The R Project for Statistical Computing

R Language Data IO
資料輸入輸出

vvn Weian Chao (趙韋安)

https://ce.nctu.edu.tw/member/teachers/23

Department of Civil Engineering, National Yang Ming Chiao Tung University, Taiwan

R: An Introduction
Accessing data.

Entering data

Small amount of data

Many methods

c, concatenate
seq
rep
scan
…

TRY it in R

R: An Introduction Accessing data.

R_data_io_a.R

R: An Introduction
Accessing data.

Print data on screen

Many methods

print
cat
sprintf → c語言
…

print

print(“any string”)
print(variable)
print(paste(“any string” , variable))

cat

```
cat(...,
file = "",
sep = " ",
fill = FALSE,
labels = NULL,
append = FALSE)
```

跳脫字元

\n: new line
\t: tab
\b: backspace

sprintf

sprintf(“format” ,…)

format

d: 整數
f: 雙精浮點數
e: 雙精指數
s: 字串
%m.n
%-

TRY it in R

R: An Introduction Accessing data.

R_data_io_b.R

R: An Introduction
Accessing data.

Importing data

R has built-in functions for importing data in many formats

Make it easy

rio

https://cran.r-project.org/web/packages/rio/rio.pdf

rio combines All of import functions into one simple utility

TXT

CSV

XLSX

JSON

Other methods

read.table
read.csv
…

```
read.table( file,
header = FALSE,
skip = 0,
comment.char = “#” ,
append = FALSE,
row.names = TRUE,
col.names = TRUE
encoding = “unknown”
)
```

R: An Introduction
Accessing data.

convert.

TRY it in R

R: An Introduction Accessing data.

R_data_io_c.R

R: An Introduction
Accessing data.

Exporting data

Methods

export
write.table
write.csv
…

TRY it in R

R: An Introduction Accessing data.

R_data_io_d.R

R: An Introduction
Accessing data.

Data in R

Data sets in package "datasets"

datasets

https://vincentarelbundock.github.io/Rdatasets/

vincentarelbundock.github.io/Rdatasets/

Rdatasets 1.0.0 Available datasets

What is this?

`Rdatasets` is a collection of over 1700 datasets that were originally distributed alongside the statistical software environment `R` and some of its add-on packages. The goal is to make these data more broadly accessible for teaching and statistical software development.

What is included?

The list of available datasets (csv and docs) is available here:

- HTML index
- CSV index

On the github repository you will also find:

- `Rdatasets.R`: `R` script to download CSV copies and HTML docs for all datasets distributed in `Base R` and a list of R packages.

github.com/vincentarelbundock/Rdatasets/tree/master/csv/datasets

Why GitHub? Team Enterprise Explore Marketplace Pricing

vincentarelbundock / Rdatasets Public

Code Issues 2 Pull requests Actions Projects Wiki Security Insights

master Rdatasets / csv / datasets /

vincentarelbundock openintro causaldata HMLdiag

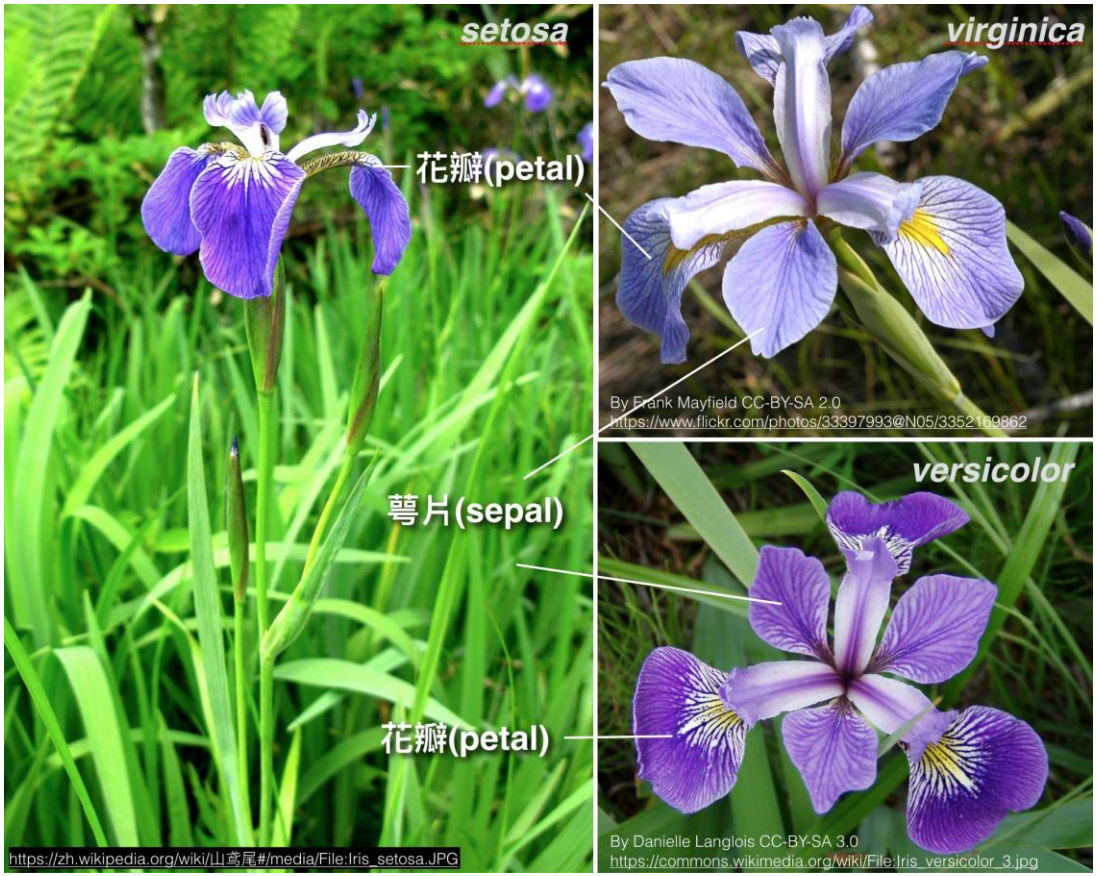
..

AirPassengers.csv	reboot
BJsales.csv	reboot
BOD.csv	reboot
CO2.csv	openintro causaldata HMLdiag
ChickWeight.csv	reboot

iris data set

-三種鳶尾花的形態特徵

Setosa, Versicolor, Virginica



TRY it in R

R: An Introduction Accessing data.

R_data_io_e.R

R: An Introduction
Basic graphics.

plot()

Basic X-Y plotting

TRY it in R

R: An Introduction
Basic graphics.

R_data_io_f.R

Setting	Description
adj	justification of text
ann	draw plot labels and titles?
bg	"background" color
bty	type of box drawn by `box()`
cex	size of text (multiplier)
cex.axis	size of axis tick labels
cex.lab	size of axis labels
cex.main	size of plot title
cex.sub	size of plot sub-title
col	color of lines and data symbols
col.axis	color of axis tick labels
col.lab	color of axis labels
col.main	color of plot title
col.sub	color of plot sub-title
fg	"foreground" color
font	font face (bold, italic) for text
font.axis	font face for axis tick labels
font.lab	font face for axis labels
font.main	font face for plot title
font.sub	font face for plot sub-title
gamma	gamma correction for colors
lab	number of ticks on axes
las	rotation of text in margins
lty	line type (solid, dashed)
lwd	line width
mgp	placement of axis ticks and tick labels
pch	data symbol type
srt	rotation of text in plot region
tck	length of axis ticks (relative to plot size)
tcl	length of axis ticks (relative to text size)
tmag	size of plot title (relative to other labels)
type	type of plot (points, lines, both)
xaxp	number of ticks on x-axis
xaxs	calculation of scale range on x-axis
xaxt	x-axis style (standard, none)
xpd	clipping region
yaxp	number of ticks on y-axis
yaxs	calculation of scale range on y-axis
yaxt	y-axis style (standard, none)

R: An Introduction
Basic graphics.
Bar charts

Simple is good

TRY it in R

R: An Introduction
Basic graphics.

R_data_io_g.R

R: An Introduction
Basic graphics.
Histograms

See what you have

Shape

Gap

Outlier

Symmetry

TRY it in R

R: An Introduction
Basic graphics.

R_data_io_h.R

R: An Introduction
Basic graphics.
Scatterplots

For visualizing the association between two quantitative variables

What to look for?

Linear

Spread

Outliers

Correlation

TRY it in R

R: An Introduction
Basic graphics.

R_data_io_i.R

R: An Introduction
Basic graphics.
Overlaying plots

Increased Information density

TRY it in R

R: An Introduction
Basic graphics.

R_data_io_j.R

Pictures first & Number later

課堂練習1: 學號-姓名-IO.R

根據氣象站資料檔weatherdata.xlsx，請試著根據資料回答以下問題:
(a)資料中標籤名稱為Rain的最大單筆雨量紀錄值是多少? 發生的日期時間為何? 共計幾筆?

課堂練習1: 學號-姓名-IO.R

(b)請試著產生以下資料格式的TXT文字檔案，檔案名稱為:

Output.txt

資料格式

第一行:共計幾筆?

第二行:最大單筆雨量

第三行:最大單筆雨量之發生日期時間