What can we do for hedge fund return data under the DandD Environment?

Daisuke Yokouchi, Ryozo Miura
Hitotsubashi University
Graduate School of
International Corporate Strategy

Outline

- Introduction to DandD Environment
- Design and implementation of a DandD client software for financial time series data through preliminary analysis of hedge fund return data.

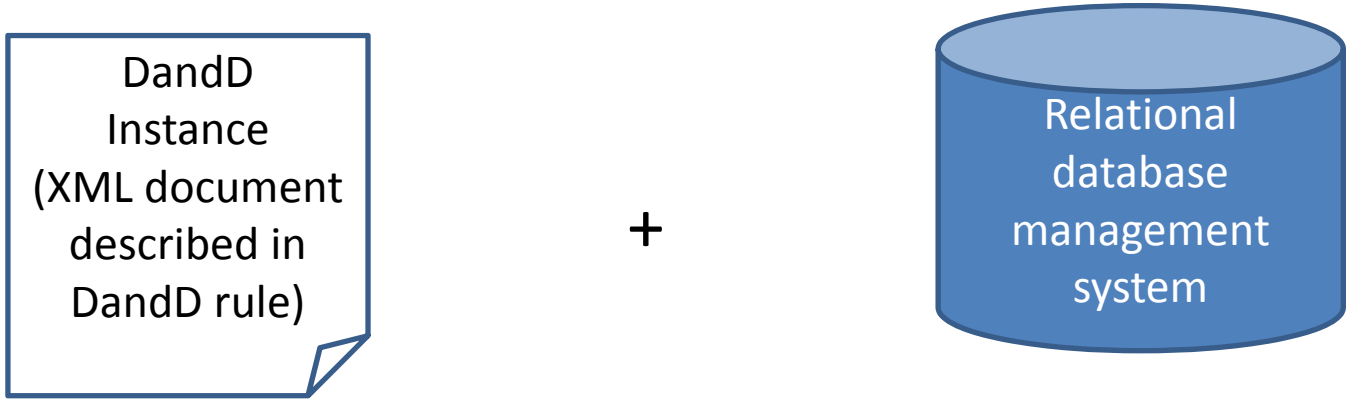
1. DandD Environment

Data and Description Project

- Development of an integrated environment for working with data
  - Establishment of a description rule for integration of data and enough background information of the data.
    - acceleration of collaboration between statisticians and researchers in different areas
  - Development software for supporting the work flow of data science
    - data acquisition, data storage, data cleaning, data analysis, model building

Project home page http://www.stat.math.keio.ac.jp/DandD/

Basic data set in DandD



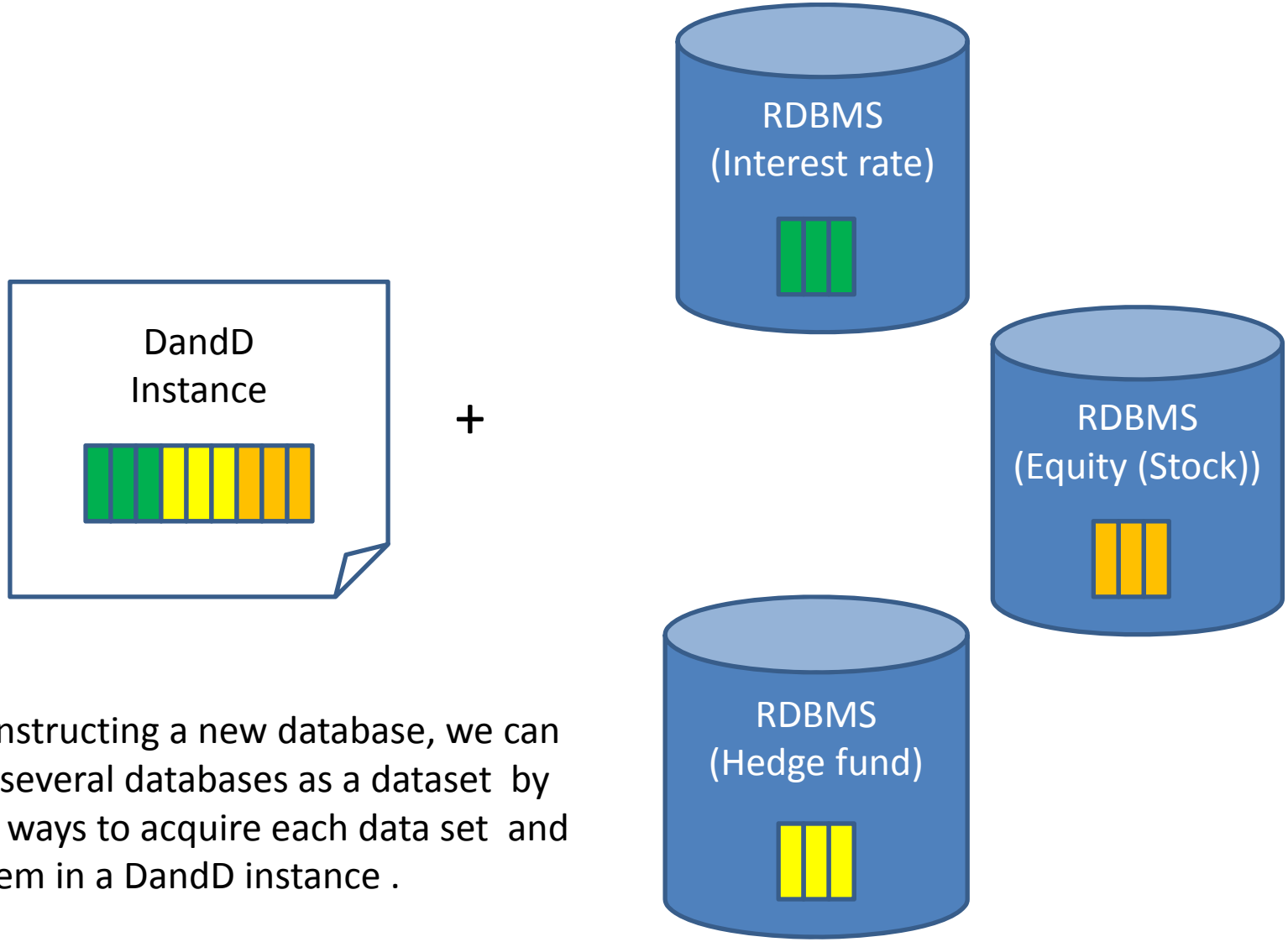
DandD instance
- background information
- variable information
- data structure information
- data acquisition information
- history

A view to data

RDBMS
- whole data

Entity of data

InterDatabase
(2001, Yokouchi, Shibata)

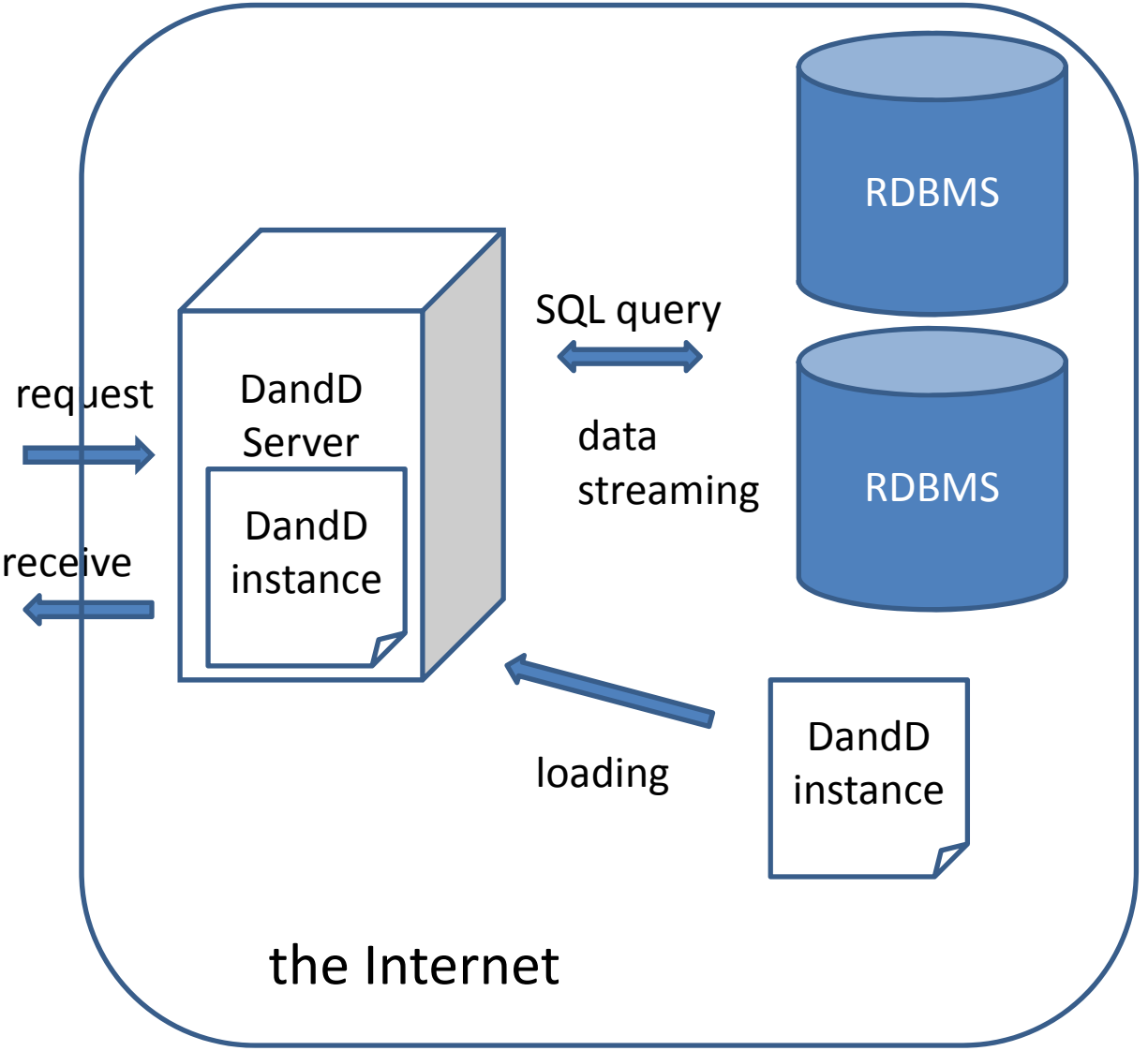


Without constructing a new database, we can use data in several databases as a dataset by embedding ways to acquire each data set and organize them in a DandD instance .

