R Language Functions for Importing and Exporting Data

John Aa. Sørensen, lektor
Section for Information Technology, DTU Diplom

16 February 2017

R Language Functions for Importing and Exporting Data.

```
> # Filename: 3_R_Data_Import_Export.R
> #
> # Objective:
> # R Language functions for importing and exporting data.
> #
> #                    TABLE OF CONTENTS
> #
> # 3.0: Enter the main R Language documentation.
> # 3.1: Setting working directory.
> # 3.2: Importing from comma separated values (csv) text file.
> # 3.3: Importing from a CSV file exported from Excel and merge.
> # 3.4: Writing dataframe to csv file in working directory.
>
```

R Language Functions for Importing and Exporting Data.

```
> #                 REFERENCES
> #
> # [Adler, 2012] Joseph Adler, "R In a Nutshell", O'Reilly, 2012.
> # [Kabacoff, 2015]  Robert I. Kabacoff, "R in Action",
> #    2'Ed, Manning Publications, 2015.
> # [R Core, 2015] R Core Team and contributors worldwide,
> #     "The R Language Manual System", from help.start().
> # [R Core, 2015a] R Core Team, "R Data import/Export"
> #     cran.r-project.org/doc/manuals/r-release/R-data.html
> # [short, 2004] Tom Short, "Short Reference Card"
> #     cran.r-project.org/doc/contrib/Short-refcard.pdf
> # [Teetor, 2011] Paul Teetor, "R Cookbook", O'Reilley, 2011.
> # [Torfs, 2014] Paul Torfs, Caludia Brauer,
> #    "A (very) Short Introduction to R",
> #  cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf
> # [Zhao, 2013] Yanchang Zhao, "R and Data Mining", Elsevier, 2013.
> # [Zhao, 2013] Yanchang Zhao, "R Reference Card for Data Mining",
> #     www.rdatamining.com/docs/r-reference-card-for-data-mining
> #
```

R Language Functions for Importing and Exporting Data.

```
> #
> ##################################################
> # 3.0: Enter the main R Language documentation.
> help.start()   # The main entry to R project documentation.
If nothing happens, you should open
http://127.0.0.1:10115/doc/html/index.html yourself
> #
>
```

R Language Functions for Importing and Exporting Data.

```
> ##########################################################
> # 3.1: Setting the working directory.
> #
> # Manually setting working directory when opening RStudio:
> # RStudio -> Session -> Set Working directory -> Select dir.
> #
> (dir <- getwd())   # Get and display working directory.
[1] "C:/W_2016_05_21/Part_3_2_ITD_kurser_v35/
BigData_Course_IDA_2016.06.16-17/R_code"
> #
> # Generate a link to file: myfilname.csv by concatenating
> # with the working directory in variable dir .
> # csv file prod_eval_1.cvs is in the working directory.
> file1 <- paste(as.character(dir),"/prod_eval_1.csv", sep="",
collapse = NULL)
> file1   # Check the path to the prod_eval_1.csv file.
[1] "C:/W_2016_05_21/Part_3_2_ITD_kurser_v35/
BigData_Course_IDA_2016.06.16-17/R_code/prod_eval_1.csv"
> #
```

R Language Functions for Importing and Exporting Data.

```
> ##########################################################
> # 3.2: Import from comma separated values (CSV) text file.
> #
> # Main refs:
> # [Kabacoff, 2015] page page 34,
> #  Section 2.3.2 "Importing data from a delimited text file"
> #
> # [R Core, 2015a] As "THE general reference".
> #
> # Read CSV data from example text file with product evaluations
> #  from users.
> # The structure is as shown in Section 2.6.1 "Create a data frame".
> # The file is organized as 5 columns comma separated observations
> #  as described below. The data columns are preceeded by one headline
> #  with names on the variables and this headline is again preceeded
> #  with 4 lines of comments.
> # This format is further described below:
> #
> # The example file prod_eval_1.csv should be included in the
> #  workspace before continuing.
>
```

R Language Functions for Importing and Exporting Data.

```
> # File format:
> # The first 4 lines of the csv file are comments.
> # Then one headline with five  names of variables observed.
> # They are further described below:
> #
> # Prod_no: Product number. Value 1 or 2.
> # Fe_Ma: The answer is from a female/male. Value F or M.
> #   S1: The product is useful.
> #   S2: The product price is acceptable.
> #   S3: The customer will, without hesitation, recommend
> #       the product to a person known well by the customer.
> #
> # The observation values in S1, S2 and S3 are Likert scaled:
> #  Value 5: Strongly agree
> #  Value 4: Agree
> #  Value 3: Neutral (Do not agree or disagree)
> #  Value 2: Disagree
> #  Value 1: Strongly disagree.
>
```

