

1_TimeSeries_GetData

September 13, 2021

```
[5]: import yfinance as yf
import os
```

```
[6]: symbols = ['MMM', 'ABT', 'ABBV', 'ACN', 'ATVI', 'AYI', 'ADBE', 'AMD', 'AAP', 'AES', 'AET',
               'AMG', 'AFL', 'A', 'APD', 'AKAM', 'ALK', 'ALB', 'ARE', 'ALXN', 'ALGN', 'ALLE',
               'AGN', 'ADS', 'LNT', 'ALL', 'GOOGL', 'GOOG', 'MO', 'AMZN', 'AEE', 'AAL', 'AEP',
               ↪
               'AXP', 'AIG', 'AMT', 'AWK', 'AMP', 'ABC', 'AME', 'AMGN', 'APH', 'APC', 'ADI', 'ANDV',
               ↪
               'ANSS', 'ANTM', 'AON', 'AOS', 'APA', 'AIV', 'AAPL', 'AMAT', 'APTV', 'ADM', 'ARNC',
               ↪
               'AJG', 'AIZ', 'T', 'ADSK', 'ADP', 'AZO', 'AVB', 'AVY', 'BHGE', 'BLL', 'BAC', 'BK',
               'BAX', 'BBT', 'BDX', 'BRK',
               ↪ 'B', 'BBY', 'BIIB', 'BLK', 'HRB', 'BA', 'BWA', 'BXP', 'BSX',
               'BHF', 'BMY', 'AVGO', 'BF',
               ↪ 'B', 'CHRW', 'CA', 'COG', 'CDNS', 'CPB', 'COF', 'CAH', 'CBOE',
               ↪
               'KMX', 'CCL', 'CAT', 'CBG', 'CBS', 'CELG', 'CNC', 'CNP', 'CTL', 'CERN', 'CF', 'SCHW',
               ↪
               'CHTR', 'CHK', 'CVX', 'CMG', 'CB', 'CHD', 'CI', 'XEC', 'CINF', 'CTAS', 'CSCO', 'C', 'CFG',
               ↪
               'CTXS', 'CLX', 'CME', 'CMS', 'KO', 'CTSH', 'CL', 'CMCSA', 'CMA', 'CAG', 'CXO', 'COP',
               ↪
               'ED', 'STZ', 'COO', 'GLW', 'COST', 'COTY', 'CCI', 'CSRA', 'CSX', 'CMI', 'CVS', 'DHI',
               ↪
               'DHR', 'DRI', 'DVA', 'DE', 'DAL', 'XRAY', 'DVN', 'DLR', 'DFS', 'DISCA', 'DISCK', 'DISH',
               ↪
               'DG', 'DLTR', 'D', 'DOV', 'DWD', 'DPS', 'DTE', 'DRE', 'DUK', 'DXC', 'ETFC', 'EMN', 'ETN',
               ↪
               'EBAY', 'ECL', 'EIX', 'EW', 'EA', 'EMR', 'ETR', 'EVHC', 'EOG', 'EQT', 'EFX', 'EQIX', 'EQR',
               ↪
               'ESS', 'EL', 'ES', 'RE', 'EXC', 'EXPE', 'EXPD', 'ESRX', 'EXR', 'XOM', 'FFIV', 'FB', 'FAST',
               ↪
               'FRT', 'FDX', 'FIS', 'FITB', 'FE', 'FISV', 'FLIR', 'FLS', 'FLR', 'FMC', 'FL', 'F', 'FTV',
               ↪
               'FBHS', 'BEN', 'FCX', 'GPS', 'GRMN', 'IT', 'GD', 'GE', 'GGP', 'GIS', 'GM', 'GPC', 'GILD',
