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.

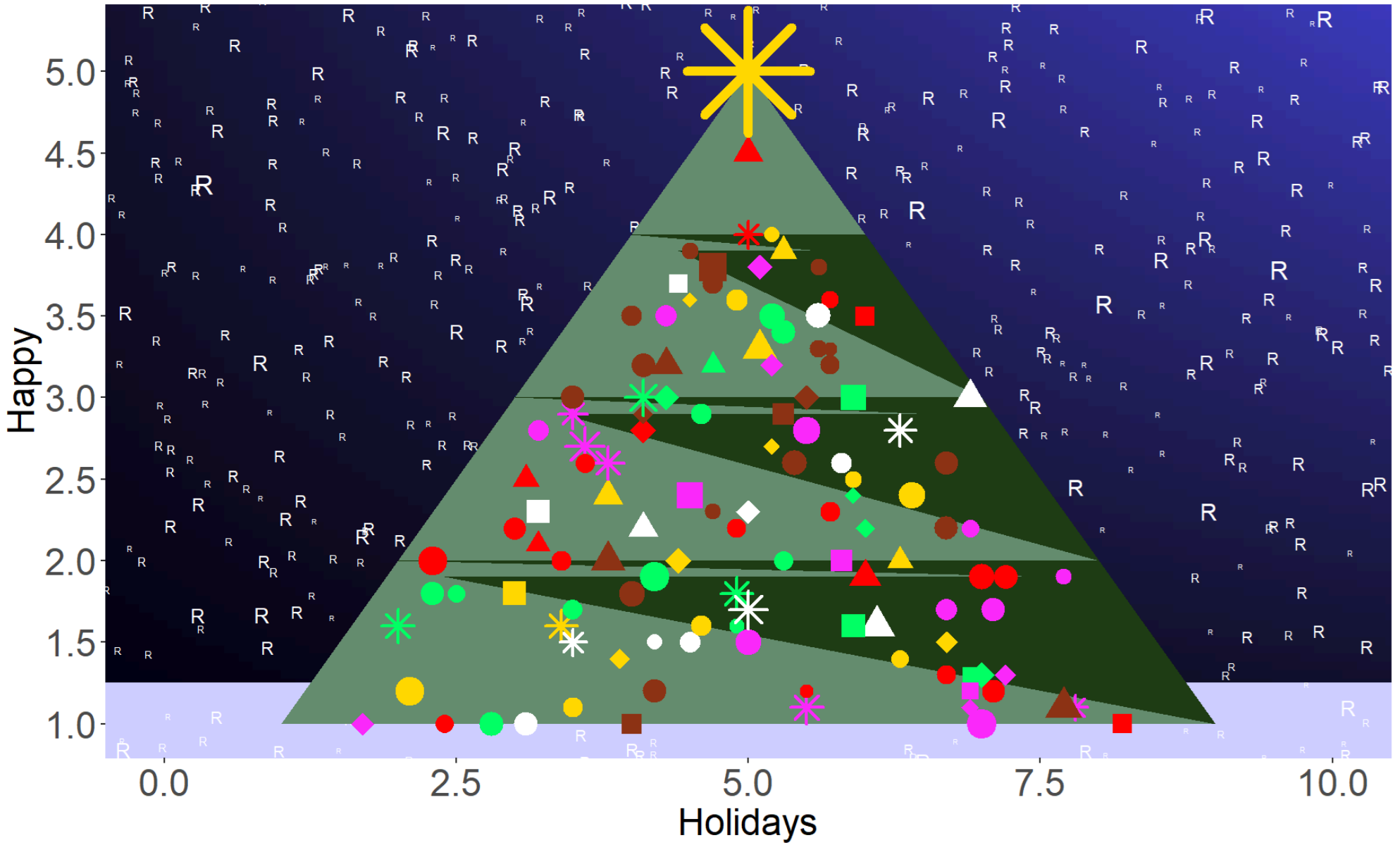
What you may expect to learn in this course:

- 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.

