

任务分解：

1. 获取股市数据，提取其中各A股代码、每日最高/最低价、成交量、成交金额、总金额等；收集股票的历史数据
2. 将数据存入数据库中，通过pymysql编写数据库插入和更新函数
3. 对以上获取数据进行聚合分析
4. 尝试服务器部署任务

整体设计：

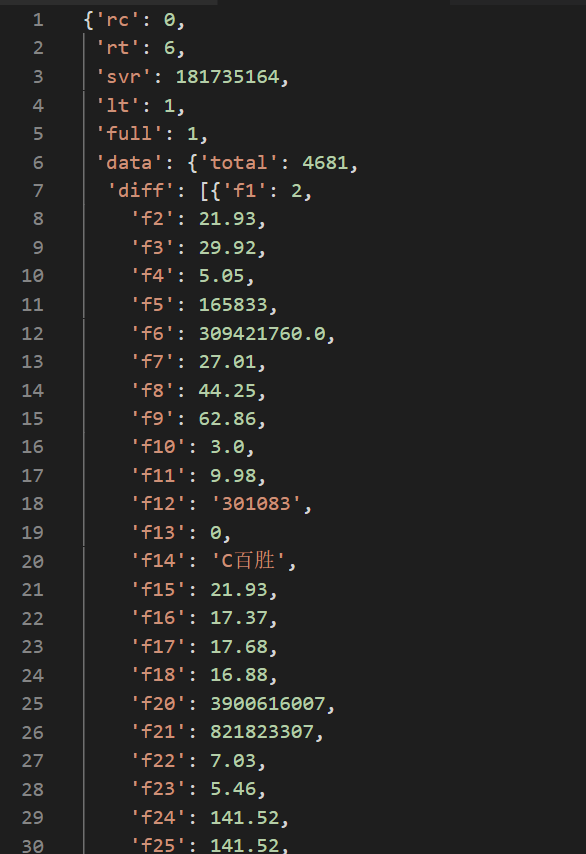
目前找到的数据源可以获取当前所有股票的信息，包括股票名称、代码、当前价格、涨幅、总手、换手、今日最低与最高、今日开盘价与昨日收盘价等信息。从中筛选出A股的必要信息进行清理与整合。

数据获取完后将其存入数据库中，并完成数据插入与更新函数，将历史数据存在History表中，当前数据存在Detail表中。定时运行更新与插入函数来记录不同时间维度上的数据。

对于以上信息可以获取到某一天收盘后的整体信息，进行聚合分析；还可以获得某一支股的从存储以来的行情趋势，从数据库中提取这些信息，在jupyter notebook中进行分析。

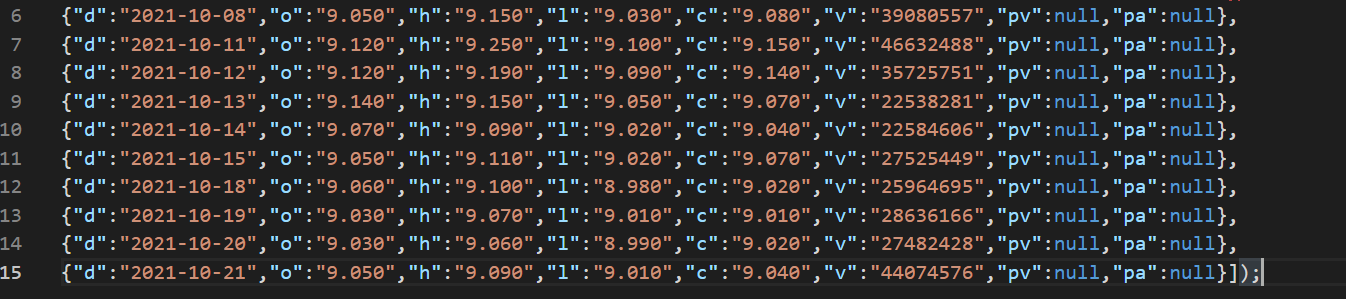
1. **数据来源**

目前在东方财富网的network中找到了一个数据链接，里面的数据格式如下：



经过和网页内信息比照获得了每一个f标签指代的信息，可以将所有的股票信息爬取下来再进行筛选整合。

新浪财经上获取到的个股历史信息如下（部分示例）：



可以利用其中每一天对应的开盘价（o），收盘价（c），最高价（h）和最低价（l）

1. **数据入库**

创建数据库、通过pymysql将数据存入本地数据库中

创建current\_details表存储实时信息并实时更新，考虑到数据量历史数据就不入库了，在聚合分析时需要分析哪只股再单独调用爬取相应数据进行分析。

1. **聚合分析**

先去了解一下聚合分析有哪些思路和手段，比如可以统计某股过去一周/半月/一个月等的均线、过去一段时间涨幅/跌幅，以及成交量的变化趋势等等...（待补充）

在jupyter notebook上写一个函数，接口是股票代码和时间段，可以给出该时间段的如上不同分析。

1. **服务部署**