RWorksheet_Gallo#4b

Alfe Kevin P. Gallo

2023-11-08

```
#1 . Using the for loop, create an R script that will display a 5x5 matrix as shown in Figure 1. It mus
vectorA <- c(1,2,3,4,5)
matrixa <- matrix(vectorA, nrow = 5,ncol = 5)

vectora <-c(0)
matrixA <- matrix(0, nrow =5,ncol =5)

for (i in length(vectorA)) {
  matrixA[i, ] <- abs(vectorA-vectorA[i])
}
print(matrixA)
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    4    3    2    1    0
```

```
#2 Print the string "*" using for() function. The output should be the same as shown in Figur

for(i in 1:5 ){
  newstar <- rep("*",i)
  print(newstar)
}
```

```
## [1] "*"
## [1] "*" "*"
## [1] "*" "*" "*"
## [1] "*" "*" "*" "*"
## [1] "*" "*" "*" "*" "*"
```

```
#3 Get an input from the user to print the Fibonacci sequence starting from the 1st input up to 500. Us

yournum <- as.integer(readline("Enter the starting Fibonacci sequence number: "))
```

```
## Enter the starting Fibonacci sequence number:
```

```
if(is.na(yournum || yournum < 0)) {
  cat("Enter something else")
} else {

inputnum <- yournum
a <- inputnum
b <- 0
```

```r
cat("Fibonacci sequence starting from", inputnum, ":\n")

repeat {
  num <- a + b
  if (num > 500){
    break
  }
  cat(num, " ")
  a <- b
  b <- num
}

cat("\n")
}
```

```
## Enter something else
```

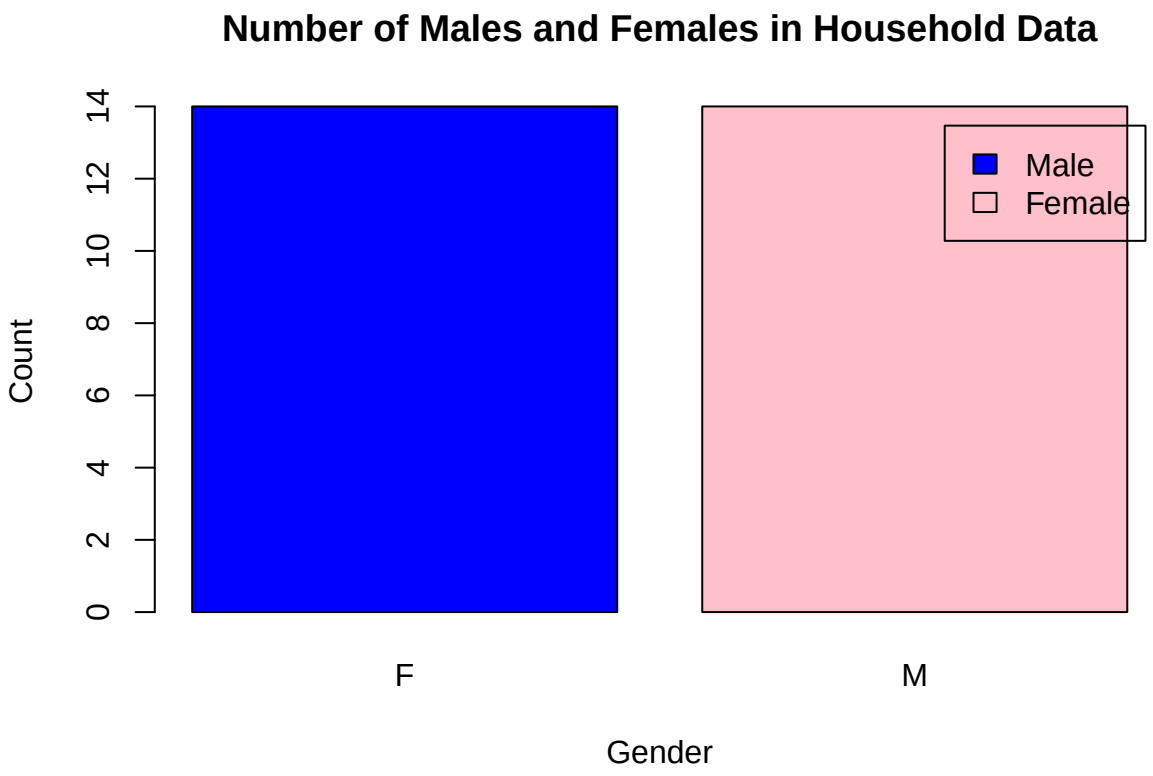
```r
#4Import the dataset as shown in Figure 1 you have created previously


#4a. What is the R script for importing an excel or a csv file? Display the first 6 rows ofthe dataset?

