Data File Comparator

---------------------------------

Overview

The Data File Comparator is a Python script that reads, compares, and processes data from two input files, and writes the results to an output file. It supports numeric, string, and date data types. This script can be extended to support additional data sources, such as SQL databases, and is able to modify the comparison method in the main function for future extensions.

---------------------------------

Features

- Data Types Supported: Numeric, String, Date

- Comparison:

- Numeric: Subtracts values from the second file from values in the first file

- String: Outputs "T" if values are the same, "F" if different

- Date: Calculates the difference in months between dates

- Extensible: Easily add support for more data sources in the future

---------------------------------

Installation

1. Ensure you have Python 3.x installed on your system.

2. Clone this repository or download the script.

---------------------------------

Usage

1. Prepare your input files:

- Each file should start with a header that indicates the data type (e.g., Num\_, Str\_, Date\_).

- Each line after the header should contain a single data entry of the specified type.

2. Place your input files (INPUT\_File1.txt and INPUT\_File2.txt) in the same directory as the script or update the file paths in the script as needed.

3. Run the script:

python data\_file\_comparato.py

---------------------------------

Example

Input Files

INPUT\_File1.txt

Num\_ Val1

10

20

30

INPUT\_File2.txt

Num\_ Val2

5

10.5

40

Output File

output\_file.txt

CompareResult

5

9.5

‐10

---------------------------------

Testing with String and Date Comparison

To test the script with string or date comparison, you can use the provided input files:

- For string comparison, use `INPUT\_File3.txt` and `INPUT\_File4.txt`.

- For date comparison, use `INPUT\_File5.txt` and `INPUT\_File6.txt`.

To do this, change the following lines in the script:

# For string comparison

input\_file1\_path = os.path.join(current\_directory, 'INPUT\_File3.txt')

input\_file2\_path = os.path.join(current\_directory, 'INPUT\_File4.txt')

# For date comparison

input\_file1\_path = os.path.join(current\_directory, 'INPUT\_File5.txt')

input\_file2\_path = os.path.join(current\_directory, 'INPUT\_File6.txt')

---------------------------------

Example

Input Files

INPUT\_File3.txt

Str\_state1

California

Florida

Texas

INPUT\_File4.txt

Str\_state2

CA

Forida

Texas

Output File

output\_file.txt

CompareResult

F

F

T

---------------------------------

Example

Input Files

INPUT\_File5.txt

Date\_birthday1

1/5/2000

2/9/1995

11/15/1970

INPUT\_File6.txt

Date\_birthday2

2/1/2000

3/1/1994

7/1/1980

Output File

output\_file.txt

CompareResult

-1

11

-116

---------------------------------

Code Description

Importing Required Modules

from datetime import datetime

import os

DataFile Class

The DataFile class handles reading from and writing to files.

- Constructor: Initializes the file path.

- readFile Method: Reads the file, determines the data type from the header, and processes the data accordingly.

- writeFile Method: Writes the results to the file.

Main Function

- compareFiles Function: Compares the two input files based on their data type and returns the comparison results.

- writeFile Function: Writes the comparison results to the output file.

---------------------------------

Script Execution

The script reads the input files, compares the data, and writes the results to the output file.

---------------------------------

Extending the Script

To add support for more data sources (e.g., SQL databases), you can extend the DataFile class and modify the readFile and writeFile methods accordingly.