Internsip | Summer 2024

IMPORTANT

You can use this **Jupyter** notebook *symbols_update.ipynb* to complete the assignment.

If you are not familiar with **Jupyter**, please feel free to develop your solution using standard **.py** file instead.

To complete the assignment, you will need to **pip install pandas** library.

Please upload all your work to public **GitHub** repository and share the link with us.

There are three .csv files that we want to upload sequentially to the database: **symbols_update_1.csv**, **symbols_update_2.csv**, and **symbols_update_3.csv**. For the sake of simplicity, the database is represented by another .csv file **database.csv**.

The goal it to complete **SymbolsUpdate** class with three functions:

- *load_new_data_from_file* : loads and processes data from **symbols_update_n.csv** file
- *save_new_data* : saves loaded data to the **database.csv** file
- *get_data_from_database* : returns the most recently updated data for every symbol in **database.csv**

```python
class SymbolsUpdate(object):
    def __init__(self ):
        self.database_file = 'database.csv'

    def load_new_data_from_file(self, file_path: str):
        pass

    def save_new_data(self, input_data: pd.DataFrame):
        pass

    def get_data_from_database(self):
        pass
```

Running the model should return the following table:

```python
su = SymbolsUpdate()
new_data = su.load_new_data_from_file('symbols_update_1.csv')
su.save_new_data(new_data)
new_data = su.load_new_data_from_file('symbols_update_2.csv')
su.save_new_data(new_data)
new_data = su.load_new_data_from_file('symbols_update_3.csv')
su.save_new_data(new_data)
```

```python
from IPython.display import Image
Image('output.png')
```

Out[3]:

	symbol	hold	country	item	item_value	updatedby	updatetime
0	AAAA	0	US	cusip	A234AC	petroineos	2024-05-10 10:37:29.609021
1	AAAA	0	US	isin	US01222911	petroineos	2024-05-10 10:37:29.609021
2	BBBB	1	Great Britain	cusip	123998	petroineos	2024-05-10 10:37:29.585001
3	BBBB	1	Great Britain	isin	GB12222201	petroineos	2024-05-10 10:37:29.585001
4	CCCC	1	US	cusip	G129111	petroineos	2024-05-10 10:37:29.585001
5	CCCC	1	US	isin	US01239811	petroineos	2024-05-10 10:37:29.585001
6	DDDD	1	Italy	cusip	78321	petroineos	2024-05-10 10:37:29.562999
7	DDDD	1	Italy	isin	IT92812323	petroineos	2024-05-10 10:37:29.562999
8	GGGG	1	Belgium	cusip	B54334AC	petroineos	2024-05-10 10:37:29.585001
9	GGGG	1	Belgium	isin	BE012568156	petroineos	2024-05-10 10:37:29.585001

Please note:

- *isin* and *cusip* columns from **symbols_update_n.csv** files are stored as item/item_value pairs in **database.csv**
- *country id* is derived from the *isin* field (e.g. US01222911 is US, GB12222201 is GB, etc...)
- The most recent update for each symbol is returned based on the **updatetime** column

Good luck!!!