

RWorksheet_#4BGanon

Ganon, Andrew Vinz C.

r Sys.Date()

#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(vectorA[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.

```
for (i in 1:5) {
  cat(rep(" ", 1 * i - 0), "\n")
}
```

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

3.

```
user_input <- 41
num2 <- 0
repeat {
  sum <- user_input + num2
  if (sum > 500) {
    break
  }
  cat(sum, " ")
  user_input <- num2
  num2 <- sum
}
```

```
## 41 41 82 123 205 328
```

```
cat("\n")
```

4.

```
housedata <- read.csv("HouseData.csv")  
housedata
```

```
##      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  
## 7         9.5   70.0      F  
## 8         9.0   71.0      F  
## 9        13.0   72.0      M  
## 10        7.5   64.0      F  
## 11        10.5   74.5      M  
## 12         8.5   67.0      F  
## 13        12.0   71.0      M  
## 14        10.5   71.0      M  
## 15        13.0   77.0      M  
## 16        11.5   72.0      M  
## 17         8.5   59.0      F  
## 18         5.0   62.0      F  
## 19        10.0   72.0      M  
## 20         6.5   66.0      F  
## 21         7.5   64.0      F  
## 22         8.5   67.0      M  
## 23        10.5   73.0      M  
## 24         8.5   69.0      F  
## 25        10.5   72.0      M  
## 26        11.0   70.0      M  
## 27         9.0   69.0      M  
## 28        13.0   70.0      M
```

```
head(housedata)
```

```
##      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
```

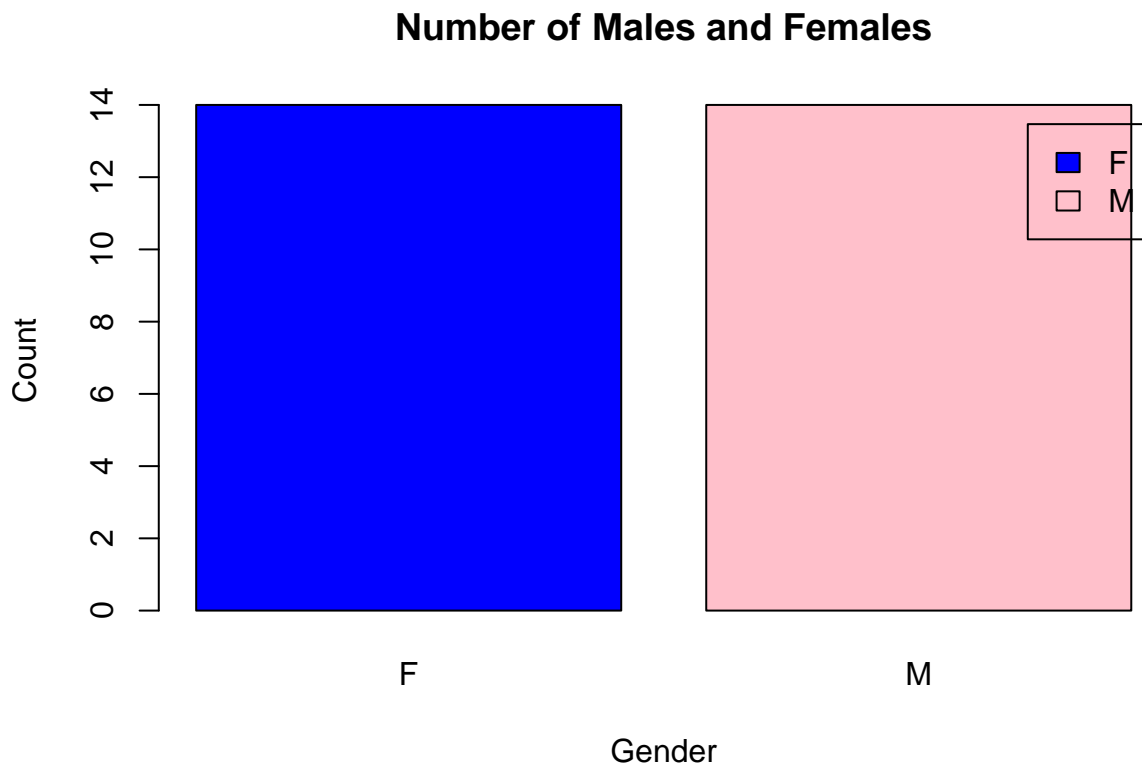
```
males <- subset(housedata, Gender == "M")  
females <- subset(housedata, Gender == "F")  
  
num_males <- nrow(males)  
num_females <- nrow(females)  
  
cat("Number of males:", num_males, "\n")
```

```
## Number of males: 14
cat("Number of females:", num_females, "\n")

## Number of females: 14

#4.c
gender_counts <- table(housedata$Gender)

barplot(gender_counts,
  main = "Number of Males and Females",
  xlab = "Gender",
  ylab = "Count",
  col = c("blue", "pink"),
  legend = names(gender_counts))
```



