

CS_102_PracticalExam

Charles Daniel Songaling

2024-03-07

```
library("warpbreaks")
```

```
data("warpbreaks")
```

A. #1. Find out, in a single command, which columns of warpbreaks are either numeric or integer. What are the data types of each column?

```
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 ...
```

#The break column is numeric.

#The wool and tension column is factor.

#2. How many observations does it have?

```
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 ...
```

#It has 54 observations.

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

```
warpbreaks$wool <- as.integer(warpbreaks)
```

```
## Error in eval(expr, envir, enclos): 'list' object cannot be coerced to type 'integer'
```

#4. Error messages in R sometimes report the underlying type of an object rather than the user-level class. Derive from the following code and error message what the underlying type. Explain what is the error all about. Do not just copy the error message that was displayed.

B. #1. Read the complete file using readLines.

```
readLines("exampleFile.txt")
```

```
## Warning in readLines("exampleFile.txt"): incomplete final line found on
## 'exampleFile.txt'
```

```
## [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"
```

#2. Separate the vector of lines into a vector containing comments and a vector containing the data. Hint: use `grepl`.

```
lines <- readLines("exampleFile.txt")
```

```
## Warning in readLines("exampleFile.txt"): incomplete final line found on
## 'exampleFile.txt'
```

```
comments <- lines[grepl("//",lines)]
data <- lines[grepl(";",lines)]
```

```
comments
```

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

```
data
```

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

#3. Extract the date from the first comment line and display on the screen “It was created data.”

```
txt <- readLines(paste("exampleFile.txt", sep = ""), n = 1)
txt
```

```
## [1] "// Survey data. Created : 21 May 2013"
```

#4. Read the data into a matrix as follows. a. Split the character vectors in the vector containing data lines by semicolon (;) using `strsplit`.

- b. Find the maximum number of fields retrieved by split. Append rows that are shorter with NA's.
- c. Use `unlist` and `matrix` to transform the data to row-column format.
- d. From comment lines 2-4, extract the names of the fields. Set these as `colnames` for the matrix you just created.