Intro to This Course; Getting Started with R

Enrico Toffalini

Why becoming an R user?

- **Mainstream in academia** for data science, increasingly used in business. *Job market advantage!*
- **Free & open-source**: wherever you go, R will be with you at no costs (unlike *MPLUS*, *MATLAB*, *SPSS*, etc.)
- **Real programming language**: difficult at the beginning, but: 1) gives you lots of flexibility; 2) has transfer on other programming languages (e.g., ***Python***).
- **Vast community support** thanks to a large and active community (also, *chatGPT*, *Lucrez-IA*, etc., know it pretty well!).
- **Huge ecosystem**, >20,000 packages on CRAN, more from other sources (e.g., GitHub), to do amazing stuff with statistical data analysis, machine learning, data visualization, developing webapps [*shiny*], writing reports and even entire books [*bookdown*, *rmarkdown*]), can integrate with *quarto*, *github*.
- Facilitates **reproducible research** by sharing code and workflows.

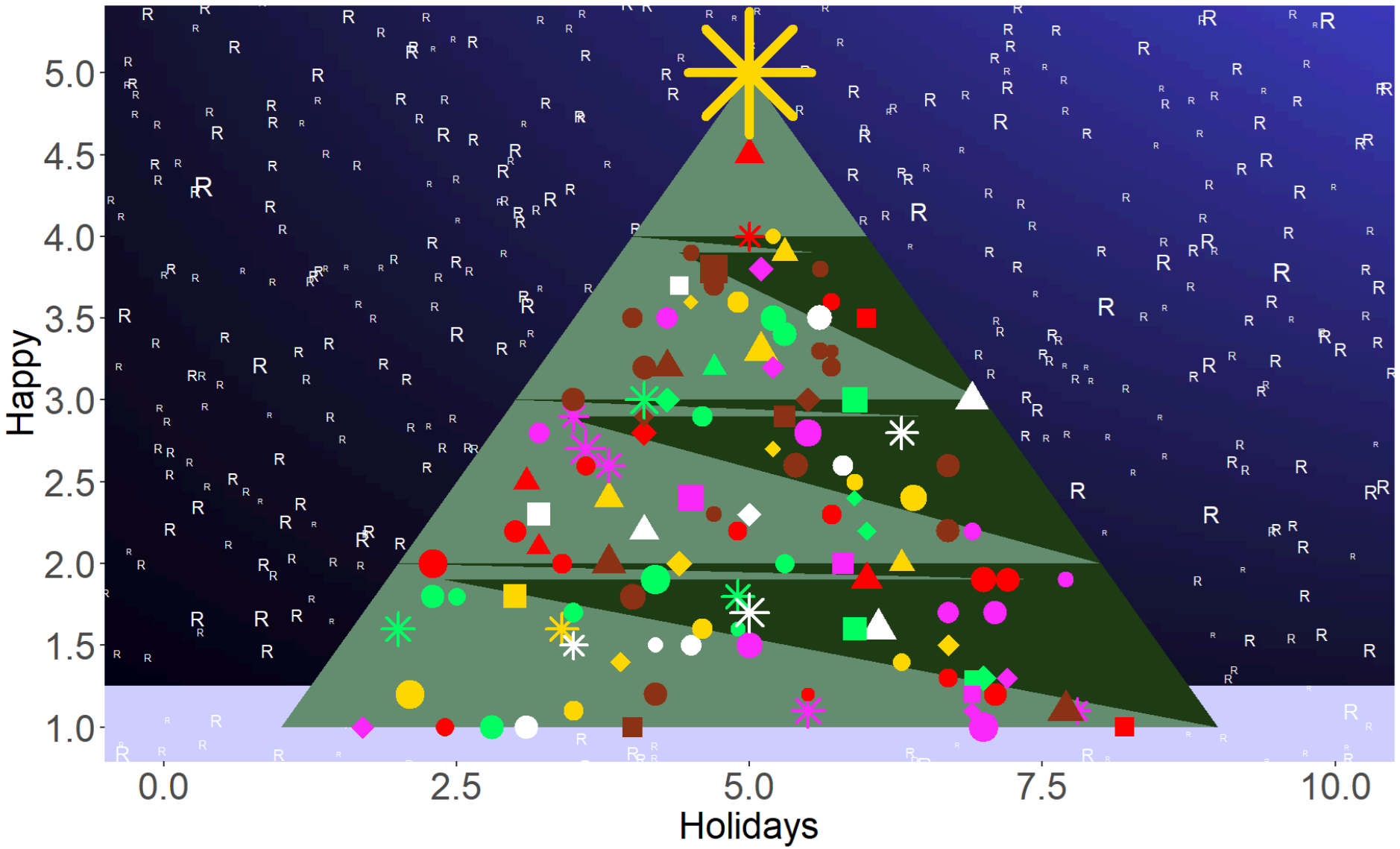
What you may expect to learn in this course:

- Executing fundamental operations and using basic functions;
- Working with essential data types and structures;
- Gaining some proficiency in managing and manipulating data with vectors and dataframes;
- Understanding some fundamental concepts of programming.

