

# RWorksheet\_guion#4b

Mikyla Grace Guion

2024-10-30

1. Using the for loop, create an R script that will display a 5x5 matrix as shown in Figure 1. It must contain vectorA = [1,2,3,4,5] and a 5 x 5 zero matrix.

```
vectorA = c(1,2,3,4,5)
matri <- matrix(0, nrow = 5, ncol = 5)
  for(i in 1:5) {
    for(j in 1:5) {
      matri[i,j] <- abs(vectorA[i] - vectorA[j])
    }
    print(matri[i,j])
  }
```

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

2. Print the string “\*” using for() function. The output should be the same as shown in Figure

```
for(i in 1:5) {
  line <- rep('*', i)
  cat(line, sep= " ")
  cat("\n")
}
```

```
## "*"
## "*" "*"
## "*" "*" "*"
## "*" "*" "*" "*"
## "*" "*" "*" "*" "*"
```

3. Get an input from the user to print the Fibonacci sequence starting from the 1st input up to 500. Use repeat and break statements. Write the R Scripts and its output.

```
printFibonacci <- function(start){
  first <- 0
  second <- 1
  next_num <- 0

  if (start == 1){
    cat(first, "", second, "", second, " ")
  }

  for (i in 0:start){
    next_num <- first + second
    first <- second
    second <- next_num
  }

  repeat{
    if (next_num > 500) break
    cat(next_num, " ")
    next_num <- first + second
    first <- second
    second <- next_num
  }
}
#start <- readline(prompt = "Enter starting term: ")
start <- 1
printFibonacci(start)

## 0 1 1 2 3 5 8 13 21 34 55 89 144 233 377
```

4.

a. What is the R script for importing an excel or a csv file? Display the first 6 rows of the dataset? Show your codes and its result

```
library(readxl)

shoe <- read_excel("shoe_size.xlsx")
head(shoe)

## # A tibble: 6 x 3
##   `Shoe Size` Height Gender
##   <dbl>   <dbl> <chr>
## 1      6.5     66    F
## 2      9      68    F
## 3      8.5    64.5  F
## 4      8.5     65    F
## 5     10.5     70    M
```

```
## 6      7      64      F
```

b. Create a subset for gender(female and male). How many observations are there in Male? How about in Female? Write the R scripts and its output.

```
female_subset <- subset(shoe, Gender == "F")
male_subset <- subset(shoe, Gender == "M")

female_count <- nrow(female_subset)
male_count <- nrow(male_subset)

female_count
```

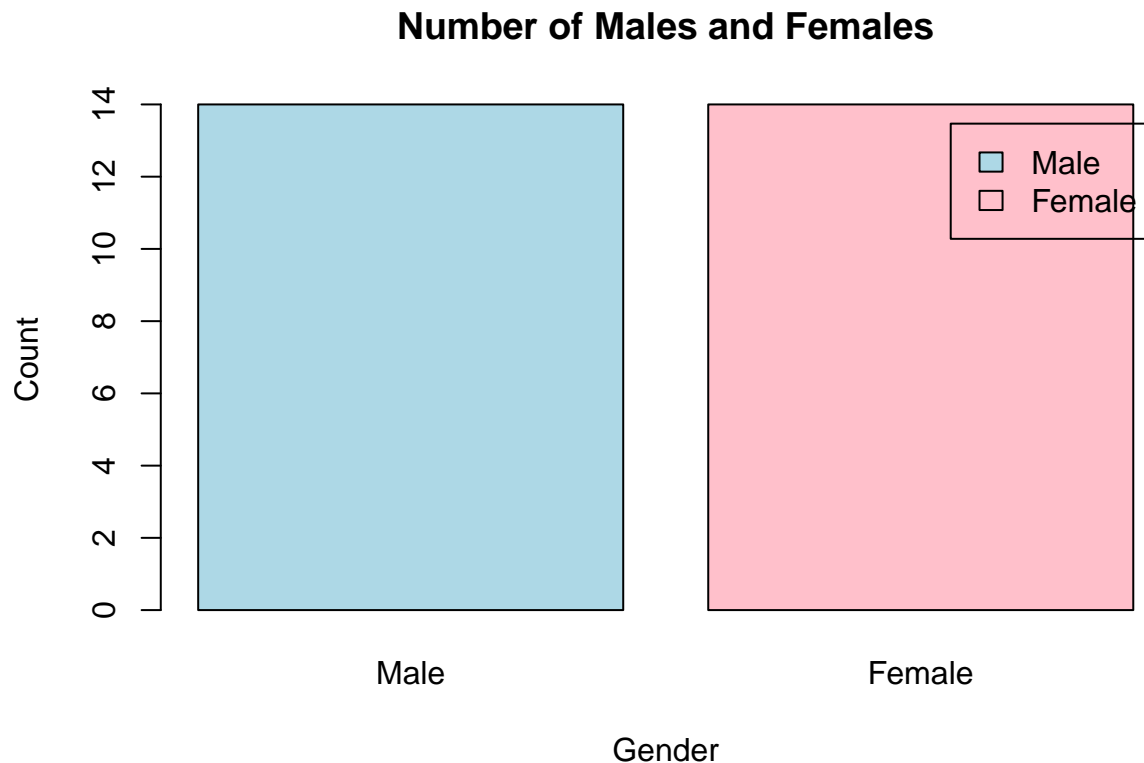
```
## [1] 14
```

```
male_count
```

```
## [1] 14
```