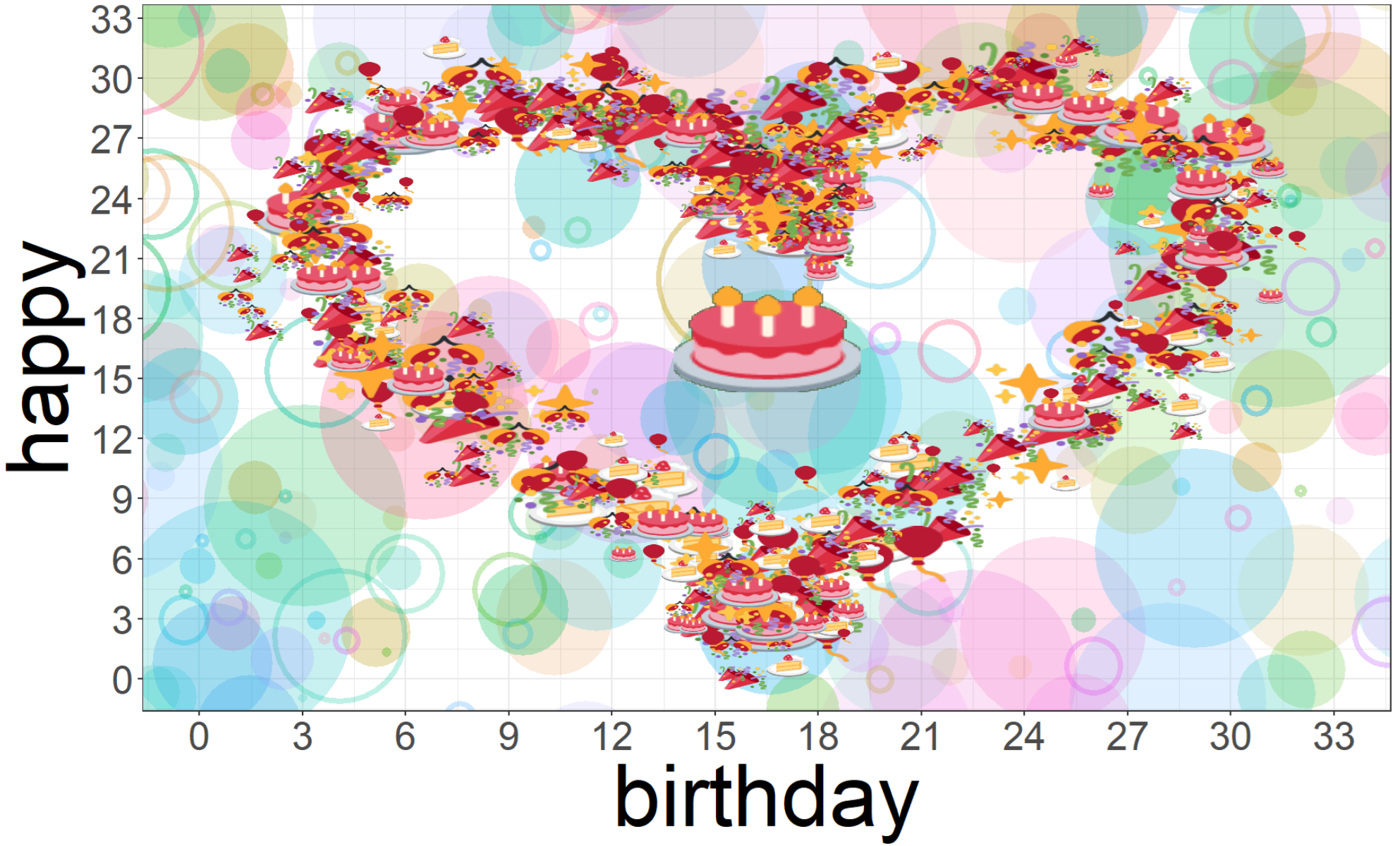
Over the next *couple of years*, following this PhD program, you may have the opportunity to learn to use R to perform at least some fundamentals about:

- 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.

