

RWorksheet_Cacho#4b

Janelle Cacho

2024-10-28

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

matrix_result

```
##      [,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("* ", i), "\n")
}
```

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

```
first <- 10 # You can change this value as needed
fibonacci <- c(first)
```

```
repeat {
  next_num <- sum(tail(fibonacci, 2))
  if (next_num > 500) break
  cat(next_num, " ")
  fibonacci <- c(fibonacci, next_num) #
}
```

```
## 10 20 30 50 80 130 210 340
```

```
cat("\nFibonacci sequence:", paste(fibonacci, collapse = ", "), "\n")
```

```
##
## Fibonacci sequence: 10, 10, 20, 30, 50, 80, 130, 210, 340
```

```

# 4a
library(readr)
sample_data <- read_csv("sample_data.csv")

## Rows: 28 Columns: 3
## -- Column specification -----
## Delimiter: ","
## chr (1): Gender
## dbl (2): ShoeSize, Height
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
data <- read_csv("sample_data.csv")
head(data)

##   ShoeSize 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

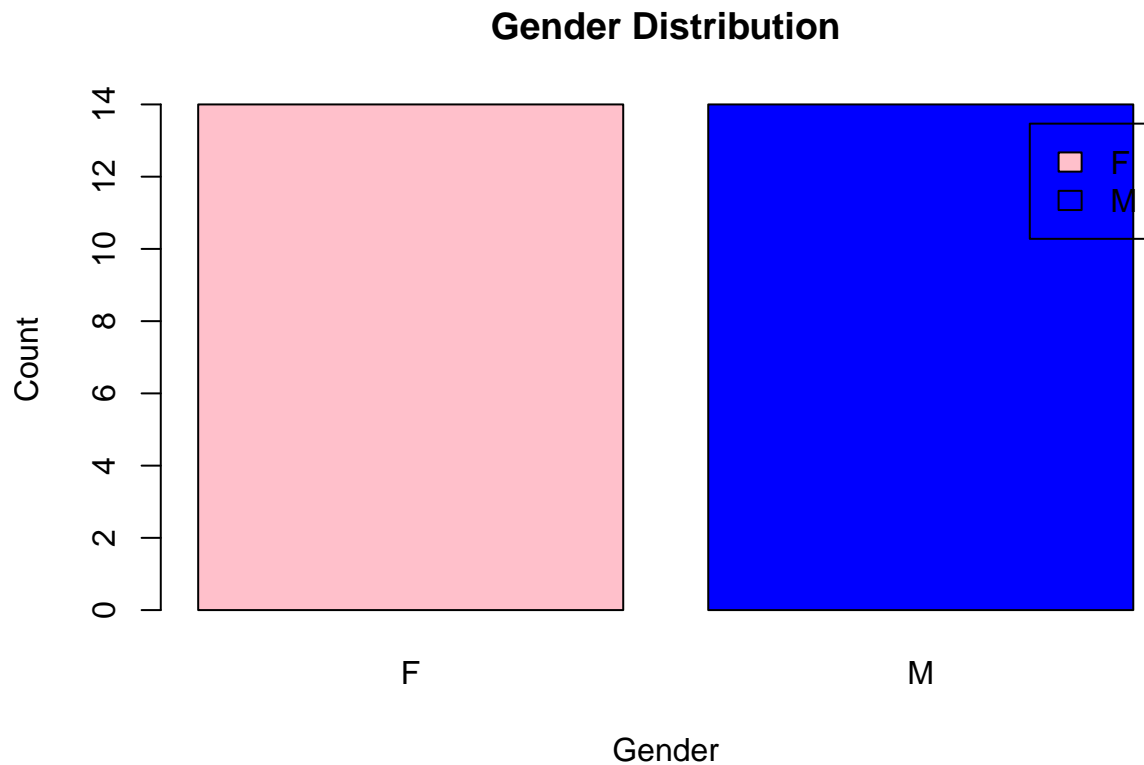
# 4b
femdata <- subset(data, Gender == "F")
maledata <- subset(data, Gender == "M")
cat("Female count:", nrow(femdata), "\n")

## Female count: 14
cat("Male count:", nrow(maledata), "\n")

## Male count: 14

# 4c
gender_count <- table(data$Gender)
barplot(gender_count, main= "Gender Distribution", col = c("pink", "blue"),
        xlab="Gender", ylab="Count", legend=TRUE)

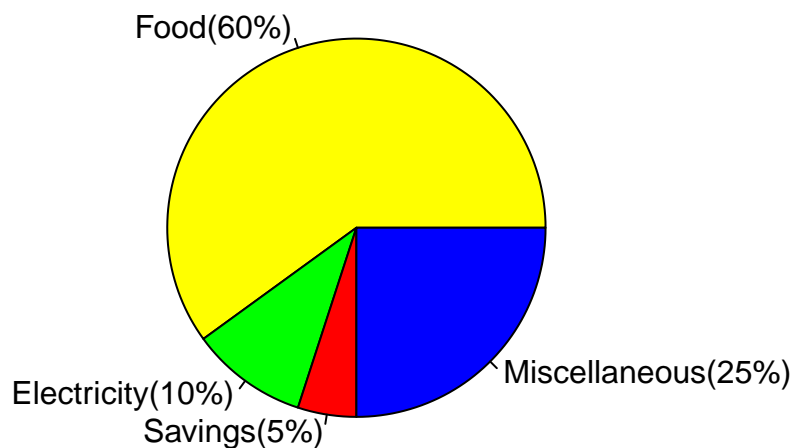
```



5

```
expenses <- c(Food = 60, Electricity = 10, Savings = 5, Miscellaneous = 25)
percent <- paste0(names(expenses), "(", round(100*expenses / sum(expenses), 1), "%)")
pie(expenses, labels= percent, col = c("yellow", "green", "red", "blue"), main = "Monthly Family Expenses")
```

Monthly Family Expenses



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 ...
```

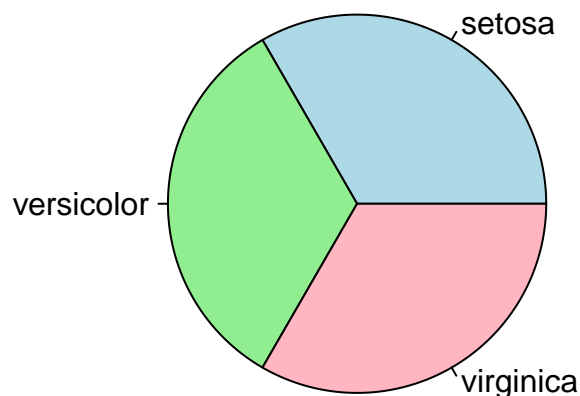
```
## $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
## $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
## $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...

# 6b
mean_values <- colMeans(iris[, c("Sepal.Length", "Sepal.Width", "Petal.Length", "Petal.Width")])
mean_values
```

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

```
# 6c
species_counts <- table(iris$Species)
pie(species_counts, main="Species Distribution", col=c("lightblue", "lightgreen", "lightpink"),
    labels=names(species_counts))
```

Species Distribution



```
# 6d
setosa <- subset(iris, Species == "setosa")
versicolor <- subset(iris, Species == "versicolor")
virginica <- subset(iris, Species == "virginica")

# Show last 6 rows of each subset
tail(setosa, 6)
```

```
##      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, 6)
```

```
##      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
```

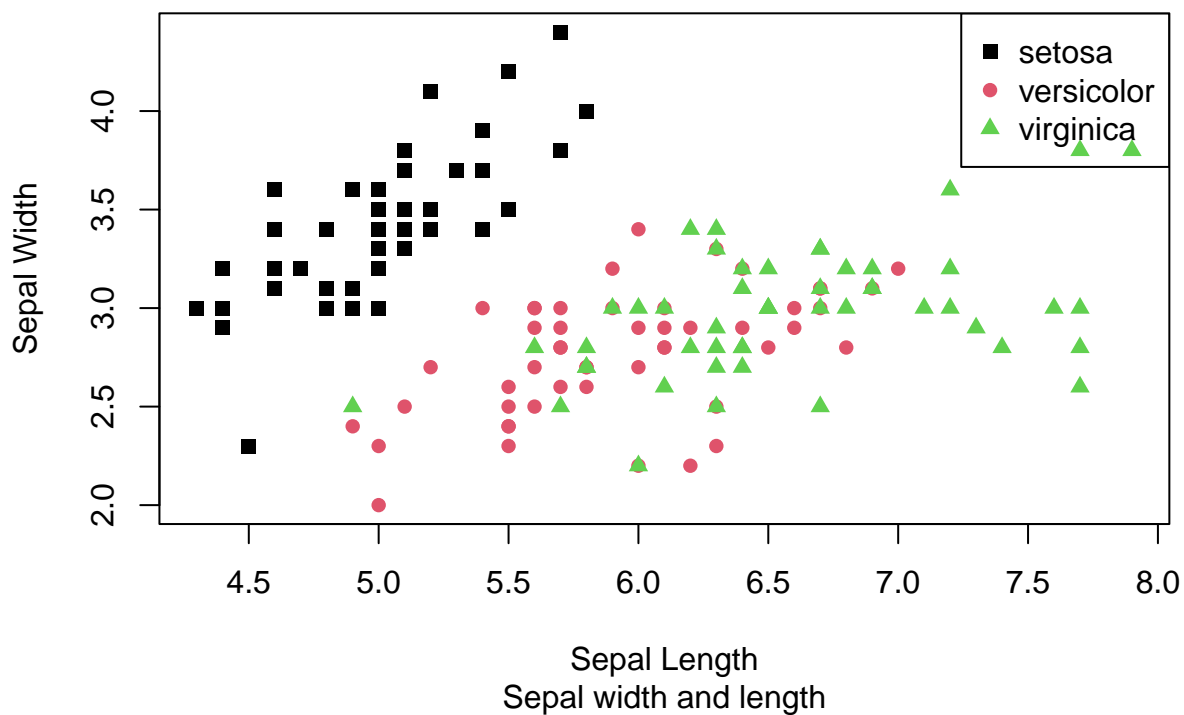
```
tail(virginica, 6)
```

```
##      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)
plot(iris$Sepal.Length, iris$Sepal.Width,
     col = iris$Species,
     pch = as.numeric(iris$Species) + 14, # Different symbols by 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 = 15:17)
```

Iris Dataset



```
# 6f
```

#In this plot, we can observe the classification of species based on sepal dimensions. Example, Setosa .

```
# 7
```

```
library(readxl)
```

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

```
## # A tibble: 3,150 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 play~ 1
## 4     5 2018-07-31 00:00:00 Charcoal Fabric I have had a lot of ~ 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 ~ 1
## 7     3 2018-07-31 00:00:00 Sandstone Fabric Without having a cel~ 1
## 8     5 2018-07-31 00:00:00 Charcoal Fabric I think this is the ~ 1
## 9     5 2018-07-30 00:00:00 Heather Gray Fabric looks great      1
## 10    5 2018-07-30 00:00:00 Heather Gray Fabric Love it! I've listen~ 1
## # i 3,140 more rows
```

```
# 7a
data$variation <- gsub("Old Name", "New Name", data$variation)
data$variation <- gsub("Black", "Charcoal Fabric", data$variation)
knitr::include_graphics("data_cs101.png")
```

	rating	date	variation	verified_reviews	fe
1	5	2018-07-31	Charcoal Fabric	Love my Echo!	
2	5	2018-07-31	Charcoal Fabric	Loved it!	
3	4	2018-07-31	Walnut Finish	Sometimes while playing a game, you can answer a questio...	
4	5	2018-07-31	Charcoal Fabric	I have had a lot of fun with this thing. My 4 yr old learns ab...	
5	5	2018-07-31	Charcoal Fabric	Music	
6	5	2018-07-31	Heather Gray Fabric	I received the echo as a gift. I needed another Bluetooth or ...	
7	3	2018-07-31	Sandstone Fabric	Without having a cellphone, I cannot use many of her featur...	
8	5	2018-07-31	Charcoal Fabric	I think this is the 5th one I've purchased. I'm working on get...	
9	5	2018-07-30	Heather Gray Fabric	looks great	
10	5	2018-07-30	Heather Gray Fabric	Love it! I've listened to songs I haven't heard since childhoo...	

