**作业内容：**

**一. 基于特质波动率的量化组合构建与计算（90分）**

**特质波动率**，即股票收益波动率中无法被系统性风险所捕捉的部分，是评估投资风险的重要指标，其通过三因子模型计算：

|  |  |
| --- | --- |
|  |  |
|  |  |

其中，是第i支股票在m月第t天的收益率，是对应的无风险收益率，是m月第t天的市场收益率，和为对应的Fama-French因子。

**特质波动率之谜**：滞后的特质波动率与预期收益率之间存在负向关系, 并把这种实证中的负向关系与理论模型中的正向关系相违背的现象视为“特质波动率之谜”。

现要求基于特质波动率构建量化策略：

（1）基于A股上市公司近10年的数据，以特质波动率为因子构建投资组合，每月最后一个交易日调仓一次，卖出当前持有股票，买入新股票。

例如：以2024年11月29日为例，该日期为24年11月最后一个交易日，我们在该天计算每一支股票的特质波动率，然后基于特质波动率值的大小，将股票分为30组，构建30个投资组合，且组合为等权重的。每一组合我们持有一个月，一直到2024年12月31日（12月的最后一个交易日），卖出当前持有的股票，并买入新的组合的股票（不考虑交易成本，且假定以卖出的资金量买入新的组合）；