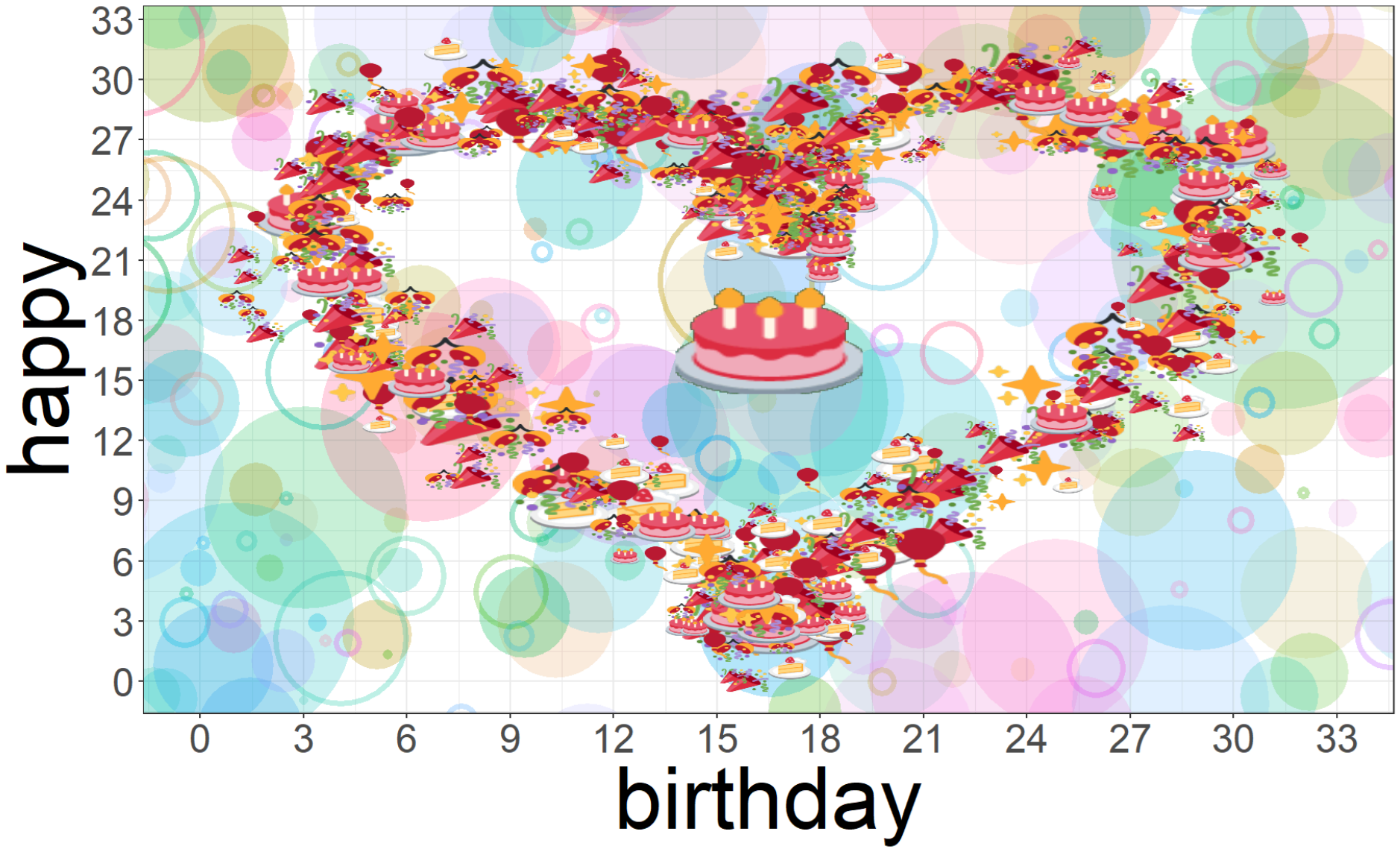
Over the next *couple of years*, following this PhD program, you may have the opportunity to learn to use R to perform at least some fundamentals about:

- Core statistical inference methods;
- (Generalized) linear (mixed-effects) modeling;
- Data visualization using *ggplot2*;
- Power analysis via data simulation ;
- Structural Equation Modeling (SEM);
- Conducting Meta-Analysis.

