

Alternative Data & Factor Investing



LU Song CFA CQF

Head of Quantitative Research



ChinaScope (Shanghai) Limited

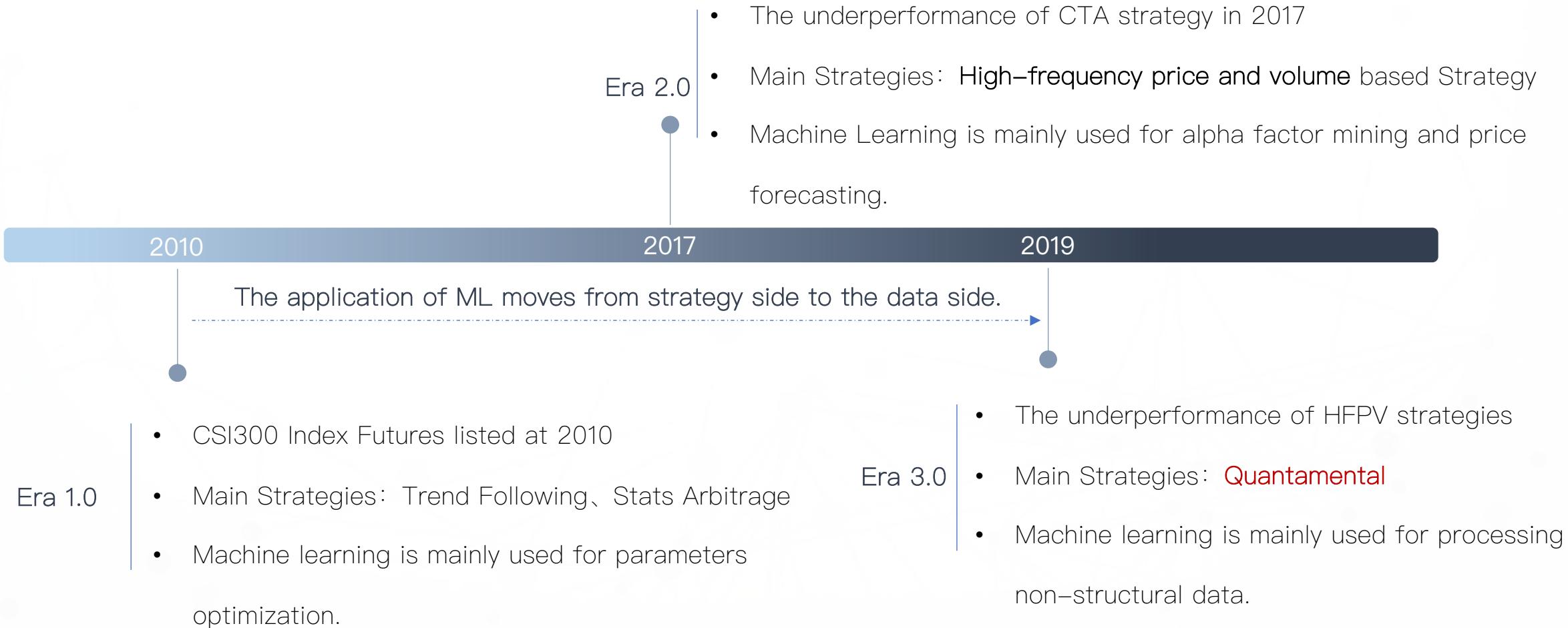
Content

Part 1. Theory

Part 2. Tools

Part 3. Application

Part 1. Theory



Traditional Data

- Fundamentals: Macroeconomic \ Financial Statement
- Market: Price and Volume \ Money Flow
- Other: Analyst Expectation \ Mutual Fund Holdings

Alternative Data

- News Analytics and Sentimental Data ★
- Industrial Chain Network Data ★
- Supply Chain Network Data ★
- ESG: Environmental, Social and Governance
- E-commerce retail data
- Geolocation\Satellite\Weather Forecasting

Q: How can we apply the alternative data in quantitative investing?

A: Shaping\Aggregating the data to be suitable for the traditional investment framework.

Q: What is the traditional investment framework
in quantitative investment field?

- 1 Timing Model: When to buy/sell?
- 2 Stock Screening Model: What to buy/sell? (Fixed Rebalance Period)

Smart Beta

Factor Model

Alternative Risk Premium

Quantamental

1950 —● Markowitz: Mean–Variance Portfolio Theory

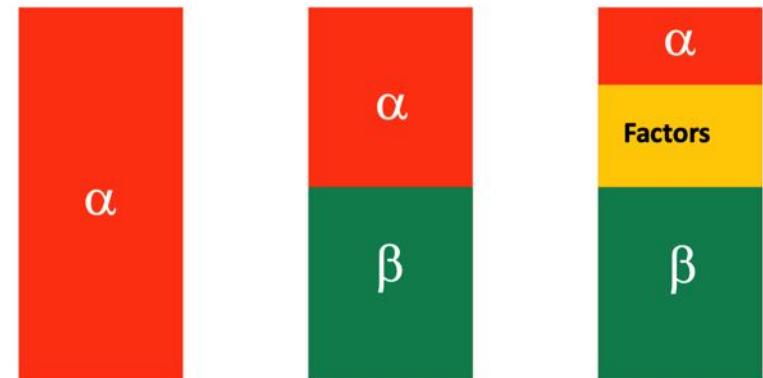
1960 —● CAPM: $E(r_i) = r_f + \beta_i[E(r_M) - r_f]$

1970 —● APT: $E(r_i) = \lambda_0 + \lambda_1 b_{i1} + \lambda_2 b_{i2} + \dots + \lambda_k b_{ik}$

1990 —● Fama–French Three Factors Model:

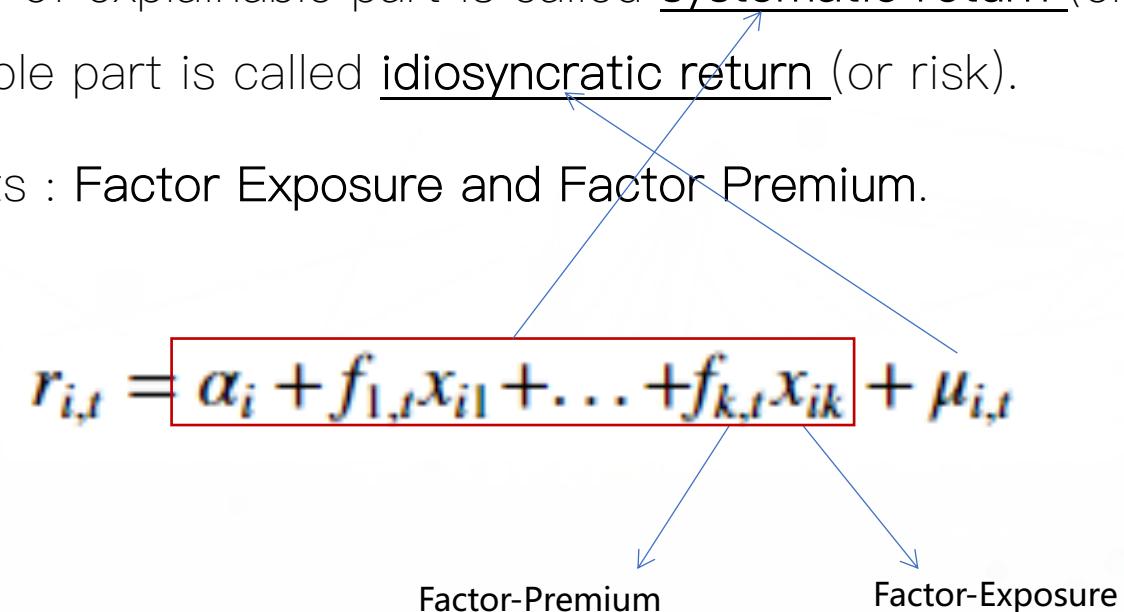
$$r_i = \alpha_i + \beta_i[E(R_m) - R_f] + b_s \cdot SMB + b_v \cdot HML + \varepsilon_i$$

2000 —● Barra Multifactor Risk Model. (Most Commonly used in practice).



Q: What is Factor Model?

- Factor Models are used to decompose the stocks return into two parts: model-explainable part and model-unexplainable part.
- The return (or risk) of explainable part is called systematic return (or risk), and the return (or risk) of unexplainable part is called idiosyncratic return (or risk).
- Two Main Concepts : Factor Exposure and Factor Premium.

$$r_{i,t} = \alpha_i + f_{1,t}x_{i1} + \dots + f_{k,t}x_{ik} + \mu_{i,t}$$


The diagram illustrates the decomposition of a stock's return ($r_{i,t}$) into its systematic component and its idiosyncratic component. The equation $r_{i,t} = \alpha_i + f_{1,t}x_{i1} + \dots + f_{k,t}x_{ik} + \mu_{i,t}$ is shown at the top. A red box highlights the term $\alpha_i + f_{1,t}x_{i1} + \dots + f_{k,t}x_{ik}$. Three blue arrows point from this highlighted term to three labels below: "Factor-Premium" (pointing to α_i), "Factor-Exposure" (pointing to $f_{1,t}x_{i1} + \dots + f_{k,t}x_{ik}$), and "Error Term" (pointing to $\mu_{i,t}$).

Q: What are Factor Exposure and Factor Premium?

- Factor exposure measures to what extent the factor will affect the stock's return.
- Factor premium measures the corresponding return of one unit factor exposure.
- At some time t , all stocks have the same factor premium but different factor exposure.
- The risk related to the factors can be hedged by long–short portfolio.

Portfolio returns can be broken down into Macro Factors, Style Factors and Pure Alpha

| Market Exposure | Active Non-Market Exposure | |
|---|---|--|
| Macro Factors | Style Factors | Pure Alpha |
| <p>Economic growth Bearing exposure to the business cycle</p> <p>Rates Bearing risk of rising rates</p> <p>Inflation Bearing risk to changes in prices</p> <p>Credit Bearing risk of company default</p> <p>Emerging markets Exposure to political and sovereign risks</p> <p>Liquidity Exposure of holding illiquid assets</p> | <p>Value Buying cheap</p> <p>Momentum Taking on trends</p> <p>Quality Seeking stable earnings</p> <p>Minimum Volatility Flight to safety</p> <p>Size Buy small</p> | <p>Security selection</p> <p>Country and industry selection</p> <p>Market and factor timing</p> |

Q: Why do we need the factor model? How the factor model can be used for?

- The factor model can help us modelling the market uncertainty by decomposing the uncertain stock return into several explainable factors.
- Forecasting the risk or return of several thousands stocks can now be turned into forecasting the exposure and return of several factors.
- More specific, the factor model can be used in two aspects: **risk model** and **return model**.

Return Model : Forecast Return

$$r_{i,t} = f_{1,t}x_{i1,t} + \dots + f_{k,t}x_{ik,t} + \mu_{i,t}$$

$$r_i = \mathbf{x}'_i \mathbf{f} + \mu_i$$

1xK Kx1

$$\mathbf{r} = \mathbf{x}' \mathbf{f} + \boldsymbol{\mu}$$

Nx1 NxK Kx1 Nx1

$$\mathbf{x}_i = (x_1, x_2, \dots, x_K)'$$

Kx1

$$\mathbf{f} = (f_1, f_2, \dots, f_k)'$$

$$\mathbf{r} = (r_1, r_2, \dots, r_N)'$$

Nx1

$$\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$$

KxN

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_N)'$$

Nx1

Risk Model : Forecast Risk

$$V(r_{i,t}) = V(f_1x_{i1,t} + \dots + f_kx_{ik,t} + \mu_{i,t})$$

$$\mathbf{V} = \mathbf{x}' \Sigma \mathbf{x} + \Lambda$$

Covariance matrix of stock return : $\mathbf{V} = Var(\mathbf{r})$

Covariance Matrix of factor return : $\Sigma = Var(\mathbf{f})$

Covariance Matrix of idiosyncratic return : $\Lambda = Var(\boldsymbol{\mu})$

Assume we have a portfolio of N stocks, we need forecast the risk (variance) of the portfolio.

Under MPT: N returns, N standard deviation, N(N-1)/2 correlation coefficients.

Under Factor Model: K factor returns, K standard deviation, K(K-1)/2 correlation coefficients.

Q: How can we apply the alternative data in quantitative investing?

$$r_{i,t} = f_{1,t}x_{i1,t} + \dots + f_{k,t}x_{ik,t} + \mu_{i,t}$$

Find more factors to make the stock return more explainable



Q: What are the steps of mining factors by using alternative data?

- Alternative Data Factorizing: The most critical step.
- Factor Correlation Testing: Test whether the factor is correlated with other factors such as momentum or size factors.
- Factor Significance Testing: IC analysis to test the correlation of the factor and the stocks' return.
- Factor Performance Testing: Back-test with the consideration of trading cost

1 Data Factorizing: The most critical step.

Don't have a standard procedure. Two different approaches:

- Top to Bottom: The factor is meaningful and has some economical logic. It expresses some sort of inefficiency of the market. We should understand the data and the model in depth.
- Bottom to top: The factor is the product of data mining and does not have any logical. Its outstanding performance is just for overfitting and survival bias. ★

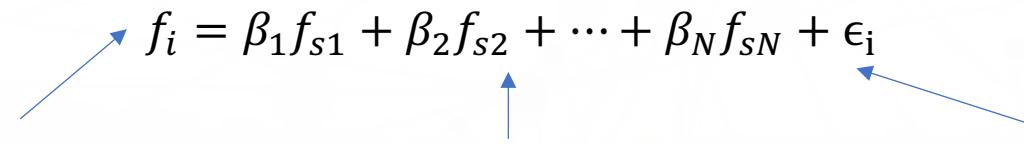
2 Factor Correlation Testing

It's not easy to find an effective factor. Probably, its good performance is only because of its high correlation with some existed factors. The high correlation may also be the reason of its bad performance.

- Style Factor Neutralization: (To strip the impact from other factors.)
 - Test the correlation of the standardized candidate factor and some common style factors.
 - If highly correlated, do the neutralization, which is also called Orthogonal Processing

$$f_i = \beta_1 f_{s1} + \beta_2 f_{s2} + \cdots + \beta_N f_{sN} + \epsilon_i$$

candidate factor Style factors Neutralized candidate factor



The diagram illustrates the decomposition of a candidate factor f_i into style factors and a neutralized component. A blue arrow points from the term ϵ_i in the equation to the text "Neutralized candidate factor". Another blue arrow points from the term f_i to the text "candidate factor". A third blue arrow points from the term f_{sN} in the equation to the text "Style factors".