（2）每一期持有的股票中不含有ST类股票；

（3）进行调仓或者买入时不包含当天涨停或跌停的股票；

（4）不包含金融行业的股票；

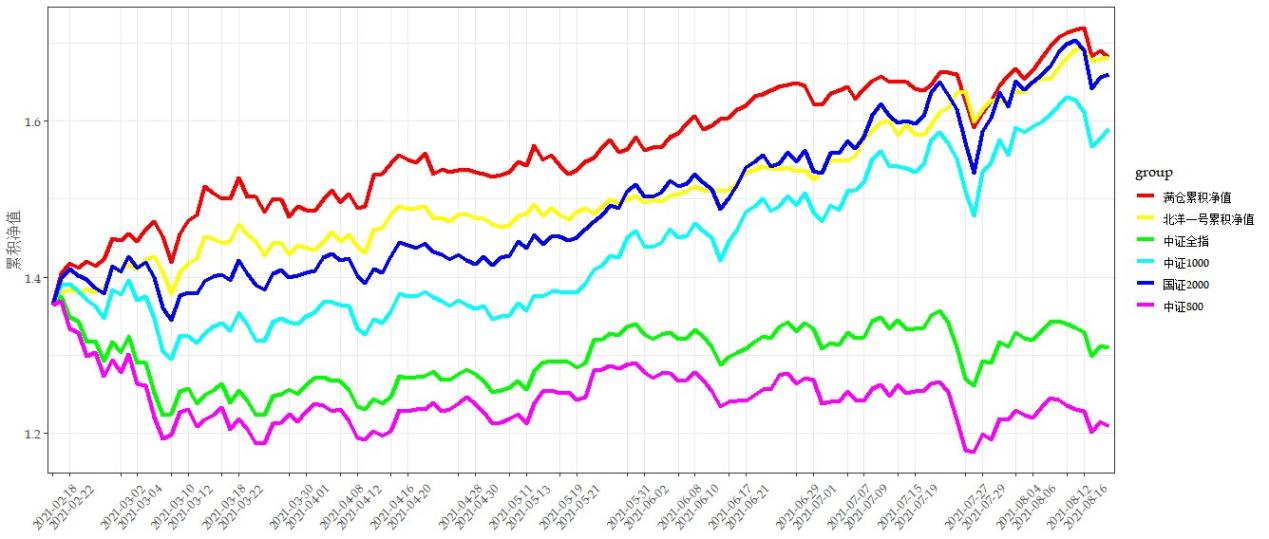
**1. 基于上述思路，要求大家构建一个基于特质波动率的量化策略类，要求：**

**（1）这个类能够完成上述投资组合的构建（有组合构建方法，体现出相应的数据筛选）（10分）**

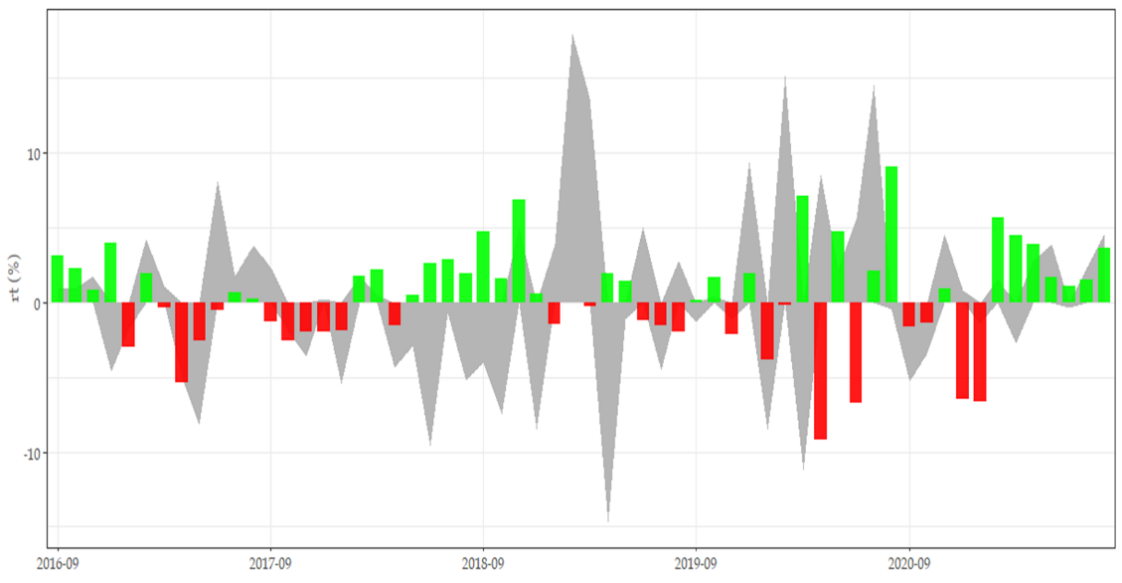
**（2）计算30个组合的累积收益率，并以折线图的形式进行可视化。（假设我们的初始净值为1）（10分）**

其中，为组合i在第t期的累积收益率，为组合i在第t期的当期收益率。

折线图要求为如下形式，呈现出30个投资组合



**（3）选择累积收益最高的投资组合，计算该组合的每月的收益，并画出对应的条形图，如下形式：（10分）**



**（4）选择累积收益最高的投资组合，作为你的最优投资组合，计算如下指标：（30分）**

a. 持有期内各期收益的算术平均，几何平均，标准差，最大值，最小值

b. 持有期内各期收益的胜率（收益大于0的比率）

c. 每期收益大于0的个股比率的平均值（每期个股的收益大于零的个股数量占该期组合中股票总量的比率，然后取各期的平均值）

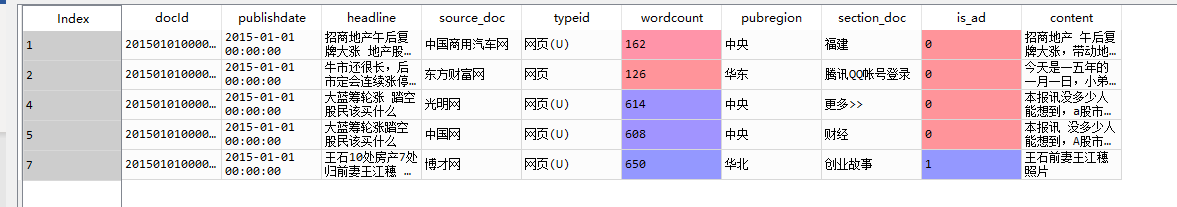
d. 年化收益率（几何平均）：对累积收益开年次方（持有总月数/12）

e. 最大回撤：最大回撤率计算公式：max(1-账户当日价值/ 当日之前账户最高价值)\*100%

**2. 考虑小市值因子长期在A股有效，现将市值与特质波动率进行交叉，具体做法，先基于特质波动率将分为6组，然后在每组内再基于市值分为6组。并基于新的量化策略组合，重新完成问题1中的计算与可视化操作。（30分）**

**二. 文本数据的筛选（10分）**

提供给大家的新闻数据的基本格式如下图：



其中，source\_doc是获取该条新闻的来源，但是，有些的新闻是这些来源转载自其他来源的新闻，现要求大家筛选出来源自《证券时报》的新闻，包括证券时报集团下的各子媒体的新闻，以及其他来源转载自证券时报的新闻。

**作业要求：**

1. 以Python作为工具来实现上述问题；

2. 以面向对象的编程思想来实现上述问题；

3. 数据处理的结果以word或PDF文件形式上交：

a. 数据整合结果，保存到excel文件中上交，截取前20行数据的截图插入到上交的word或PDF文件中，并对相应结果进行简要的说明（如：问题1的整理结果共包含xx支股票，xx条数据）；

b. 可视化结果，将python画出的图保存为png或者jpg等图片格式，插入到word或者PDF文件中，并标识号说明；

4. 上述数据处理的python程序上交，程序保存为.py文件，程序中要加好注释。

5. 数据整合完之后的形式：



6. 作业以小组成员的姓名和学号命名，[发送至助教邮箱python\_2425@163.com](mailto:发送至助教邮箱python_2425@163.com)。