

法律声明

本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

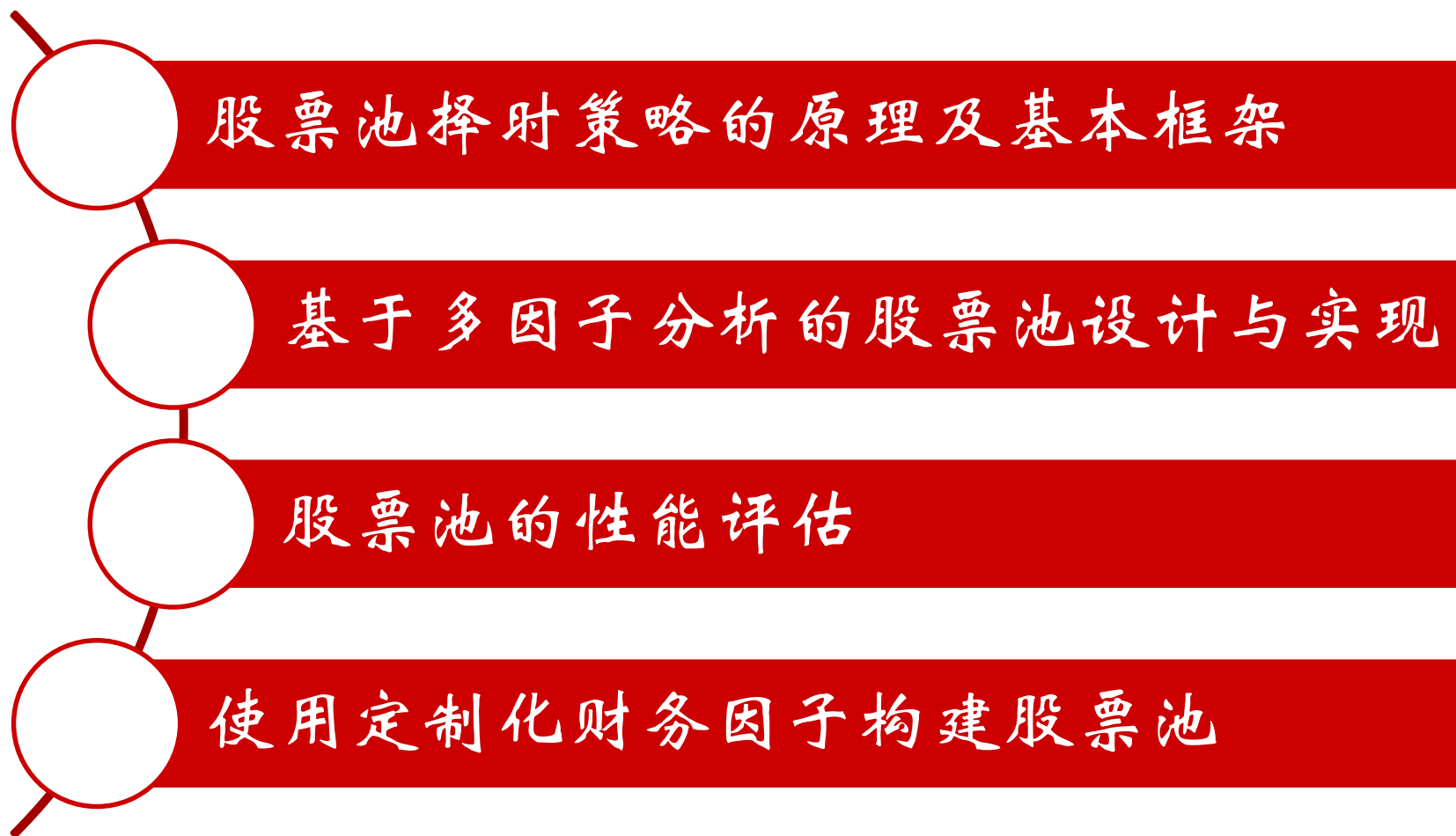


关注 小象学院

《进阶！量化交易实战：迭代式的量化策略研发》第3期

第3课：基于多因子分析的股票池开发 主 讲：刘英斐

课程大纲



股票池择时策略的原理及基本框架

一个交易系统的要素

- 市场：买卖什么？
- 头寸规模：买卖多少？
- 入市：什么时候买卖？
- 止损：什么时候放弃一个亏损的头寸？
- 退出：什么时候退出一个盈利的头寸？
- 战术：怎么买卖？

