Worksheet#4a

Lorie Mae Tupaz BSIT 2B

2024-10-19

shoe_data <- data.frame(ShoeSize = c(6.5, 9.0, 8.5, 8.5, 10.5, 7.0, 9.5, 9.0, 13.0, 7.5, 10.5, 8.5, 12.0, 10.5, 13.0, 11.5, 8.5, 5.0, 10.0, 6.5, 7.5, 8.5, 10.5, 8.5, 10.5, 11.0, 9.0, 13.0), Height = c(66.0, 68.0, 64.5, 65.0, 70.0, 64.0, 70.0, 71.0, 72.0, 64.0, 74.5, 67.0, 71.0, 71.0, 77.0, 72.0, 59.0, 62.0, 72.0, 66.0, 64.0, 67.0, 73.0, 69.0, 72.0, 70.0, 69.0, 70.0), Gender = c("F", "F", "F", "F", "M", "F", "F", "F", "M", "F", "M", "F", "M", "M", "M", "M", "F", "F", "M", "F", "F", "M", "M", "F", "M", "M", "M", "M"))

print(shoe_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 #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

#1.a

#Shoe Size ranges from 5.0 to 13.0, indicating a variety of sizes.

#Height varies from 59.0 to 77.0 inches, reflecting a diverse range of heights.

#Gender is categorized as "M" for male and "F" for female, allowing for comparisons between the two groups.

#1.b female_subset <- shoe_data[shoe_data$Gender == "F",]

male_subset <- shoe_data[shoe_data$Gender == "M",]

print(female_subset) print(male_subset)

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 6 7.0 64.0 F 7 9.5 70.0 F 8 9.0 71.0 F 10 7.5 64.0 F 12 8.5 67.0 F 17 8.5 59.0 F 18 5.0 62.0 F 20 6.5 66.0 F 21 7.5 64.0 F 24 8.5 69.0 F

ShoeSize Height Gender

#5 10.5 70.0 M #9 13.0 72.0 M #11 10.5 74.5 M #13 12.0 71.0 M #14 10.5 71.0 M #15 13.0 77.0 M #16 11.5 72.0 M #19 10.0 72.0 M #22 8.5 67.0 M #23 10.5 73.0 M #25 10.5 72.0 M #26 11.0 70.0 M #27 9.0 69.0 M #28 13.0 70.0 M

#1.c mean_shoe_size <- mean(shoe_data$ShoeSize)$mean_h eight < -mean(shoe_d ata$Height)

Use paste() to concatenate the strings, avoiding special characters

paste("Mean Shoe Size:", mean_shoe_size) paste("Mean Height:", mean_height)

#Mean Shoe Size: 9.410714 #Mean Height: 68.57143

#1.d colors <- ifelse(shoe_data$Gender == "M", "blue", "pink")

plot(shoe_data$ShoeSize, shoe_data$Height, main = "Shoe Size vs Height", xlab = "Shoe Size", ylab = "Height", pch = 19, col = colors)

legend("topright", legend = c("Male", "Female"), col = c("blue", "pink"), pch = 19)

#Yes, there is a relationship between shoe size and height because taller individuals generally have larger feet. This correlation exists due to proportionality, growth patterns, and genetic factors. As people grow taller, their foot size often increases to maintain balance and support their body structure.

#Factors

#2

months <- c("March", "April", "January", "November", "January", "September", "October", "September", "November", "August", "January", "November", "November", "February", "May", "August", "July", "December", "August", "August", "September", "November", "February", "April")

factor_months_vector <- factor(months)

print(factor_months_vector)

#[1] March April January November January September October September November August January
#[12] November November February May August July December August August September November #[23] February April
#Levels: April August December February January July March May November October September

#3 months_vector <- c("March", "April", "January", "November", "January", "September", "October", "September", "November", "August", "January", "November", "November", "February", "May", "August", "July", "December", "August", "August", "September", "November", "February", "April")

factor_months_vector <- factor(months_vector)

summary(months_vector) summary(factor_months_vector)

#Length Class Mode # 24 character character

April August December February January July March May November October September

2 4 1 2 3 1 1 1 5 1 3

#No, they are not equally useful.

#months_vector shows all the months as they appear, including duplicates.

#factor_months_vector summarizes the unique months and their frequencies, making it easier to analyze the data.

#4 factor_data <- c("East", "West", "North", "West", "West", "North", "West", "West")

new_order_data <- factor(factor_data, levels = c("East", "West", "North"))

print(new_order_data)

summary(new_order_data)

#[1] East West North West West North West West #Levels: East West North

#East West North # 1 5 2

#5

#5.a setwd("C:/Users/ASITSD/Documents/CS101/Rworksheet#4_Tupaz/Rworksheet#4a") # Using forward slashes

#5.b data <- read.table("import_march.csv", header = TRUE, sep = ",")

print(head(data))

Students Strategy.1 Strategy.2 Strategy.3

#1 Male 8 10 8 #2 4 8 6 #3 0 6 4 #4 Female 14 4 15 #5 10 2 12 #6 6 0 9