

# Пакетна организация, променливи, основни математически операции в R и типове данни

## Статистическа обработка на данни с R

Пламен Петров и Тодор Балабанов

Център за обучение  
Институт по информационни и комуникационни технологии  
Българската академия на науките  
*p.petrov@iit.bas.bg todorb@iinf.bas.bg*

03.V.2020

## Acknowledgments

These teaching materials are funded by Velbazhd Software LLC and it is partially supported by the Bulgarian Ministry of Education and Science (contract D01–205/23.11.2018) under the National Scientific Program “Information and Communication Technologies for a Single Digital Market in Science, Education and Security (ICTinSES)”, approved by DCM # 577/17.08.2018.

# Съдържание

- 1 **Инсталиране на пакети**
  - Пример с пакета `coefplot`
  - Връзки между пакетите
  - Зареждане и премахване на пакети
- 2 **Основни математически операции**
  - Изпълнение на команди с операции за пресмятане
  - Свойства на операциите
- 3 **Типове и променливи**
  - Създаване на променливи
  - Премахване на променливи
- 4 **Представяне на информацията**
  - Числени стойности
  - Символни низове
  - Астрономическо време
  - Логически стойности
- 5 **Заклучение**
  - Дискусия

# Инсталиране на пакети

Пример с пакета coefplot

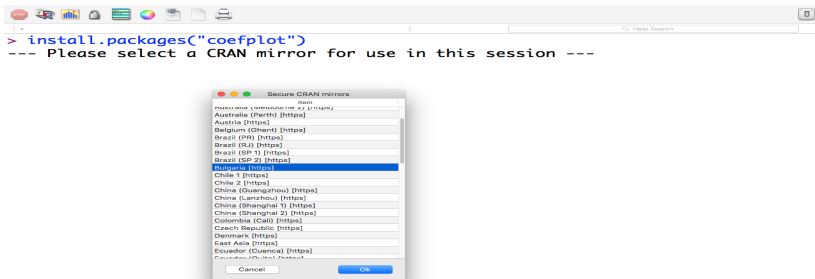
# Команда за инсталиране на пакета coefplot



```
> install.packages("coefplot")
```

Пример с пакета `coefplot`

## Избор на сървър за изтегляне на пакета



# Зависимости между пакетите

```

> install.packages("coefplot")
--- Please select a CRAN mirror for use in this session ---
also installing the dependencies 'colorspace', 'RColorBrewer', 'dichromat',
'munsell', 'labeling', 'stringi', 'bindr', 'jsonlite', 'yaml', 'fansi', 'utf8',
'digest', 'gtable', 'lazyeval', 'rlang', 'scales', 'viridisLite', 'withr',
'Rcpp', 'stringr', 'assertthat', 'bindrcpp', 'glue', 'pkgconfig', 'R6',
'tidysselect', 'BH', 'plogr', 'htmlwidgets', 'htmltools', 'zoo', 'xts', 'cli',
'crayon', 'pillar', 'ggplot2', 'plyr', 'reshape2', 'useful', 'dplyr', 'dygraphs',
'tibble', 'magrittr', 'purrr'

There is a binary version available but the source version is later:
  binary source needs_compilation
xts 0.10-2 0.11-0 TRUE

Do you want to install from sources the package which needs compilation? (Yes/no/
cancel) Yes

```

## Резултат от инсталацията на пакета

```
include -I"/Library/Frameworks/R.framework/Versions/3.5/Resources/library/zoo/
include" -I/usr/local/include -fPIC -Wall -g -O2 -c unique.time.c -o
unique.time.o
clang -dynamiclib -Wl,-headerpad_max_install_names -undefined dynamic_lookup -
single_module -multiply_defined suppress -L/Library/Frameworks/R.framework/
Resources/lib -L/usr/local/lib -o xts.so add_class.o any.o attr.o binsearch.o
coredata.o diff.o dimnames.o endpoints.o extract_col.o init.o isOrdered.o
isXts.o merge.o migrate_index_attr.o na.o period_apply.o period_arithmetic.o
period_quantile.o rbind.o rollfun.o runSum.o startofyear.o subset.o subset.old.o
toperiod.o totalcols.o tryXts.o unique.time.o -F/Library/Frameworks/
R.framework/.. -framework R -Wl,-framework -Wl,CoreFoundation
installing to /Library/Frameworks/R.framework/Versions/3.5/Resources/library/
xts/libs
** R
** data
** inst
** byte-compile and prepare package for lazy loading
** help
*** installing help indices
** building package indices
** installing vignettes
** testing if installed package can be loaded
* DONE (xts)

The downloaded source packages are in
'/private/var/folders/dx/x_v7h9fj6w10mcd11km9txz80000gq/T/RtmphPCUT9/
downloaded_packages'
```