```

```

    ↪ 'GPN', 'GS', 'GT', 'GWW', 'HAL', 'HBI', 'HOG', 'HRS', 'HIG', 'HAS', 'HCA', 'HCP', 'HP', 'HSIC',
    ↪ 'HSY', 'HES', 'HPE', 'HLT', 'HOLX', 'HD', 'HON', 'HRL', 'HST', 'HPQ', 'HUM', 'HBAN', 'HII',
    ↪ 'IDXX', 'INFO', 'ITW', 'ILMN', 'IR', 'INTC', 'ICE', 'IBM', 'INCY', 'IP', 'IPG', 'IFF', 'INTU',
    ↪ 'ISRG', 'IVZ', 'IQV', 'IRM', 'JEC', 'JBHT', 'SJM', 'JNJ', 'JCI', 'JPM', 'JNPR', 'KSU', 'K', 'KEY',
    ↪ 'KMB', 'KIM', 'KMI', 'KLAC', 'KSS', 'KHC', 'KR', 'LB', 'LLL', 'LH', 'LRCX', 'LEG', 'LEN', 'LUK',
    ↪ 'LLY', 'LNC', 'LKQ', 'LMT', 'L', 'LOW', 'LYB', 'MTB', 'MAC', 'M', 'MRO', 'MPC', 'MAR', 'MMC', 'MLM',
    ↪ 'MAS', 'MA', 'MAT', 'MKC', 'MCD', 'MCK', 'MDT', 'MRK', 'MET', 'MTD', 'MGM', 'KORS', 'MCHP', 'MU',
    ↪ 'MSFT', 'MAA', 'MHK', 'TAP', 'MDLZ', 'MON', 'MNST', 'MCO', 'MS', 'MOS', 'MSI', 'MYL', 'NDAQ',
    ↪ 'NOV', 'NAVI', 'NTAP', 'NFLX', 'NWL', 'NFX', 'NEM', 'NWSA', 'NWS', 'NEE', 'NLSN', 'NKE', 'NI',
    ↪ 'NBL', 'JWN', 'NSC', 'NTRS', 'NOC', 'NCLH', 'NRG', 'NUE', 'NVDA', 'ORLY', 'OXY', 'OMC', 'OKE',
    ↪ 'ORCL', 'PCAR', 'PKG', 'PH', 'PDCO', 'PAYX', 'PYPL', 'PNR', 'PBCT', 'PEP', 'PKI', 'PRGO', 'PFE',
    ↪ 'PCG', 'PM', 'PSX', 'PNW', 'PXD', 'PNC', 'RL', 'PPG', 'PPL', 'PX', 'PCLN', 'PFG', 'PG', 'PGR',
    ↪ 'PLD', 'PRU', 'PEG', 'PSA', 'PHM', 'PVH', 'QRVO', 'PWR', 'QCOM', 'DGX', 'RRC', 'RJF', 'RTN', 'O',
    ↪ 'RHT', 'REG', 'REGN', 'RF', 'RSG', 'RMD', 'RHI', 'ROK', 'COL', 'ROP', 'ROST', 'RCL', 'CRM', 'SBAC',
    ↪ 'SCG', 'SLB', 'SNI', 'STX', 'SEE', 'SRE', 'SHW', 'SIG', 'SPG', 'SWKS', 'SLG', 'SNA', 'SO', 'LUV',
    ↪ 'SPGI', 'SWK', 'SBUX', 'STT', 'SRCL', 'SYK', 'STI', 'SYMC', 'SYF', 'SNPS', 'SYY', 'TROW', 'TPR',
    ↪ 'TGT', 'TEL', 'FTI', 'TXN', 'TXT', 'TMO', 'TIF', 'TWX', 'TJX', 'TMK', 'TSS', 'TSCO', 'TDG', 'TRV',
    ↪ 'TRIP', 'FOXA', 'FOX', 'TSN', 'UDR', 'ULTA', 'USB', 'UAA', 'UA', 'UNP', 'UAL', 'UNH', 'UPS', 'URI',
    ↪ 'UTX', 'UHS', 'UNM', 'VFC', 'VLO', 'VAR', 'VTR', 'VRSN', 'VRSK', 'VZ', 'VRTX', 'VIAB', 'V', 'VNO',
    ↪ 'VMC', 'WMT', 'WBA', 'DIS', 'WM', 'WAT', 'WEC', 'WFC', 'HCN', 'WDC', 'WU', 'WRK', 'WY', 'WHR', 'WMB',
    ↪ 'WLTW', 'WYN', 'WYNN', 'XEL', 'XRX', 'XLNX', 'XL', 'XYL', 'YUM', 'ZBH', 'ZION', 'ZTS']

```

```
[7]: symbols.append('SPY')
```

```
[8]: if not os.path.exists('data'):
      os.mkdir('data')
```

```
[9]: '''
      Next, I create a folder called Data, which is where we will store our
      ↪downloaded data.
      '''

      '''
      Next, we have a loop that looks at all the symbols and symbols list one by one.
      '''
      for symbol in symbols:
          '''
          Inside this loop, we check whether or not a CSFI file already exists for
          ↪our symbol in the data folder.
          '''
          if not os.path.exists(f"data/{symbol}.csv"):
              '''
              If it doesn't, then we proceed. Next, we call the function we have to
              ↪download to download stock data for the current symbol of used
              '''
              data = yf.download(symbol, start="2010-01-01", end="2018-12-31")
              '''
              So when we get back our data frame, we want to check it size.
              If the size is greater than zero, then we save the data frame in the
              ↪second loop.
              '''
              if data.size > 0:
                  data.to_csv(f"data/{symbol}.csv")
              else:
                  print("Not saving...")
```

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed

1 Failed download:

- AGN: No data found, symbol may be delisted

Not saving...

[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed

1 Failed download:

- APC: No data found, symbol may be delisted

Not saving...

[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed
[*****100%*****]	1 of 1 completed

[*****100%*****] 1 of 1 completed

1 Failed download:

- ARNC: Data doesn't exist for startDate = 1262284200, endDate = 1546194600

Not saving...

[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed

1 Failed download:

- BHGE: No data found, symbol may be delisted

Not saving...

[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed

1 Failed download:

- BBT: No data found, symbol may be delisted

Not saving...

[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed

1 Failed download:

- BRK.B: No data found, symbol may be delisted

Not saving...

[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed

1 Failed download:

- BF.B: No data found for this date range, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- CBG: No data found for this date range, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- CBS: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- CELG: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- CTL: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- CHK: Data doesn't exist for startDate = 1262284200, endDate = 1546194600

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

[illegible]

1 Failed download:

- CX0: No data found, symbol may be delisted

Not saving...

[illegible]


```

[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed

```

1 Failed download:

- FLIR: No data found, symbol may be delisted

Not saving...

```

[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed

```

1 Failed download:

- GGP: No data found for this date range, symbol may be delisted

Not saving...

```

[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed

```

1 Failed download:

- HRS: No data found, symbol may be delisted

Not saving...

```

[*****100%*****] 1 of 1 completed

```

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

[illegible]

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

```
1 Failed download:
- LLL: No data found, symbol may be delisted
Not saving...
```

```
1 Failed download:
- LUK: No data found for this date range, symbol may be delisted
Not saving...
```

11

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- KORS: No data found for this date range, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- MON: Data doesn't exist for startDate = 1262284200, endDate = 1546194600

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- MYL: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```


[illegible]

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

[illegible]

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- STI: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- SYMC: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- TIF: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- TMK: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- TSS: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- FOXA: Data doesn't exist for startDate = 1262284200, endDate = 1546194600
Not saving...

```
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- FOX: Data doesn't exist for startDate = 1262284200, endDate = 1546194600
Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- UTX: No data found, symbol may be delisted
Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- VAR: No data found, symbol may be delisted
Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- VIAB: No data found, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- HCN: No data found for this date range, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- WYN: No data found for this date range, symbol may be delisted

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

1 Failed download:

- XL: Data doesn't exist for startDate = 1262284200, endDate = 1546194600

Not saving...

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

```
[12]: '''
Delete files which have <10 lines
'''
```

```

for symbol in symbols:
    if os.path.exists(f"data/{symbol}.csv"):
        s = open(f"data/{symbol}.csv").readlines()
        if len(s) < 10:
            os.system(f"rm data/{symbol}.csv")

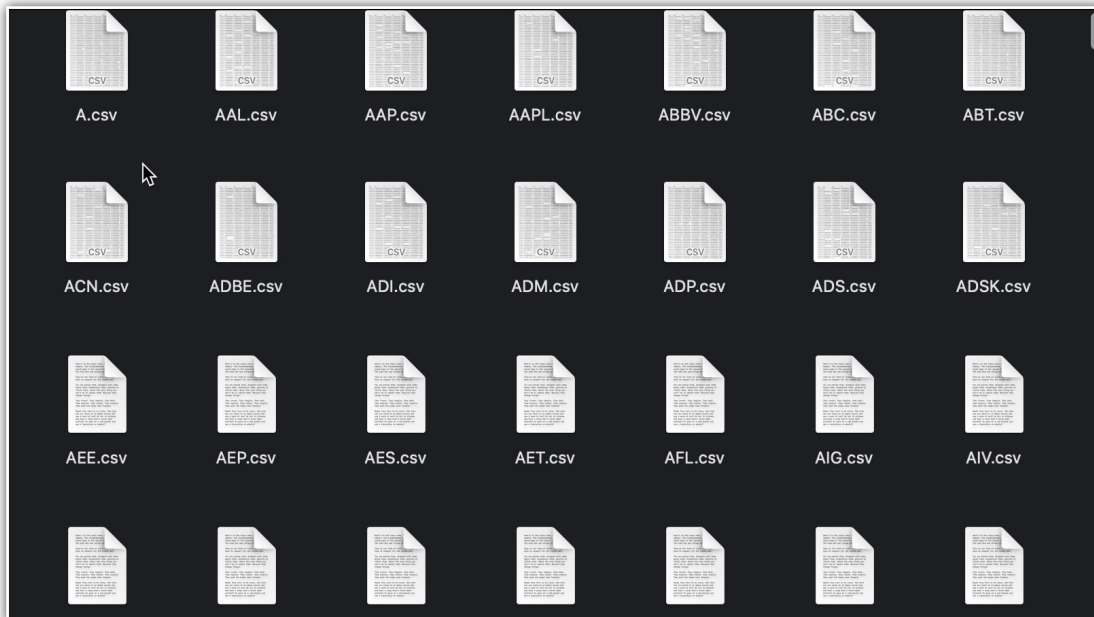
```

```

[13]: #Download the below Files
from IPython.display import Image
Image(filename='/Users/subhasish/GIT/Interstellar/SB-AI-DEV/ML/SB/TimeSeries/
↳Lazy Programmers/Image/f1.jpg')

```

[13]:

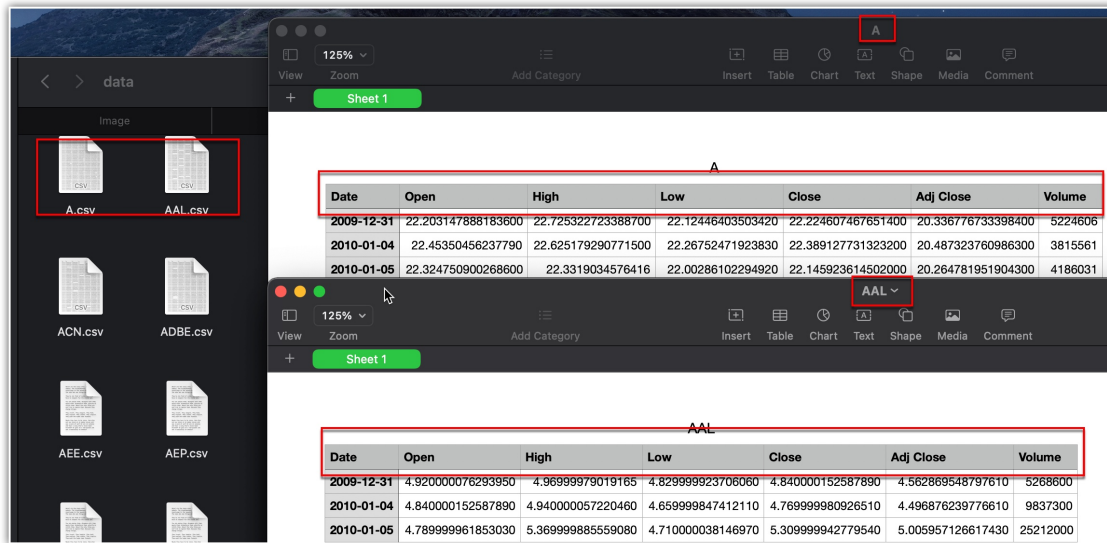


```

[14]: '''
You might wonder why did I previously save each data frame in a separate CSV,
↳rather than just combining
them all into the same CSV?
Well, the reason is, as you'll see later, they all have the exact same column,
↳names and dates.
'''
from IPython.display import Image
Image(filename='/Users/subhasish/GIT/Interstellar/SB-AI-DEV/ML/SB/TimeSeries/
↳Lazy Programmers/Image/2021-09-13_18-13-06.jpg')

```

[14]:



[16]: '''
 What we would like to do before combining the data is to have a name column_
 ↳ that tells us which stock
 that row of data refers to, and that's what this script does.
 '''

```
import pandas as pd
from glob import glob
```

[27]: '''
 Next we get all the file path for all the CSV's in the data folder.
 '''
 files = glob('data/*.csv')
 '''
 Next, I instantiate a variable called full_df to none, and this is the variable_
 ↳ that will store our
 data frame as we accumulate data from each of the CSV.
 '''
 full_df = None
 '''
 Next, we enter a loop to loop through all the CSV.
 Inside this loop, we load in the current CSV by calling pd.read_csv(f).
 '''

```
for f in files:
    if not os.path.exists("sp500full.csv"):
        print(f)
        df = pd.read_csv(f)
```

```

    #get the symbol name from CSV file name and set and Dataframe name
    ↪column
    symbol = f.split('/')[1].split('.')[0]
    '''
    Next, we create a new column called Name, and we assign it the value
    ↪symbol, which we just obtained.
    This will assign the same value to every row in the current data frame.
    '''
    df['Name'] = symbol
    '''
    Next, we check whether full_df is none if it is, then we can just
    ↪assign full deserve to be the current
    data frame.
    Otherwise we call the append function passing in the current data frame
    ↪D.F. and assign the result to
    '''
    if full_df is None:
        full_df = df
    else:
        full_df = full_df.append(df, ignore_index=True)
#we save the dataframe
if not os.path.exists("sp500full.csv"):
    full_df.to_csv('sp500full.csv', index=False)

```

```

data/CSCO.csv
data/UAL.csv
data/TROW.csv
data/ISRG.csv
data/PRGO.csv
data/TPR.csv
data/DVN.csv
data/MRO.csv
data/BA.csv
data/VRTX.csv
data/GILD.csv
data/NLSN.csv
data/EQIX.csv
data/MDT.csv
data/V.csv
data/QRVO.csv
data/A.csv
data/MO.csv
data/SWKS.csv
data/MCHP.csv
data/CDNS.csv
data/WLTW.csv
data/CHTR.csv

```

data/EIX.csv
data/BBY.csv
data/WBA.csv
data/HCA.csv
data/AJG.csv
data/DTE.csv
data/C.csv
data/T.csv
data/CF.csv
data/DISH.csv
data/MGM.csv
data/HUM.csv
data/CBOE.csv
data/CFG.csv
data/WU.csv
data/APH.csv
data/SYY.csv
data/MSI.csv
data/FCX.csv
data/ADM.csv
data/LH.csv
data/PKI.csv
data/LNT.csv
data/BAC.csv
data/LNC.csv
data/PSX.csv
data/GPN.csv
data/SRCL.csv
data/PPG.csv
data/IRM.csv
data/IQV.csv
data/ESS.csv
data/NOV.csv
data/NAVI.csv
data/HAL.csv
data/STZ.csv
data/FLS.csv
data/DXC.csv
data/ADI.csv
data/F.csv
data/HOG.csv
data/ADBE.csv
data/TDG.csv
data/ULTA.csv
data/ARE.csv
data/SYK.csv
data/CB.csv
data/TSN.csv

data/FLR.csv
data/PEP.csv
data/PEG.csv
data/LLY.csv
data/COST.csv
data/REG.csv
data/NWS.csv
data/LOW.csv
data/MDLZ.csv
data/FMC.csv
data/XEL.csv
data/AIZ.csv
data/PDCO.csv
data/CERN.csv
data/MET.csv
data/FTV.csv
data/DLR.csv
data/XRAY.csv
data/SCG.csv
data/FAST.csv
data/TJX.csv
data/SNA.csv
data/MPC.csv
data/D.csv
data/CA.csv
data/MRK.csv
data/STX.csv
data/NOC.csv
data/BXP.csv
data/KHC.csv
data/IPG.csv
data/UNP.csv
data/ALLE.csv
data/ABBV.csv
data/ORCL.csv
data/ECL.csv
data/ETR.csv
data/EBAY.csv
data/SBUX.csv
data/IR.csv
data/AMT.csv
data/INTU.csv
data/DRE.csv
data/CMA.csv
data/PG.csv
data/CAT.csv
data/MCD.csv
data/MNST.csv

data/AMZN.csv
data/INTC.csv
data/PNR.csv
data/GLW.csv
data/BDX.csv
data/KMI.csv
data/PWR.csv
data/APTV.csv
data/EXR.csv
data/HOLX.csv
data/EXPD.csv
data/GM.csv
data/TXN.csv
data/VRSK.csv
data/SJM.csv
data/TMO.csv
data/OXY.csv
data/RL.csv
data/CCI.csv
data/MMM.csv
data/MOS.csv
data/HSY.csv
data/JNPR.csv
data/DHI.csv
data/ED.csv
data/ES.csv
data/ADSK.csv
data/IP.csv
data/EXPE.csv
data/KO.csv
data/PCAR.csv
data/WDC.csv
data/PYPL.csv
data/NEE.csv
data/UPS.csv
data/LEG.csv
data/EMR.csv
data/MSFT.csv
data/ANSS.csv
data/CTAS.csv
data/UDR.csv
data/WEC.csv
data/AME.csv
data/HP.csv
data/IT.csv
data/ACN.csv
data/VRSN.csv
data/EW.csv

data/FL.csv
data/CMG.csv
data/AWK.csv
data/COO.csv
data/SHW.csv
data/HPQ.csv
data/AMAT.csv
data/CCL.csv
data/MLM.csv
data/AVY.csv
data/AAP.csv
data/ATVI.csv
data/EA.csv
data/DE.csv
data/SPG.csv
data/AMD.csv
data/KLAC.csv
data/NDAQ.csv
data/URI.csv
data/WHR.csv
data/PNC.csv
data/KMX.csv
data/WRK.csv
data/BIIB.csv
data/NVDA.csv
data/CHRW.csv
data/ROP.csv
data/IDXX.csv
data/EXC.csv
data/HES.csv
data/HD.csv
data/ALB.csv
data/VLO.csv
data/AON.csv
data/ZTS.csv
data/FDX.csv
data/DG.csv
data/HIG.csv
data/SIG.csv
data/CMS.csv
data/COL.csv
data/CAG.csv
data/INCY.csv
data/SCHW.csv
data/HSIC.csv
data/AZO.csv
data/AXP.csv
data/HPE.csv

data/DFS.csv
data/SEE.csv
data/HRL.csv
data/SO.csv
data/FRT.csv
data/ZBH.csv
data/CME.csv
data/XOM.csv
data/AMP.csv
data/AMG.csv
data/CVX.csv
data/CMCSA.csv
data/PCG.csv