3 Factor Significance Testing

Two approaches of doing the test:

- Information Coefficient (IC) :
 - Normal IC: the cross-sectional correlation of the factor value and stock return
 - Rank IC: the cross-sectional correlation of the factor value ranking and stock return ranking.
- T-Statistics:

$$\bullet \quad r_i = \alpha + \beta f_i + \epsilon_i$$



The t-statistics of the coefficient

4 Factor Performance Testing

Two approaches of doing the test:

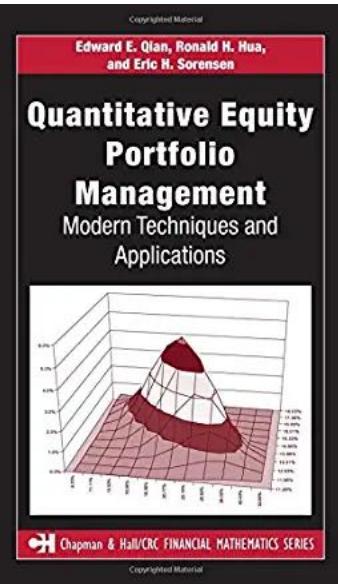
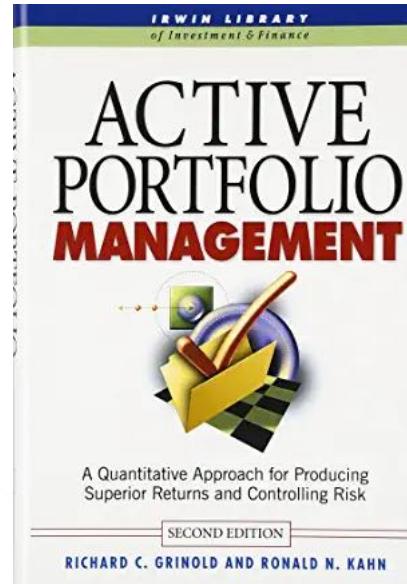
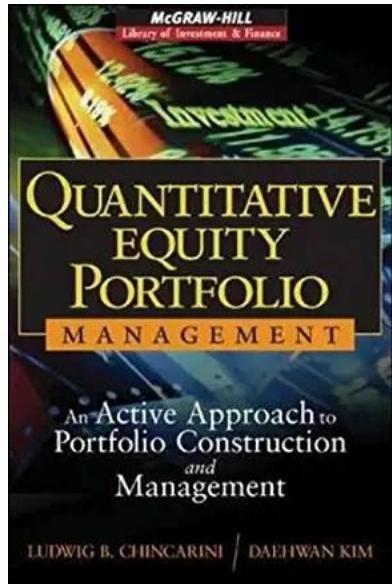
- Quantile Return Analysis:
 - Split the stock into several groups, and calculate the performance of each group.
 - A good factor will clearly distinguish the potential stock with better performance from those with worse performance.
- Quantile Turnover Analysis:
 - Factor Autocorrelation
 - Actual Turnover get from the former step

Part 2. Tools

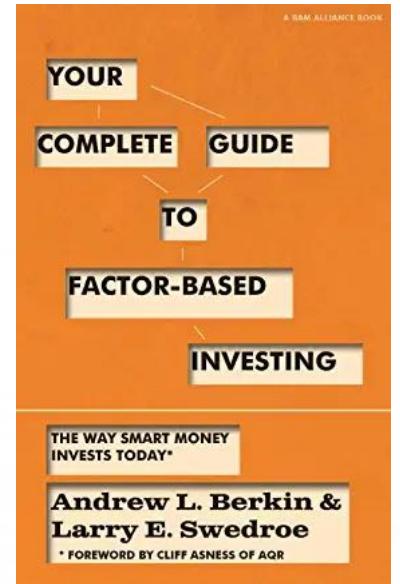
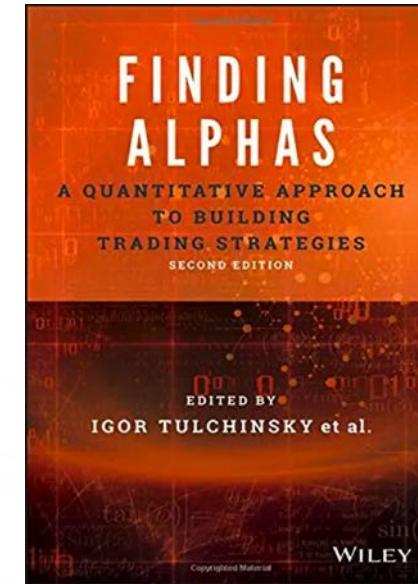
Tools —— Learning Resources

Construct a solid knowledge of factor modeling with these books:

Theoretical



Practical



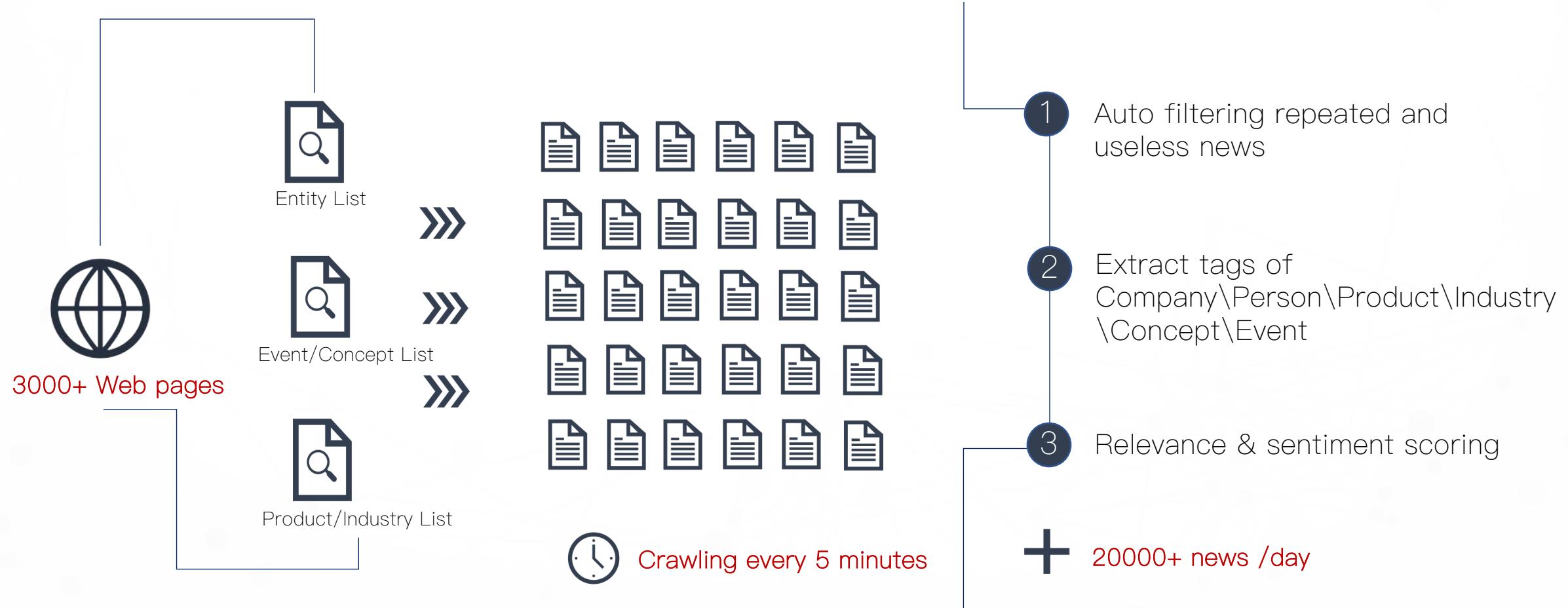
Recommended reading order



- Research Tools: Jupyter Lab
- Basic Data Source: WindPy\Tushare\JQData
- Data Storage: MySql (SqlAlchemy)\Mongo (Pymongo)
- Data Manipulation: Pandas\Numpy\Json
- Statistical Model: Statsmodels\Scipy\CvxOpt\Scikit-Learn
- Factor Testing: AlphaLens
- Strategy BackTesting: BackTrader\Pyfolio

Part 3. Application

News Analytics Data – SmarTag Intelligence News System



CSCI service: a comprehensive interpretation of Fosun's happy travel

Positive 92.80%

Scoring the article's sentimental

"Revisit city" brand release: deep definition of happy "revisit city" brand is a brand-new tourism destination brand launched by Fosun travel culture after successfully building Sanya Atlantis. The English name of "revisit city" is foliday Town, which is an important brand of "destination" one of the three business sectors of Fosun travel culture. In April 2018, Atlantis achieved great success after its launch in Sanya. Based on this, Fosun tourism and culture will create a new tourism destination. In the future, more tourism destination projects will be named after "Fosun city", which also marks the upgrading of leisure and vacation business under Fosun tourism and culture. Designed by Gensler, the world's No.1 design company, Taicang revisit city is planned to be a modern international resort town, which includes two tourism core IPS of alpine ice and snow world and revisit Sports Park, as well as two leisure and vacation contents of Club Med joyview City Resort Hotel and European style business street. The "revisit city" has tourism products with multi-national characteristics, and has established international communities. In the future, it will create the mode of global tourism + community living, or really empower the small cities that can live, upgrade the scene of leisure and vacation, truly break through the single scenic spot sightseeing experience, and upgrade the vacation mode to global vacation and full-time vacation. As a brand-new tourism destination brand under Fosun travel culture, Taicang Fuyou city will still maintain the one-stop service characteristics of its products, and the concept of "immersive experience" will be more prominent in the Fuyou City, which combines many high-quality products and characteristics, so that tourists do not have to rush to the next station, which can carry all the happiness. In 2018, five new resorts in Hokkaido, such as tomamu and Anji, were opened in the world. In 2019-2021, 11 new club med resorts are planned to be opened in the world. At the same time, Club Med joyview products are launched in China to position peripheral tourism and further enrich the tourism and holiday industry. Its differentiated positioning and "one price all inclusive" mode will lead the development of the domestic leisure and holiday market.

Fosun Tourism [HK] Pos 83.40% CSC [SHA: A-Share] [HK] Bond Neu 63.20% Guo Guangchang Neu 91.00% Hotels, Resorts and Cruise Lines

Operation of Hotels and Resorts

Tourist Services

Tourism Concept

Taicang

Sanya

Intelligent Abstract

⌚ 2019-11-13 13:04:00 source:新浪网

Add to Favorites

Entity Extraction & Sentimental Scoring:
Company/Person

Concept Tag

Location Tag

Industrial Tag

Product Tag (Powered by ChinaScope SAM Database)

1 Basic Information

```
{
  'newsInfo': {
    'newsId': 6261912,
    'newsTitle': '内外交困，业绩下滑，分众传媒能否走出泥淖',
    'newsTs': '2019-08-01T00:03:36+0800',
    'newsUrl':
      'https://new.qq.com/omn/20190731/20190731A0UR3P00.html',
    'newsSource': '腾讯网',
    'newsSummary': '',
    'whitelistFlag': 0,
    'newsExtId':
      'c41596c69a4fdb3dc2b6f563393ef365'
  }
}
```

● Time of Crawling

Used to construct factor

2 Tags : Entity\Product\Industry\Event\Concept

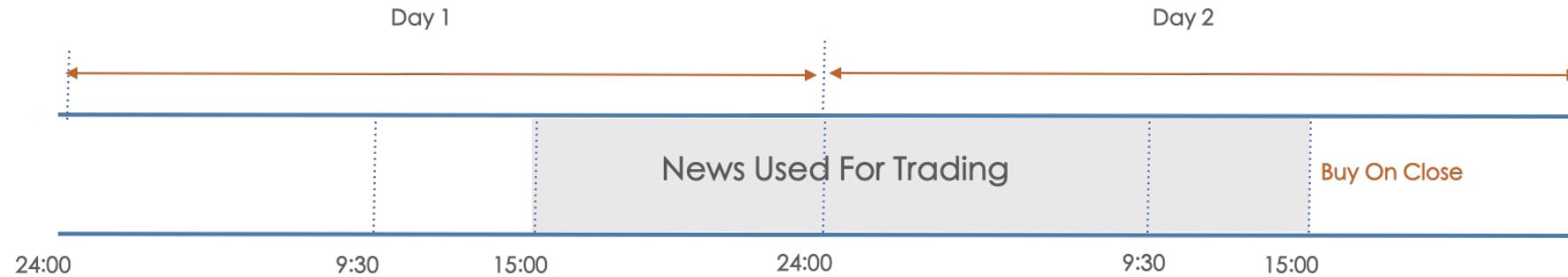
```
{
  'newsTags': [
    {'itemType': 'Product',
      'itemName_cn': '电梯媒介代理',
      'itemName': 'Elevator Advertising Agent Services',
      'itemId': 'ME001011020200201',
      'itemAlgoVersion': 'v0.4.0.9',
      'ItemRelevance': 0.07},
    {'itemType': 'Company',
      'itemName_cn': '分众传媒',
      'itemName': 'Focus Media',
      'itemId': 'CSF0000001061',
      'itemAlgoVersion': 'v0.4.0.17',
      'ItemExtId': '002027_SZ_EQ',
      'ItemRelevance': 0.92},
    {'itemType': 'Event',
      'itemName_cn': '业绩下降',
      'itemName': 'Performance Decline',
      'itemId': 'CD008019',
      'itemAlgoVersion':
        'v0.4.0.32'},
    {'itemType': 'Industry',
      'itemName_cn': '广告',
      'itemName': 'Advertising',
      'itemId': 'ME001',
      'itemAlgoVersion': 'v0.4.0.9',
      'ItemRelevance': 0.17}
  ]
}
```

3 Sentimental: For News&Entity

```
{
  'emotionInfos': [
    {'emotionEntity': 'News',
      'entityRefId': '6261912',
      'emotionIndicator': 2,
      'emotionWeight': 0.94,
      'emotionDetail':
        '{2=0.9390000000000001, 1=5.0E-4, 0=0.0605}'},
    {'emotionEntity': 'Company',
      'entityRefId': 'CSF0000001061',
      'entityName_cn': '分众传媒',
      'entityName': 'Focus Media',
      'entityCode': '002027 SZ EQ',
      'emotionIndicator': 2,
      'emotionWeight': 0.74,
      'emotionDetail':
        '{0=0.1854, 1=0.0763, 2=0.7383}'},
    {'emotionEntity': 'Industry',
      'entityRefId': 'ME001',
      'entityName_cn': '广告',
      'entityName': 'Advertising',
      'entityCode': '002027 SZ EQ',
      'emotionIndicator': 2,
      'emotionWeight': 0.74,
      'emotionDetail':
        '{0=0.1854, 1=0.0763, 2=0.7383}'}
  ]
}
```

