# От вектора к текстам в R

Konstantin Ivanin

24 сентября 2017 г

## Работа со строками

Загрузим библиотеки

```
library(stringr)
library(tidyverse)
library(tidyr)
library(readr)
library(stringi)
library(stringr)
```

#### Чет-нечет

Напишите функцию is.odd(), которая возвращает значение TRUE, если число символов в строке нечетно, FALSE, если число символов в строке четно.

```
is.odd <- function(str) {
   return(nchar(str) %% 2 != 0)
}
is.odd(c("odd", "even", ""))</pre>
```

## [1] TRUE FALSE FALSE

#### Искусственные данные по средней продолжительности сна

В данных по продолжительности сна (1.1\_sleep\_hours.csv) две переменных: код испытуемого и среднее время сна. Попробуйте сделать следующий график:

```
sleep.data <- read_csv('data/1.1_sleep_hours.csv')
sleep.data %>%
  separate(code, c('sex', 'year'), sep = '__') %>%
  ggplot(aes(sleep_hours, fill = sex)) +
  geom_density(alpha = 1/3) +
  ggtitle(' ') +
  xlab(' ') +
  ylab(' ')
```

```
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
                     0x7f
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
            0xe1
## Warning in grid.Call(L stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L stringMetric, as.graphicsAnnot(x$label)):
           0xae
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
            0xed
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L stringMetric, as.graphicsAnnot(x$label)):
            0xa4
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L stringMetric, as.graphicsAnnot(x$label)):
           0x9e
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
                     0x7f
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
                     0x80
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
                     0x7f
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
            0xb7
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
           0xd2
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
            0xe1
```

```
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
            8bx0
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
            0xdc
## Warning in grid.Call(L stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L stringMetric, as.graphicsAnnot(x$label)):
            0xfe
##
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)):
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xdf
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xfe
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0x9e
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xd0
```

```
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xd8
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0x9e
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0x7f
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xdc
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0x9e
```

```
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0x7f
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xed
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xdf
```

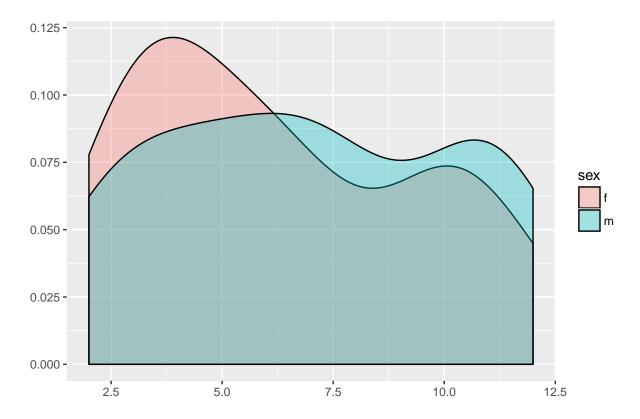
```
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xd0
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xde
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0x2
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xde
```

```
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0x9e
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0x7f
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  9xd8
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  9xd8
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xe1
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xe1
```

```
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xe1
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xae
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
                        0xae
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xdf
```

```
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0x9e
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
                        0xdf
## $y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
                        0x9e
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
                        0xfe
## $y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0xb7
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0x7f
```

```
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xd0
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xd8
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
                  0xde
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0x2
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
##
                  0x7f
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
                        0xb7
## $y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
## Warning in grid.Call.graphics(L text, as.graphicsAnnot(x$label), x$x, x
                        0xd8
## $y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, :
                        0xde
```



## Алфавитный порядок

Как было сказано выше, "алфавитный" порядок разный в разных локалях. Напишите функцию ordering\_in\_locales, которая на вход принимает два символа, а возвращает таблицу с информацией о том, в скольких локалях порядок соответствует поданному в функцию, а в скольких порядок обратный:

```
ordering_in_locales <- function(ch1, ch2) {
   str <- c(ch1, ch2)
   loc <- stri_locale_list()
   return(table(sapply(loc, function(x) {paste(stringr::str_sort(str, locale = x), collapse = '_')})))
}
ordering_in_locales('i', '')</pre>
```

```
##
## _i i_
## 29 668
```

#### Функция для проверки полиндромности

Напишите функцию is.palindrome, которая будет проверять, является ли слово полиндромом.

```
is.palindrome <- function(str) {
    #return(str == stri_reverse(str))
    return(str == paste(rev(str_split(str, pattern = '', simplify = T)), collapse = ''))
}
is.palindrome(' ')
## [1] TRUE
is.palindrome(' ')</pre>
## [1] FALSE
```

### Функция для зеркального отражения размера букв

Напишите функцию  $mirror_case()$ , которая в строке все большие буквы заменяет на маленькие, а все маленькие – на большие.

```
mirror_case <- function(str) {
   str <- str_split(str, pattern = '', simplify = T)
   up <- grepl('[[:upper:]]', str)
   low <- grepl('[[:lower:]]', str)
   str[up] <- str_to_lower(str[up])
   str[low] <- str_to_upper(str[low])
   return(paste(str, collapse = ''))
}
mirror_case(' ')</pre>
```

## [1] " "