Assignment 2\_1

1. How to Import SAS XPORT files into R with the foreign package?

read.xport (foreign)

read.xport(file)

1. How to Import SAS Files into R with the Haven package?

install.packages("haven")

library(haven)

dat = read\_sas("path to file", "path to formats catalog")

Eg: path <- [system.file](https://rdrr.io/r/base/system.file.html)("examples", "iris.sas7bdat", package = "haven")

[read\_sas](https://rdrr.io/cran/haven/man/read_sas.html)(path)

1. How to read Weka Attribute-Relation File Format (ARFF) files in R?

Library(foreign)

read.arff(file)

Eg: [read.arff](https://rdrr.io/cran/RWeka/man/read.arff.html)([system.file](https://rdrr.io/r/base/system.file.html)("arff", "contact-lenses.arff",

package = "RWeka"))

1. How to read a heavy csv/tsv file using readr package?

# Read a txt file, named "mtcars.txt"

my\_data <- read\_tsv("mtcars.txt")

# Read a csv file, named "mtcars.csv"

my\_data <- read\_csv("mtcars.csv")

Assignment 2\_2

1. Define matrix mymat by replicating the sequence 1:5 for 4 times and transforming into a matrix, sum over rows and columns.

b<-matrix(rep(1:5), nrow=5, ncol=4);b

colSums(b)

rowSums(b)

Assignment 2\_3

df1 = data.frame(CustId = c(1:6), Product = c(rep("TV", 3), rep("Radio", 3)))

df2 = data.frame(CustId = c(2, 4, 6), State = c(rep("Texas", 2), rep("NYC", 1)))

df1 #left table

df2 #right table

For the above given data frames and tables perform the following operations:

1. Return only the rows in which the left table have match

2. Return all rows from both tables, join records from the left which have matching keys in the right table.

merge(x=df1\_1, y=df2\_1, by = "CustId",all.x=TRUE)

merge(x=df1\_1, y=df2\_1, by = "CustId", all=TRUE)

Assignment 2\_4

1. x <- c(‘data.science.in.R’, ‘machine.learning.in.R’)

Perform the below string Operation:

• Replace the period character "." within each string with another character i.e. "-" minus sign.

Gsub(‘\\.’, ‘-‘, x) or gsub(‘[.]’, ‘-‘, x)

2. x <- c('data.science.in.R','machine.learning.in.R')

Perform the below String Operation:

• Append again with “-“ minus sign character at the start of the each string and finally concatenate all the string within the vector to form a final single string and assigning it the object.

#Replaces .with minus

gsub('[.]','-',x)

gsub('\\.','-',x)

# paste two strings in a vector

x\_1<-paste('-',x)

paste(x\_1[1],x\_1[2]) #without space

paste0(x\_1[1],x\_1[2]) #with space

Assignment 2\_5

1. States = rownames(US Arrests)

Get states names with ‘w’.

States[grep("^w",rownames(USArrests))]

Get states names with ‘W’.

States[grep("^W",rownames(USArrests))]

1. Prepare a Histogram of the number of characters in each US state.

Hist(States$rowames)

hist(nchar(States), main = "Histogram",

xlab = "number of characters in US State names")