- 0: Neutral
- 1: Positive
- 2: Negative

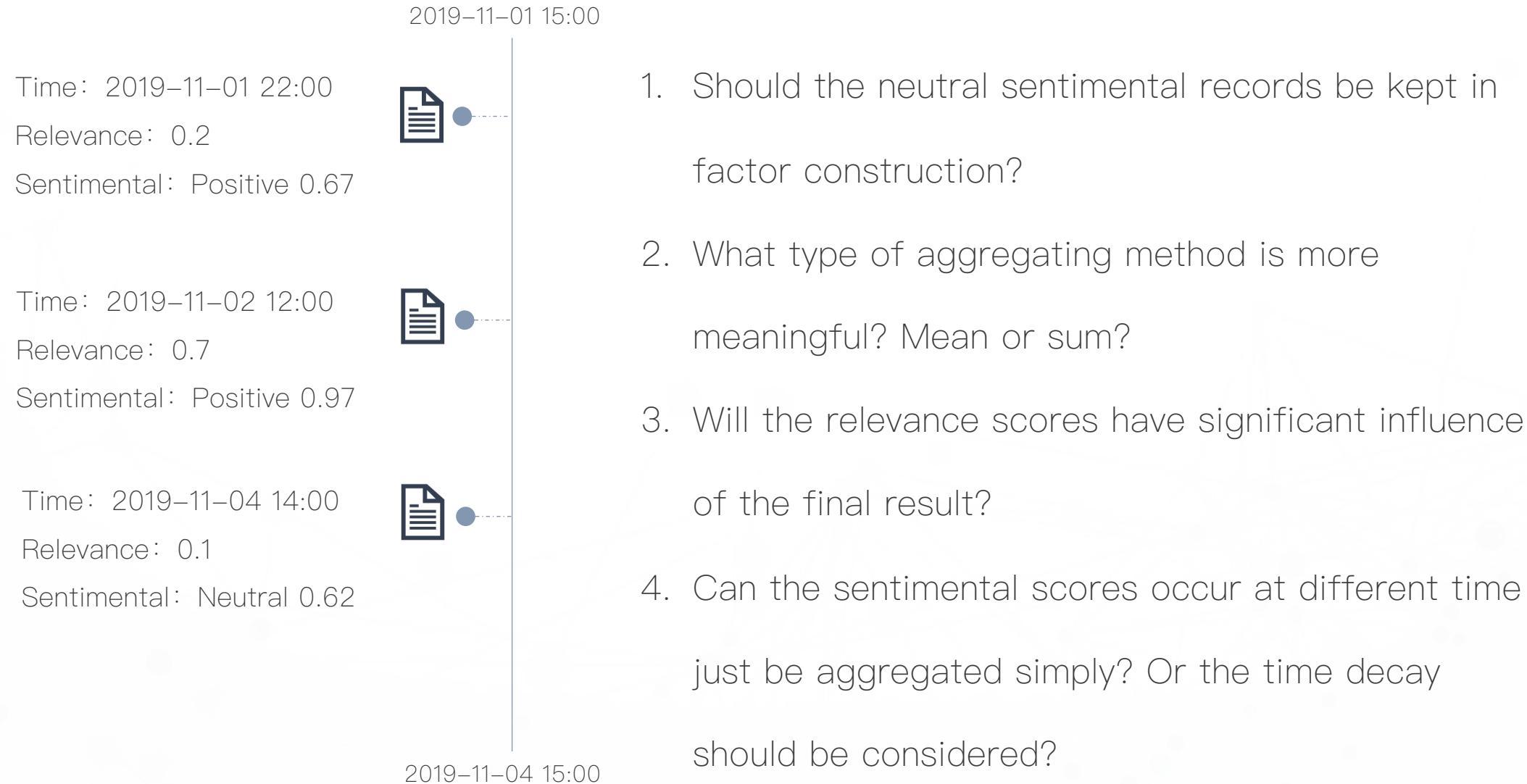
Sentimental Data To Sentimental Factor



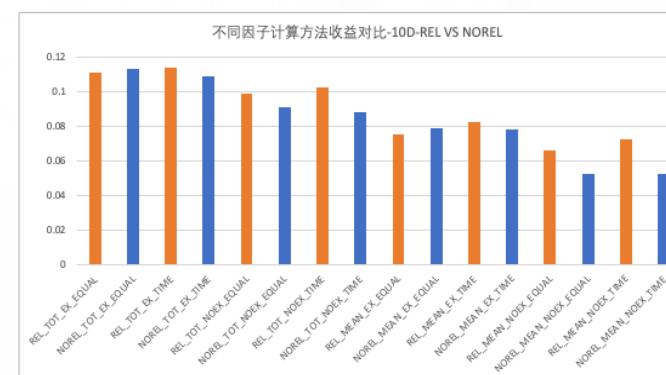
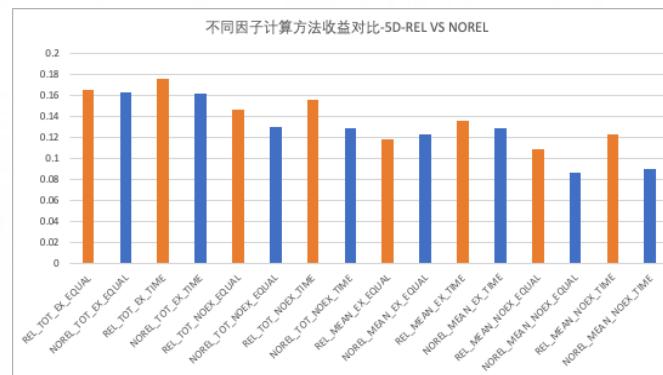
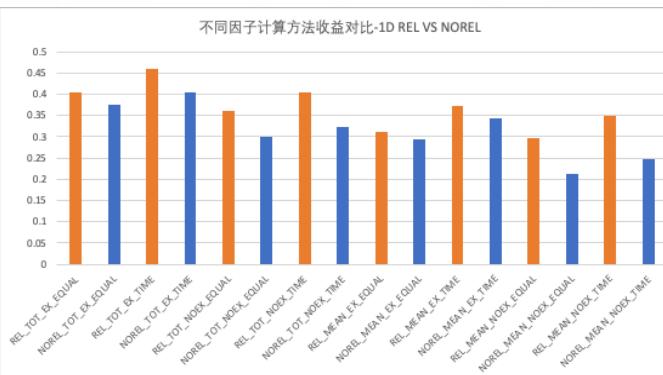
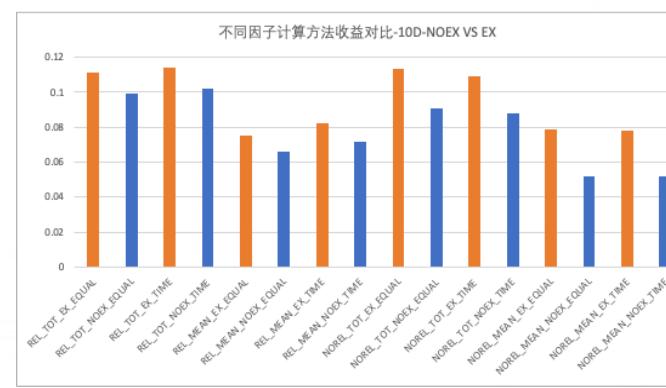
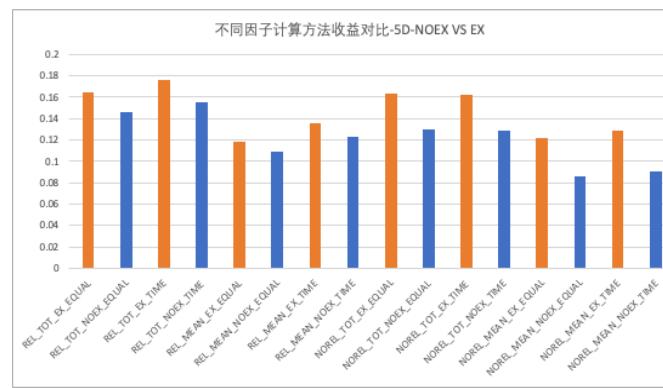
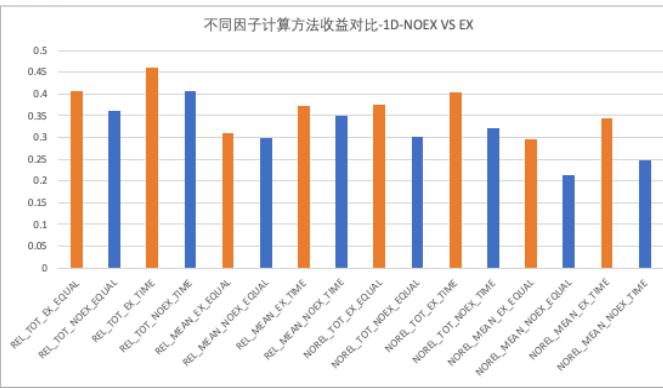
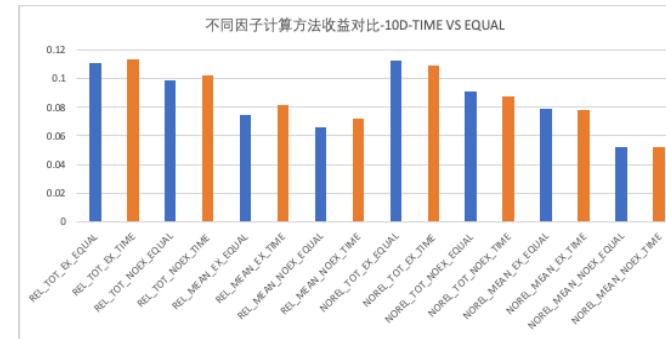
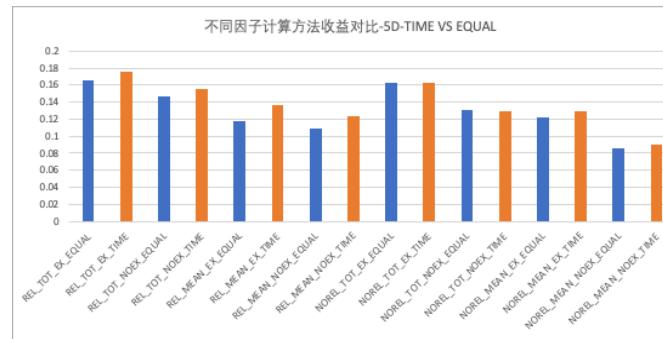
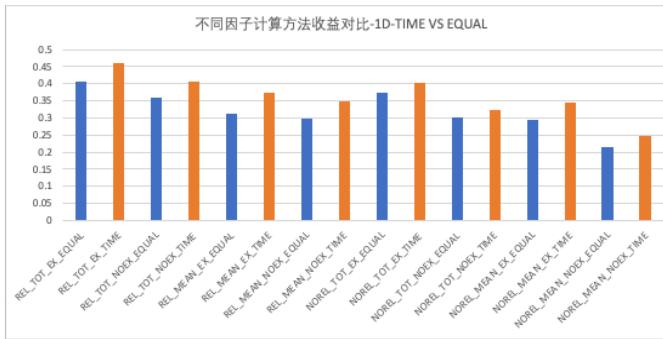
- Daily Sentimental Factor:

Aggregate the sentimental data from 15:00 of the last trading day to 15:00 of the current trading day.

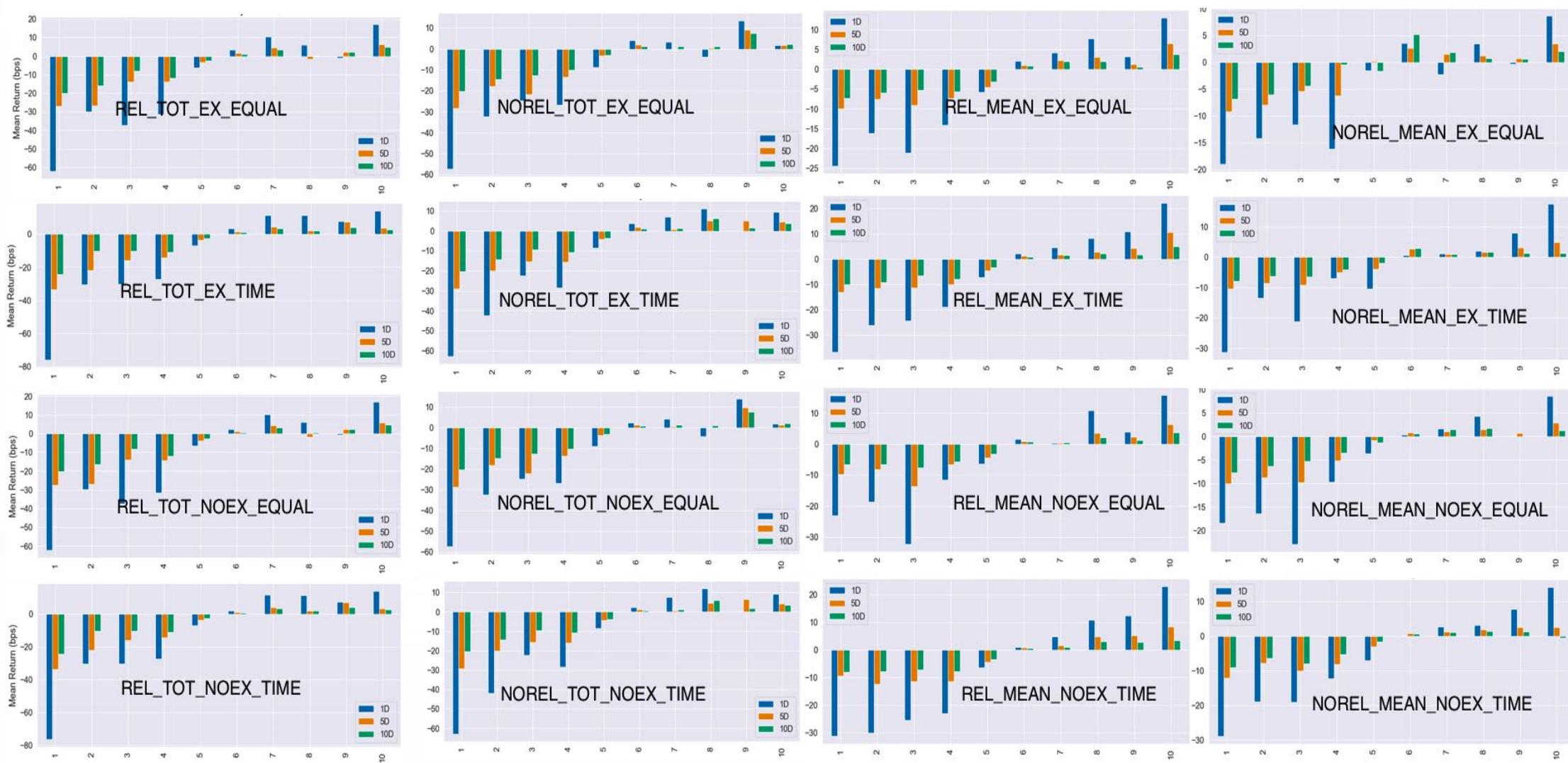
- So the problem is How to aggregate this data. We using the following method.