DandD client server system (Yokouchi, Shibata, 2004)

A DandD Client

ex. DandD Browser



2. Design of ICS FinAnalyzer through preliminary analysis of hedge fund return data

Barrier to data analysis

- Handling of a large amount of financial data is not so easy.
  - For example, we often fail to import gigabyte data into statistical software without processing by program.
  - In many cases, we can not operate such data without use of full fledged database management system.
  - Data cleaning is also difficult owing to size of the data

For beginners, it is even difficult to retrieve data of a fund, and browse as a chart.

It is necessary to develop software helping to handle and analyze a large amount of financial data stored in relational database management system.

ICS FinAnalyzer (ICSFA)

- Basic concepts
  - ICSFA
    - provides us with a simple interface for exploring data stored in relational database management systems.
    - provides us with enough background information of the data.
- Requirements
  - ICSFA
    - is a DandD client.
    - can send a DandD instance to a DandD Server and receives the data described in the instance.
    - can display the data as a parallel coordinate plot (a textile plot) or some typical plots for financial data.
    - can revises SQL query to the financial database in a DandD instance according to user's request by mouse action through the plot or direct input of SQL query.
      - This makes it possible to obtain any subset from the data.

Then, we will discuss design of ICSFA through analysis of hedge fund return data

What is hedge fund ?

- A Hedge Fund
  - is an investment fund.
  - is open to a limited range of investors that is permitted by regulators to undertake a wider range of activities than other investment funds.

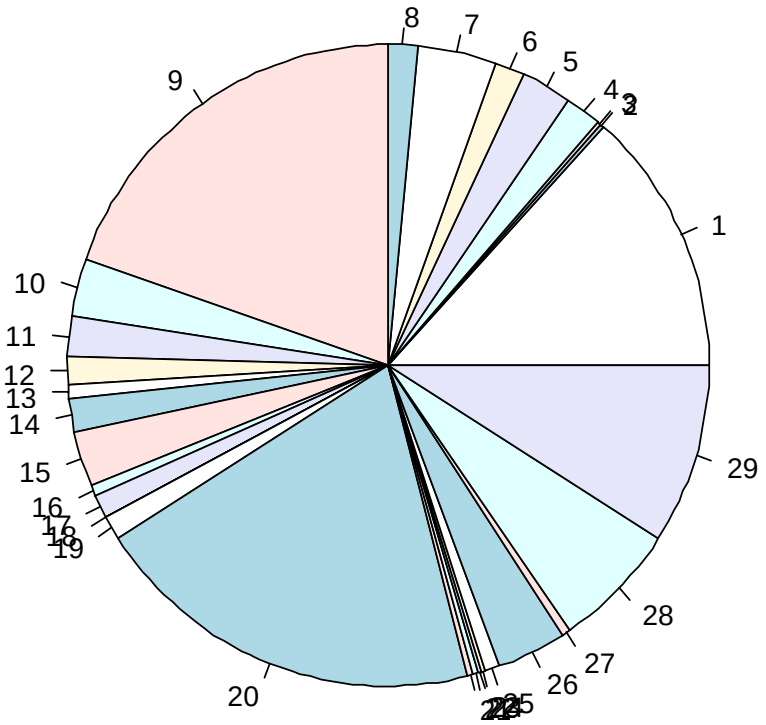
As the name implies, hedge funds often seek to offset potential losses in the principal markets they invest in by hedging their investments using a variety of methods, most notably short selling. However, the term "hedge fund" has come to be applied to many funds that do not actually hedge their investments.

Database

- The Center for International Securities and Derivatives Markets (CISDM) Hedge Fund Database by University of Massachusetts Amherst
  - 11,101 Hedge funds (including CTA) information
  - 662,040 monthly reports by the hedge funds from June 1972 to July 2007

[1] ""
[2] "Capital Structure Arbitrage"
[3] "Conservative"
[4] "Convertible Arbitrage"
[5] "Discretionary"
[6] "Distressed Securities"
[7] "Emerging Markets"
[8] "Equity Long Only"
[9] "Equity Long/Short"
[10] "Equity Market Neutral"
[11] "Event Driven Multi Strategy"
[12] "Fixed Income"
[13] "Fixed Income - MBS"
[14] "Fixed Income Arbitrage"
[15] "Global Macro"
[16] "Invest Funds in Parent Company"
[17] "Market Neutral"
[18] "Market Timing"
[19] "Merger Arbitrage"
[20] "Multi Strategy"
[21] "Opportunistic"
[22] "Option Arbitrage"
[23] "Other Relative Value"
[24] "Regulation D"
[25] "Relative Value Multi Strategy"
[26] "Sector"
[27] "Short Bias"
[28] "Single Strategy"
[29] "Systematic"

Proportion of hedge fund strategies in CISDM

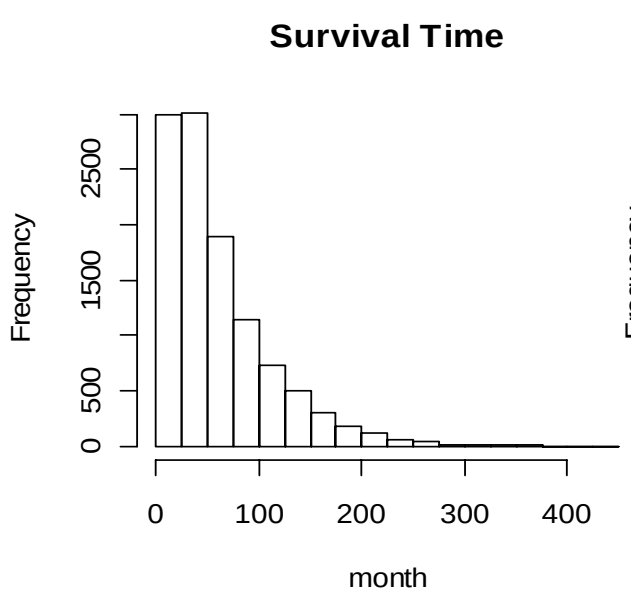


Each hedge fund recorded in CISDM database have one of 29 strategies.

The blank [1] “” means a hedge fund don’t report its investment strategy.