ShoeSdata <- read.csv("Shoe_Sizes")

#4b . Create a subset for gender(female and male). How many observations are there in Male? How about i

MaleSub <- subset(ShoeSdata, Gender == "M")
FemaleSub <- subset(ShoeSdata, Gender == "M")

cat("The number of observation in male subset:", nrow(MaleSub),"\n")
```

```
## The number of observation in male subset: 14
```

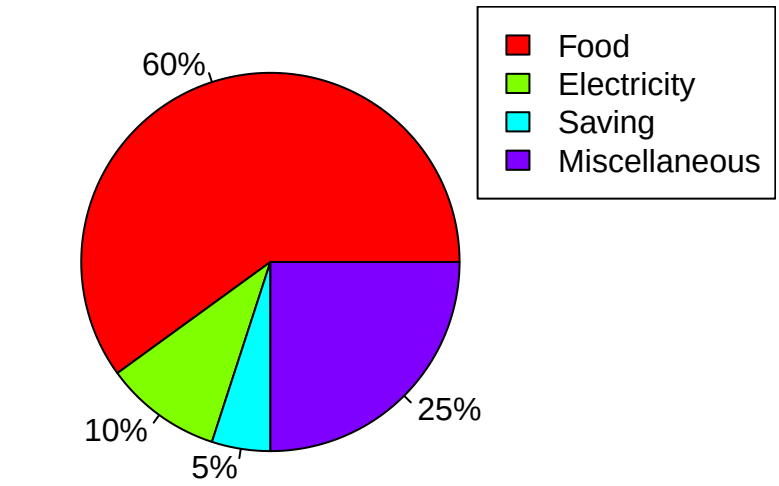
```r
cat("The number of observation in female subset:", nrow(FemaleSub),"\n")
```

```
## The number of observation in female subset: 14
```

```r
#4c  Create a graph for the number of males and females for Household Data. Use plot(),chart type = bar

GenderRibil <- table(ShoeSdata$Gender)


barplot(GenderRibil,
        main = "Number of Males and Females in Household Data",
        xlab = "Gender",
        ylab = "Count",
        col = c("blue", "pink"),
        legend.text = c("Male", "Female"))
```




```
#5 a. Create a piechart that will include labels in percentage.Add some colors and title of the chart.

Dela_Cruz_Fam <- c(60,10,5,25)
 pie(Dela_Cruz_Fam,labels = paste0(Dela_Cruz_Fam,"%"),
     main = "Expenses of Dela Cruz Family",
     col = rainbow(length(Dela_Cruz_Fam)))

 legend("topright",legend =c("Food","Electricity","Saving","Miscellaneous"),
        fill = rainbow(length(Dela_Cruz_Fam)))
```




```
#6 Use the iris dataset
```

```r
#6a a. Check for the structure of the dataset using the str() function. Describe what youhave seen in t

data(iris)
str(iris)
```

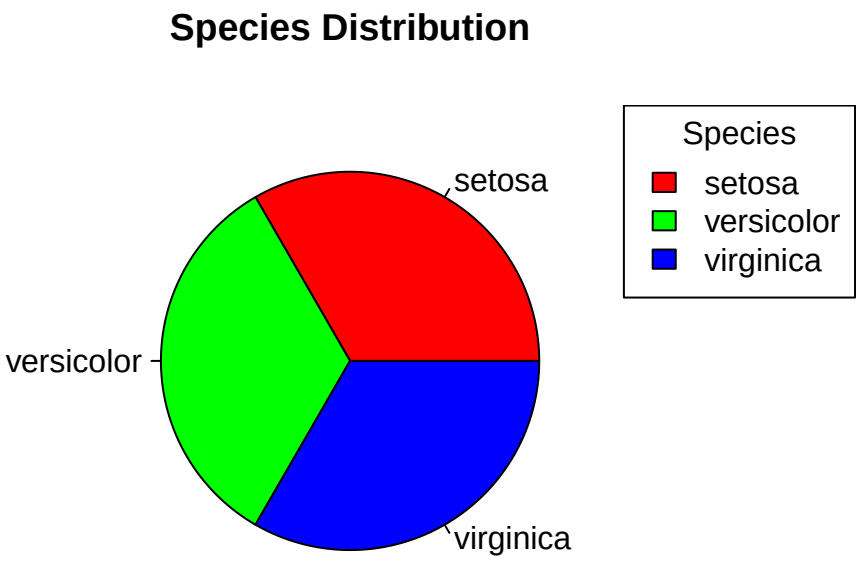
```
## 'data.frame':    150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
##  $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
##  $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
##  $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
##  $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```r
# I saw a data frame of 5 variables which include the numeric measurements for Sepal.Length, Sepal.Widt
```

```r
#6b. Create an R object that will contain the mean of the sepal.length, sepal.width,petal.length,and pe

data(iris)

iris_mean <- colMeans(iris[, 1:4])