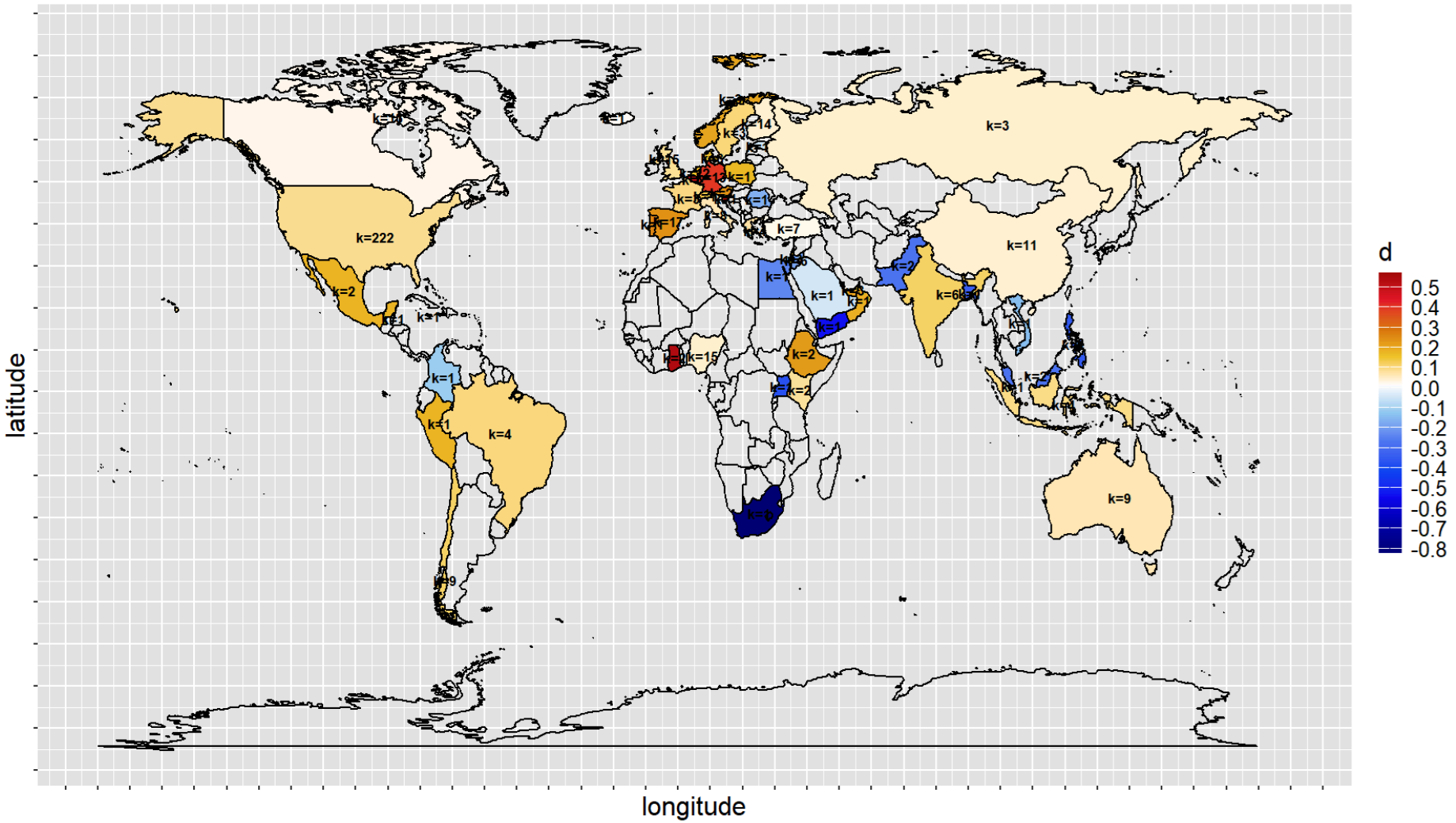
you may even create greeting cards



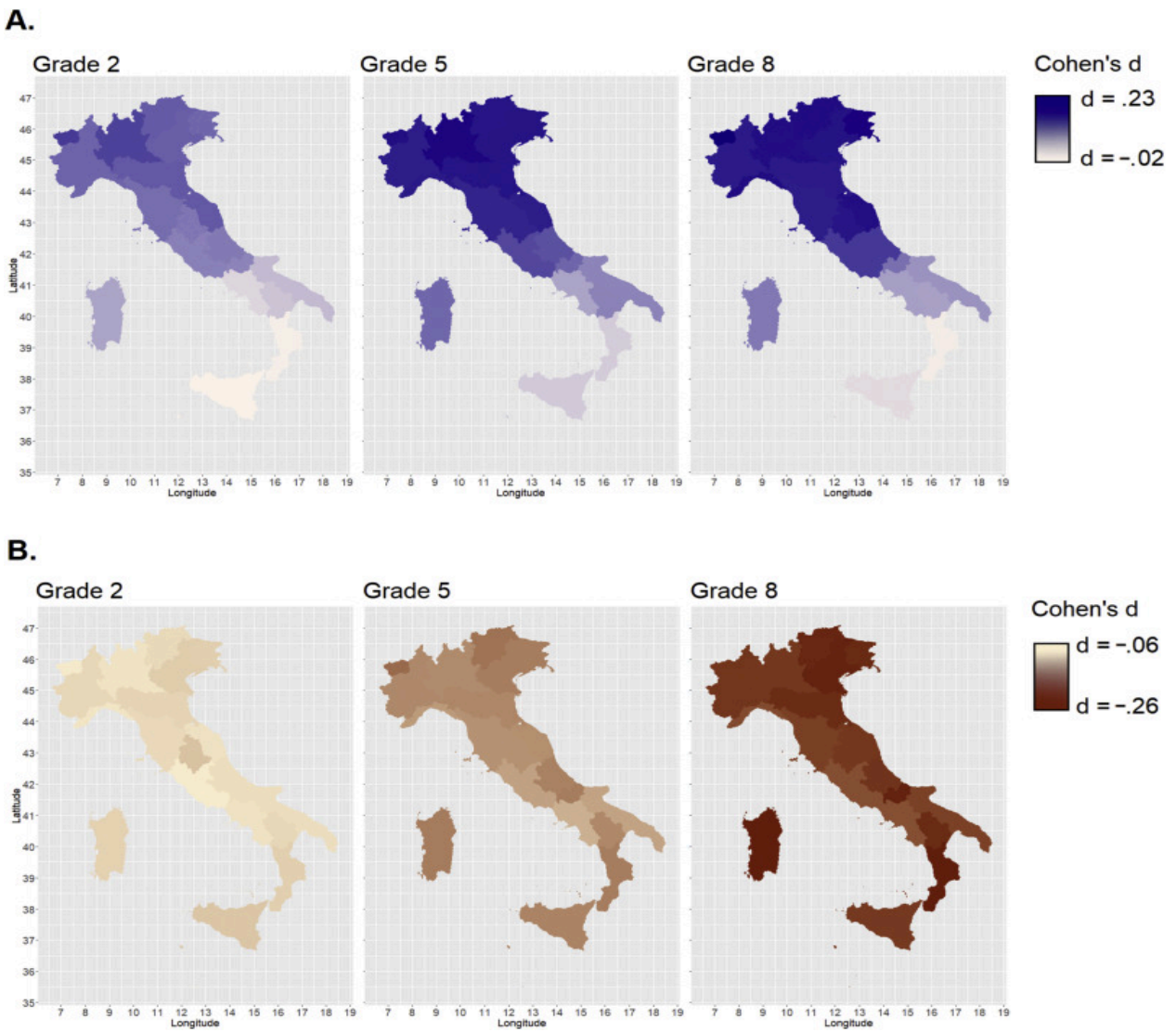
you may even create greeting cards



