**Data Engineering Assignment – 10**

**NESTED JSON TO CSV CONVERSION:**

The first step is to read the JSON file as a python dict object. The read\_json() function is used for the task, which taken the file path along with the extension as a parameter and returns the contents of the JSON file as a python dict object.

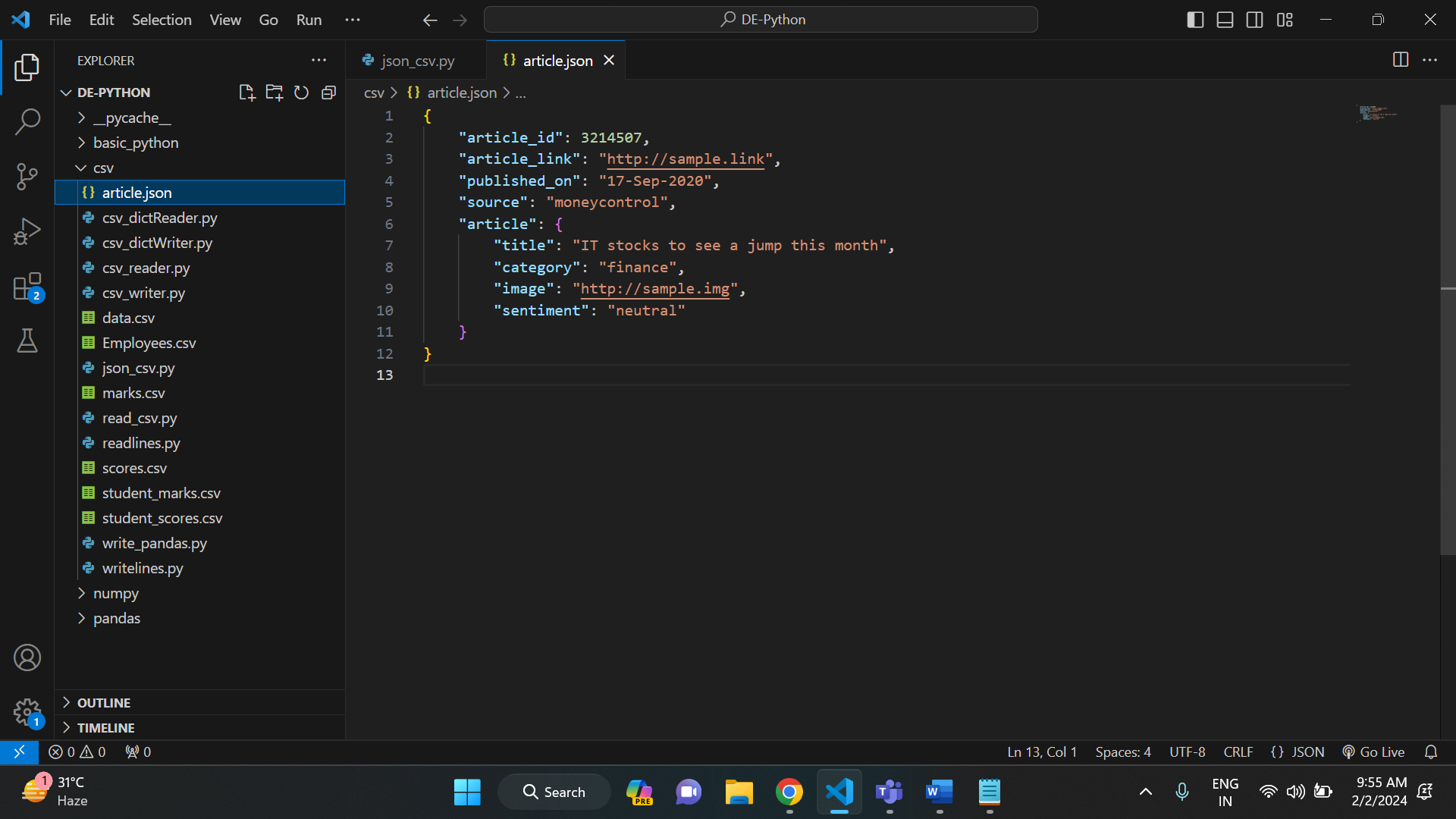
We normalize the dict object using the normalize\_json() function. It checks for the key-value pairs in the dict object.

The desired CSV data is created using the generate\_csv\_data() function. This function concatenates each record using a comma (,).

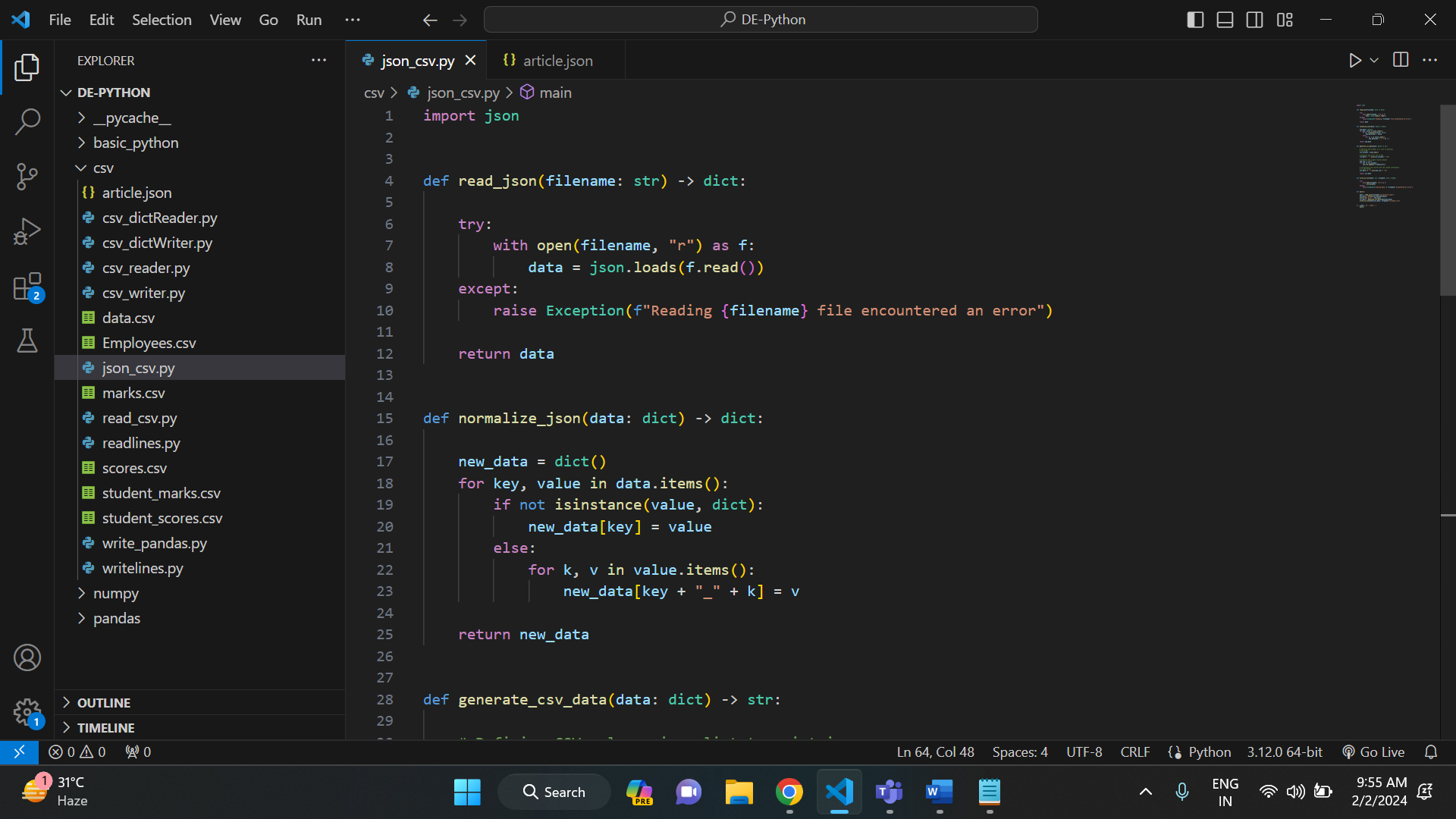
In the final step, we write the CSV data generated in the earlier step to a preferred location provided through the filepath parameter.

