Assignment_3, Div: CSAI-B, Roll No.: 37

Atharva Salitri

25-01-2025

Exercise 1

```
DF1<-read.csv("hair_eye_color_CSV.csv",header = TRUE)
attach(DF1)
```

```
## The following objects are masked from DF1 (pos = 3):
##
##     Eye.Color, Hair.Color, Person.No.
```

```
# 1  How many people have brown eye color?
sum(Eye.Color=="Brown")
```

```
## [1] 10
```

```
# 2 How many people have Blonde hair?
sum(Hair.Color=="Blonde")
```

```
## [1] 6
```

```
# 3 How many Brown haired people have Black eyes?
sum(Hair.Color=="Brown" & Eye.Color=="Black")
```

```
## [1] 2
```

```
# 4 What is the percentage of people with Green eyes?
(sum(Eye.Color=="Green")/nrow(DF1))*100
```

```
## [1] 10
```

```
# 5 What percentage of people have red hair and Blue eyes?
(sum(Eye.Color=="Blue" & Hair.Color=="Red")/nrow(DF1))*100
```

```
## [1] 5
```

Exercise 2

```
data <- read.csv("hair_eye_color_CSV.csv")

# 1. How many people have brown eye color?
num_brown_eyes <- sum(data$Eye.Color == "Brown")
cat("Number of people with brown eyes:", num_brown_eyes, "\n")
```

```
## Number of people with brown eyes: 10
```

```
# 2. How many people have Blonde hair?
num_blonde_hair <- sum(data$Hair.Color == "Blonde")
cat("Number of people with blonde hair:", num_blonde_hair, "\n")
```

```
## Number of people with blonde hair: 6
```

```
# 3. How many Brown-haired people have Black eyes?
num_brown_hair_black_eyes <- sum(data$Hair.Color == "Brown" & data$Eye.Color == "Black")
cat("Number of brown-haired people with black eyes:", num_brown_hair_black_eyes, "\n")
```

```
## Number of brown-haired people with black eyes: 2
```

```
# 4. What is the percentage of people with Green eyes?
num_green_eyes <- sum(data$Eye.Color == "Green")
percentage_green_eyes <- (num_green_eyes / nrow(data)) * 100
cat("Percentage of people with green eyes:", percentage_green_eyes, "%\n")
```

```
## Percentage of people with green eyes: 10 %
```

```
# 5. What percentage of people have red hair and Blue eyes?
num_red_hair_blue_eyes <- sum(data$Hair.Color == "Red" & data$Eye.Color == "Blue")
percentage_red_hair_blue_eyes <- (num_red_hair_blue_eyes / nrow(data)) * 100
cat("Percentage of people with red hair and blue eyes:", percentage_red_hair_blue_eyes, "%
\n")
```

```
## Percentage of people with red hair and blue eyes: 5 %
```

Exercise 3

```
data <- read.csv("germination_CSV.csv")

# 1. What is the average number of seeds germinated for the uncovered boxes with a level of w
atering equal to 4?
avg_seeds_uncovered_watering_4 <- mean(data$germinated[data$Box == "Uncovered" & data$water_a
mt == 4], na.rm = TRUE)
cat("Average number of seeds germinated for uncovered boxes with watering level 4:", avg_seed
s_uncovered_watering_4, "\n")
```

```
## Average number of seeds germinated for uncovered boxes with watering level 4: 78
```

```
# 2. What is the median value for the data of covered boxes?
median_seeds_covered <- median(data$germinated[data$Box == "Covered"], na.rm = TRUE)
cat("Median number of seeds germinated for covered boxes:", median_seeds_covered, "\n")
```

```
## Median number of seeds germinated for covered boxes: 45
```

```
# 3. Analysis of association
# Association of levels of watering with the number of germinating seeds in case of covered a
nd uncovered boxes
covered_correlation <- cor(data$water_amt[data$Box == "Covered"],
                           data$germinated[data$Box == "Covered"], use = "complete.obs")
cat("Correlation between watering level and number of seeds germinating for covered boxes:",
covered_correlation, "\n")
```

```
## Correlation between watering level and number of seeds germinating for covered boxes: -0.6
527353
```

```
uncovered_correlation <- cor(data$water_amt[data$Box == "Uncovered"],
                             data$germinated[data$Box == "Uncovered"], use = "complete.obs")
cat("Correlation between watering level and number of seeds germinating for uncovered boxe
s:", uncovered_correlation, "\n")
```

```
## Correlation between watering level and number of seeds germinating for uncovered boxes: -
0.050549
```

```
# Association of the number of germinating seeds with the fact that the boxes were covered or
uncovered
avg_seeds_covered <- mean(data$germinated[data$Box == "Covered"], na.rm = TRUE)
avg_seeds_uncovered <- mean(data$germinated[data$Box == "Uncovered"], na.rm = TRUE)
cat("Average seeds germinated for covered boxes:", avg_seeds_covered, "\n")
```

```
## Average seeds germinated for covered boxes: 47.29167
```

```
cat("Average seeds germinated for uncovered boxes:", avg_seeds_uncovered, "\n")
```

```
## Average seeds germinated for uncovered boxes: 47.95833
```