Sentimental Factor Test



Sentimental Factor Test



Conclusion

- The neutral sentimental records should be filtered out.
- The relevance score should be included in the factor calculation.
- The time decay of sentimental plays a import role in the factor performance.
- The total-score factor is good at screening negative sentimental stocks. The mean-score factor is good at screening positive sentimental stocks.

Sentimental Data To Sentimental Factor

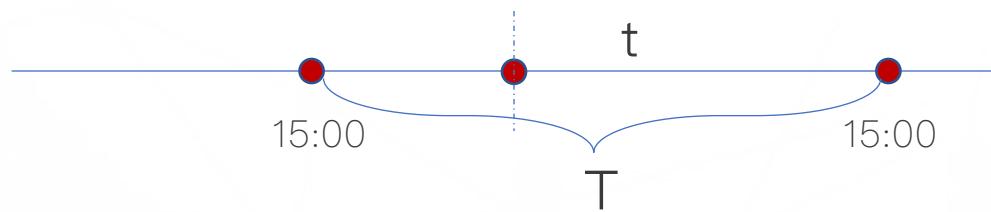
1 Raw sentimental score:

| | |
|-------------|--------------|
| 0: Neutral | 0: Neutral |
| 1: Positive | 1: Positive |
| 2: Negative | -1: Negative |

The score of stock i in news j is:

$$score_j^i = emotionIndicator_j^i * emotionWeight_j^i * itemRelevance_j^i * 100$$

2 Considering time decay:

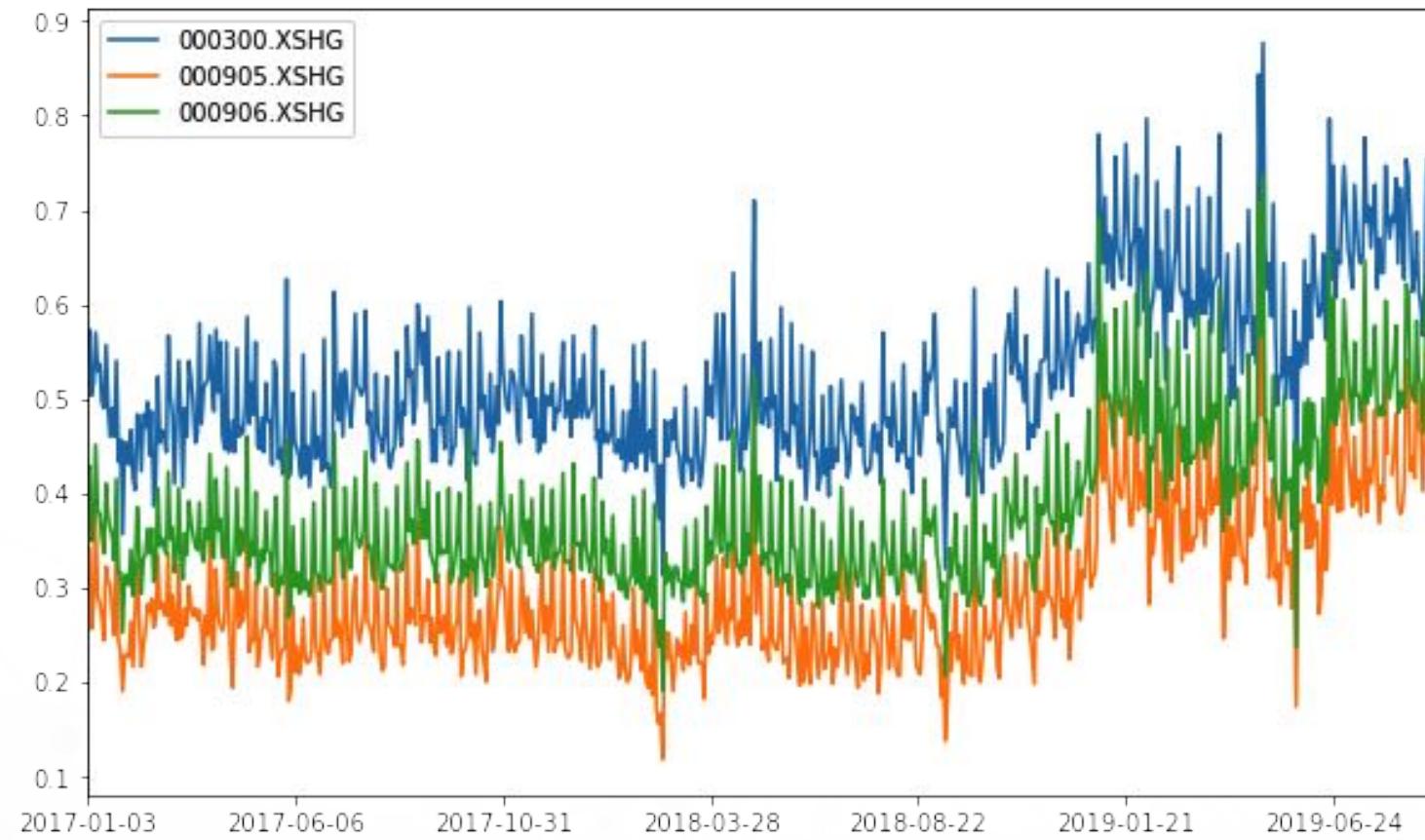


$$Decay_j = 1 / \exp(\frac{t}{T})$$

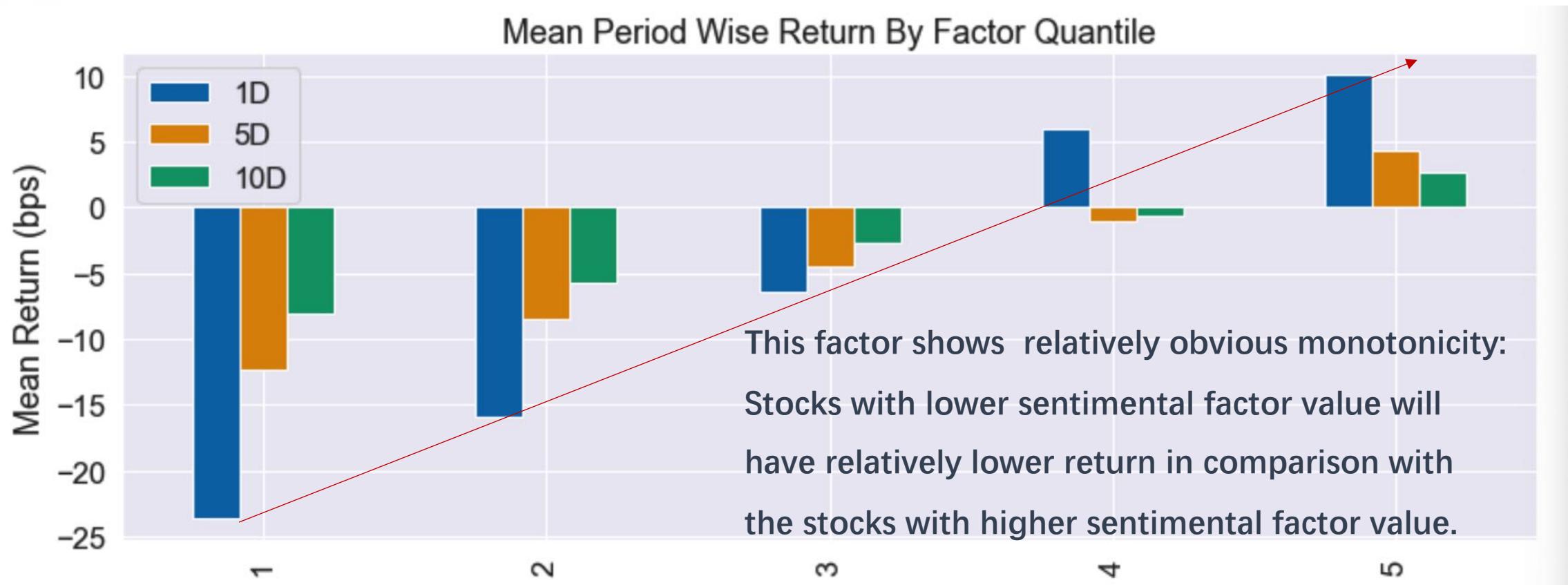
3 Aggregating sentimental score:

$$Factor_i = \left(\sum_{j=1}^N score_j^i * Decay_j \right) / N \quad \text{if emontionIndictor of news j is not 0}$$