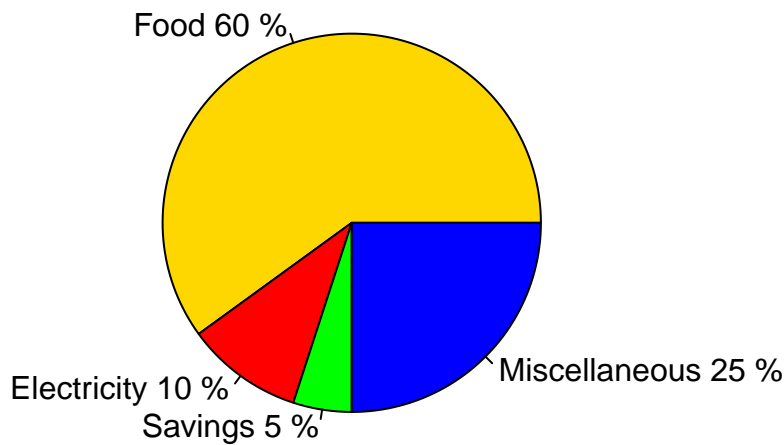
5.

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

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

pie(expense,
  labels = paste(labels, percentages, "%"), main = "Monthly Expenses of Dela Cruz Family",
  col = c("gold", "red", "green", "blue")
)
```

Monthly Expenses of Dela Cruz Family



6.

```
data(iris)

#6a.
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 ...

# Output description: The iris dataset is a data frame with 150 observations and 5 variables.
# The variables are:
# Species: Factor with 3 levels "setosa", "versicolor", and "virginica"
# Sepal.Length: numeric, representing the sepal length in centimeters
# Sepal.Width: numeric, representing the sepal width in centimeters
# Petal.Length: numeric, representing the petal length in centimeters
# Petal.Width: numeric, representing the petal width in centimeters

#b.

mean_measurements <- c(
  mean(iris$Sepal.Length),
  mean(iris$Sepal.Width),
  mean(iris$Petal.Length),
  mean(iris$Petal.Width)
)
names(mean_measurements) <- c("Sepal.Length", "Sepal.Width", "Petal.Length", "Petal.Width")
print(mean_measurements)

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

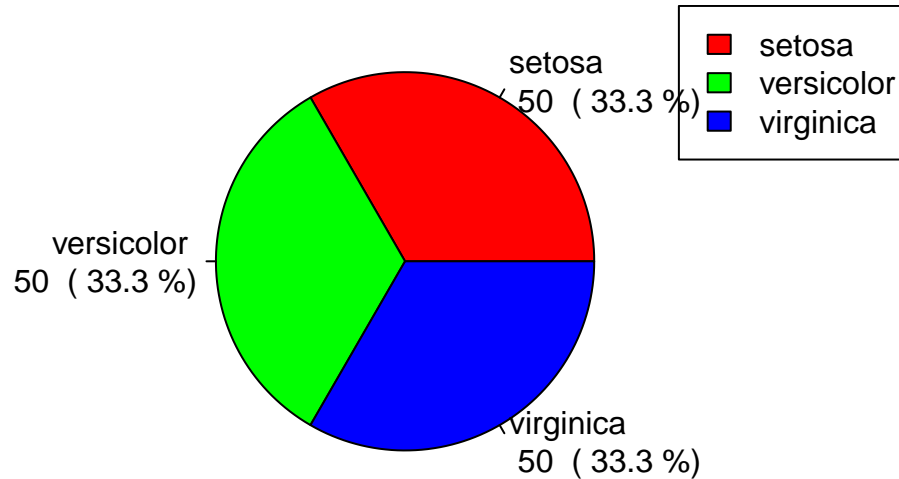
#c
species_counts <- table(iris$Species)
pie(species_counts,
```