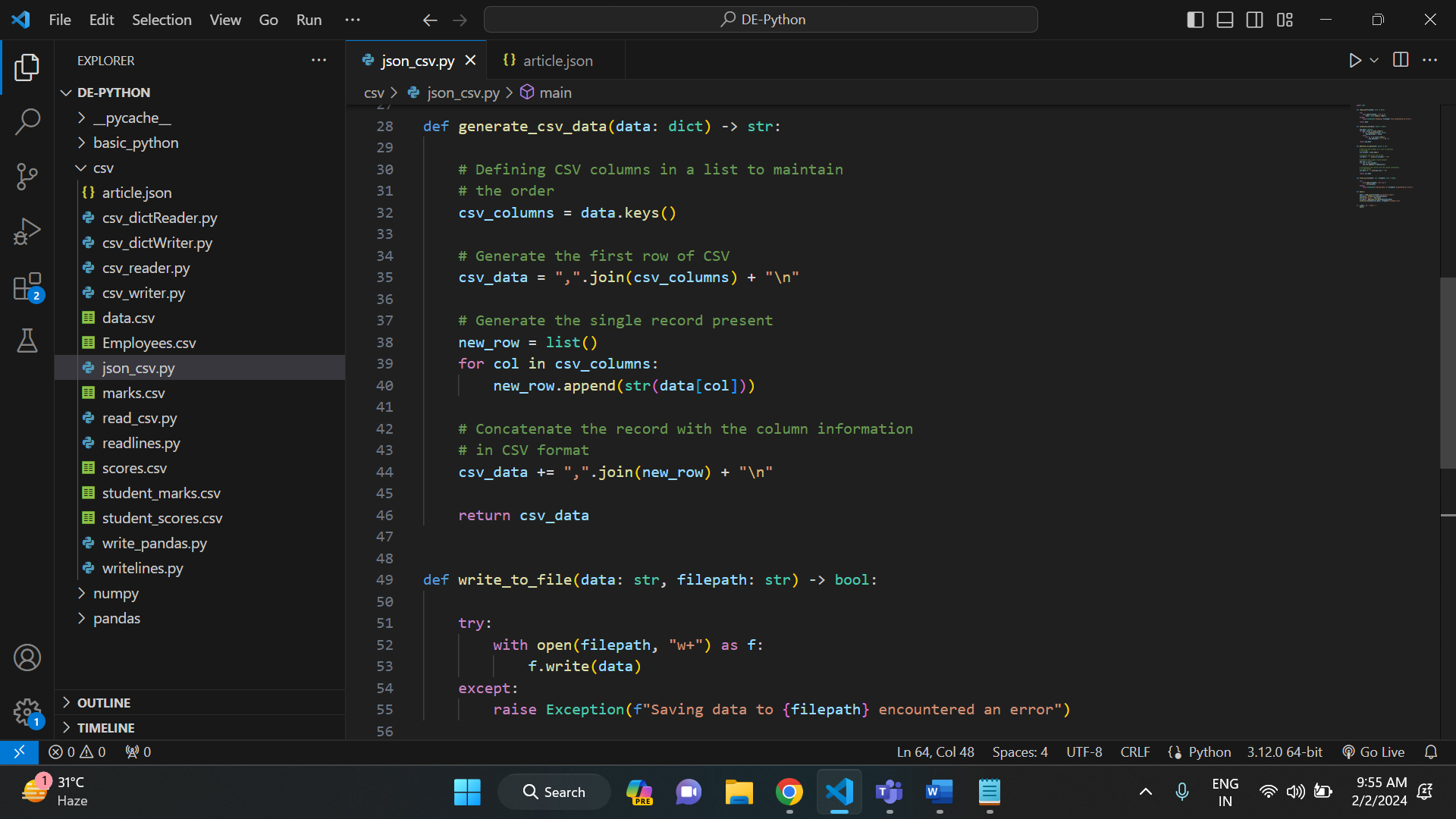
I have used article. Json file which is a nested json for the above program.

