

STAT475/675 TUT01

www.sfu.ca/~zza115/teaching

Zhiyang Zhou (zhiyang_zhou@sfu.ca)

2018-01-09

R is free

- Open-source
- Citation
- Installation
 - download and install BASE *R* from <https://cran.r-project.org>
 - download and install *Rstudio* from <https://www.rstudio.com>
 - download and install packages via *Rstudio*
- Working directory
- Packages
 - installation: `install.packages()`
 - loading: `library()`
- Help manual: `help()`, `?`, google, etc.

R is powerful

- Data types: let `str()` or `class()` tell you
 - numbers (integer, real, or complex)
 - characters (“abc”)
 - logical (TRUE or FALSE)
 - date & time
 - factor (commonly encountered in this course)
 - NA (different from Inf, “”, 0, NaN etc.)
- Data structures: let `str()` or `class()` tell you
 - vector: an ordered collection of the same data type
 - matrix: two-dimensional collection of the same data type
 - array: more than two dimensional collection of the same data type
 - data frame: collection of vectors of same length but of arbitrary data types
 - list: collection of arbitrary objects
- Data input and output
 - create
 - * vector: `c()`, `seq()`, `rep()`
 - * matrix: `matrix()`, `cbind()`, `rbind()`
 - * data frame
 - output: `write.table()`, `write.csv()`, `write.xlsx()`
 - import: `read.table()`, `read.csv()`, `read.xlsx()`
 - * header: whether or not assume variable names in first row
 - * stringsAsFactors: whether or not convert character string to factors

- `scan()`: a more general way to input data
- Simple data manipulation
- Basic plots
 - strip chart, histogram, box plot, scatter plot
 - Package `ggplot2` (RECOMMENDED)

***R* is not cheap — by Carl James Schwarz**

- NO quality control
- requires statistical sophistication
- time-consuming to become a master