In [5]:
```
pip install tushare
```

```
Collecting tushare
  Downloading tushare-1.2.77-py3-none-any.whl (130 kB)
     |████████████████████████████████| 130 kB 6.0 M
B/s
Collecting bs4
  Using cached bs4-0.0.1-py3-none-any.whl
Collecting simplejson
  Downloading simplejson-3.17.6-cp39-cp39-macosx_11_0_arm64.whl (73 kB)
     |████████████████████████████████| 73 kB 14.4 M
B/s
Requirement already satisfied: pandas in /opt/homebrew/Cellar/jupyterlab/
3.2.4/libexec/lib/python3.9/site-packages (from tushare) (1.3.5)
Collecting lxml
  Downloading lxml-4.7.1.tar.gz (3.2 MB)
     |████████████████████████████████| 3.2 MB 21.7
MB/s
  Preparing metadata (setup.py) ... done
Requirement already satisfied: requests in /opt/homebrew/Cellar/jupyterla
b/3.2.4/libexec/lib/python3.9/site-packages (from tushare) (2.26.0)
Requirement already satisfied: websocket-client>=0.57.0 in /opt/homebrew/
Cellar/jupyterlab/3.2.4/libexec/lib/python3.9/site-packages (from tushar
e) (1.2.1)
Collecting beautifulsoup4
  Using cached beautifulsoup4-4.10.0-py3-none-any.whl (97 kB)
Requirement already satisfied: pytz>=2017.3 in /opt/homebrew/Cellar/jupyt
erlab/3.2.4/libexec/lib/python3.9/site-packages (from pandas->tushare) (2
021.3)
Requirement already satisfied: python-dateutil>=2.7.3 in /opt/homebrew/Ce
llar/jupyterlab/3.2.4/libexec/lib/python3.9/site-packages (from pandas->t
ushare) (2.8.2)
Requirement already satisfied: numpy>=1.20.0 in /opt/homebrew/lib/python
3.9/site-packages (from pandas->tushare) (1.21.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /opt/homebrew/Cel
lar/jupyterlab/3.2.4/libexec/lib/python3.9/site-packages (from requests->
tushare) (1.26.7)
Requirement already satisfied: certifi>=2017.4.17 in /opt/homebrew/Cella
r/jupyterlab/3.2.4/libexec/lib/python3.9/site-packages (from requests->tu
share) (2021.10.8)
Requirement already satisfied: charset_normalizer~=2.0.0 in /opt/homebre
w/Cellar/jupyterlab/3.2.4/libexec/lib/python3.9/site-packages (from reque
sts->tushare) (2.0.7)
Requirement already satisfied: idna<4,>=2.5 in /opt/homebrew/Cellar/jupyt
erlab/3.2.4/libexec/lib/python3.9/site-packages (from requests->tushare)
(3.3)
Requirement already satisfied: six>=1.5 in /opt/homebrew/Cellar/six/1.16.
0_2/lib/python3.9/site-packages (from python-dateutil>=2.7.3->pandas->tus
hare) (1.16.0)
Collecting soupsieve>1.2
  Downloading soupsieve-2.3.1-py3-none-any.whl (37 kB)
Building wheels for collected packages: lxml
  Building wheel for lxml (setup.py) ... done
  Created wheel for lxml: filename=lxml-4.7.1-cp39-cp39-macosx_12_0_arm6
4.whl size=1494527 sha256=e7ded5457d00c7a7f1ec03d06e92964f173d70bb41c9dc5
cc950f37a7bd4b7f8
  Stored in directory: /Users/andy/Library/Caches/pip/wheels/b2/0f/42/b06
```

```
ea5234bf22bd3f4bf2d60a0dcdf4d4b2e709435d3ffb3c3
Successfully built lxml
Installing collected packages: soupsieve, beautifulsoup4, simplejson, lxm
l, bs4, tushare
Successfully installed beautifulsoup4-4.10.0 bs4-0.0.1 lxml-4.7.1 simplej
son-3.17.6 soupsieve-2.3.1 tushare-1.2.77
Note: you may need to restart the kernel to use updated packages.
```

In [6]:
```
pip install tqdm
```

```
Collecting tqdm
  Using cached tqdm-4.62.3-py2.py3-none-any.whl (76 kB)
Installing collected packages: tqdm
Successfully installed tqdm-4.62.3
Note: you may need to restart the kernel to use updated packages.
```

