RWorksheet#4B_Sapan

Leorenze Marc Sapan

2025-12-18

```r
# 1.)

numVector <- c(1, 2, 3, 4, 5)
differenceMatrix <- matrix(0, nrow = 5, ncol = 5)

for (rowIndex in 1:5) {
  for (colIndex in 1:5) {
    differenceMatrix[rowIndex, colIndex] <- abs(rowIndex - colIndex)
  }
}

print(differenceMatrix)



# 2.)

for (row in 1:5) {
  for (starCount in 1:row) {
    cat("* ")
  }
  cat("\n")
}



# 3.)

firstValue <- as.numeric(readline(prompt = "Enter first number: "))
secondValue <- as.numeric(readline(prompt = "Enter second number: "))

cat(firstValue, secondValue, " ")

repeat {
  fibonacciNext <- firstValue + secondValue
  if (fibonacciNext > 500) break

  cat(fibonacciNext, " ")
  firstValue <- secondValue
  secondValue <- fibonacciNext
}
```

```r
# BASIC GRAPHICS

# 4.)

shoeData <- read.csv("shoe_size.csv")
head(shoeData)

# 4b.)

femaleData <- subset(shoeData, Gender == "F")
maleData <- subset(shoeData, Gender == "M")

cat("Number of Female observations:", nrow(femaleData), "\n")
cat("Number of Male observations:", nrow(maleData), "\n")


# 4c.)

genderFrequency <- table(shoeData$Gender)

barplot(
  genderFrequency,
  main = "Number of Males and Females in Shoe Size Data",
  xlab = "Gender",
  ylab = "Count",
  col = c("pink", "lightblue"),
  legend.text = c("Female", "Male")
)


# 5.)

expenseValues <- c(60, 10, 5, 25)
expenseCategories <- c("Food", "Electricity", "Savings", "Miscellaneous")

expensePercent <- round(expenseValues / sum(expenseValues) * 100)
expenseLabels <- paste(expenseCategories, expensePercent, "%")

pie(
  expenseValues,
  labels = expenseLabels,
  col = c("gold", "skyblue", "lightgreen", "tomato"),
  main = "Monthly Income Distribution of Dela Cruz Family"
)


# 6.)

data(iris)

# 6a.)
str(iris)
```

```r
# 6b.)

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

# 6c.)

speciesFrequency <- table(iris$Species)

pie(
  speciesFrequency,
  labels = names(speciesFrequency),
  col = c("lightgreen", "skyblue", "pink"),
  main = "Distribution of Iris Species"
)

legend(
  "topright",
  legend = names(speciesFrequency),
  fill = c("lightgreen", "skyblue", "pink")
)

# 6d.)

setosaData <- subset(iris, Species == "setosa")
versicolorData <- subset(iris, Species == "versicolor")
virginicaData <- subset(iris, Species == "virginica")

tail(setosaData)
tail(versicolorData)
tail(virginicaData)

# 6e.)

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

plot(
  iris$Sepal.Length,
  iris$Sepal.Width,
  col = c("red", "green", "blue")[iris$Species],
  pch = 19,
  main = "Iris Dataset",
  sub = "Sepal Width and Length",
  xlab = "Sepal Length (cm)",
  ylab = "Sepal Width (cm)"
)

legend(
  "topright",
  legend = levels(iris$Species),
  col = c("red", "green", "blue"),
  pch = 19
)
```

```r
# 7.)

install.packages("readxl")
library(readxl)

alexaData <- read_excel("alexa_file.xlsx")

# 7a.)
alexaData$variation <- gsub("Black  ", "Black ", alexaData$variation)
alexaData$variation <- gsub("White  ", "White ", alexaData$variation)

install.packages("dplyr")
library(dplyr)

variationSummary <- alexaData %>%
  count(variation)

# 7b.)

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

# 7c.)

variationValues <- variationSummary$n
names(variationValues) <- variationSummary$variation

barplot(
  variationValues,
  main = "Total Number of Each Alexa Variation",
  xlab = "Variations",
  ylab = "Total Count",
  col = rainbow(length(variationValues)),
  las = 2
)

# 7d.)

blackVariationData <- variationSummary %>% filter(grepl("Black", variation))
whiteVariationData <- variationSummary %>% filter(grepl("White", variation))

par(mfrow = c(1, 2))

barplot(
  blackVariationData$n,
  names.arg = blackVariationData$variation,
  main = "Black Variants",
  col = "black",
  las = 2
)

barplot(
```

```
  whiteVariationData$n,
  names.arg = whiteVariationData$variation,
  main = "White Variants",
  col = "gray",
  las = 2
)
```