c. Create a graph for the number of males and females for Household Data. Use plot(), chart type = barplot. Make sure to place title, legends, and colors. Write the R scripts and its result.

```
gender_counts <- table(shoe$Gender)
barplot(gender_counts,
        main = "Number of Males and Females",
        xlab = "Gender",
        ylab = "Count",
        col = c("lightblue", "pink"),
        legend = c("Male", "Female"),
        names.arg = c("Male", "Female"))
```



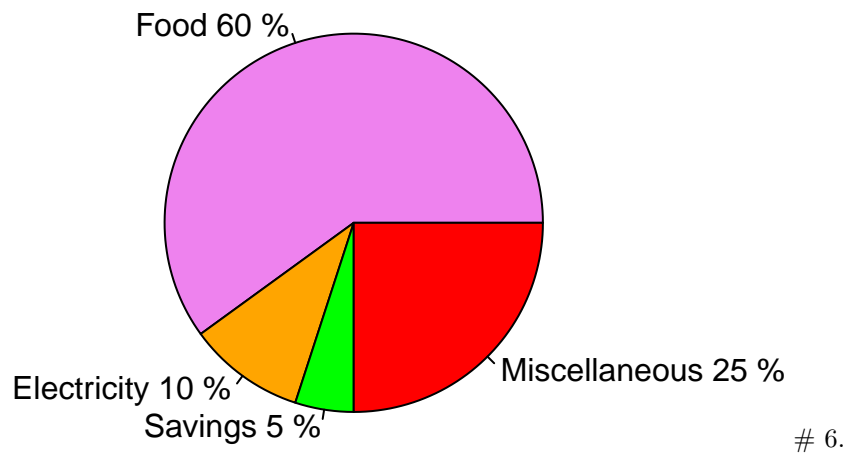
a. Create a piechart that will include labels in percentage. Add some colors and title of the chart. Write the R scripts and show its output.

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

percentages <- round((expenses / sum(expenses)) * 100)
labels <- paste(categories, percentages, "%")

pie(expenses,
    labels = labels,
    col = c("violet", "orange", "green", "red"),
    main = "Household Expenses Distribution")
```

## Household Expenses Distribution



a. Check for the structure of the dataset using the `str()` function. Describe what you have seen in the output.

```
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 ...
```

b. Create an R object that will contain the mean of the `sepal.length`, `sepal.width`, `petal.length`, and `petal.width`. What is the R script and its result?

