ESTP R - Introduction and Installation

What is R

- A programming language for statistical computing, founded in 1993
- Open source
- Developed and maintained by the *R-core*, a small team of dedicated volunteers
- Extendable via packages hosted on **CRAN**, the Comprehensive R Archive Network
- As of June 2022, there are 18564 (!) packages on CRAN
- The most popular IDE (editor) is **RStudio**

Base R and packages

Base R (so the plain R language) is sufficient for most statistical computations and visualizations R packages extend the functionality of R. Which to use?

- An extensive list of packages organized by tasks: **Task Views**
- A list of recommended software for Official Statistics **Awesome official statistical software**
- Packages maintained by RStudio are generally recommended (contained in the umbrella package called **tidyverse**) (https://www.rstudio.com/products/rpackages/)

R and other programming languages

R (and Python) are widely recommended languages for statistics and data science. Why?

- Both R and Python are high level programming languages; few lines of code are required for a computation (C requires more code, and Assembly even more)
- Statistical computing can be done with SAS, SPSS, and Excel, but this is discouraged because of limited functionality and dependency on private companies
- Matlab is very similar to R, but is commercial and has a much smaller community (and therefore less add-on packages)
- Julia is a relatively new data science language, much faster than R and Python, but is still maturing

R or Python?

- Python is a general-purpose programming language, whereas R is more focused on statistics and data science
- Python and its core libraries (numpy, scipy, etc.) tend to be more stable than R. On the other hand, they are less evolving.
- (Almost) everything that can be done in R can be done in Python and vice versa
- Installation of R is easier for end-users
- Generally speaking, Python is ahead in machine learning and R in data exploration and visualization

Installation

R:

- Go to: https://cran.r-project.org and choose your operating system
- Follow the instructions

RStudio:

- Go to: https://www.rstudio.com/products/rstudio/ and select either Desktop or Server
- Follow the instructions

Installation of packages can be done inside RStudio