iris_mean
```

```
## Sepal.Length  Sepal.Width Petal.Length  Petal.Width
##     5.843333     3.057333     3.758000     1.199333
```

```r
#6c. Create a pie chart for the Species distribution. Add title, legends, and colors. Write the R scrip
data(iris)

species <- table(iris$Species)

pie(species, labels = names(species),
    col = rainbow(length(species)),
    main = "Species Distribution")

legend("topright", legend = names(species),
       fill = rainbow(length(species)), title = "Species")
```



```r
#6d. Subset the species into setosa, versicolor, and virginica. Write the R scripts and show the last s
data(iris)

setosa_sub <- subset(iris, Species == "setosa")
versicolor_sub <- subset(iris, Species == "versicolor")
virginica_sub <- subset(iris, Species == "virginica")

#to display the last 6 rows of each species

tail(setosa_sub)
```

```
##    Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 45          5.1         3.8          1.9         0.4  setosa
## 46          4.8         3.0          1.4         0.3  setosa
## 47          5.1         3.8          1.6         0.2  setosa
## 48          4.6         3.2          1.4         0.2  setosa
## 49          5.3         3.7          1.5         0.2  setosa
## 50          5.0         3.3          1.4         0.2  setosa
```

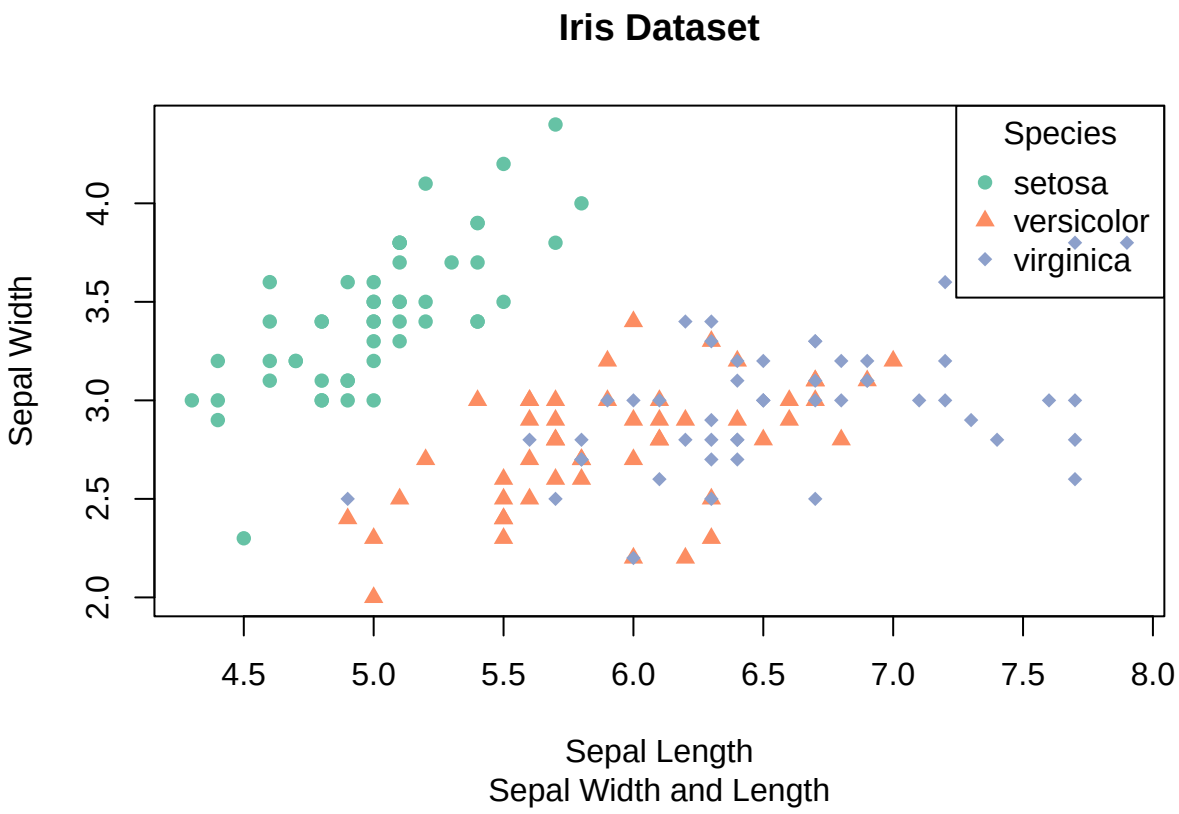
```r
tail(versicolor_sub)
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
## 95           5.6         2.7          4.2         1.3 versicolor
## 96           5.7         3.0          4.2         1.2 versicolor
## 97           5.7         2.9          4.2         1.3 versicolor
## 98           6.2         2.9          4.3         1.3 versicolor
## 99           5.1         2.5          3.0         1.1 versicolor
## 100          5.7         2.8          4.1         1.3 versicolor
```

```r
tail(virginica_sub)
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 145          6.7         3.3          5.7         2.5 virginica