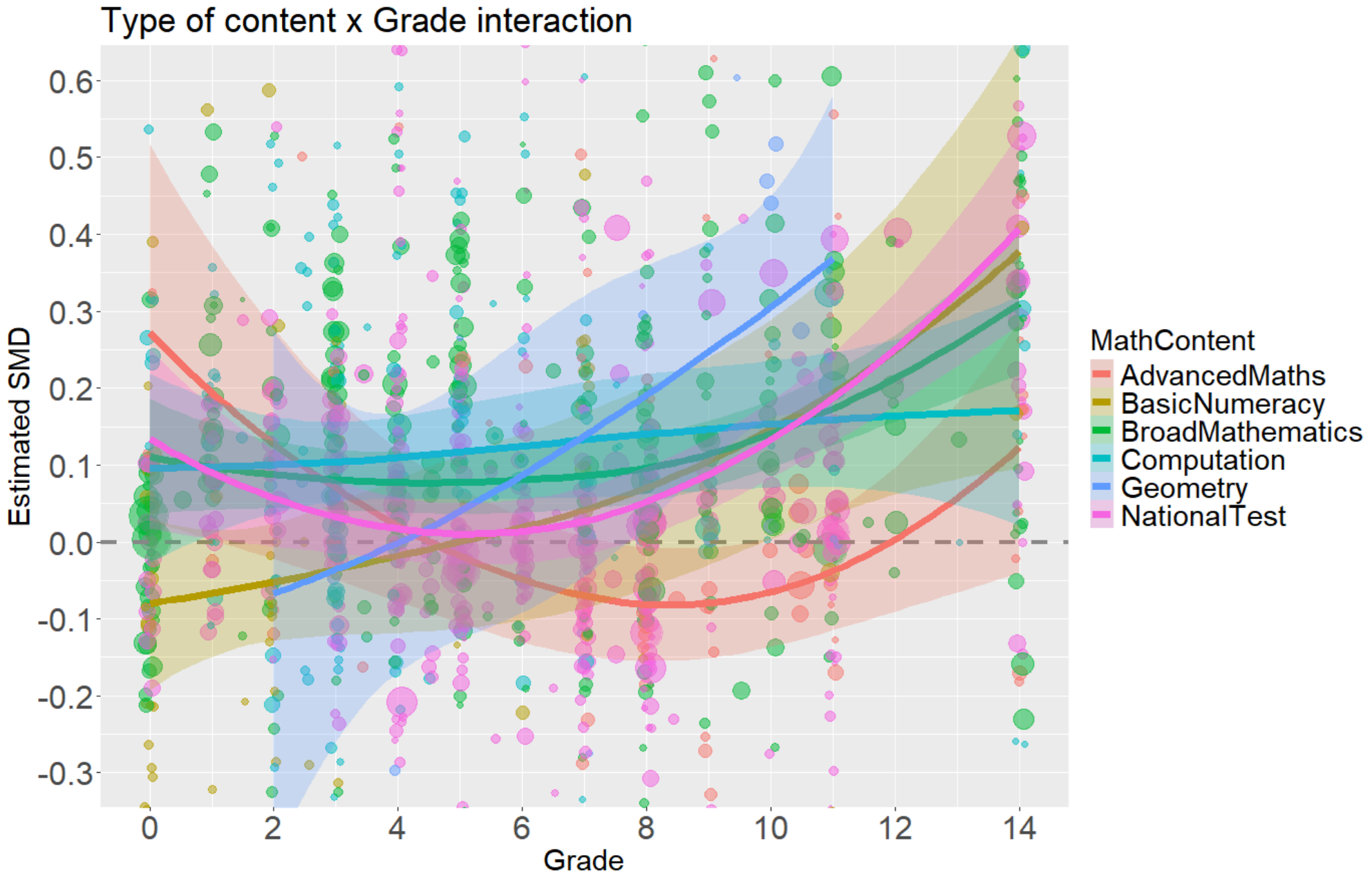
or like fancy infographics



or like fancy infographics



perform classical data analysis


Type of content x Grade interaction
Estimated SMD
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
0
2
4
6
8
10
12
14
Grade
MathContent
AdvancedMaths
BasicNumeracy
BroadMathematics
Computation
Geometry
NationalTest

