Rworksheet_Mabalina#4b

Kristian B. Mabalina

2024-10-29

1

```
vector <-c(1,2,3,4,5)
m1 <-matrix(nrow =5, ncol = 5)

for(i in 1:5){
  for(j in 1:5){
     m1[i, j] <- vector[i] + vector[j]
  }
}
print(m1)
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    2    3    4    5    6
## [2,]    3    4    5    6    7
## [3,]    4    5    6    7    8
## [4,]    5    6    7    8    9
## [5,]    6    7    8    9   10
```

2

```
num_rows <- 5
for (i in 1:num_rows) {
  cat(rep("*", i), "\n")
}
```

```
## *
## * *
## * * *
## * * * *
## * * * * *
```

3

```
#start <- as.integer(readline(prompt = "Enter the starting number for the Fibonacci sequence: "))
start <- 5
a <- 0
b <- 1
repeat {
  fib <- a + b
```

```
  a <- b
  b <- fib
  if (fib >= start) {
    cat(fib, " ")
  }
  if (fib >= 500) {
    break
  }
}
```

```
## 5  8  13  21  34  55  89  144  233  377  610
```

4

```
shoes <- read.csv("4bnum3.csv")
shoes
```

```
##    Shoe.size.1 Height.1 Gender.1
## 1          6.5     66.0        F
## 2          9.0     68.0        F
## 3          8.5     64.5        F
## 4          9.0     65.0        F
## 5         10.5     70.0        M
## 6          7.5     64.0        F
## 7          9.0     70.0        F
## 8          9.5     66.0        F
## 9         10.0     72.0        M
## 10         7.5     60.0        F
## 11        10.5     74.5        M
## 12         6.5     72.0        F
## 13        12.0     71.0        M
## 14        10.5     71.0        M
## 15        13.0     77.0        M
## 16        11.5     72.0        F
## 17         8.5     59.0        F
## 18         5.0     62.0        F
## 19        10.0     72.0        M
## 20         6.5     66.0        F
## 21         7.5     68.0        F
## 22        10.5     67.0        M
## 23         8.5     73.0        M
## 24         8.5     69.0        F
## 25        10.5     72.0        M
## 26        11.0     72.0        M
## 27         9.0     69.0        M
## 28        13.0     70.0        M
```

#b

```
femalesubset <- subset(shoes, Gender.1 == "F")
malesubset <- subset(shoes, Gender.1 == "M")
femalecount <- nrow(femalesubset)
malecount <- nrow(malesubset)
cat("Female:", femalecount, "\n")
```

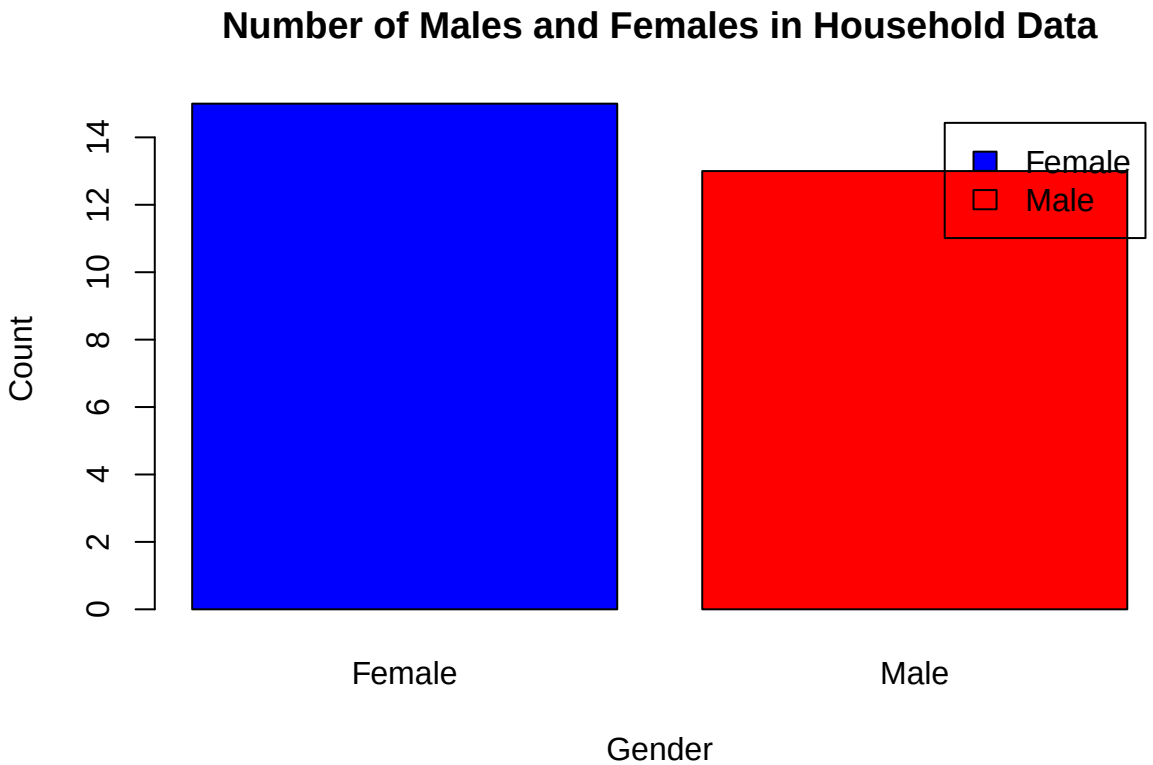
```
## Female: 15
```

```
cat("Male:", malecount, "\n")
```

```
## Male: 13
```

