Домашнее задание

# Задачи

*Операции в задачах 2–10 должны быть выполнены с помощью функций из системы библиотек tidyverse.*

## Задача 1

Загрузите данные из файла beasts.csv в R в датафрейм beasts. При необходимости выберите подходящий разделитель столбцов и десятичный разделитель. Проверьте, что все столбцы имеют корректные типы, запросив структуру датафрейма.

*Ответ:*

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --  
## v ggplot2 3.3.6 v purrr 0.3.4   
## v tibble 3.1.8 v dplyr 1.0.10  
## v tidyr 1.2.1 v stringr 1.4.1   
## v readr 2.1.2 v forcats 0.5.2   
## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

setwd("C:/Users/alexandra\_c/Desktop/Уроки R/Практические задания/06-10 Практическая работа")  
  
beasts <- read.csv("beasts.csv")  
str(beasts)

## 'data.frame': 81 obs. of 8 variables:  
## $ X : int 0 1 2 3 4 5 6 7 8 9 ...  
## $ Name : chr "Acromantula" "Basilisk" "Chimaera" "Antipodean Opaleye" ...  
## $ Class : chr "Acromantula" "Basilisk" "Chimaera" "Dragon" ...  
## $ Classification: chr "XXXXX" "XXXXX" "XXXXX" "XXXXX" ...  
## $ Colour : chr "Jet-Black" "Green" "Golden" "Pearly" ...  
## $ Eye : chr "Black" "Yellow" "White" "Multi-coloured" ...  
## $ Native : chr "Island of Borneo" "Greece" "Greece" "New Zealand" ...  
## $ Size : num 180 600 NA NA 300 216 360 600 NA 180 ...

## Задача 2

Добавьте в датафрейм beasts столбец Dragon, состоящий из 0 и 1, где 1 соответствует драконам (значение Dragon в Class).

*Ответ:*

beasts <- beasts %>% mutate(Dragon = ifelse(Class == "Dragon", 1 ,0))

## Задача 3

Добавьте в датафрейм beasts столбец LogSize, содержащий натуральные логарифмы размера зверя.

*Ответ:*

beasts <- beasts %>% mutate(LogSize = log(Size))

## Задача 4

Добавьте в датафрейм beasts столбец Danger со следующими значениями:

* High: если класс Министерства Магии равен 4 или 5 (XXXX или XXXXX);
* Medium: если класс Министерства Магии равен 3 (XXX);
* Low: если класс Министерства Магии равен 1 или 2 (X и XX).

**Подсказка.** Можно несколько раз применить функцию ifelse() или написать свою функцию для преобразований.

*Ответ:*

danger\_class <- function(beasts){  
 a <- c()  
 for (i in beasts$Classification){  
 if (i == "XXXX" | i == "XXXXX"){  
 a <- c(a, "High")  
 } else if (i == "XXX"){  
 a <- c(a, "Medium")  
 } else {  
 a <- c(a, "Low")  
 }  
 }  
 a  
}  
  
beasts <- beasts %>% mutate(Danger = danger\_class(beasts))

## Задача 5

Добавьте в датафрейм beasts столбец DangerNum с числовым значением класса Министерства Магии (X — 1, XX — 2 и так далее до XXXXX).

**Подсказка.** Для определения длины строки может пригодиться базовая функция nchar().

*Ответ:*

beasts <- beasts %>% mutate(DangerNum = nchar(Classification))

## Задача 6

Выберите строки, соответствующие драконам, и сохраните эти строки в датафрейм dragons.

*Ответ:*

dragons <- beasts %>% filter(Dragon == 1)

## Задача 7

Выберите строки, соответствующие существам с классом Министерства Магии не ниже 4 и имеющим жёлтые глаза (Yellow). Сохраните эти строки в датафрейм dang\_yellow.

*Ответ:*

dang\_yellow <- beasts %>% filter(DangerNum >= 4 & Eye == "Yellow")

## Задача 8

Выберите строки, соответствующие существам, родиной которых считается Америка (North America, South America) или Европа (Europe). Сохраните их в датафрейм AmEu.

**Пояснение.** Это означает, что слова *North America*, *South America* и *Europe* должны присутствовать в ячейке в столбце Native, в том числе в составе перечня стран. Например, в AmEu должны уйти существа со значениями Europe, Europe|North America, Europe|Russia|North America|South America и подобные.

*Ответ:*

AmEu <- beasts %>% filter(str\_detect(Native,   
 c("North America", "South America", "Europe")))  
?select

## запускаю httpd сервер помощи... готово

## Задача 9

Сгруппируйте строки по значениям столбца Eye и выведите для каждой группы число существ в каждой группе, а также долю таких существ от общего числа существ в датафрейме.

*Ответ:*

beasts %>% group\_by(Eye) %>% summarise(N= n(),  
 `Share` = N / nrow(beasts))

## # A tibble: 18 x 3  
## Eye N Share  
## <chr> <int> <dbl>  
## 1 "" 1 0.0123  
## 2 "Black" 11 0.136   
## 3 "Blue" 2 0.0247  
## 4 "Blue-green" 1 0.0123  
## 5 "Brilliant purple" 1 0.0123  
## 6 "Brown" 3 0.0370  
## 7 "Deep red" 1 0.0123  
## 8 "Green" 1 0.0123  
## 9 "Grey" 1 0.0123  
## 10 "Hazel" 1 0.0123  
## 11 "Multi-coloured" 1 0.0123  
## 12 "No data" 34 0.420   
## 13 "Orange" 1 0.0123  
## 14 "Red" 3 0.0370  
## 15 "Varies" 8 0.0988  
## 16 "White" 2 0.0247  
## 17 "Yellow" 8 0.0988  
## 18 "Yellow|White|Green" 1 0.0123

## Задача 10

Сгруппируйте строки по значениям столбца Classification и выведите для каждой группы минимальный, средний и максимальный размер существа.

*Ответ:*

class\_group <- beasts %>% group\_by(Classification) %>%  
 summarise(`Min\_size` = min(Size, na.rm = TRUE),  
 `Mean\_size` = mean(Size, na.rm = TRUE),  
 `Max\_size` = max(Size, na.rm = TRUE))

## Задача 11

С помощью библиотеки stargazer сформируйте выдачу, совпадающую по содержанию с таблицей из задачи 10. Выгрузите полученную выдачу в файл с расширением .htm. К выгруженной в файл таблице добавьте заголовок *Fantastic beasts*.

*Ответ:*

library(stargazer)

##   
## Please cite as:

## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.

## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

stargazer(class\_group,  
 summary = FALSE,  
 type = "html",  
 out = "Fantastic\_beasts.htm",  
 title = "Fantastic beasts")

##   
## <table style="text-align:center"><caption><strong>Fantastic beasts</strong></caption>  
## <tr><td colspan="5" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"></td><td>Classification</td><td>Min\_size</td><td>Mean\_size</td><td>Max\_size</td></tr>  
## <tr><td colspan="5" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left">1</td><td>X</td><td>10</td><td>10</td><td>10</td></tr>  
## <tr><td style="text-align:left">2</td><td>XX</td><td>0.05</td><td>16.2928571428571</td><td>60</td></tr>  
## <tr><td style="text-align:left">3</td><td>XXX</td><td>8</td><td>38.4</td><td>120</td></tr>  
## <tr><td style="text-align:left">4</td><td>XXXX</td><td>12</td><td>105</td><td>180</td></tr>  
## <tr><td style="text-align:left">5</td><td>XXXXX</td><td>180</td><td>353.333333333333</td><td>600</td></tr>  
## <tr><td colspan="5" style="border-bottom: 1px solid black"></td></tr></table>