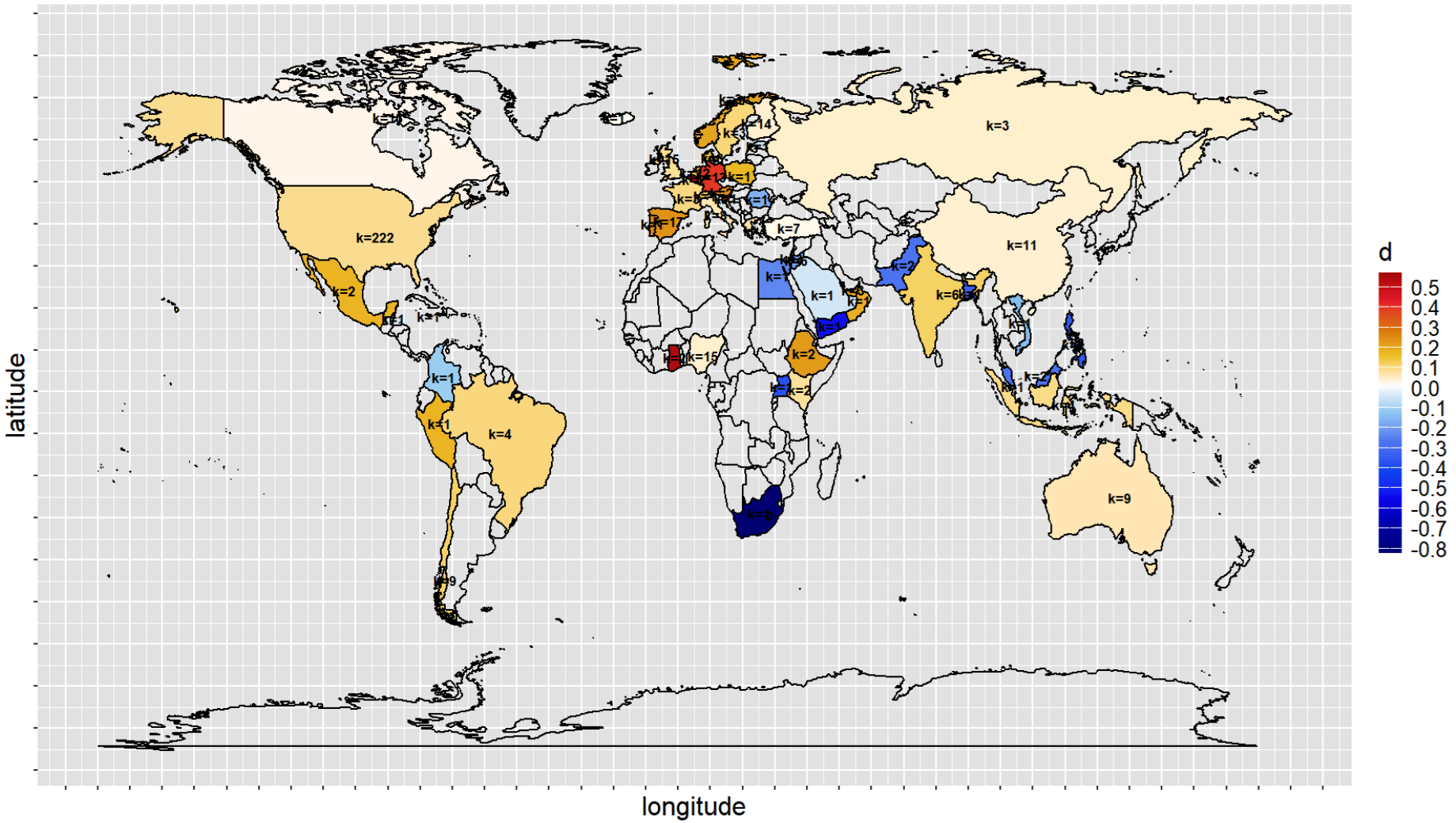
you may even create greeting cards



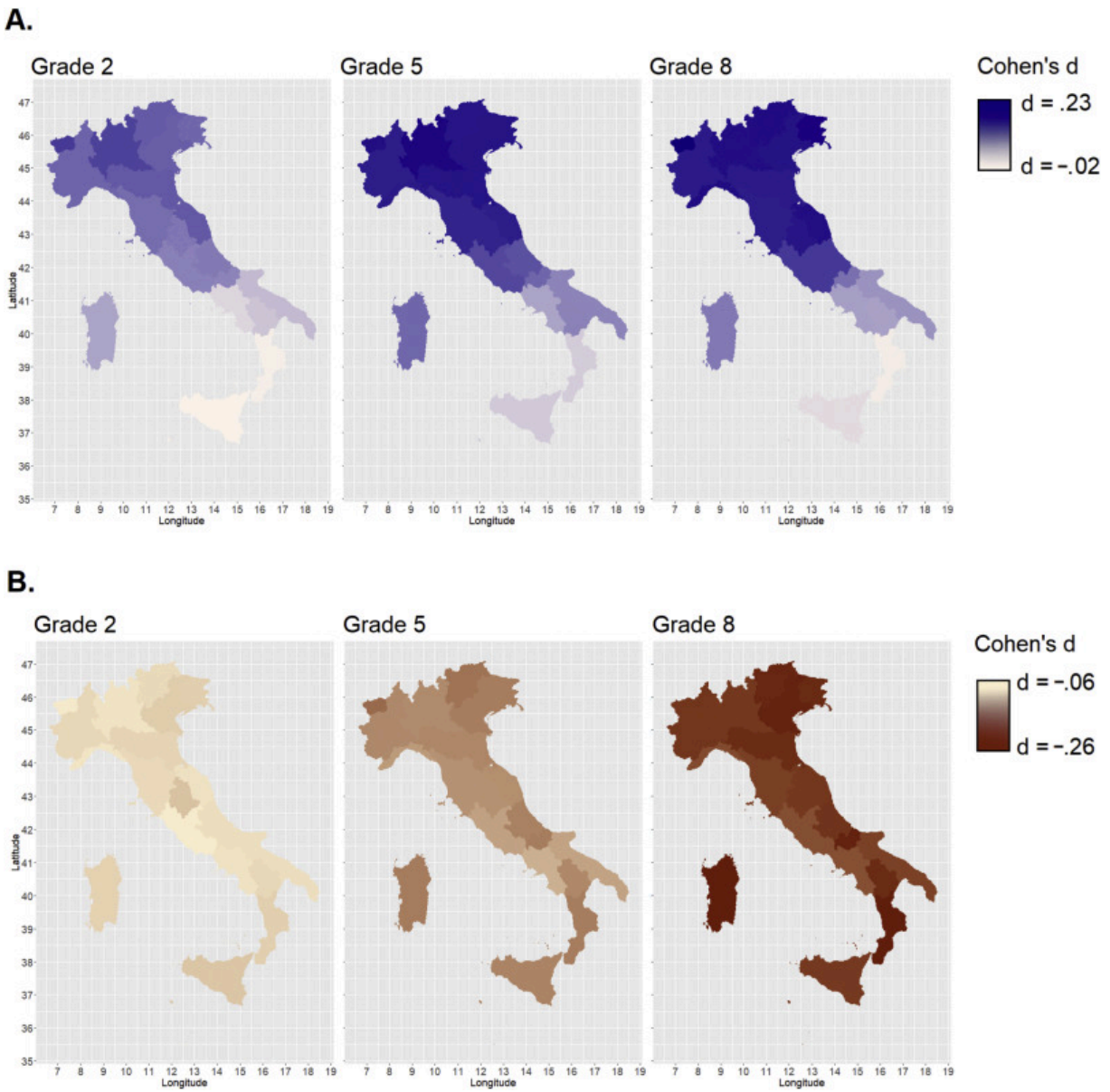
you may even create greeting cards



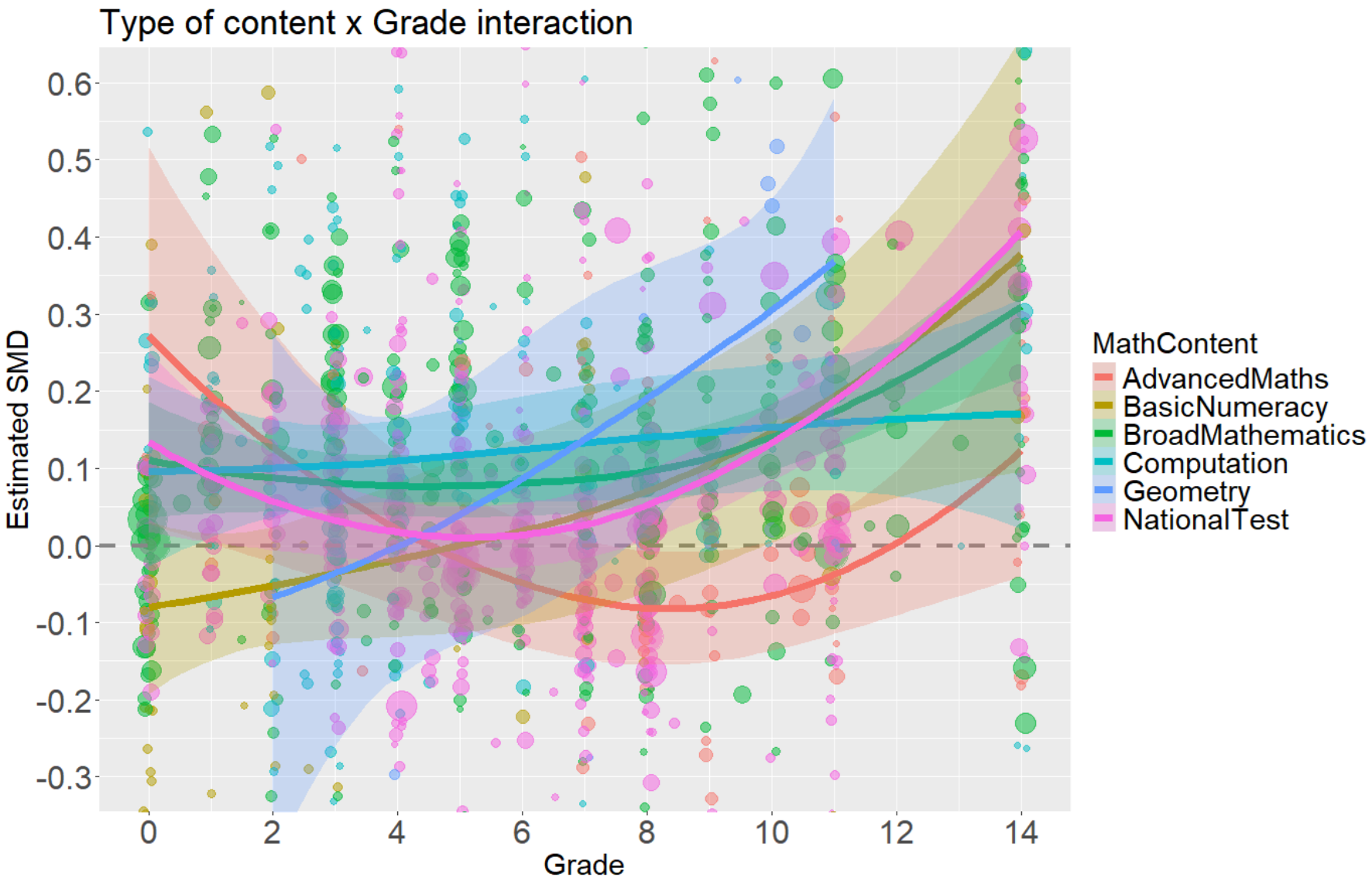
or like fancy infographics



or like fancy infographics



perform classical data analysis



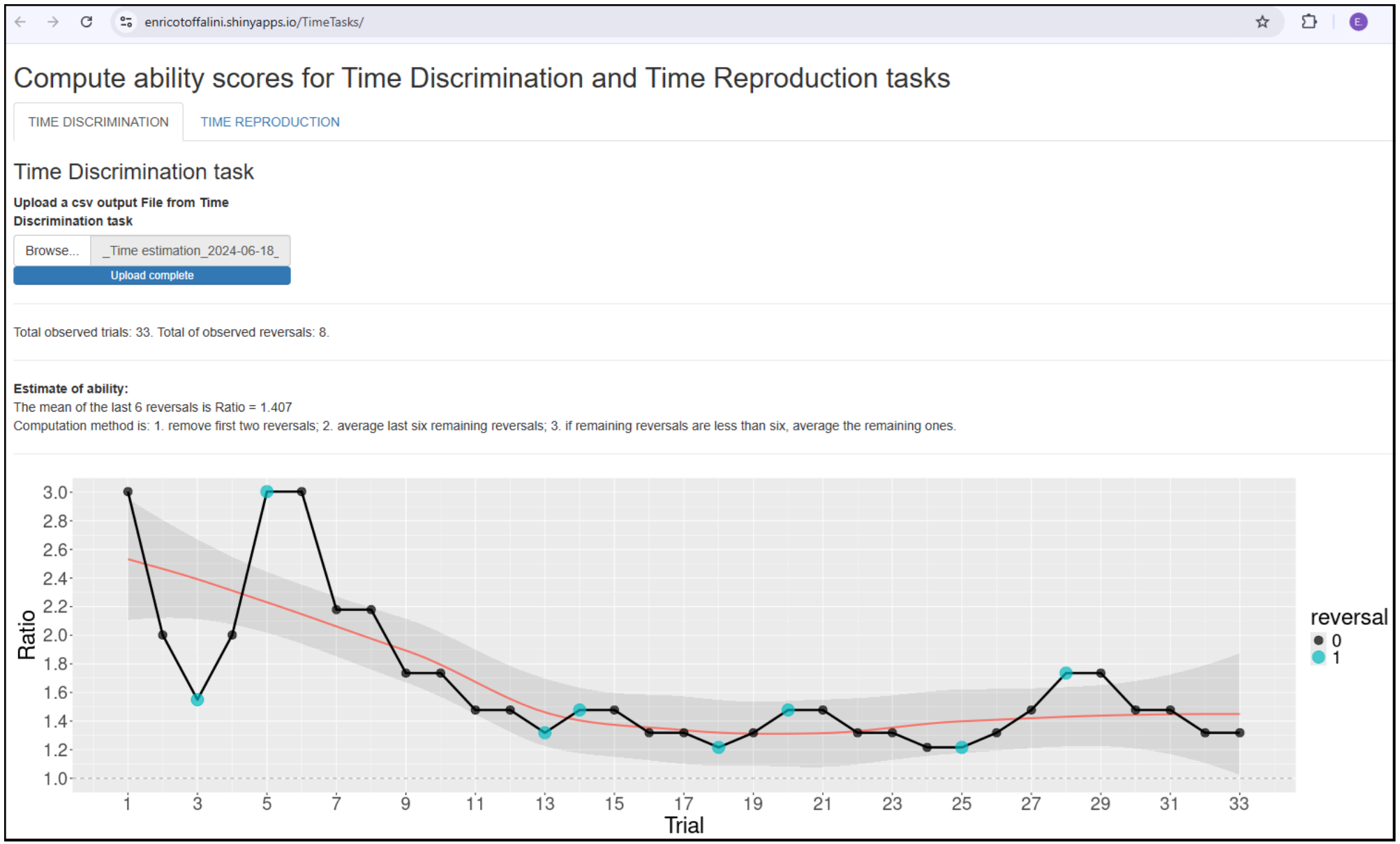
you may create interactive webapps with Shiny

see **Shiny gallery**

here's a couple of recent real examples from **Psicostat** members:

- this game-like shiny app developed for the science4all event in Padova; see here some explanation in Italian
- practical ad-hoc shiny app for scoring experimental data collected by students

you may create interactive webapps with Shiny



or entire websites and books

examples of other resources that can be **created within the R ecosystem**, integrating other tools such as *GitHub* and *Quarto*:

- this very course support material is a website in its own right
- this very course textbook is a book/website
- this book by Daniël Lakens explaining Statistical Inference

R and its *Integrated Development Environment*(s)

