一、研究思路：

1. 根据研报内容计算191个因子值

这一部分，我们严格按照研报关于因子的计算方式实现的。

1. 将因子的异常值进行处理，并做标准化处理
2. 对于分钟线的因子值进行行业中性化处理，保留残差值作为新的因子值
3. 利用每分钟的因子值，求沪深300成分股在每个小时内的因子值的均值
4. 利用行情数据，求每只股票在每个小时内的均价，并按均价计算每只股票在8个小时内的收益值
5. 对时间数据进行遍历，计算每个小时的所有股票的因子均值和未来8个小时所有股票的收益率的相关系数。遍历完成后，可以得到每个因子的IC时间序列。
6. 根据每个因子的IC时间序列，参考IC的稳健性，根据IC值选择效果较为显著的因子
7. 对效果显著的因子进行回测

二 修正后的计划：

1. 根据研报内容计算191个因子值

这一部分，我们严格按照研报关于因子的计算方式实现的；

1. 对所有因子进行正交标准化处理，即行业中性，具体做法：通过因子值对行业因子进行多元回归，得到因子残差值，研报中还对风格因子进行了剔除；
2. 对于不同的预测周期，通过多元回归方程的到因子的191个因子收益率时间序列；
3. 通过过去T时间因子收益均值作为下期预测值，乘以当期每个因子的残差截面值，得到alpha收益截面的预测向量；
4. 类似第四步，对当期的收益和因子值进行多元回归得到实际的alpha收益截面向量；
5. 求因子IC,即预测值和实际值的相关系数；
6. 求ICIR，即IC值的T统计量，判断是否显著；
7. 得到有效因子，按照第四步的方式预测未来收益，即进行回测阶段，包括进行资金等权分配，扣除成本，停牌处理等，得到最后的模型结果。