股票池的引入

- 买上证50中，最近突破5日均线的股票

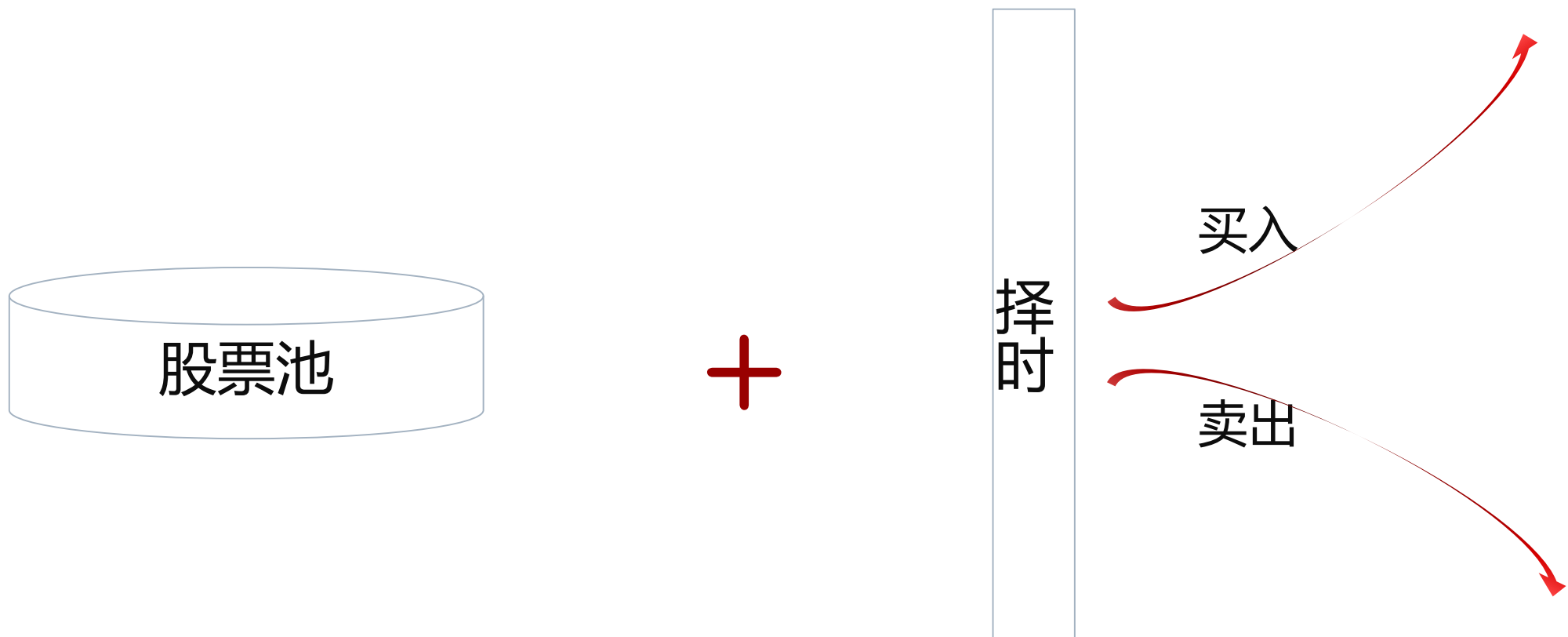


- 2017年是上证50的好年份
- 突破5日移动均线说明近期发力上涨

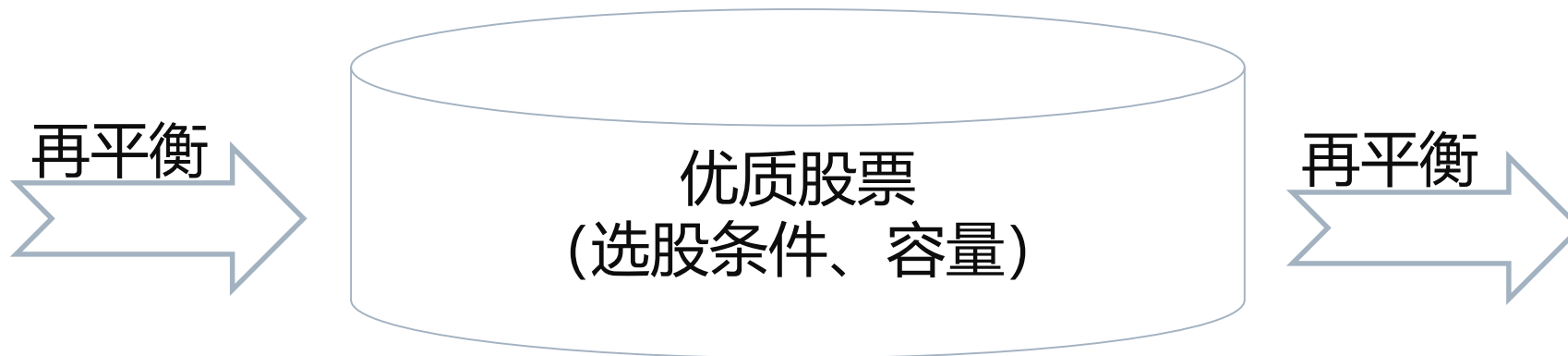


- 从优质股票的集合中，选择近期开始发力上涨的股票

股票池择时



股票池



股票池实例 – 低估值

- 选股条件
 - PE: 0~30
- 容量
 - 100
- 再平衡周期
 - 10个交易日

问题:

1. PE能选出低估值的股票吗?
2. 再平衡周期设定多少合适?

股票池实例 – 小市值低估值

□ 选股条件

- PE: 0~30

- 市值: 10亿~100亿

□ 容量

- 50

□ 再平衡周期

- 10个交易日

问题:

2017年以后的小市值股票还值得买入吗?

股票池实例 – 小市值低估值高质量

□ 选股条件

- PE: 0~30
- 市值: 10亿~100亿
- ROE: 10%~40%

□ 容量

- 50

□ 再平衡周期

- 10个交易日

问题:

ROE可以反映企业的质量，它一定是越大越好吗？

基于多因子分析的股票池设计与实现

巴菲特的投资密码

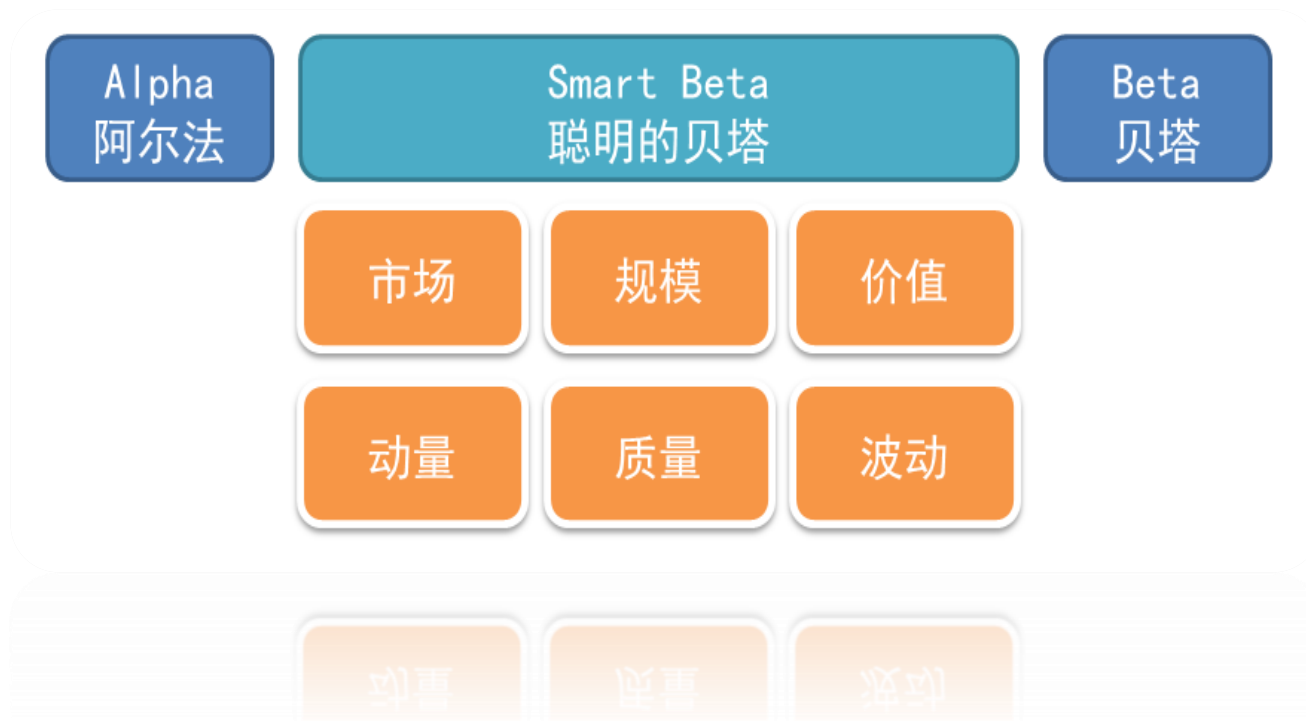
- 巴菲特反复说“用平常的价钱买一家很棒的公司远远强过用很棒的价钱买一家平常的公司”
- 什么是“很棒”的公司，巴菲特没有给出很清晰的量化标准
- 学界近年来发现“质量”可以被具体量化而成为一个新的风险因子
- 而且，质量投资也对多种资产有效，可以和价值投资，动量投资互为补充，组合成更加有利的策略。



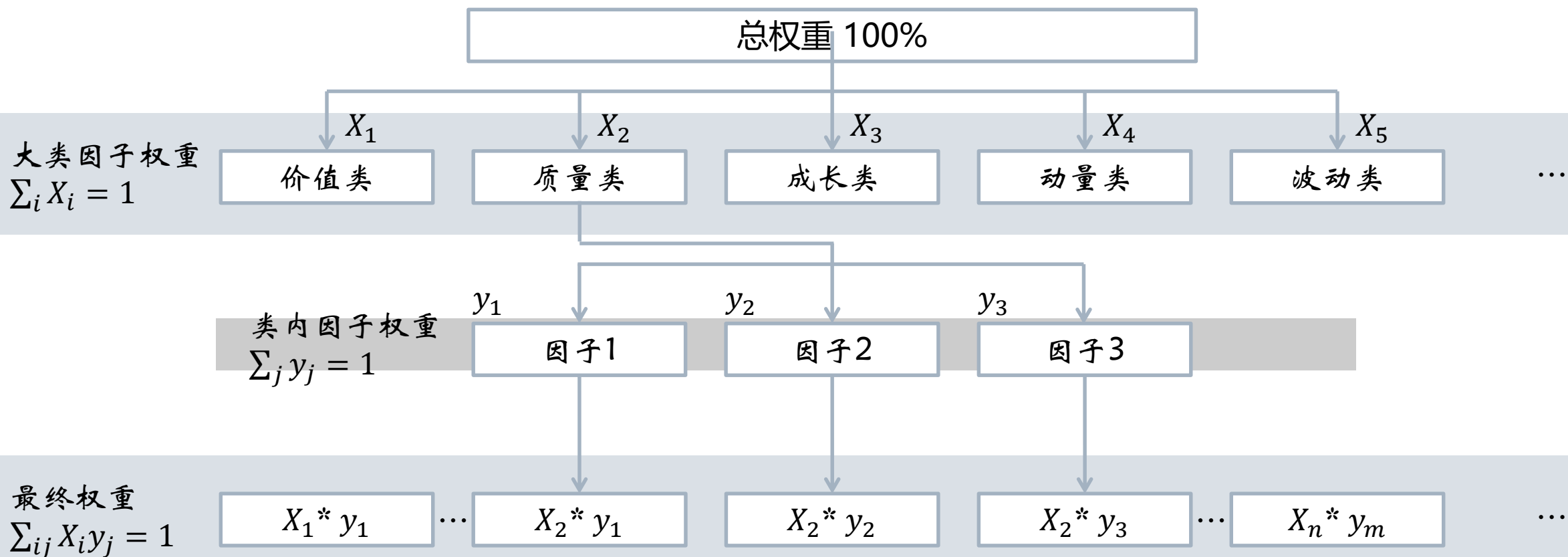
因子量化投资策略的组合

策略	买入操作	评价
规模投资	买入小盘股票	成熟市场小盘股的正alpha在学术成果公开后逐渐消失，仅可取得与风险匹配的回报
价值投资	买入低估值的股票	低估值的公司未必是高质量的公司，价格可能长期低迷，落入价值陷阱
动量投资	买入过去一段时间价格上涨的股票	需要判断市场方向，执行难度大
质量投资	买入高质量的股票	高质量体现在高利润、高成长、稳定和优良管理，价格可能偏高，但有正alpha
规模+价值	买入低估值的小盘股	过去很长时间跑赢低估值的大盘股
价值+动量	买入因外部事件的发生而价格开始逐渐上涨的低估值股票	互补关系，可以享受双重好处
价值+质量	买入具有合理价格的高质量股票	单纯的价值投资可能错过这类好股票
价值+质量+动量	买入价格开始上涨，但其估值仍然低于其高质量应得价格的股票	互补关系，可以享受三重好处

6因子模型来解释巴菲特



多因子模型的选股框架



对每只股票：

- 根据单因子分层进行打分
- 根据分层打分和因子权重进行股票得分的综合计算
- 根据综合打分进行排序，选取得分前10%股票进入股票池

典型的因子候选

大类因子	具体因子	因子描述
估值	EP	净利润(TTM)/总市值
估值	EPcut	扣除非经常性损益后净利润(TTM)/总市值
估值	BP	净资产/总市值
估值	SP	营业收入(TTM)/总市值
估值	NCFP	净现金流(TTM)/总市值
估值	OCFP	经营性现金流(TTM)/总市值
估值	DP	近 12 个月现金红利(按除息日计)/总市值
估值	G/PE	净利润(TTM)同比增长率/PE_TTM
成长	Sales_G_q	营业收入(最新财报, YTD)同比增长率
成长	Profit_G_q	净利润(最新财报, YTD)同比增长率
成长	OCF_G_q	经营性现金流(最新财报, YTD)同比增长率
成长	ROE_G_q	ROE(最新财报, YTD)同比增长率
财务质量	ROE_q	ROE(最新财报, YTD)
财务质量	ROE_ttm	ROE(最新财报, TTM)
财务质量	ROA_q	ROA(最新财报, YTD)
财务质量	ROA_ttm	ROA(最新财报, TTM)
财务质量	grossprofitmargin_q	毛利率(最新财报, YTD)
财务质量	grossprofitmargin_ttm	毛利率(最新财报, TTM)
财务质量	profitmargin_q	扣除非经常性损益后净利润率(最新财报, YTD)
财务质量	profitmargin_ttm	扣除非经常性损益后净利润率(最新财报, TTM)
财务质量	assetturnover_q	资产周转率(最新财报, YTD)
财务质量	assetturnover_ttm	资产周转率(最新财报, TTM)
财务质量	operationcashflowratio_q	经营性现金流/净利润(最新财报, YTD)
财务质量	operationcashflowratio_ttm	经营性现金流/净利润(最新财报, TTM)

杠杆	financial_leverage	总资产/净资产
杠杆	debtequityratio	非流动负债/净资产
杠杆	cashratio	现金比率
杠杆	currentratio	流动比率
市值	ln_capital	总市值取对数
动量反转	HAlpha	个股 60 个月收益与上证综指回归的截距项
动量反转	return_Nm	个股最近 N 个月收益率, N=1, 3, 6, 12
动量反转	wgt_return_Nm	个股最近 N 个月内用每日换手率乘以每日收益率求算术平均值, N=1, 3, 6, 12
动量反转	exp_wgt_return_Nm	个股最近 N 个月内用每日换手率乘以函数 $\exp(-x_i/N/4)$ 再乘以每日收益率求算术平均值, x_i 为该日距离截面日的交易日的个数, N=1, 3, 6, 12
波动率	std_FF3factor_Nm	特质波动率——个股最近 N 个月内用日频收益率对 Fama French 三因子回归的残差的标准差, N=1, 3, 6, 12
波动率	std_Nm	个股最近 N 个月的日收益率序列标准差, N=1, 3, 6, 12
股价	ln_price	股价取对数
beta	beta	个股 60 个月收益与上证综指回归的 beta
换手率	turn_Nm	个股最近 N 个月内日均换手率(剔除停牌、涨跌停的交易日), N=1, 3, 6, 12
换手率	bias_turn_Nm	个股最近 N 个月内日均换手率除以最近 2 年内日均换手率(剔除停牌、涨跌停的交易日)再减去 1, N=1, 3, 6, 12
情绪	rating_average	wind 评级的平均值
情绪	rating_change	wind 评级(上调家数-下调家数)/总数
情绪	rating_targetprice	wind 一致目标价/现价-1
股东	holder_avgpctchange	户均持股比例的同比增长率
技术	MACD	经典技术指标(释义可参考百度百科), 长周期取 30 日, 短周期取 10 日, 计算 DEA 均线的周期(中周期)取 15 日
技术	DEA	
技术	DIF	
技术	RSI	经典技术指标, 周期取 20 日
技术	PSY	经典技术指标, 周期取 20 日
技术	BIAS	经典技术指标, 周期取 20 日

资料来源: Wind, 华泰证券研究所

单因子检验 - 零投资组合检验

Step1

- 选定研究因子，并使用每一时期的该因子敞口对股票进行排序，对排序结果进行五（或十）分为划割

Step2

- 在第一个五分位里构造相同权重的股票组合，同时在最后一个五分位里相同权重的股票组合，第一个五分位是根据该因子排序的前1/5的股票，最后一个五分位是根据该因子的最后1/5的股票

Step3

- 计算两个股票组合的月度收益，并计算两个组合收益率的差值，该差值即为零投资组合的收益。这里的零投资组合是通过买进最顶层划分中的投资组合并卖空最底层划分中的投资组合得到的，它被称为零投资组合是因为从理论上讲构造该组合不需要资本。

Step4

- 在计算过零投资组合的收益后，我们进行统计检验，以考察等权重股票组合与零投资组合收益是否存在显著不同

多因子模型的原则



因子尽量少

从系统复杂度的角度考虑，用尽量少的因子来描述尽量多的收益回报差别，既降低了系统的复杂度，又降低了过拟合的风险



因子间尽量独立

如果采用大量相关性很强的因子来描述收益回报的差别，不但造成冗余，还会带来系统复杂度的显著增加，没有必要



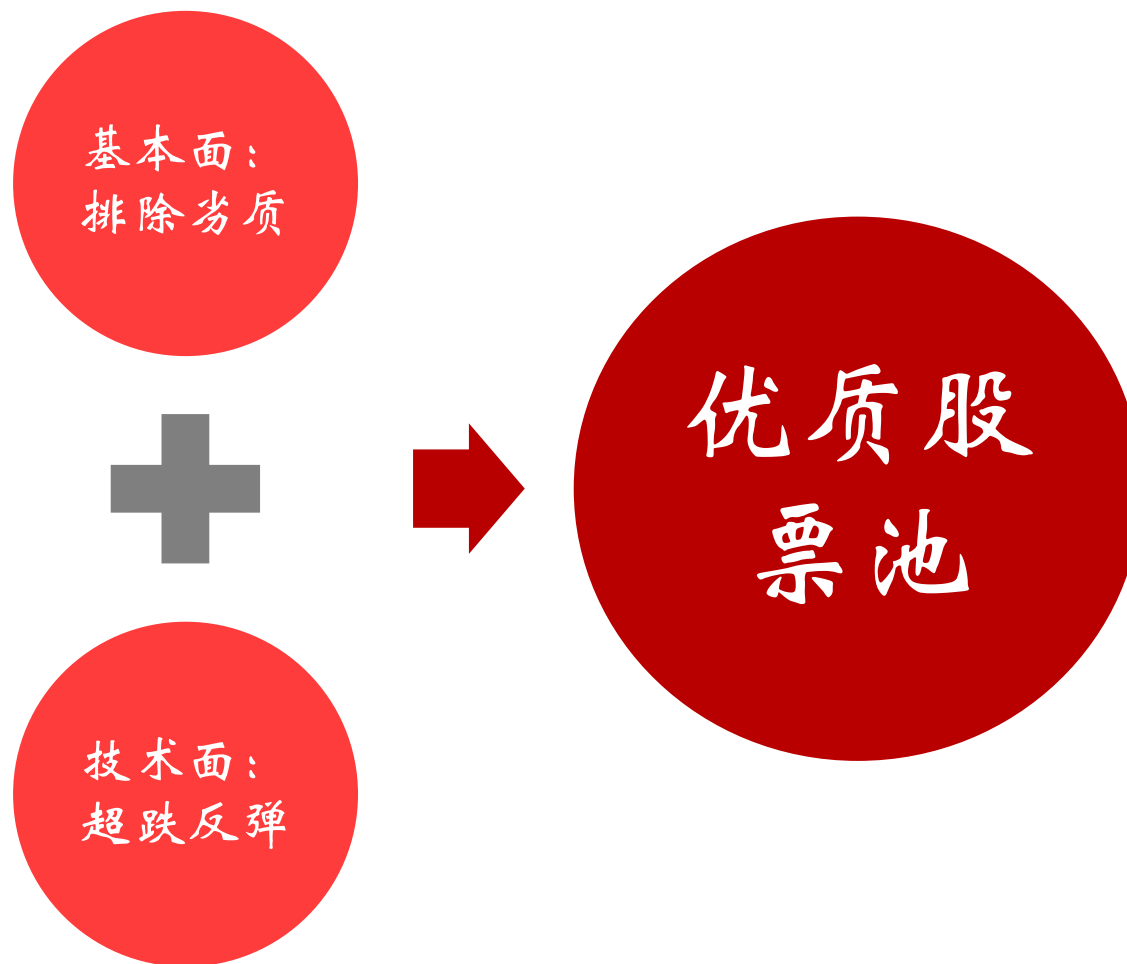
时间尽可能长

因子的回报不能用偶然的巧合来解释，所以必须要经过足够长时间的检验，至少要跨越牛熊周期

休息一下
10分钟后回来

股票池的性能评估

股票池设计



基本面

- ❑ 剔除总市值排名最小的10%的股票
- ❑ 剔除PE(TTM) 小于0或大于100的股票
- ❑ 剔除ST股票

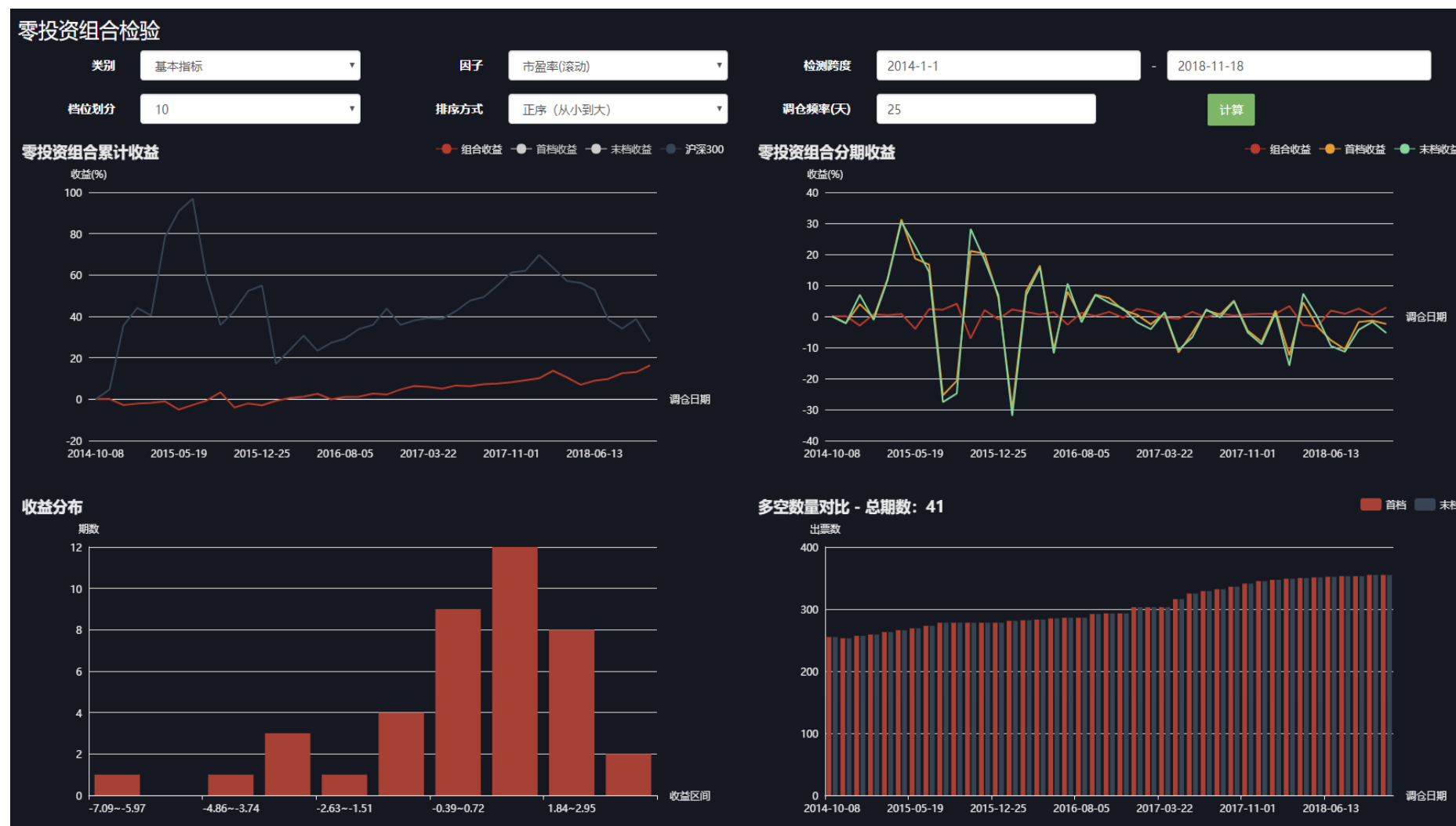
技术面

□ 取25日跌幅前10%的股票

单因子检验 - 市值



单因子检验 - PE



股票池定义

□ 选股条件

- 剔除总市值排名最小的10%的股票
- 剔除PE(TTM) 小于0或大于100的股票
- 剔除ST股票
- 取25日跌幅前10%的股票

□ 再平衡周期

- 25个交易日

□ 容量

- 无限制

股票池检验

□ 股票池第一次构建

- 买入股票池中所有的股票

□ 股票池每次再平衡

- 卖出持仓股中已经不在股票池中的股票
- 买入股票池中新纳入的股票

□ 仓位

- 平均分配可用资金

源代码 – 财务数据

```
df = get_fundamentals(query(valuation,indicator), context.current_dt.strftime("%Y-%m-%d"))
raw_data = []
for index in range(len(df['code'])):
    raw_data_item = {