Ex. Survival Time of Hedge funds in CISDM database

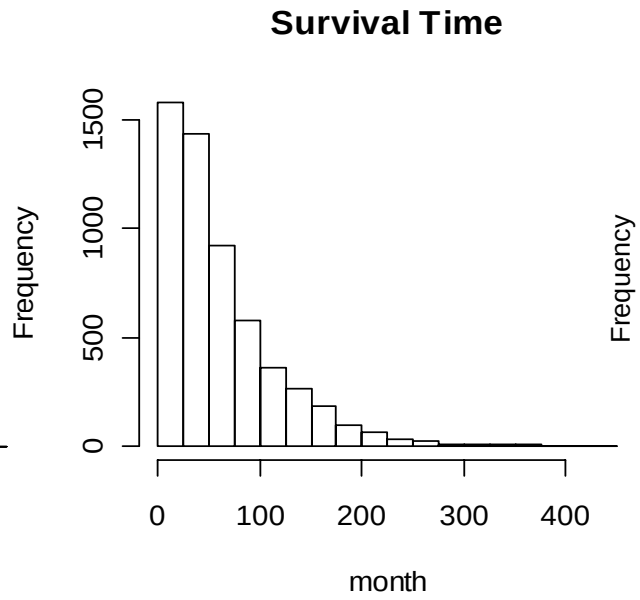
ALL



mean 60.29772

standard deviation 50.9629

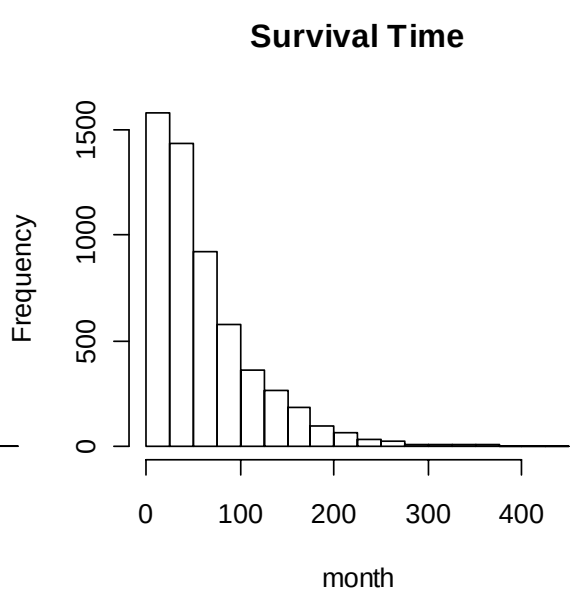
living until July2007



mean 61.27041

standard deviation 53.80356

not living by July2007



mean 59.29919

standard deviation 47.85614

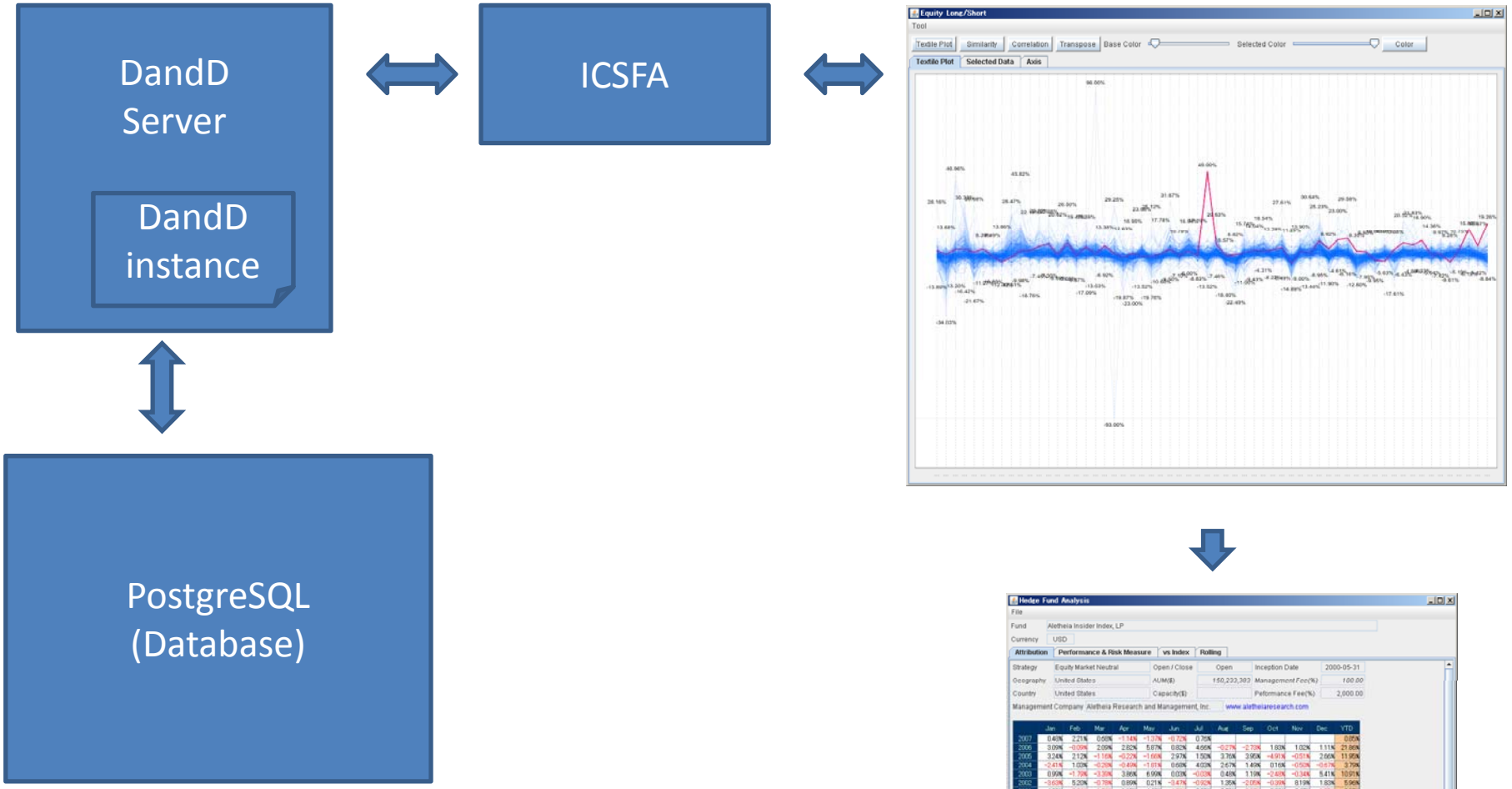
Major differences among three histograms are not found.

Detailed browsing of the data is necessary . Then, we started with construction of ICSFA for detailed examination of characteristics of hedge fund returns.

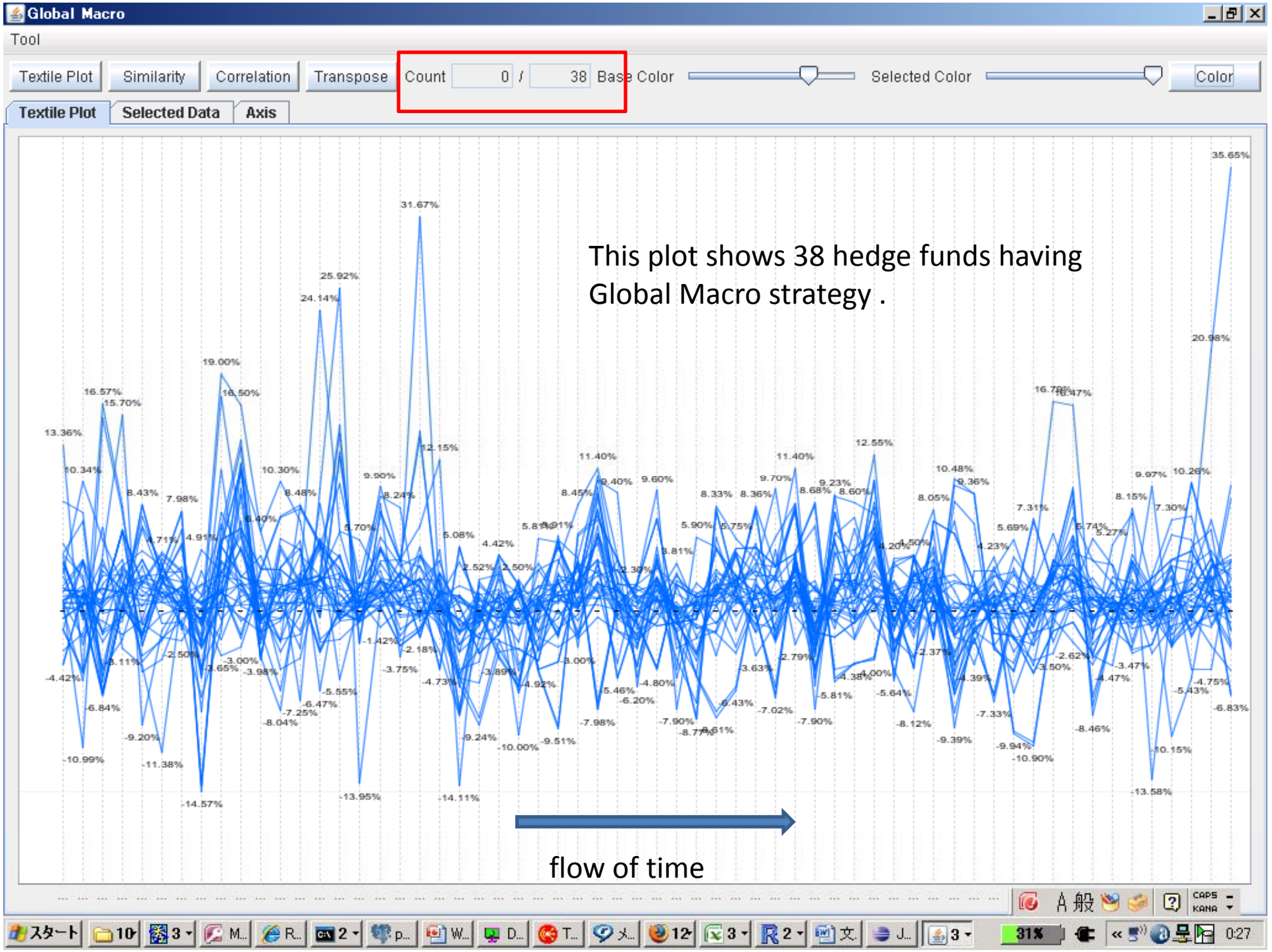
Our aim of data analysis

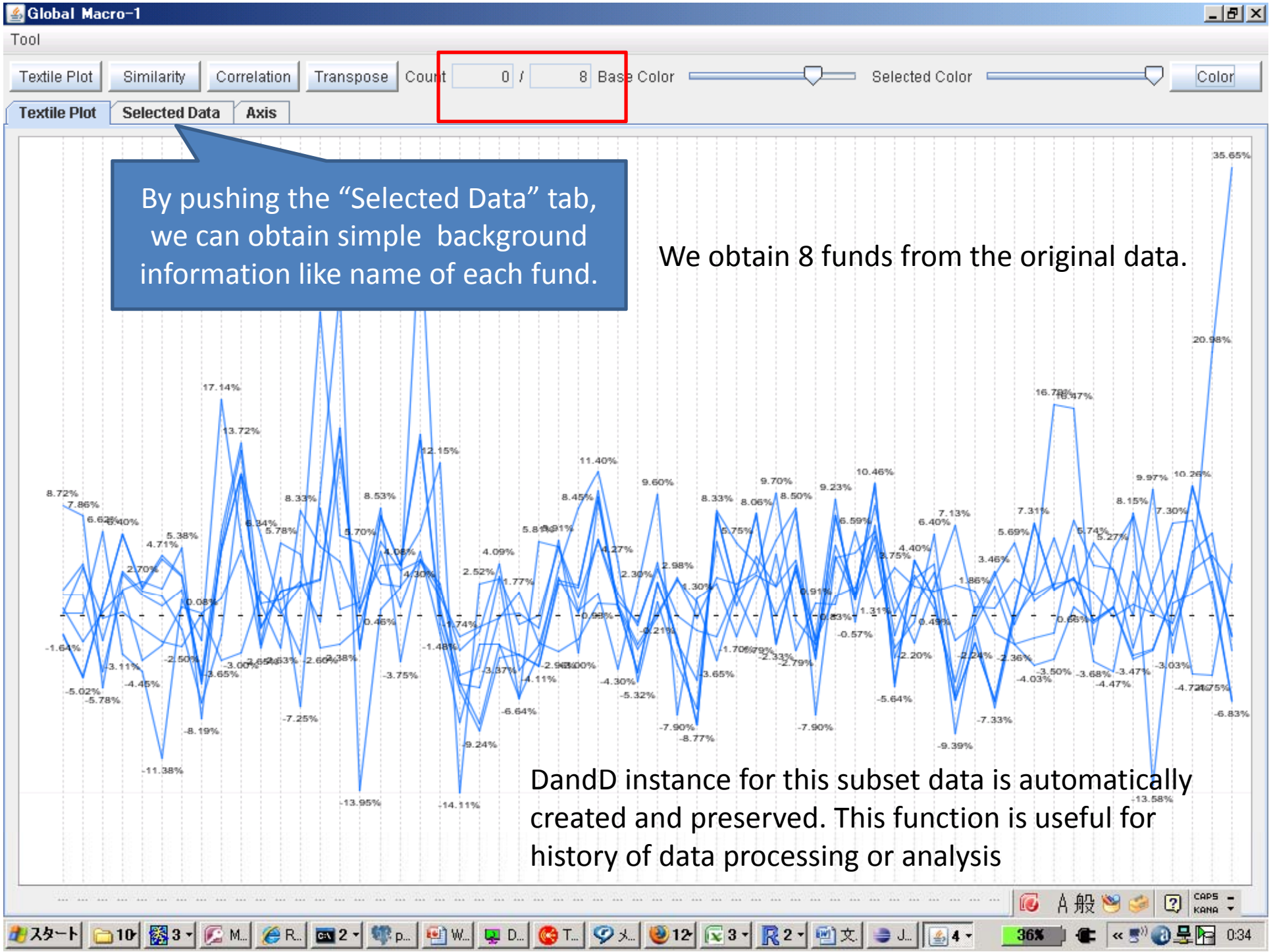
- We want to know
  - differences both living funds and dead ones.
    - How did attributes of hedge fund affect life and death of hedge funds ?
  - styles of hedge funds.
    - How hedge funds invest?

