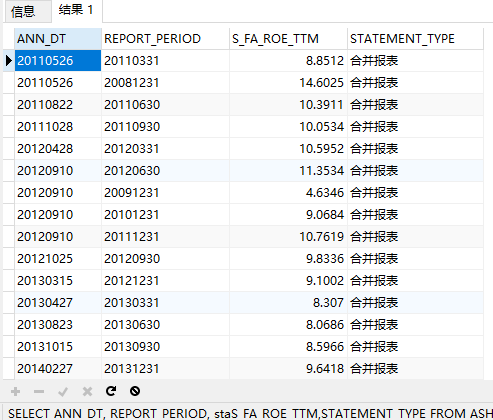
**ROE选股**

**不稳定问题的原因及改进**

1. 不稳定问题的原因
   1. 数据库中存在同一时间的重复数据。

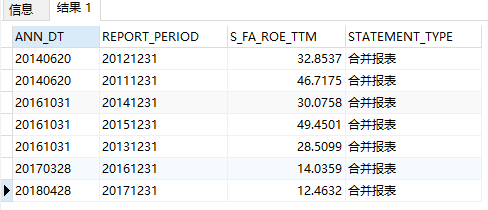
如图，在同一日期下，存在相同“发布日期”ANN\_DT、但是不同“报告截止日期”的记录。

整批取出数据时，数据排列顺序是不固定的，所以当去重drop\_duplicates时，每次drop的记录不相同，



* 1. 数据库中缺失数据需要剔除。
  2. 计算ROE\_stable时，没有按时间排序。

1. 改进
   1. 提取数据时，选用Report\_Period为12月31日的记录（即年报记录）



发现公布日期仍然很不规律。

采用折中办法，统一将次年3月31日选为年报公布日期。

* 1. 剔除缺失数据。

1. 效果
   1. 选股数量：41
   2. 选股记录：

600066.SH 600188.SH 603288.SH 600276.SH 002352.SZ 300070.SZ 600309.SH 000540.SZ 601166.SH 002236.SZ 000568.SZ 000157.SZ 000963.SZ 002024.SZ 601877.SH 002415.SZ 000983.SZ 002146.SZ 600570.SH 300144.SZ 002572.SZ 002081.SZ 600177.SH 002739.SZ 000858.SZ 000792.SZ 002602.SZ 600547.SH 002475.SZ 002385.SZ 601933.SH 300059.SZ 601021.SH 000895.SZ 000651.SZ 600406.SH 002411.SZ 300408.SZ 600153.SH 002142.SZ 002558.SZ

* 1. 最新选股：

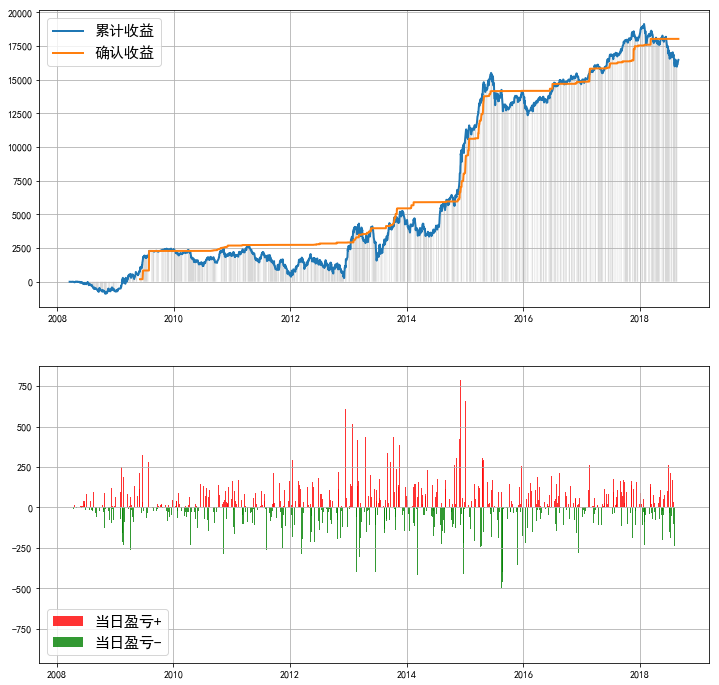


* 1. 定投策略回顾：

根据在每个行业内选出的股票，择其roe最高者进行定投。

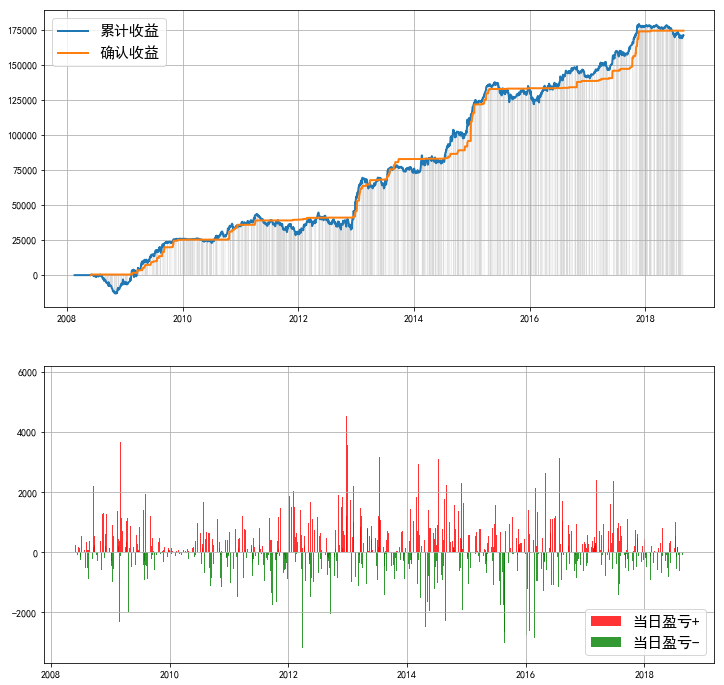
每周定投一直持续，直至止盈。止盈后，才重新审视当期该行业选择的股票，对其进行定投。如果一直没有止盈，定投一直持续。

* 1. 收益率曲线



1. 比对

之前的收益率曲线



与之前的比对发现，原先的收益率曲线更加平稳。

主要原因是：原先无意中使用了包含了季报的ROE数据，因而调整频率较快。而目前改进后，只使用年报数据，在次年3月底公布年报数据后决定调仓。相当于选股频率降到了每年。

其次，原先选出的股票更多，风险更分散。