        'code'      :df['code'][index],
        'market_cap':df['market_cap'][index],
        'pe_ratio'  :df['pe_ratio'][index]
    }
    raw_data.append(raw_data_item)
```

源代码 – 基本面

按照财务信息中的总市值降序排序

```
raw_data = sorted(raw_data, key = lambda item: item['market_cap'], reverse=True)
```

剔除总市值排名最小的10%的股票

```
filtered_market_cap = raw_data[:int(len(raw_data) * 0.9)]
```

剔除PE TTM 小于0或大于100的股票

```
filtered_pe = []
```

```
for stock in filtered_market_cap:
```

```
    if stock['pe_ratio'] == None or math.isnan(stock['pe_ratio']) or  
        float(stock['pe_ratio']) < 0 or float(stock['pe_ratio']) > 100:
```

```
        continue
```

```
    filtered_pe.append(stock['code'])
```

源代码 – 技术面

加载标的的涨跌幅信息

```
change_pct_dict_list = load_change_pct_data(context,filtered_pe)
```

按照涨跌幅升序排序

```
change_pct_dict_list = sorted(change_pct_dict_list,key = lambda  
    item:item['change_pct'], reverse=False)
```

取跌幅前10%的股票

```
change_pct_dict_list = change_pct_dict_list[0:(int(len(change_pct_dict_list)*0.1))]
```

源代码 – 合成股票池

获取最终的股票池

```
g.stock_pool = []
```

```
for stock in change_pct_dict_list:
```

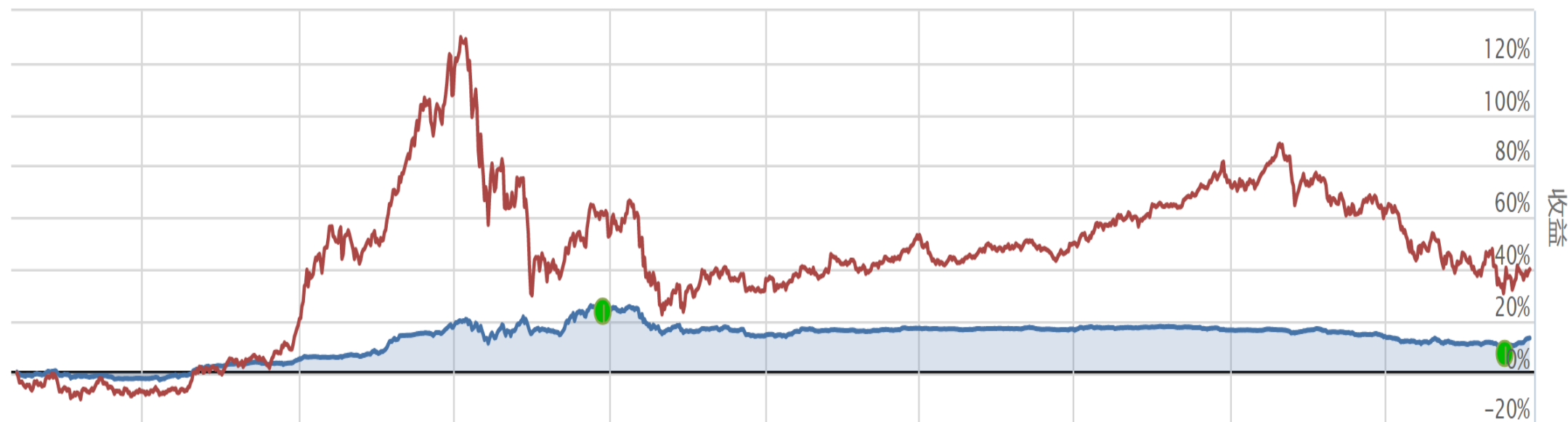
```
    g.stock_pool.append(stock['code'])
```


股票池检验结果

策略收益	策略年化收益	基准收益	Alpha	Beta	Sharpe	胜率	盈亏比	最大回撤 ?	其他指标
13.21%	2.64%	39.81%	-0.020	0.189	-0.225	0.528	1.636	12.947%	>

缩放: 1个月 1年 全部 ■ 策略收益 ■ 基准收益 ■ 超额收益 ● 普通轴 ● 对数轴 □ 超额收益 ?

时间: 2013-12-30 - 2018-11-16



股票池优化探讨

- 引入质量因子：ROE
- 引入资金面因子：连续N日主力资金净流入
- 引入综合财务指标：财务优秀

使用定制化财务因子构建股票池

财务分析分类

风险

运营效率

盈利能力及质量

前景与评估

风险

- ☐ 资产负债率
- ☐ 净资产负债率
- ☐ 流动比率
- ☐ 速动比率
- ☐ ...

运营效率

- ☐ 总资产周转率
- ☐ 固定资产周转率
- ☐ 存货周转率
- ☐ 应收账款周转率
- ☐ ...

盈利能力及质量

- 净资产收益率
- 总资产收益率
- 每股收益
- 销售收现比率
- 净利润现金含量比率
- 毛利率、净利率
- ...

前景与评估

- 营业收入增长率
- 净利润增长率
- 总资产增长率
- 净资产增长率
- ...

财务评分卡框架

类型	权重	总分	财务指标	分数
风险	10%	100	资产负债率	40
			流动比率	60
运营效率	25%	100	总资产周转率	50
			存货周转率	50
盈利能力及质量	40%	100	净资产收益率	20
			每股收益	30
			销售收现比率	20
			净利润现金含量比率	30
前景与评估	25%	100	营业收入增长率	50
			净利润增长率	50

盈利能力及质量评价因子

因子评分 = 风险得分*10% + 运营效率得分*25% + 盈利能力及质量得分*40% + 前景与评估得分*25%

	总评分排名占比	评级
财务评级	前10%	A级
	前10%—30%	B级
	中30%—60%	C级
	后60%—80%	D级
	后80%—100%	E级

财务指标计算-资产负债率

财务指标	指标值	得分
资产负债率	40%—60%	40
	30%—40%	30
	0—30%	20
	60-80%	10
	80%—100%	5

财务指标计算-流动比率

财务指标	指标值排名占比	得分
流动比率	前10%	60
	前10%—30%	50
	中30%—60%	40
	后60%—80%	25
	后80%—100%	10

财务指标计算-总资产周转率

财务指标	指标值排名占比	得分
总资产周转率	前10%	50
	前10%—30%	40
	中30%—60%	30
	后60%—80%	20
	后80%—100%	10

财务指标计算-存货周转率

财务指标	指标值排名占比	得分
存货周转率	前10%	50
	前10%—30%	40
	中30%—60%	30
	后60%—80%	20
	后80%—100%	10

财务指标计算-净资产收益率

财务指标	指标值排名占比	得分
净资产收益率	前10%	20
	前10%—30%	15
	中30%—60%	10
	后60%—80%	5
	后80%—100%	1

财务指标计算-每股收益

财务指标	指标值排名占比	得分
每股收益	前10%	30
	前10%—30%	25
	中30%—60%	15
	后60%—80%	10
	后80%—100%	5

财务指标计算-销售收现比率

财务指标	指标值	得分
销售收现比率	大于等于1	20
	小于1	10

财务指标计算-净利润现金含量比率

财务指标	指标值排名占比	得分
净利润现金含量比率	前10%	30
	前10%—30%	25
	中30%—60%	15
	后60%—80%	10
	后80%—100%	5

财务指标计算-营业收入增长率

财务指标	指标值排名占比	得分
营业收入增长率	前10%	50
	前10%—30%	40
	中30%—60%	30
	后60%—80%	20
	后80%—100%	10

财务指标计算-净利润增长率

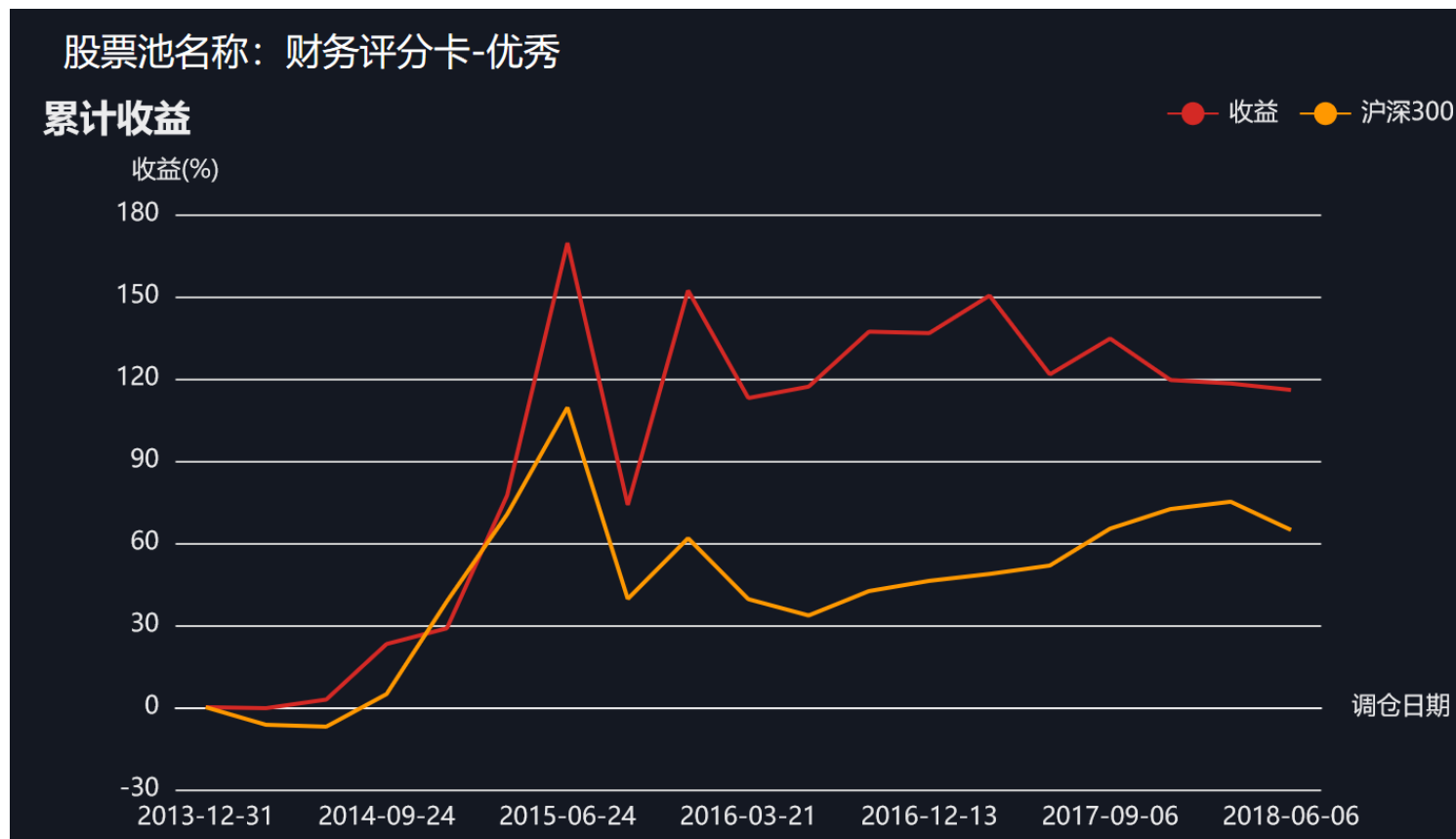
财务指标	指标值排名占比	得分
净利润增长率	前10%	50
	前10%—30%	40
	中30%—60%	30
	后60%—80%	20
	后80%—100%	10

实例-恒生电子

资产负债率	35%	30
流动比率	排名 35%	40
得分	$(30+40)*0.1=7$ 分	
总资产周转率	排名 25%	40
存货周转率	排名 40%	30
得分	$(40+30)*0.25=17.5$ 分	
净资产收益率	排名 50%	10
每股收益	排名 32%	15
销售收现比率	大于 1	20
净利润现金含量比率	排名 46%	15
得分	$(10+15+20+15)*0.4=24$ 分	
营业收入增长率	排名 32%	30
净利润增长率	排名 70%	20
得分	$(30+20)*0.25=12.5$ 分	
总得分	$7+17.5+24+12.5=61$ ，排名 35%，C 级	

股票池实例

财务评级：A级 股票池调整周期：60交易日



THANKS