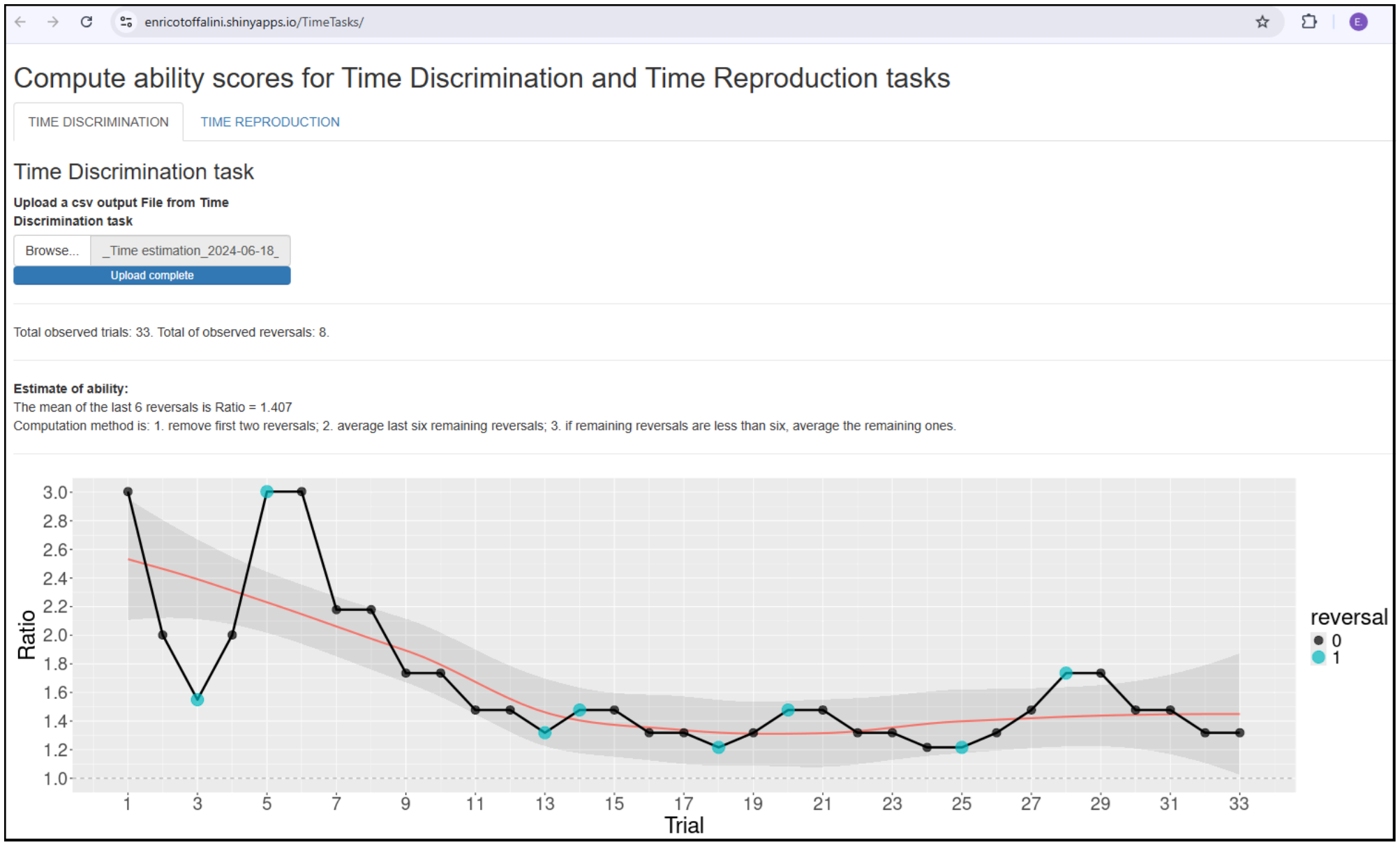
you may create interactive webapps with Shiny

