RWorksheet_asenjo#4b

Samuel Asenjo

2024-10-28

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 - vectorA[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.

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

for(i in v){
  cat(rep("*", i),"\n")
}
```

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

3.

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

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

```
a <- 0
b <- 1
if (!is.na(start_num) < 0) {
1
```

```r
cat("Please enter a non-negative starting number.\n")
} else {
repeat {
if (!is.na(start_num) && a >= start_num) {
cat(a, "\n")
}
temp <- a + b
a <- b
b <- temp
if (!is.na(start_num) && a > 500) {
break
}
}
}
```

```
## Please enter a non-negative starting number.
```

4. a.

```r
Shoesizes <-read.csv("/cloud/project/Worksheet 4/shoesizes.csv")

head(Shoesizes)
```

```
##   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.

```r
male_data <- subset(Shoesizes, Gender == "M")
female_data <- subset(Shoesizes, Gender == "F")

num_males <- nrow(male_data)
num_females <- nrow(female_data)

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

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

```r
cat("Number of observations for Female: ", num_females, "\n")
```

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

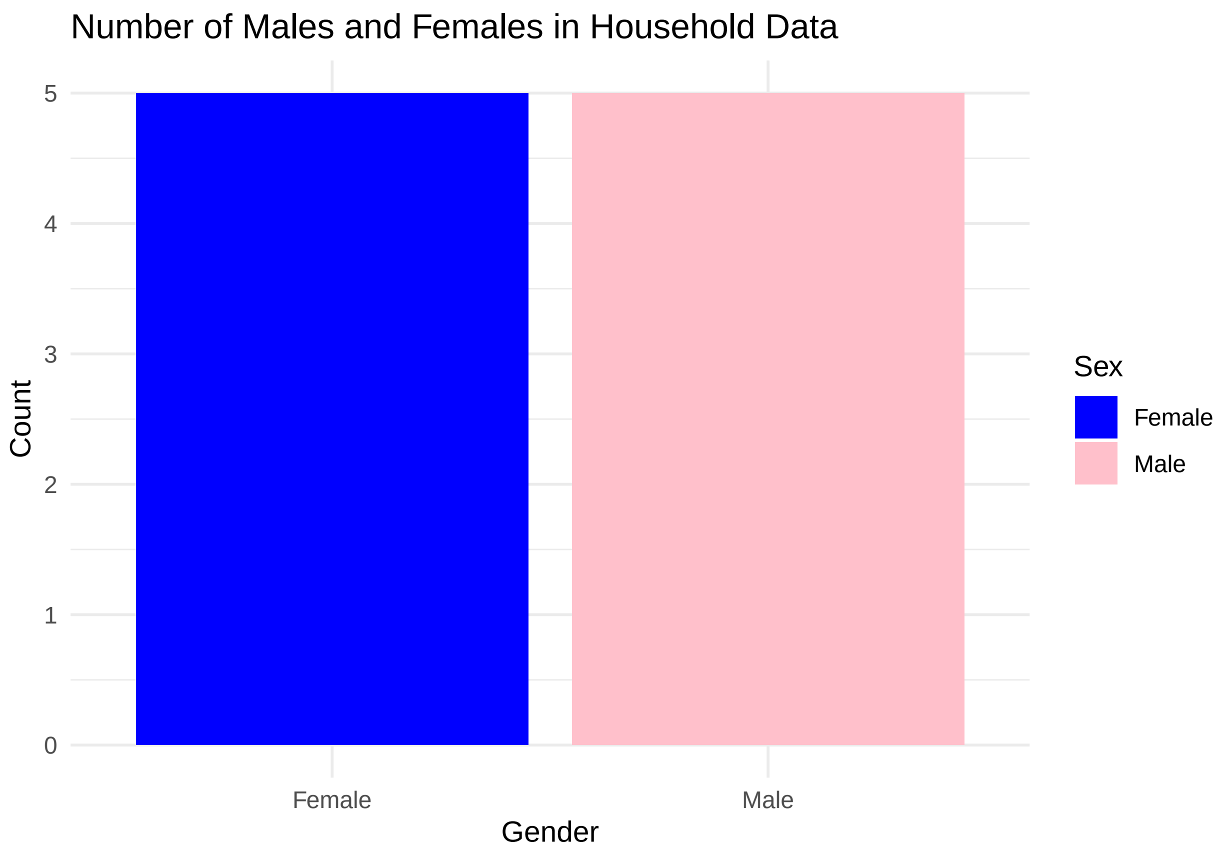
c.