R Language Functions for Importing and Exporting Data.

```
> ?read.table()       # The R manual for the function.
> #
> product_eval_1 <- read.table(file="prod_eval_1.csv",
+ header=TRUE,
+   colClasses =c("numeric", "character", "numeric", "numeric",
+   "numeric"),
+   skip=4, sep="," )
> str(product_eval_1)    # Display structure.
'data.frame': 24 obs. of  5 variables:
 $ Prod_no: num  2 1 2 2 1 2 2 1 2 2 ...
 $ Fe_Ma  : chr  " F" " M" " F" " F" ...
 $ S1     : num  3 5 3 3 5 3 3 5 3 3 ...
 $ S2     : num  4 4 4 4 4 4 4 4 4 4 ...
 $ S3     : num  5 2 5 5 2 5 5 2 5 5 ...
> #
```

R Language Functions for Importing and Exporting Data.

```
> #
> # The skip=4, skips top 4 lines of csv file when reading,
> # thus allowing these lines to be used for comments only.
> head(product_eval_1, n=10)  # Display 10 lines from the top.
   Prod_no Fe_Ma S1 S2 S3
1        2     F  3  4  5
2        1     M  5  4  2
3        2     F  3  4  5
4        2     F  3  4  5
5        1     M  5  4  2
6        2     F  3  4  5
7        2     F  3  4  5
8        1     M  5  4  2
9        2     F  3  4  5
10       2     F  3  4  5
>
```

R Language Functions for Importing and Exporting Data.

```
> ##########################################################
> # 3.2: Read additional CSV file, gen. from Excel in CSV format.
> #
> product_eval_2 <- read.table(file="prod_eval_2.csv",
+ header=TRUE,
+       colClasses =c("numeric", "character", "numeric", "numeric",
+ "numeric"), skip=4, sep=";" )
> str(product_eval_2)  # Display structure.
'data.frame': 1 obs. of  5 variables:
 $ Prod_no: num 2
 $ Fe_Ma  : chr "F"
 $ S1     : num 3
 $ S2     : num 4
 $ S3     : num 5
```

R Language Functions for Importing and Exporting Data.

```
> # sep=";", is separator in csv file gen. from Excel export
> # to a csv file.
> product_eval_2   # Display all entries.
  Prod_no Fe_Ma S1 S2 S3
1       2     F  3  4  5
> #
> # Merging two data frames.
> #
> product_eval_all <- rbind(product_eval_1, product_eval_2)
> tail(product_eval_all, n=3)  # Check new observation in bottom.
   Prod_no Fe_Ma S1 S2 S3
23       1     M  5  4  2
24       2     F  3  4  5
25       2     F  3  4  5
> str(product_eval_all) # Check 1 more observation (row).
'data.frame': 25 obs. of  5 variables:
 $ Prod_no: num  2 1 2 2 1 2 2 1 2 2 ...
 $ Fe_Ma  : chr  " F" " M" " F" " F" ...
 $ S1     : num  3 5 3 3 5 3 3 5 3 3 ...
 $ S2     : num  4 4 4 4 4 4 4 4 4 4 ...
 $ S3     : num  5 2 5 5 2 5 5 2 5 5 ...
```

R Language Functions for Importing and Exporting Data.

```
> ##############################################################
> # 3.4: Writing a data frame to csv file in working directory.
> #
> write.csv(product_eval_all, file = "product_eval_all.csv")
> #
>
```

References I

Joseph Adler (2012)

R in a Nutshell

OReilly

Robert I. Kabacoff (2015)

R in Action

Manning Publications 2'Ed.

R Core Team and contributors worldwide (2015)

The R Language Manual System

CRAN e.g. via RStudio

Tom Short, (2004)

Short Reference Card

CRAN cran.r-project.org/doc/contrib/Short-refcard.pdf

Paul Teetor

R Cookbook

O'Reilley

References II

Paul Torfs, Caludia Brauer

A (very) Short Introduction to R.

CRAN cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf

Yanchang Zhao

R and Data Mining.

Elsevier 2013.

Yanchang Zhao

R Reference Card for Data Mining.

www.rdatamining.com

www.rdatamining.com/docs/r-reference-card-for-data-mining.pdf