RWorksheet_Celestra#4b

Kenneth Celestra

2024-10-30

1.

```
vectorA <- c(1, 2, 3, 4, 5)

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

for (i in 1:5) {
  for (j in 1:5) {
    matrixA[i, j] <- abs(i-j)
  }
}

print(matrixA)
```

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

2.

```
# Number of rows
num_rows <- 5


for (i in 1:num_rows)
  cat(strrep(' "*"', i), "\n")
```

```
##  "*"
##  "*" "*"
##  "*" "*" "*"
##  "*" "*" "*" "*"
##  "*" "*" "*" "*" "*"
```

3.

```
# start <- as.integer(readline(prompt = "Enter the starting number for Fibonacci sequence: "))

start <- 1
a <- start
b <- 1

repeat {
  if (a >= start) {
```

```
    cat(a, " ")
  }

  next_number <- a + b
  a <- b
  b <- next_number

  if (a > 500) {
    break
  }
}
```

```
## 1  1  2  3  5  8  13  21  34  55  89  144  233  377
```

4. a

```
data <- read.csv("Shoe_sizes.csv")
head(data)
```

```
##   Show.Size Height Gender
## 1       6.5   66.0      F
## 2       9.0   68.0      F
## 3       8.5   64.5      F
## 4       8.5   65.0      F
## 5      10.5   70.0      M
## 6       7.0   64.0      F
```

b

```
maledata <- subset(data, Gender == "M")
femaledata <- subset(data, Gender == "F")

num_male <- nrow(maledata)
num_female <- nrow(femaledata)

cat("Number of Male observations:", num_male, "\n")
```

```
## Number of Male observations: 14
```

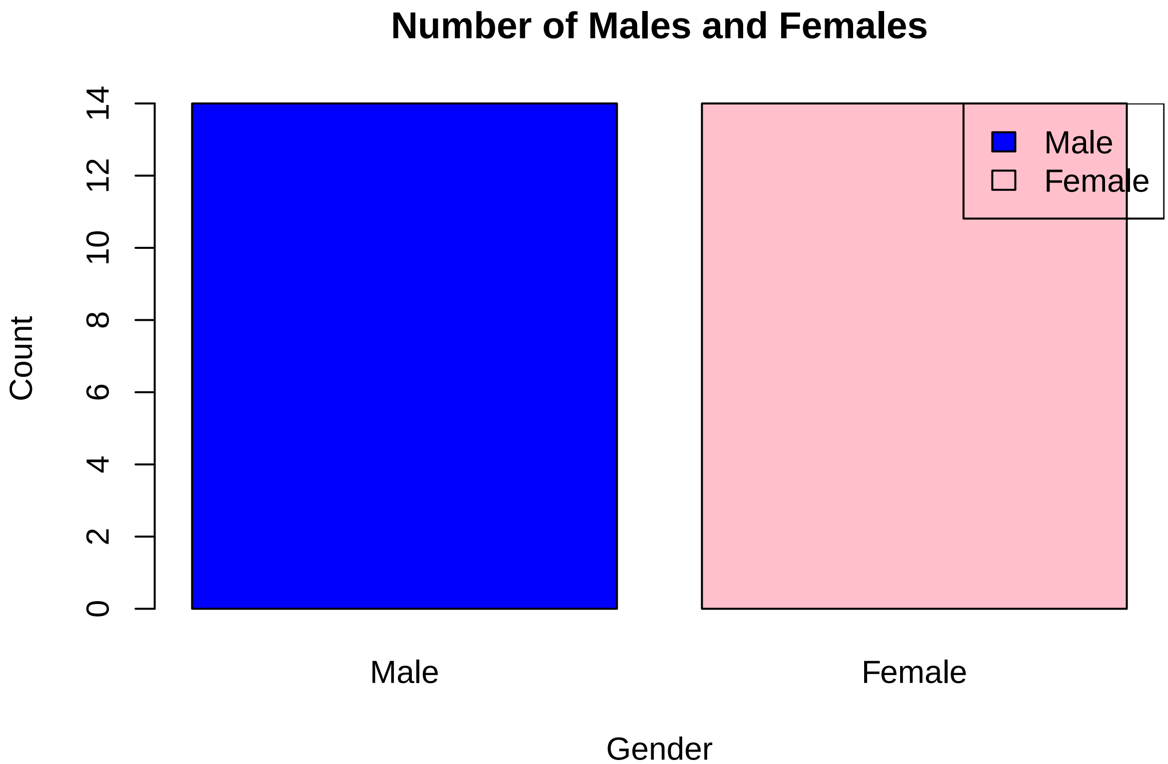
```
cat("Number of Female observations:", num_female, "\n")
```

```
## Number of Female observations: 14
```