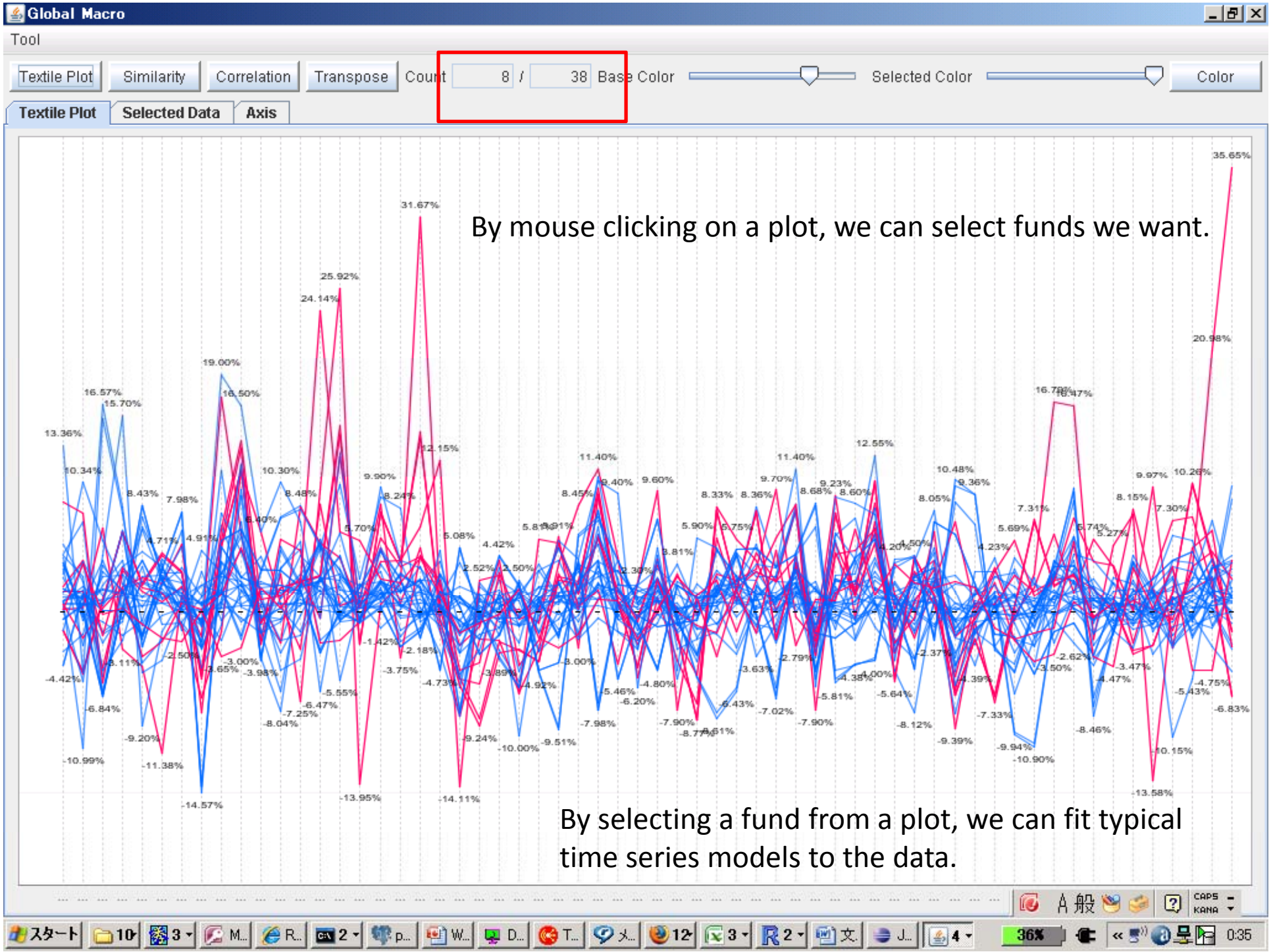
To begin with, we start with design of tools for these analysis, and implementation of them on ICSFA.

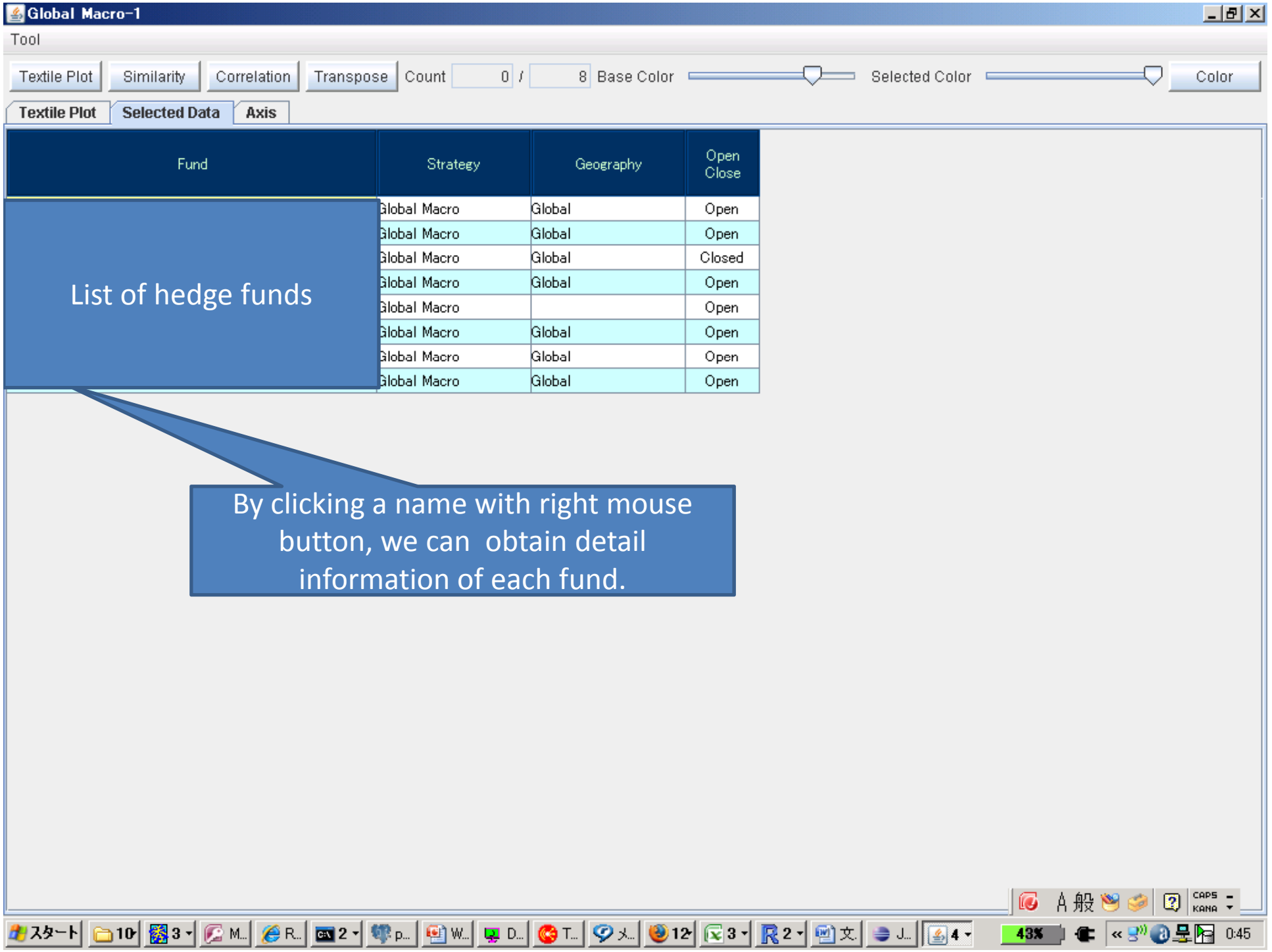


ICSFA adopts textile plot for an interface to handle the data. Through the interface, we can examine performance of individual hedge fund, or performances of hedge funds of each strategy.


Global Macro
Tool
Textile Plot
Similarity
Correlation
Transpose
Count 0 / 38
Base Color
Selected Color
Color
Textile Plot
Selected Data
Axis
This plot shows 38 hedge funds having Global Macro strategy .
flow of time


Global Macro-1
Tool
Textile Plot
Similarity
Correlation
Transpose
Count
0 / 8
Base Color
Selected Color
Color
Textile Plot
Selected Data
Axis
By pushing the “Selected Data” tab, we can obtain simple background information like name of each fund.
We obtain 8 funds from the original data.
DandD instance for this subset data is automatically created and preserved. This function is useful for history of data processing or analysis


Global Macro
Tool
Textile Plot
Similarity
Correlation
Transpose
Count
8 / 38
Base Color
Selected Color
Color
Textile Plot
Selected Data
Axis
By mouse clicking on a plot, we can select funds we want.
By selecting a fund from a plot, we can fit typical time series models to the data.


Global Macro-1
Tool
Textile Plot
Similarity
Correlation
Transpose
Count
0
/
8
Base Color
Selected Color
Color
Textile Plot
Selected Data
Axis
Fund
Strategy
Geography
Open Close
Global Macro
Global
Open
Global Macro
Global
Open
Global Macro
Global
Closed
Global Macro
Global
Open
Global Macro
Open
Global Macro
Global
Open
Global Macro
Global
Open
Global Macro
Global
Open
List of hedge funds
By clicking a name with right mouse button, we can obtain detail information of each fund.

Hedge Fund Analysis

File

Fund: hedge fund name

Currency: USD

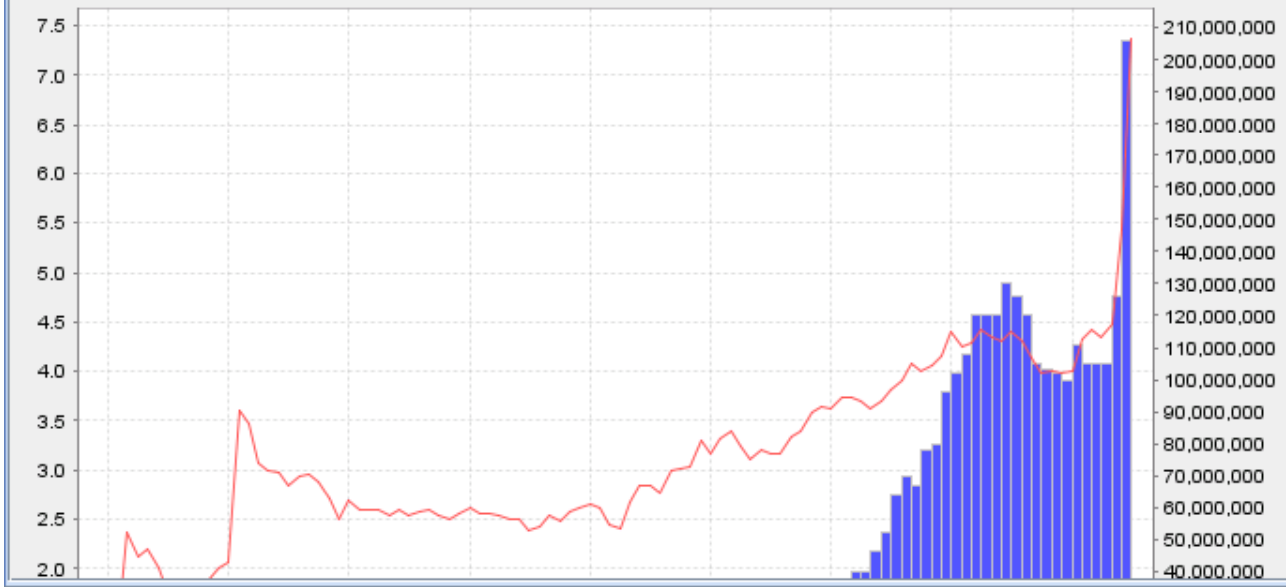
Attribution | Performance & Risk Measure | vs Index | Rolling

Strategy	Global Macro	Open / Close	Open	Inception Date	1999-01-31
Geography	Global	AUM($)	206,000,000	Management Fee(%)	1.00
Country	United States	Capacity($)		Peformance Fee(%)	20.00
Management Company	Balestra Capital, Ltd.				

Detailed background information of the fund.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	YTD
2007	0.55%	8.15%	1.80%	-1.61%	3.33%	20.98%	35.65%						84.70%
2006	6.28%	-3.39%	0.69%	3.34%	-2.07%	-0.71%	2.28%	-2.36%	-4.03%	-3.50%	0.66%	-0.61%	-3.88%
2005	-0.53%	3.64%	-0.08%	-1.22%	-1.78%	1.88%	3.19%	2.34%	4.39%	-1.99%	1.50%	2.19%	14.11%
2004	-3.75%	4.46%	2.49%	-4.76%	-4.18%	3.09%	-0.71%	-0.05%	5.41%	1.41%	5.89%	1.12%	10.12%
2003	1.50%	-1.20%	-6.96%	-1.65%	11.19%	6.34%	0.47%	-3.11%	8.50%	0.38%	0.97%	8.53%	26.06%
2002	2.54%	-2.34%	-0.18%	-0.69%	-0.89%	-0.44%	-4.43%	1.96%	4.08%	-2.03%	3.50%	1.86%	2.61%
2001	7.64%	-3.40%	-0.07%	-0.30%	-1.97%	2.27%	-2.18%	1.78%	0.76%	-2.34%	-1.37%	1.70%	2.06%
2000	3.65%	74.19%	-3.90%	-11.28%	-2.73%	-0.51%	-4.37%	2.83%	0.63%	-2.45%	-5.52%	-8.08%	24.89%
1999	25.44%	9.17%	72.98%	-10.06%	3.47%	-9.11%	-12.87%	-7.34%	-0.32%	-5.69%	23.02%	6.99%	100.15%

This table shows numeric return data.



スタート | 10 | 3 | M... | R... | 2 | p... | W... | D... | T... | 12 | 3 | 2 | 6 | 48% | 0:54

Hedge Fund Analysis

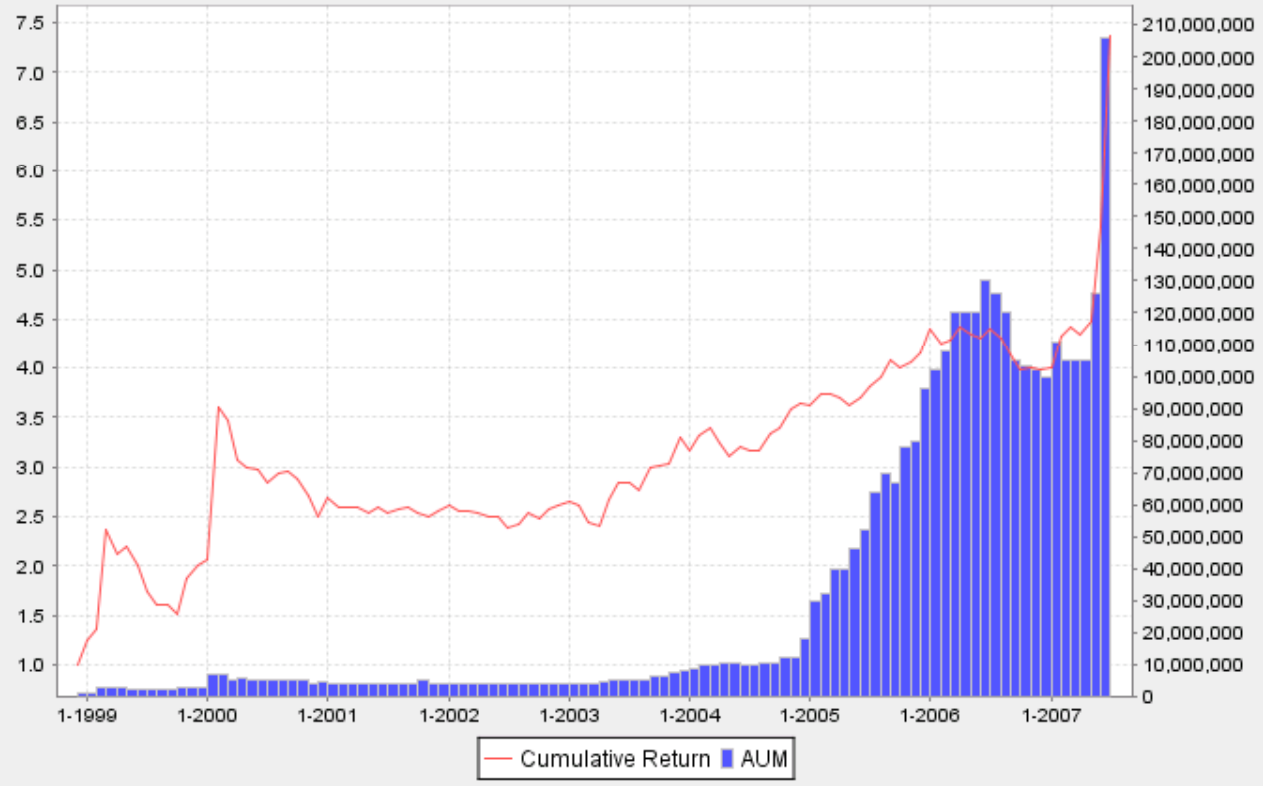
File

Fund hedge fund name

Currency USD

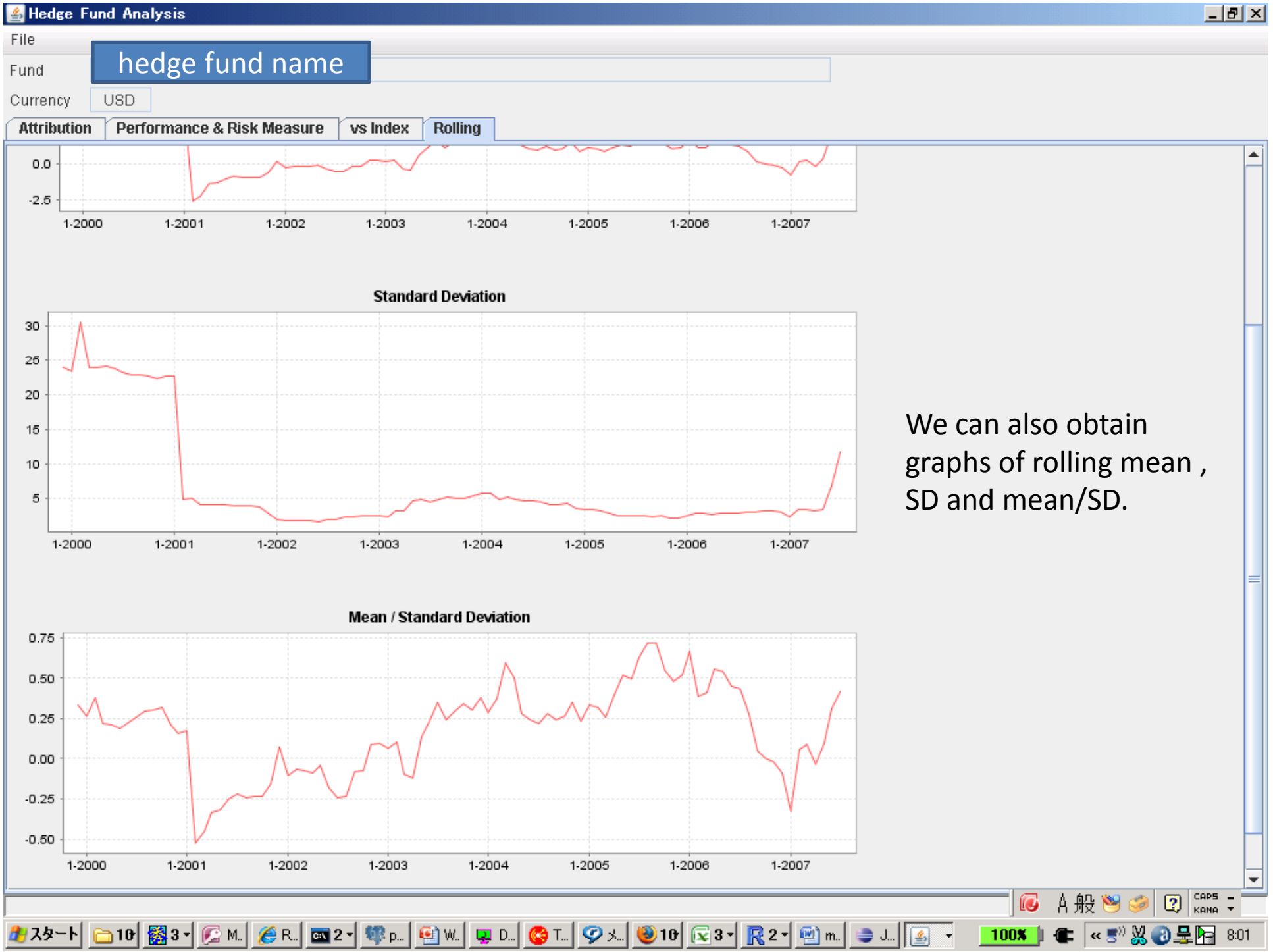
Attribution | Performance & Risk Measure | vs Index | Rolling

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	YTD
2007	0.55%	8.15%	1.80%	-1.61%	3.33%	20.98%	35.65%						84.70%
2006	6.28%	-3.39%	0.69%	3.34%	-2.07%	-0.71%	2.28%	-2.36%	-4.03%	-3.50%	0.66%	-0.61%	-3.88%
2005	-0.53%	3.64%	-0.08%	-1.22%	-1.78%	1.88%	3.19%	2.34%	4.39%	-1.99%	1.50%	2.19%	14.11%
2004	-3.75%	4.46%	2.49%	-4.76%	-4.18%	3.09%	-0.71%	-0.05%	5.41%	1.41%	5.89%	1.12%	10.12%
2003	1.50%	-1.20%	-6.96%	-1.65%	11.19%	6.34%	0.47%	-3.11%	8.50%	0.38%	0.97%	8.53%	26.06%
2002	2.54%	-2.34%	-0.18%	-0.69%	-0.89%	-0.44%	-4.43%	1.96%	4.08%	-2.03%	3.50%	1.86%	2.61%
2001	7.64%	-3.40%	-0.07%	-0.30%	-1.97%	2.27%	-2.18%	1.78%	0.76%	-2.34%	-1.37%	1.70%	2.06%
2000	3.65%	74.19%	-3.90%	-11.28%	-2.73%	-0.51%	-4.37%	2.83%	0.63%	-2.45%	-5.52%	-8.08%	24.89%
1999	25.44%	9.17%	72.98%	-10.06%	3.47%	-9.11%	-12.87%	-7.34%	-0.32%	-5.69%	23.02%	6.99%	100.15%



In this graph, the red line shows cumulative return of the fund and the blue bar chart shows amount of asset under management(AUM).

AUM is about 210 times as much as AUM at inception.


Hedge Fund Analysis
File
hedge fund name
Fund
Currency
USD
Attribution
Performance & Risk Measure
vs Index
Rolling
0.0
-2.5
1-2000
1-2001
1-2002
1-2003
1-2004
1-2005
1-2006
1-2007
Standard Deviation
30
25
20
15
10
5
Mean / Standard Deviation
0.75
0.50
0.25
0.00
-0.25
-0.50
We can also obtain graphs of rolling mean , SD and mean/SD.

Hedge Fund Analysis

File

Fund

hedge fund name

Currency USD

Attribution | Performance & Risk Measure | vs Index | Rolling

By pushing this tab, we can obtain result of performance and risk measurement of the fund.

2005	-0.53%	3.64%	-0.08%	-1.22%	-1.78%	1.88%	3.19%	2.34%	4.39%	-1.99%	1.50%	2.19%	14.11%
2004	-3.75%	4.46%	2.49%	-4.76%	-4.18%	3.09%	-0.71%	-0.05%	5.41%	1.41%	5.89%	1.12%	10.12%
2003	1.50%	-1.20%	-6.96%	-1.65%	11.19%	6.34%	0.47%	-3.11%	8.50%	0.38%	0.97%	8.53%	26.06%
2002	2.54%	-2.34%	-0.18%	-0.69%	-0.89%	-0.44%	-4.43%	1.96%	4.08%	-2.03%	3.50%	1.86%	2.61%
2001	7.64%	-3.40%	-0.07%	-0.30%	-1.97%	2.27%	-2.18%	1.78%	0.76%	-2.34%	-1.37%	1.70%	2.06%
2000	3.65%	74.19%	-3.90%	-11.28%	-2.73%	-0.51%	-4.37%	2.83%	0.63%	-2.45%	-5.52%	-8.08%	24.89%
1999	25.44%	9.17%	72.98%	-10.06%	3.47%	-9.11%	-12.87%	-7.34%	-0.32%	-5.69%	23.02%	6.99%	100.15%

Performance Measure

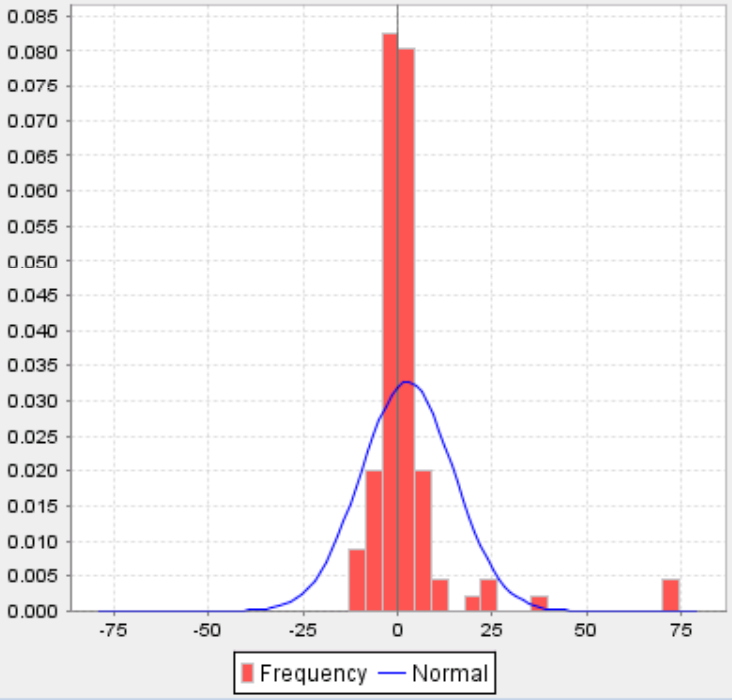
Cumulative Return	636.16	%
Best Monthly Return	74.19	%
Worst Monthly Return	-12.87	%
% of Positive Month	51.46	%
Annualized Return	26.18	%
Sharpe Ratio	0.18	
Calmar Ratio	0.73	

Risk Measure

Annualized Volatility	41.88	%	
Maximum Drawdown	35.80	%	
	90%	95%	99%
VaR(%)	13.0	17.4	25.6
Historical VaR(%)	4.6	7.9	12.8
CVaR(%)			
Historical CVaR(%)	8.1	10.2	12.8
Cornish Fisher VaR(%)	1.3	-7.3	-33.9
CDD(%)	32.4	33.6	35.7

