

# RWorksheet\_calvario#4b.Rmd

Jolien

2024-10-30

1. Using the for loop, create an R script that will display a 5x5 matrix as shown in Figure 1. It must contain vectorA = [1,2,3,4,5] and a 5 x 5 zero matrix.

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

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

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

zero\_matrix

```
##      [,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. Print the string “\*” using for() function.

```
n <- 5

for (i in 1:n) {
  output <- paste(rep("*", i), collapse = " ")
  cat(output, "\n")
}
```

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

3. Get an input from the user to print the Fibonacci sequence starting from the 1st input up to 500. Use repeat and break statements. Write the R Scripts and its output.

```
start_value <- as.numeric(readline(prompt = "Enter a starting number for the Fibonacci sequence: "))

a <- 0
b <- 1

cat("Fibonacci sequence starting from", start_value, "up to 500:\n")

repeat {
  next_fib <- a + b

  if (next_fib > 500) {
    break
  }

  if (next_fib >= start_value) {
    cat(next_fib, "\n")
  }
  a <- b
  b <- next_fib
}
```

4. Import the dataset as shown in Figure 1 you have created previously.

a. What is the R script for importing an excel or a csv file? Display the first 6 rows of the dataset? Show your codes and its result

```
data <- read.csv("shoe_size.csv")

head(data)
```

```
##   Shoe.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. Create a subset for gender(female and male). How many observations are there in Male? How about in Female? Write the R scripts and its output.

```

male_data <- subset(data, Gender == "Male")
female_data <- subset(data, Gender == "Female")

num_male <- nrow(male_data)
num_female <- nrow(female_data)

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

## Number of Male observations: 0
cat("Number of Female observations:", num_female, "\n")

## Number of Female observations: 0

```

c. Create a graph for the number of males and females for Household Data. Use `plot()`, chart type = `barplot`. Make sure to place title, legends, and colors. Write the R scripts and its result.

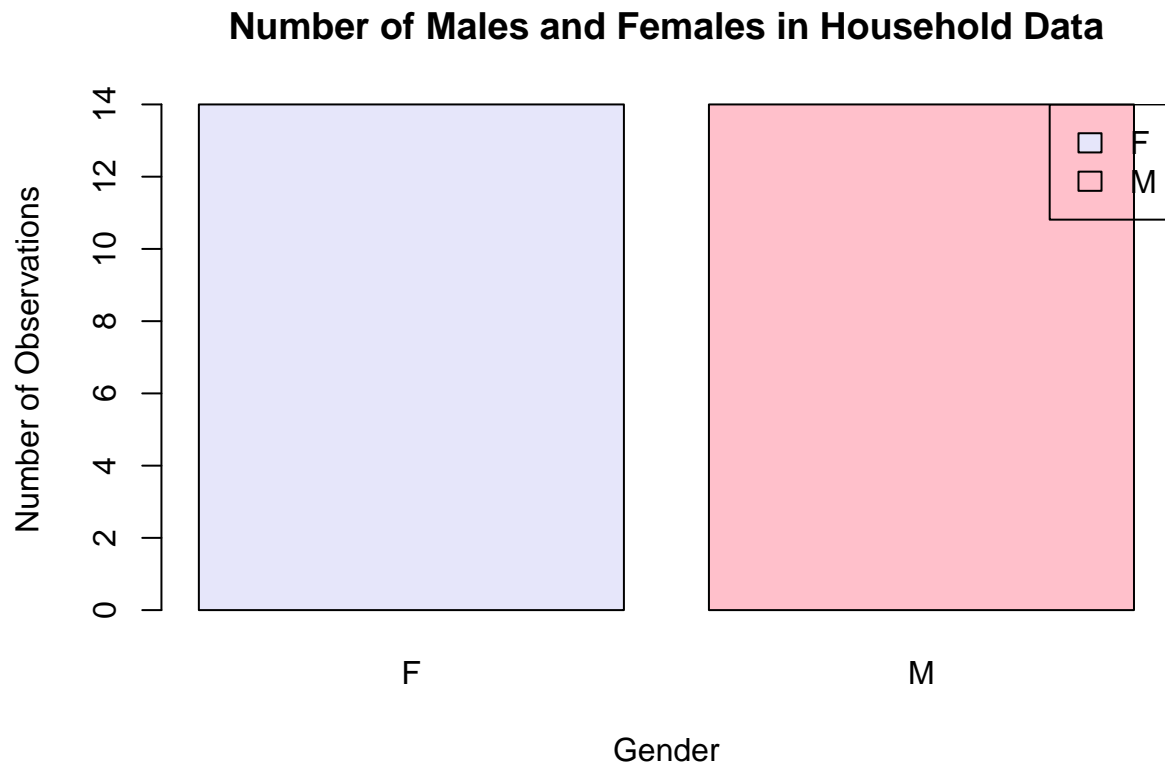
```

gender_counts <- table(data$Gender)

barplot(gender_counts,
        main = "Number of Males and Females in Household Data",
        xlab = "Gender",
        ylab = "Number of Observations",
        col = c("lavender", "pink"),
        beside = TRUE)

legend("topright",
      legend = names(gender_counts),
      fill = c("lavender", "pink"))

```



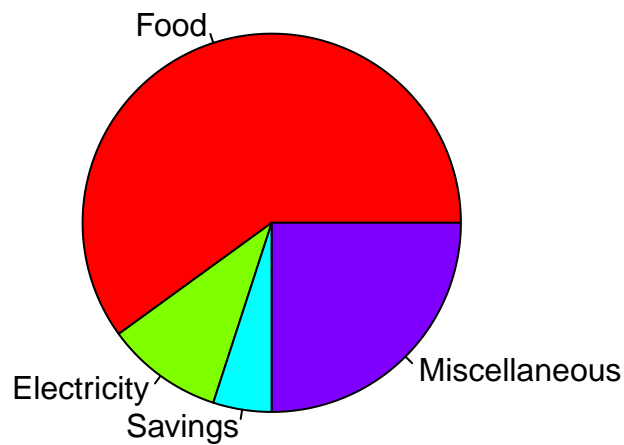
5. The monthly income of Dela Cruz family was spent on the following:

a.

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

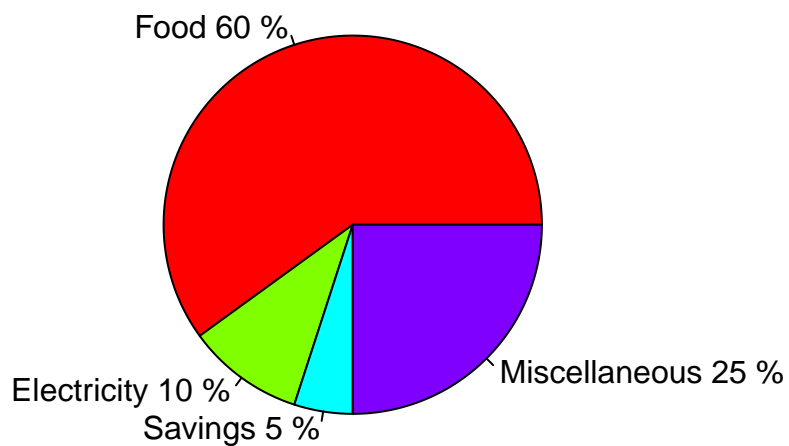
pie(expenses,
    labels = labels,
    main = "Monthly Income Distribution of Dela Cruz Family",
    col = rainbow(length(expenses)))
```

## Monthly Income Distribution of Dela Cruz Family



```
percentages <- round(expenses / sum(expenses) * 100, 1)
labels_with_percentages <- paste(labels, percentages, "%")
pie(expenses,
    labels = labels_with_percentages,
    main = "Monthly Income Distribution of Dela Cruz Family",
    col = rainbow(length(expenses)))
```

## Monthly Income Distribution of Dela Cruz Family



6. Use the iris dataset.

a.

```
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.

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

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

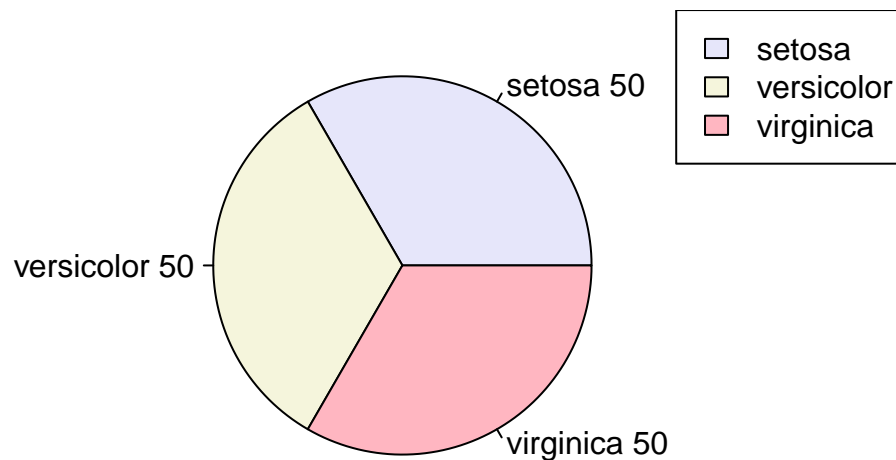
c.

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

pie(species_counts,
    main = "Species Distribution in Iris Dataset",
    col = c("lavender", "beige", "lightpink"),
    labels = paste(names(species_counts), species_counts))

legend("topright", legend = names(species_counts), fill = c("lavender", "beige", "lightpink"))
```

### Species Distribution in Iris Dataset



d.

