

python爬虫项目

地址:

[http://www.shfe.com.cn/statements/dataview.html?paramid=delaymarket_all]

目标:

- 抽取数据转换成 `pandas.DataFrame`
- 按照涨跌幅进行排序
- 按照交易量进行排序
- 代码提交到Git
- 完成说明文档

学习内容:

- http协议通讯流程
- 基础的文本处理
- 正则表达式
- DataFrame基础操作

Tips:

- Chrome -- 更多工具 --- 开发者工具 -- 观察数据流
- `import re`
- json处理