```
MeanSeL <- mean(iris$Sepal.Length)
MeanSeW <- mean(iris$Sepal.Width)
MeanPeL <- mean(iris$Petal.Length)
MeanPeW <- mean(iris$Petal.Width)
```

```
MeanSeL
```

```
## [1] 5.843333
```

```
MeanSeW
```

```
## [1] 3.057333
```

```
MeanPeL
```

```
## [1] 3.758
```

```
MeanPeW
```

```
## [1] 1.199333
```

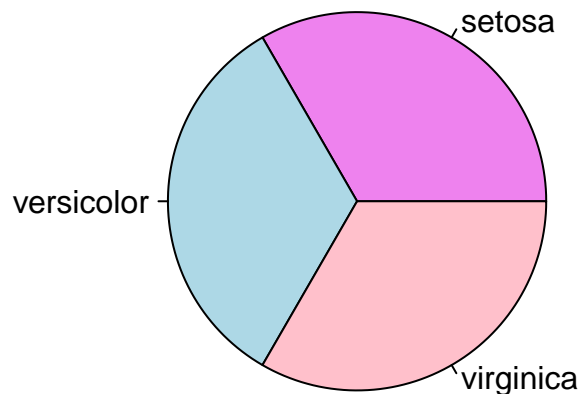
c. Create a pie chart for the Species distribution. Add title, legends, and colors. Write the R script and its result.

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

pie(species_counts,
    main = "Iris Species Distribution",
    legend = names(species_counts),
    col = c("violet", "lightblue", "pink"))
```

```
## Warning in text.default(1.1 * P$x, 1.1 * P$y, labels[i], xpd = TRUE, adj =
## ifelse(P$x < : "legend" is not a graphical parameter
## Warning in text.default(1.1 * P$x, 1.1 * P$y, labels[i], xpd = TRUE, adj =
## ifelse(P$x < : "legend" is not a graphical parameter
## Warning in text.default(1.1 * P$x, 1.1 * P$y, labels[i], xpd = TRUE, adj =
## ifelse(P$x < : "legend" is not a graphical parameter
## Warning in title(main = main, ...): "legend" is not a graphical parameter
```

### Iris Species Distribution



# d. Subset the species into setosa, versicolor, and virginica.  
Write the R scripts and show the last six (6) rows of each species.

```
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

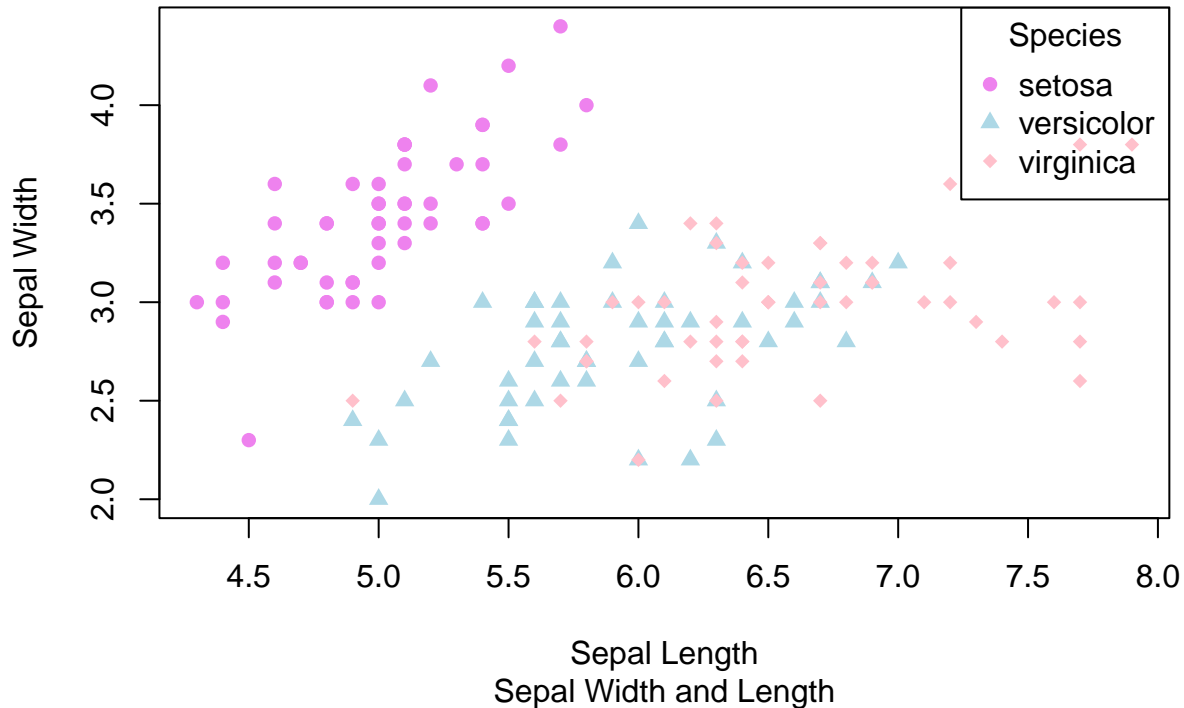
e. Create a scatterplot of the sepal.length and sepal.width using the different species(setosa,versicolor,virginica). Add a title = “Iris Dataset”, subtitle = “Sepal width and length, labels for the x and y axis, the pch symbol and colors should be based on the species.

