



105-1: EE4052
計算機程式設計
Computer Programming

Unit 06: 檔案資料輸入與輸出

連 豐 力

臺大電機系

Sep 2016 - Jan 2017



大綱

計算機程式設計 - 2016F
Chap 06: 資料輸入與輸出
Feng-Li Lian @ NTU-EE

- 取得檔案的資料：輸入與輸出
- 取得內建資料庫
- 取得其他資料庫的資料
- 取得網頁的資料

檔案資料輸入與輸出

3



設定工作目錄

計算機程式設計 - 2016F
Chap 06: 資料輸入與輸出
Feng-Li Lian @ NTU-EE

- `rm(list = ls())`
- `ls()`

- `mywd <- "L:/DataWD"`

- `setwd(mywd)`

- `getwd()`

- 4

- % 儲存物件的名稱與其值 (內容)
- `x <- 1:10`
- `y <- matrix(1:6, nrow = 2, ncol = 3)`
- `ls()`
- `dump(c("x", "y"), file = "mydump.txt"`)
- `rm(x); rm(y)`
- `ls()`
- `source(file = "mydump.txt")`
- `ls()`

- % 儲存物件的值 (內容)
- `x <- 1:10`
- `y <- matrix(1:6, nrow = 2, ncol = 3)`
- `dput(y, file = "mydput.txt")`
- `newy <- dget("mydput.txt")`
- `newy`

- % 螢幕輸出轉到檔案

- sink("mysink.txt")

- x

- y

- sink()

- x

- y

- dataf <- iris[c(1, 2, 51, 52, 101, 102), c(1, 2, 5)]

- dataf <- edit(dataf)

- write.table(dataf, "mydataf.txt")

- df <- read.table("mydataf.txt")

- sf0 <- read.table("mydataf0.txt") % 去掉 "" 的資料

- sf1 <- read.table("mydataf1.txt", header = FALSE)

 % 去掉標題的資料

- sf2 <- read.table("mydataf2.txt", header = TRUE, row.names = NULL)

 % 去掉列名稱的資料

- sf3 <- read.table("mydataf3.txt", header = FALSE, row.names = NULL)

 % 去掉標題與列名稱

- % csv: comma separated values
- read.table("mydataf2.csv")
- read.table("mydataf2.csv", header = TRUE)
- read.table("mydataf2.csv", header = TRUE, sep = ",")
- read.table("mydataf2.csv", header = TRUE, sep = ",")
- sf2_csv <- read.table("mydataf2.csv", sep = ",", header = TRUE)
- sf2_csv2 <- read.csv("mydataf2.csv")

- % 設定不同的目錄
- tphdata <- read.table("L:/DataWD/Typhoon-01.txt", header = TRUE)
- tphdata

- % 讀取大量資料的文字檔
- `scandata <- scan("scanlist.txt", list(Sepal.Length = 0, Sepal.Width = 0, Species = ""))`

- % 讀取資料
- `Ins_c1 <- read.csv("Insurance.csv")`
- `head(Ins_c1)`

- `Ins_c2 <- read.table("Insurance.csv")`
- `head(Ins_c2)`

- `Ins_c3 <- read.table("Insurance.csv", header = TRUE, sep=",")`
- `head(Ins_c3)`

- `Ins_t1 <- read.table("Insurance.txt")`
- `head(Ins_t1)`

- `Ins_t2 <- read.table("Insurance.txt", header = TRUE, sep="")`
- `head(Ins_t2)`

- R Programming/Importing and exporting data
 - https://en.wikibooks.org/wiki/R_Programming/Importing_and_exporting_data

- Getting Data From An Online Source
 - <https://www.r-bloggers.com/getting-data-from-an-online-source/>

內建資料庫

- % R 內建資料集 : datasets
- % 近百個資料集，涵蓋：醫學、自然、社會、人體等資料

- data(package = "datasets")
- help(AirPassengers)
- ?AirPassengers
- **AirPassengers** % Monthly Airline Passenger Numbers 1949-1960
- summary(AirPassengers)

- data(package = .packages(all.available = TRUE))

- % CO2 Carbon Dioxide Uptake in Grass Plants
- % uspop Populations Recorded by the US Census
- % Titanic Survival of passengers on the Titanic
- % women Average Heights and Weights for American Women

- % CO2 Carbon Dioxide Uptake in Grass Plants
- CO2
- summary(CO2)

- % uspop Populations Recorded by the US Census
- uspop
- summary(uspop)

- % Titanic Survival of passengers on the Titanic
- Titanic
- summary(Titanic)

- % women Average Heights and Weights for American Women
- women
- summary(women)

- % MASS
- library(MASS)
- data(Insurance)

- ?Insurance
- head(Insurance)
- tail(Insurance)
- dim(Insurance)
- names(Insurance)
- attributes(Insurance)

- class(Insurance\$District)
- class(Insurance\$Age)
- class(Insurance\$Holders)
- levels(Insurance\$Age)

- % arules 軟體套件中，Groceries 資料集

- install.packages(arules)
- library(arules)
- data(Groceries)

- ?Groceries

- Groceries[1:10]

- inspect(Groceries[1:10]) % 10位消費者購物車中的商品

其他資料庫的資料

19



網路上的資料庫

計算機程式設計 - 2016F
Chap 06: 資料輸入與輸出
Feng-Li Lian @ NTU-EE

- The R Datasets Package
 - <https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/00Index.html>
- UC Irvine Machine Learning Repository
 - <https://archive.ics.uci.edu/ml/datasets.html>
- The Free Datasets at r-dir.com
 - <http://r-dir.com/reference/datasets.html>
- Rdatasets: An archive of datasets distributed with R
 - <http://vincentarelbundock.github.io/Rdatasets>
- Datasets in R packages (IAState)
 - http://www.public.iastate.edu/~hofmann/data_in_r_sortable.html

- 20

網頁的資料

21



讀取資料 - 從網路的檔案

計算機程式設計 - 2016F
Chap 06: 資料輸入與輸出
Feng-Li Lian @ NTU-EE

- Reading a CSV-file from an URL
- The number of police officers in Scotland over time
- http://www.quandl.com/api/v1/datasets/EUROSTAT/CRIM_PLCE_42.csv
- `read.csv("http://www.quandl.com/api/v1/datasets/EUROSTAT/CRIM_PLCE_42.csv")`

- 和訊網 萬科 股票的相關金融資料
- `http://stockdata.stock.hexun.com/2008en/zxcwzb.aspX?stockid=000002&type=1&date=2013.06.30`
- `install.packages("XML")`
- `library(XML)`
- `url = "http://stockdata.stock.hexun.com/2008en/zxcwzb.aspX?stockid=000002&type=1&date=2013.06.30"`
- `tales1 = readHTMLTable(url)`
- `tales1`

- 和訊網 萬科 股票的相關金融資料
- `ls()`
- `names(tales1)`
- `tales1[[1]]`
- `tales1[[2]]`
- `tales1[[3]]`
- `tales1[[4]]`
- `aa <- tales1[[2]]`
- `aa`
- `class(aa)`
- `aa[1]`
- `aa[2]`
- `aa[3]`

- 課程網頁
- <http://cc.ee.ntu.edu.tw/~fengli/Teaching/Computer/index.html>

- `install.packages("RCurl")`
- `library(RCurl)`

- `URL <- "http://cc.ee.ntu.edu.tw/~fengli/Teaching/Computer/index.html"`

- `webdata <- getURLContent(URL)`

- `hold <- strsplit(webdata, "")[[1]]`

作業

HW05：檔案資料輸入與輸出

計算機程式設計 - 2016F
Chap 06: 資料輸入與輸出
Feng-Li Lian @ NTU-EE

On 10/25, 2016

- 假設您預計購買一個手機，
從網路上找到一些手機的資料，
如下表所示：

廠牌	Brand	Apple	HTC	HTC	ASUS	ASUS
型號	Model	iPhone7	OneM8	OneS9	ZenFoneDeluxe	ZenFoneZoom
價格 (元)	Price	24500	21900	9990	8990	15990
螢幕 (吋)	Screen	4.7	5	5	5.5	5.5
重量 (克)	Weight	138	160	158	170	185
記憶體 (GB)	GB	32	16	32	16	64
日期 (年月日)	Date	20160916	20140328	20160617	20150827	20151201

- 這筆資料，已經整理成兩個檔案：(請到課程網站下載到您的工作目錄)
 - HW_U05_Phone.txt
 - HW_U05_Phone.csv

- 27

HW05：檔案資料輸入與輸出

計算機程式設計 - 2016F
Chap 06: 資料輸入與輸出
Feng-Li Lian @ NTU-EE

On 10/25, 2016

- 編輯一個程式於 .R 檔，完成下面的工作：
 - 用 `read.table` 取讀取 HW_U05_Phone.txt 中的資料，指定到：`myphonetxt`
 - 用 `read.csv` 取讀取 HW_U05_Phone.csv 中的資料，指定到：`myphonecsv`
 - 用 `read.table` 或 `read.csv` 去取讀取下面網址的檔案，指定到：`myphoneweb`
`http://cc.ee.ntu.edu.tw/~fengli/Teaching/Computer/HW_U05_Phone.web`
 - 比較一下：`myphonetxt`，`myphonecsv` 與 `myphoneweb`，這三個物件，
看看：內容是否相同？格式是否相同？
 - 如果不同，可否重新讀入，
或者重新指定到新的物件，使得這三個的內容與格式都相同。
 - 建立一個 5x3 的矩陣 (matrix)：`Number`，放置五個手機的價格，螢幕，重量
三種資料
- 把執行的過程，以及產生的數據等，整理到報告檔 (pptx or pdf)。

- 28

HW05：檔案資料輸入與輸出

On 10/25, 2016

- 繳交下面檔案，檔案名稱：[HW05_學號](#)
 - 主要指定檔案：[HW05_B01921001.R](#)
 - 報告檔案：[HW05_B01921001.pptx](#) 或者 [.pdf](#)
- 繳交方式與期限：
 - E-mail 上面四個檔案到：ntucp105f@gmail.com
 - E-mail 主旨：[HW05_B01921001](#) (就是，作業編號_您的學號)
 - 繳交期限：[10/26 \(Wed\), 2016, 11pm 以前](#)
 - [HW05](#)，每位繳交自己的作業，或者兩人繳交一份
 - 兩人完成作業者，請註明工作模式，例如：
 - A 進行指令測試與報告編輯，然後給 B 重複測試
 - A 進行指令測試，B 進行報告編輯
 - A 進行指令測試與報告編輯，B 僅看過報告，但沒有實際測試
 - A 進行指令測試與報告編輯，B 沒有參與指令測試，也沒有看過報告