```
setosa_subset <- iris[iris$Species == "setosa", ]
versicolor_subset <- iris[iris$Species == "versicolor", ]
virginica_subset <- iris[iris$Species == "virginica", ]

setosa_last_six <- tail(setosa_subset, 6)
```

```

versicolor_last_six <- tail(versicolor_subset, 6)
virginica_last_six <- tail(virginica_subset, 6)

print("Last six rows of Setosa:")

## [1] "Last six rows of Setosa:"
print(setosa_last_six)

##      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

print("Last six rows of Versicolor:")

## [1] "Last six rows of Versicolor:"
print(versicolor_last_six)

##      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

print("Last six rows of Virginica:")

## [1] "Last six rows of Virginica:"
print(virginica_last_six)

##      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.

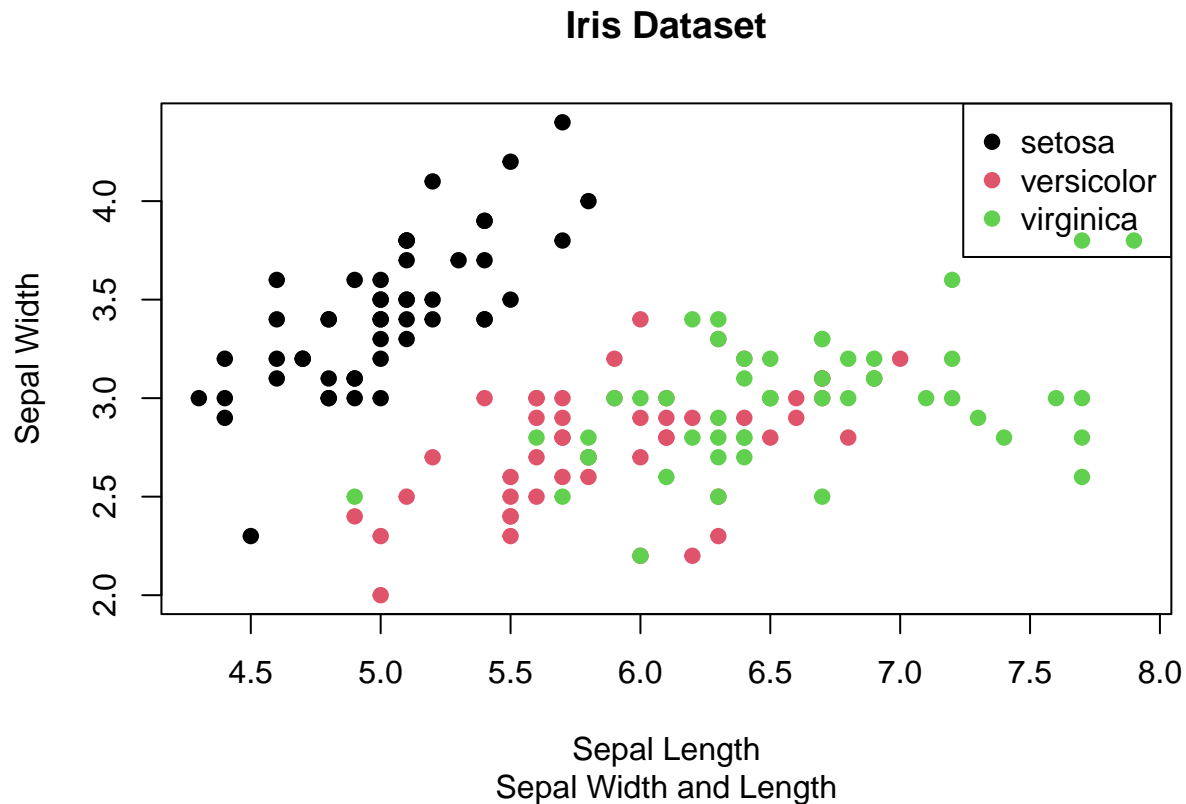
```

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

plot(iris$Sepal.Length, iris$Sepal.Width,
     col = iris$Species,
     pch = 19,
     main = "Iris Dataset",
     sub = "Sepal Width and Length",
     xlab = "Sepal Length",
     ylab = "Sepal Width")

```

```
legend("topright", legend = levels(iris$Species), col = 1:3, pch = 19)
```



f. Interpret the result.

The resulting plot visually displays the relationship between sepal length and sepal width for the different species of iris. Each species will appear in distinct colors, allowing for easy visual differentiation. This can help in understanding how sepal dimensions vary among the species. You may observe clustering of points by species, indicating differences in sepal dimensions that could be useful for classification or further analysis. For instance, Setosa typically has smaller sepals compared to the other two species.