

# RWorksheet\_Sobusa#4b.Rmd

Nexon Sobusa

2024-10-28

```
# 1. Creating a 5x5 matrix with vectorA = [1,2,3,4,5] and a 5x5 zero 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[j] - i)
  }
}
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. Printing "*" using a for loop
for (i in 1:5) {
  cat(rep("*", i), "\n")
}
```

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

```
# 3. Fibonacci sequence up to 500 using repeat and break statements
```

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

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

```
if (is.na(n) || n < 0) {
  cat("Please enter a valid positive integer.\n")
} else {
  a <- 0
  b <- 1
  repeat {
    if (a >= n) {
      cat(a, " ")
    }
    temp <- a + b
    a <- b
    b <- temp
  }
}
```

```

    if (a > 500) break
  }
}

```

```
## Please enter a valid positive integer.
```

```
# 4. Importing, subsetting, and plotting gender data
```

```
library(readxl)
```

```
dataset <- read.csv("DATA.csv")
```

```
male_subset <- subset(dataset, Gender == "M")
```

```
female_subset <- subset(dataset, Gender == "F")
```

```
cat("Number of Males:", nrow(male_subset), "\n")
```

```
## Number of Males: 14
```

```
cat("Number of Females:", nrow(female_subset), "\n")
```

```
## Number of Females: 14
```

```
# Plotting the gender data
```

```
gender_data <- read.csv("DATA.csv")
```

```
str(gender_data)
```

```
## 'data.frame': 28 obs. of 3 variables:
```

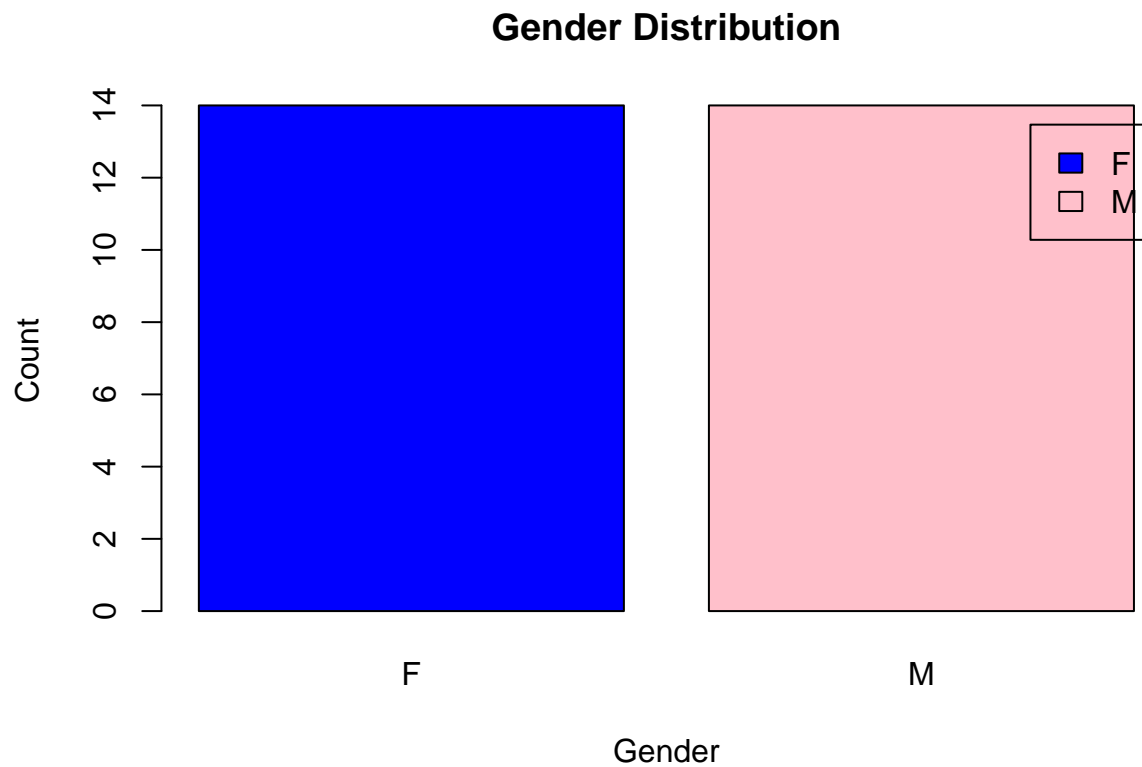
```
## $ Shoe.size: num 6.5 9 8.5 8.5 10.5 7 9.5 9 13 7.5 ...
```

```
## $ Height : num 66 68 64.5 65 70 64 70 71 72 64 ...
```

```
## $ Gender : chr "F" "F" "F" "F" ...
```