```
colors <- c("setosa" = "violet", "versicolor" = "lightblue", "virginica" = "pink")
pch_symbols <- c("setosa" = 16, "versicolor" = 17, "virginica" = 18)
```

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

```
legend("topright", legend = levels(iris$Species),
     col = c("violet", "lightblue", "pink"),
     pch = c(16, 17, 18),
     title = "Species")
```

## Iris Dataset



f. Interpret the result. The scatter plot shows the sepal width of each specie on the y axis and sepal length on the x axis. It also shows different shapes and color to represent each specie. When looking at the scatter plot, there is a significant difference between the sizes of the setosa from the versicolor and virginica. Versicolor and virginica species overlap with each other which means they are quite similar with their size.

7.

a. Rename the white and black variants by using `gsub()` function.

```
library(readxl)
library(knitr)

alexa_data <- read_excel("alexa_file.xlsx")

alexa_data$variation <- gsub("Black Dot", "BlackDot", alexa_data$variation)
alexa_data$variation <- gsub("Black Plus", "BlackPlus", alexa_data$variation)
alexa_data$variation <- gsub("Black Show", "BlackShow", alexa_data$variation)
alexa_data$variation <- gsub("Black Spot", "BlackSpot", alexa_data$variation)

# Fix "White" variants
alexa_data$variation <- gsub("White Dot", "WhiteDot", alexa_data$variation)
alexa_data$variation <- gsub("White Plus", "WhitePlus", alexa_data$variation)
alexa_data$variation <- gsub("White Show", "WhiteShow", alexa_data$variation)
alexa_data$variation <- gsub("White Spot", "WhiteSpot", alexa_data$variation)

alexa_data$variation[1052:2000]

##      [1] "WhiteSpot" "BlackSpot" "BlackSpot" "BlackSpot" "BlackSpot" "WhiteSpot"
```



[illegible]

[illegible]

[illegible]