c

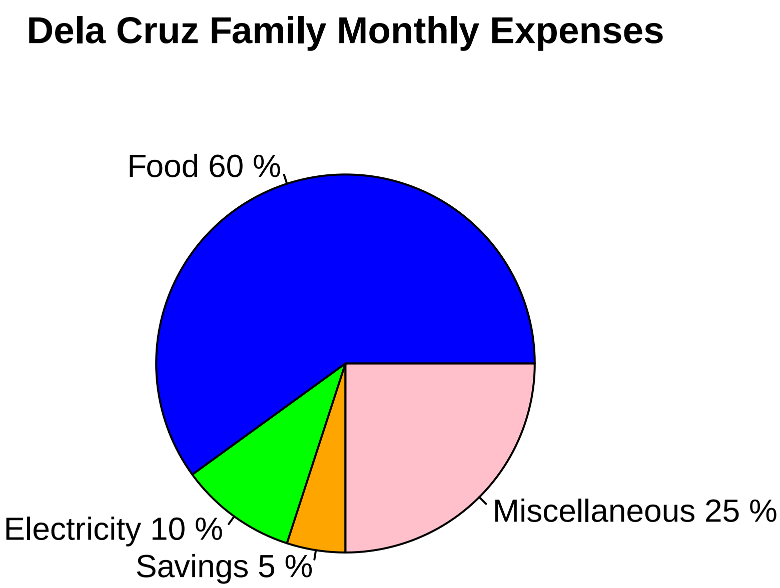
```
# Sample data for Household Data with counts of males and females
gender_counts <- table(data$Gender)

barplot(
  gender_counts,
  main = "Number of Males and Females",
  col = c("blue", "pink"), # Colors for bars
  names.arg = c("Male", "Female"),
  xlab = "Gender",
  ylab = "Count",
)
legend(
  "topright",
```

```
  legend = c("Male", "Female"),
  fill = c("blue", "pink")
)
```



5. a

```
expenses <- c(Food = 60, Electricity = 10, Savings = 5, Miscellaneous = 25)

percent_labels <- paste(names(expenses), round(expenses / sum(expenses) * 100, 1), "%")

pie(
  expenses,
  labels = percent_labels,
  col = c("blue", "green", "orange", "pink"),
  main = "Dela Cruz Family Monthly Expenses"
)
```



6.

```r
data(iris)
```

a

```r
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 ...
```

The Sepal.Length representing the length of the sepal (in centimeters). Sepal.Width representing the width of the sepal. Petal.Length representing the length of the petal. Petal.Width representing the width of the petal. Species A factor variable with 3 levels: “setosa”, “versicolor”, and “virginica”, representing three species of iris flowers.

b

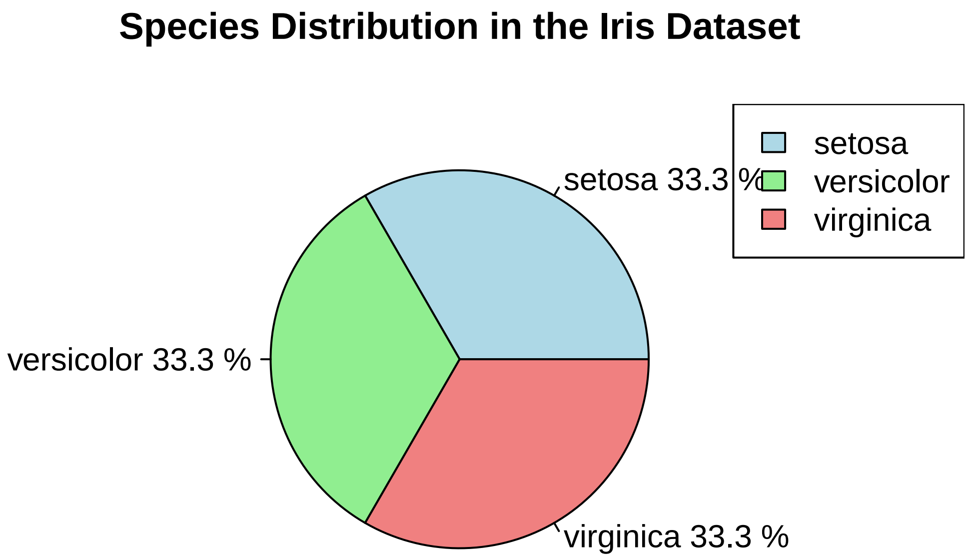
```r
mean_values <- colMeans(iris[, 1:4])

mean_values
```

