

RWorksheet_Aguirre#4b

Ryza Faith Aguirre

2024-10-28

Using Loop Function 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] <-ifelse(j>=i, vectorA[abs(j-i)+1],0)
  }
}

matrixA
```

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

2.

```
for (i in 1:5) {
  for (j in 1:i) {
    cat("*")
  }
  cat("\n")
}
```

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

3.

```
start_pos <- as.integer(readline(prompt = "Enter the starting position: "))
```

```
## Enter the starting position:
```

```

fib1 <- 0
fib2 <- 1
count <- 1
fib_sequence <- c()
repeat {
  fib_next <- fib1 + fib2

  if (!is.na(start_pos) && count >= start_pos && fib_next <= 500) {
    fib_sequence <- c(fib_sequence, fib_next)
  }

  fib1 <- fib2
  fib2 <- fib_next
  count <- count + 1

  if (fib_next > 500) break
}

print(fib_sequence)

```

NULL

4a.

```

library(readr)
shoe_Sizes <- read_csv("~/DataScience/CS101/worksheet 4b/Shoe Sizes.csv")

head(shoe_Sizes)

```

```

##   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

```

4b.

```

male <- subset(shoe_Sizes, Gender == "M")
female <- subset(shoe_Sizes, Gender == "F")
num_males <- nrow(male)
num_females <- nrow(female)

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

```

Number of Male observations: 14

```

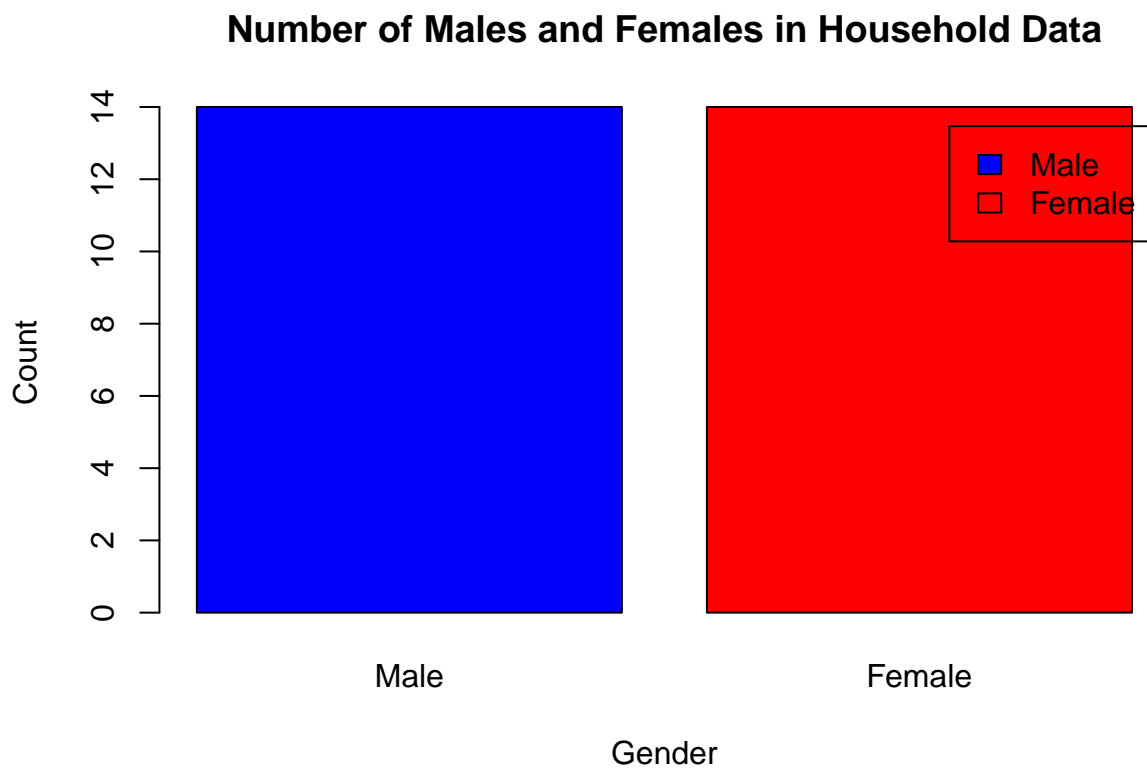
cat("Number of Female observations:", num_females, "\n")

```

Number of Female observations: 14

4c.

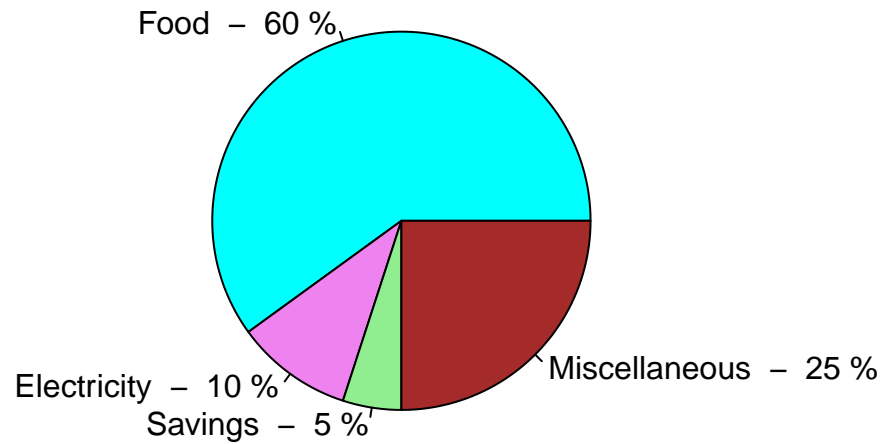
```
gender_counts <- c(num_males, num_females)
names(gender_counts) <- c("Male", "Female")
barplot(gender_counts,
        main = "Number of Males and Females in Household Data", xlab = "Gender", ylab = "Count", col = c("Male", "Female"))
```



5.

