R software

Course title - Intitulé du cours	R software
Level / Semester - Niveau /semestre	M2 / S1
School - Composante	Ecole d'Economie de Toulouse
Teacher - Enseignant responsable	THIBAULT LAURENT
Other teacher(s) - Autre(s) enseignant(s)	
Other teacher(s) - Autre(s) enseignant(s)	
Other teacher(s) - Autre(s) enseignant(s)	
Other teacher(s) - Autre(s) enseignant(s)	
Other teacher(s) - Autre(s) enseignant(s)	
Lecture Hours - Volume Horaire CM	
TA Hours - Volume horaire TD	
TP Hours - Volume horaire TP	8
Course Language - Langue du cours	Anglais
TA and/or TP Language - Langue des TD et/ou TP	Anglais

Teaching staff contacts - Coordonnées de l'équipe pédagogique :

Thibault.Laurent@tse-fr.eu – T206

Several means of interaction are possible: after the classes, by email or at my office with prior appointment.

Course Objectives - Objectifs du cours :

This course is open to beginner, intermediate or advanced students with the software and programming language **R**. It will allow participants to deal with particular situations in data science (string data management/text mining, messy data, big data, etc.). Participants will learn how to optimize their computation time by using either the specificities of the software as the vectorized functions or by using specific packages like **Rcpp** (for programming in **C++** language) or from the **tidyverse** universe. An introduction to parallel computing will be given in order to use the maximum capacities of the machines. Finally, an overview of data visualization will be given by using the packages **ggplot2** and **plotly**. The course is structured in 4 sessions. Each session will be illustrated by practical work in the form of a case study.

Prerequisites - Pré requis :

R version: >= 4.1.0

R packages needed:

Available on CRAN : "foreign", "jsonlite", "readr", "readxl", "sas7bdat", "XML", "reticulate", "data.table", "ff", "microbenchmark", "Matrix", "classInt", "glue", "stringr", "wordcloud", "gplots", "pryr", "Rcpp", "tidyverse", "DSR", "Amelia", "DMwR", "missForest", "naniar", "sp", "parallel", "rpart", "snow", "snowFT", "broom", "gapminder", "ggcorrplot", "ggpol", "ggridges", "kableExtra", "knitr", "plotly", "RColorBrewer", "shiny", "stargazer", "survival", "survminer", "visreg", "vcd", "zoo"

Available on Github: "hadley/emo", "edwindj/ffbase"

Practical information about the sessions - Modalités pratiques de gestion du cours :

During the sessions, the students can bring their own laptop or tablet or use the computers in the room.

In order to respect their teacher and class mates, the students are expected not to be more than 5 minutes late.

Grading system - Modalités d'évaluation :

This course is exempt from note

Bibliography/references - Bibliographie/références :

the R team (2021). *An introduction to R*. https://cran.r-project.org/doc/manuals/R-intro.pdf

Thibault Laurent (2021). *R avancé*. http://www.thibault.laurent.free.fr/cours/R_avance/

Hadley Wickham (2019). *Advanced R*, Chapman & Hall. http://adv-r.had.co.nz/

Session planning - Planification des séances :

* session 1 : Data Management with **R**

* session 2 : Advanced programming with **R**

* session 3 : Parallel computing with **R**

* session 4 : Data visualization with **R**

Distance learning – Enseignement à distance :

Distance learning can be provided when necessary by implementing, for example: / En cas de nécessité, un enseignement à distance sera assuré en mobilisant, par exemple :

- Recorded lectures (videos) / Vidéo enregistrée de la présentation du matériel pédagogique
- Remote (online) tutorials (classes) / TP/TD à distance