## 146          6.7         3.0          5.2         2.3 virginica
## 147          6.3         2.5          5.0         1.9 virginica
## 148          6.5         3.0          5.2         2.0 virginica
## 149          6.2         3.4          5.4         2.3 virginica
## 150          5.9         3.0          5.1         1.8 virginica
```

```r
#6e. Create a scatterplot of the sepal.length and sepal.width using the different species(setosa,versic
data(iris)

iris$Species <- as.factor(iris$Species)

colors <- c("setosa" = "#66c2a5", "versicolor" = "#fc8d62", "virginica" = "#8da0cb")
symbols <- c("setosa" = 16, "versicolor" = 17, "virginica" = 18)

plot(iris$Sepal.Length, iris$Sepal.Width,
     col = colors[iris$Species],
     pch = symbols[iris$Species],
     main = "Iris Dataset",
     sub = "Sepal Width and Length",
     xlab = "Sepal Length",
     ylab = "Sepal Width")

legend("topright",legend = levels(iris$Species),col= colors, pch = symbols, title = "Species")
```




```
#6e
#by factoring the species, it will be represents as a categories in R.
```

```
#7.
library(readxl)

alexa <- read_excel("alexa_file.xlsx")

alexa
```

```
## # A tibble: 3,150 x 5
##    rating date                variation           verified_reviews      feedback
##     <dbl> <dttm>              <chr>               <chr>                    <dbl>
##  1      5 2018-07-31 00:00:00 Charcoal Fabric     Love my Echo!                1
##  2      5 2018-07-31 00:00:00 Charcoal Fabric     Loved it!                    1
##  3      4 2018-07-31 00:00:00 Walnut Finish       Sometimes while play~        1
##  4      5 2018-07-31 00:00:00 Charcoal Fabric     I have had a lot of ~        1
##  5      5 2018-07-31 00:00:00 Charcoal Fabric     Music                        1
##  6      5 2018-07-31 00:00:00 Heather Gray Fabric I received the echo ~        1
##  7      3 2018-07-31 00:00:00 Sandstone Fabric    Without having a cel~        1
##  8      5 2018-07-31 00:00:00 Charcoal Fabric     I think this is the ~        1
##  9      5 2018-07-30 00:00:00 Heather Gray Fabric looks great                  1
## 10      5 2018-07-30 00:00:00 Heather Gray Fabric Love it! I've listen~        1
## # i 3,140 more rows
```

```
#7a. Rename the white and black variants by using gsub() function.
