Tutorial for AP Lab - Assignment 4

Subject Code: CS 215
Assignment submission date: 06/04/2023

Importing and Exporting data in R

There are some basic functions given in R to read and write data. Some of the most important functions are listed below. The following links are some tutorials for data processing.

1. https://www.geeksforgeeks.org/file-handling-in-r-programming/
2. https://cran.r-project.org/doc/manuals/r-release/R-data.html
3. https://cran.r-project.org/doc/manuals/r-release/R-intro.html#Reading-data-from-files
4. https://www.youtube.com/watch?v=vbVRhr8qexQ

Importing and Exporting data in R

1. load() - load the datasets written with save
2. data(x) - loads specified data sets
3. library(x) - load add-on packages
4. read.table(file) - reads a file in table format and creates a data frame from it

Importing and Exporting data in R

1. read.csv("filename",header=TRUE) id. but with defaults set for reading comma-delimited files
2. read.delim("filename",header=TRUE) id. but with defaults set for reading tab-delimited files
3. read.fwf(file,widths,header=FALSE,sep="",as.is=FALSE) - read a table of fixed width formatted data into a ’data.frame’

Importing and Exporting data in R

1. write.table(x,file="",row.names=TRUE,col.names=TRUE, sep=" ") prints x after converting to a data frame; if quote is TRUE, character or factor columns are surrounded by quotes ("); sep is the field separator; eol is the end-of-line separator; na is the string for missing values; use col.names=NA to add a blank column header to get the column headers aligned correctly for spreadsheet input
2. sink(file) output to file, until sink()

How to read .gz file?

gzfile(description, open = "", encoding = getOption("encoding"),

compression = 6)

Internally gzfile() (see connections) is used to read (write) chunks to (from) the gzip file. If the process is interrupted before completed, the partially written output file is automatically removed.

For gzfile the description is the path to a file compressed by gzip: it can also open for reading uncompressed files and those compressed by bzip2, xz or lzma.

How to read .bz file?

bzfile(description, open = "", encoding = getOption("encoding"),

compression = 9)

Returns the number of (input/compressed) bytes read.

For bzfile the description is the path to a file compressed by bzip2.

How to read URL?

```
url(description, open = "", blocking = TRUE,
    encoding = getOption("encoding"),
    method = getOption("url.method", "default"),
    headers = NULL)
```

For url the description is a complete URL including scheme (such as ‘http://’, ‘https://’, ‘ftp://’ or ‘file://’). Method "internal" is that available since connections were introduced but now mainly defunct.

How to read MS Excel, Binary and ASCII Files?

To import MS Excel file one can use:

```
my_data <- read.table(file = "clipboard",
                      sep = "\t", header=TRUE)
```

or

```
library("readxl")

# xls files

my_data <- read_excel("my_file.xls")

# xlsx files

my_data <- read_excel("my_file.xlsx")
```

How to read MS Excel, Binary and ASCII Files?

To read Binary file one can use:

```
read.filename <- file("/web/com/binmtcars.dat", "rb")
```

```
Other examples:
```

```
column.names <- readBin(read.filename, character(),  n = 3)
```

```
read.filename <- file("/web/com/binmtcars.dat", "rb")
```

```
bindata <- readBin(read.filename, integer(),  n = 18)
```

How to read MS Excel, Binary and ASCII Files?

The main functions used in R to import data from ASCII files are read.table and read.csv to read data in a tabular form, and readLines to read lines from a text file. The only difference between read.table and read.csv is that in the later the default separator is a comma.

```
data <- read.csv("input.csv")
```

Further readings

1. https://www.tutorialspoint.com/r/r_csv_files.htm
2. https://cran.r-project.org/doc/contrib/Short-refcard.pdf
3. https://stat.ethz.ch/R-manual/R-devel/library/base/html/connections.html
4. http://venus.ifca.unican.es/Rintro/dataReading.html#:~:text=many%20different%20formats!-,ASCII%20data%20files,The%20only%20difference%20between%20read.
5. https://cran.r-project.org/doc/manuals/r-release/R-data.html
6. https://stat.ethz.ch/R-manual/R-devel/library/base/html/write.html

Thank you!