

Introduction

Programming in R for Data Science

Anders Stockmarr, Kasper Kristensen, Anders Nielsen

A first session

```
> x<-rnorm(100)
> head(x)

[1] -0.2596588 -0.5439287 -0.3976459 -0.8051366 -0.8854298
[6] -0.1317834

> mean(x)

[1] 0.04613127

> sd(x)

[1] 1.013185

> min(x)

[1] -3.207916

> max(x)

[1] 2.550923
```

R as a calculator

Basic operations

```
> 2+2
```

```
[1] 4
```

```
> 7*17
```

```
[1] 119
```

```
> sqrt(9)
```

```
[1] 3
```

```
> 3^3
```

```
[1] 27
```

```
> log(7)
```

```
[1] 1.94591
```

```
> log10(7)
```

```
[1] 0.845098
```

Precision

```
> sin(pi/2)
```

```
[1] 1
```

```
> pi
```

```
[1] 3.141593
```

```
> options(digits=22)
```

```
> pi
```

```
[1] 3.1415926535897931
```

Infinity or not defined, and missings

```
> 1/0
```

```
[1] Inf
```

```
> 2*Inf
```

```
[1] Inf
```

```
> -1/0
```

```
[1] -Inf
```

```
> 0/0
```

```
[1] NaN
```

```
> c(1,2,3,NA,5)
```

```
[1] 1 2 3 NA 5
```

```
> mean(c(1,2,3,NA,5))
```

```
[1] NA
```

Assignments to variables

```
> x <- 5
```

```
> x
```

```
[1] 5
```

```
> # x=5 can be used; not recommended
```

```
> x*x
```

```
[1] 25
```

```
> y <- x+5
```

```
> ls()
```

```
[1] "x" "y"
```

```
> rm(x)
```

```
> ls()
```

```
[1] "y"
```

Getting help

Different sources

- ▶ The internal help functions
- ▶ Manuals
- ▶ Cheat sheets
- ▶ Mailing lists
- ▶ Google
- ▶ <http://stackoverflow.com/>
- ▶ Books
- ▶ Local **R** users

Internal help function

- ▶ If we know what function we need help with, then type:

```
> ?mean      # shorthand for help(mean)
```
- ▶ If we just want to see an example

```
> example(mean)
```
- ▶ Often we don't know exactly what we are looking for

```
> ??"fitting linear model"    # shorthand for  
>                             # help.search("fitting linear model")
```

- ▶ Available on-line <http://www.r-project.org>, but main ones are also part of the installation, type:
`> help.start()`
- ▶ "An Introduction to R" and "R Data Import/Export" are worth looking at.

Cheat sheet

- ▶ May be useful to mount near you desk when starting with **R**
- ▶ Several can be found at <http://cran.r-project.org/other-docs.html>

Mailing lists

- ▶ **R** has an extremely active user base.
- ▶ The mailing lists are very helpful, you can access many at <https://www.r-project.org/mail.html>... but users prefer that you read and think before you pose questions.
- ▶ Also they easily smell if you are asking for an answer to a homework question.
- ▶ The archive is a goldmine of knowledge <http://tolstoy.newcastle.edu.au/R/>.

Do you know stack overflow?

- ▶ Main site at: `http://stackoverflow.com/`
- ▶ Find 'tags' and **R**
- ▶ The rating often gives you a high quality answer
- ▶ Example `http://stackoverflow.com/questions/9508518/why-are-these-numbers-not-equal`

Books

- ▶ There are many
- ▶ Here is a list of 150 **R** –books:
<http://www.r-project.org/doc/bib/R-books.html>
- ▶ Most are fairly specialized

Local R users

- ▶ By far the optimal source of information.
- ▶ Not only are they close by, but on top of explanations, they may even provide you with a piece of code that nearly does the job.
- ▶ Remember this when someones comes to you for R help.