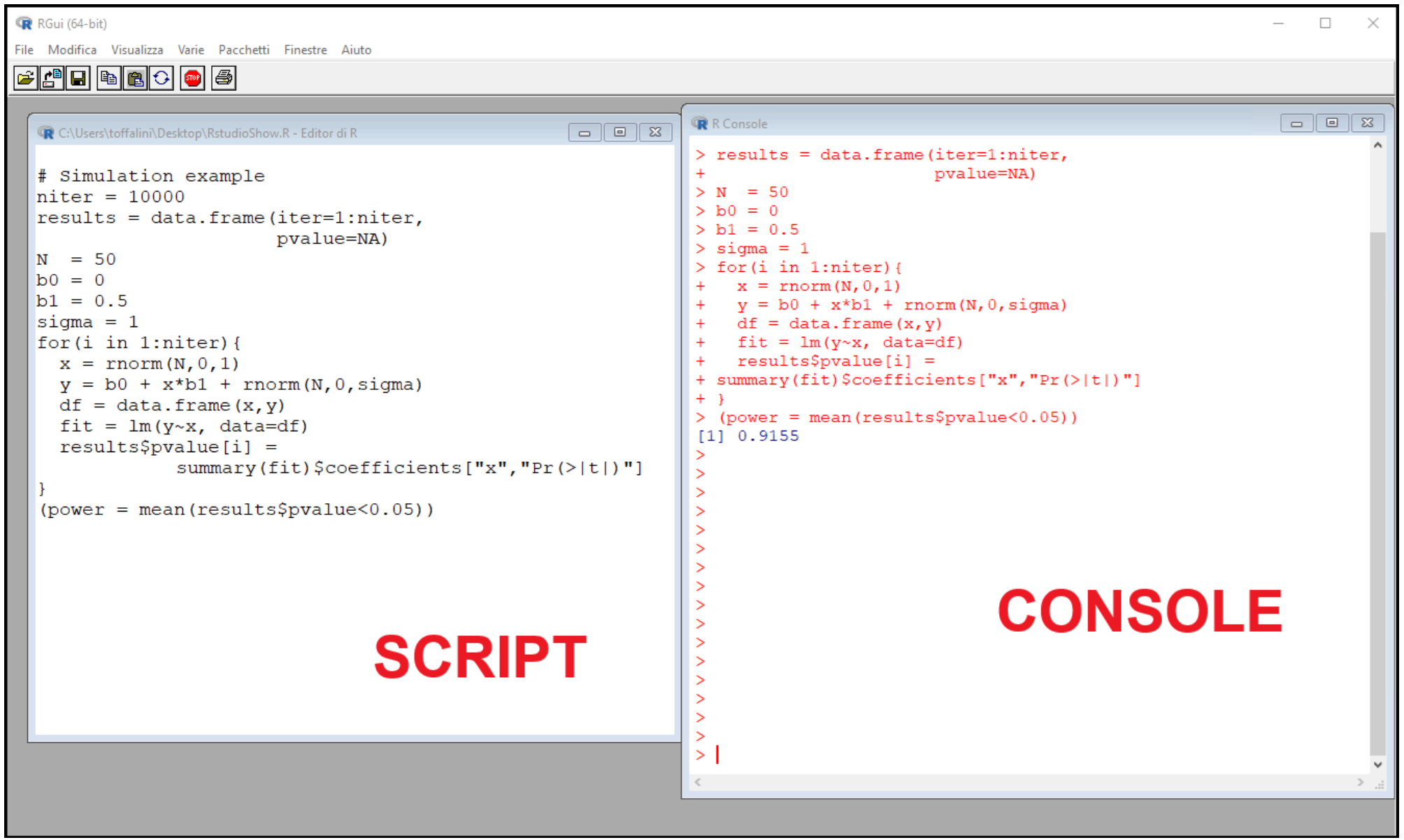
Make sure you install:

- **R** as the programming language interpreter and the basic environment and packages
