一、构造预测因子

A股市场有一类投资者被称为技术流派，主要根据股票的K线图走势来分析股票及预测股价，比如红三兵、圆弧底、“V”字型底部、反转十字星、“W”底（双重底）、头肩底股价等。我们以“红三兵”为例，来构造预测因子。

红三兵的定义如下：

1.连续三天阳线；

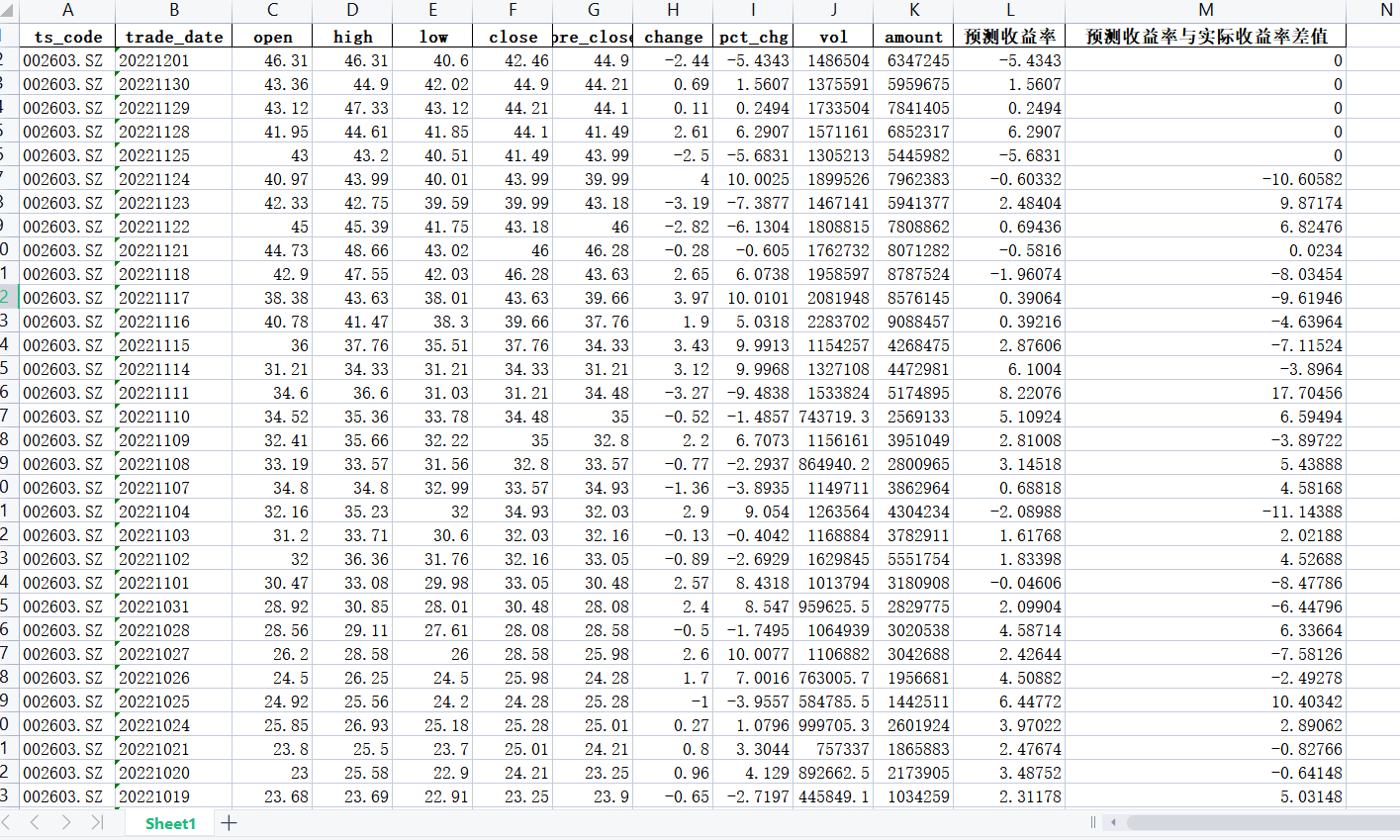
2.第二、三天开盘价在前一天阳线实体之内；

3.每天收盘价接近当天最高点；

4.三根阳线实体部分近似等长。

我们按照上述规则，定义Python函数，来判断是否出现“红三兵”的形态。

接下来，我们对以岭药业的日K线数据进行判断，找出其在考察期内，每天是否出现“红三兵”形态。如果某日出现“红三兵”形态，就给该股票打1分，即预测下一个交易日该股票的收益率为3%，否则给该股票打0分，用五日移动平均收益率来预测下一个交易日该股票的收益率。预测结果如下图所示：