```
expenses <- c(Food = 60, Electricity = 10, Savings = 5, Miscellaneous = 25)
colors <- c("cyan", "violet", "lightgreen", "brown")
pie(expenses, labels = paste(names(expenses), " - ", round((expenses / sum(expenses)) * 100, 1), "%"), col = colors)
```

Dela Cruz Family Monthly Income Distribution



6a.

```
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 ...
```

6b.

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

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

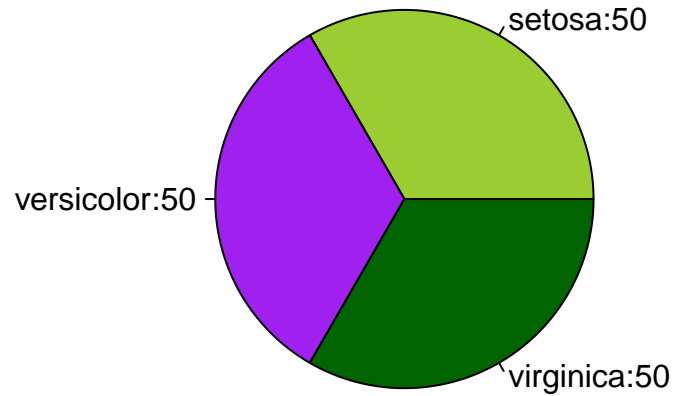
6c.

```
species_count <- table(iris$Species)
species_labels <- paste(names(species_count), species_count, sep = ":")
```

```
color <-c("yellowgreen", "purple", "darkgreen")

pie(species_count, labels= species_labels, col= color, main="Species Distribution in Iris DataSet")
```

Species Distribution in Iris DataSet



6d.