#black
alexa$variation <- gsub("Black Dot", "BlackDot", alexa$variation)
```

```
alexa$variation <- gsub("Black Plus", "BlackPlus", alexa$variation)
alexa$variation <- gsub("Black Show", "BlackShow", alexa$variation)
alexa$variation <- gsub("Black Spot", "BlackSpot", alexa$variation)

#white
alexa$variation <- gsub("White Dot", "WhiteDot", alexa$variation)
alexa$variation <- gsub("White Plus", "WhitePlus", alexa$variation)
alexa$variation <- gsub("White Show", "WhiteShow", alexa$variation)
alexa$variation <- gsub("White Spot", "WhiteSpot", alexa$variation)

alexa
```

```
## # A tibble: 3,150 x 5
##    rating date                variation           verified_reviews      feedback
##     <dbl> <dttm>              <chr>               <chr>                    <dbl>
##  1      5 2018-07-31 00:00:00 Charcoal Fabric     Love my Echo!                1
##  2      5 2018-07-31 00:00:00 Charcoal Fabric     Loved it!                    1
##  3      4 2018-07-31 00:00:00 Walnut Finish       Sometimes while play~        1
##  4      5 2018-07-31 00:00:00 Charcoal Fabric     I have had a lot of ~        1
##  5      5 2018-07-31 00:00:00 Charcoal Fabric     Music                        1
##  6      5 2018-07-31 00:00:00 Heather Gray Fabric I received the echo ~        1
##  7      3 2018-07-31 00:00:00 Sandstone Fabric    Without having a cel~        1
##  8      5 2018-07-31 00:00:00 Charcoal Fabric     I think