```
knitr::include_graphics("/cloud/project/GUION_CODE/RWorksheet_guion#4/yuh.png")
```

	rating	date	variation	verified_reviews	feedback
1222	5	2018-07-27	BlackSpot	Love my spot, it now acts as my new alarm.	1
1223	5	2018-07-27	WhiteSpot	This is the one of my favorite Amazon devices. Hav...	1
1224	5	2018-07-27	BlackSpot	Wasn't so sure about bringing &#34;smart technolo...	1
1225	5	2018-07-27	WhiteSpot	Love it	1
1226	1	2018-07-27	BlackSpot	You need a Harvard law degree to operate this thing...	0
1227	5	2018-07-26	BlackSpot	Love it getting use to it and Alexa is getting use to ...	1
1228	2	2018-07-26	BlackSpot	product turns on randomly and sometimes at night	0
1229	5	2018-07-26	WhiteSpot	I used it as an alarm clock i lov it	1
1230	3	2018-07-26	BlackSpot	It was much smaller than I was expecting for the cost.	1
1231	5	2018-07-26	BlackSpot	I'm amazed	1
1232	5	2018-07-26	WhiteSpot	Alexa wakes me up everyday for work to whatever s...	1
1233	5	2018-07-26	BlackSpot	Fun so far...still learning how it all works	1
1234	5	2018-07-26	BlackSpot	I purchased this on prime day mostly as a present f...	1
1235	5	2018-07-26	BlackSpot	We love the echo spot I use it on my nightstand for ...	1
1236	5	2018-07-26	BlackSpot	It was easy set up. I use it more than I thought.	1
1237	1	2018-07-26	BlackSpot	I would love this but there is no way to stop the scre...	0
1238	5	2018-07-26	BlackSpot	I am a small business owner with no staff. With the ...	1
1239	5	2018-07-26	BlackSpot	Bedroom clock , ask questions, weather reportsAlar...	1
1240	4	2018-07-26	WhiteSpot	Speakers are not as loud as Google Home	1
1241	2	2018-07-26	WhiteSpot	I haven't figured out how to make or receive calls. ...	0

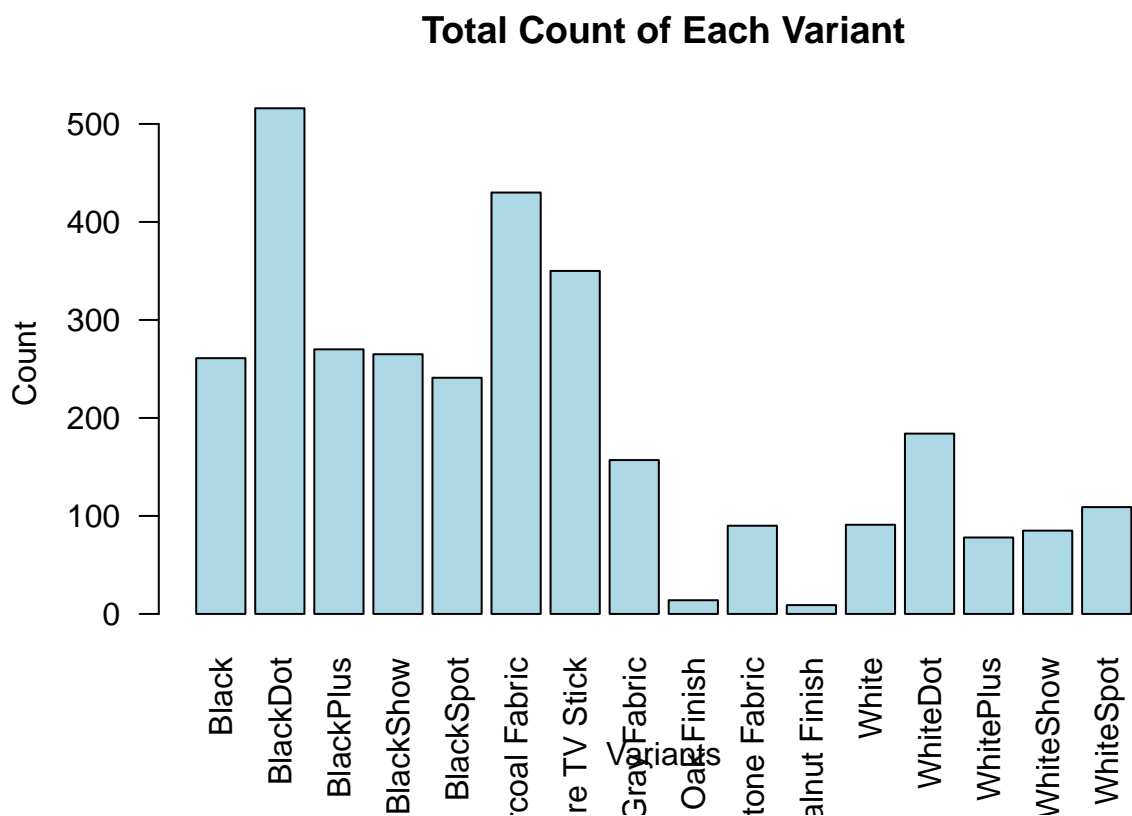
# b. b. Get the total number of each variations and save it into another object. Save the object as variations.RData. Write the R scripts. What is its result?

```
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
variant_counts <- alexa_data %>%
  count(variation)
save(variant_counts, file = "variations.RData")
```

c. From the `variations.RData`, create a `barplot()`. Complete the details of the chart which include the title, color, labels of each bar.

```
barplot(
  variant_counts$n,
  names.arg = variant_counts$variation,
  col = "lightblue",
  main = "Total Count of Each Variant",
  xlab = "Variants",
  ylab = "Count",
  las = 2
)
```



# d. Create a `barplot()` for the black and white variations. Plot it in 1 frame, side by side. Complete the details of the chart.

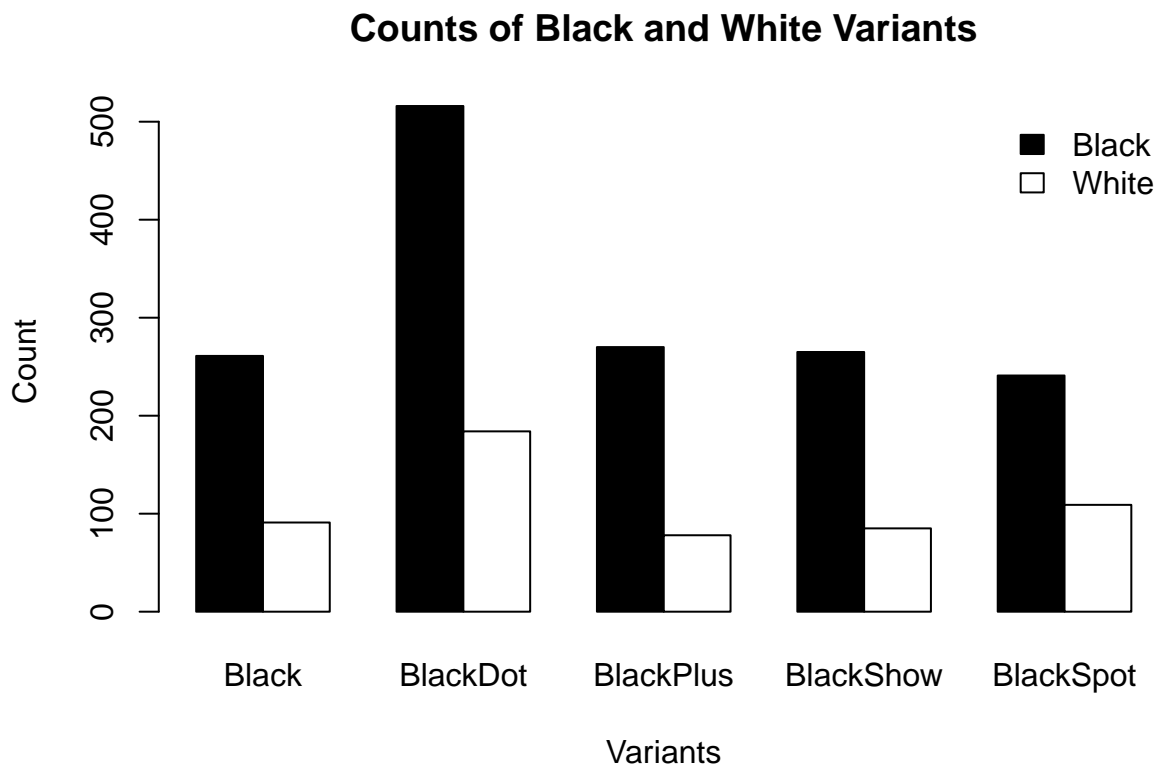
```
black_white_counts <- variant_counts %>%
  filter(grepl("Black|White", variation))

barplot(
  matrix(black_white_counts$n, nrow = 2, byrow = TRUE),
  beside = TRUE,
  names.arg = gsub("^(Black|White)\\s", "", black_white_counts$variation[1:(nrow(black_white_counts)/2)]),
  col = c("black", "white"),
  main = "Counts of Black and White Variants",
  xlab = "Variants",
  ylab = "Count",
)
```

```

legend.text = c("Black", "White"),
args.legend = list(x = "topright", bty = "n")
)

```



