Introduction

Christophe Lalanne

Fall 2017

Overview

Tools

Overview

Organization

Lectures are scheduled each Tuesday, starting November, 7th.

This will be a mix of R programming and statistical modeling, ranging from data crunching to data mining and machine learning.

- Nov, 7: R basics
- Nov, 14: Graphics with `ggplot2`
- Nov, 15*: R dplyr and data.table
- Nov, 28: R and database, data mining
- Dec, 5: Statistical modeling in R
- Dec, 12: Reporting with R Markdown and Shiny
- Dec, 19: Final project

Synopsis

- R programming, statistical methods, hand-on practicals
- R Markdown report
- Combined use of practicals (20%), intermediate project (40%) and final assessment (40%).

Tools

Ressources

Ressources are available in the following Git-versioned repository: https://bitbucket.org/chlalanne/rstats-esme

You can fork the repository using `http` or `ssh` (`git@bitbucket.org:chlalanne/rstats-esme.git`) and keep posted with `git pull`. You can submit Issues or PRs if you like.

Additional tools

- R and Rstudio
- Some R packages (and their dependencies which shall be installed automatically): tidyverse, ggplot2, data.table, leaflet, ...
- "R for Data Science" [R4DS] (Wickham and Grolemund 2017), (maybe) "Advanced R" [AR] (Wickham 2014), "An Introduction to Statistical Learning" [ISLR] (James et al. 2013)

The following command will download required packages for you:

```
source("get_packages.R")
```

Eventually, you may want to install on your local machine Git, jq and csvkit, as well as a good text editor (other than Notepad).[1]

[1] I will not be of any help in case you are running Windows on your computer.

References

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R*. Springer New York. http://www-bcf.usc.edu/~gareth/ISL/.

Wickham, Hadley. 2014. *Advanced R*. Chapman & Hall/CRC. https://adv-r.hadley.nz.

Wickham, Hadley, and Garrett Grolemund. 2017. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media, Inc. http://r4ds.had.co.nz.