

Resample

Contents

- [Resample](#)
- [FinanceDataReader](#)

Resample 은 시간데이터를 다른 시간 단위로 변경하고 싶을 때 활용합니다. 예를 들면, 초 단위 데이터를 일단위 혹은 월단위 데이터로 변경 할 수 있습니다. 연습을 위하여 시간 레벨의 데이터가 필요합니다. 시간레벨 데이터는 FinanceDataReader 패키지에서 제공하는 일봉 데이터를 활용하겠습니다. FinanceDataReader 는 이승준님이 금융자료 분석을 하시는 분들을 위하여 만들어 주신 정말 유용한 패키지입니다. 자세한 내용은 아래 링크에 설명이 되어 있습니다.

<https://financedata.github.io/posts/finance-data-reader-users-guide.html> 또한, 이승준님이 Pycon 에서 엑셀에 비하여 파이썬의 장점에 대하여 강연하시는 내용이 유튜브에 있습니다.
https://www.youtube.com/watch?v=w7Q_eKN5r-I

FinanceDataReader 를 import 합니다. DataReader 함수에 종목코드, 시작일, 종료일을 인수로 넣어주면 아래와 같이 일봉데이터를 리턴합니다. 출력해보면 Date 가 index 로 되어 있음을 알 수 있습니다.

```
import FinanceDataReader as fdr

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

stock_data.head().style.set_table_attributes('style="font-size: 12px"') # head 메소드는 처음 5 row 만 출력합니다.
```

	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

각 월별 종가의 평균, 최대값, 최소값을 알아봅니다. 월별로 요약하면 index 에는 월의 마지막 날 짜가 되는 것을 유의하세요. head 메소드로 출력을 5 열로 제한합니다. pd.options 로 소숫점 이하는 보이지 않도록 합니다. 시간이 index 가 되어 있을 때 resample 이 가능합니다.

```
import pandas as pd
pd.options.display.float_format = '{:,.0f}'.format
stock_data.resample('M')
['Close'].agg(['mean', 'max', 'min']).head().style.set_table_attributes('style="font-size: 12px"') # 처음 5개만 출력
```

	mean	max	min
Date			
2021-01-31 00:00:00	86565.000000	91000	82000
2021-02-28 00:00:00	83127.777778	85300	81600
2021-03-31 00:00:00	82072.727273	84000	80900
2021-04-30 00:00:00	83586.363636	86000	81500
2021-05-31 00:00:00	80521.052632	83200	78500

주별로 요약할 수 도 있습니다. 이번에는 resample('W') 라고 해 줍니다. Resample 이 정말 유용한 기능이라는 것을 직감하셨을 것으로 생각합니다. 역시 한 주(월요일 ~ 일요일)의 마지막날이 Index 로 들어가 있습니다. 디폴트는 일요일입니다.

```
pd.options.display.float_format = '{:,.0f}'.format
stock_data.resample('W')
['Close'].agg(['mean','max','min']).head().style.set_table_attributes('style="font-size: 12px"')
```

	mean	max	min
Date			
2021-01-10 00:00:00	84160.000000	88800	82200
2021-01-17 00:00:00	89800.000000	91000	88000
2021-01-24 00:00:00	86820.000000	88100	85000
2021-01-31 00:00:00	85480.000000	89400	82000
2021-02-07 00:00:00	83600.000000	84600	82500