Lab - 6

1 Reading and writing files

To work with available real data in most cases we need to read the data from a file, typically a '.csv' or '.txt' file. Before doing anything you need to know how the data are provided. Basically you need to know the following:

- Do you have a **header** for the data?
- How the columns are separated in each line? Commonly used delimiter are: semi colon(;), space or a comma(,)
- What character was used to specify the missing data?

I have the following data in a file named 'file.txt':

```
person age sex
Ramesh 22 M
Priyanka 19 F
Vinodh 17 M
Ravi 19 M
Rstudio 12 na
Palak 20 F
```

```
mydata <- read.table(file="file.txt",header=T, sep=" ",na.strings="na")
mydata[,1]

## [1] Ramesh   Priyanka Vinodh   Ravi     Rstudio  Palak
## Levels: Palak Priyanka Ramesh Ravi Rstudio Vinodh
```

Note that I have a header to the file. Data are separated with space and 'na' is used to denote missing data. Reading such file can be done easily as follows: Now let us read a csv file:

```
my.csvtable <- read.csv(file="biostats.csv",sep=",",header=T,stringsAsFactors=F)
my.csvtable

##    Name Sex Age Height Weight
## 1  Alex   M  41     74    170
## 2  Bert   M  42     68    166
## 3  Carl   M  32     70    155
## 4  Dave   M  39     72    167
## 5  Elly   F  30     66    124
## 6  Fran   F  33     66    115
## 7  Gwen   F  26     64    121
## 8  Hank   M  30     71    158
## 9  Ivan   M  53     72    175
## 10 Jake   M  32     69    143
## 11 Kate   F  47     69    139
## 12 Luke   M  34     72    163
```

```
## 13 Myra   F  23     62     98
## 14 Neil   M  36     75    160
## 15 Omar   M  38     70    145
## 16 Page   F  31     67    135
## 17 Quin   M  29     71    176
## 18 Ruth   F  28     65    131

class(my.csvtable)

## [1] "data.frame"

names(my.csvtable)

## [1] "Name"   "Sex"    "Age"    "Height" "Weight"
```

You can also read files from a website:

```
my.url <- "https://people.sc.fsu.edu/~jburkardt/data/csv/deniro.csv"
my.urldata <- read.csv(my.url)

## Error in file(file, "rt"):  cannot open the connection to 'https://people.sc.fsu.edu/
```

Well, that error message is for my system. It should be fine for you. Now let us write a data frame to a file.

```
my.csvtable <- my.csvtable[c(1,3,2,4,5)]
write.csv(x=my.csvtable,file="somenewfile.csv",
sep=",",row.names=F,quote = F,col.names = T)
```

```
Name,Age,Sex,Height,Weight
Alex,41,M,74,170
Bert,42,M,68,166
Carl,32,M,70,155
Dave,39,M,72,167
Elly,30,F,66,124
Fran,33,F,66,115
Gwen,26,F,64,121
Hank,30,M,71,158
Ivan,53,M,72,175
Jake,32,M,69,143
Kate,47,F,69,139
Luke,34,M,72,163
Myra,23,F,62,98
Neil,36,M,75,160
Omar,38,M,70,145
Page,31,F,67,135
Quin,29,M,71,176
Ruth,28,F,65,131
```