

Pickle

Pickle 은 사전적으로 절여서 저장해 놓는다는 말인데요. 파이썬에서 데이터를 저장해 놓을 때 쓰는 패키지입니다. 파이썬 언어로 만들어진 데이터는 RAM 메모리에 존재합니다. 따라서, 컴퓨터가 꺼지면 자동으로 데이터가 사라지게 됩니다. 그래서, 저는 pickle 를 이용해서 데이터 작업 중간에 데이터를 저장합니다. 파이썬 DataFrame 의 저장은 csv, excel, json 등 다양한 형식으로 저장할 수 있으나, 파이썬의 데이터 타입을 손상시키지 않고, 원형대로 저장하고 불러올 수 있는 pickle 이 제일 편리합니다. 삼성전자 일봉데이터를 가져와서 피클로 저장해 보겠습니다.

```
import FinanceDataReader as fdr

code = '005930' # 삼성전자
stock_data = fdr.DataReader(code, start='2021-01-03', end='2021-12-31')

stock_data.to_pickle('stock_data.pkl') # 디렉토리를 지정하지 않으면 현재 작업 폴더에 저장
이 됩니다.
```

이번에는 저장된 pickle 파일을 불러와 출력해 보겠습니다. read_pickle 을 이용하면 데이터가 손상되지 않고, 원형 그대로 복원되었음을 알 수 있습니다.

```
import pandas as pd
stock_data = pd.read_pickle('stock_data.pkl')
stock_data.head().style.set_table_attributes('style="font-size: 12px"')
```

	Open	High	Low	Close	Volume	Change
Date						
2021-01-04 00:00:00	81000	84400	80200	83000	38655276	0.024691
2021-01-05 00:00:00	81600	83900	81600	83900	35335669	0.010843
2021-01-06 00:00:00	83300	84500	82100	82200	42089013	-0.020262
2021-01-07 00:00:00	82800	84200	82700	82900	32644642	0.008516
2021-01-08 00:00:00	83300	90000	83000	88800	59013307	0.071170

pickle 모듈을 이용하여 binary 파일로 저장하는 것도 가능합니다. 특히 pickle 모듈로 파일을 저장하고 읽을 때는 저장하는 환경의 Pandas 버전과 읽는 환경의 Pandas 버전이 동일해야 에러가 발생하지 않습니다.

```
import pickle

with open('stock_data.pkl', 'wb') as file:    # Binary 파일로 저장
    pickle.dump(stock_data, file)

with open('stock_data.pkl', 'rb') as file:    # 저장된 binary 파일 읽기
    stock_data = pickle.load(file)
```

```
stock_data.head().style.set_table_attributes('style="font-size: 12px"')
```

	Open	High	Low	Close	Volume	Change
Date						
2021-01-04 00:00:00	81000	84400	80200	83000	38655276	0.024691
2021-01-05 00:00:00	81600	83900	81600	83900	35335669	0.010843
2021-01-06 00:00:00	83300	84500	82100	82200	42089013	-0.020262
2021-01-07 00:00:00	82800	84200	82700	82900	32644642	0.008516
2021-01-08 00:00:00	83300	90000	83000	88800	59013307	0.071170

