#####################Программирование на языке R. Занятие 4 #####################

######################## Работа с данными в R #######################

## 1) Загрузка данных из файла в рабочую среду.

# 1.1 Загрузка файла по строкам - функция readLines()

setwd("E:/Документы/Работа/Методические материалы по дисциплинам/Программирование на R/Практики - ВВ/Лабораторное занятие № 4-20221018")

# Перед тем, как открыть новый файл через R - оцените его структуру при помощи редактора Notepad++.

# Желательно включить режим отображения специальных символов.

readLines("synapse.txt")

readLines("bison\_data.csv")

read.csv("bison\_data.csv")

is.data.frame(read.csv("bison\_data.csv"))

# 1.2

readLines("bison\_data.txt")

is.character(readLines("bison\_data.txt"))

read.table("bison\_data.txt")

read.table("bison\_data.txt", header = TRUE)

names(read.table("bison\_data.txt"))

names(read.table("bison\_data.txt", header = TRUE))

# 1.3 Продолжим работать с загрузкой таблиц

read.table("iris\_data\_2.txt", header = TRUE)

read.table("iris\_data\_2.txt", header = TRUE, sep = "+")

df <- read.table("iris\_data\_2.txt", header = TRUE, sep = "+")

str(df)

df <- read.table("iris\_data\_2.txt", header = TRUE, sep = "+", dec = ",", as.is = FALSE)

str(df)

##### 2) Как получить доступ к данным в объектах различного типа?

### 2.1. Векторы.

val <- c("a", "b", "c", "d", "e", "f")

val[5]

val[-1]

val[-2]

val[4:6]

val[-(4:6)]

val[length(val)]

val[length(val)-2]

val[val > "c"]

v <- val > "c"

v

val[v]

val[val %in% c("a", "e")]

names(val) <- LETTERS[1:6]

val

val["A"]

val[c("A", "D")]

### 2.2. Как получить доступ к данным в объектах различного типа? Матрицы.

x <- matrix(data=1:6, nrow=2, ncol=3)

x

x[1, 3]

x[2, 2]

x[1, ]

x[, 3]

x[,]

x[x < 3]

### 2.3. Доступ к элементам таблицы

df[,]

df[1,]

df[,1]

df[1:10,]

df[1:10,2:3]

df$Sepal.Length

## 2.4. Отбор элементов по логическому условию

df[df$Species == "setosa",]

subset(df, subset = (Species == "setosa"), select = c(Sepal.Length, Sepal.Width))

#### УПРАЖНЕНИЕ 1.

#### Используйте функцию subset, чтобы отобрать из объекта df в объект df2 строки,

#### содержащие описание ирисов virginica с параметром Sepal.Length 6.5 и выше.

#### Запишите в переменную df3 колонки Petal.Length, Petal.Width и Species для ирисов,

#### относящихся к виду virginica или имеющих Petall.Width меньше единицы.

#### КОНЕЦ УПРАЖНЕНИЯ

#### УПРАЖНЕНИЕ 2.

#### Вы можете использовать функции rbind() и cbind(), чтобы комбинировать несколько таблиц вместе

#### и увеличивать в них количество строк (rbind(), таблицы "ставятся друг на друга")

#### либо столбцов (cbind(), таблицы "ставятся рядом друг с другом").

#### а. Используйте rbind(), чтобы создать таблицу combi1, содержащую общие для df2 и df3 колонки.

#### б. Создайте таблицу combi2, содержащую первые 20 строк таблицы df3 и первые 20 строк из

#### первых двух колонок df2 (именно в таком порядке).

#### в. Выполните следующий код:

cbind(df2[1:20,], df2[21:40,])

#### Объясните, почему в колонки были добавлены NA.

#### КОНЕЦ УПРАЖНЕНИЯ

#### 3) Описание и аннотирование данных в таблице

comment(df) <- "Dataset, created by Ronald Fisher. Content: Iris sepals and petals parameters"

cat(comment(df))

str(df)

#### comment() является частным случаем использования функции attr(). добавим с её помощью комментарии

#### для отдельных колонок.

attr(x = df$Sepal.Length, which = "comment") <- "Column with length of sepals"

str(df)

#### Добавление баннера

install.packages("bannerCommenter")

library(bannerCommenter)

banner(x="This is the banner comment.")

attr(df, which = "comment") <- banner(x="This is the banner comment.")

cat(comment(df))

#### 4) Вспоминаем rownames() и colnames(). Они принимают вектор с именами строк и колонок

#### соответственно. Длина вектора должна соответствовать числу строк и колонок

#### УПРАЖНЕНИЕ 3

#### Отберите 20 любых строк из таблицы df в объект renames. Измените в нём имена колонок на

#### произвольные строки текста, а имена строк - на случайные буквы из встроенного вектора LETTERS (функция sample()).

#### КОНЕЦ УПРАЖНЕНИЯ.

#### 5) Скомбинируем три таблицы в список при помощи функции list()

l <- list(df, df2, df3)

l[1]

is.data.frame(l[1])

l[[1]]

is.data.frame(l[[1]])

l[[1]][1:10, 1:3]

str(l)

names(l) <- c("df1", "df2", "df3")

str(l)

l$df1

l$df3

##### 6) Сохранение данных.

## Запись рабочего окружения на диск

save.image(file = "environment.Rdata")

# Убираем все объекты из окружения

rm(list = ls())

# Загружаем объекты назад

load("environment.Rdata")

## Запись объекта на жёсткий диск в определённом формате часто осуществляется через функцию,

## которая начинается с write

write.table(x = renames, file = "renames.txt")

## Аргументы: quote - заключать ли character и factor в кавычки.

## sep и dec - знак разделителя и десятичного знака.

## rownames и colnames позволяют записывать имена строк и колонок в файл, соответственно.

## Опробуйте их.