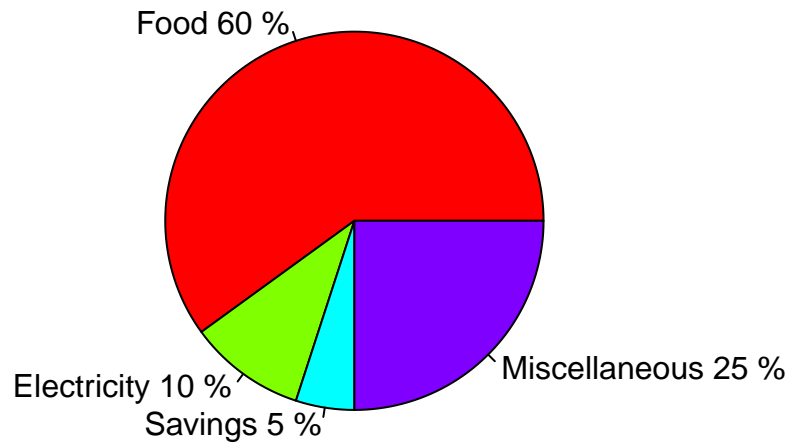
```
gender_counts <- table(gender_data$Gender)
```

```
barplot(gender_counts, main = "Gender Distribution", col = c("blue", "pink"),
        legend = rownames(gender_counts), xlab = "Gender", ylab = "Count")
```



```
# 5. Pie chart for Dela Cruz family expenses
expenses <- c(60, 10, 5, 25)
labels <- c("Food", "Electricity", "Savings", "Miscellaneous")
pie(expenses, labels = paste(labels, round(expenses / sum(expenses) * 100), "%"),
    col = rainbow(length(expenses)), main = "Dela Cruz Family Monthly Expenses")
```

## Dela Cruz Family Monthly Expenses



*# 6. Using the iris dataset*

```
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 ...
```

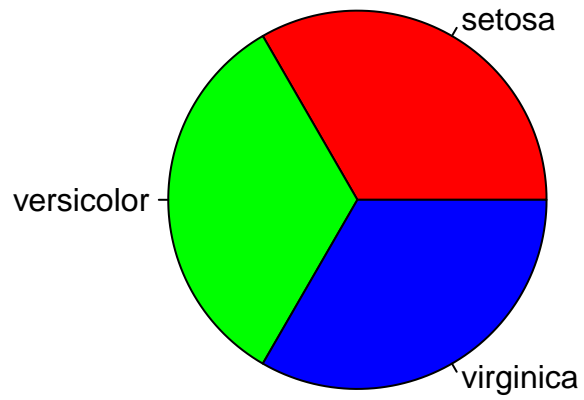
```
means <- colMeans(iris[, 1:4])
print(means)
```

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

*# Pie chart for species distribution*

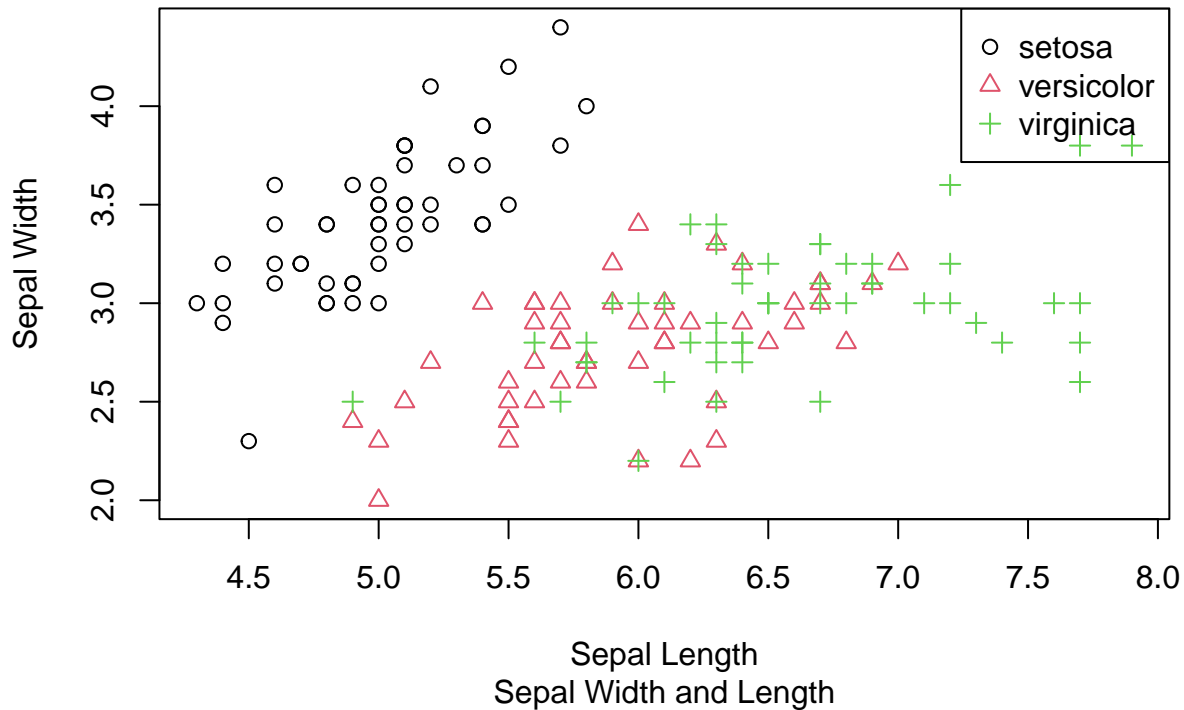
```
species_counts <- table(iris$Species)
pie(species_counts, main = "Species Distribution", col = rainbow(length(species_counts)),
    labels = names(species_counts))
```

