Lab 2

Branden Walter

2023-05-18

Question 0

Open a new R Markdown file.

Part A

Delete all sample code from the R markdown file. - Done

Part B

For each question, insert a new chunk and label the question number. - Done

Question 1

Download the data set FlowerData posted under Lab Content in Brightspace (under Lab 2) and save it to whatever file you are using for this course.

Part A

Read the FlowerData file into R and call it F data.

```
Fdata <- read.csv("FlowerData.csv")
```

Part B

Is F data a matrix or a data frame? How did you know that? - It is a data frame

```
class(Fdata)
```

```
## [1] "data.frame"
```

Part C

Create a matrix called FlowerMatrix which contains the numerical columns of F data. Show the first few rows of the matrix.

```
FlowerMatrix <- as.matrix(Fdata[,sapply(Fdata, is.numeric)])
head(FlowerMatrix)
```

```
##      Age..days. Height..cm.
## [1,]         31         5.0
## [2,]         48        16.0
## [3,]         39        12.5
## [4,]         29         6.0
## [5,]         32         4.0
## [6,]         37         7.0
```

Part D

Re-name the columns of Flower Matrix to be: Age (in days) and Height (in cm)

```
colnames(FlowerMatrix) <- c("Age (in days)", "Height (in cm)")
FlowerMatrix
```

```
##       Age (in days) Height (in cm)
##  [1,]            31            5.0
##  [2,]            48           16.0
##  [3,]            39           12.5
##  [4,]            29            6.0
##  [5,]            32            4.0
##  [6,]            37            7.0
##  [7,]            37            8.0
##  [8,]            26            5.5
##  [9,]            41           10.0
## [10,]            34            8.5
## [11,]            38           12.0
## [12,]            40           18.0
## [13,]            45           16.0
## [14,]            40           12.5
## [15,]            43            9.5
## [16,]            33            7.0
## [17,]            35            6.0
## [18,]            39            6.5
## [19,]            37           12.5
## [20,]            32           13.0
## [21,]            31           10.5
## [22,]            36           11.0
## [23,]            41           17.0
## [24,]            39           15.5
## [25,]            31            9.5
## [26,]            33           10.0
## [27,]            33           11.0
## [28,]            28            5.5
## [29,]            35           13.5
## [30,]            37           16.0
## [31,]            42           14.0
## [32,]            45           17.0
## [33,]            37           16.5
## [34,]            46           18.0
## [35,]            37           17.5
## [36,]            44           19.0
## [37,]            44           14.0
## [38,]            37            8.0
## [39,]            29           10.0
## [40,]            37            9.0
## [41,]            36           12.0
## [42,]            47           15.5
## [43,]            38           16.0
## [44,]            40           16.0
## [45,]            45           17.0
```

Part E

Re-name the rows of FlowerMatrix using the Individuals column from F data. Show the first few rows of the matrix again. Note the difference between the two display.

```
rownames(FlowerMatrix) <- Fdata$Individuals
head(FlowerMatrix)
```

```
##      Age (in days) Height (in cm)
## [1,]            31            5.0
## [2,]            48           16.0
## [3,]            39           12.5
## [4,]            29            6.0
## [5,]            32            4.0
## [6,]            37            7.0
```

Question 2

Use the FlowerMatrix matrix: ## Part A Determine the average age of the flowers in the data set

```
mean(FlowerMatrix[, "Age (in days)"])
```

```
## [1] 37.42222
```

Part B

Determine the average height of the flowers in the data set.

```
mean(FlowerMatrix[, "Height (in cm)"])
```

```
## [1] 11.87778
```

Part C

Determine the first individual in the matrix which has the largest height and its height

```
which.max(FlowerMatrix[, "Height (in cm)"])
```

```
## [1] 36
```

```
FlowerMatrix[which.max(FlowerMatrix[, "Height (in cm)"]), "Height (in cm)"]
```

```
## Height (in cm)
##             19
```

Part D

Determine the first individual in the matrix which is the youngest and its age.

```
which.min(FlowerMatrix[, "Age (in days)"])
```

```
## [1] 8
```

```
FlowerMatrix[which.min(FlowerMatrix[, "Age (in days)"]), "Age (in days)"]
```

```
## Age (in days)
##            26
```

Part E

What colour are the flowers in your answer to parts (c) and (d)?

Question 3

Part A

In R console, display the data set Fdata, not FlowerData. There are two potential ”issues” in the data set. Where are they? - The different variable types could be an issue.ex. , , etc.

```
Fdata
```

```
##    Individual Age..days. Height..cm. Colour
## 1           a         31         5.0 purple
## 2           b         48        16.0 yellow
## 3           c         39        12.5    red
## 4           d         29         6.0    red
## 5           e         32         4.0    red
## 6           f         37         7.0 yellow
## 7           g         37         8.0 yellow
## 8           h         26         5.5 yellow
## 9           i         41        10.0 purple
## 10          j         34         8.5 purple
## 11          k         38        12.0     pr
## 12          l         40        18.0 yellow
## 13                    45        16.0 yellow
## 14          n         40        12.5 yellow
## 15          o         43         9.5    red
## 16          p         33         7.0 yellow
## 17          q         35         6.0 yellow
## 18          r         39         6.5 yellow
## 19          s         37        12.5    red
## 20          t         32        13.0 purple
## 21          u         31        10.5 yellow
## 22          v         36        11.0    red
## 23          w         41        17.0    red
## 24          x         39        15.5 yellow
## 25          y         31         9.5 yellow
## 26          z         33        10.0 yellow
## 27         aa         33        11.0 yellow
## 28         bb         28         5.5    red
## 29         cc         35        13.5    red
## 30         dd         37        16.0 yellow
## 31         ee         42        14.0    red
## 32         ff         45        17.0    red
## 33         gg         37        16.5    red
## 34         hh         46        18.0    red
## 35         ii         37        17.5 yellow
## 36         jj         44        19.0 yellow
## 37         kk         44        14.0 yellow
## 38         ll         37         8.0 yellow
## 39         mm         29        10.0 yellow
## 40         nn         37         9.0    red
## 41         oo         36        12.0 purple
## 42         pp         47        15.5 purple
## 43         qq         38        16.0 yellow
## 44         rr         40        16.0    red
## 45         ss         45        17.0 purple
```

Part B

Now you TA will show you how to ”fix” them? - Type cast

```
str(Fdata)
```

```
## 'data.frame':    45 obs. of  4 variables:
##  $ Individual : chr  "a" "b" "c" "d" ...
##  $ Age..days. : int  31 48 39 29 32 37 37 26 41 34 ...
##  $ Height..cm.: num  5 16 12.5 6 4 7 8 5.5 10 8.5 ...
##  $ Colour     : chr  "purple" "yellow" "red" "red" ...
```

Question 4

Once you make sure all the code works in the R markdown file, knit it to either an HTML or Word file. Make sure the file contains all answers to the questions. Then open the knitted file and print it as a PDF file. The name of the file should be Stat123 Lab02 YourLastName.pdf. Then submit the pdf file to the appropriate Brightspace folder.

- Done