```r
library(ggplot2)

HouseHoldData <-read.csv("/cloud/project/Worksheet 4/HouseholdData.csv")


ggplot(HouseHoldData, aes(x=Sex, fill=Sex)) +
  geom_bar() +
  ggtitle("Number of Males and Females in Household Data") +
  xlab("Gender") +
```

```r
  ylab("Count") +
  scale_fill_manual(values = c("blue", "pink")) +
  theme_minimal()
```



5.

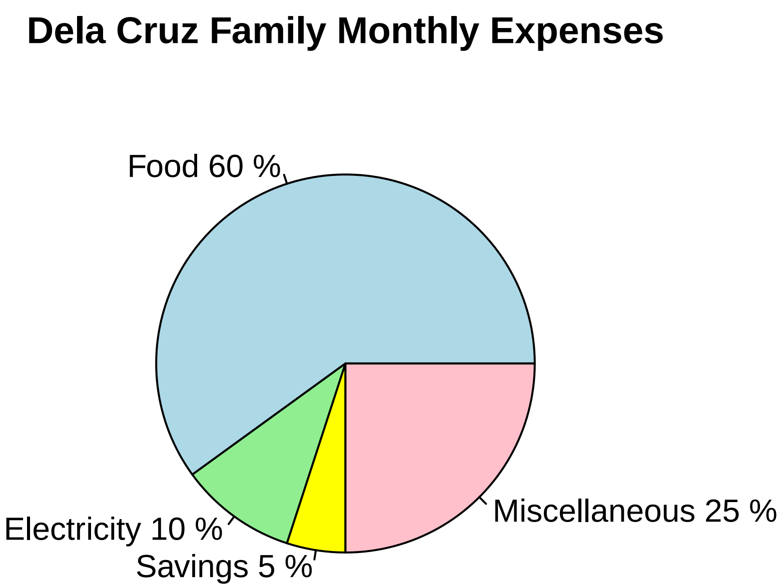
```r
expenses <- c(Food = 60, Electricity = 10, Savings = 5, Miscellaneous = 25)

percentages <- round(100 * expenses / sum(expenses), 1)

labels <- paste(names(expenses), percentages, "%")

colors <- c("lightblue", "lightgreen", "yellow", "pink")

pie(expenses, labels = labels, col = colors, main = "Dela Cruz Family Monthly Expenses")
```



6.

```
data(iris)
```

a. The output shows number of variables and objects, and rows and columns

```
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.

```
means <- colMeans(iris[, c("Sepal.Length", "Sepal.Width", "Petal.Length", "Petal.Width")])

means
```

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

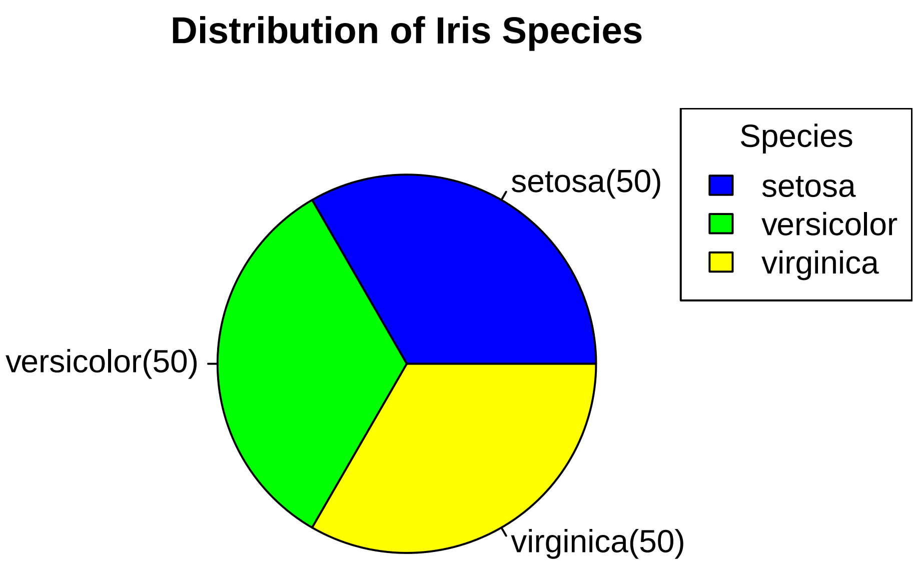
c.