## Species Distribution



```
# Scatterplot of sepal length vs. sepal width by species
plot(iris$Sepal.Length, iris$Sepal.Width, col = iris$Species,
     pch = as.numeric(iris$Species), main = "Iris Dataset",
     sub = "Sepal Width and Length", xlab = "Sepal Length", ylab = "Sepal Width")
legend("topright", legend = levels(iris$Species), col = 1:3, pch = 1:3)
```

## Iris Dataset



*# 7. Cleaning and renaming variants in the Alexa data*

```
library(readxl)
```

```
alex_data <- read_excel("alex_data.xlsx")
print(colnames(alex_data))
```

```
## [1] "rating"      "date"        "variation"    "verified_reviews"
```

```
## [5] "feedback"
```

```
str(alex_data)
```

```
## tibble [3,150 x 5] (S3: tbl_df/tbl/data.frame)
```

```
## $ rating      : num [1:3150] 5 5 4 5 5 5 3 5 5 5 ...
```

```
## $ date        : POSIXct[1:3150], format: "2018-07-31" "2018-07-31" ...
```

```
## $ variation   : chr [1:3150] "Charcoal Fabric" "Charcoal Fabric" "Walnut Finish" "Charcoal Fabr"
```

```
## $ verified_reviews: chr [1:3150] "Love my Echo!" "Loved it!" "Sometimes while playing a game, you c"
```

```
## $ feedback    : num [1:3150] 1 1 1 1 1 1 1 1 1 1 ...
```

```
head(alex_data)
```

```
## # A tibble: 6 x 5
```

```
##   rating date          variation      verified_reviews      feedback
##   <dbl> <dtm>          <chr>          <chr>          <dbl>
## 1     5 2018-07-31 00:00:00 Charcoal Fabric Love my Echo!          1
## 2     5 2018-07-31 00:00:00 Charcoal Fabric Loved it!              1
## 3     4 2018-07-31 00:00:00 Walnut Finish  Sometimes while playi~ 1
## 4     5 2018-07-31 00:00:00 Charcoal Fabric I have had a lot of f~ 1
## 5     5 2018-07-31 00:00:00 Charcoal Fabric Music                  1
```

## 6 5 2018-07-31 00:00:00 Heather Gray Fabric I received the echo a~

1

```
if ("Variant" %in% colnames(alexa_data)) {  
  unique_variants <- unique(alexa_data$Variant)  
  print(unique_variants)  
  
  alexa_data$Variant <- gsub("Black Dot", "BlackDot", alexa_data$Variant)  
  alexa_data$Variant <- gsub("Black Plus", "BlackPlus", alexa_data$Variant)  
  alexa_data$Variant <- gsub("Black Show", "BlackShow", alexa_data$Variant)  
  alexa_data$Variant <- gsub("Black Spot", "BlackSpot", alexa_data$Variant)  
  alexa_data$Variant <- gsub("White Dot", "WhiteDot", alexa_data$Variant)  
  alexa_data$Variant <- gsub("White Plus", "WhitePlus", alexa_data$Variant)  
  alexa_data$Variant <- gsub("White Show", "WhiteShow", alexa_data$Variant)  
  alexa_data$Variant <- gsub("White Spot", "WhiteSpot", alexa_data$Variant)  
  
  cleaned_snippet <- head(alexa_data)  
  print(cleaned_snippet)  
} else {  
  warning("Column 'Variant' does not exist in the data frame.")  
}
```

## Warning: Column 'Variant' does not exist in the data frame.