Advanced programming codes

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import numpy as np

data = {

'gene': ['Gene1', 'Gene2', 'Gene3', 'Gene4', 'Gene5'],

'logFC': [2.5, -1.8, 0.5, -2.1, 1.2], # log fold change

'pvalue': [0.001, 0.05, 0.2, 0.0001, 0.01] # p-values

}

df = pd.DataFrame(data)

df['neg\_log\_pvalue'] = -np.log10(df['pvalue'])

# Create a Volcano plot

plt.figure(figsize=(8,6))

plt.scatter(df['logFC'], df['neg\_log\_pvalue'], c=(df['pvalue'] < 0.05) & (abs(df['logFC']) > 1), cmap='coolwarm', edgecolors='k')

plt.title('Volcano Plot')

plt.xlabel('Log2 Fold Change')

plt.ylabel('-Log10 P-Value')

# Highlight significant points

plt.axhline(y=-np.log10(0.05), color='black', linestyle='--') # p-value cutoff

plt.axvline(x=1, color='gray', linestyle='--') # fold change cutoff

plt.axvline(x=-1, color='gray', linestyle='--') # fold change cutoff

plt.colorbar(label='Significance')

plt.show()

import matplotlib.pyplot as plt

import numpy as np

import pandas as pd

# Generate some mock data for demonstration

np.random.seed(0)

n = 100 # Number of data points

log2\_fold\_changes = np.random.randn(n) # Random log2 fold changes

p\_values = np.random.uniform(0, 1, n) # Random p-values

# Create a DataFrame

df = pd.DataFrame({

'log2FoldChange': log2\_fold\_changes,

'pValue': p\_values

})

# Calculate -log10(pValue) for the y-axis

df['negLog10PValue'] = -np.log10(df['pValue'])

# Set a threshold for significance

significance\_threshold = 0.05 # p-value threshold for significance

df['significant'] = df['pValue'] < significance\_threshold

# Create the volcano plot

plt.figure(figsize=(10, 6))

plt.scatter(df['log2FoldChange'], df['negLog10PValue'],

c=df['significant'], cmap='coolwarm', alpha=0.7)

plt.title("Volcano Plot for Alzheimer's Disease")

plt.xlabel("Log2 Fold Change")

plt.ylabel("-Log10 P-Value")

# Highlight significant points (in red)

plt.axhline(y=-np.log10(significance\_threshold), color='black', linestyle='--')

plt.axvline(x=0, color='black', linestyle='--')

# Show the plot

plt.colorbar(label='Significance')

plt.show()

#R CODE

library(dplyr)

alzheimers\_genes <- c("APP", "PSEN1", "PSEN2", "MAPT", "APOE", "CLU", "BIN1")

schizophrenia\_genes <- c("GRIN2A", "ZNF804A", "CACNA1C", "APOE", "COMT", "NRG1", "BIN1")

common\_genes <- intersect(alzheimers\_genes, schizophrenia\_genes)

print("Common genes:")

print(common\_genes)

write.table(common\_genes, file = "common\_genes.txt", row.names = FALSE, col.names = FALSE, quote = FALSE)