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

c

```r
species_counts <- table(iris$Species)

percent_labels <- paste(names(species_counts), round(species_counts / sum(species_counts) * 100, 1), "%

pie(
  species_counts,
  labels = percent_labels,
  col = c("lightblue", "lightgreen", "lightcoral"), # Colors for each species
  main = "Species Distribution in the Iris Dataset" # Title
)
```

```r
legend(
  "topright",
  legend = names(species_counts),
  fill = c("lightblue", "lightgreen", "lightcoral")
)
```



d

```r
setosa_subset <- subset(iris, Species == "setosa")
versicolor_subset <- subset(iris, Species == "versicolor")
virginica_subset <- subset(iris, Species == "virginica")

print(tail(setosa_subset, 6))
```

```
##    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
print(tail(versicolor_subset, 6))
```

```
##     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
```

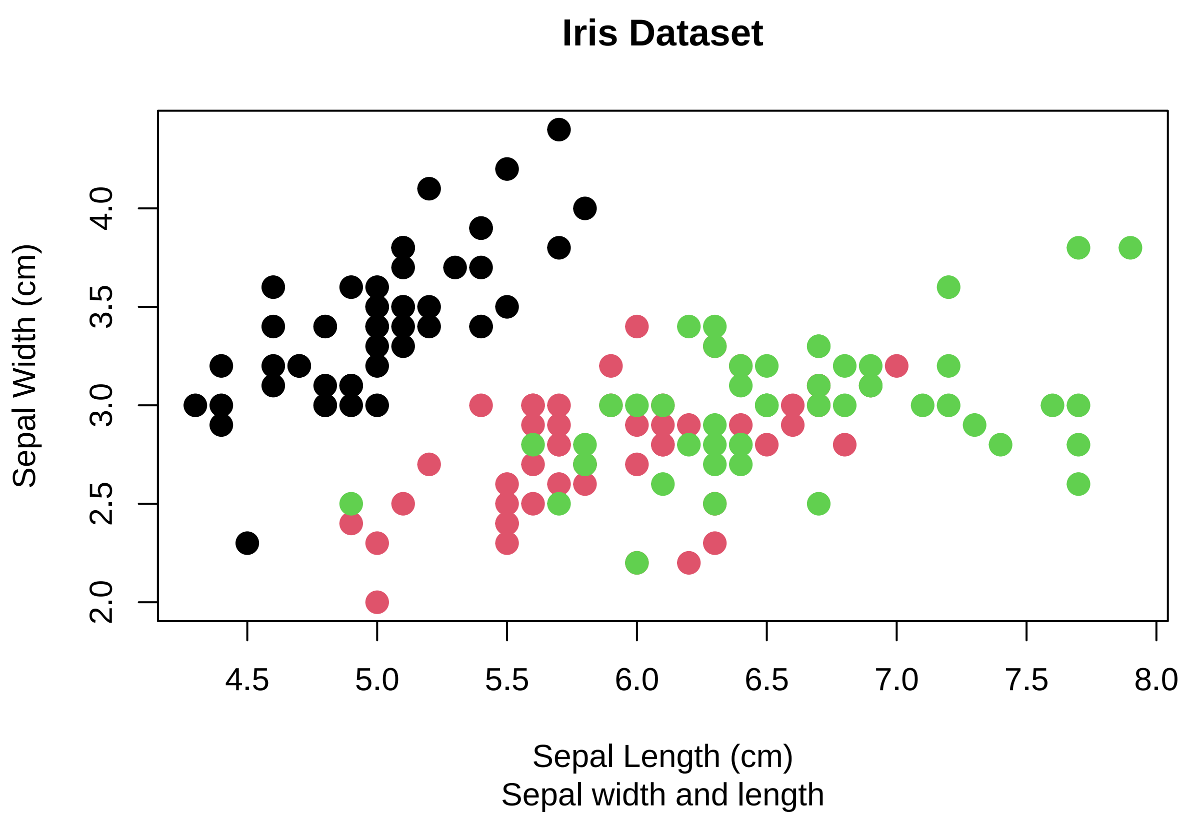
```r
print(tail(virginica_subset, 6))
```

```
##     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

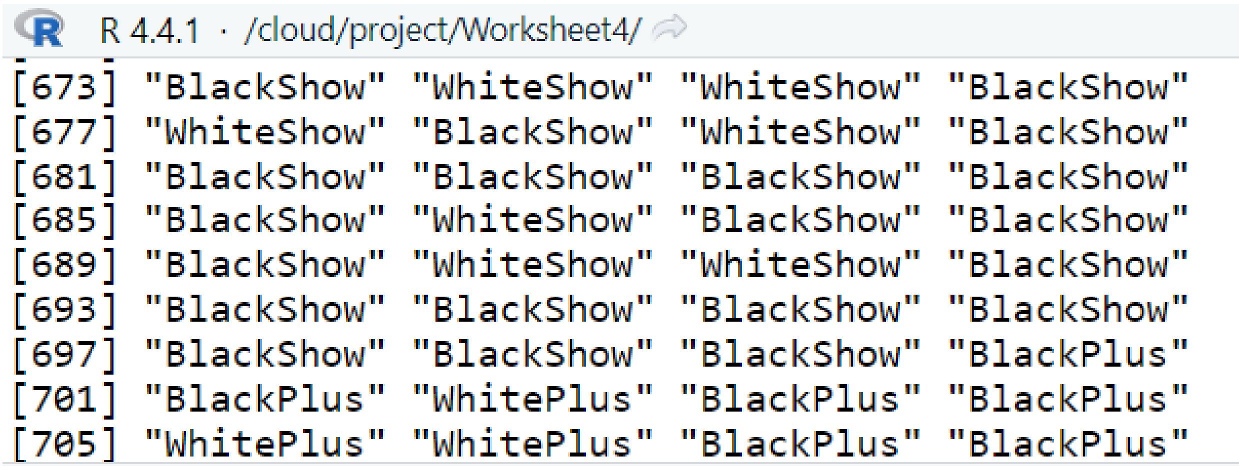
```
iris$Species <- as.factor(iris$Species)

plot(
  iris$Sepal.Length,
  iris$Sepal.Width,
  main = "Iris Dataset",
  sub = "Sepal width and length",
  xlab = "Sepal Length (cm)",
  ylab = "Sepal Width (cm)",
  pch = 21,
  col = iris$Species,
  bg = iris$Species,
  cex = 1.5
)
```



f The scatterplot depicts the relationship between Sepal.Length and Sepal.Width for three iris species: setosa, versicolor, and virginica. Setosa is clearly clustered in the lower left quadrant, with smaller dimensions, distinguishing it from the other species. Versicolor exhibits a broader range of sizes, overlapping with virginica, which generally has larger measurements. 7. a

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

alexa_data$variation <- gsub("Black  Dot", "BlackDot", alexa_data$variation)
alexa_data$variation <- gsub("Black  Plus", "BlackPlus", alexa_data$variation)
```

```r
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)

knitr::include_graphics("/cloud/project/Worksheet4/Screenshot 2024-11-02 153147.png")
```



b

```r
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
```

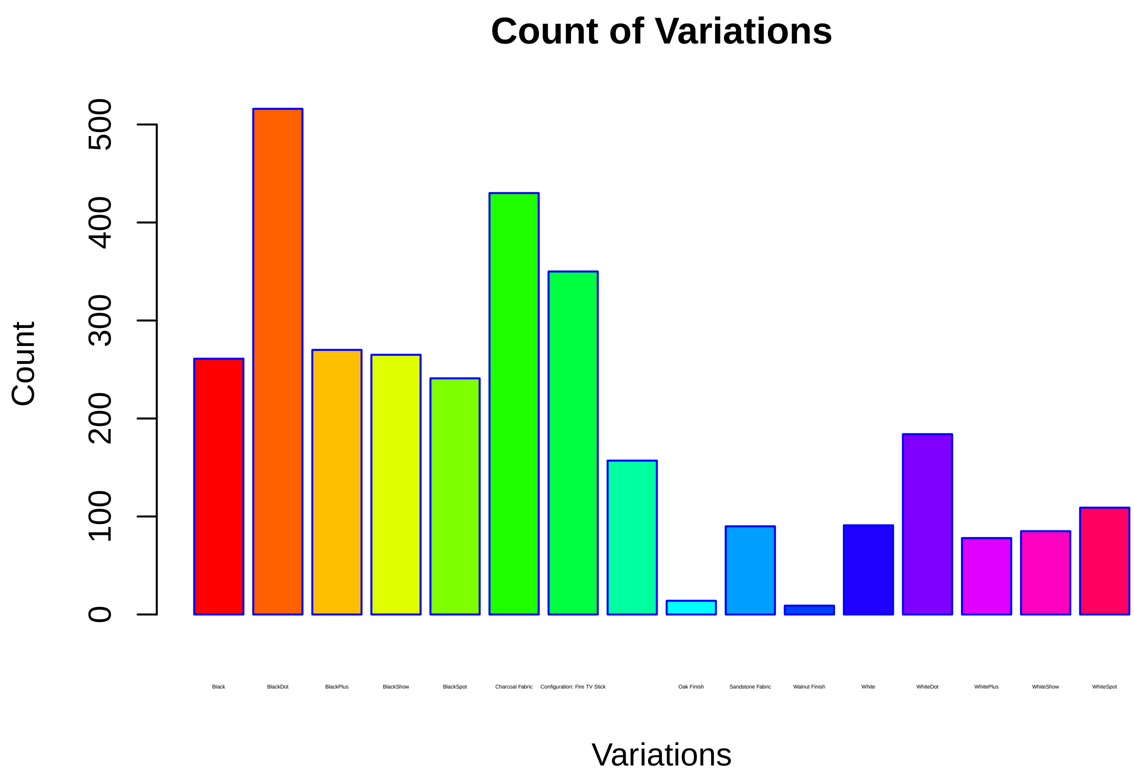
```r
variations.RData <- alexa_data %>%
  count(alexa_data$variation)

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

print(variations.RData)
```

```
## # A tibble: 16 x 2
##    `alexa_data$variation`           n
##    <chr>                        <int>
##  1 Black                          261
##  2 BlackDot                       516
##  3 BlackPlus                      270
##  4 BlackShow                      265
##  5 BlackSpot                      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 WhiteDot                       184
##  14 WhitePlus                       78
##  15 WhiteShow                       85
##  16 WhiteSpot                      109
```

c

```r
barplot(
  variations.RData$n,
  names.arg = variations.RData$`alexa_data$variation`,
  cex.names = 0.2,
  main = "Count of Variations",
  xlab = "Variations",
  ylab = "Count",
  col = rainbow(length(variations.RData$n)),
  border = "blue"
)
```



d

```r
bw_variations <- variations.RData %>%
  filter(grepl("^Black|^White", `alexa_data$variation`))

par(mfrow = c(1, 2))

barplot(
  bw_variations$n[bw_variations$`alexa_data$variation` %in% c("Black", "BlackPlus", "BlackShow", "BlackS
```

```
  names.arg = bw_variations$`alexa_data$variation`[bw_variations$`alexa_data$variation` %in% c("Black",
  cex.names = 0.2,
  main = "Black Variations",
  xlab = "Variations",
  ylab = "Count",
  col = "black",
  border = "blue"
)

barplot(
  bw_variations$n[bw_variations$`alexa_data$variation` %in% c("White", "WhitePlus", "WhiteShow", "White
  names.arg = bw_variations$`alexa_data$variation`[bw_variations$`alexa_data$variation` %in% c("White",
  cex.names = 0.2,
  main = "White Variations",
  xlab = "Variations",
  ylab = "Count",
  col = "white",
  border = "blue"
)
```