#c

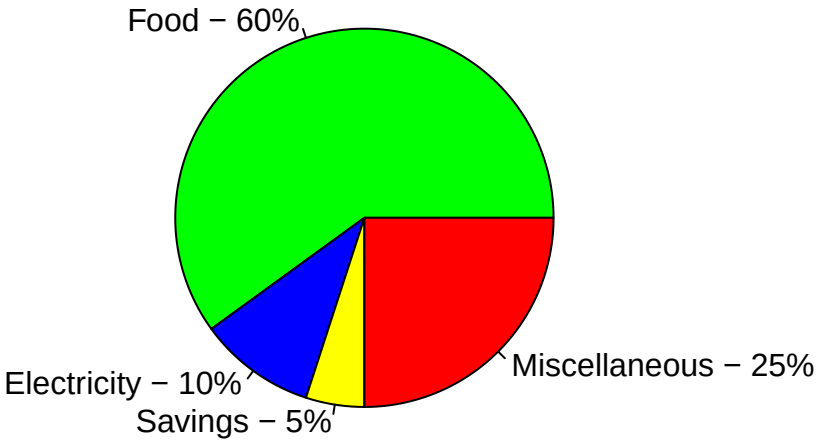
```
gendercounts <- table(shoes$Gender.1)
barplot(gendercounts,
        main = "Number of Males and Females in Household Data",
        xlab = "Gender",
        ylab = "Count",
        col = c("blue", "red"),
        legend = c("Female", "Male"),
        names.arg = c("Female", "Male"))
```



5

```
expenses <- c(Food = 60, Electricity = 10, Savings = 5, Miscellaneous = 25)
percent_labels <- paste0(names(expenses), " - ", round(expenses / sum(expenses) * 100), "%")
pie(expenses,
    labels = percent_labels,
    col = c("green", "blue", "yellow", "red"),
    main = "Dela Cruz Family Monthly Expenses")
```

Dela Cruz Family Monthly Expenses



6

```
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 ...
```

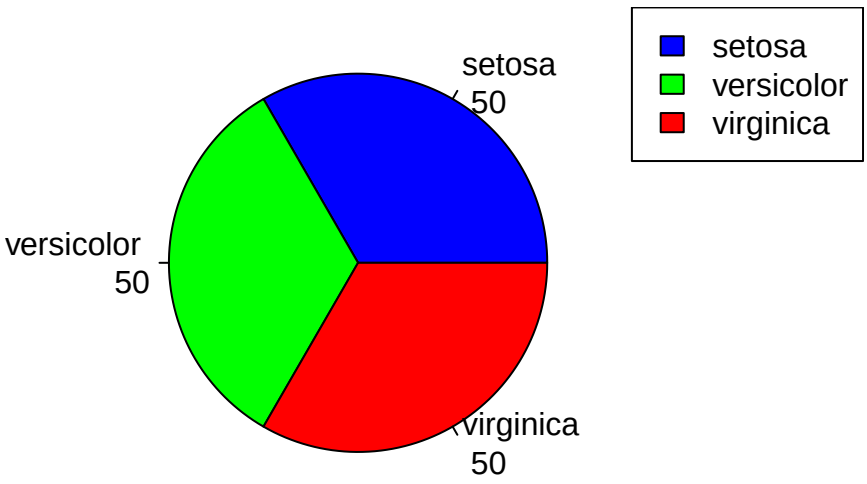
#b

```
colMeans(iris[, 1:4])
```

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

#c

```
species_counts <- table(iris$Species)
pie(species_counts,
    main = "Species Distribution in Iris Dataset",
    col = c("blue", "green", "red"),
    labels = paste(names(species_counts), "\n", species_counts))
legend("topright",
       legend = names(species_counts),
       fill = c("blue", "green", "red"))
```



#d

```
setosa <- subset(iris, Species == "setosa")
versicolor <- subset(iris, Species == "versicolor")
virginica <- subset(iris, Species == "virginica")
tail(setosa)
```

```
##    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
```

```
tail(versicolor)
```

```
##     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
```

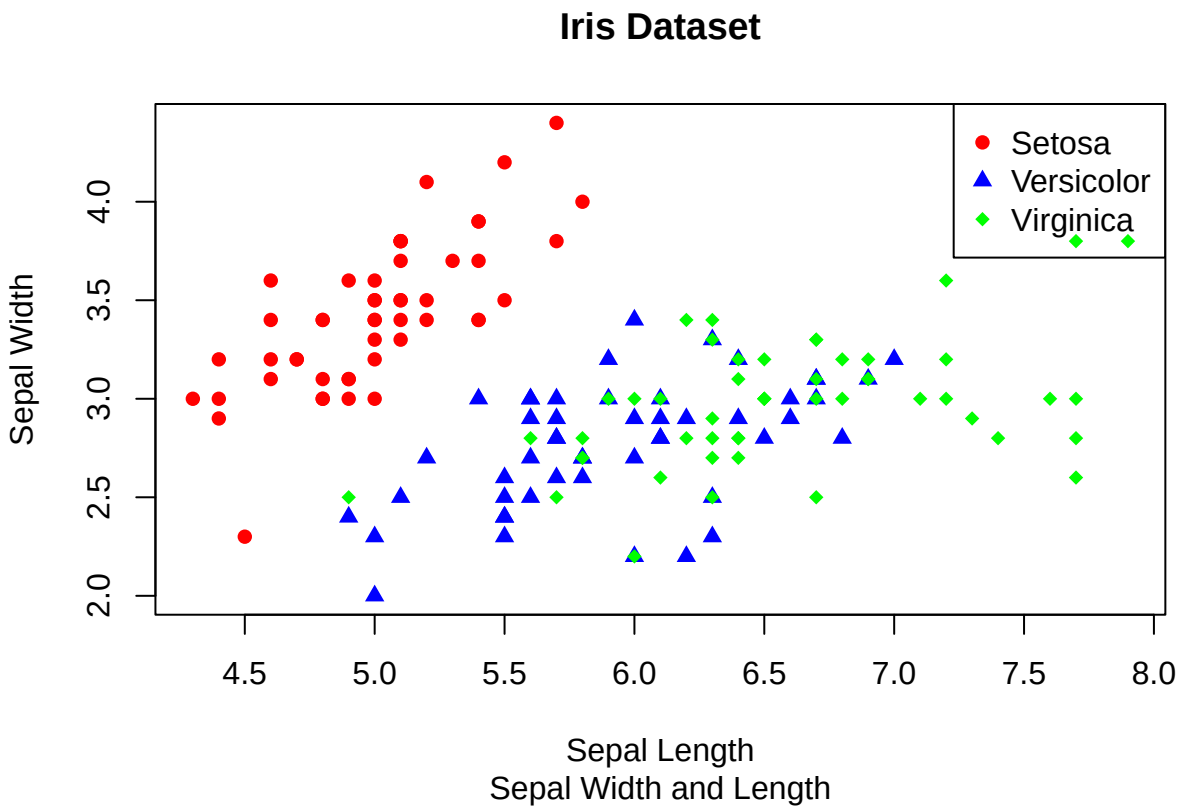
```
tail(virginica)
```

```
##     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
```

#e

```
colors <- c("setosa" = "red", "versicolor" = "blue", "virginica" = "green")
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 = c("Setosa", "Versicolor", "Virginica"),
       col = c("red", "blue", "green"),
       pch = c(16, 17, 18))
```



#f #–The data structure proved suitable for both initial exploratory analysis and in-depth statistical modeling. #–Mean values offered a quick snapshot of the key characteristics of the iris flowers. #–The pie chart effectively showcased the species distribution, highlighting setosa as the dominant species. #–Subsetting the dataset enabled a closer examination of each species, aiding in focused, species-specific analysis. #–The scatterplot revealed the correlation between sepal length and width, visually distinguishing species and allowing further investigation into their relationships.

7

```
library(readxl)
alexa_data <- read_excel("alexa_file.xlsx")
unique(alexa_data$variation)
```

```
##  [1] "Charcoal Fabric"          "Walnut Finish"
##  [3] "Heather Gray Fabric"      "Sandstone Fabric"
##  [5] "Oak Finish"               "Black"
##  [7] "White"                    "Black  Spot"
##  [9] "White  Spot"              "Black  Show"
```

```
## [11] "White  Show"                  "Black  Plus"
## [13] "White  Plus"                  "Configuration: Fire TV Stick"
## [15] "Black  Dot"                   "White  Dot"
```

```
alexa_data$variation <- gsub("Black Dot", "BlackDot", alexa_data$variation)
alexa_data$variation <- gsub("Black Plus", "BlackPlus", alexa_data$variation)
alexa_data$variation <- gsub("Black Show", "BlackShow", alexa_data$variation)
alexa_data$variation <- gsub("Black Spot", "BlackSpot", alexa_data$variation)
alexa_data$variation <- gsub("White Dot", "WhiteDot", alexa_data$variation)
alexa_data$variation <- gsub("White Plus", "WhitePlus", alexa_data$variation)
alexa_data$variation <- gsub("White Show", "WhiteShow", alexa_data$variation)
alexa_data$variation <- gsub("White Spot", "WhiteSpot", alexa_data$variation)
unique(alexa_data$variation)
```

```
##  [1] "Charcoal Fabric"              "Walnut Finish"
##  [3] "Heather Gray Fabric"          "Sandstone Fabric"
##  [5] "Oak Finish"                   "Black"
##  [7] "White"                        "Black  Spot"
##  [9] "White  Spot"                  "Black  Show"
## [11] "White  Show"                  "Black  Plus"
## [13] "White  Plus"                  "Configuration: Fire TV Stick"
## [15] "Black  Dot"                   "White  Dot"
```

#b

```
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
```

```
variation_counts <- alexa_data %>%
  count(variation)
print(variation_counts)
```

