LAB Logbook

Lab 1

Code-

import numpy as np

# Last two digits of SID

last\_two\_digits = 71

# Create a vector with np.arange of size 71

vector = np.arange(last\_two\_digits)

# Convert the vector to a 2D array with 1 row

matrix\_2d = vector.reshape(1, -1)

# Print the vector

print("1D Vector:")

print(vector)

# Print the 2D array

print("\n2D Array with 1 row:")

print(matrix\_2d)

# Check the shape of the 2D array

print("\nShape of the 2D Array:")

print(matrix\_2d.shape)

# Check the dimession of the array

print("\nDimession of the Array:")

print(matrix\_2d.ndim)

Output-

A screenshot of a computer

Description automatically generatedA white rectangular object with black lines

Description automatically generated with medium confidence

Lab 2

Code-

import pandas as pd

# Load the DataFrame (replace this with the path to your file)

df = pd.read\_csv('adult\_data\_mini.csv')

# Step 1: Set n (last digit of SID) to 1

n = 1

# Step 2: Group by "relationship" and "hours-per-week" before reduction

grouped\_original = df.groupby(['relationship', 'hours-per-week']).size().reset\_index(name='count')

print("Grouped by relationship and hours-per-week (before reduction):")

print(grouped\_original)

# Step 3: Reduce "hours-per-week" values by n (1)

df['hours-per-week'] = df['hours-per-week'] - n

# Step 4: Group by "relationship" and reduced "hours-per-week"

grouped\_reduced = df.groupby(['relationship', 'hours-per-week']).size().reset\_index(name='count')

print("\nGrouped by relationship and reduced hours-per-week:")

print(grouped\_reduced)

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

Lab 3

Code-

pip install pandas seaborn matplotlib

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

# Load the CSV file (update the file path if necessary)

file\_path = 'C:/Users/falgu/Downloads/telecom\_churn.csv'

telecom\_data = pd.read\_csv(file\_path)

# Create a bicolour scatter plot using 'Account length' and 'Total day minutes'

plt.figure(figsize=(9, 7))

sns.scatterplot(data=telecom\_data, x='Account length', y='Total day minutes', hue='Churn', palette='coolwarm')

# Add labels and a title to the plot

plt.title('Interaction between Account Length and Total Day Minutes', fontsize=14)

plt.xlabel('Account Length', fontsize=12)

plt.ylabel('Total Day Minutes', fontsize=12)

# Display the plot

plt.show()

A screenshot of a computer

Description automatically generated

A screen shot of a graph

Description automatically generated

Lab 4

Lab 5

Lab 6

Lab 7

Lab 8

Lab 9

Lab 10

Lab 11

Lab 12