Data Types in R

2024-01-09

#Data Types  
#Vectors  
a = c(1,4,7,5,4.0,7) #numerical vector  
a

## [1] 1 4 7 5 4 7

b = c("What's", "up", "dude?") #character vector  
b

## [1] "What's" "up" "dude?"

c = c(T,T,F,F,T) #logical vector  
c

## [1] TRUE TRUE FALSE FALSE TRUE

# To generate a numeric matric  
y = matrix(31:50, nrow = 5, ncol = 4, byrow = T)  
y

## [,1] [,2] [,3] [,4]  
## [1,] 31 32 33 34  
## [2,] 35 36 37 38  
## [3,] 39 40 41 42  
## [4,] 43 44 45 46  
## [5,] 47 48 49 50

#Another eg.  
cells = c(1,44,55,33)  
rnames = c("R1", "R2")  
cnames = c("C1", "C2")  
  
mymat = matrix(cells, nrow = 2, ncol = 2, byrow = T, dimnames = list(rnames, cnames))  
mymat

## C1 C2  
## R1 1 44  
## R2 55 33

mymat[1,2]

## [1] 44

mymat[2,]

## C1 C2   
## 55 33

mymat[,1]

## R1 R2   
## 1 55

#Data Frames  
d = c(1,2,3,4)  
e = c("red","yellow","steelblue",NA)  
f = c(T,F,T,F)  
mydata = data.frame(d,e,f)  
mydata

## d e f  
## 1 1 red TRUE  
## 2 2 yellow FALSE  
## 3 3 steelblue TRUE  
## 4 4 <NA> FALSE

names(mydata) = c("Id","Color", "Passes")  
mydata

## Id Color Passes  
## 1 1 red TRUE  
## 2 2 yellow FALSE  
## 3 3 steelblue TRUE  
## 4 4 <NA> FALSE

mydata$Color

## [1] "red" "yellow" "steelblue" NA

#lists (collection varied objects)  
w = list(name = "freddy", mynum = a, mymat = y, age = 21.61)  
w

## $name  
## [1] "freddy"  
##   
## $mynum  
## [1] 1 4 7 5 4 7  
##   
## $mymat  
## [,1] [,2] [,3] [,4]  
## [1,] 31 32 33 34  
## [2,] 35 36 37 38  
## [3,] 39 40 41 42  
## [4,] 43 44 45 46  
## [5,] 47 48 49 50  
##   
## $age  
## [1] 21.61

#Factor  
# variable gender with 20 "males" entries and 30 "female" entires  
gender = c(rep("male", 20), rep("female", 30))  
#labelling is done alphabatically  
gender

## [1] "male" "male" "male" "male" "male" "male" "male" "male"   
## [9] "male" "male" "male" "male" "male" "male" "male" "male"   
## [17] "male" "male" "male" "male" "female" "female" "female" "female"  
## [25] "female" "female" "female" "female" "female" "female" "female" "female"  
## [33] "female" "female" "female" "female" "female" "female" "female" "female"  
## [41] "female" "female" "female" "female" "female" "female" "female" "female"  
## [49] "female" "female"

gender = factor(gender)  
gender

## [1] male male male male male male male male male male   
## [11] male male male male male male male male male male   
## [21] female female female female female female female female female female  
## [31] female female female female female female female female female female  
## [41] female female female female female female female female female female  
## Levels: female male

summary(gender)

## female male   
## 30 20

#importing excel file  
library(readxl)

## Warning: package 'readxl' was built under R version 4.3.2

dir()

## [1] "09-01-2024.docx" "09-01-2024.R" "09-01-2024.Rmd" "23-01-2024(ML).R"  
## [5] "mydata.csv" "mydata.txt" "sampfile.xlsx"

data1 = read\_excel("sampfile.xlsx")  
data1

## # A tibble: 4 × 3  
## Name `Reg. No.` Marks  
## <chr> <chr> <dbl>  
## 1 Ross 23MDT1007 95  
## 2 Joey 23MDT1018 89  
## 3 Monica 23MDT1018 99  
## 4 Chandler 23MDT1012 91

#create a data frame from scratch and export  
age = c(25,30,56)  
gender = c("M","F","F")  
weight = c(160,110,120)  
mydata1 = data.frame(age,gender,weight)  
mydata1

## age gender weight  
## 1 25 M 160  
## 2 30 F 110  
## 3 56 F 120

write.table(mydata1, "mydata.txt", sep = "\t")  
write.csv(mydata1, "mydata.csv")  
dir()

## [1] "09-01-2024.docx" "09-01-2024.R" "09-01-2024.Rmd" "23-01-2024(ML).R"  
#­# [5] "mydata.csv" "mydata.txt" "sampfile.xlsx"