Home

Welcome to **Introduction to Statistics with R**! In this course, we will walk through **R**, a programming language popularly used in computational sciences. Please bring your laptop to the course. More importantly, please make sure that you install the following software on your laptop prior to showing up! If you have any problem installing these software, contact me or the IT support of your department to ensure you have the software you need for the course. The software are as follows:

- **R Statistical Software (https://cran.r-project.org/bin/windows/base/)**
- **RStudio Desktop IDE (https://www.rstudio.com/products/rstudio/)**
- **R Markdown package (https://bookdown.org/yihui/rmarkdown/installation.html)**

Of the three, probably RMarkdown is the most difficult one to install, if you are not familiar with R and RStudio. The nice thing about RMarkdown is that you can also install it within RStudio, simply by opening an RMarkdown file, **as shown in this silent video (https://www.youtube.com/watch?v=QaKCirYknS8)** !

The course is designed for absolute beginners, who had never installed R on their laptop nor heard of it! Yet, it can also be useful for those who have been already working with this computer language.

Within a 3-day program, the course is intended to:

1. Introduce R as a programming language:

- Become familiar with R and RStudio environments
- Understand R objects and their characteristics
- Understand different data structures in R
- Learn to write and organize R syntax and implement reproducible analyses within RStudio
- Basic statistical analyses

2. Introduce R's applications for data management

- Descriptive statistics in R
- Learn data cleaning and data management
- Carry out data quality checks
- Data transformation
- Using R loops for data management or analysis.

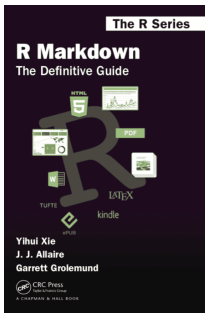
3. Introduce R graphical grammar for data visualization

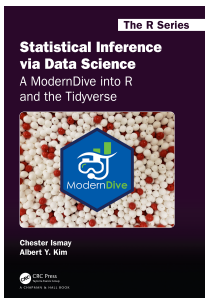
- Grammar of graphics with GGPLOT2 package

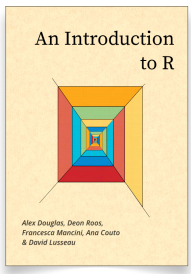
Course materials

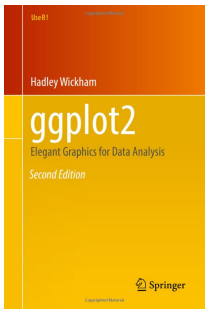
The course materials are all available publicly and there is no need for you to buy any book! This is one of the benefits of using open-source software that are maintained by the scientific community. You don't have to pay for neither the software nor the books! During the course, I will present some of the contents of the following books:

1. **R Markdown** (https://bookdown.org/yihui/rmarkdown/)
2. **Statistical Inference via data science** (https://moderndive.com/)
3. **An Introduction to R** (https://intro2r.com/)
4. **ggplot2** (https://ggplot2-book.org/)









Luckily, R has so many excellent free-ebooks. If the books above are not enough, here are some more (optional):

- **R for Data Science** (https://r4ds.had.co.nz/index.html)
- **An introduction to R** (https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf)