1. 通过郑州商品交易所的网站，抓取2021年全年的持仓排名数据。



答：使用selenium进行爬虫，程序为code1-1.py。数据按照月份进行整合保存，命名为month\_{}\_result\_data.csv中。在爬虫过程中，受访问限制或网络因素，会导致网页响应超时，程序中断，在程序第80行，可将循环起始改为中断月份，重新开始下载数据。

注：文件中的msedgedriver.exe为selenium需要配置的驱动，请勿删除或修改，此外，电脑需安装Edge浏览器，以便数据爬虫。

1. 根据每日持仓排名表，统计过去一年中，在甲醇（MA）中，哪一家期货公司的总计成交量最大？合计多少手？



如2022/2/8日，甲醇上成交量最大的是东证期货，211529手。

答：计算的文件code1-2.py。甲醇MA期货2021年总成交量最大的期货公司是东证期货，总计85824525手。

1. （附加题，不做要求）跟踪成交量占比高的公司的增减仓行为（考察“持买/卖仓量”的增减量），是否能带来额外的预测力？请自行寻找相应数据进行分析。

答：对于预测力的描述，我选择的是信息系数（Information Coefficient, IC）和信息比率（Information Ratio, IR）进行表述。IC越大，因子预测股票收益的能力越强；IR则描述IC的稳定性。IC和IR的计算程序在IC.ipynb的文件中。

对于数据，因获取的数据有限，仅收录了甲醇MA, 玻璃FG, 纯碱SA, 硅铁SF, 锰硅SM, PTA, 尿素UR, 动力煤ZC，8种期货进行计算。每一个期货的价格，均采用的是主力合约888，对价格进行了前复权平滑处理，数据在feature\_price.csv中。

从结果中可以看出，增仓量的10日IC超过0.02的比例为0.48，IR为-0.07765，说明增仓量有一定的预测能力。减仓量的10日IC超过0.02的比例为0.408，IR为-0.12047，其预测能力弱于增仓量。

根据增减仓量作为交易信号，采用CTA策略中的布林带策略，当买仓量上涨，卖仓量上涨，则进行买入，且突破10日均线时，进行买入；当卖仓量下降，卖仓量下降，且突破10日均线时，进行平仓。程序在MA888.ipynb中，此程序选择甲醇MA888为标的，进行交易。在未考虑滑点、手续费、止盈止损等限制后，累计收益为24.85%，最大单笔亏损3.05%，一共进行了16次交易。由于时间有限，本人仅适用简单的以持买/卖仓量进行CTA策略的设计，如有机会，希望可以进行进一步的探索。

下图为买卖点的标识、价格走势以及通道图。

