Lab Expt 1:

# Combining and Merging datasets

df1 <- data.frame(ID = 1:5, Name = c("A", "B", "C", "D", "E"))

df2 <- data.frame(ID = 3:7, Age = c(25, 28, 22, 30, 26))

# Merging on ID

merged\_df <- merge(df1, df2, by = "ID", all = TRUE)

print(merged\_df)

# Reshaping data (pivoting)

library(tidyr)

long\_df <- gather(merged\_df, key = "Variable", value = "Value", -ID)

print(long\_df)

# Pivoting back to wide format

wide\_df <- spread(long\_df, key = "Variable", value = "Value")

print(wide\_df)

Lab Expt 2

# String Manipulation

strings <- c("Data\_Analysis", "Computational\_Statistics")

library(stringr)

# Replace underscores with spaces

strings <- str\_replace\_all(strings, "\_", " ")

print(strings)

# Regular Expressions

emails <- c("user1@example.com", "user2@gmail.com")

valid\_emails <- grep("@example.com", emails, value = TRUE)

print(valid\_emails)

Lab expt 4

data <- c(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)

# Central Tendency

mean\_value <- mean(data)

median\_value <- median(data)

mode\_value <- as.numeric(names(sort(table(data), decreasing=TRUE)[1]))

# Dispersion Measures

sd\_value <- sd(data)

var\_value <- var(data)

mad\_value <- mad(data)

quartile\_deviation <- IQR(data) / 2

list(mean = mean\_value, median = median\_value, mode = mode\_value,

sd = sd\_value, variance = var\_value, MAD = mad\_value, quartile\_dev = quartile\_deviation)