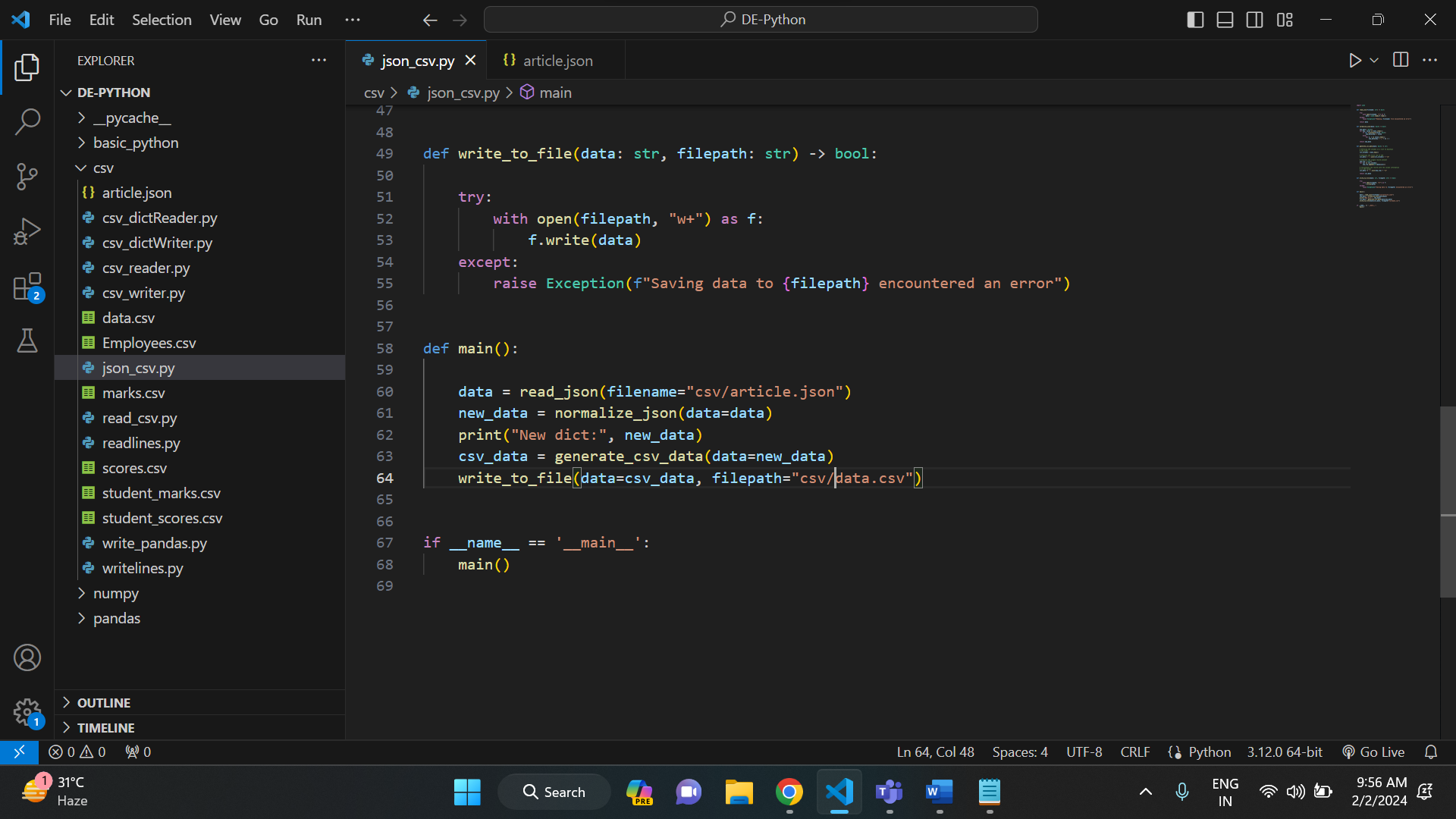
ARTICLE. JSON:



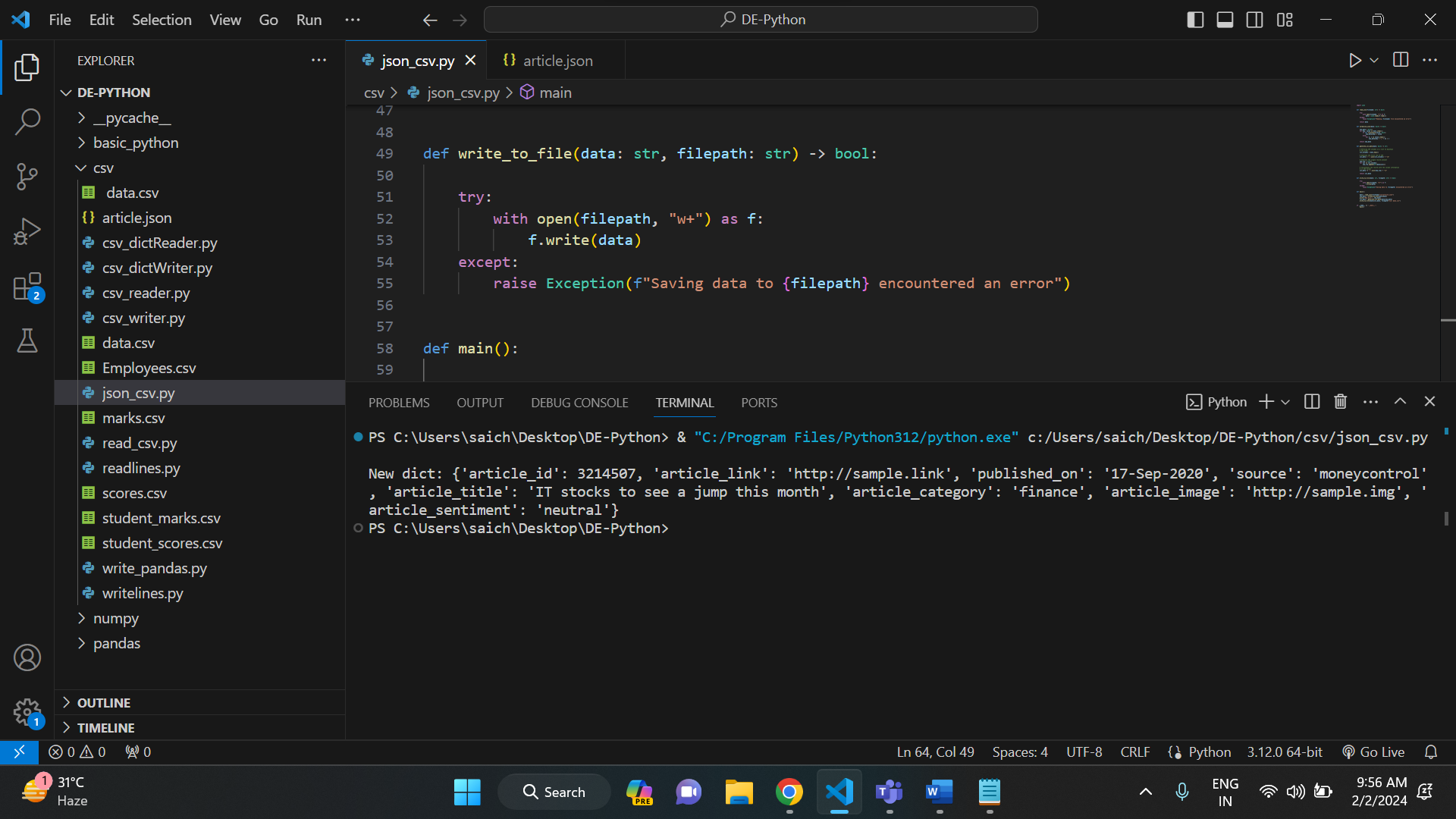
CODE FOR CONVERSION:



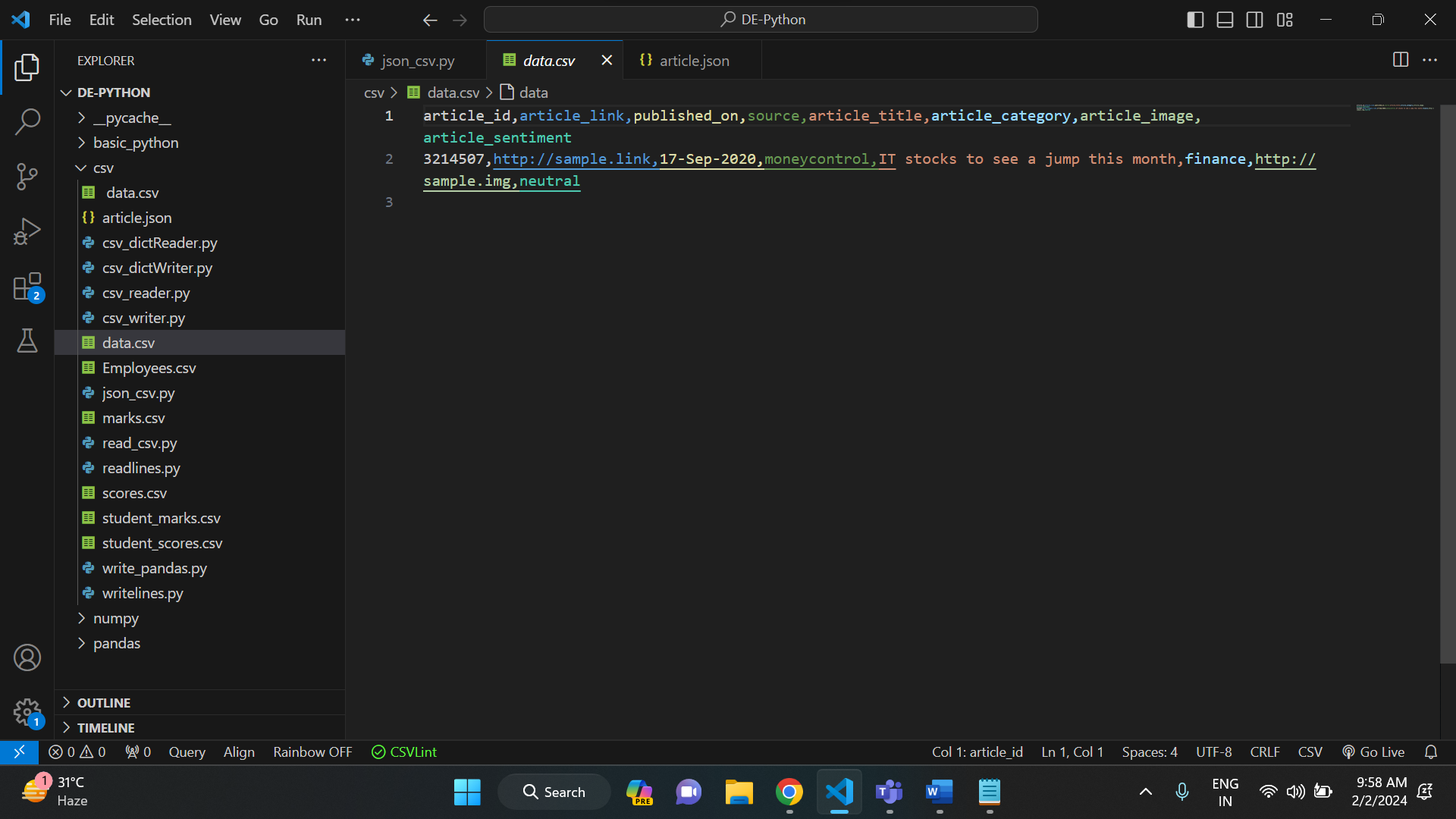




OUTPUT:



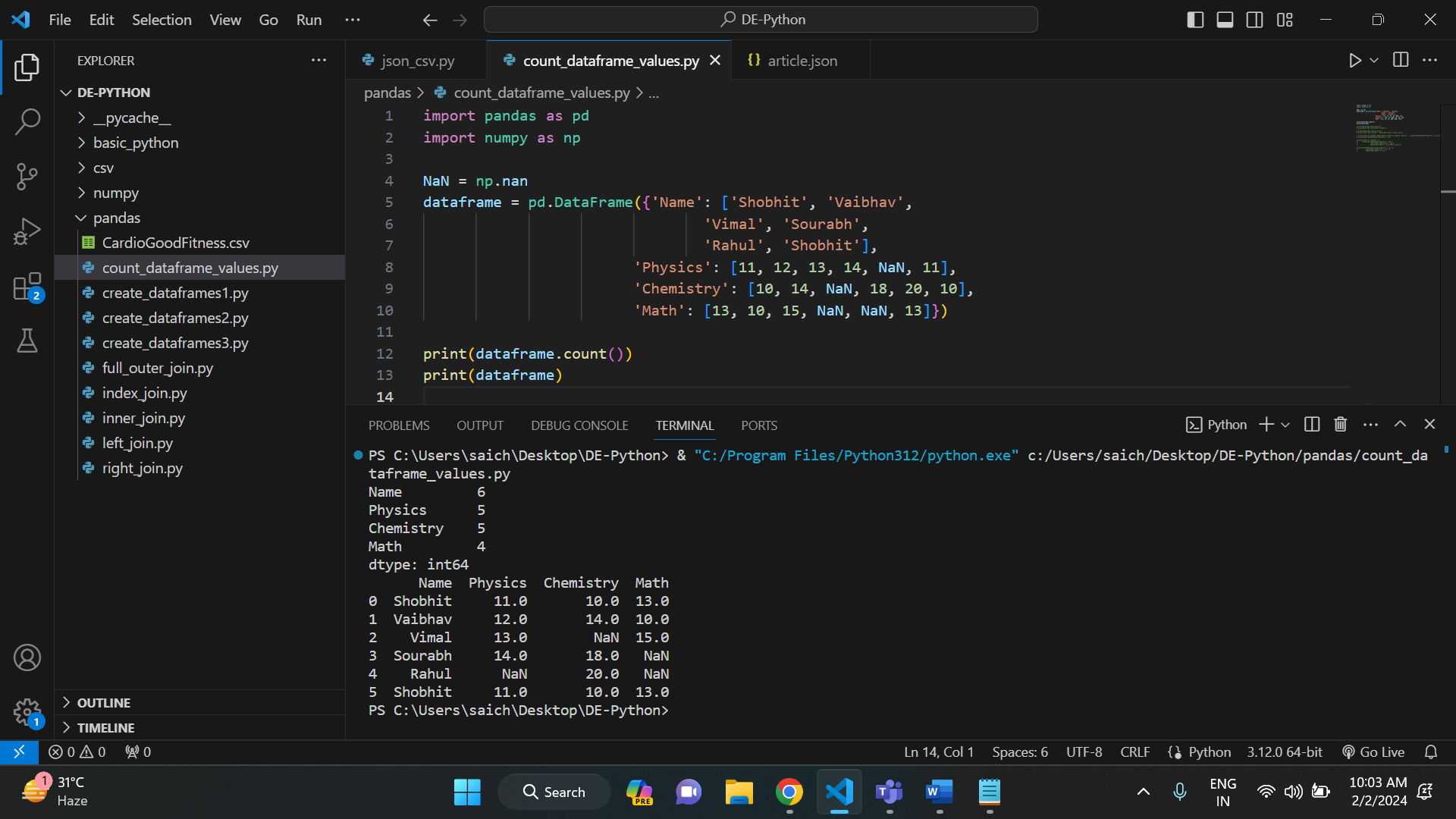
And below is the created data. csv file.



