PractExam

Natalie Joy Loredo

2024-03-07

```
#A. Load the built-in warpbreaks dataset.

data("warpbreaks")
warpbreaks
```

```
##    breaks wool tension
## 1      26    A       L
## 2      30    A       L
## 3      54    A       L
## 4      25    A       L
## 5      70    A       L
## 6      52    A       L
## 7      51    A       L
## 8      26    A       L
## 9      67    A       L
## 10     18    A       M
## 11     21    A       M
## 12     29    A       M
## 13     17    A       M
## 14     12    A       M
## 15     18    A       M
## 16     35    A       M
## 17     30    A       M
## 18     36    A       M
## 19     36    A       H
## 20     21    A       H
## 21     24    A       H
## 22     18    A       H
## 23     10    A       H
## 24     43    A       H
## 25     28    A       H
## 26     15    A       H
## 27     26    A       H
## 28     27    B       L
## 29     14    B       L
## 30     29    B       L
## 31     19    B       L
## 32     29    B       L
## 33     31    B       L
## 34     41    B       L
## 35     20    B       L
## 36     44    B       L
## 37     42    B       M
```

```
## 38     26    B       M
## 39     19    B       M
## 40     16    B       M
## 41     39    B       M
## 42     28    B       M
## 43     21    B       M
## 44     39    B       M
## 45     29    B       M
## 46     20    B       H
## 47     21    B       H
## 48     24    B       H
## 49     17    B       H
## 50     13    B       H
## 51     15    B       H
## 52     15    B       H
## 53     16    B       H
## 54     28    B       H
```

```
#1.
str(warpbreaks)
```

```
## 'data.frame':    54 obs. of  3 variables:
##  $ breaks : num  26 30 54 25 70 52 51 26 67 18 ...
##  $ wool   : Factor w/ 2 levels "A","B": 1 1 1 1 1 1 1 1 1 1 ...
##  $ tension: Factor w/ 3 levels "L","M","H": 1 1 1 1 1 1 1 1 1 2 ...
```

```
#A2. How many observations does it have?

#Answer: 54 obsrvations
```

```
#A3.
typeof(warpbreaks$breaks)
```

```
## [1] "double"
```

```
typeof(warpbreaks$wool)
```

```
## [1] "integer"
```

```
typeof(warpbreaks$tension)
```

```
## [1] "integer"
```

```
#B1. Read the complete file using readLines.

file <- file("exampleFile.txt")

readFile <- readLines(file)
readFile
```

```
## [1] "// Survey data. Created : 21 May 2013"
## [2] "// Field 1: Gender"
## [3] "// Field 2: Age (in years)"
## [4] "// Field 3: Weight (in kg)"
## [5] "M;28;81.3"
## [6] "male;45;"
## [7] "Female;17;57,2"
## [8] "fem.;64;62.8"
```

```r
#B2. Use grepl

Comments <- readFile[grepl("^//", readFile)]
Comments
```

```
## [1] "// Survey data. Created : 21 May 2013"
## [2] "// Field 1: Gender"
## [3] "// Field 2: Age (in years)"
## [4] "// Field 3: Weight (in kg)"
```

```r
datavec <- readFile[!grepl("^//", readFile)]
datavec
```

```
## [1] "M;28;81.3"      "male;45;"       "Female;17;57,2" "fem.;64;62.8"
```

```r
#B3. Extract the date from the first comment line and display on the screen "It was created data."

subCom <- Comments[1]
date <- gsub ("// Survey data. Created : ", "", subCom)
date
```

```
## [1] "21 May 2013"
```

```r
cat("It was created, ", date)
```

```
## It was created,  21 May 2013
```

```r
#4. Read the data into a matrix
#B4a.
vecsplit <- (strsplit(datavec, ";"))
vecsplit
```

```
## [[1]]
## [1] "M"    "28"   "81.3"
##
## [[2]]
## [1] "male" "45"
##
## [[3]]
## [1] "Female" "17"     "57,2"
##
## [[4]]
## [1] "fem." "64"   "62.8"
```

```r
#4b.
maxvec <- max(lengths(vecsplit))
maxvec
```

```
## [1] 3
```

```r
row_append <- lapply(vecsplit, function(x) c(x, rep(NA, maxvec - length(x))))
row_append
```

```
## [[1]]
## [1] "M"    "28"   "81.3"
##
## [[2]]
## [1] "male" "45"   NA
##
```

```
## [[3]]
## [1] "Female" "17"     "57,2"
##
## [[4]]
## [1] "fem." "64"   "62.8"
```

```
#B4c

unlist_data <- unlist(row_append)
unlist_data
```

```
##  [1] "M"      "28"     "81.3"   "male"   "45"     NA       "Female" "17"
##  [9] "57,2"   "fem."   "64"     "62.8"
```

```
data_matrix <- matrix(unlist_data, ncol = 4, nrow = 3,
                      dimnames = list(c("row1", "row2", "row3")))
data_matrix
```

```
##      [,1]   [,2]   [,3]     [,4]
## row1 "M"    "male" "Female" "fem."
## row2 "28"   "45"   "17"     "64"
## row3 "81.3" NA     "57,2"   "62.8"
```

```
#B4d.

field_names <- Comments[2:4]
extra_field_name <- gsub("//", "", field_names)
rownames(data_matrix) <- extra_field_name
data_matrix
```

```
##                          [,1]   [,2]   [,3]     [,4]
##  Field 1: Gender         "M"    "male" "Female" "fem."
##  Field 2: Age (in years) "28"   "45"   "17"     "64"
##  Field 3: Weight (in kg) "81.3" NA     "57,2"   "62.8"
```