- **RStudio** is the IDE of election to make writing R code easy

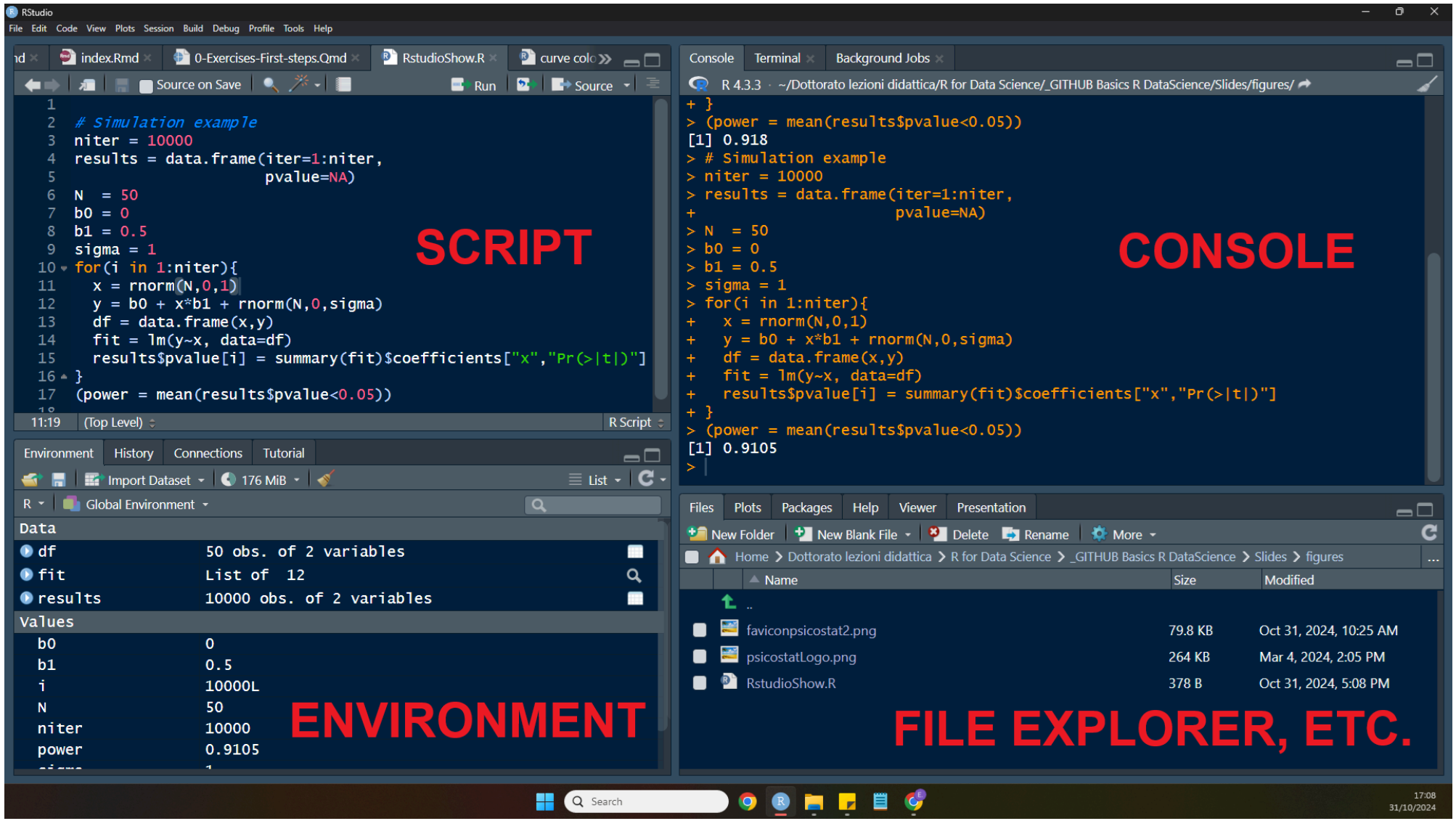
Interesting alternatives to installing RStudio:

- **Positron** (based on MS Visual Studio)
- **Posit.cloud** (fully online, actually RStudio)
- **Google Colab** (fully online, make sure to set *R runtime type*; actually a Jupyter notebook)

R Console (just base R)



R Studio (full IDE)



Some R packages that you will or may need in the future (1/3)

Package	Used for what	Example(s) of functions
`base` (base R)	Basic functions	`sum`, `mean`, `sqrt`, `abs`, `c`, `data.frame`, `summary`, `scale`, `plot`, `+`, `-`
`stats` (base R)	Basic statistical calculations and functions	`sd`, `cor`, `cor.test`, `t.test`, `lm`, `glm`, `AIC`, `rnorm`, `rbinom`
`graphics` (base R)	Basic statistical calculations and functions	`plot`, `boxplot`, `hist`, `barplot`

(You may actually use these "base" packages very often without even realizing that they are packages)

Some R packages that you will or may need in the future (2/3)

Package	Used for what	Example(s) of functions
`effectsize`	Compute different effect sizes	`cohens_d`, `hedges_g`, `cohens_f`, `d_to_r`
`lme4`	Fitting (generalized) (non-)linear mixed-effects models	`lmer`, `glmer`, `ranef`
`performance`	Useful tools for models	`check_collinearity`, `r2_nagelkerke`, `icc`
`effects`	Display effects for various statistical models	`allEffects`
`emmeans`	Estimate marginal means for various models	`emmeans`

Some R packages that you will or may need in the future (3/3)

Package	Used for what	Example(s) of functions
`ggplot2`	Create beautiful plots using The Grammar of Graphics	`ggplot`, `geom_point`, `geom_line`, ...
`lavaan`	Structural Equation Models (SEM)	`sem`, `cfa`
`semTools`	Useful tools for SEMs	`compRelSEM`, `measEq.syntax`
`metafor`	Perform meta-analysis	`rma`, `rma.mv`, `forest`, `funnel`, `regtest`
`brms`	Fitting practically any Bayesian model via MCMC with STAN	`brm`, `set_prior`
`blavaan`	Fitting Bayesian SEMs	`bcfa`, `bsem`

Let’s Test the Environment!

Let’s run a few commands in RStudio to familiarize with its console and see if the installation works properly

```
rnorm(10) # draw 10 random values from a Standard Normal distribution
```

```
 [1] -0.6790134  0.5173610 -0.9890813 -0.6538885 -2.8993340 -1.1826100
 [7]  0.3177626 -0.9854805  0.5988272 -1.2425435
```

```
?rnorm # open the help tab for the "rnorm" function
```

```
round( rnorm(10, mean=100, sd=15) ) # draw 10 values from IQ distribution, round them
```

```
 [1]  89 129  86  95 111  94 100 104  76 101
```

```
install.packages("psych") # install a package from CRAN
```

```
library(psych) # load the newly installed package

fisherz(rho=0.5) # use it to transform a correlation into a Fisher's z
```

```
[1] 0.5493061
```