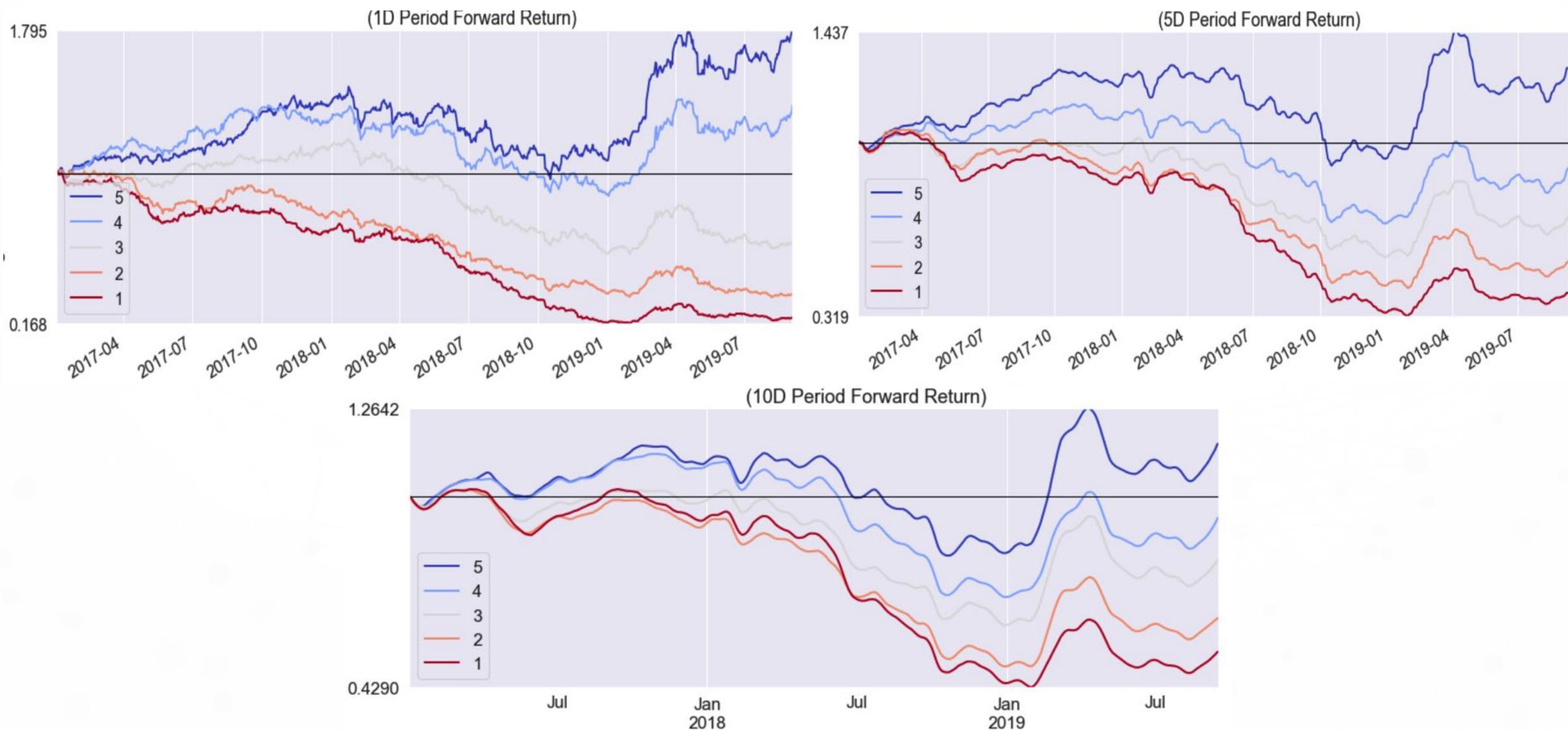
Stock Coverage of Main Index



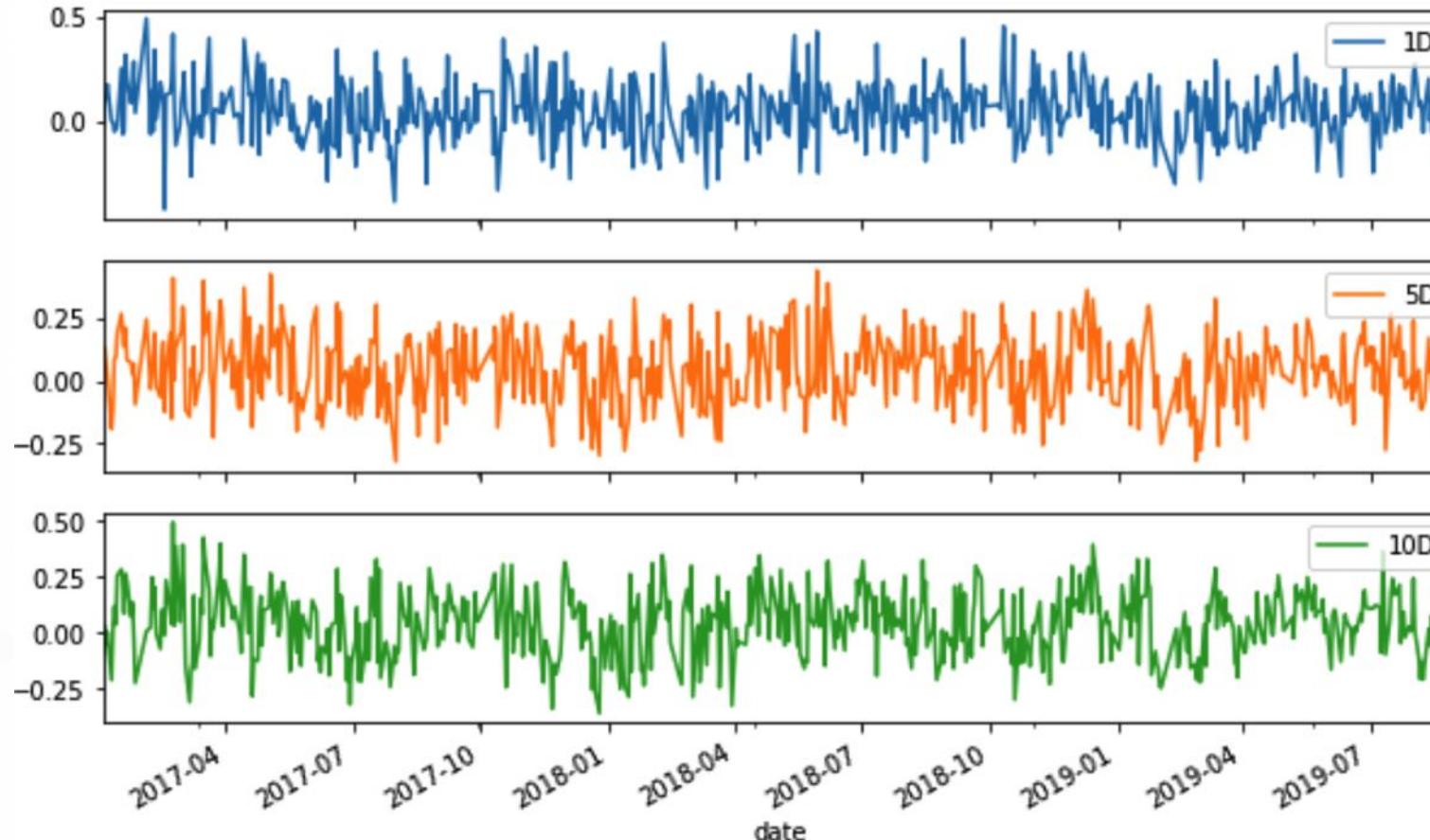
Sentimental Factor Testing within the components of CSI800 index.



Factor Testing – Return Analysis – Cumulative Return for Each Holding Period

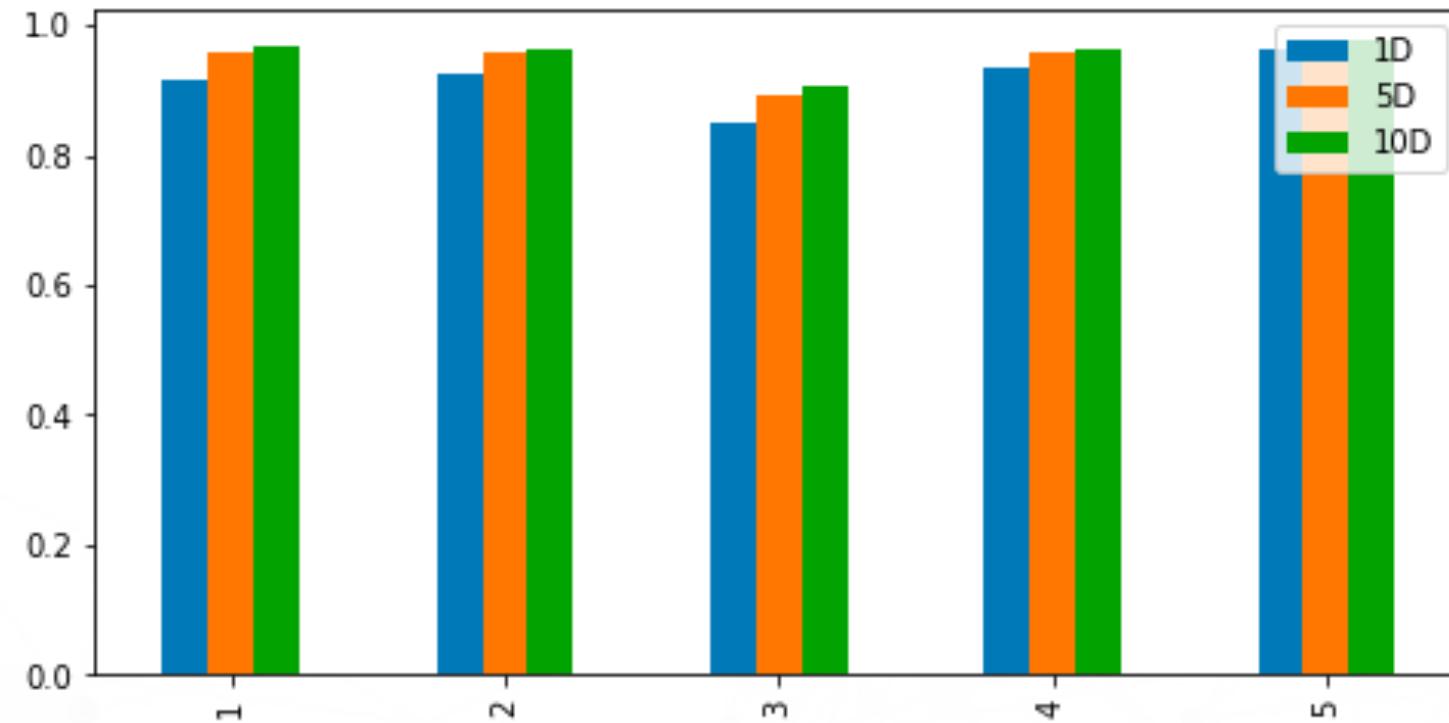


Factor Testing – IC Analysis



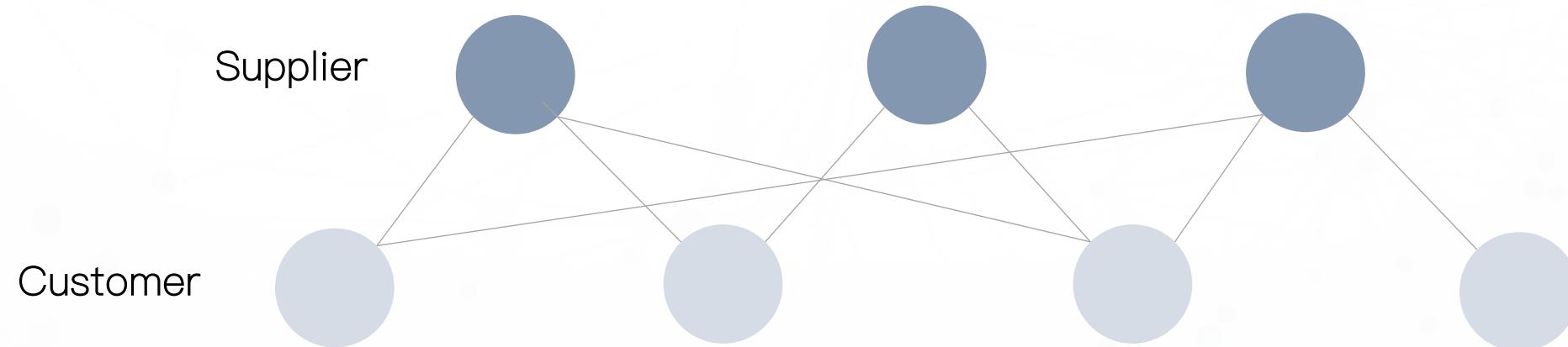
| Period | IC Mean | IC Std |
|--------|---------|--------|
| 1D | 4.31% | 13.7% |
| 5D | 4.40% | 13.9% |
| 10D | 3.86% | 14.5% |

Factor Testing – Turnover Analysis



Definition:

- Assume that the stock price of a listed company will be affected by its downstream companies, which are their customers, to some extent.
- We defined this factor as customer–momentum:
 - which means the customer's stock price fluctuation will transfer and affect the supplier's stock price



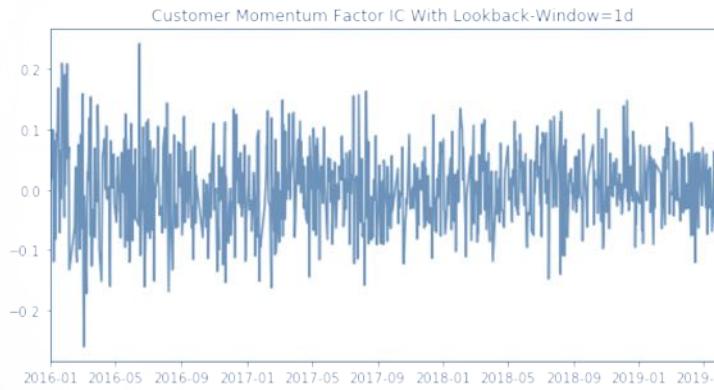
The Steps:

- Only keep the data where the customer is listed on China–A, HK, US market;
- Match the report–date to the report publish date, and set the data–available date as the next calendar day after the financial report publish date;
- At the end of each holding period, get the newest customer list for each subject company, if the customer was released more than two years ago, do not include it on its customer list.
- Calculate the mean return of each subject’s customers for the former holding period, this return is the customer–momentum factor;
- Conduct factor analysis for the customer–momentum factor.

$$\text{AggCustMom}_i(t', t) \triangleq \sum_{j \in \mathbb{C}_i} w_{ij} \text{Mom}_j(t', t), \quad \text{Mom}_j(t', t) \triangleq \frac{P_j(t)}{P_j(t')} - 1, \quad \forall j \in \mathbb{C}_i.$$

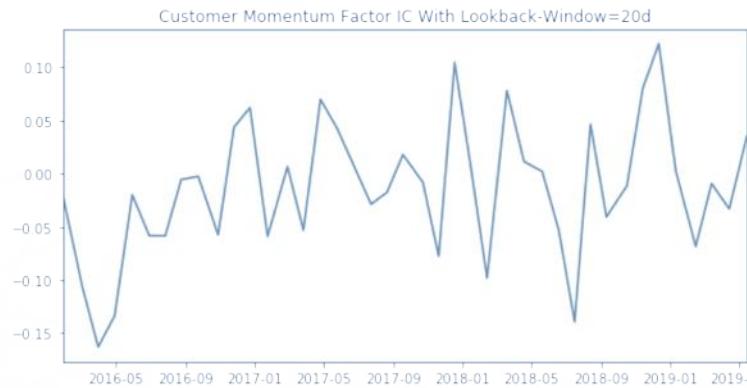
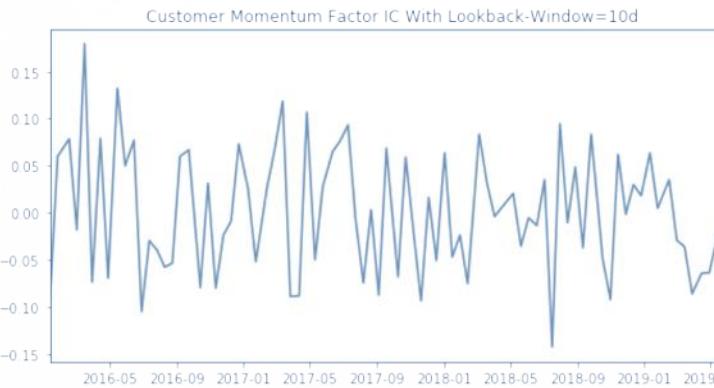
Supply–Chain Network Data: Customer Momentum

IC Series of Different Holding–Period



IC Abs Value Mean

| 1d | 5d | 10d | 20d |
|-------|-------|-------|-------|
| 0.051 | 0.052 | 0.054 | 0.050 |