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

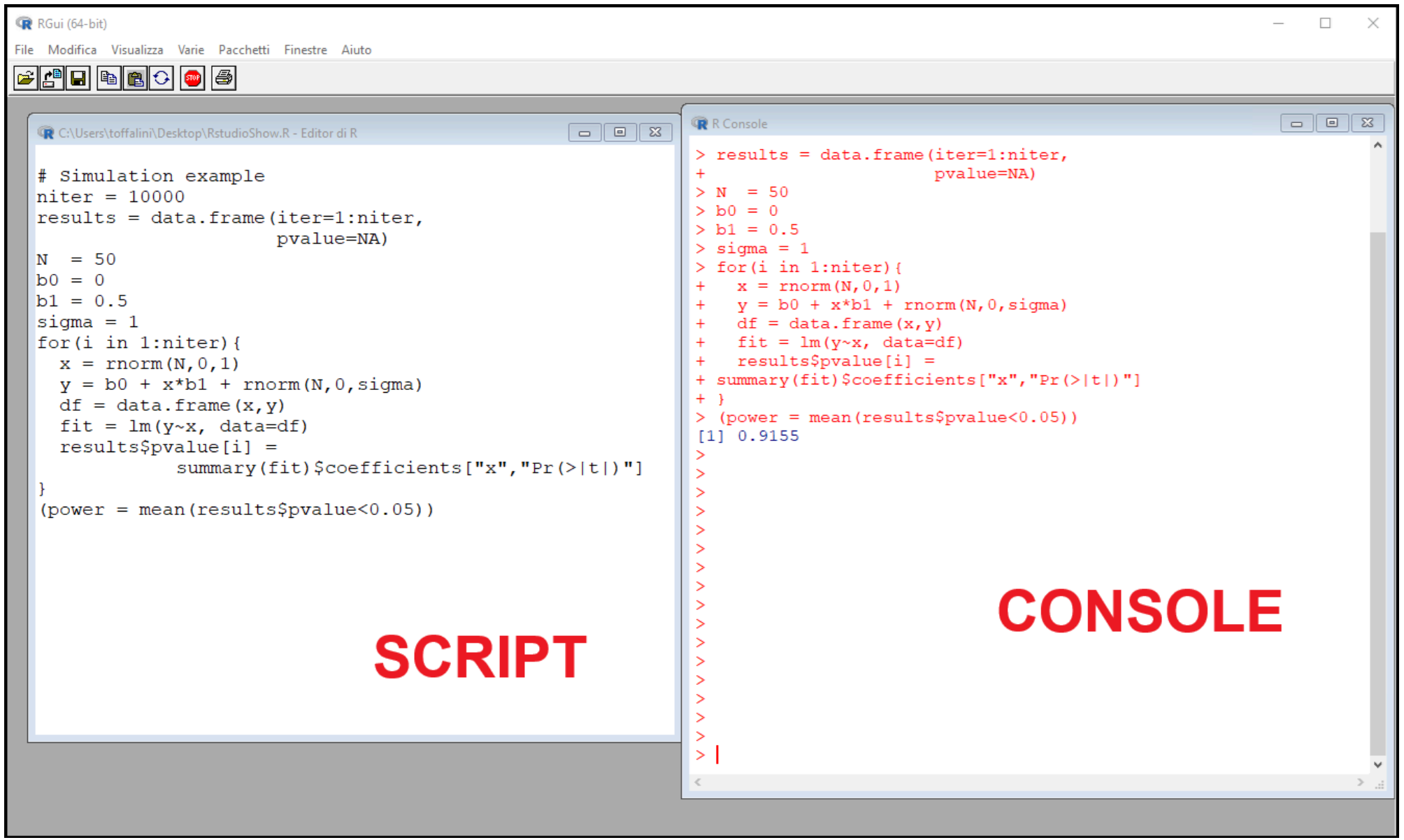
- this entire course is a website in its own right
- the course textbook is a book/website
- also see this wonderful book by Daniël Lakens explaining Statistical Inference

of course, these resources integrate other tools such as GitHub and Quarto, but they can be created within the R ecosystem