data/CSRA.csv
data/PNW.csv
data/ICE.csv
data/NFX.csv
data/CTXS.csv
data/TRIP.csv
data/BEN.csv
data/DISCK.csv
data/UHS.csv
data/EMN.csv
data/SBAC.csv
data/ROK.csv
data/NRG.csv
data/NSC.csv
data/NKE.csv
data/FIS.csv
data/VTR.csv
data/MAS.csv
data/RF.csv
data/TAP.csv
data/MAR.csv
data/XYL.csv
data/CMJ.csv
data/FB.csv
data/MTD.csv
data/KR.csv
data/PLD.csv
data/IBM.csv
data/USB.csv
data/BSX.csv
data/LKQ.csv
data/FBHS.csv
data/ITW.csv
data/EOG.csv
data/PVH.csv

data/KMB.csv
data/SPGI.csv
data/NEM.csv
data/WFC.csv
data/ANDV.csv
data/EL.csv
data/GS.csv
data/GD.csv
data/CNP.csv
data/PM.csv
data/RE.csv
data/MCO.csv
data/CLX.csv
data/CAH.csv
data/HRB.csv
data/DGX.csv
data/AVB.csv
data/DIS.csv
data/GE.csv
data/HII.csv
data/ALL.csv
data/ETN.csv
data/ALGN.csv
data/NFLX.csv
data/LEN.csv
data/BHF.csv
data/FITB.csv
data/GWW.csv
data/NTRS.csv
data/CVS.csv
data/AOS.csv
data/FE.csv
data/ABC.csv
data/JPM.csv
data/ABT.csv
data/OMC.csv
data/COF.csv
data/TSCO.csv
data/PH.csv
data/HST.csv
data/AYI.csv
data/JBHT.csv
data/MAC.csv
data/COP.csv
data/DHR.csv
data/COG.csv
data/MAT.csv
data/CNC.csv

data/MCK.csv
data/TXT.csv
data/MTB.csv
data/SPY.csv
data/DISCA.csv
data/AKAM.csv
data/RMD.csv
data/GOOGL.csv
data/PAYX.csv
data/ALK.csv
data/DRI.csv
data/ILMN.csv
data/AAL.csv
data/XLNX.csv
data/MAA.csv
data/MMC.csv
data/EVHC.csv
data/GT.csv
data/FFIV.csv
data/VNO.csv
data/CINF.csv
data/VMC.csv
data/TWX.csv
data/SRE.csv
data/ORLY.csv
data/IVZ.csv
data/RCL.csv
data/PXD.csv
data/COTY.csv
data/SNPS.csv
data/GOOG.csv
data/YUM.csv
data/EQT.csv
data/KSS.csv
data/PFE.csv
data/AIV.csv
data/AVGO.csv
data/DUK.csv
data/REGN.csv
data/CL.csv
data/VFC.csv
data/UA.csv
data/VZ.csv
data/JCI.csv
data/ESRX.csv
data/AMGN.csv
data/TEL.csv
data/ADP.csv

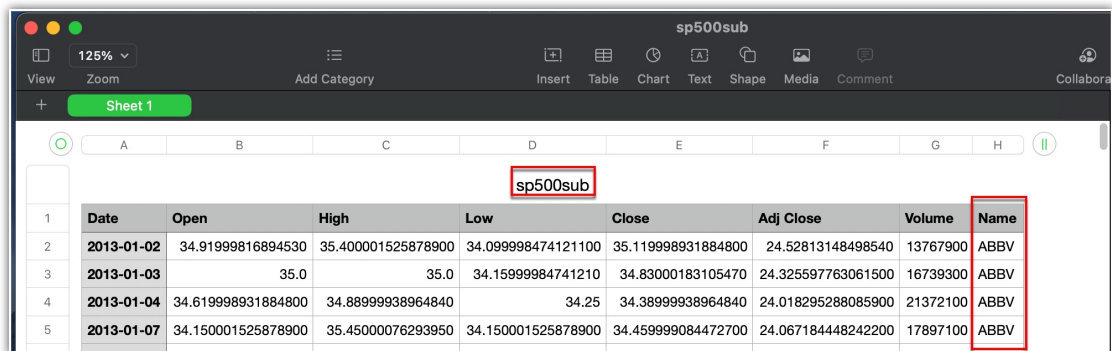
data/AET.csv
data/LB.csv
data/STT.csv
data/RRC.csv
data/RSG.csv
data/IFF.csv
data/ANTM.csv
data/GPS.csv
data/BLL.csv
data/QCOM.csv
data/LYB.csv
data/GIS.csv
data/PHM.csv
data/ROST.csv
data/LUV.csv
data/ALXN.csv
data/XEC.csv
data/MS.csv
data/CPB.csv
data/OKI.csv
data/BK.csv
data/SYF.csv
data/CHD.csv
data/SLG.csv
data/MHK.csv
data/INFO.csv
data/DAL.csv
data/APA.csv
data/K.csv
data/JWN.csv
data/AFL.csv
data/ADS.csv
data/CSX.csv
data/NI.csv
data/PFG.csv
data/NCLH.csv
data/ZION.csv
data/RJF.csv
data/HBAN.csv
data/UNH.csv
data/PRU.csv
data/GPC.csv
data/FISV.csv
data/WMB.csv
data/EQR.csv
data/PBCT.csv
data/KSU.csv
data/DVA.csv

data/AIG.csv
data/MA.csv
data/HBI.csv
data/HON.csv
data/O.csv
data/NWSA.csv
data/AES.csv
data/SLB.csv
data/XRX.csv
data/TGT.csv
data/AAPL.csv
data/MKC.csv
data/WY.csv
data/APD.csv
data/GRMN.csv
data/AEE.csv
data/HLT.csv
data/DLTR.csv
data/HAS.csv
data/WMT.csv
data/NTAP.csv
data/KIM.csv
data/BAX.csv
data/LMT.csv
data/KEY.csv
data/UNM.csv
data/BMY.csv
data/PSA.csv
data/WYNN.csv
data/RHI.csv
data/EFX.csv
data/NUE.csv
data/PKG.csv
data/CTSH.csv
data/SWK.csv
data/MU.csv
data/TRV.csv
data/L.csv
data/AEP.csv
data/CI.csv
data/SNI.csv
data/JNJ.csv
data/WM.csv
data/DOV.csv
data/FTI.csv
data/M.csv
data/CRM.csv
data/PGR.csv

