RWorksheet_Subosa#4b.Rmd

Gian Adree Subosa

2024-10-30

#1. Using the for loop, create an R script that will display a 5x5 matrix

```
vectorA <- c(1, 2, 3, 4, 5)
matrix5x5 <- matrix(0, nrow = 5, ncol = 5)

for (i in 1:5) {
  for (j in 1:5) {
    matrix5x5[i, j] <- abs(vectorA[i] - vectorA[j])
  }
}

print(matrix5x5)
```

```
##      [,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

```
asteris <- 5

for(i in 1:asteris) {
  cat(rep("*", i), "\n")
}
```

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

#3. Fibonacci Sequence

```
val <- 25

a <- 0
b <- 1

if (val <= a) {
  cat(a, " ")
}
if (val <= b) {
  cat(b, " ")
```

```
}

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

  a <- b
  b <- next_number
}
```

```
## 34  55  89  144  233  377
```

#4. Import the dataset

```
#a.) Import the file. Display the first 6 rows of the dataset
my_data <- read.csv("shoe_height_gender_data.csv")
head(my_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)
m_sub <- subset(my_data, Gender == "M")
f_sub <- subset(my_data, Gender == "F")

num_Male <- nrow(m_sub)
num_Female <- nrow(f_sub)

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

```
## Number of Male observations:  13
```

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

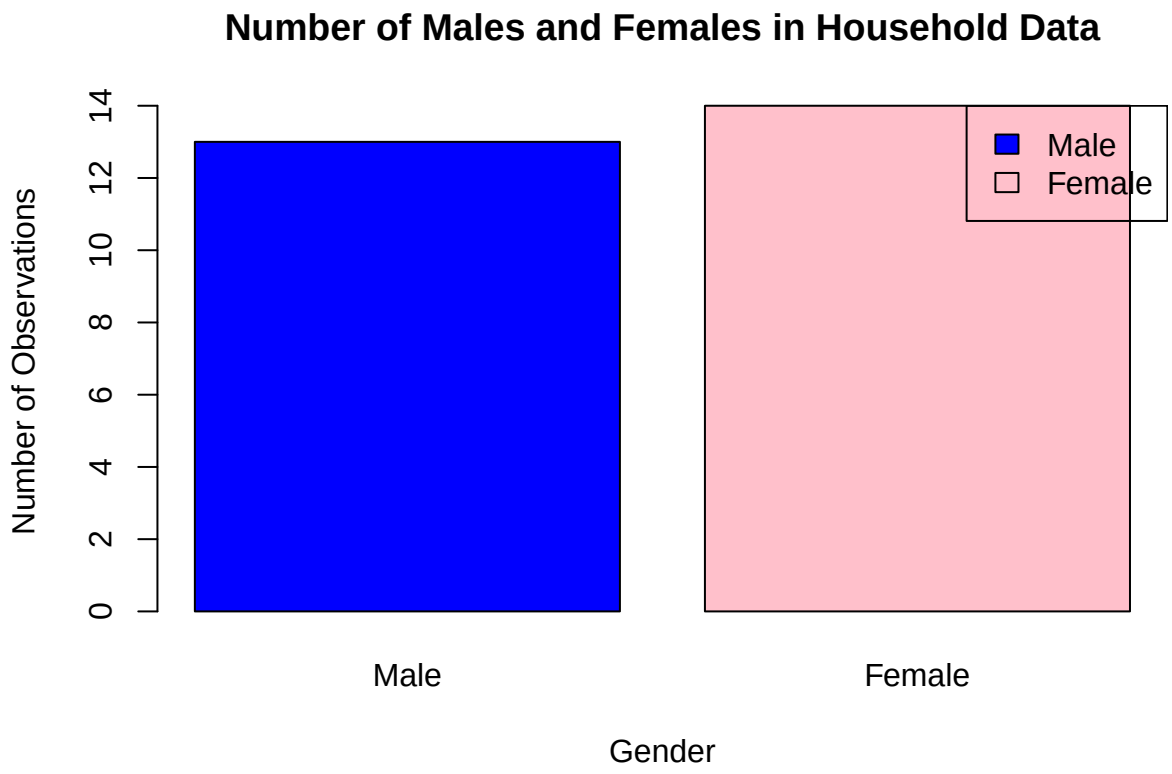
```
## Number of Female observations:  14
```

```
#c.) Create a graph for the number of males and females for Household Data
male_count <- nrow(subset(my_data, Gender == "M"))
female_count <- nrow(subset(my_data, Gender == "F"))

gender_count <- data.frame(
  Gender = c("Male", "Female"),
  Count = c(num_Male, num_Female)
)

barplot(
  height = gender_count$Count,
```

```
  names.arg = gender_count$Gender,
  col = c("blue", "pink"),
  main = "Number of Males and Females in Household Data",
  xlab = "Gender",
  ylab = "Number of Observations",
  legend.text = TRUE,
  beside = TRUE
)
legend("topright", legend = gender_count$Gender, fill = c("blue", "pink"))
```



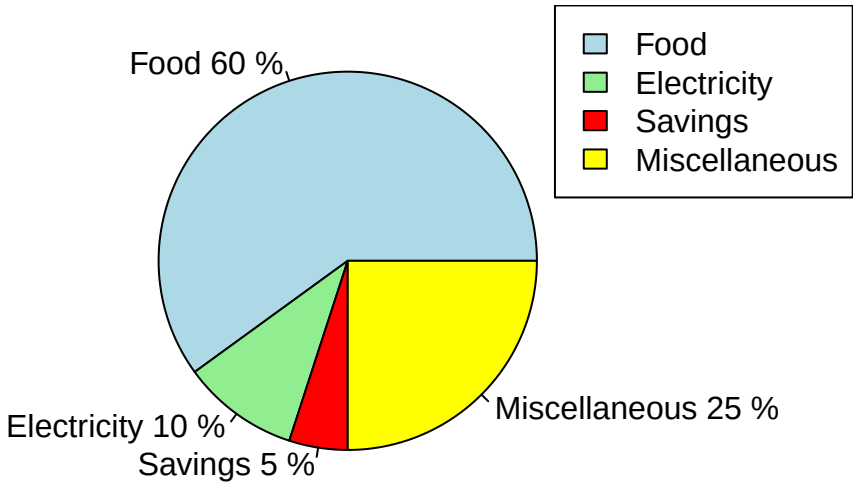
#5. Monthly income of DeLaCruz family

```
# Data for Dela Cruz family's monthly income distribution
expenses <- c(60, 10, 5, 25)
labels <- c("Food", "Electricity", "Savings", "Miscellaneous")

#a.) Create a pie chart that will include labels in percentage
percentages <- round(expenses / sum(expenses) * 100, 1)
labels_with_percentages <- paste(labels, percentages, "%", sep = " ")
colors <- c("lightblue", "lightgreen", "red", "yellow")

pie(
  expenses,
  labels = labels_with_percentages,
  col = colors,
  main = "Monthly Income Distribution of Dela Cruz Family"
)
legend("topright", legend = labels, fill = colors)
```

Monthly Income Distribution of Dela Cruz Family



#6. The iris dataset

```
#a.) Check for the structure of the dataset using the str() function
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 ...
```

```
#The dataset shows its dimensions which are 150 observations and 5 variables.
```

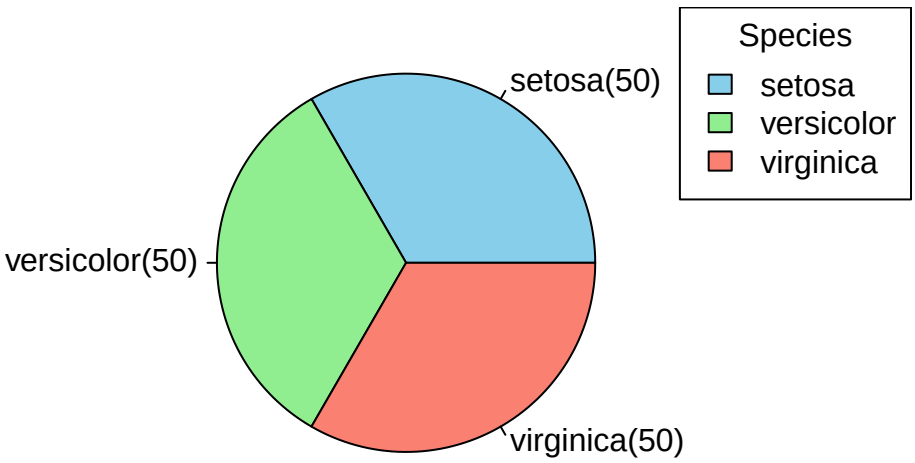
```
#b.) Create an R object that will contain the mean of the sepal.length, sepal.width,petal.length,and pe
mean_values <- colMeans(iris[, 1:4])
print(mean_values)
```