**COUNTING VALUES IN PANDAS DATAFRAME:**

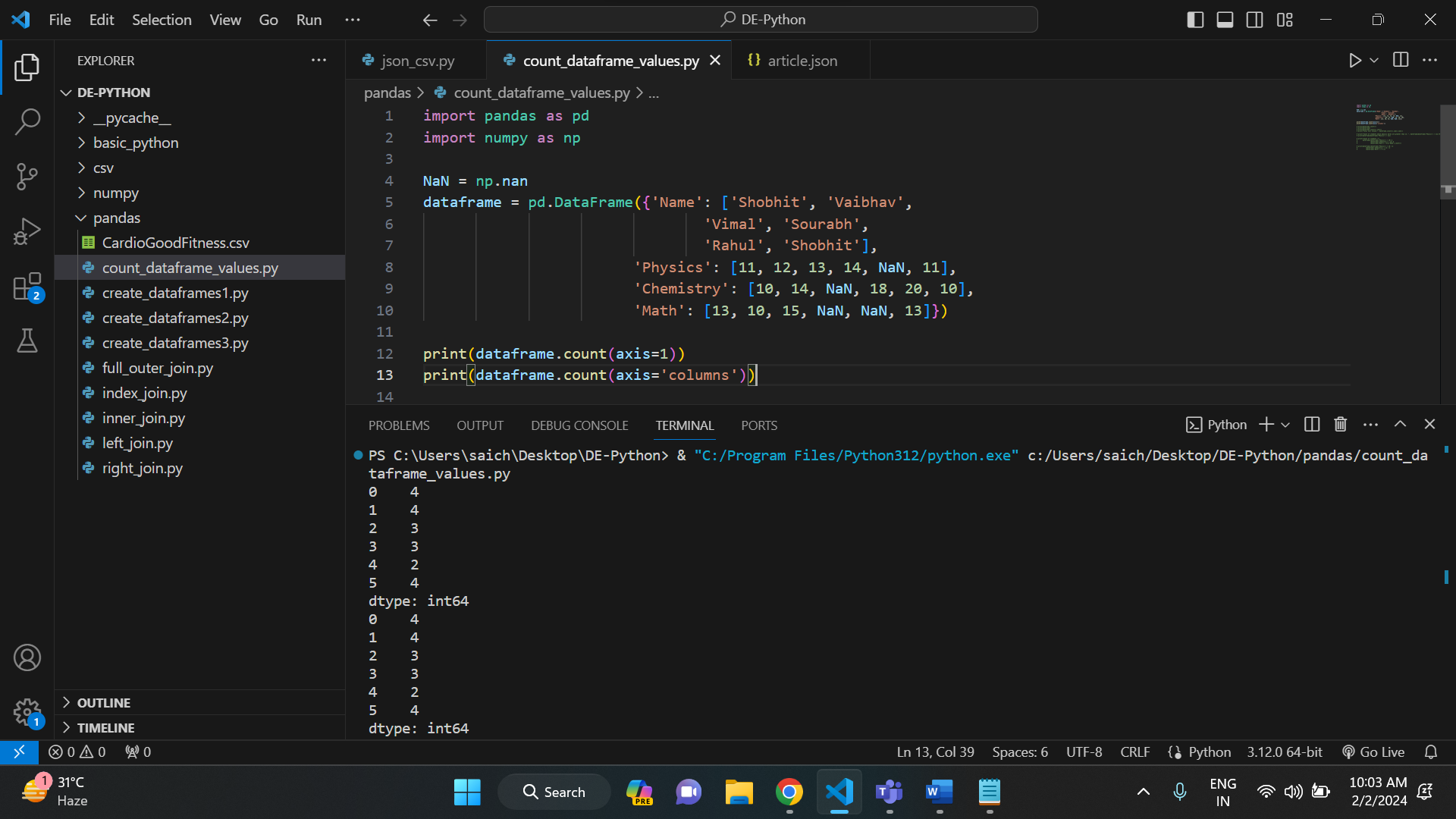
We have students and their marks in different subjects. The dataframe also contains null values. And below is the program to count the values within the dataframe.

Example: 1



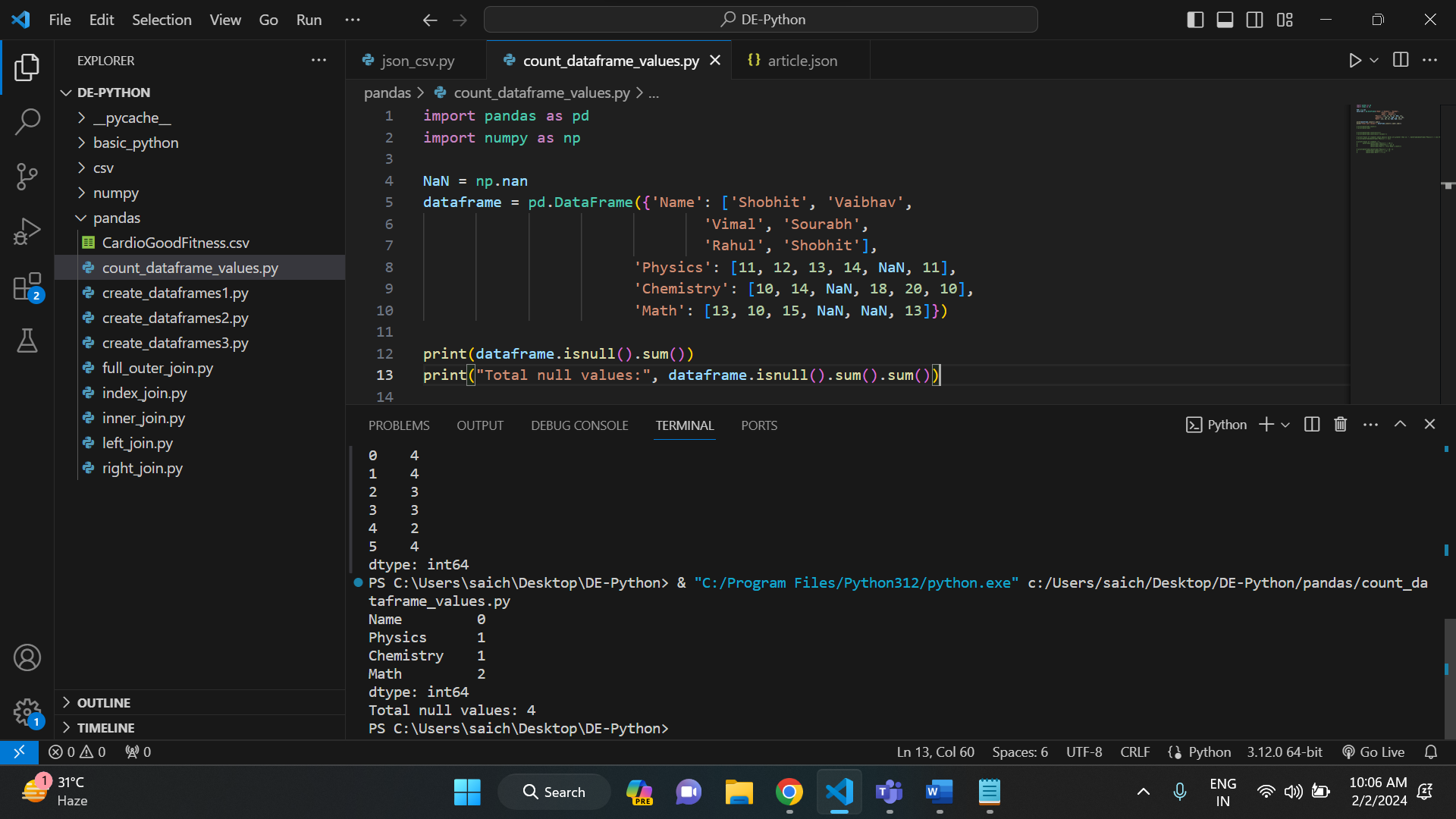
Example: 2

In this example, it finds count of values in the row.