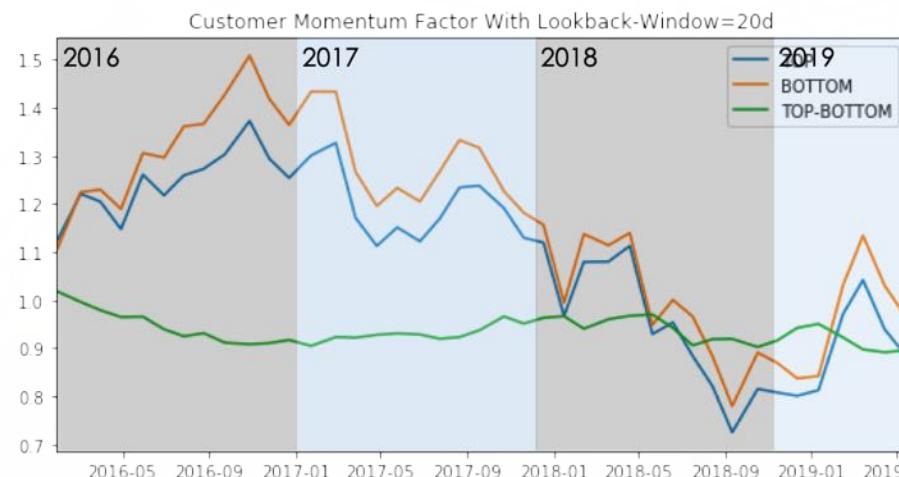
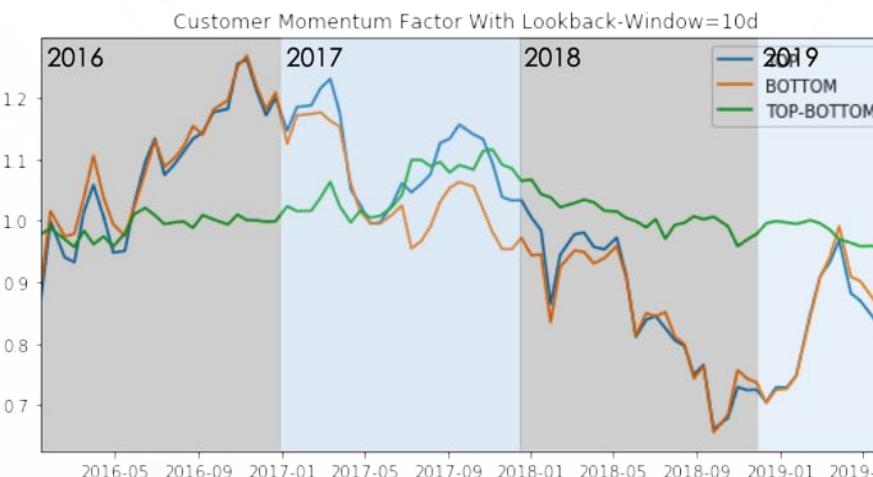
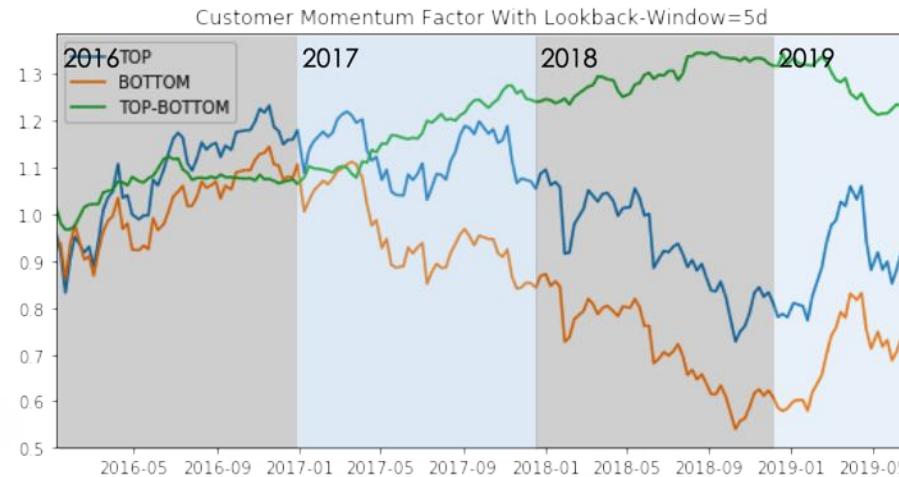
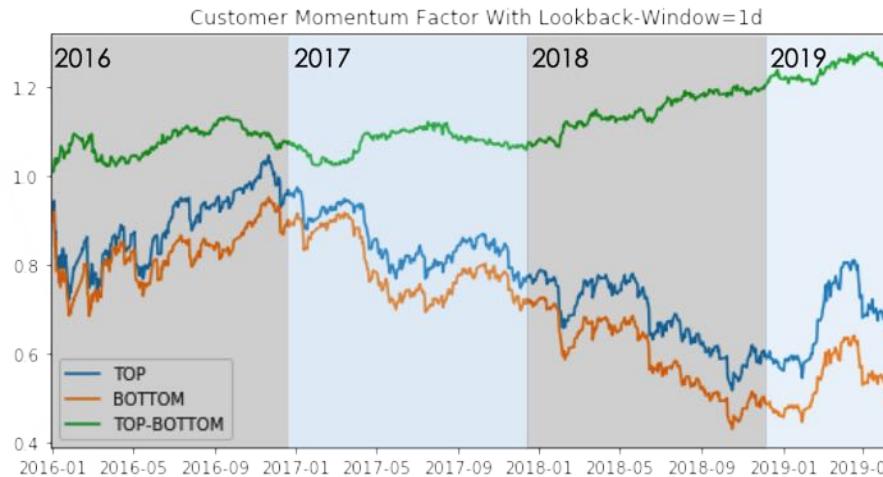


IC-IR

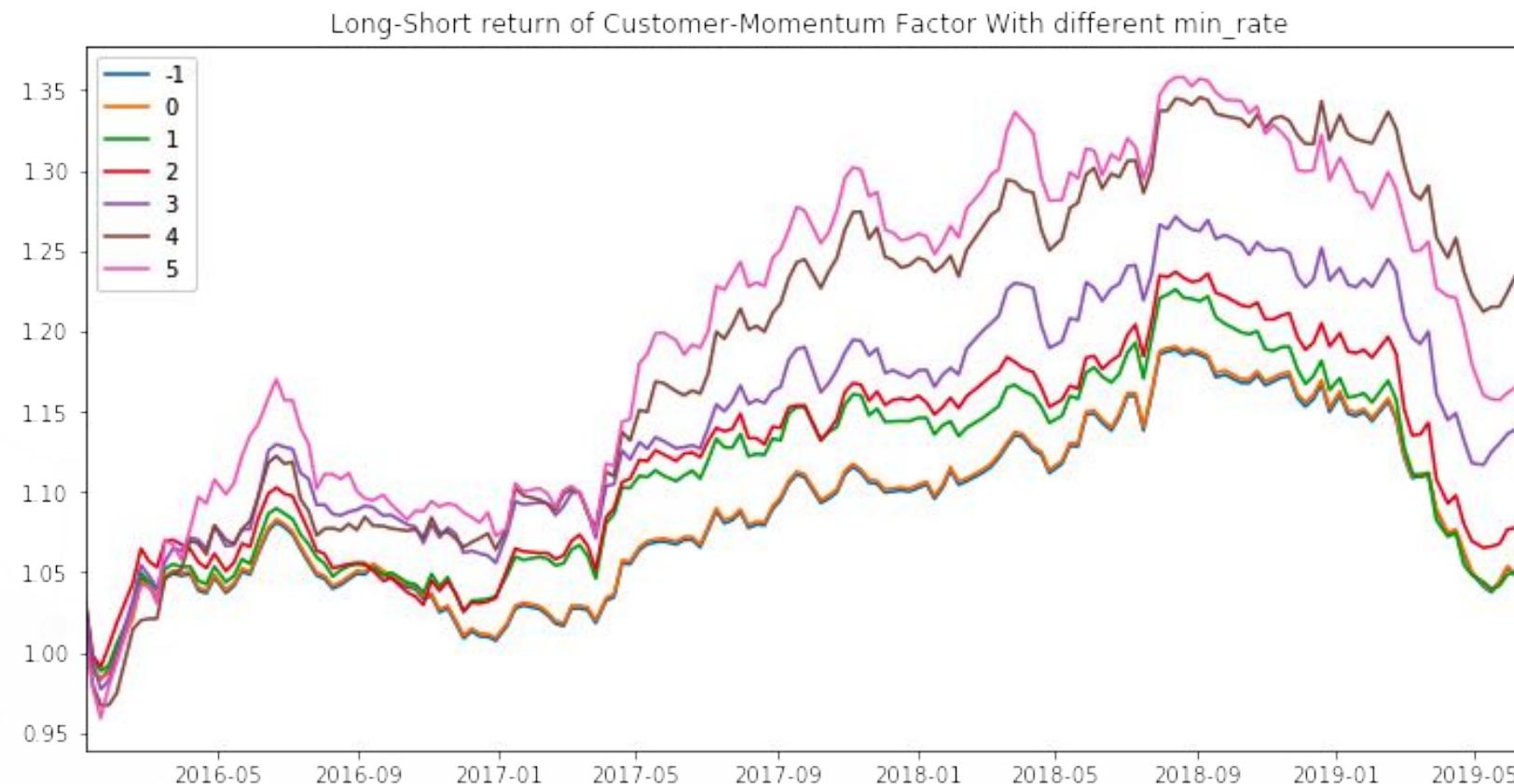
| 1d | 5d | 10d | 20d |
|-------|-------|-------|--------|
| 0.017 | 0.132 | 0.010 | -0.223 |

Supply–Chain Network Data: Customer Momentum

, long–short performance of different holding–period



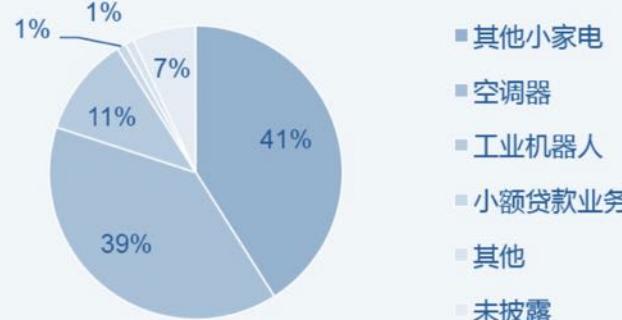
Performance is highly correlated with the rate threshold.



- The listed companies will disclose their main products at their financial report.
- We processed 100 thousands unstandardized products into 4000 standardized products.

示例

美的集团产品收入占比 (2017年度)



| 原始披露 | 数库标准 | 收入占比 |
|-----------|--------|------|
| 消费电器 | 其他小家电 | 41% |
| 暖通空调 | 空调器 | 39% |
| 机器人及自动化系统 | 工业机器人 | 11% |
| 利息收入 | 小额贷款业务 | 1% |

Industrial Chain Network Data

产品名称:

标准产品分类:

- 能源设备与服务
- 石油、天然气与消费用燃料
- 石油与天然气
- 煤与消费用燃料
- 煤炭的采选
- 低质煤
- 煤泥
- 原煤
- 无烟煤
- 褐煤
- 烟煤
- 焦煤
- 筛选煤
- 粉煤
- 洗精煤
- 其他煤炭产品
- 煤化工
- 其他煤化工产品
- 煤焦化
- 焦炭
- 冶金焦
- 焦炉煤气
- 焦油
- 煤液化
- 液化燃料油
- 煤气化
- 水煤浆
- 生物柴油
- 化学制品
- 建筑材料

选择科目 期间 货币单位 货币币种 市场

显示所占比例

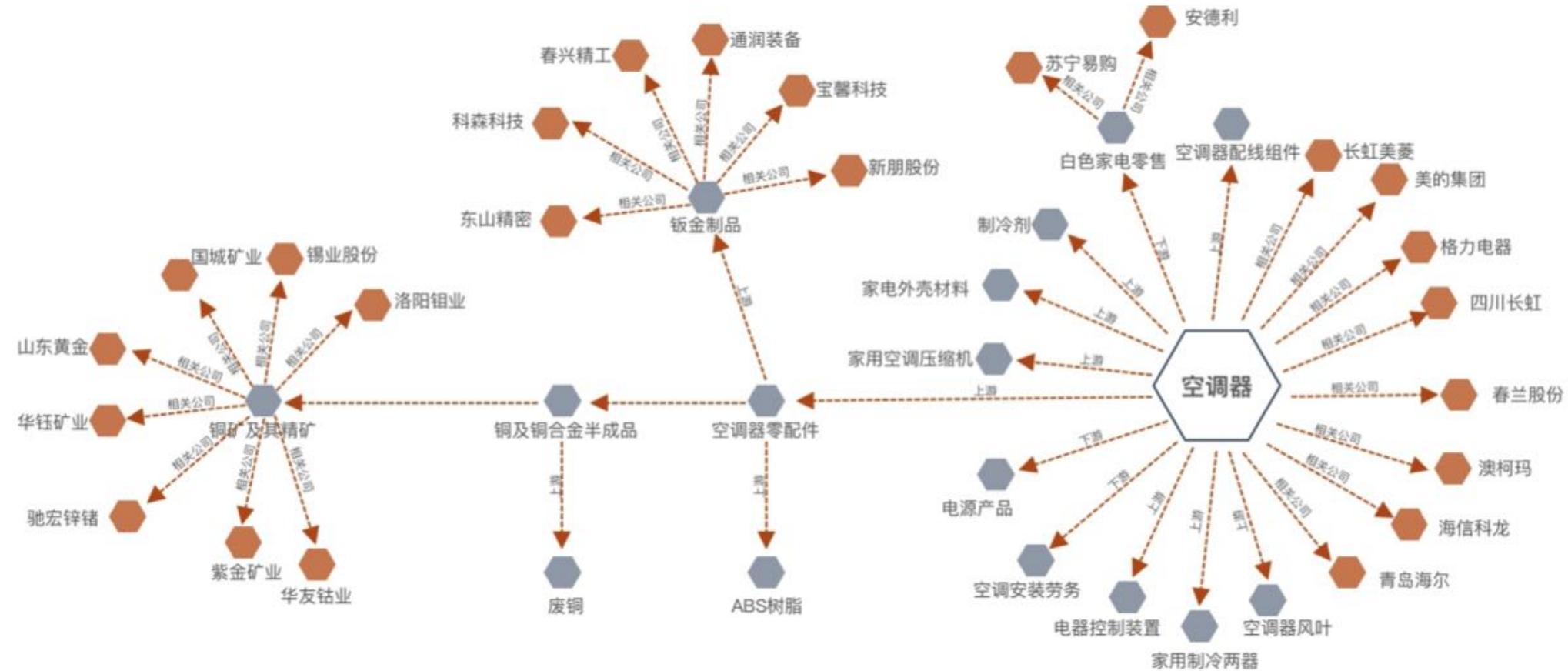
收入 石油、天然气与消费用燃料 > 煤炭的采选 > 原煤 2013 >>

| 证券简称 | 证券代码 | 2018A | 占比 | 2017A | 占比 | 2016A |
|---------------|--------------|-------------|------|-----------|------|-----------|
| 首钢资源 | 00639_HK_EQ | + 323,757 | 100% | 289,176 | 100% | 162,093 |
| 飞尚无烟煤 | 01738_HK_EQ | + 123,415 | 100% | 102,295 | 100% | 55,786 |
| 南方能源 | 01573_HK_EQ | + 64,119 | 100% | 64,243 | 100% | 69,099 |
| MONGOL MINING | 00975_HK_EQ | + 403,864 | 99% | 303,962 | 98% | 82,863 |
| 伊泰B股 | 900948_SH_EQ | + 3,502,511 | 89% | 3,434,300 | 93% | 2,061,858 |
| 陕西煤业 | 601225_SH_EQ | + 4,563,489 | 80% | 4,036,388 | 79% | 2,791,202 |
| 露天煤业 | 002128_SZ_EQ | + 654,902 | 80% | 613,354 | 81% | 428,791 |
| 中国油气控股 | 00702_HK_EQ | + 29,115 | 77% | 34,785 | 84% | 24,726 |
| 平庄能源 | 000780_SZ_EQ | + 167,841 | 75% | 229,686 | 82% | 166,122 |
| 兰花科创 | 600123_SH_EQ | + 536,527 | 63% | 450,866 | 60% | 264,382 |
| 红阳能源 | 600758_SH_EQ | + 363,461 | 50% | 390,315 | 51% | 312,110 |
| 恒鼎实业 | 01393_HK_EQ | + 46,977 | 44% | 34,728 | 43% | 21,716 |
| *ST大洲 | 000571_SZ_EQ | + 66,281 | 42% | 57,937 | 36% | 34,162 |
| *ST凯迪 | 000939_SZ_EQ | + 72,979 | 30% | 80,000 | 15% | 58,464 |
| 永泰能源 | 600157_SH_EQ | + 435,653 | 20% | 483,664 | 22% | 375,270 |
| 中煤能源 | 601898_SH_EQ | + 885,200 | 9% | 778,600 | 10% | -- |
| 云天化 | 600096_SH_EQ | + 90,247 | 2% | 86,998 | 2% | 40,067 |

35 家相关公司



示例



Barra CNE5 Model: 1 Country Factor, 10 Style Factors, 20+ Industrial Factors

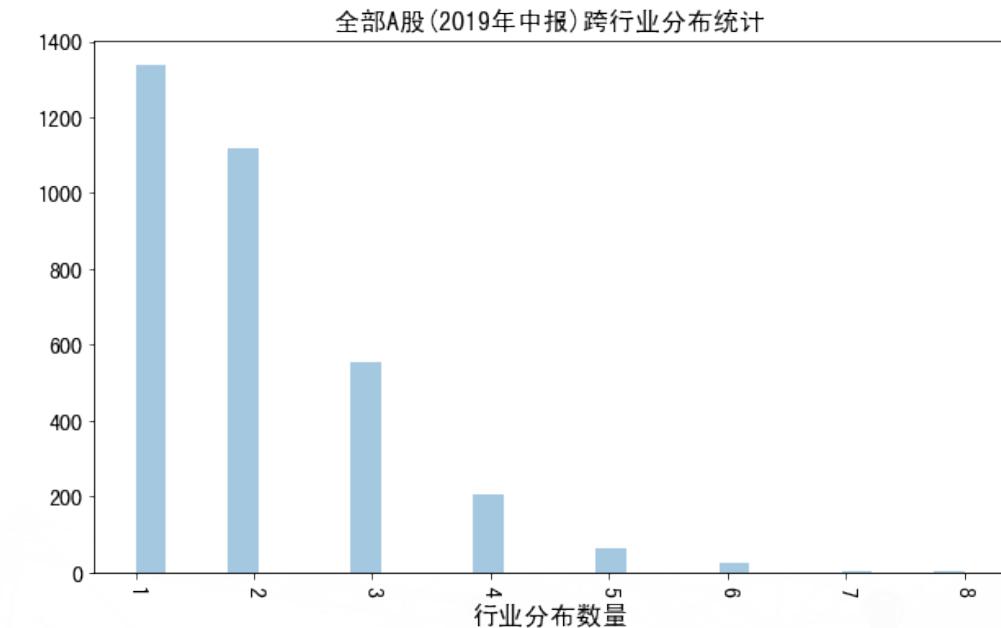
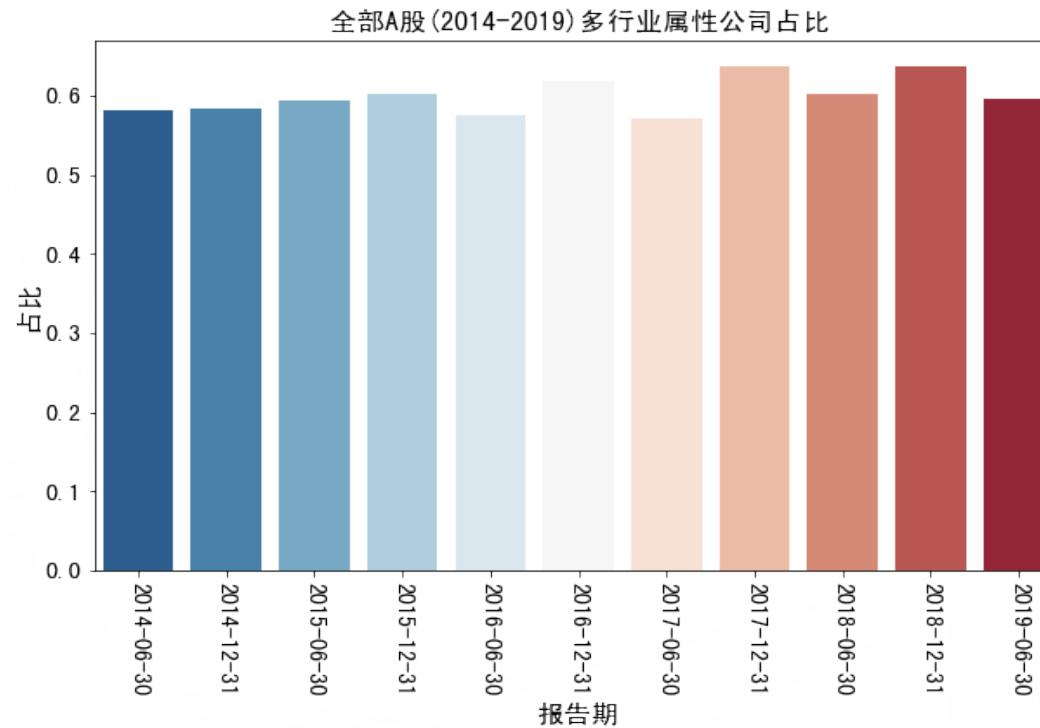
$$\begin{bmatrix} r_1 - r_f \\ r_2 - r_f \\ \vdots \\ r_N - r_f \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} f_C + \begin{bmatrix} X_1^{I_1} \\ X_2^{I_1} \\ \vdots \\ X_N^{I_1} \end{bmatrix} f_{I_1} + \cdots + \begin{bmatrix} X_1^{I_P} \\ X_2^{I_P} \\ \vdots \\ X_N^{I_P} \end{bmatrix} f_{I_P} + \begin{bmatrix} X_1^{S_1} \\ X_2^{S_1} \\ \vdots \\ X_N^{S_1} \end{bmatrix} f_{S_1} + \cdots + \begin{bmatrix} X_1^{S_Q} \\ X_2^{S_Q} \\ \vdots \\ X_N^{S_Q} \end{bmatrix} f_{S_Q} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}$$

- Using traditional industry category data, the company can only belong to one industry.
- But, actually, one company can have several industrial exposure in the same time.

For example, Shanshan 600884.SH

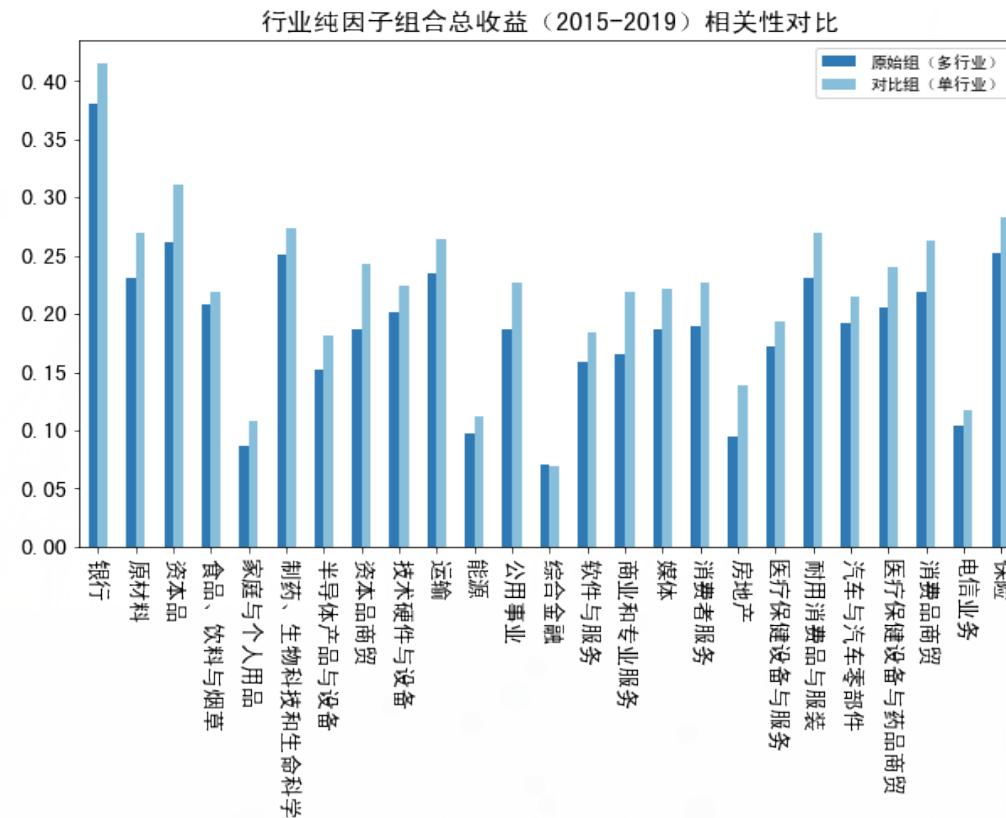
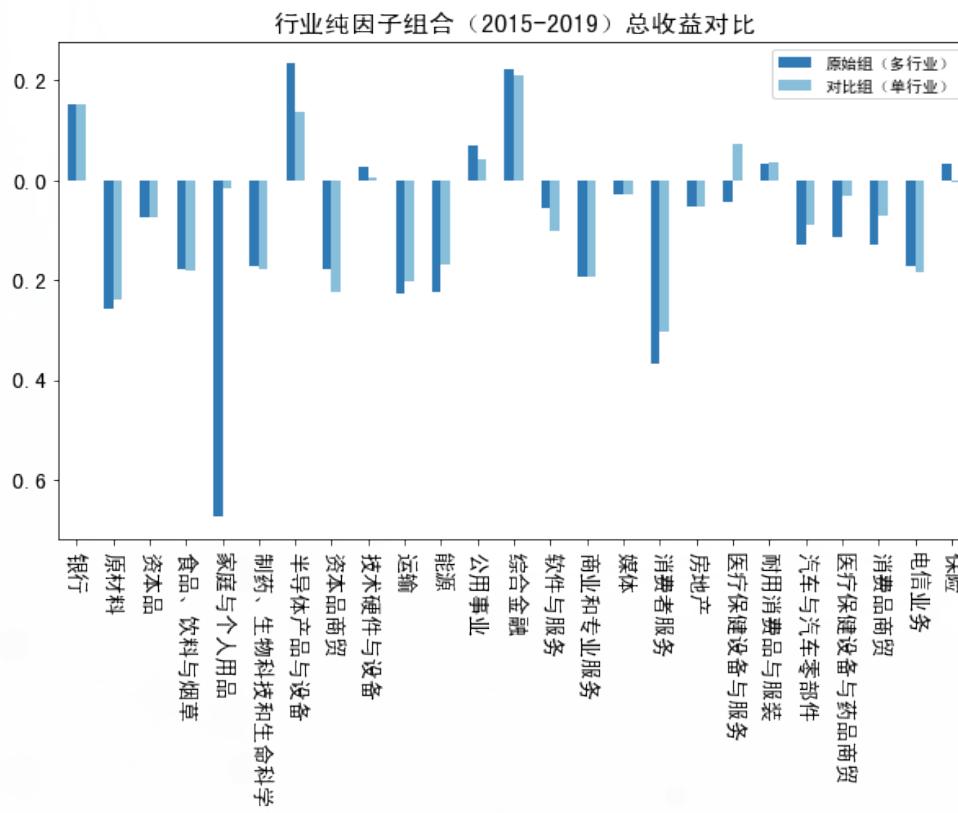
| 产品(SAM一级) | 收入占(%) | 产品对应二级行业 | 行业属性占比(%) |
|-----------|--------|----------|-----------|
| 商业服务 | 6.61 | 商业与专业服务 | 6.80 |
| 服饰、鞋类与奢侈品 | 9.16 | 耐用消费品与服装 | 9.42 |
| 汽车 | 1.40 | 汽车与汽车零部件 | 1.44 |
| 综合金融服务 | 1.73 | 综合金融 | 1.77 |
| 电子制造 | 78.35 | 技术硬件与设备 | 80.57 |

Industrial Chain Network Data —— Multi-Industrial Exposure Factor in Risk Model



Industrial Chain Network Data —— Multi-Industrial Exposure Factor in Risk Model

- Original Group: A company can have multiple industrial exposure.
- Contrast Group: A company can only belong to one industry which it has most exposure.