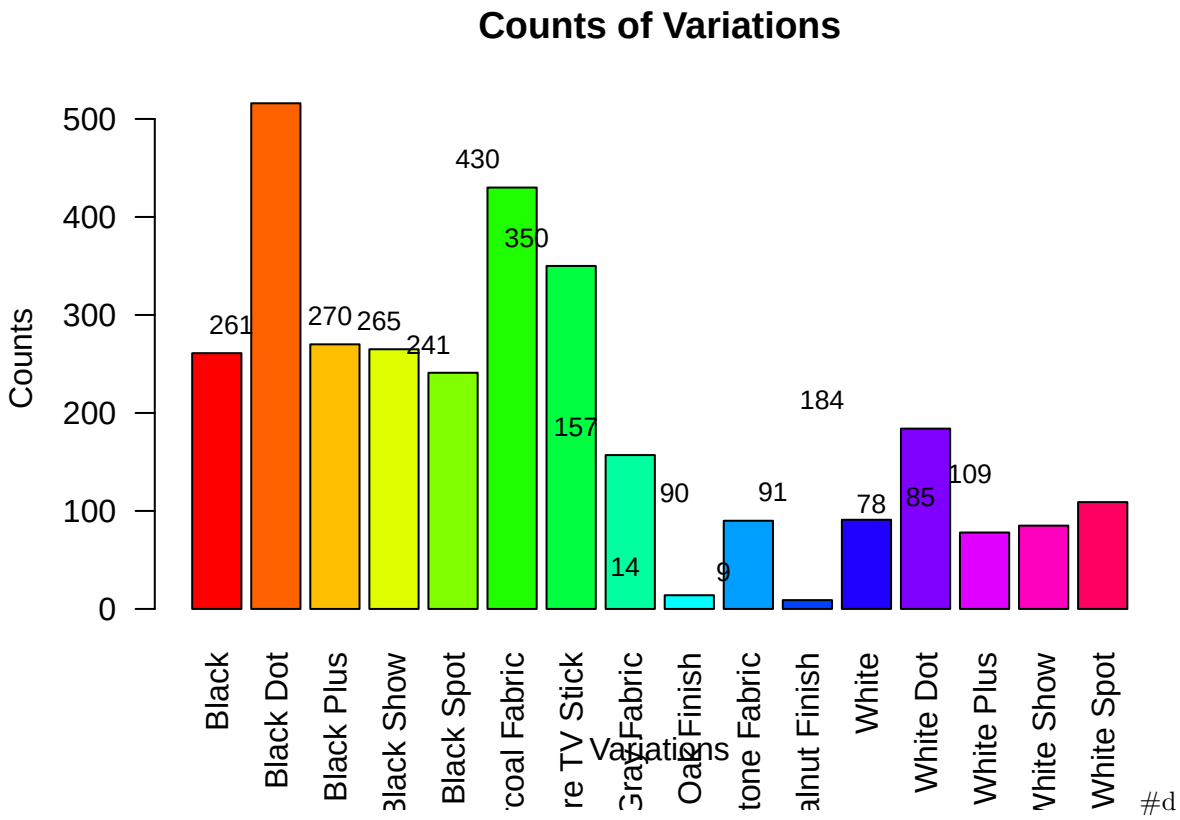
```
## # A tibble: 16 x 2
##    variation                        n
##    <chr>                        <int>
##  1 Black                          261
##  2 Black  Dot                     516
##  3 Black  Plus                    270
##  4 Black  Show                    265
##  5 Black  Spot                    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(variation_counts, file = "variations.RData")
```

#c

```
library(dplyr)
load("variations.RData")
variation_counts$variation <- gsub(" +", " ", variation_counts$variation)
variation_counts$variation <- trimws(variation_counts$variation)
bar_data <- variation_counts$n
bar_names <- variation_counts$variation
barplot(
  bar_data,
  main = "Counts of Variations",
  col = rainbow(length(bar_data)),
  names.arg = bar_names,
  xlab = "Variations",
  ylab = "Counts",
  las = 2,
  border = "black"
)
text(x = seq_along(bar_data), y = bar_data, labels = bar_data, pos = 3, cex = 0.8, col = "black")
```



#d

```
library(ggplot2)
library(dplyr)
load("variations.RData")
variation_counts$variation <- gsub(" +", " ", variation_counts$variation)
```

```r
variation_counts$variation <- trimws(variation_counts$variation)
bw_variations <- variation_counts %>%
  filter(grepl("Black|White", variation))
bar_data <- as.matrix(bw_variations$n)
bar_names <- bw_variations$variation
barplot(
  bar_data,
  beside = TRUE,
  main = "Counts of Black and White Variations",
  col = c("black", "gray", "lightgray", "white"),
  names.arg = bar_names,
  xlab = "Variations",
  ylab = "Counts",
  las = 2,
  border = "black"
)
text(x = seq_along(bar_data), y = bar_data, labels = bar_data, pos = 3, cex = 0.8, col = "black")
```