Apple公司股价数据 - 时间序列

导库

```
import pandas as pd
import matplotlib.pyplot as plt
```

```
apple = pd.read_csv('data/APPLE_STOCK.csv')
apple.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7407 entries, 0 to 7406
Data columns (total 7 columns):
Date            7407 non-null object
       Open        7407 non-null float64
  High          7407 non-null float64
  Low           7407 non-null float64
   Close        7407 non-null float64
 Volume         7407 non-null int64
  Adj Close     7407 non-null float64
dtypes: float64(5), int64(1), object(1)
memory usage: 405.1+ KB
```

转换日期列数据，并设置其为索引列

```
apple.Date = pd.to_datetime(apple.Date)
apple = apple.set_index('Date')
apple.head(3)
```

```
<tr style="text-align: right;">
  <th></th>
  <th>Open</th>
  <th>High</th>
  <th>Low</th>
  <th>Close</th>
  <th>Volume</th>
  <th>Adj Close</th>
</tr>
<tr>
  <th>Date</th>
  <th></th>
  <th></th>
  <th></th>
  <th></th>
  <th></th>
  <th></th>
</tr>
```

```
<tr>
  <th>2014-01-22</th>
  <td>550.91</td>
  <td>557.29</td>
  <td>547.81</td>
  <td>551.51</td>
  <td>13570900</td>
  <td>551.51</td>
</tr>
<tr>
  <th>2014-01-21</th>
  <td>540.99</td>
  <td>550.07</td>
  <td>540.42</td>
  <td>549.07</td>
  <td>11733100</td>
  <td>549.07</td>
</tr>
<tr>
  <th>2014-01-17</th>
  <td>551.48</td>
  <td>552.07</td>
  <td>539.90</td>
  <td>540.67</td>
  <td>15240700</td>
  <td>540.67</td>
</tr>
```

查看每一列的数据类型

```
apple.columns = [col.strip() for col in apple.columns]
```

```
apple.dtypes
```

```
Open         float64
High         float64
Low          float64
Close        float64
Volume         int64
Adj Close    float64
dtype: object
```

有重复的日期吗?

```
apple.index.is_unique
```

```
True
```

找到每个月的最后一个交易日(business day)

```
apple_month = apple.resample('BM').mean()
apple_month.head(3)
```

```
<tr style="text-align: right;">
  <th></th>
  <th>Open</th>
  <th>High</th>
  <th>Low</th>
  <th>Close</th>
  <th>Volume</th>
  <th>Adj Close</th>
</tr>
<tr>
  <th>Date</th>
  <th></th>
  <th></th>
  <th></th>
  <th></th>
  <th></th>
  <th></th>
</tr>
```

```
<tr>
  <th>1984-09-28</th>
  <td>26.981250</td>
  <td>27.333125</td>
  <td>26.606250</td>
  <td>26.738750</td>
  <td>4.807300e+06</td>
  <td>2.948750</td>
</tr>
<tr>
  <th>1984-10-31</th>
  <td>25.035652</td>
  <td>25.313478</td>
  <td>24.780435</td>
  <td>24.806957</td>
  <td>5.559409e+06</td>
  <td>2.736957</td>
</tr>
<tr>
  <th>1984-11-30</th>
  <td>24.545238</td>
  <td>24.782857</td>
  <td>24.188095</td>
  <td>24.236190</td>
  <td>5.749562e+06</td>
  <td>2.674286</td>
</tr>
```

数据集中最早的日期和最晚的日期相差多少天?

```
(apple.index.max() - apple.index.min()).days
```

```
10729
```

按照时间顺序可视化Adj Close值

```
apple_open = apple['Adj Close'].plot(title='Apple Stock')
fig = apple_open.get_figure()
fig.set_size_inches(16, 9)
plt.show()
```