Example: 3

In this example, I am trying to find the count null values in our dataframe.



Example: 4

In this example, I want to count no of students whose physics marks are greater than 11.



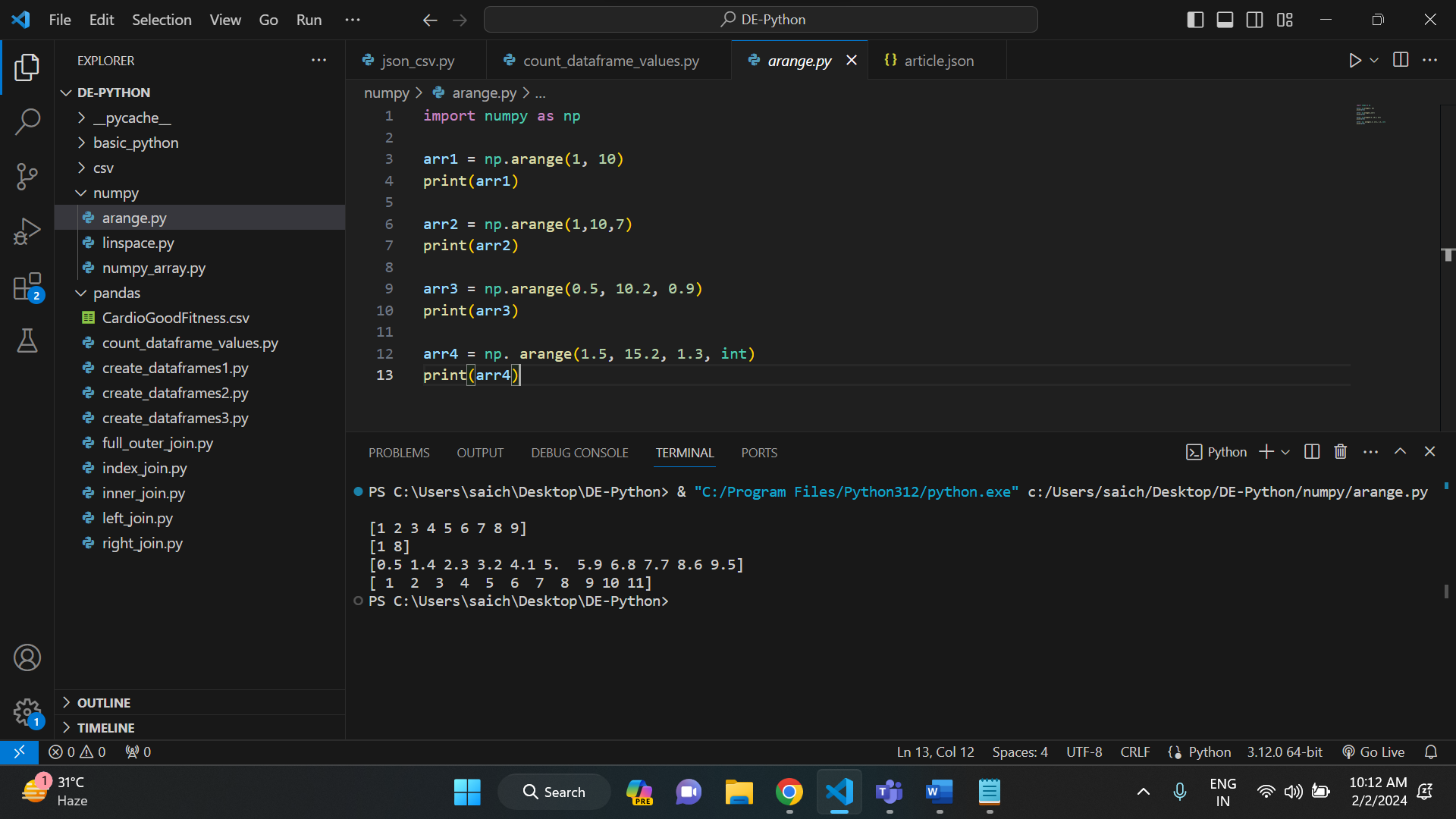
**CREATING ARRAYS WITH EVENLY SPACED VALUES:**

* **ARANGE ():**

arange returns evenly spaced values within a given interval.

Syntax: arange([start,] stop[, step], [, dtype=None])

Example:

