

R_manipulation

Gabriele Degola

October, 2021

```
library(dplyr)
```

Kindergarten

Data loading

Data are copy-pasted from the PDF and saved in R variables.

```
group1 <- "score in petite section 1,36 0,24 0,22 0,11 1,25 0,24 0,35 0,24 1,60 0,63
score in moyenne section 2,07 3,60 2,57 4,29 5,67 1,67 2,43 4,72 7,67 3,83
score in grande section 7,63 5,79 7,38 8,50 5,75 4,83 5,38 7,75 7,54 8,38
0,68 0,00 1,44 1,86 3,47 0,58 0,00 0,82 2,75 3,14 1,58 2,83 1,92 0,83
5,13 2,31 1,00 3,74 6,64 3,39 2,36 6,75 5,13 8,00 7,08 6,01 6,25 2,61
7,50 6,17 5,68 7,25 7,54 5,64 6,46 7,50 8,75 9,00 7,25 7,42 6,58 8,25
2,47 1,13 1,72 0,61 0,35 1,11 1,24 2,36 0,50
3,61 2,67 5,24 1,56 1,89 4,06 3,11 4,00 2,89
7,46 7,21 7,50 8,17 6,96 8,00 6,92 5,71 7,00"
group2 <- "score in petite section 1,57 1,36 1,25 2,81 0,86 0,47 0,35 1,81 1,63 1,36
score in moyenne section 2,26 3,68 6,26 3,88 6,54 5,25 6,00 6,67 4,96 2,57
score in grande section 4,42 6,63 6,88 7,17 7,00 7,33 7,63 6,54 8,54 5,21
0,35 1,75 0,35 2,58 1,58 1,67 0,60 1,63 1,06 0,00 2,17 0,25 0,00 2,67
1,64 5,83 6,42 6,88 4,46 1,94 5,50 1,71 3,33 2,28 5,67 5,32 1,56 2,32
4,46 7,00 6,50 6,00 7,25 5,47 7,63 3,83 4,92 5,96 6,33 7,13 4,61 4,13
2,33 1,13 1,22 1,33 1,00 0,83
4,17 6,18 4,32 6,50 3,00 2,67
7,00 7,25 7,88 8,50 5,21 7,00"
groups <- c(group1, group2)
```

Data manipulation

```
data <- data.frame()
for (i in 1:length(groups)) {
  group <- groups[i]
  # Split data according to newlines
  group <- as.data.frame(strsplit(group, "\n"))
  # Replace ',' with '.' (decimal numbers) and remove "titles"
  group <- group %>% mutate_all(~ gsub(",", ".", .))
  group <- group %>% mutate_all(~ gsub("[A-Za-z]", "", .))
  # Put together scores of each section
  petite <- paste(group[1,], group[4,], group[7,])
  moyenne <- paste(group[2,], group[5,], group[8,])
  grande <- paste(group[3,], group[6,], group[9,])
}
```

```

# Split values
petite <- as.numeric(unlist(strsplit(petite, " ")))
petite <- petite[!is.na(petite)]
moyenne <- as.numeric(unlist(strsplit(moyenne, " ")))
moyenne <- moyenne[!is.na(moyenne)]
grande <- as.numeric(unlist(strsplit(grande, " ")))
grande <- grande[!is.na(grande)]
print(petite)
print(moyenne)
print(grande)
# Convert to dataframe
# data <- rbind(data, setNames(data.frame(petite), names(data)))
# data <- rbind(data, setNames(data.frame(moyenne), names(data)),
  ↪ setNames(data.frame(moyenne), names(data)))
}

## [1] 1.36 0.24 0.22 0.11 1.25 0.24 0.35 0.24 1.60 0.63 0.68 0.00 1.44 1.86 3.47
## [16] 0.58 0.00 0.82 2.75 3.14 1.58 2.83 1.92 0.83 2.47 1.13 1.72 0.61 0.35 1.11
## [31] 1.24 2.36 0.50
## [1] 2.07 3.60 2.57 4.29 5.67 1.67 2.43 4.72 7.67 3.83 5.13 2.31 1.00 3.74 6.64
## [16] 3.39 2.36 6.75 5.13 8.00 7.08 6.01 6.25 2.61 3.61 2.67 5.24 1.56 1.89 4.06
## [31] 3.11 4.00 2.89
## [1] 7.63 5.79 7.38 8.50 5.75 4.83 5.38 7.75 7.54 8.38 7.50 6.17 5.68 7.25 7.54
## [16] 5.64 6.46 7.50 8.75 9.00 7.25 7.42 6.58 8.25 7.46 7.21 7.50 8.17 6.96 8.00
## [31] 6.92 5.71 7.00
## [1] 1.57 1.36 1.25 2.81 0.86 0.47 0.35 1.81 1.63 1.36 0.35 1.75 0.35 2.58 1.58
## [16] 1.67 0.60 1.63 1.06 0.00 2.17 0.25 0.00 2.67 2.33 1.13 1.22 1.33 1.00 0.83
## [1] 2.26 3.68 6.26 3.88 6.54 5.25 6.00 6.67 4.96 2.57 1.64 5.83 6.42 6.88 4.46
## [16] 1.94 5.50 1.71 3.33 2.28 5.67 5.32 1.56 2.32 4.17 6.18 4.32 6.50 3.00 2.67
## [1] 4.42 6.63 6.88 7.17 7.00 7.33 7.63 6.54 8.54 5.21 4.46 7.00 6.50 6.00 7.25
## [16] 5.47 7.63 3.83 4.92 5.96 6.33 7.13 4.61 4.13 7.00 7.25 7.88 8.50 5.21 7.00

```