```

main = "Iris Species Distribution",
col = c("red", "green", "blue"),
labels = paste(names(species_counts), "\n", species_counts, " (", round(species_counts / sum(species_counts) * 100, 1), "%)", sep = "", collapse = "\n"),
legend("topright", legend = names(species_counts), fill = c("red", "green", "blue"))

```

Iris Species Distribution



```

#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)
```

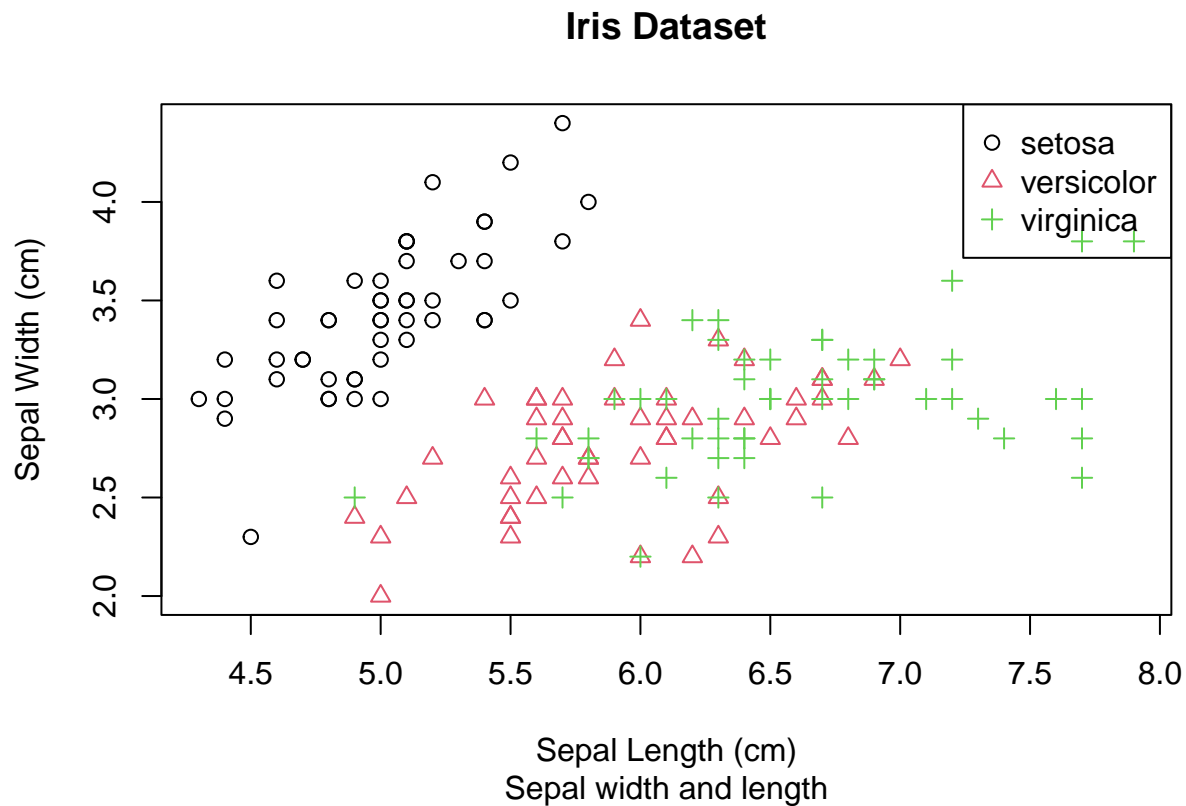
```

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



#f. The scatterplot illustrates a clear distinction between the three iris species based on sepal length.

```
library(readxl)

alexa_data <- read_excel("alexa_file.xlsx")
```

