

# 爬取新浪财经现金流量表

股票代码: 600357 / 600358 / 600359, 保存到 D:\datasets。

```
In [1]: import os
import csv
import requests
from bs4 import BeautifulSoup
```

```
In [2]: save_dir = r'D:\datasets'
os.makedirs(save_dir, exist_ok=True)
headers = {'User-Agent': 'Mozilla/5.0'}
```

```
In [3]: def download(stockid):
    url = f'https://money.finance.sina.com.cn/corp/go.php/vFD_CashFlow/stockid/{stockid}'
    r = requests.get(url, headers=headers, timeout=30)
    r.encoding = 'gb2312'
    soup = BeautifulSoup(r.text, 'html.parser')
    table = soup.find('table', id='ProfitStatementNewTable0')

    rows = []
    for tr in table.find_all('tr'):
        cells = [c.get_text(strip=True) for c in tr.find_all(['td', 'th'])]
        if cells:
            rows.append(cells)

    company = rows[0][0].split('(')[0]
    path = os.path.join(save_dir, f'{company}({stockid}) 现金流量表.csv')
    with open(path, 'w', newline='', encoding='utf-8-sig') as f:
        csv.writer(f).writerows(rows)
    print('保存: ', path)
```

```
In [4]: for code in ['600357', '600358', '600359']:
    download(code)
```

保存: D:\datasets\承德钒钛(600357) 现金流量表.csv  
保存: D:\datasets\国旅联合(600358) 现金流量表.csv  
保存: D:\datasets\新农开发(600359) 现金流量表.csv