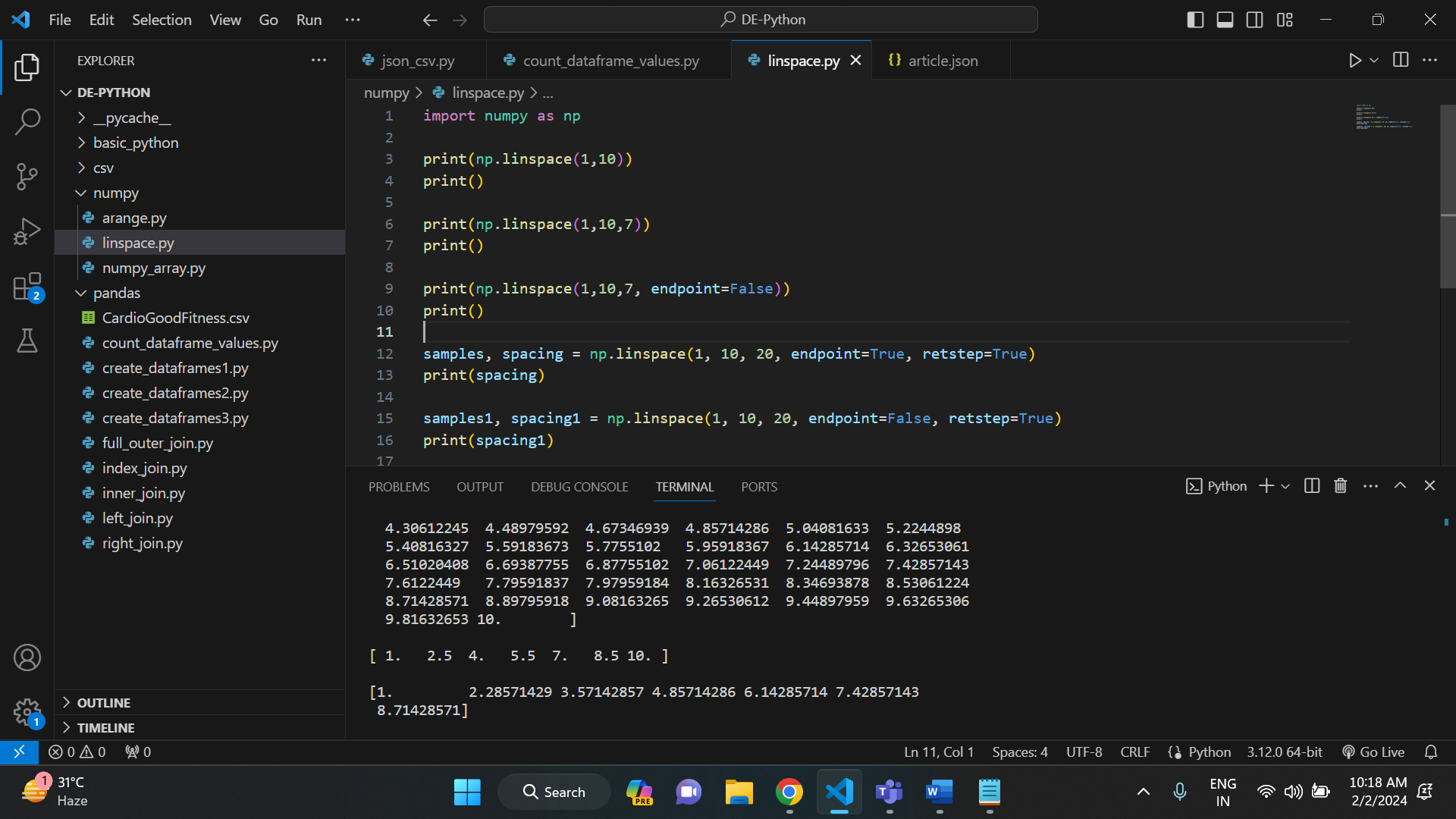
****

* **LINSPACE ():**

Linspace returns a ndarray, consisting of 'num' equally spaced samples in the closed interval [start, stop] or the half-open interval [start, stop).

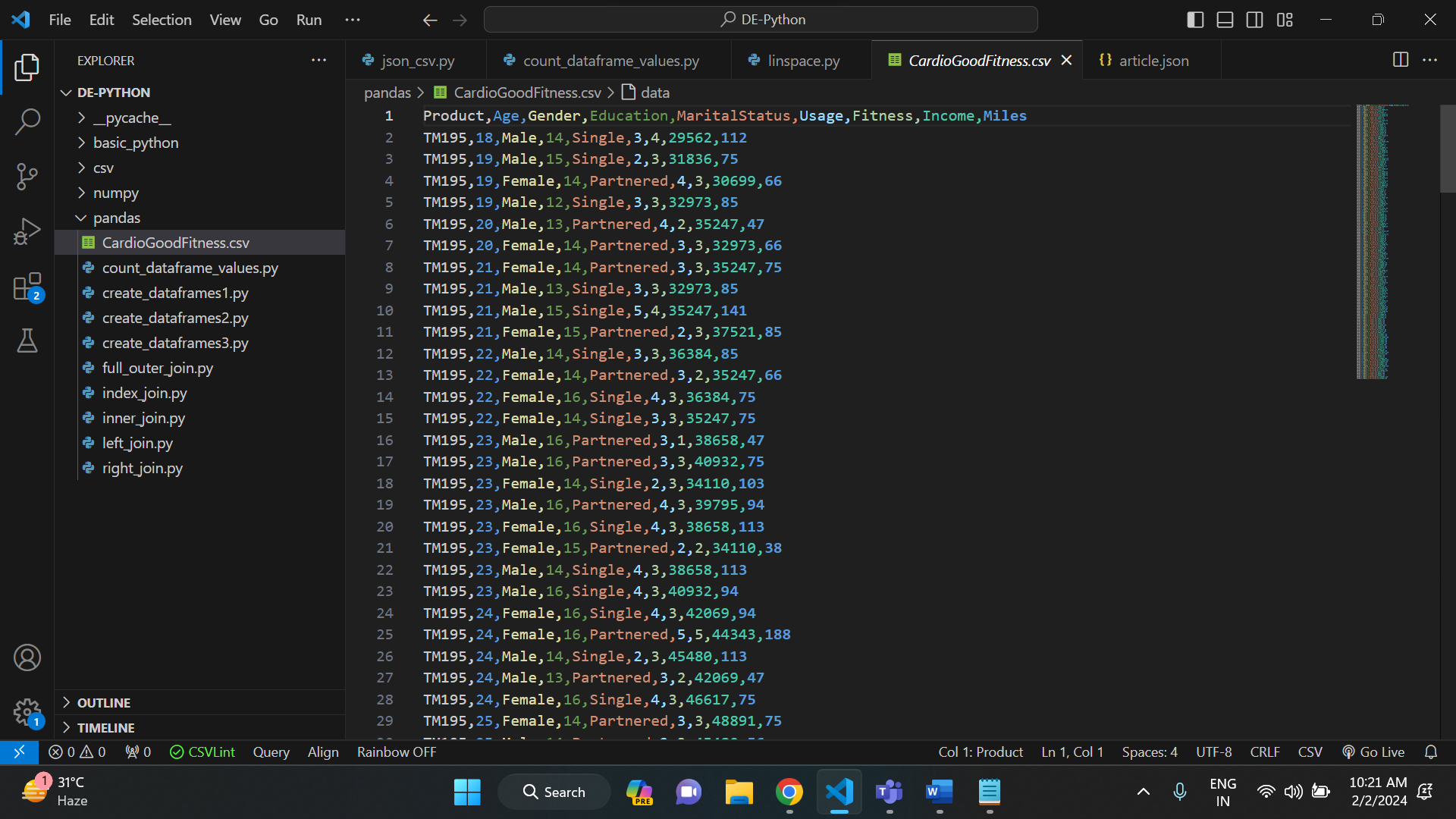
Syntax: linspace(start, stop, num=50, endpoint=True, retstep=False)

Example:



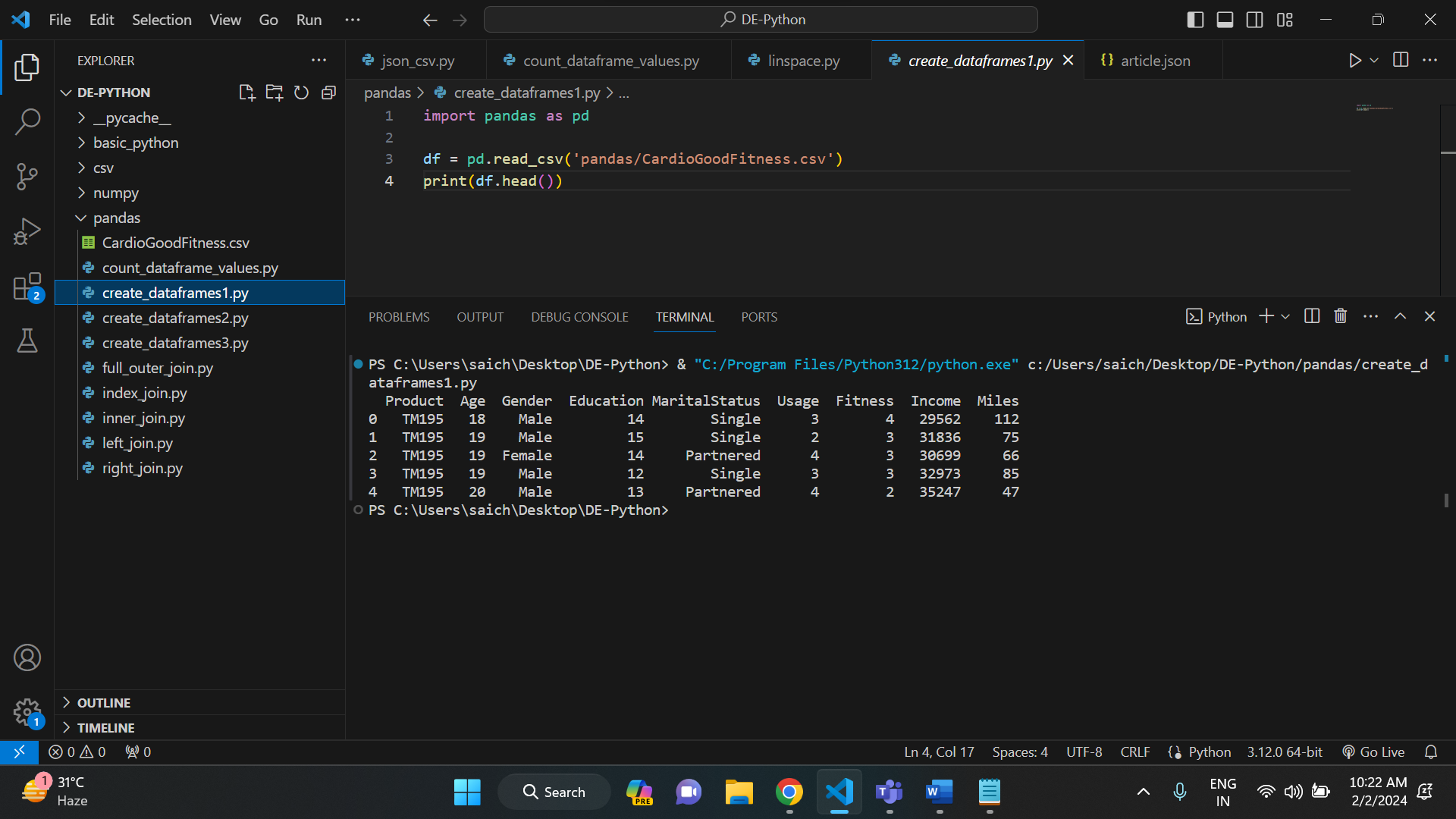
**CREATING DATAFRAMES IN PYTHON USING PANDAS:**

There are different ways of creating data frames using pandas. For this topic, I have use CardioGoodFitness. Csv file. And below is that file.

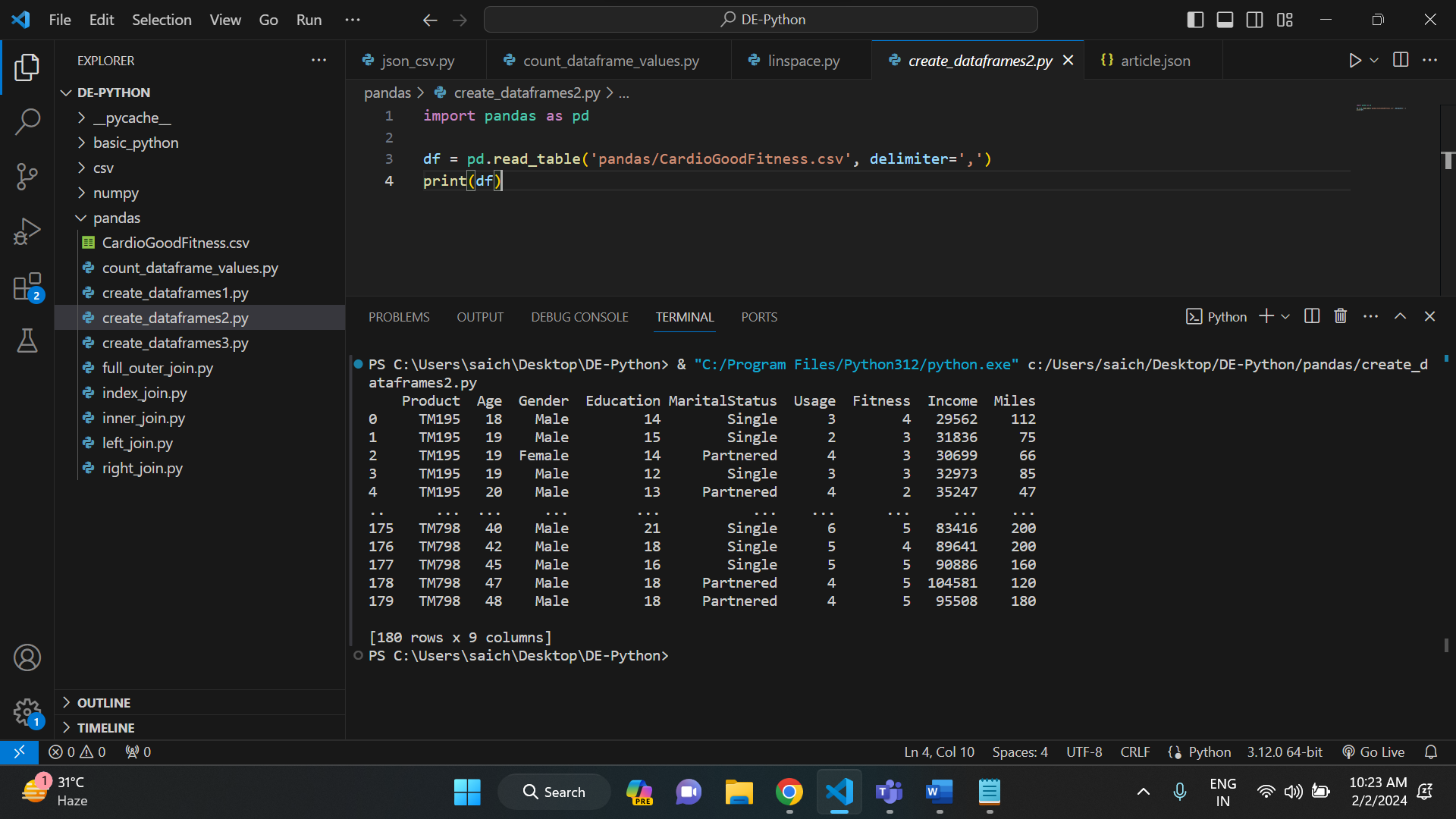


Now, the methods to create data frames are:

* Read\_csv ():



* Read\_table ():



* Using csv module:

