BOX plot with outliers

ggplot(length\_of\_hospital\_stay, aes(x = SUBJECT, y = Value, fill = SUBJECT)) +

geom\_boxplot(outlier.color = "red", outlier.shape = 16, outlier.size = 2) +

theme\_minimal() +

labs(title = "Boxplot with Outliers", x = "Care Condition", y = "Length of Stay (days)") +

scale\_fill\_manual(values = c("lightblue", "lightpink")) +

expand\_limits(y = c(0, max(length\_of\_hospital\_stay$Value) + 2) # Extend y-axis upwards

Histogram with overlay curve

hist\_data <- hist(length\_of\_hospital\_stay$Value,

plot = FALSE) # Keep as frequencies (default behavior)

# Extend the y-axis by adding space above the tallest bar

y\_max <- max(hist\_data$counts) +40 # Adjust "+ 5" to add more space

# Plot the histogram with frequency counts

hist(length\_of\_hospital\_stay$Value,

main = "Histogram",

xlab = "Length of Stay",

col = "lightblue",

border = "black",

ylim = c(0, y\_max))

# Calculate mean and standard deviation

mean\_val <- mean(length\_of\_hospital\_stay$Value)

sd\_val <- sd(length\_of\_hospital\_stay$Value)

# Overlay the normal curve scaled to frequency

curve(dnorm(x, mean = mean\_val, sd = sd\_val) \* length(length\_of\_hospital\_stay$Value) \* diff(hist\_data$breaks)[1],

col = "red",

lwd = 2,

add = TRUE)

T test

t\_test\_result <- t.test(length\_of\_hospital\_stay$Value ~ length\_of\_hospital\_stay$SUBJECT)

print(t\_test\_result)