# Sumithra Class 3 Assignment - DEG Classification (Modified Thresholds)

# Load required packages

library(dplyr)

# Define a function to classify DEGs with stricter criteria

classify\_DEGs <- function(data) {

data <- data %>%

mutate(

DEG\_class = case\_when(

padj < 0.01 & logFC > 1.5 ~ "Highly Upregulated",

padj < 0.01 & logFC < -1.5 ~ "Highly Downregulated",

padj < 0.05 & abs(logFC) > 1 ~ "Moderately Significant",

TRUE ~ "Not Significant"

)

)

return(data)

}

# Files to process

files <- c("DEGs\_Data\_1.csv", "DEGs\_Data\_2.csv")

# For-loop to process both files automatically

for (f in files) {

# Read CSV

df <- read.csv(f)

# Apply classification function

classified\_df <- classify\_DEGs(df)

# Save output with new naming convention

output\_name <- paste0("classified\_", gsub(".csv", "\_class3.csv", f))

write.csv(classified\_df, output\_name, row.names = FALSE)

cat("Processed:", f, " -> Saved as:", output\_name, "\n")

}

# Class 3 Assignment – DEG Classification (Sumithra)

## Task

- Write a function to classify DEGs.

- Apply stricter thresholds for classification.

- Process two input files automatically with a for-loop.

## Classification Rules

- \*\*Highly Upregulated\*\*: padj < 0.01 and logFC > 1.5

- \*\*Highly Downregulated\*\*: padj < 0.01 and logFC < -1.5

- \*\*Moderately Significant\*\*: padj < 0.05 and |logFC| > 1

- \*\*Not Significant\*\*: All other cases

## Input Files

- `DEGs\_Data\_1.csv`

- `DEGs\_Data\_2.csv`

## Output Files

- `classified\_DEGs\_Data\_1\_class3.csv`

- `classified\_DEGs\_Data\_2\_class3.csv`

## How to Run

1. Open `Sumithra\_class3\_Assignment.R` in RStudio.

2. Ensure `DEGs\_Data\_1.csv` and `DEGs\_Data\_2.csv` are in the working directory.

3. Run the script.

4. Outputs will be generated automatically in the same folder.