```
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
```

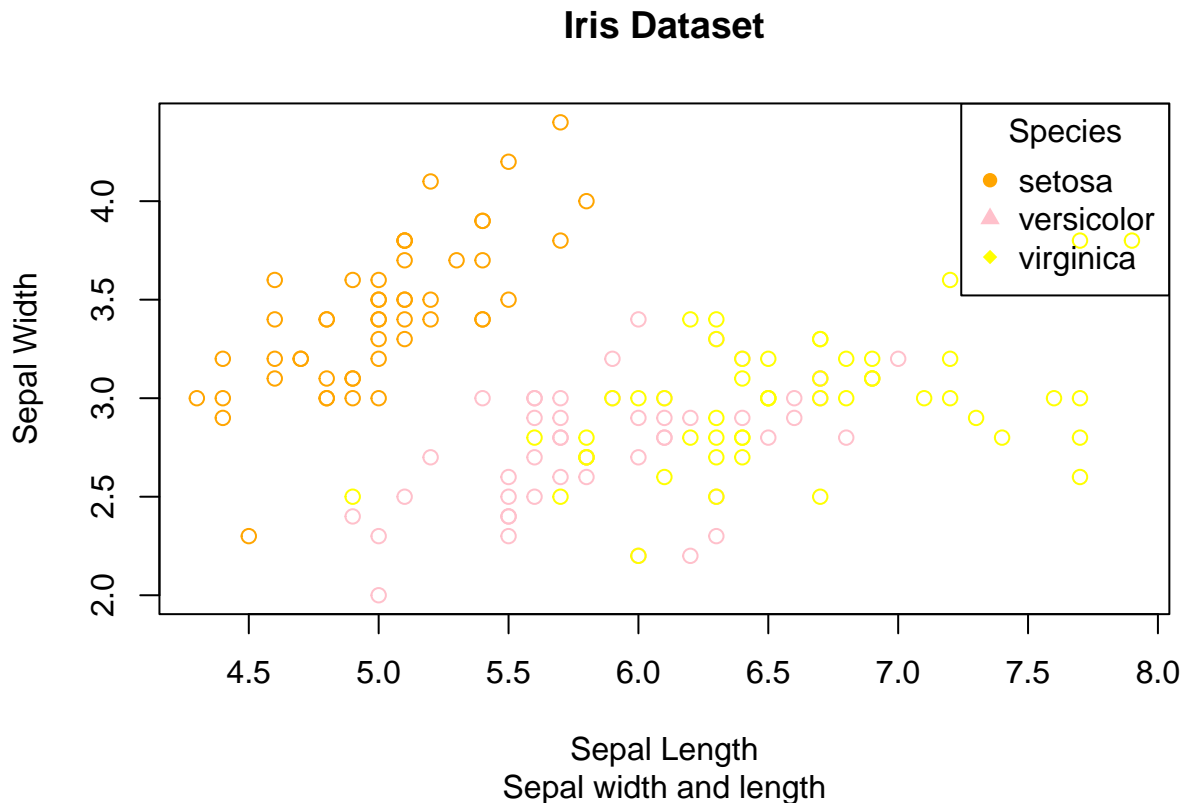
6e.

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

```
Species_color <- c("setosa"="orange","versicolor"="pink", "virginica"= "yellow")
species_pch <- c("setosa"= 16, "versicolor"= 17, "virginica"= 18)
```

```
plot(iris$Sepal.Length, iris$Sepal.Width, col= Species_color[iris$Species],
     xlab = "Sepal Length",
     ylab= "Sepal Width", main ="Iris Dataset", sub= "Sepal width and length")
```

```
legend ("topright", legend=levels(iris$Species), col= Species_color, pch = species_pch, title = "Species")
```



6f. Interpretation of the Scatterplot Result

The scatterplot shows the relationship between Sepal.Length and Sepal.Width for the three species. Setosa points (in orange) tend to have shorter sepal lengths and widths. Versicolor points (in pink) are more spread out and overlap slightly with virginica. Virginica points (in yellow) have larger sepal lengths and widths, distinguishing them from setosa and overlapping with some versicolor points

7.

```
library(readxl)
data <- read_excel("~/DataScience/CS101/worksheet 4b/alexa_file.xlsx")

data$variation <- gsub("Black Dot", "Black Dot", data$variation)
data$variation <- gsub("White Plus", "White Plus", data$variation)

head(data$variation)
```

```
## [1] "Charcoal Fabric"      "Charcoal Fabric"      "Walnut Finish"
## [4] "Charcoal Fabric"      "Charcoal Fabric"      "Heather Gray Fabric"
```