this is the ~        1
##  9      5 2018-07-30 00:00:00 Heather Gray Fabric looks great                  1
## 10      5 2018-07-30 00:00:00 Heather Gray Fabric Love it! I've listen~        1
## # i 3,140 more rows
```

```
#7b . Get the total number of each variations and save it into another object. Save the object as varia
library("dplyr")
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
var_TOTAL <- alexa %>%
  count(alexa$variation)

var_TOTAL
```

```
## # A tibble: 16 x 2
##    `alexa$variation`                n
##    <chr>                        <int>
##  1 Black                          261
##  2 Black  Dot                    516
##  3 Black  Plus                   270
##  4 Black  Show                   265
##  5 Black  Spot                   241
##  6 Charcoal Fabric               430
##  7 Configuration: Fire TV Stick  350
```

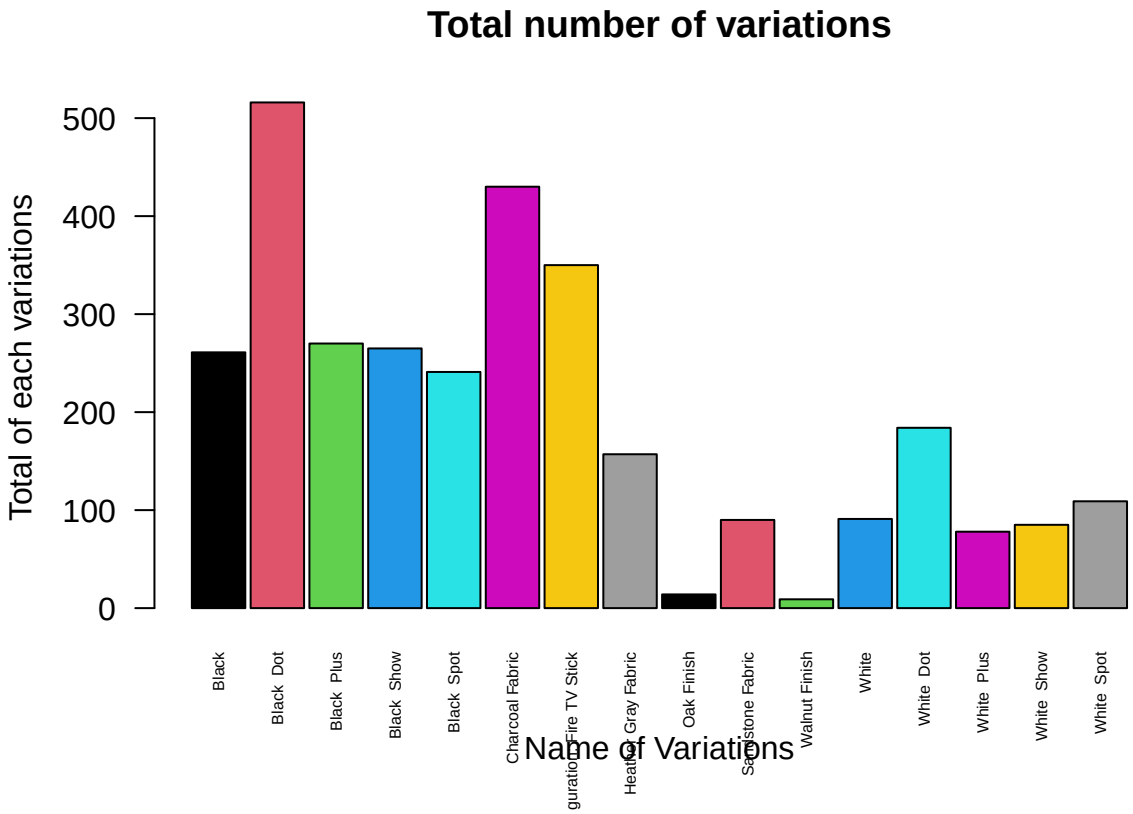
```
##  8 Heather Gray Fabric            157
##  9 Oak Finish                      14
## 10 Sandstone Fabric                90
## 11 Walnut Finish                    9
## 12 White                           91
## 13 White  Dot                     184
## 14 White  Plus                     78
## 15 White  Show                     85
## 16 White  Spot                    109
```

```
save(var_TOTAL, file= "variations.RData")
```

```
#7c.  From the variations.RData, create a barplot(). Complete the details of the chart which include th
