PracticalExam

Erikka Jane Parrenas

2024-03-07

```
#A. Load the built-in warpbreaks dataset.
#1a. Find out, in a single command, which columns of warpbreaks are either numeric or integer. What are
  summary(warpbreaks)
```

```
##      breaks      wool   tension
##  Min.   :10.00   A:27   L:18
##  1st Qu.:18.25   B:27   M:18
##  Median :26.00          H:18
##  Mean   :28.15
##  3rd Qu.:34.00
##  Max.   :70.00
```

```
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
```

```
#1a. Find out, in a single command, which columns of warpbreaks are either numeric or integer. What are
typeof(warpbreaks$breaks)
```

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

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

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

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

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

```
# column wool and tension are integers.
# while column breaks are double
```

#2a. How many observations does it have? #it has 54 observations

#3a. Is numeric a natural data type for the columns which are stored as such? Convert to integer when necessary. #yes

```
#B. Load the exampleFile.txt
#1b. Read the complete file using readLines.

file <- file("exampleFile.txt")

  readExamFile <- readLines(file)
  readExamFile
```

```
## [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"
```

```
#2b. Separate the vector of lines into a vector containing comments and a vector containing the data. H
comment <- readExamFile[grepl("^//", readExamFile)]
comment
```

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

```
vec <- readExamFile[grepl("^//", readExamFile)]
vec
```

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

```
#3b
subsetCom <- comment[1]

date <- gsub("// Survey data. Created :", "", subsetCom)
date
```

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

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

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

```
#B4a
vec_Split <- (strsplit(vec, ";"))
vec_Split
```

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

```
#B4b
max_Vec <- max(length(vec_Split))
max_Vec
```

```
## [1] 4
```

```
appendRow <- lapply(vec_Split, function(x) c(x,rep(NA, max_Vec - length(x))))
appendRow
```

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

```
#B4c
unlisDat <- unlist(appendRow)
unlisDat
```

```
##  [1] "// Survey data. Created : 21 May 2013"
##  [2] NA
##  [3] NA
##  [4] NA
##  [5] "// Field 1: Gender"
##  [6] NA
##  [7] NA
##  [8] NA
##  [9] "// Field 2: Age (in years)"
## [10] NA
## [11] NA
## [12] NA
## [13] "// Field 3: Weight (in kg)"
## [14] NA
## [15] NA
## [16] NA
```

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

```
## Warning in matrix(unlisDat, ncol = 4, nrow = 3, dimnames = list(c("row1", :
## data length [16] is not a sub-multiple or multiple of the number of rows [3]
```

```
dat_matrix
```

```
##      [,1]                                    [,2]
## row1 "// Survey data. Created : 21 May 2013" NA
## row2 NA                                      "// Field 1: Gender"
## row3 NA                                      NA
##      [,3]                          [,4]
## row1 NA                            NA
## row2 NA                            NA
## row3 "// Field 2: Age (in years)" NA
```