In [3]:
```
pip install pandas
```

```
Collecting pandas
  Downloading pandas-1.3.5.tar.gz (4.7 MB)
     |████████████████████████████████| 4.7 MB 5.0 M
B/s
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Preparing metadata (pyproject.toml) ... done
Requirement already satisfied: numpy>=1.20.0 in /opt/homebrew/lib/python
3.9/site-packages (from pandas) (1.21.4)
Requirement already satisfied: python-dateutil>=2.7.3 in /opt/homebrew/Ce
llar/jupyterlab/3.2.4/libexec/lib/python3.9/site-packages (from pandas)
(2.8.2)
Requirement already satisfied: pytz>=2017.3 in /opt/homebrew/Cellar/jupyt
erlab/3.2.4/libexec/lib/python3.9/site-packages (from pandas) (2021.3)
Requirement already satisfied: six>=1.5 in /opt/homebrew/Cellar/six/1.16.
0_2/lib/python3.9/site-packages (from python-dateutil>=2.7.3->pandas) (1.
16.0)
Building wheels for collected packages: pandas
  Building wheel for pandas (pyproject.toml) ... done
  Created wheel for pandas: filename=pandas-1.3.5-cp39-cp39-macosx_12_0_a
rm64.whl size=10245250 sha256=5507a715d28f81c3e04731e831bd9f17905a4f94b3f
97e7372b5ae8b44ff094c
  Stored in directory: /Users/andy/Library/Caches/pip/wheels/46/1f/09/be8
c6f216f000b48aaef3009dc7017707a1b18ef30ba548b8d
Successfully built pandas
Installing collected packages: pandas
Successfully installed pandas-1.3.5
Note: you may need to restart the kernel to use updated packages.
```

In [100]:
```
# Import libraries
import pandas as pd
import tushare as ts
from datetime import date
from tqdm import tqdm
```

In [101]:
```python
def download_ticker_mapping(tushare_token, ticker_path=None):
    ts.set_token(tushare_token)
    pro=ts.pro_api()
    df_ticker=pro.stock_basic(exchange="",list_status="L",fields="")

    if ticker_path is not None:
        df_ticker.to_excel(ticker_path,index=False,encoding="utf_8_sig")

    return df_ticker
#download_ticker_mapping("600848", ticker_path=None)
```

In [117]:
```python
all_data=ts.get_hist_data('603208')
new_stock=all_data.iloc[::-1]
#all_data.iloc[::-1].reset_index().to_csv("yingliu.csv", encoding='utf-8',
new_stock
```

本接口即将停止更新，请尽快使用Pro版接口：https://waditu.com/document/2

Out[117]:

date	open	high	close	low	volume	price_change	p_change	ma5	ma10	ma20	v_n
2019-06-19	31.42	32.00	31.15	30.94	10196.39	0.05	0.16	31.150	31.150	31.150	10196
2019-06-20	31.02	32.06	31.89	30.89	7192.10	0.74	2.38	31.520	31.520	31.520	8694
2019-06-21	31.97	32.45	32.00	31.73	5794.39	0.11	0.34	31.680	31.680	31.680	7727
2019-06-24	32.58	34.00	33.88	32.11	14843.00	1.88	5.88	32.230	32.230	32.230	9506
2019-06-25	33.97	33.97	33.28	32.70	8129.08	-0.60	-1.77	32.440	32.440	32.440	9230
...	...	...	...	...	...	...	...	...	...	...	...
2021-12-13	62.00	63.27	61.00	60.50	19318.03	-1.94	-3.08	62.408	58.073	57.978	26770
2021-12-14	61.00	61.00	59.44	59.10	12903.49	-1.56	-2.56	62.168	58.674	57.851	24027
2021-12-15	59.97	59.97	58.28	57.83	10597.00	-1.16	-1.95	61.208	59.067	57.760	19902
2021-12-16	58.79	59.80	59.67	57.72	12643.37	1.39	2.38	60.266	59.769	57.804	15967
2021-12-17	59.50	60.49	58.75	58.03	10406.95	-0.92	-1.54	59.428	60.331	57.731	13173

610 rows × 14 columns

The data shown here is the up-to-date historical data of mainland china stock market, where i just need to input the token/stock number of the specific stock, then I can have the datatable for the entire stock from 2019 to 2021, which inclused the open, high, close, low, and etc information.

