

菜鸟学Python趣味实战项目-股票

项目的要求

1.Pre-study

这次涉及pandas的很多操作，大家要提前看一下pandas的书，建议看"用Python进行数据分析这本书"，星球里面都有，而且都有最新的版本，建议看中文和英文，中文用来看，英文用来搜索。

1).包的准备

安装tushare,pandas,matplotlib的包

a).如果要单独安装的话，下面是我安装的版本

```
leoxindeMacBook-Pro:bin leoxin$ pip3 show pandas Name: pandas Version: 0.20.1
```

```
leoxindeMacBook-Pro:bin leoxin$ pip3 show tushare Name: tushare Version: 1.2.21
```

```
leoxindeMacBook-Pro:bin leoxin$ pip3 show matplotlib Name: matplotlib Version: 2.0.2
```

b).还有一个更简单的方法就是直接下载anaconda里面会有很多默认数据分析的内置包，不过它的版本比较新！

2.要用tushare获取股票的数据

a).我们先拿到所有券商股的列表

b).接着我们获取所有券商股的历史数据，比如我们可以拿到它近1年的200多个交易日的数据

```
import tushare as ts
def get_all_securities():
    df_all=ts.get_industry_classified()
    df_finance=df_all[df_all['c_name']=='金融行业']
    return df_finance[df_finance['name'].str.contains('证')]
get_all_securities()
```

	code	name	c_name
1846	600030	中信证券	金融行业
1848	600109	国金证券	金融行业
1850	600369	西南证券	金融行业
1855	600837	海通证券	金融行业
1856	600958	东方证券	金融行业
1857	600999	招商证券	金融行业
1862	601198	东兴证券	金融行业
1867	601377	兴业证券	金融行业
1869	601555	东吴证券	金融行业
1872	601688	华泰证券	金融行业

1873	601788	光大证券	金融行业
1875	601901	方正证券	金融行业
1882	000562	宏源证券	金融行业
1884	000686	东北证券	金融行业
1886	000728	国元证券	金融行业
1887	000750	国海证券	金融行业
1888	000776	广发证券	金融行业
1889	000783	长江证券	金融行业
1891	002500	山西证券	金融行业
1892	002673	西部证券	金融行业
1893	002736	国信证券	金融行业

上面的不是很全或者用页面爬取：<http://www.shdjt.com/flsort.asp?lb=991036>

获取所有的券商列表

3.进行简单的数据预处理

我们主要是保留成交量，涨幅

4.数据可视化

我们把所有的数据数据进行对比分析，用不同的可视化图像进行对比 比如可用用直方图，散点图，折线图和箱体图进行对比