# Зареждане на пакета coefplot



```
> require( coefplot )
Loading required package: coefplot
Loading required package: ggplot2
> |
```

## Премахване на пакета coefplot от общата памет



```
> require( coefplot )  
Loading required package: coefplot  
Loading required package: ggplot2  
> detach( package:coefplot )  
>
```

# Основни математически операции

Изпълнение на команди с операции за пресмятане

## Примерни аритметични операции



```
> 1 + 1
[1] 2
> 2 + 2 * 2
[1] 6
> 5 / 3
[1] 1.666667
>
```

## Кардиналност и асоциативност

Събиране - бинарна операция

$$1 + 1$$

Минус - унарна операция

$$-5$$

Лява асоциативност

$$2 + 2 + 2$$

Дясна асоциативност

$$a = b = 2$$

# Контекстна зависимост на операциите

## Контекстна зависимост при събиране

"abc" + "def"

2 + 2

## Контекстна зависимост при делене

5 / 3

5.0 / 3.0

# Приоритет на операцияте

## Аритметични операции

$$2 + 2 * 2$$

## Смяна на приоритета

$$(2 + 2) * 2$$

# Типове и променливи



# Операции за присвояване

## Задаване на стойност

```
a = 1
```

```
b <- 2
```

```
c = d = 3
```

```
e <- f <- 4
```

```
assign("g", 5)
```

```
a += 6
```

```
b -= 7
```

## Алтернативи за присвояване

### Два различни записа

```
median(x = 1:10)
```

```
median(x <- 1:10)
```

В първия случай променливата  $x$  не остава в глобалната памет, а изчезва, докато при втория случай променливата  $x$  остава в глобалната памет след извикването.

# Почистване на глобалната памет

Конкретна променлива или цялата памет

```
rm( a )
```

```
rm( list=ls() )
```

# Представяне на информацията

## Проверка за типа на променливата

### Функция class

```
a = 5
class( a )
[1] "numeric"
```

### Функция is.numeric

```
is.numeric( a )
[1] TRUE
```

## Използване на целочислен тип

Типовете numeric и integer са различни

```
b <- 6L
class( b )
[1] "integer"
is.numeric( b )
[1] TRUE
is.integer( b )
[1] TRUE
is.integer( a )
[1] FALSE
```

# Представяне на текст

## Символни низове в R

```
s1 <- "Simple text."
[1] "Simple text."
class( s1 )
[1] "character"
s2 <- factor( "Another simple text." )
[1] Another simple text.
Levels: Another simple text.
class( s2 )
[1] "factor"
```

## Дължина на данните

### Символни низове и числа

```
nchar( s1 )
```

```
[1] 12
```

```
nchar( s2 )
```

```
Error in nchar(s2) : 'nchar()' requires a character vector
```

```
nchar( a )
```

```
[1] 1
```



## Представяне на дати в паметта

### Тип Date

```
d1 <- as.Date("1979-04-21")
[1] "1979-04-21"
class( d1 )
[1] "Date"
as.numeric( d1 )
[1] 3397
```

# Представяне дати и часове в паметта

## Тип POSIXct

```
d2 <- as.POSIXct("1980-02-12 05:25")
```

```
[1] "1980-02-12 05:25:00 EET"
```

```
class( d2 )
```

```
[1] "POSIXctPOSIXt"
```

```
as.numeric( d2 )
```

```
[1] 319173900
```

# Булеви стойности

## Логически тип данни

```
x1 <- TRUE
[1] TRUE
class( x1 )
[1] "logical"
is.logical( x1 )
[1] TRUE
as.numeric( x1 )
[1] 1
```



# Алтернативни константи

По-добре е да се ползва пълният запис

```
x2 <- F
[1] FALSE
class( x2 )
[1] "logical"
is.logical( x2 )
[1] TRUE
as.numeric( x2 )
[1] 0
```

## Резултат от логически ирази

### Операции за сравнение

2 == 3

[1] FALSE

5 != 6

[1] TRUE

"Peter" == "Ivan"

[1] FALSE

# Заклучение

## Въпроси и отговори

Благодаря за вниманието!