```

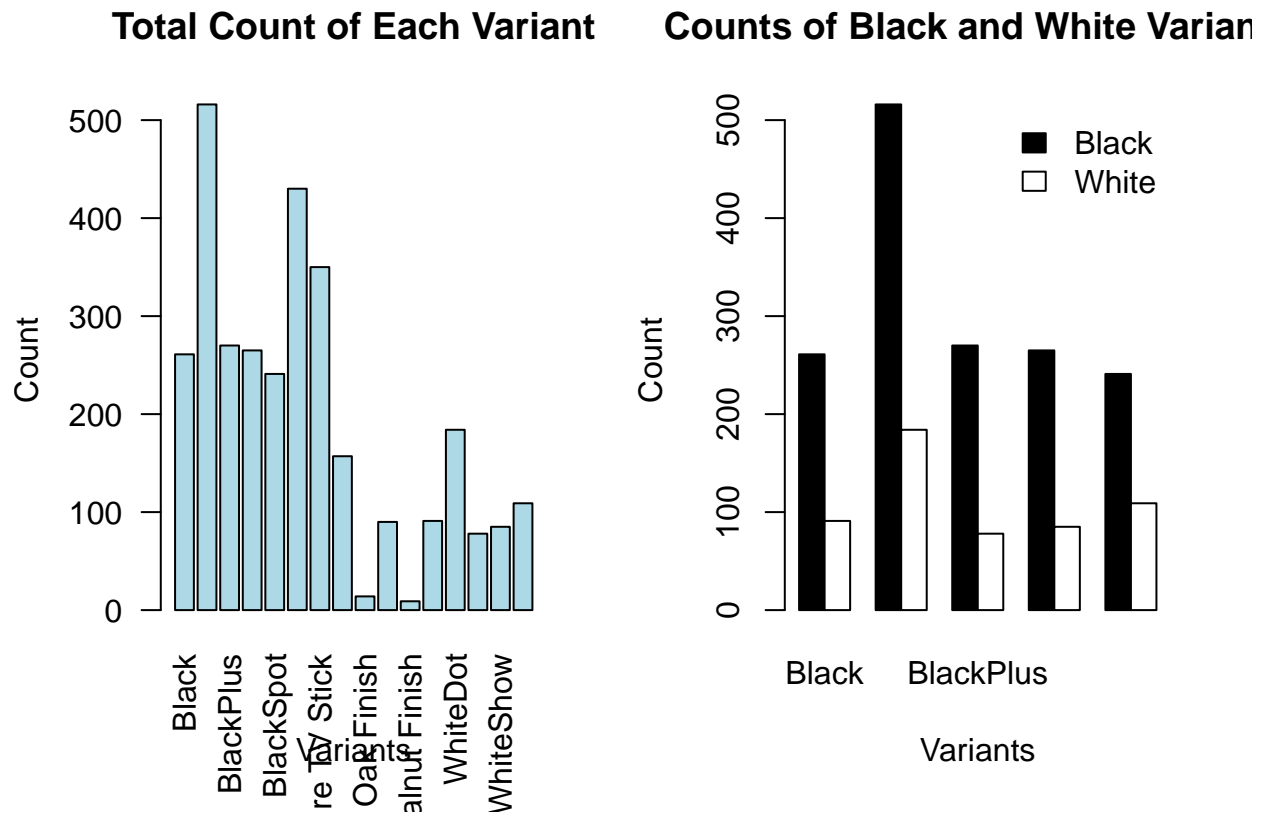
par(mfrow = c(1, 2))

barplot(
  variant_counts$n,
  names.arg = variant_counts$variation,
  col = "lightblue",
  main = "Total Count of Each Variant",
  xlab = "Variants",
  ylab = "Count",
  las = 2
)

black_white_counts <- variant_counts %>%
  filter(grepl("Black|White", variation))

barplot(
  matrix(black_white_counts$n, nrow = 2, byrow = TRUE),
  beside = TRUE,
  names.arg = gsub("(Black|White)\\s", "", black_white_counts$variation[1:(nrow(black_white_counts)/2)]),
  col = c("black", "white"),
  main = "Counts of Black and White Variants",
  xlab = "Variants",
  ylab = "Count",
  legend.text = c("Black", "White"),
  args.legend = list(x = "topright", bty = "n")
)

```



The first graph shows all variants in a single sequence and its easier to see the total count of each variant in one glance. The second graph groups the black and white variations side-by-side for each type which allows for a direct comparison between black and white variants of each type.