```
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_count <- data %>%
  count(variation)
save(variation_count, file = "variations.RData")
print(variation_count)
```

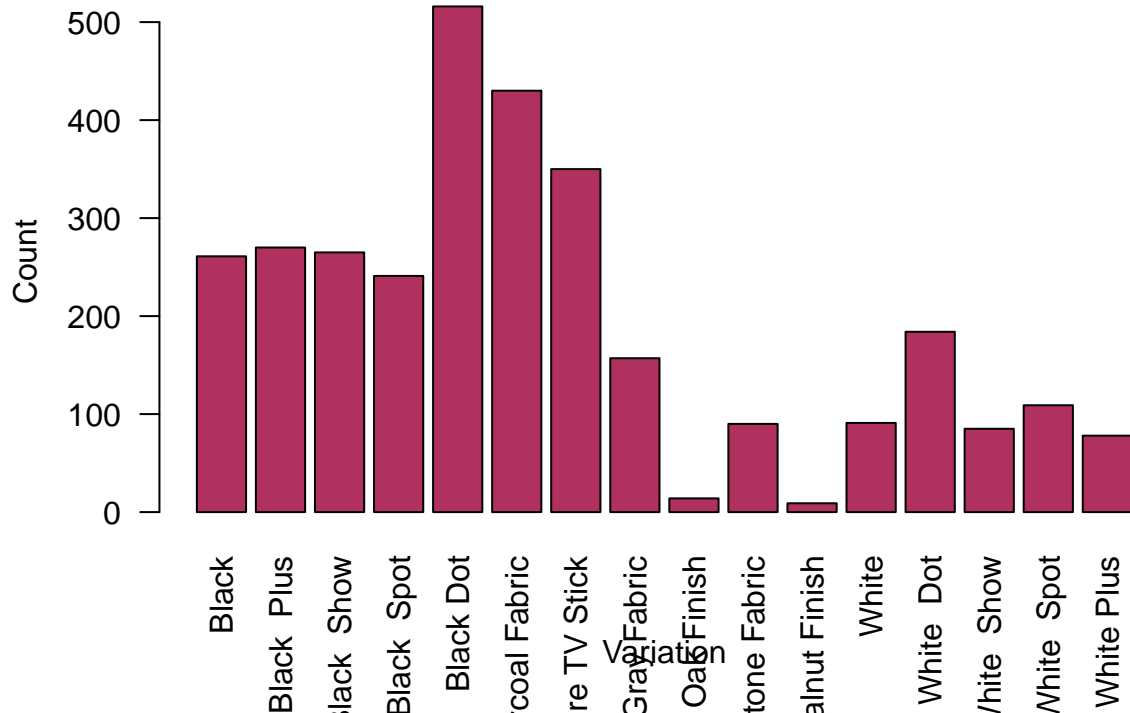
```
## # A tibble: 16 x 2
##   variation      n
##   <chr>      <int>
## 1 Black      261
## 2 Black Plus 270
## 3 Black Show 265
## 4 Black Spot 241
## 5 Black Dot  516
## 6 Charcoal Fabric 430
## 7 Configuration: Fire TV Stick 350
## 8 Heather Gray Fabric 157
## 9 Oak Finish   14
## 10 Sandstone Fabric 90
## 11 Walnut Finish 9
## 12 White       91
## 13 White Dot   184
## 14 White Show  85
## 15 White Spot  109
## 16 White Plus  78
```

7b.

```
library(dplyr)
load("variations.RData")

barplot(variation_count$n, names.arg = variation_count$variation, col = "maroon",
  main = "Variation Count", xlab = "Variation", ylab = "Count", las = 2)
```


Variation Count



```
library(readxl)
library(dplyr)

alexa_file <- "~/DataScience/CS101/worksheet 4b/alexa_file.xlsx"
data <- read_excel(alexa_file)

data$variation <- gsub("Black Dot", "Black Dot", data$variation)
data$variation <- gsub("Black Plus", "Black Plus", data$variation)
data$variation <- gsub("Black Show", "Black Show", data$variation)
data$variation <- gsub("Black Spot", "Black Spot", data$variation)
data$variation <- gsub("White Dot", "White Dot", data$variation)
data$variation <- gsub("White Plus", "White Plus", data$variation)
data$variation <- gsub("White Show", "White Show", data$variation)
data$variation <- gsub("White Spot", "White Spot", data$variation)

black_white_data <- data %>%
  filter(grepl("Black|White", variation))

variation_count <- black_white_data %>%
  count(variation)

black_variations <- variation_count %>% filter(grepl("Black", variation))
white_variations <- variation_count %>% filter(grepl("White", variation))

black_counts <- setNames(black_variations$n, gsub("Black ", "", black_variations$variation))
white_counts <- setNames(white_variations$n, gsub("White ", "", white_variations$variation))
```

```

common_variations <- c("Dot", "Plus", "Show", "Spot")

black_counts <- black_counts[common_variations]
black_counts[is.na(black_counts)] <- 0

white_counts <- white_counts[common_variations]
white_counts[is.na(white_counts)] <- 0

barplot_matrix <- rbind(black_counts, white_counts)

barplot(barplot_matrix, beside = TRUE, col = c("purple", "blue"),
        main = "Count of Black and White Variations",
        xlab = "Variation Type", ylab = "Count",
        legend = rownames(barplot_matrix), args.legend = list(x = "topright"))

```