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

colors <- c("blue", "green", "yellow")

pie(species_distribution,
    main = "Distribution of Iris Species",
    col = colors,
    labels = paste(names(species_distribution), "(", species_distribution, ")", sep=""))

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



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
```

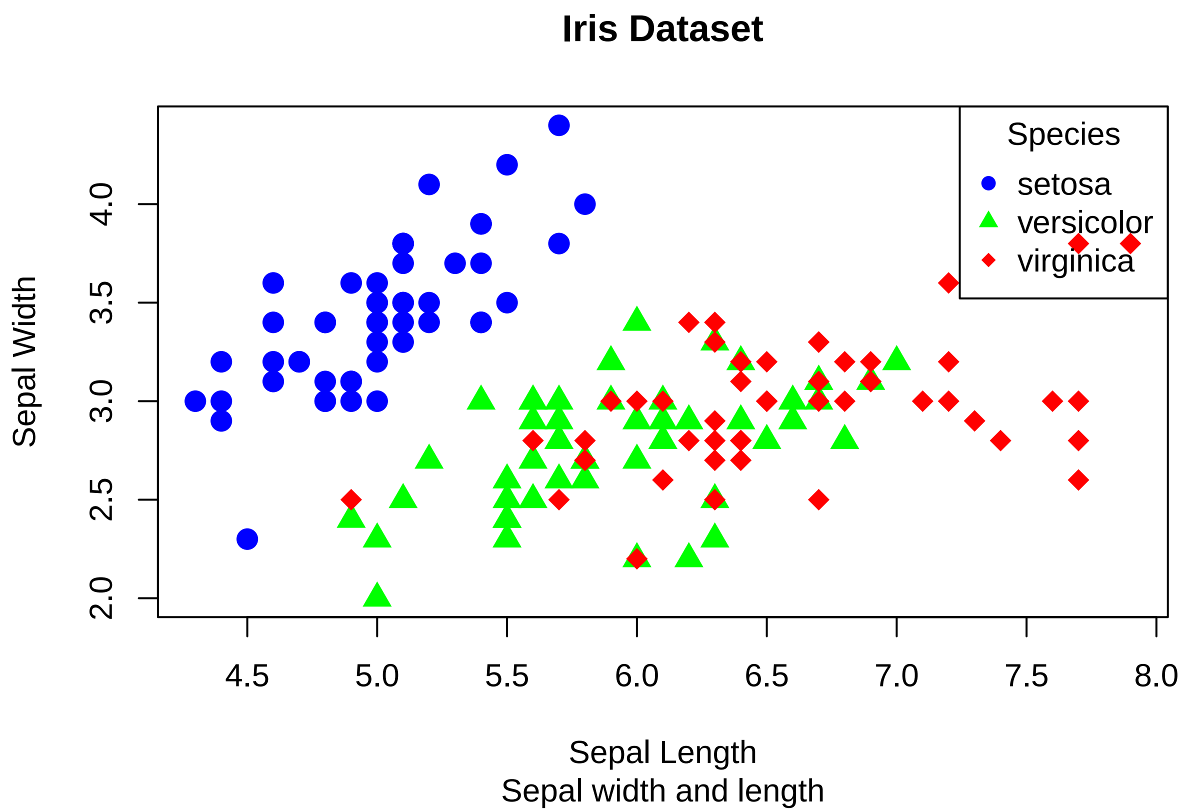
e.

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

colors <- c("setosa" = "blue", "versicolor" = "green", "virginica" = "red")
symbols <- c("setosa" = 16, "versicolor" = 17, "virginica" = 18)

plot(iris$Sepal.Length, iris$Sepal.Width,
     col = colors[iris$Species],
     pch = symbols[iris$Species],
     xlab = "Sepal Length",
     ylab = "Sepal Width",
     main = "Iris Dataset",
     sub = "Sepal width and length",
     cex = 1.5)

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



f. The scatterplot shows clear separation between setosa and the other two species based on Sepal Length and Sepal Width. Setosa has distinctively shorter and wider sepals, forming a separate cluster. Versicolor and virginica overlap more, particularly in sepal width, but virginica tends to have longer sepals. Overall, there is a slight negative correlation, where longer sepals tend to be narrower, especially in virginica.