Distribution

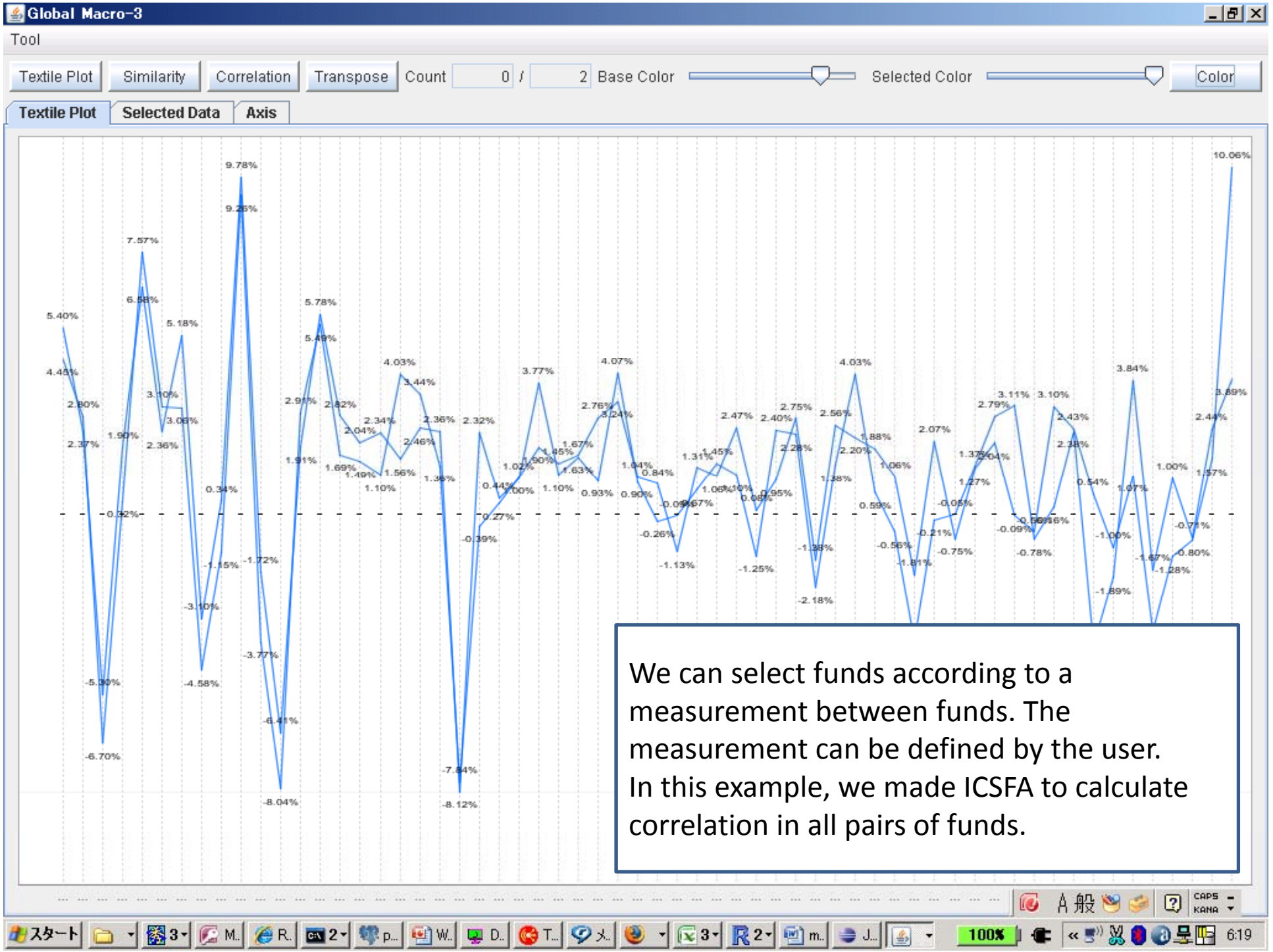
Mean(%)	2.50	Standard Deviation(%)	12.09
Skewness	4.33	Kurtosis	22.73

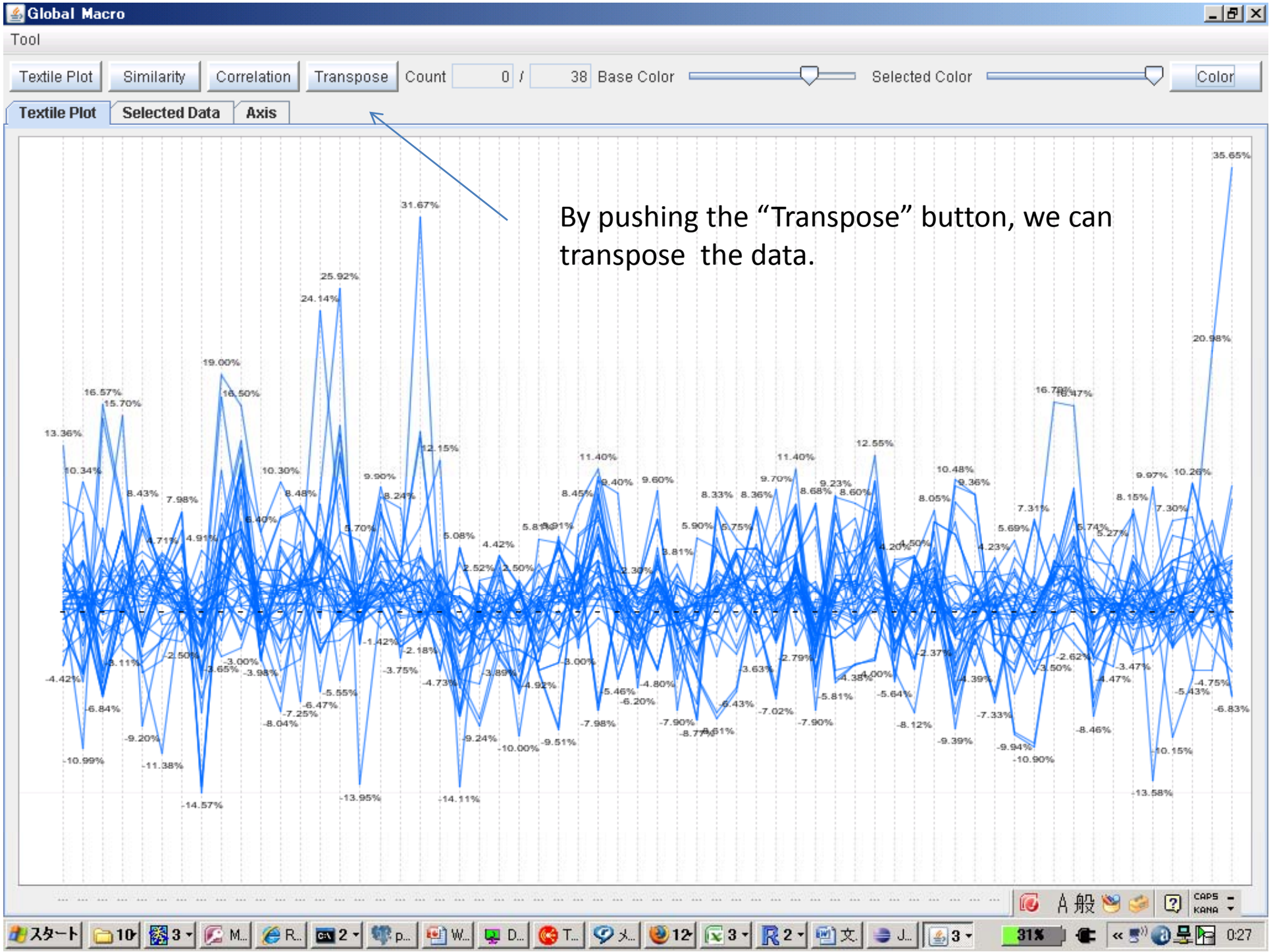


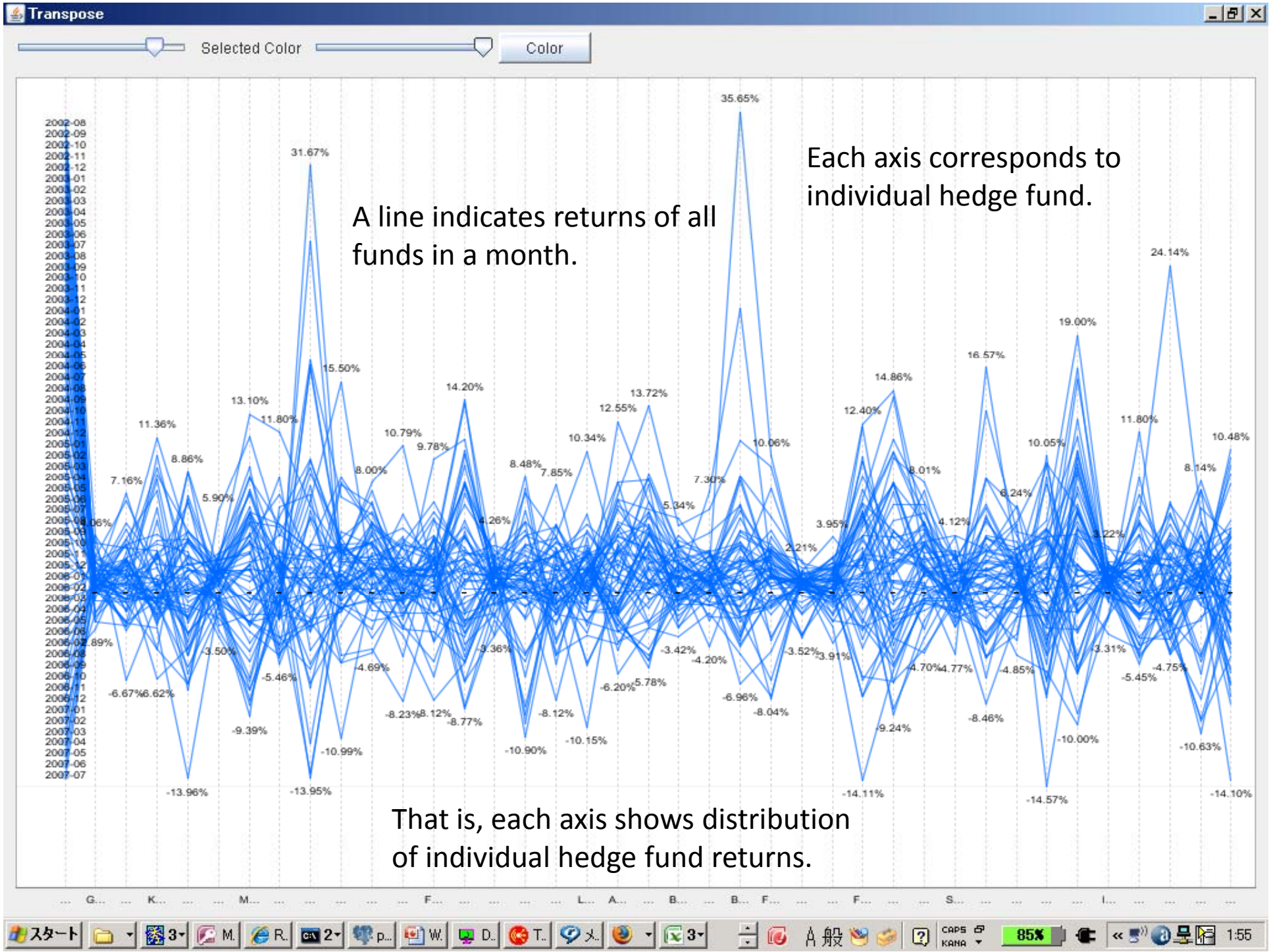
Moments of returns of all funds in CISDM database