load("variations.RData")
var_TOTAL
```

```
## # A tibble: 16 x 2
##    `alexa$variation`                n
##    <chr>                        <int>
##  1 Black                          261
##  2 Black  Dot                     516
##  3 Black  Plus                    270
##  4 Black  Show                    265
##  5 Black  Spot                    241
##  6 Charcoal Fabric                430
##  7 Configuration: Fire TV Stick   350
##  8 Heather Gray Fabric            157
##  9 Oak Finish                      14
## 10 Sandstone Fabric                90
## 11 Walnut Finish                    9
## 12 White                           91
## 13 White  Dot                     184
## 14 White  Plus                     78
## 15 White  Show                     85
## 16 White  Spot                    109
```

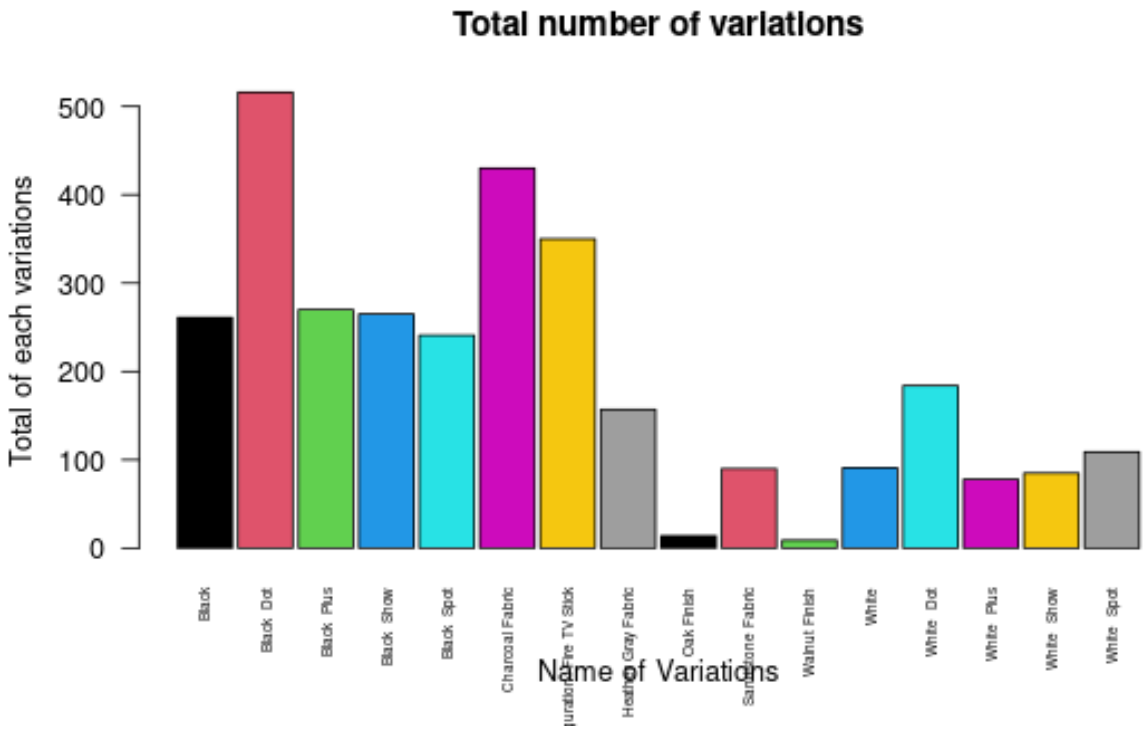
```
namevar <- var_TOTAL$`alexa$variation`

alexaplot <- barplot(var_TOTAL$n,
                     names.arg = namevar,
                     main = "Total number of variations",
                     xlab = "Name of Variations",
                     ylab = "Total of each variations",
                     col = 1:16,
                     space = 0.1,
                     cex.names = 0.5,
                     las = 2)
```




```
png("alexaplot.png")
dev.off()
```

```
## pdf
##   2
```

```
knitr::include_graphics("/cloud/project/Worksheet#4/alexaplot.png")
```



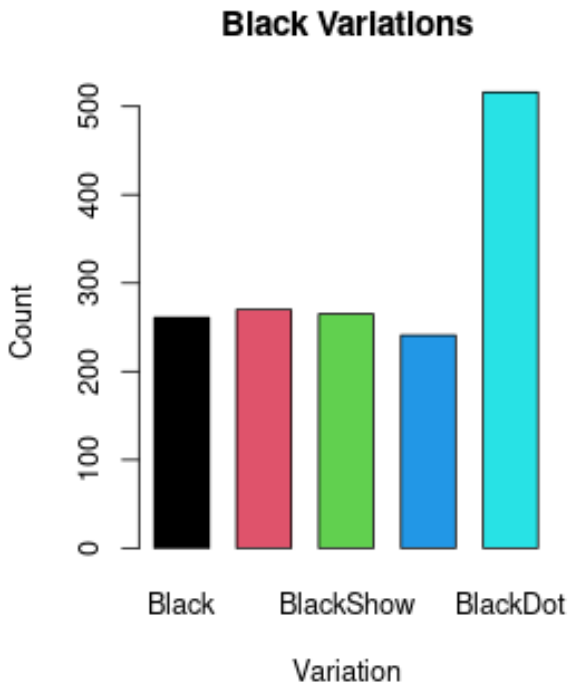

```
#7d.  Create a barplot() for the black and white variations. Plot it in 1 frame, side byside. Complete
library(RColorBrewer)

par(mfrow = c(1,2))


blackvarplot <- barplot(height = c(261,270,265,241,516),
                        names.arg = c("Black","BlackPlus","BlackShow","BlackSpot","BlackDot"), main = "
                        col = 1:5,
                        space = 0.5,
                        xlab = "Variation",
                        ylab = "Count")

png("blackvarplot.png")
dev.off()
```

```
## pdf
##   2
```

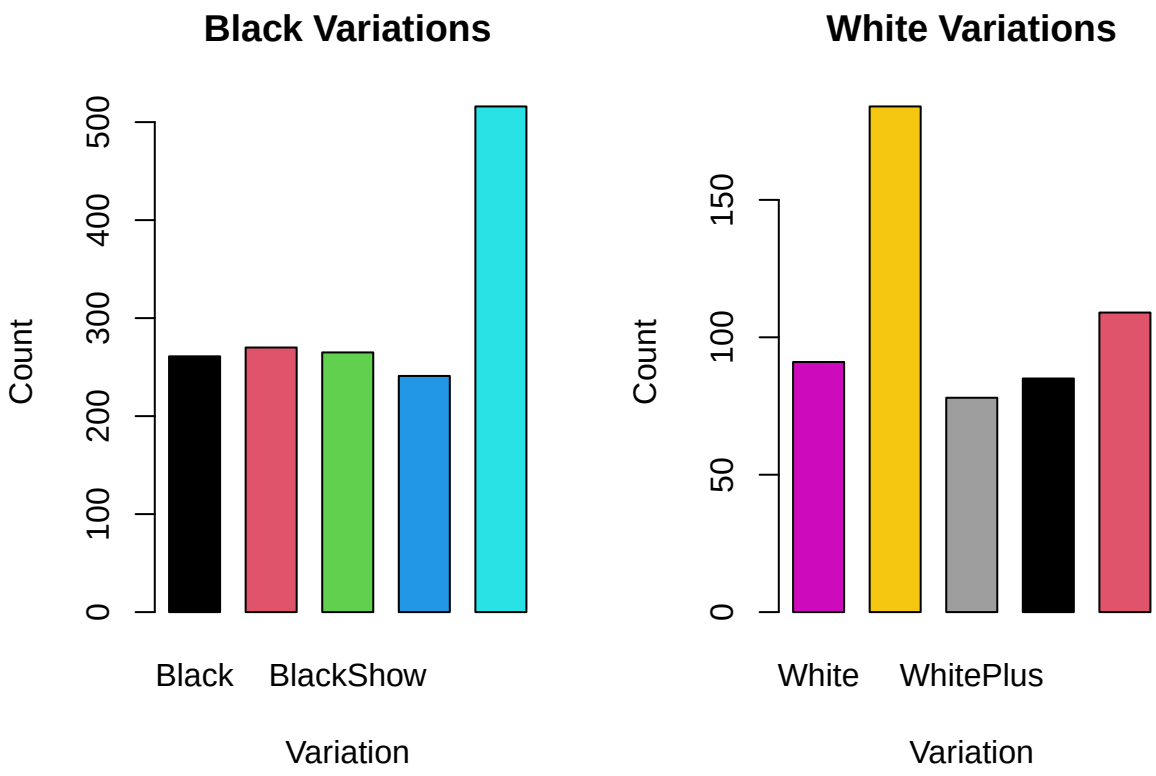
```
getwd()
```

```
## [1] "/cloud/project/Worksheet#4"
```

```
knitr::include_graphics("/cloud/project/Worksheet#4/blackvarplot.png")
```



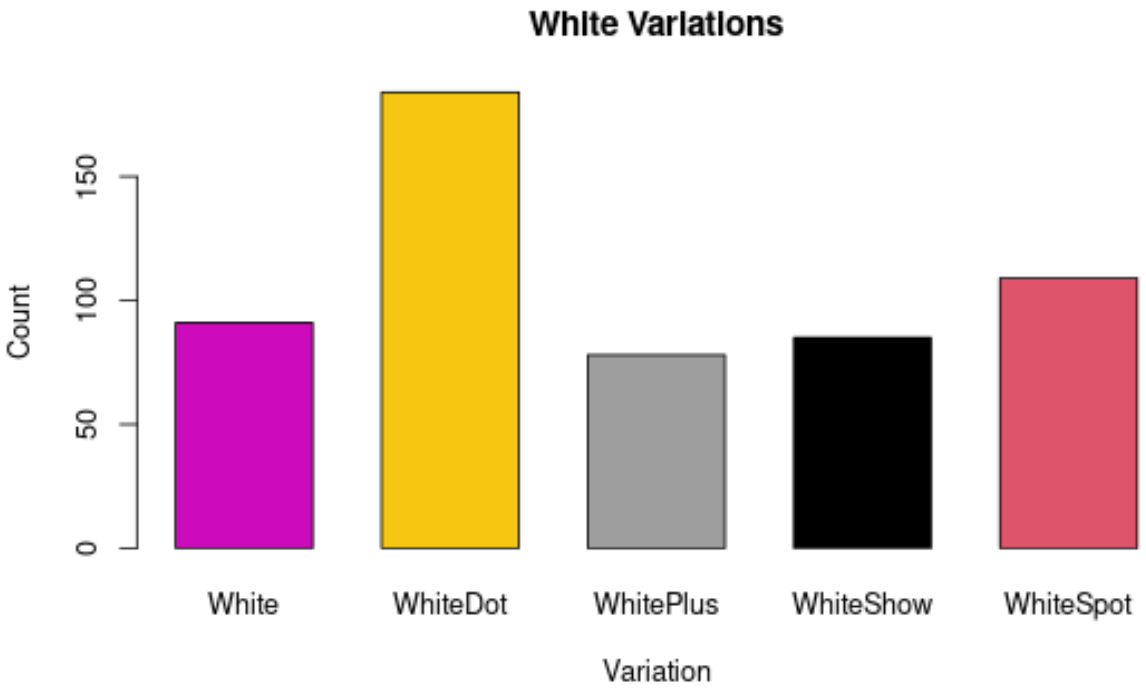

```
whitevarplot <- barplot(height = c(91,184,78,85,109),
                        names.arg = c("White", "WhiteDot", "WhitePlus", "WhiteShow", "WhiteSpot"),
                        main = "White Variations",
                        space = 0.5,
                        col = 6:10,
                        xlab = "Variation",
                        ylab = "Count",)
```




```
png("whitevarplot.png")
dev.off()
```

```
## pdf
##   2
```

```
knitr::include_graphics("/cloud/project/Worksheet#4/whitevarplot.png")
```


White Variations
Count
0
50
100
150
White
WhiteDot
WhitePlus
WhiteShow
WhiteSpot
Variation