TDI PYTHON FOR DATA ANALYTICS WEEK 3: WORKING WITH DATA

Welcome to Python Week 3!

In this project, you'll advance your skills by learning how to handle files, read and write CSV files, and dive into NumPy, a powerful library for numerical computing. This week's assignment will focus on the following key topics:

Topics Covered:

- File Handling
  - Reading from and writing to files
  - Working with CSV files using the “csv” module
- Introduction to NumPy
  - Understanding NumPy and its importance
  - Creating and manipulating NumPy arrays
  - Basic array operations

Dataset: Titanic Analysis

We will use the Titanic dataset you uploaded. This dataset contains information about passengers, such as age, class, fare, and survival status. You’ll be working with this data to understand basic file operations and analysis using NumPy.

Learning Resources:

Understanding File Handling in Python

Working with text data

Introduction to NumPy

Understanding Numpy

Questions:

File Handling & CSVs:

1. Reading a Text File:

Write a Python script that opens a file in read mode and prints its contents. You can use any text file for practice.

2. Writing to a File:

Create a Python script that writes the string "Hello, Data Analysis!" to a new text file.

3. Reading the Titanic Dataset (CSV):

Use the “csv” module to read the Titanic dataset. Print the first 5 rows of the dataset.

4. Counting Records:

After reading the Titanic CSV file, count the total number of passengers.

5. Calculating Average Fare:

Write a script that calculates the average fare paid by passengers. Use the Titanic CSV data for this calculation.

6. Survivors by Class:

Write a Python script that calculates the number of survivors per passenger class (1st, 2nd, and 3rd class).

Introduction to NumPy:

7. NumPy Basics:

Import the NumPy library and create a 1D array containing the following numbers: “[10, 15, 20, 25]”. Print the array.

8. Array Operations:

Using the array from the previous question, perform the following operations:

- Multiply each element by 2.

- Subtract 5 from each element.

9. 2D Array Creation:

Create a 2D NumPy array with shape (3, 3) using the following data:

“[[1, 2, 3], [4, 5, 6], [7, 8, 9]]”

Print the array.

10. Titanic Data with NumPy:

Convert the "Fare" column of the Titanic dataset to a NumPy array and calculate:

- The minimum fare

- The maximum fare

- The mean fare

11. Array Indexing:

Use array indexing to extract the first row and first column of the 2D array created earlier. Print the results.

12. Random Array:

Create a NumPy array with 10 random numbers between 0 and 1. Print the array and find the maximum value.

13. Practice: Survivor Analysis with NumPy:

- Use the Titanic dataset and convert the "Survived" column to a NumPy array.

- Calculate the percentage of passengers who survived.

Submission Instructions:

Submit your Python scripts or Jupyter notebook with answers to these tasks on our designated platform. Be sure your code is clean, well-documented, and easy to follow.