```
# 7b
install.packages("dplyr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
```

```
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 <- data %>%
  count(variation)

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

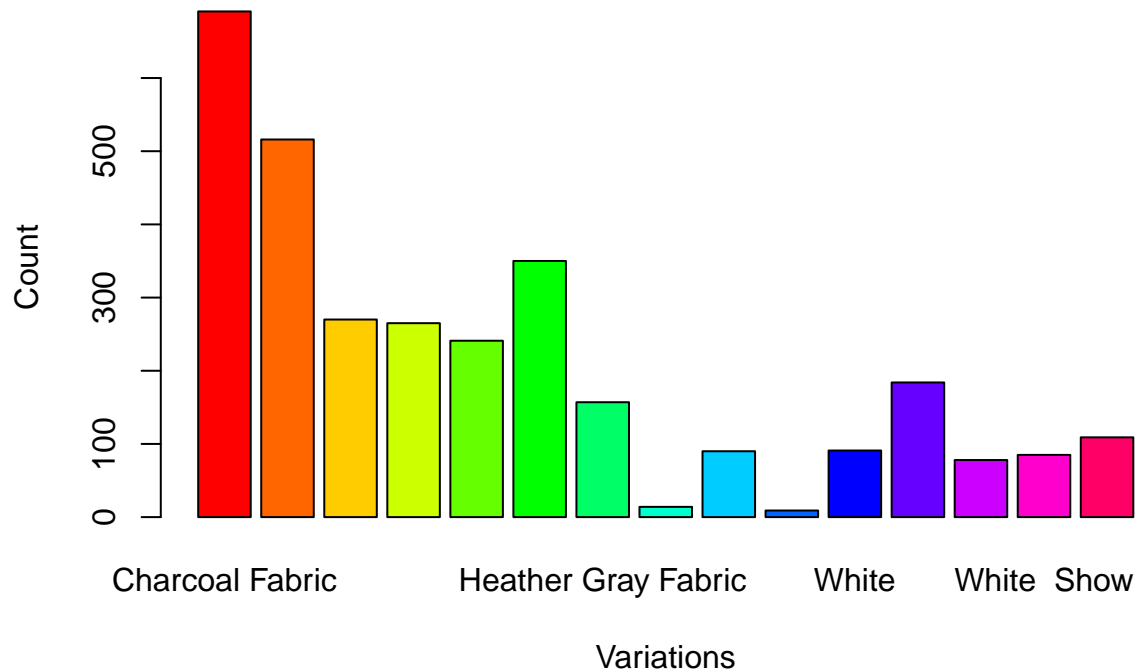
variation_counts

## # A tibble: 15 x 2
##   variation          n
##   <chr>          <int>
## 1 Charcoal Fabric    691
## 2 Charcoal Fabric Dot    516
## 3 Charcoal Fabric Plus    270
## 4 Charcoal Fabric Show    265
## 5 Charcoal Fabric Spot    241
## 6 Configuration: Fire TV Stick    350
## 7 Heather Gray Fabric    157
## 8 Oak Finish         14
## 9 Sandstone Fabric     90
## 10 Walnut Finish        9
## 11 White              91
## 12 White Dot         184
## 13 White Plus         78
## 14 White Show         85
## 15 White Spot        109

# 7c
load("variations.RData")

barplot(variation_counts$n, names.arg = variation_counts$variation,
  col = rainbow(length(variation_counts$variation)),
  main = "Total Counts of Variations",
  xlab = "Variations",
  ylab = "Count")
```

Total Counts of Variations



```
# 7d
black_variants <- variation_counts %>% filter(grepl("Black", variation))
white_variants <- variation_counts %>% filter(grepl("White", variation))

par(mfrow = c(1, 2)) # 1 row, 2 columns

if (nrow(black_variants) > 0) {
  barplot(black_variants$n, names.arg = black_variants$variation,
    col = rainbow(length(black_variants$variation)),
    main = "Black Variants",
    xlab = "Total Numbers",
    ylab = "Variants")
} else {
  plot.new() # Create an empty plot
  title(main = "No Black Variants Found")
}

if (nrow(white_variants) > 0) {
  barplot(white_variants$n, names.arg = white_variants$variation,
    col = rainbow(length(white_variants$variation)),
    main = "White Variants",
    xlab = "Total Numbers",
    ylab = "Variants")
} else {
  plot.new() # Create an empty plot
  title(main = "No White Variants Found")
}
```


No Black Variants Found

White Variants