The data come from tushare which is a chinese company, they first post their open-source github code online, and created a package called tushare to track stock information, it will update everyday. website in chinese: https://waditu.com/document/2 (https://waditu.com/document/2) github repository: https://github.com/waditu/Tushare (https://github.com/waditu/Tushare)

It is a tool that realizes the process of data collection, cleaning and processing to data storage of financial data such as stocks/futures, and meets the needs of financial quantitative analysts and people studying data analysis in terms of data acquisition. It is characterized by a wide range of data coverage and it is simple to use the packages and datatable.

The biasness maybe the data does not cover the holistics to a stock, some stocks are very old but the datatable only covers the all the information from 2019/4/23. I have not yet come up with any censoring related to the datatable, however, it is a chinese company that will come under chinese policy's rule.

```
In [120]: new_stock.loc["2019-06-19"]
```

```
Out[120]: open              31.42
          high              32.00
          close             31.15
          low               30.94
          volume         10196.39
          price_change       0.05
          p_change           0.16
          ma5               31.15
          ma10              31.15
          ma20              31.15
          v_ma5          10196.39
          v_ma10         10196.39
          v_ma20         10196.39
          turnover           3.31
          Name: 2019-06-19, dtype: float64
```

```
In [125]: s = pd.DataFrame(
              {"StartDate": pd.date_range("2019-06-19", "2021-12-17", freq="D")}
          )
          #s["StartDate"].astype(str)[0]==new_stock["StartDate"][0]
          #==pd.Timestamp(new_stock["StartDate"][0]).date()
          len(s["StartDate"].astype(str))
          new_stock.loc["2019-06-19"][0]
          print(s["StartDate"].astype(str)[1])
```

```
2019-06-20
```

In [126]:
```python
new_s= pd.DataFrame(
    {"StartDate": s["StartDate"].astype(str)}
)
all_info=[]
temp=list(new_stock.index)
for all_t in s["StartDate"].astype(str):
    for weekday in temp:
        if all_t==weekday:
            all_info.append(new_stock.loc[weekday][0])
            temp.pop(0)
            break
        else:
            all_info.append(0)
            break

for i in range(len(all_info)):
    if all_info[i]==0:
        all_info[i]=all_info[i-1]

all_info
```

```
59.48,
58.8,
58.8,
58.8,
60.13,
58.0,
56.2,

56.88,
57.28,
57.28,
57.28,
54.9,
53.88,
53.43,
54.52,
52.72,
52.72,
52.72,
53.12,
55.5
```

In [127]:
```python
new_s["open"]=all_info
new_s.to_csv("jiangshan.csv", encoding='utf-8', index=False)
```

In [128]:
```python
len(new_s)
```

Out[128]: 913

In [51]:
```python
import datetime
pd.date_range(start='6/18/2019', end='12/16/2021')
```

Out[51]:
```
DatetimeIndex(['2019-06-18', '2019-06-19', '2019-06-20', '2019-06-21',
               '2019-06-22', '2019-06-23', '2019-06-24', '2019-06-25',
               '2019-06-26', '2019-06-27',
               ...
               '2021-12-07', '2021-12-08', '2021-12-09', '2021-12-10',
               '2021-12-11', '2021-12-12', '2021-12-13', '2021-12-14',
               '2021-12-15', '2021-12-16'],
              dtype='datetime64[ns]', length=913, freq='D')
```

In [49]:
```python
list(s["StartDate"])
```
```
 Timestamp('2019-10-13 00:00:00'),
 Timestamp('2019-10-19 00:00:00'),
 Timestamp('2019-10-20 00:00:00'),
 Timestamp('2019-10-26 00:00:00'),
 Timestamp('2019-10-27 00:00:00'),
 Timestamp('2019-11-02 00:00:00'),
 Timestamp('2019-11-03 00:00:00'),
 Timestamp('2019-11-09 00:00:00'),
 Timestamp('2019-11-10 00:00:00'),

 Timestamp('2019-11-16 00:00:00'),
 Timestamp('2019-11-17 00:00:00'),
 Timestamp('2019-11-23 00:00:00'),
 Timestamp('2019-11-24 00:00:00'),
 Timestamp('2019-11-30 00:00:00'),
 Timestamp('2019-12-01 00:00:00'),
 Timestamp('2019-12-07 00:00:00'),
 Timestamp('2019-12-08 00:00:00'),
 Timestamp('2019-12-14 00:00:00'),
 Timestamp('2019-12-15 00:00:00'),
 Timestamp('2019-12-21 00:00:00'),
```

In [ ]: