

The Task

You are given two data sheets:

1. DummyTransactions.csv
2. Currency2Year.csv

In the DummyTransactions.csv file, take note of the following 3 columns:

A	F	G	H	I
	Transactions			
Transaction Date	Transaction amount	Transaction ID	Supplier city (optional)	Currency
1/1/21	15280	1	Newcastle	AUD
1/1/21	3654	2	Singapore	SGD
6/1/21	7387	3		SGD
10/1/21	446	4		SGD
10/1/21	3177	5		SGD
14/1/21	2049	6	Singapore	SGD
14/1/21	1036	7	Singapore	SGD
15/1/21	5956	8		SGD

Each row is a transaction, with the amount of money spent on that transaction, in a particular currency, on a particular date.

In the *Currency2Year.csv* file, you will find each row to be the currency conversion factor from the specific currency to USD, on a particular date.

Date	Currency	Factor
4/5/20	AED	0.272257826
5/5/20	AED	0.272242261
6/5/20	AED	0.272249673
7/5/20	AED	0.272253453
8/5/20	AED	0.272242261
9/5/20	AED	0.272245892

(amount_in_currency * **Factor** = amount_in_USD)

Your goal is to convert all prices in the first datasheet under *Transaction amount* into USD prices.

Note that each transaction occurs on a particular date, so you will have to find the historical USD conversion factor that is closest to the transaction date, and apply that.

Here is what I mean by being “closest to the transaction date”:

Let’s define the date of a particular transaction as **T**.

We find a series of rows in the second datasheet where there are USD conversion factors of said transaction's currency. We define the dates of these rows to be $t_0, t_1, t_2, t_3, t_4 \dots t_n$ where $t_{n-1} < t_n$

T is closest to t_x if $t_x \leq T < t_{x+1}$, and we should use the conversion factor on date t_x to compute the said transaction's USD amount.

The output of your program should be a modified *DummyTransactions.csv*, where the **Transaction amount** column gets converted to their corresponding USD values, and as a courtesy please change all values in the **Currency** column to 'USD'.

Grading

The correctness of the output will give you half the scores.

I am interested in how you optimise your code and make the program run as fast as possible, which contributes to another half of the scores.

Other Requirements

- Use Python3
- You are **allowed** to use *pandas* and/or *numpy*
- Write your code in a git repository so I can see your commit history.
- Submit your source code folder in a zip file.

Best of luck!