

```

1  ##朝日新聞・東大共同調査:2014年衆院選-16年参院選世論調査
2  dat.vote <- ("http://www.masaki.j.u-tokyo.ac.jp/utas/2014_2016UTASV20161004.csv")
3  dat.vote <- read.table(dat.vote, header = T, sep = ",",
4                        fileEncoding = "CP932", #Mac user
5                        na = c("66", "99", "999")) #66:非該当,99.999:無回答
6
7  str(dat.vote)
8  colnames(dat.vote)
9  #質問票
10 quest <- ("http://www.masaki.j.u-tokyo.ac.jp/utas/2014_2016UTASV_codebook20161004.docx")
11 library(textreadr)
12 file <- download(quest)
13 txt <- read_docx(file)
14 write.csv(txt, "data/quest.txt")
15
16 ##UC Irvine Machine Learning Repository
17 ##UCIが提供するデータを利用する.url = "http://archive.ics.uci.edu/ml"
18 ##利用にあたってUCIからデータをダウンロードしてRに読み込む
19 ##事前にダウンロードしフォルダに格納しておいても解析に問題はない
20
21 ###データがフォルダごと圧縮(.zip)されている場合
22 ##Bank data,Number of Instances:45211,Number of Attributes:17
23 temp <- tempfile()
24 download.file("https://archive.ics.uci.edu/ml/machine-learning-databases/00222/bank.zip", temp)
25 fb <- unzip(temp)
26 fb #フォルダの中身確認
27 dat.full <- read.table(fb[1], header=T, sep=";", #第1引数:列番号
28                      stringsAsFactors = F) #文字列をfactor型にしない,=Tでfactor型に
29 dat.b <- read.table(fb[3], header=T, sep=";",stringsAsFactors = F)
30 str(dat.full)
31 colnames(dat.full)
32 str(dat.b)
33 colnames(dat.b)
34 unlink(temp);rm(temp);unlink(fb);rm(fb)
35 #This dataset is public available for research.
36 #The details are described in [Moro et al., 2014].
37 #[Moro et al., 2014] S. Moro, P. Cortez and P. Rita.
38 #A Data-Driven Approach to Predict the Success of Bank Telemarketing.
39 #Decision Support Systems, Elsevier, 62:22-31, June 2014
40
41 ###WHITE & RED WINE Data
42 #Attribute Information:
43 #For more information, read [Cortez et al., 2009].
44 #Input variables (based on physicochemical tests):
45 #1 - fixed acidity
46 #2 - volatile acidity
47 #3 - citric acid
48 #4 - residual sugar
49 #5 - chlorides
50 #6 - free sulfur dioxide
51 #7 - total sulfur dioxide
52 #8 - density
53 #9 - pH
54 #10 - sulphates
55 #11 - alcohol
56 #Output variable (based on sensory data):
57 #12 - quality (score between 0 and 10)
58 white.wine <- ("https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/winequality-white.csv")
59 dat.white <- read.table(white.wine, header = T, sep = ";")
60 str(dat.white)
61 colnames(dat.white)
62 red.wine <- ("https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/winequality-red.csv")
63 dat.red <- read.table(red.wine, header = T, sep = ";")
64 str(dat.red)
65 colnames(dat.red)
66 #Source:
67 #Paulo Cortez, University of Minho, Guimarães, Portugal,
68 #http://www3.dsi.uminho.pt/pcortez
69 #A. Cerdeira, F. Almeida, T. Matos and J.
70 #Reis, Viticulture Commission of the Vinho Verde Region(CVRVV),
71 #Porto, Portugal @2009
72
73 ###Optical Recognition of Handwritten Digits
74 #V65列は0~9の数字のラベル,
75 #V65列以外は28*28画素(ピクセル)のグレースケールを表す0~255の値
76 #Number of Instances
77 #optdigits.tra:Training 3823
78 #optdigits.tes:Testing 1797
79 #Number of Attributes:64 +1(calss)
80 d.tes <- ("https://archive.ics.uci.edu/ml/machine-learning-databases/optdigits//optdigits.tes")
81 dat.tes <- read.table(d.tes, sep = ",",)
82 d.tra <- ("https://archive.ics.uci.edu/ml/machine-learning-databases/optdigits//optdigits.tra")
83 dat.tra <- read.table(d.tra, sep = ",",)
84 str(dat.tra)
85 colnames(dat.tra)
86 str(dat.tes)
87 colnames(dat.tes)
88 table(dat.tes$V65)
89 table(dat.tra$V65)
90 #Source:
91 #E. Alpaydin, C. Kaynak

```

```
91 #Department of Computer Engineering
92 #Bogazici University, 80815 Istanbul Turkey
93 #alpaydin '@' boun.edu.tr
94
95
```