**项目介绍**

**1.项目简介**  
 本项目以510300沪深300基金与燃油主连期货为标的，使用python回测了一个分钟频率的、基于XMA函数的均值回归量化策略。在项目过程中，我完成了数据清洗，构建并优化策略指标，搭建回测系统等工作。以下是该项目的亮点。

1. **项目亮点**
2. **对原始数据的数据清理**

本项目的一分钟频率数据均来自Tushare网站，但这些原始数据无法直接用于回测。分标的看，510300沪深300基金的一分钟数据仅提供不复权版本，而燃油期货则只有具体合约数据，缺乏主力连续合约数据。为解决510300基金的数据问题，在查阅了复权价格与不复权价格的关系后，手动记录了历次分红的时间节点和金额，并利用Python将不复权价格转换为前复权价格。对于燃油期货数据，我通过Tushare提供的每日主力合约目录构建了“燃油主力合约”。此外，因Tushare网站的燃油期货数据在部分时间段存在价格完全不变的问题，为保证回测的准确性，我删除了这些存在问题的日期数据。

**（2）基于XMA函数的指标的构建与优化**

本项目的量化策略基于通达信股票软件中的XMA函数，因此首先需将XMA函数转化为Python自定义函数。由于这一函数无法在互联网上直接找到，我根据通达信软件和网络上相关文字描述，结合自身理解，将其成功转换为Python代码。总体而言，XMA函数是对过去N分钟价格的加权平均，其主要难点在于计算当前时点之前的XMA时会用到“未来数据”，即需要将当前数据加入计算中。而当“未来数据”用完时，算法需恢复为常规加权方式来计算加权平均值。项目的具体策略中应用了双层XMA复合函数，因此需要在当前时点下计算过去的XMA值，一定程度上增加了复杂性。

在初步实现中，我使用了双层for循环来复刻XMA函数，导致整个策略代码形成了4层for循环结构。尽管这一实现能够理想地计算出每个时点上过去每一天的XMA值，但由于1分钟频率的数据量过大，代码运行效率不足，无法实际生成策略指标。为此，我通过减少for循环次数进行优化，仅计算每个时点当天的XMA值。我将双层XMA函数转换为两个独立函数，并以复合函数形式实现XMA，最终将循环结构简化为3层for循环。通过减少循环次数，我成功利用优化后的函数生成了量化策略的指标。

**（3）构造回测系统**

为评估本策略的盈利能力与风险，本项目构建了一个包含多种要素的回测系统，其中包括交易费用、交易时间限制、止损机制和精确收益率等内容。一方面，为了纯粹反映策略的盈利能力并隔离日间风险，本研究对每天开盘后25分钟与收盘前一小时的交易作出了限制；另一方面，为更精确地反映策略的收益率表现，本项目采用了多种收益率计算方式，包括（当期开盘价-上期开盘价）/上期开盘价、（当期收盘价-上期开盘价）/上期开盘价、（当期开盘价-上期收盘价）/上期收盘价、（当期收盘价-上期收盘价）/上期收盘价等四种收益率来计算累计收益率，而非仅使用一般的收盘价收益率。该方法更贴近策略在实际运行中的收益表现。

**（4）策略盈利性**

本研究在燃油主连期货的回测中展现了较好的盈利能力。在不考虑交易费用的情况下，从2020年1月2日至2024年10月31日，该策略的累计收益率达到231%，年化收益率为35%，最大回撤为24%。在考虑万分之一的交易费用后，累计收益率降至18%，年化收益率为4%，最大回撤为35%。