#7a.

```
unique(alexa_data$variation)
```

```
## [1] "Charcoal Fabric"      "Walnut Finish"
## [3] "Heather Gray Fabric"  "Sandstone Fabric"
## [5] "Oak Finish"           "Black"
## [7] "White"                "Black Spot"
## [9] "White Spot"           "Black Show"
## [11] "White Show"           "Black Plus"
## [13] "White Plus"           "Configuration: Fire TV Stick"
```

```
## [15] "Black Dot" "White Dot"

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

alexa_data$variation <- gsub("White Dot", "White Dot", alexa_data$variation)
alexa_data$variation <- gsub("White Plus", "White Plus", alexa_data$variation)
alexa_data$variation <- gsub("White Show", "White Show", alexa_data$variation)
alexa_data$variation <- gsub("White Spot", "White Spot", alexa_data$variation)
unique(alexa_data$variation)

## [1] "Charcoal Fabric" "Walnut Finish"
## [3] "Heather Gray Fabric" "Sandstone Fabric"
## [5] "Oak Finish" "Black"
## [7] "White" "Black Spot"
## [9] "White Spot" "Black Show"
## [11] "White Show" "Black Plus"
## [13] "White Plus" "Configuration: Fire TV Stick"
## [15] "Black Dot" "White Dot"

head(alexa_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

#7b.
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_counts <- alexa_data %>% count(variation)
print(variation_counts)

## # A tibble: 16 x 2
##   variation      n
##   <chr>      <int>
## 1 Black      261
## 2 Black Dot  516
## 3 Black Plus 270
## 4 Black Show 265
```

```
## 5 Black Spot                241
## 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 Plus               78
## 15 White Show               85
## 16 White Spot               109
```

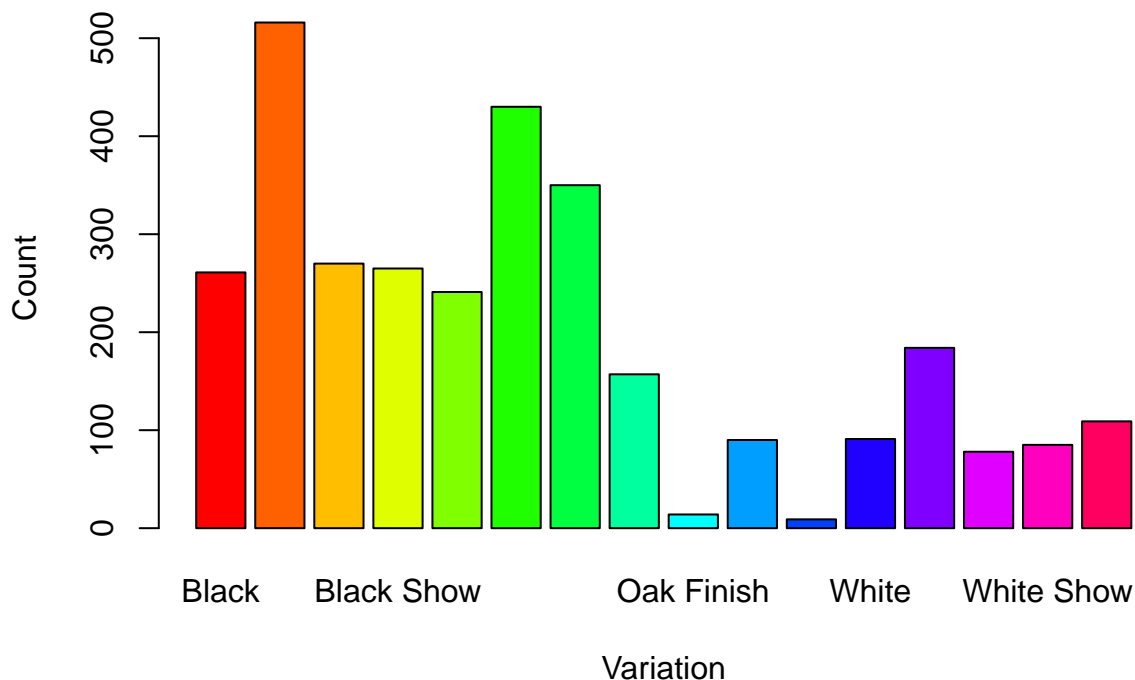
```
save(variation_counts, file = "variations.RData")
```

```
#7c.
```

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

```
barplot(variation_counts$n,
        names.arg = variation_counts$variation,
        main = "Alexa Variations",
        xlab = "Variation",
        ylab = "Count",
        col = rainbow(nrow(variation_counts))
)
```

Alexa Variations



```
#7d.
```

```
black_variants <- variation_counts[grepl("Black", variation_counts$variation), ]
white_variants <- variation_counts[grepl("White", variation_counts$variation), ]
```

```
par(mfrow = c(1, 2))
```



```

barplot(black_variants$n,
        names.arg = black_variants$variation,
        main = "Black Variants",
        xlab = "Variation",
        ylab = "Count",
        col = rainbow(nrow(black_variants)))
)

barplot(white_variants$n,
        names.arg = white_variants$variation,
        main = "White Variants",
        xlab = "Variation",
        ylab = "Count",
        col = rainbow(nrow(white_variants)))
)

```

