

Selecting representative constituents in CSI 300 Index

Author: Yikun Hu

23 March 2020

Selecting representative constituents

- Fully index tracking
 - Advantage
 - 100% mimic
 - No tracking error
 - Disadvantage
 - High trading cost
 - Large fund
 - High time cost
 - High opportunity cost
 - Not flexibility for a fund manager's preference
- Partial index tracking
 - Advantage
 - Low trading cost
 - Small fund
 - Flexibility for a fund manager's preference
 - Disadvantage
 - Not 100% mimic
 - Tracking error exists

Data acquisition

Constituents in CSI 300 Index from 1st of Jan 2010 to 1st of Jan 2020.

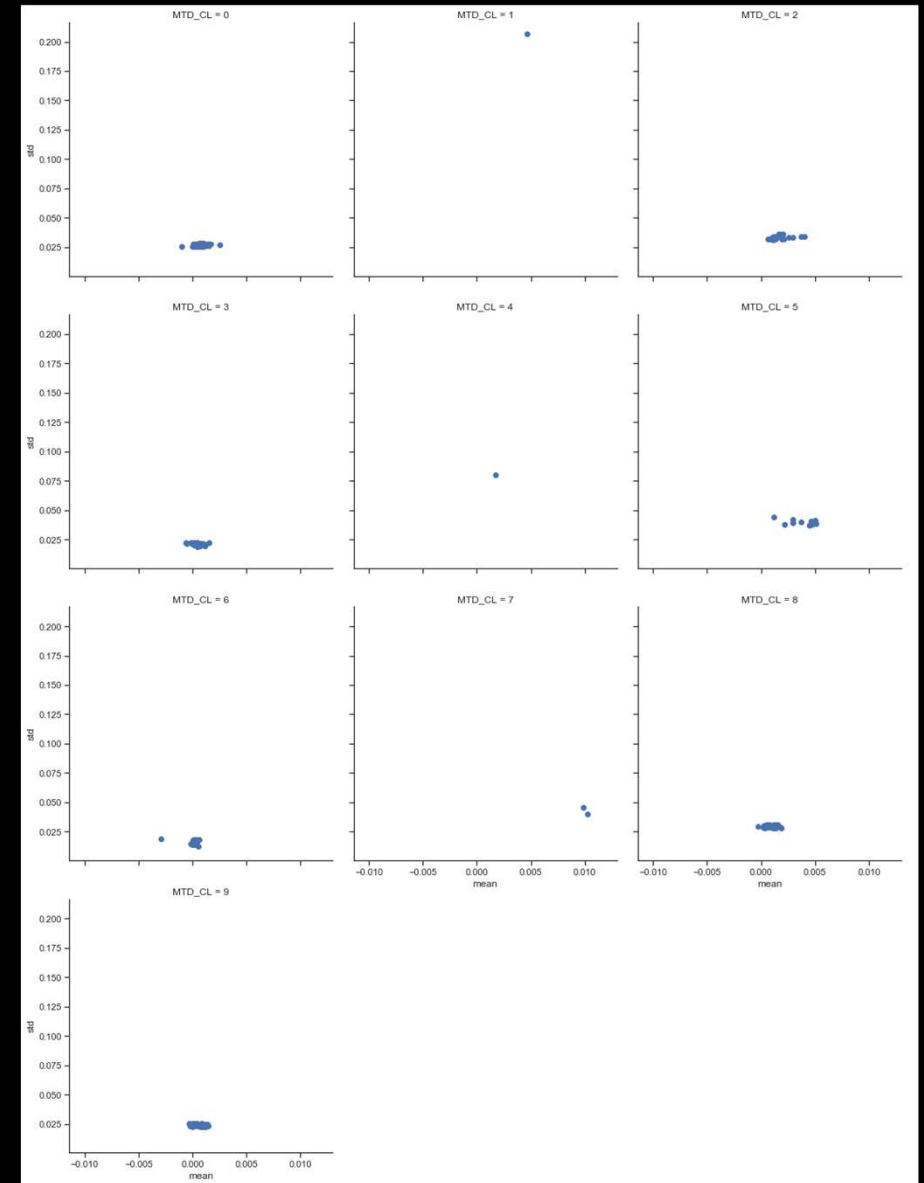
- Raw data
 - Market data
 - Closing price
 - 300 constituents
 - 2447 trading days
 - Financial ratios
 - ROE (return of equity), ROA (return of assets), Operating profit to net profit, Net profit to total revenue, Operating cash flow to revenue of
 - 292 constituents (because of lack of data on database)
- Cleaned data
 - Market data
 - Closing price
 - 300 constituents
 - 2447 trading days
 - Financial ratios
 - ROE (return of equity), ROA (return of assets), Operating profit to net profit, Net profit to total revenue, Operating cash flow to revenue of
 - 285 constituents (because of lack of data on database)

Exploratory Data Analysis

Market data

The number of clusters: 10

Mean-variance distributions

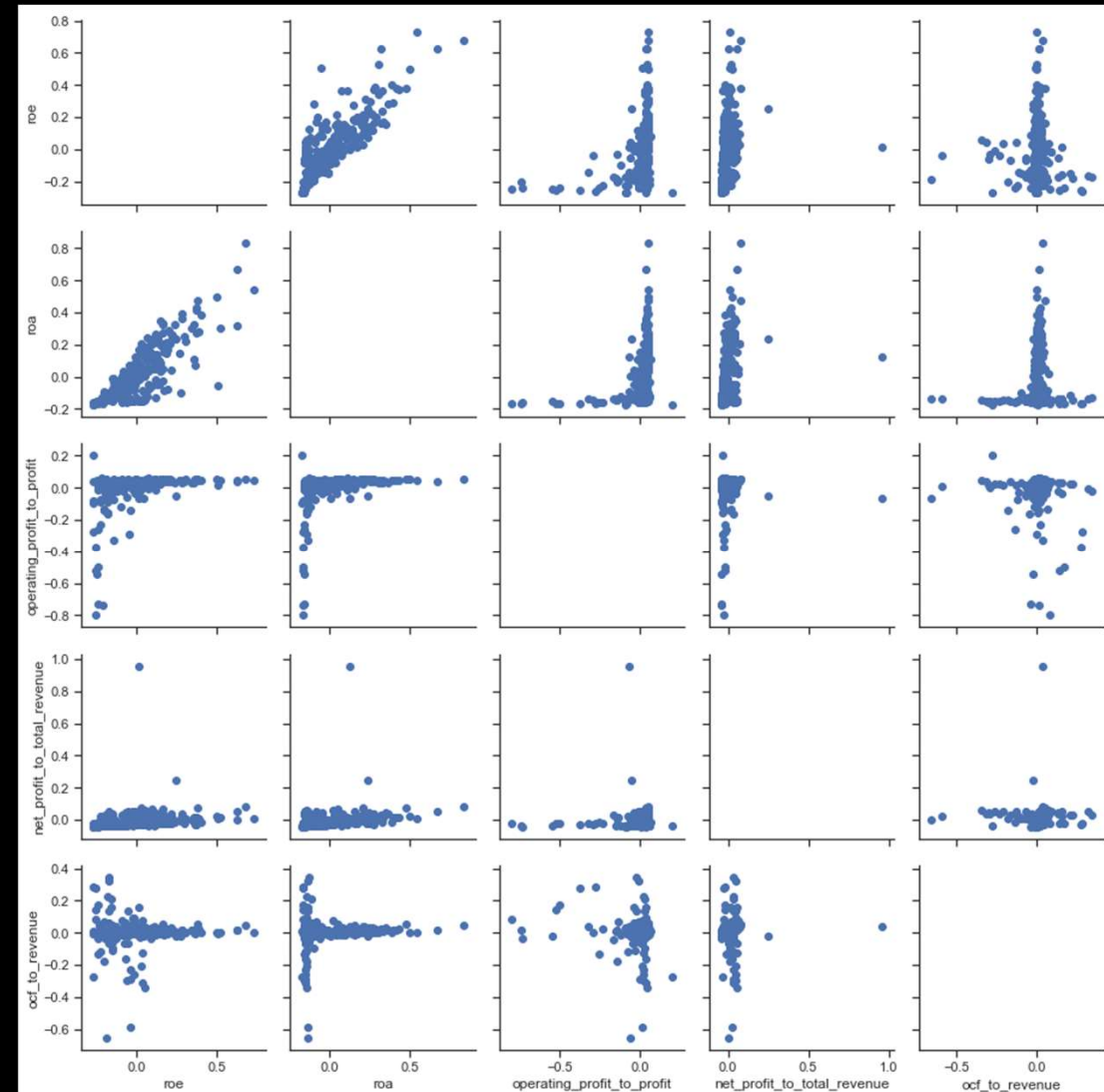


Exploratory Data Analysis

Financial data

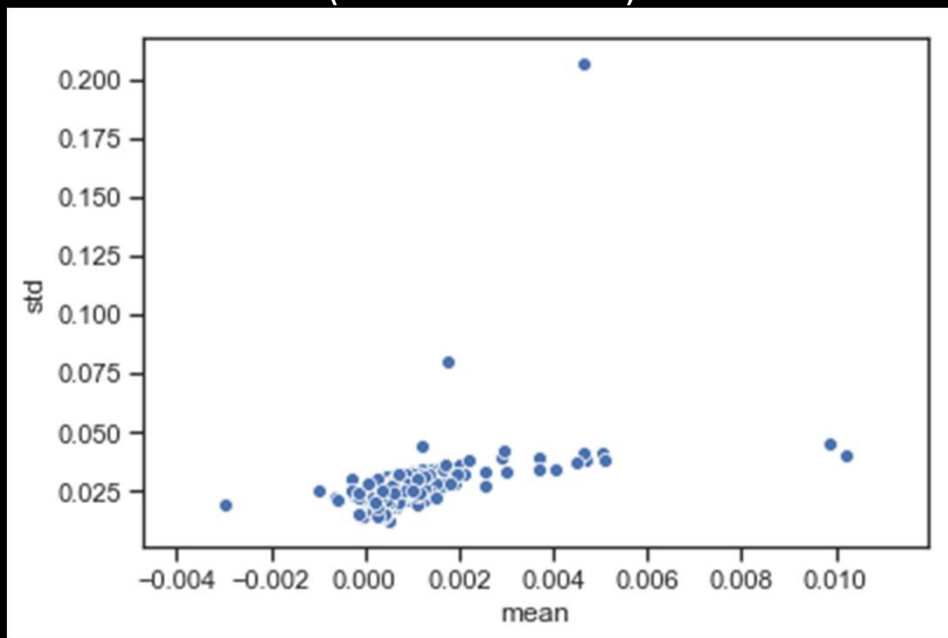
The number of clusters: 10

Pairwise relationship

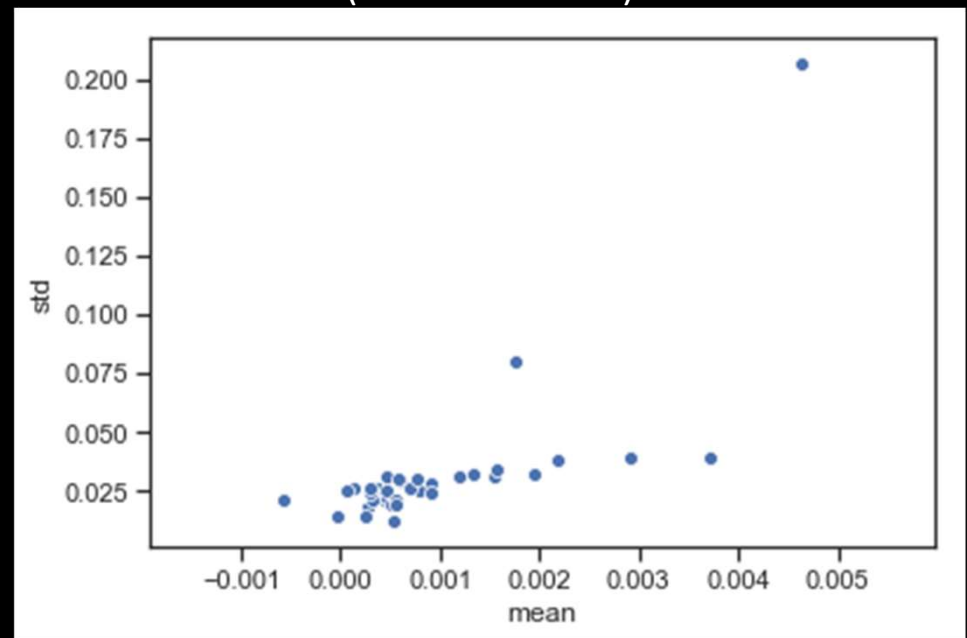


Exploratory Data Analysis

Mean-variance distribution of constituents in
CSI 300 Index
(300 constituents)



Mean-variance distribution of constituents in
the tracking index
(36 constituents)



Conclusion

- Reduced number of constituents for tracking an index
- Understand mean-variance distribution of constituents by clusters
- Understand relationship between financial ratios
- Find out representatives for tracking CSI 300 with specified financial characteristic.

Future directions

- Using larger size of constituents
- Add more market data for clustering
- Add more financial ratios for clustering
- Create a portfolio partially tracking CSI 300 Index with specified weights and financial ratio preference
- Create specified index consisting of constituents with similar financial characteristics