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

```
#c.) Create a pie chart for the Species distribution
data(iris)
species_counts <- table(iris$Species)
colors <- c("skyblue", "lightgreen", "salmon")

pie(species_counts,
    main = "Species Distribution in Iris Dataset",
    col = colors,
    labels = paste(names(species_counts), "(", species_counts, ")", sep = "")
)

legend("topright",
       legend = names(species_counts),
       fill = colors,
       title = "Species")
```




```
#d.) Subset the species into setosa, versicolor, and virginica
data(iris)
setosa <- subset(iris, Species == "setosa")
versicolor <- subset(iris, Species == "versicolor")
virginica <- subset(iris, Species == "virginica")

tail(setosa, 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
```

```
tail(versicolor, 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
```

```
tail(virginica, 6)
```

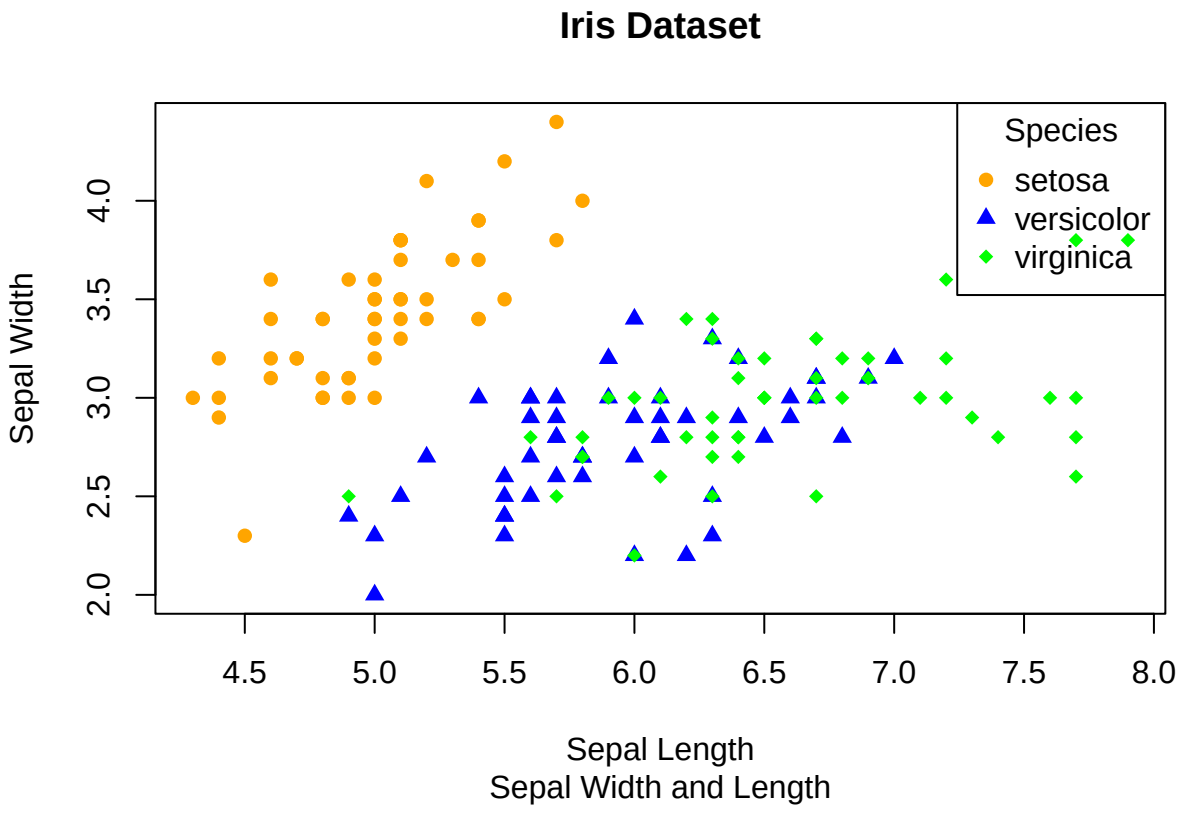
```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 145          6.7         3.3          5.7         2.5 virginica
## 146          6.7         3.0          5.2         2.3 virginica
## 147          6.3         2.5          5.0         1.9 virginica
## 148          6.5         3.0          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.) Create a scatterplot
data(iris)
```

```r
iris$Species <- as.factor(iris$Species)
species_colors <- c("setosa" = "orange", "versicolor" = "blue", "virginica" = "green")
species_symbols <- c("setosa" = 16, "versicolor" = 17, "virginica" = 18)

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

legend("topright",
       legend = levels(iris$Species),
       col = c("orange", "blue", "green"),
       pch = c(16, 17, 18),
       title = "Species")
```




```r
#f.) Interpret the result

#This scatterplot clearly illustrates the differences in sepal diameters between the three iris flower
```