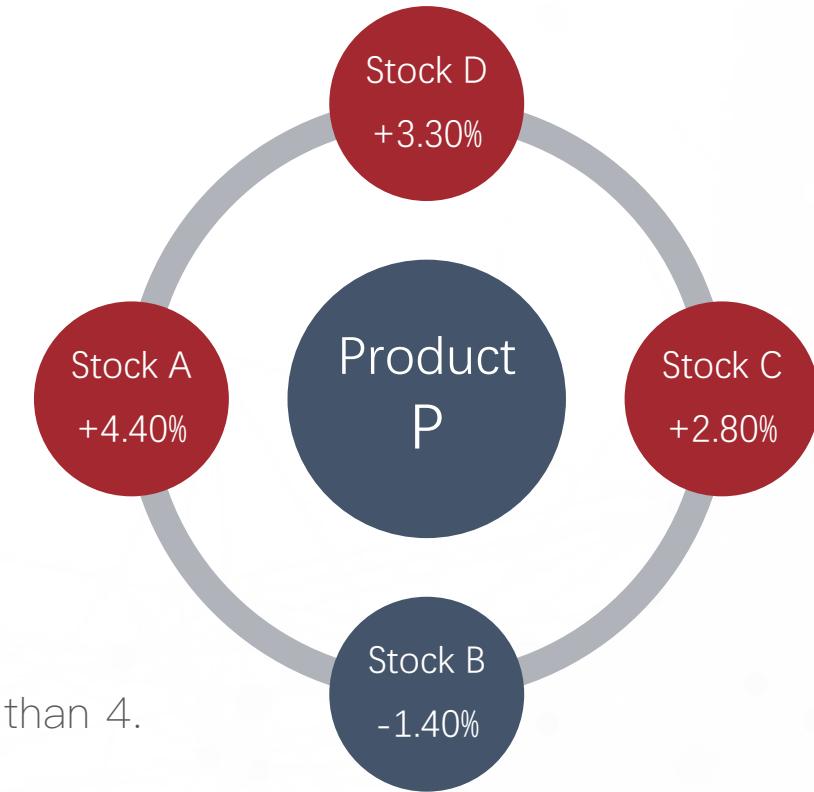
Industrial Chain Co-movement Strategy

❖ Strategy's Logic

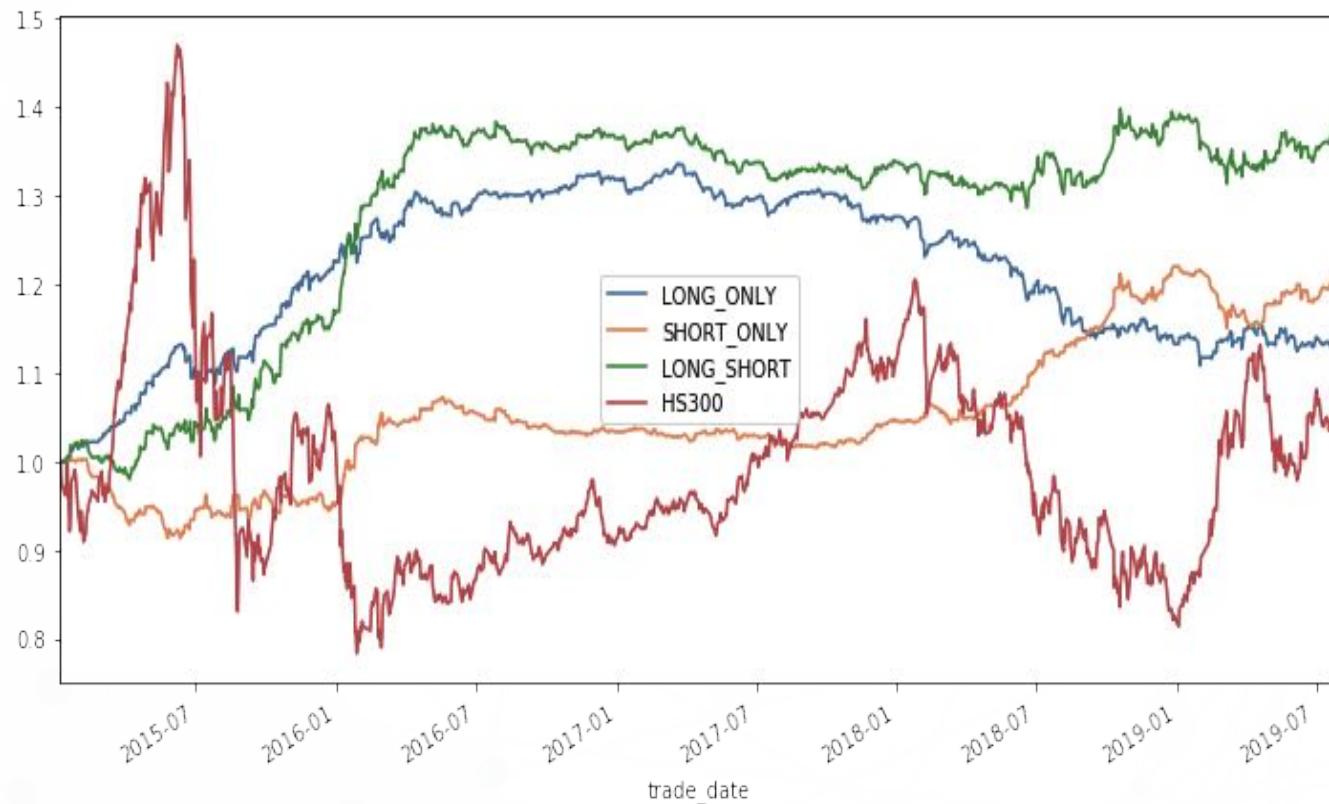
- The share price of those companies whose core business are similar will be influenced by some common factors.
- When most companies' stock prices move in the same direction, the left companies will catch up in few days.

⚙ Parameters Setting

- Filter the product node under which the number of the companies is greater than 4.
- When $\frac{3}{4}$ of the companies' price moves in the direction, buy or short the stocks which move in the opposite direction and hold these stocks for one day.

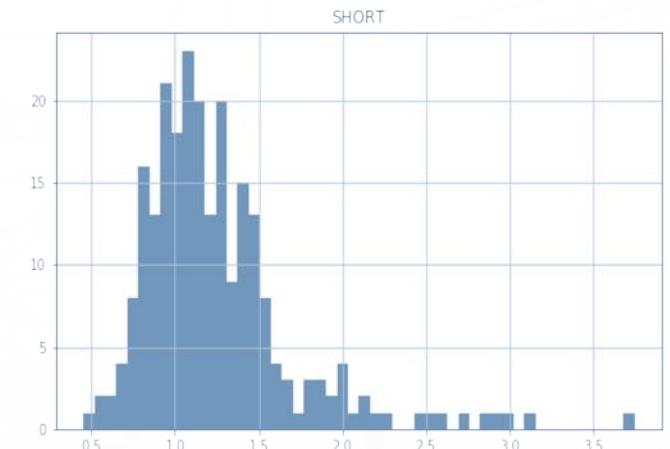
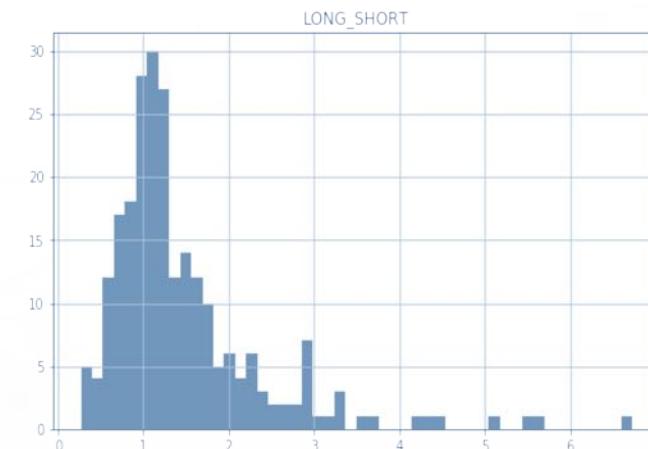
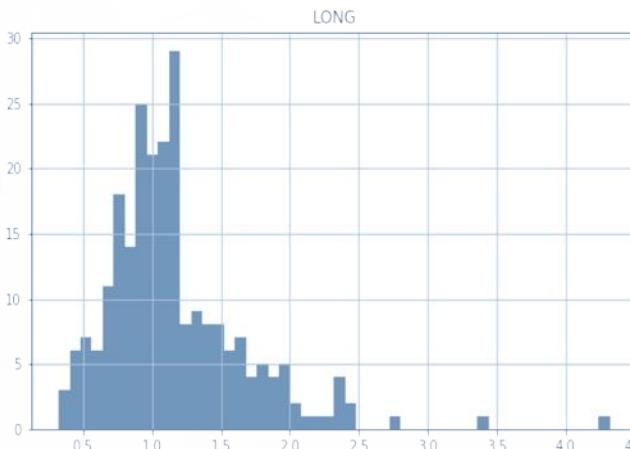


Industrial Chain Co-movement Strategy



| | LONG | SHORT | LONG_SHORT |
|---------------------|---------|--------|------------|
| Annual return | 3.00% | 4.30% | 7.50% |
| Cumulative returns | 13.80% | 20.00% | 37.00% |
| Annual volatility | 5.60% | 5.50% | 6.70% |
| Sharpe ratio | 0.56 | 0.78 | 1.11 |
| Calmar ratio | 0.18 | 0.46 | 1.07 |
| Stability | 0.01 | 0.78 | 0.51 |
| Max drawdown | -17.10% | -9.20% | -7.00% |
| Omega ratio | 1.13 | 1.21 | 1.24 |
| Sortino ratio | 0.79 | 1.32 | 1.71 |
| Skew | -0.34 | 1.47 | 0.29 |
| Kurtosis | 5.89 | 12.01 | 4.38 |
| Tail ratio | 1.22 | 1.18 | 1.27 |
| Daily value at risk | -0.70% | -0.70% | -0.80% |
| Alpha | 0.02 | 0.05 | 0.07 |
| Beta | 0.13 | -0.13 | -0.01 |
| Win Ratio | 53% | 48% | 52% |

- Most product nodes do not have positive return.
- The distribution is seriously positively skewed which means the product nodes which have good performance will always better than others.
- Tune the model with these changes: set a rolling window, pick the 10 best-performed product node in the former window, then only trade these product nodes in the following window.



Industrial Chain Co-movement Strategy

滚动选取前5个交易日，收益最好的产品



| | LONG | SHORT | LONG_SHORT | LONG_SHORT_1 |
|---------------------|---------|---------|------------|--------------|
| Annual return | 25.80% | 6.30% | 32.80% | 33.90% |
| Cumulative returns | 172.60% | 30.50% | 245.50% | 258.20% |
| Annual volatility | 11.90% | 9.40% | 14.60% | 14.10% |
| Sharpe ratio | 1.99 | 0.69 | 2.02 | 2.15 |
| Calmar ratio | 2.08 | 0.42 | 1.43 | 1.87 |
| Stability | 0.87 | 0.68 | 0.88 | 0.94 |
| Max drawdown | -12.40% | -14.90% | -23.00% | -18.20% |
| Omega ratio | 1.56 | 1.19 | 1.46 | 1.48 |
| Sortino ratio | 3.42 | 1.13 | 3.4 | 3.71 |
| Skew | 0.81 | 1.14 | 0.59 | 0.63 |
| Kurtosis | 4.93 | 10.25 | 3.73 | 2.57 |
| Tail ratio | 1.46 | 1.22 | 1.49 | 1.38 |
| Daily value at risk | -1.40% | -1.20% | -1.70% | -1.70% |
| Alpha | 0.23 | 0.08 | 0.3 | 0.3 |
| Beta | 0.15 | -0.17 | -0.04 | -0.02 |
| Win Ratio | 56% | 52% | 56% | 55% |

Industrial Chain Co-movement Strategy

| 2015-2019 胜率前 20 的产品节点 | | | | | |
|------------------------|---------|-------------|---------|-------------|---------|
| LONG | | SHORT | | LONG_SHORT | |
| 其他塑料包装薄膜 | 100.00% | 其他互联网广告 | 100.00% | 其他互联网广告 | 100.00% |
| 化肥与农用药剂 | 100.00% | 钻采设备及其零配件 | 83.33% | 电能管理系统及设备 | 88.89% |
| 电能管理系统及设备 | 100.00% | 电能管理系统及设备 | 80.00% | 应用系统集成服务 | 75.00% |
| 其他互联网广告 | 100.00% | 建筑材料贸易 | 80.00% | 电影院 | 75.00% |
| 风电机组零部件 | 87.50% | 铅矿石以及铅精矿 | 80.00% | 有线电视传输网络 | 69.70% |
| 太阳能发电 | 83.33% | 应用系统集成服务 | 78.57% | 触摸显示屏 | 68.42% |
| 兽用疫苗 | 81.25% | 资产管理 (非信托类) | 75.00% | 发电机组及相关设备 | 67.86% |
| 锂离子电池材料 | 80.77% | 通信设备 | 75.00% | 化肥与农用药剂 | 66.67% |
| 电梯零部件 | 80.00% | 工程咨询服务 | 75.00% | 建筑材料贸易 | 66.67% |
| 电影院 | 76.92% | 电影院 | 71.43% | 工程咨询服务 | 65.52% |
| 计算机整机 | 75.00% | 酱油 | 70.00% | 酱油 | 65.00% |
| 有线电视传输网络 | 75.00% | IC 卡 | 69.23% | 风电机组零部件 | 64.71% |
| 触摸显示屏 | 73.33% | 进出口贸易 | 69.23% | 涤纶长丝 | 64.13% |
| 航空航天与国防信息技术 | 72.73% | 发电机组及相关设备 | 68.18% | 网络优化服务 | 63.64% |
| 应用系统集成服务 | 70.00% | 摩托车制造 | 66.67% | 摩托车制造 | 63.16% |
| 电视机 | 69.23% | 原材料贸易 | 66.67% | 资产管理 (非信托类) | 62.50% |
| 保健食品 | 68.00% | 网络优化服务 | 66.67% | 锂离子电池材料 | 62.00% |
| 涤纶长丝 | 66.67% | 混凝土 | 65.79% | 医药中间体 | 60.75% |
| 发电机组及相关设备 | 66.67% | 触摸显示屏 | 65.22% | 系统集成服务 | 60.18% |
| 保健护理机构 | 66.67% | 酒店及度假村经营 | 65.12% | 包装机械 | 60.00% |

| 2015-2019 收益率前 20 的产品节点 | | | | | |
|-------------------------|---------|-------------|---------|-------------|---------|
| LONG | | SHORT | | LONG_SHORT | |
| 体外诊断试剂 | 219.68% | 酒店、度假村与豪华游轮 | 185.94% | 酒店、度假村与豪华游轮 | 402.90% |
| 金融行业应用软件 | 151.43% | 锂离子电池材料 | 160.08% | 经纪业务 | 216.30% |
| 其他信息科技服务 | 105.70% | 百货商店 | 153.45% | 商品房开发 | 202.53% |
| 有色金属 | 103.09% | 其他煤炭产品 | 127.52% | 装修工程 | 186.78% |
| 医药中间体 | 103.00% | 超市 | 121.55% | 金融行业应用软件 | 160.02% |
| 应用系统集成服务 | 91.33% | 装修工程 | 116.95% | 体外诊断试剂 | 159.41% |
| 其他中成药 | 90.14% | 经纪业务 | 98.10% | 有色金属 | 148.79% |
| 航空公司 | 89.41% | 专业市场经营 | 97.93% | 其他商品化工产品 | 146.11% |
| 稀土永磁材料 | 88.64% | 工业机械 | 90.72% | 铝加工制品 | 140.20% |
| 锂离子电池 | 80.97% | 商品房开发 | 89.68% | 安全软件 | 133.32% |
| 酒店、度假村与豪华游轮 | 75.88% | 铝加工制品 | 84.57% | 百货商店 | 131.13% |
| 安全软件 | 73.14% | 物业管理 | 74.74% | 其他中成药 | 122.51% |
| 其他商品化工产品 | 68.99% | 物流服务 | 74.02% | 医药中间体 | 113.20% |
| 电子化学产品 | 67.15% | 电动机及其零部件 | 70.99% | 证券经纪业务 | 111.59% |
| 经纪业务 | 59.67% | 幕墙工程 | 70.11% | 幕墙工程 | 110.19% |
| 商品房开发 | 59.50% | 女装 | 69.20% | 系统集成服务 | 107.91% |
| 表面贴装式 LED | 57.24% | 系统集成服务 | 69.17% | 锂离子电池材料 | 106.27% |
| 中成药 | 57.03% | 交通运输系统集成 | 66.59% | LED 照明灯具 | 105.44% |
| 绿化工程 | 54.48% | 污水处理工程服务 | 61.46% | 物业管理 | 101.33% |
| 酒店及度假村经营 | 52.95% | 其他水产加工制品 | 59.41% | 其他水产加工制品 | 99.05% |

Thanks!

Contact Us:

Tel: (86) 021-36359362

Email: business@chinascope.com