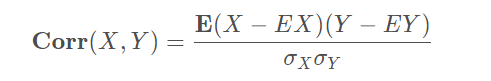


二、预测收益率分析

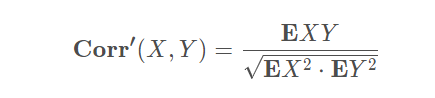
既已得到对未来收益的预测，那么，很自然的，下一步就是测试一下这个预测准不准了。

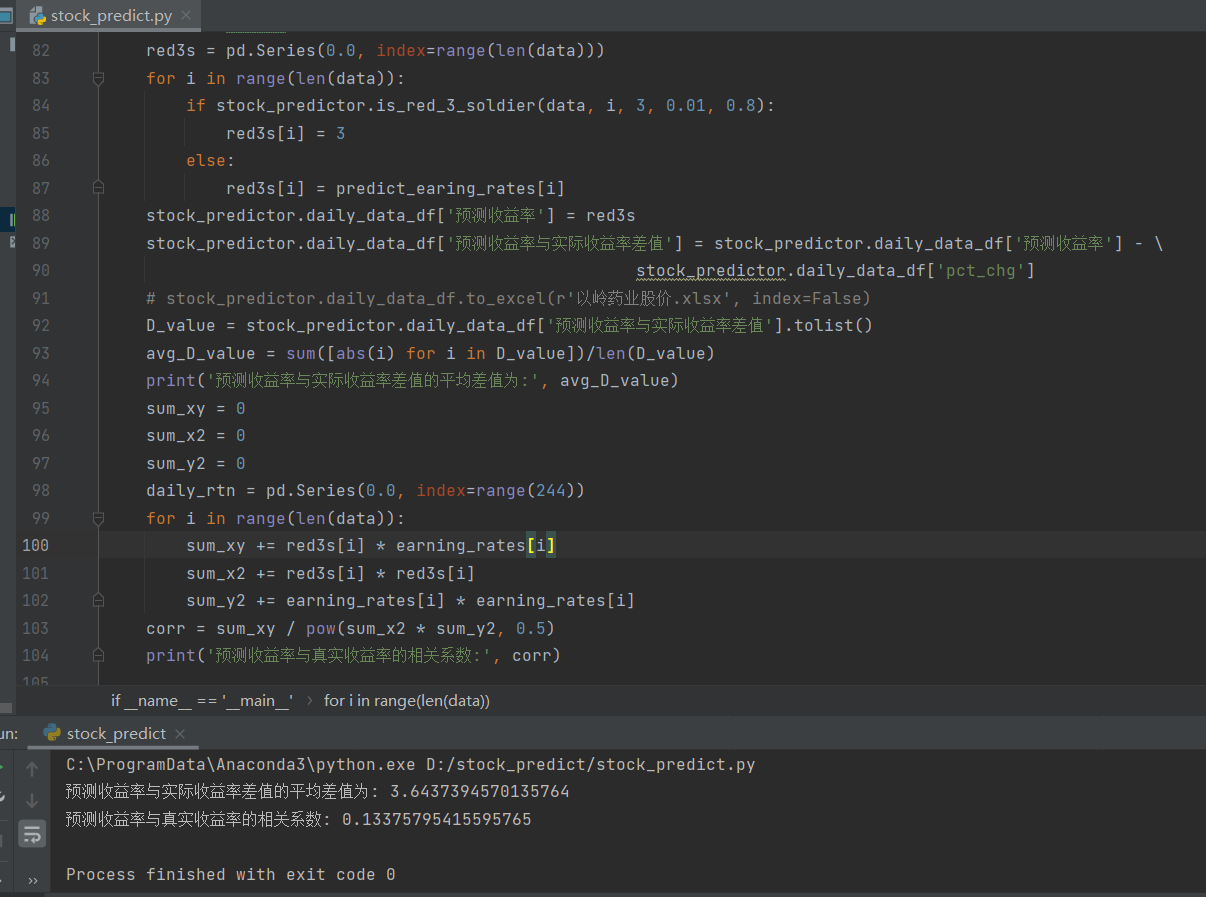
我们采用一个简单的方法，对预测模型的好坏，做个初步的判断。

我们将计算预测收益率与真实收益率的相关系数。当然，相关系数的定义多样，最通行的算法是：

/

然而我们通常会采用一个简化算法：





根据相关系数计算公式，我们进行python代码编写。运行代码，得到计算结果如上图所示，预测收益率与实际收益率差值的平均差值为: 3.6437394570135764

预测收益率与真实收益率的相关系数: 0.13375795415595765。

我们知道，相关系数r的绝对值一般在0.8以上,认为A和B有强的相关性 0.3到0.8之间,可以认为有弱的相关性，0.3以下,认为没有相关性。同时，预测收益率和实际收益率的平均差值达到了3%，而A股每天涨跌幅限制是10%,这个差异已经算是很大了。通过以上分析可以得知，这个预测结果并不理想，同时也证明用技术分析或者K线形态来预测股价这种方法并不可靠。

因为影响股价的因素有很多，比如宏观经济因素，行业因素，公司自身因素（盈利能力，经验状况）等等。而我们所选的以岭药业作为连花清瘟的生产产商，受疫情因素的影响，公司的股价也是节节攀升，近一年走出了翻倍行情，所以用K线形态对该股进行股价预测并不能达到理想效果。