Alpha因子构建注意事项：

1. Is\_valid确定股票是否合格

目前我们有几种is\_valid来作为股票筛选器，是以下条件的组合：股票上市未超过120交易日、股票摘牌、股票在该时期是ST或ST\*、股票停牌、股票复牌未超过20天、股票暂停上市、股票恢复上市未超过60天、股票在当日涨跌停

1. 该股票当日是否应该有数据

原则上该股票在某日是否有数据应该由因子编写者确定，由于因子不同于交易和回测，即使停牌也可以有因子值，这么做的原因是考虑后续进行横截面操作时，对于全市场的影响，编写者要考虑在停牌等时段时的因子值是否有意义。另外，对于某些时间序列的因子，编写者也要考虑踢掉不合适的时段，如计算Beta时将复牌后20天的数据剔除，再进行回归，若此时数据缺失过多则赋值Nan。在后面的流程中，合法天合法股票因子值将进行填充，目前主要使用行业平均值进行填充。

1. 未来函数问题

未来函数是编写因子中最重要的问题，因为一旦加入了未来函数会使效果变好，而进行实盘时则无法使用。查找未来函数除了编写时尽量使用rolling等方法不用到后面的数据，还可以检查最新日期是否能计算出因子值。

1. 异常值处理

计算因子时出现的异常值需要做一定处理，如分母为0, 得到Inf等，计算因子时的outlier也不应过多。计算完因子后应查看一下分布直方图，确定分布合理。

1. 初始化与后续更新一致性

计算因子要考虑第一次全量计算和后续更新计算的一致性，由于后续更新时只会截断到所需的最长时间，如MTM\_20d需要20天，此时有一些细节可能导致不同。在过去因子中发现pct\_change这个函数在rolling计算nan后的第一个数，使用的是nan后的第一个数与nan前的第一个数的变化率，若更新时的头部有nan会导致更新和初始化不一致，此时可以考虑使用shift进行操作。