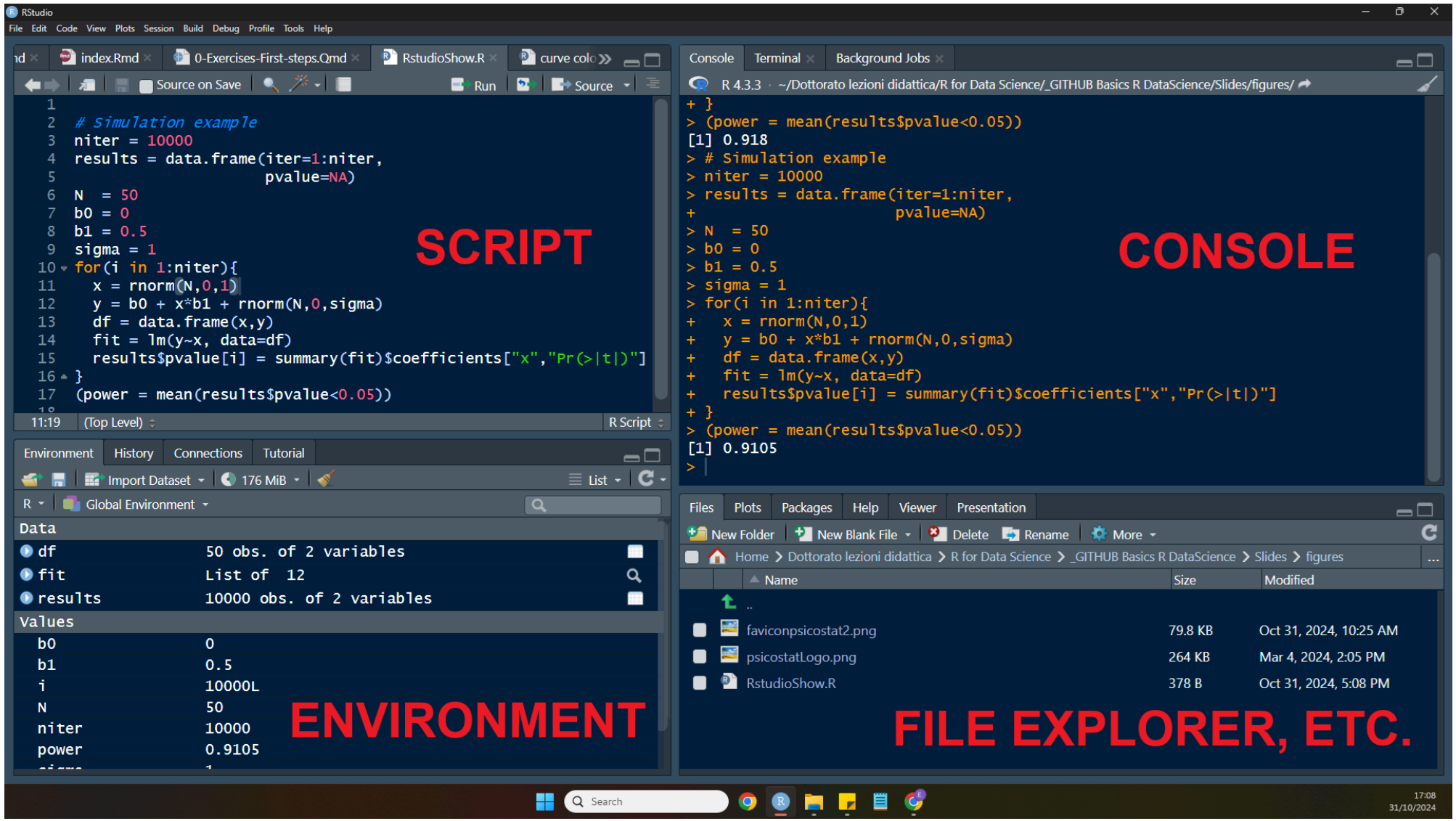
install R and Rstudio

first of all, for getting started, follow the instructions in *Chapter 1* of *Introduction2R* to ensure that both R and RStudio are installed

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	`boxplot`, `hist`, `barplot`
`effectsize`	Compute different effect sizes	`cohens_d`, `hedges_g`, `cohens_f`, `d_to_r`

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

Package	Used for what	Example(s) of functions
`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`
`ggplot2`	Create beautiful plots using The Grammar of Graphics	`ggplot`, `geom_*`

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

Package	Used for what	Example(s) of functions
`lavaan`	Structural equation modeling (SEM)	`cfa`, `sem`
`semTools`	Useful tools for SEMs	`reliability`
`metafor`	Perform meta-analysis	`rma`, `rma.mv`, `forest`, `funnel`, `regtest`
`brms`	Fitting practically any Bayesian model via MCMC with STAN	`brm`, `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]  0.80615400 -0.58547491  1.10525657 -0.40379478 -0.98313925  0.06539615
 [7]  0.81912540  1.60196277  2.44925450 -0.44644926
```

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

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

```
 [1]  92 103 140  99  86  86  99  73  97  83
```

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

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

fisherz(rho=0.9) # use it to transform a correlation into a Fisher's z
```

```
[1] 1.472219
```