	2	2	2	0
163	2	2	3	2	3	0	2	2	2	2	2	0
175	2	2	2	2	2	2	0	2	2	2	2	0
177	2	2	2	2	2	2	2	0	2	2	2	2
180	2	2	2	2	2	2	2	2	0	2	2	2
183	2	2	2	2	2	2	2	2	2	0	0	0
189	2	2	2	2	0	0	0	2	2	0	0	0
650	2	2	0	0	0	0	0	0	0	0	0	0
1579	2	2	2	2	2	2	2	2	2	2	2	2

Showing 1 to 15 of 27 entries

Console Terminal x

C:/Users/Lukas/Downloads/

```
> str(df)
'data.frame': 27 obs. of 27 variables:
 $ X151 int 0 2 2 2 2 2 2 2 2 2 ...
 $ X153 int 2 0 2 2 2 2 2 2 2 2 ...
 $ X155 int 2 2 0 2 3 3 2 2 2 2 ...
 $ X160 int 2 2 2 0 2 2 2 2 2 2 ...
 $ X162 int 2 2 3 2 0 3 2 2 2 2 ...
 $ X163 int 2 2 3 2 3 0 2 2 2 2 ...
 $ X175 int 2 2 2 2 2 2 0 2 2 2 ...
 $ X177 int 2 2 2 2 2 2 2 0 2 2 ...
 $ X180 int 2 2 2 2 2 2 2 2 0 2 ...
 $ X183 int 2 2 2 2 2 2 2 2 2 0 ...
 $ X189 int 2 2 2 2 0 0 0 2 2 0 ...
```

Environment History Connections

Import Dataset

Global Environment

Data

df 27 obs. of 27 variables

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5
<input type="checkbox"/> acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
<input type="checkbox"/> antiword	Extract Text from Microsoft Word Documents	1.1
<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"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.2
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<input type="checkbox"/> beeper	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.66.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2.2
<input type="checkbox"/> bit	A Class for Vectors of 1-Bit Booleans	1.1-13
<input type="checkbox"/> bit64	A S3 Class for Vectors of 64bit Integers	0.9-7
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> blob	A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS')	1.1.1
<input type="checkbox"/> brew	Templating Framework for Report Generation	1.0-6
<input type="checkbox"/> ca	Simple, Multiple and Joint Correspondence Analysis	0.70

Data conversions: matrix to DF

- Inconvenient process
- **Argument** `stringsAsFactors` is important – otherwise any text becomes **factor variable**
- **Data class** of matrix will be **preserved**
 - No automatic recognition built in on conversion
- If matrix is unnamed, column names are **assigned automatically**

```
mat2df <- as.data.frame(mat)
```

```
mat2df <- as.data.frame(mat, stringsAsFactors = FALSE)
```

Names of rows and columns

- **Column names** are an issue in data frames
- Data frame columns **must** have proper **column (variable) names**
 - Names **starting with number are not allowed** (default transformation behavior to “X#”)
 - Missing column names lead to **automatically generated** names (V1, V2, ..., Vn)
 - Conversion (e.g. from list) might result into crazy column names

Names of rows and columns

- If issue with names is present, **renaming** needs to follow
- Functions `rownames()` and `colnames()` allow to access existing names
- Renaming is **counterintuitive** – we need to assign **vector with names** to the **function result**
- It must be of a **same length** as number of columns

```
colnames(df) <- c("var1" , "var2" , ... , "varN")
```


Names of rows and columns

- Names may be **generated automatically** using `paste()` or `paste0()` functions – these functions collapse text bits together
 - Function `paste()` adds **spaces** between texts
 - Function `paste0()` collapses texts **without spaces**

```
colnames(df) <- paste0("actor" , 1:27)
```

- **In case of a matrix, each dimension** needs to be named **separately**

```
colnames(mat) <- paste0("actor" , 1:27)
```

```
colnames(mat) <- rownames(mat)
```

Data export

- Export of entire workspace
- Export of particular objects
 - R native files
 - Structured data (tables)
 - Unstructured data

Data export: saving entire workspace

- Function `save.image()`
 - Rdata format
 - No additional arguments required

```
save.image(file = "workspace_2018_07_04.rdata")
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

df x

Filter

	X151	X153	X155	X160	X162	X163	X175	X177	X180	X183	X189	X650
151	0	2	2	2	2	2	2	2	2	2	2	2
153	2	0	2	2	2	2	2	2	2	2	2	2
155	2	2	0	2	3	3	2	2	2	2	2	2
160	2	2	2	0	2	2	2	2	2	2	2	2
162	2	2	3	2	0	3	2	2	2	2	2	0
163	2	2	3	2	3	0	2	2	2	2	2	0
175	2	2	2	2	2	2	0	2	2	2	2	0
177	2	2	2	2	2	2	2	0	2	2	2	2
180	2	2	2	2	2	2	2	2	0	2	2	2
183	2	2	2	2	2	2	2	2	2	0	0	0
189	2	2	2	2	0	0	0	2	2	0	0	0
650	2	2	0	0	0	0	0	0	0	0	0	0
1579	2	2	2	2	2	2	2	2	2	2	2	2

Showing 1 to 15 of 27 entries

Console Terminal x

C:/Users/Lukas/Downloads/

> View(df)

> |

Environment History Connections

Import Dataset

Global Environment

Data

df 27 obs. of 27 variables

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5
<input type="checkbox"/> acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
<input type="checkbox"/> antiword	Extract Text from Microsoft Word Documents	1.1
<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"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.2
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<input type="checkbox"/> beepr	Easily Play Notification Sounds on any Platform	1.2
<input type="checkbox"/> BH	Boost C++ Header Files	1.66.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2.2
<input type="checkbox"/> bit	A Class for Vectors of 1-Bit Booleans	1.1-13
<input type="checkbox"/> bit64	A S3 Class for Vectors of 64bit Integers	0.9-7
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> blob	A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS')	1.1.1
<input type="checkbox"/> brew	Templating Framework for Report Generation	1.0-6
<input type="checkbox"/> ca	Simple, Multiple and Joint Correspondence Analysis	0.70

Data export: saving particular object as Rdata

- Function `save()`
- Rdata is **R native format**
 - Preserves **the object as is**, with its object name
 - Allows to **avoid export and import issues**

```
save(df,  
      file = "BALIBOMBING_export.rdata")
```

Data export: saving tabular data

- Function `write.table()`
- Object which is exported has to be specified
- Otherwise similar to `read.table()`

```
write.table(df,  
            "BALIBOMBING_export.csv",  
            sep = ",",  
            row.names = TRUE,  
            col.names = TRUE,  
            fileEncoding = "UTF-8")
```

Data export: saving unstructured data

- Function `writelnLines()`
- Similar to `readLines()`
- Basically no arguments

```
writelnLines(df,  
             "BALIBOMBING_export.txt")
```

Other input and output options

- Built-in CSV-specific functions
 - `read.csv()` and `write.csv()` functions
 - `read.csv()` requires less arguments than `read.table()` but has more limited functionality
- Libraries “xlsx” or “openxlsx” for XLSX input
 - Both packages contain functions `read.xlsx()` and `write.xlsx()` – similar to `read.table()`
- Package “readtext” designed to read unstructured text data (PDFs, Word documents, etc.)
- Other packages for specific formats (XML, HTML, Json, SQL, ...)

Practice 1

- Download the package “MEB433_434_03A_practice.zip” into your computer
 - Unpack it into any folder
 - Set the folder as a working directory
- Import the data “nrg_122a.csv” into R
 - Explore and adjust the data set in a text editor, if necessary
 - Achieve the following
 - Dataset has proper column names
 - Missing values are treated correctly (find function arguments)
 - All columns are **numeric** (there’s a catch here)

Practice 2

- Get rid of the summary rows in the data frame
- Get rid of the missing values for the last year
- Extract only the last 3 years into a separate object
- Export the subset in rdata format
- Export the subset in CSV format