```
data/WAT.csv
data/BWA.csv
data/LRCX.csv
data/NWL.csv
data/UAA.csv
data/BLK.csv
data/PPL.csv
```

```
[21]: from IPython.display import Image
Image(filename='/Users/subhasish/GIT/Interstellar/SB-AI-DEV/ML/SB/TimeSeries/
↳Lazy Programmers/Image/2021-09-13_18-30-52.jpg')
```

[21]:



	A	B	C	D	E	F	G	H
				sp500sub				
1	Date	Open	High	Low	Close	Adj Close	Volume	Name
2	2013-01-02	34.91999816894530	35.400001525878900	34.099998474121100	35.119998931884800	24.52813148498540	13767900	ABBV
3	2013-01-03		35.0	34.15999984741210	34.83000183105470	24.325597763061500	16739300	ABBV
4	2013-01-04	34.619998931884800	34.88999938964840		34.25	34.38999938964840	21372100	ABBV
5	2013-01-07	34.150001525878900	35.45000076293950	34.150001525878900	34.459999084472700	24.067184448242200	17897100	ABBV

```
[24]: import random
'''
Next, we have a set called Small Symbol's, which consists of a few stocks that
↳I'll be using in this
course, along with a few others from different industries.
'''
small_symbols = {'MMM', 'ABT', 'ABBV', 'ACN', 'ATVI', 'ADBE', 'AMD', 'AAP',
↳'AES',
'AFL', 'AKAM', 'IBM', 'GOOG', 'SBUX', 'AAPL', 'SPY'}

'''
So the point of this is to choose a few stocks from other industries that are
↳hopefully uncorrelated
So anyway, since I do want a few more stocks to work with, with the help of the
↳random module, I
sample one hundred more stocks from our symbol's list.
'''
sample = random.sample(symbols, 100)
```

```
[25]: '''
So the next step is to convert our sample into a set and then do a union of our
↳two sets.
```

```
As you know, sets can only have unique items, so any duplicates are now gone.
'''
small_symbols = small_symbols | set(sample)
```

```
[26]: full_df = None
for symbol in small_symbols:
    f = f"data/{symbol}.csv"
    if os.path.exists(f):
        df = pd.read_csv(f)
        df['Name'] = symbol
        if full_df is None:
            full_df = df
        else:
            full_df = full_df.append(df, ignore_index=True)

full_df.to_csv('sp500sub.csv', index=False)
```

```
[28]: # At the end we get two file which we will work on
from IPython.display import Image
Image(filename='/Users/subhasish/GIT/Interstellar/SB-AI-DEV/ML/SB/TimeSeries/
↳Lazy Programmers/Image/2021-09-13_18-42-10.jpg')
```

[28]:



[]: