

R Data Reading

Young W. Lim

2018-02-10 Sat

1 Introduction

- References
- R Data Reading
- CSV
- Excel
- Database
- Other Tools
- R Binary Files
- Data in R
- Read from Web Sites

"R for Everyone - Advanced Analytics and Graphic" J. P. Lander

I, the copyright holder of this work, hereby publish it under the following licenses: GNU head Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled GNU Free Documentation License.

CC BY SA This file is licensed under the Creative Commons Attribution ShareAlike 3.0 Unported License. In short: you are free to share and make derivative works of the file under the conditions that you appropriately attribute it, and that you distribute it only under a license compatible with this one.

- CSVs (comma separated values)
- Excel data
- Databases
- other tools
- R binary files
- Data included with R
- Extract from Web Sites

Comma Separated Values

```
Url <- http://www.jaredlander.com/data/Tomato%20First.csv"
tomato <- read.table(file=Url, header=TRUE, sep=" ");
head(tomato)
```

```
x <- 10:1
y <- -4:5
q <- c("H", "F", "B", "C", "R", "L", "BB", "T", "CK", "S")
DF <- data.frame(First=x, Second=y, Sport=q,
stringsAsFactors=FALSE)
DF$Sport
```

Converting into a CSV file

- convert an Excel file into a CSV file
- gdata, XLConnect, xlsReadWrite, RODBC

```
require(RODBC)

db <- odbcConnect("QV Training")
orderedTB <- sqlQuery(db, "SELECT * FROM Orders",
  stringAsFactors=FALSE)
detailTB <- sqlQuery(db, "SELECT * FROM [Order Details]",
  stringAsFactors=FALSE)
longQuery <- "SELECT * FROM Orders, [Order Details]
WHERE Orders.OrderID = [Order Details].OrderID"
detailsJoin <- sqlQuery(db, longQuery,
  stringsAsFactors=FALSE)
```

Read from other tools

read.spss	SPSS
read.dta	Stata
read.ssd	SAS
read.octave	Octave
read.mtp	Minitab
read.systat	Systat

R Binary Files

```
save(tomato, file="data/tomato.rdata")
rm(tomato)
head(tomato)
load("data/tomato.rdata")
head(tomato)
n <- 0
r <- 1:10
w <- data.frame(n, r)
n
w
save(n, r, w file="data/multiple.rdata")
rm(n, r, w)
n
r
w
```

Data included in R

```
require(ggplot2)
data(diamonds)
head(diamonds)
```

```
require(XML)

Url <- "http://www/jaredlander.com/2012/02
/another-kind-of-super-bowl-pool")
bowlPool <- readHTMLTable(Url, which=1, header=FALSE,
  stringsAsFactors=FALSE)
bowlPool
```