	Strategy	Mean	SD	Skew	Kurt
1		0.007471	0.061303	4.345923	205.0127
2	Capital Structure Arbitrage	0.008614	0.013041	3.474266	34.78332
3	Conservative	0.007072	0.015282	−0.56906	11.95907
4	Convertible Arbitrage	0.008756	0.027995	1.487238	62.81736
5	Discretionary	0.014622	0.089664	8.155201	264.9909
6	Distressed Securities	0.011574	0.04051	0.676274	32.36574
7	Emerging Markets	0.013766	0.07957	0.709518	41.31701
8	Equity Long Only	0.012324	0.067362	0.056994	18.1374
9	Equity Long/Short	0.012009	0.067942	39.7254	5213.074
10	Equity Market Neutral	0.007051	0.033475	1.414135	41.86964
11	Event Driven Multi Strategy	0.011425	0.038583	0.642048	38.88433
12	Fixed Income	0.006823	0.022988	0.443012	13.93962
13	Fixed Income - MBS	0.009183	0.030608	0.864339	77.12253
14	Fixed Income Arbitrage	0.00615	0.02466	−4.17692	78.85956
15	Global Macro	0.009719	0.05483	1.203638	27.10782
16	Invest Funds in Parent Company	0.007146	0.021552	−0.23614	4.516506
17	Market Neutral	0.006849	0.017447	0.484736	18.4647
18	Market Timing	0.008538	0.029564	−0.12511	1.289813
19	Merger Arbitrage	0.007777	0.033112	20.44728	1086.558
20	Multi Strategy	0.008406	0.038352	119.2845	29289.89
21	Opportunistic	0.008343	0.037701	1.005266	8.471425
22	Option Arbitrage	0.006398	0.029542	1.451407	14.03414
23	Other Relative Value	0.009485	0.035798	1.803127	37.08924
24	Regulation D	0.015711	0.039871	9.919054	170.1813
25	Relative Value Multi Strategy	0.008159	0.027886	−1.50599	90.30763
26	Sector	0.014052	0.078061	1.042876	15.66722
27	Short Bias	0.002533	0.078479	0.155383	11.06083
28	Single Strategy	0.008205	0.040328	1.294828	26.13972
29	Systematic	0.011087	0.078994	2.450358	36.564

Returns of hedge fund seem to be asymmetric ally distributed with heavy tail .


Global Macro-3
Tool
Textile Plot
Similarity
Correlation
Transpose
Count
0
/
2
Base Color
Selected Color
Color
Textile Plot
Selected Data
Axis
We can select funds according to a measurement between funds. The measurement can be defined by the user. In this example, we made ICSFA to calculate correlation in all pairs of funds.


Global Macro
Tool
Textile Plot
Similarity
Correlation
Transpose
Count
0
/
38
Base Color
Selected Color
Color
Textile Plot
Selected Data
Axis
By pushing the “Transpose” button, we can transpose the data.


Transpose
Selected Color
Color
A line indicates returns of all funds in a month.
Each axis corresponds to individual hedge fund.
That is, each axis shows distribution of individual hedge fund returns.

Is a hedge fund dead or alive?

- Stopping reports by hedge fund to CISDM means
  - the hedge fund is bankrupt .
  - the hedge fund stops submitting performance report although it is still alive, because, for example, it earned much money.

In some previous work of hedge fund data analysis, the two reason for stopping report were not distinguished.

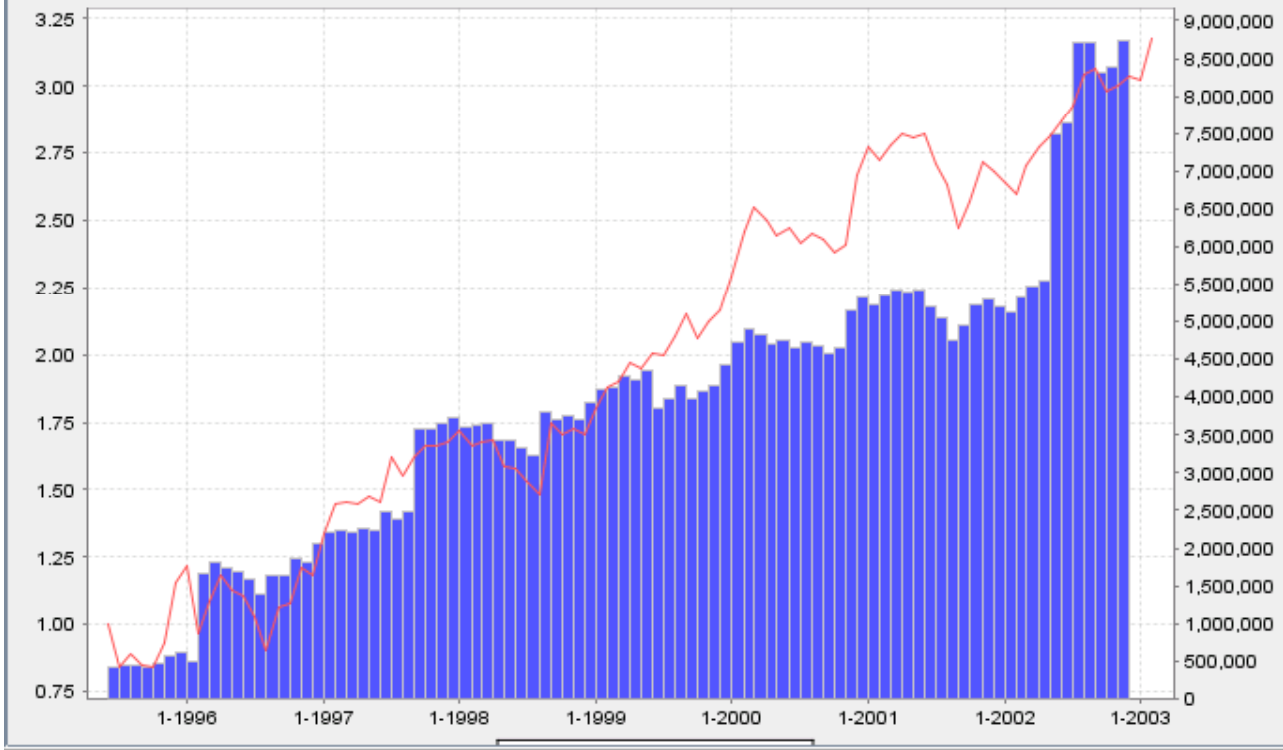
To distinguish living hedge funds from dead funds, we also need to browse and check both performance and amount of asset under management of individual fund.

Dead

Strategy Global Macro ☑ Only not Entried

Hedge Fund	Current	Alive	Dead	Unknown
List of hedge funds		Alive	Dead	Unknown
		Alive	Dead	Unknown
		Alive	Dead	Unknown
		Alive	Dead	Unknown
		Alive	Dead	Unknown

2001	3.94%	-1.89%	2.07%	1.67%	-0.52%	0.40%	-4.06%	-2.81%	-6.12%	4.04%	5.67%	-1.31%	0.39%
2000	6.08%	7.00%	3.85%	-1.86%	-2.18%	0.90%	-2.03%	1.27%	-0.73%	-2.07%	1.20%	10.71%	23.39%
1999	5.50%	4.42%	0.85%	3.93%	-1.29%	2.98%	-0.14%	3.21%	4.26%	-4.26%	2.83%	2.00%	26.65%
1998	2.58%	-3.31%	0.71%	0.33%	-5.41%	-0.61%	-3.07%	-3.09%	17.67%	-2.53%	1.48%	-1.16%	1.82%
1997	13.08%	8.21%	0.39%	-0.39%	1.86%	-1.20%	11.23%	-4.10%	4.33%	2.52%	0.01%	0.89%	41.76%
1996	5.56%	-20.67%	12.39%	8.49%	-4.47%	-1.61%	-6.97%	-12.22%	17.33%	1.33%	12.58%	-2.32%	2.47%
1995							-15.72%	5.34%	-4.70%	-0.74%	11.07%	23.78%	15.46%



The ICSFA provides us with a tool for checking performance and AUM of individual fund which stopped reporting before the last day of the data.
User's judgment as dead or alive can be recorded as auxiliary data.

This fund seems to have grown favorably. we think that it should not regarded as dead fund.

Style analysis

- There are a long tradition of characterizing investment funds according to parameters estimated via a linear model of returns.
  - ex. Capital Asset Pricing Model (CAPM)

$$E(R_{i,t}) - r_t = \beta_t (E(R_{M,t}) - r_t)$$

- ICSFA provides us with functions to fit asset pricing models or user defined models to the data.

Residual

Index S&P 500 INDEX

Selection of market index

Hedge Fund	Strategy	Geography	R2	α	α p	β	β P
	Equity Long/Short	United States	0.54	0.39%	0.52	1.89	8.18
	Equity Long/Short	United States					
	Equity Long/Short	UK					
	Equity Long/Short	North America					
	Equity Long/Short	Western Europe					
	Equity Long/Short	North America					
	Equity Long/Short	Global					
	Equity Long/Short	Japan					
	Equity Long/Short	United States	0.27	1.00%	2.16	0.67	4.66
	Equity Long/Short	United States	0.49	0.34%	0.87	0.91	7.44
	Equity Long/Short	North America	0.06	1.35%	4.75	0.17	1.89
	Equity Long/Short	Japan	0.03	0.39%	1.10	-0.15	-1.35
	Equity Long/Short	Global	0.09	1.64%	4.18	0.28	2.36
	Equity Long/Short	North America	0.27	0.47%	1.48	0.45	4.64
	Equity Long/Short	United States	0.51	0.30%	1.34	0.53	7.71

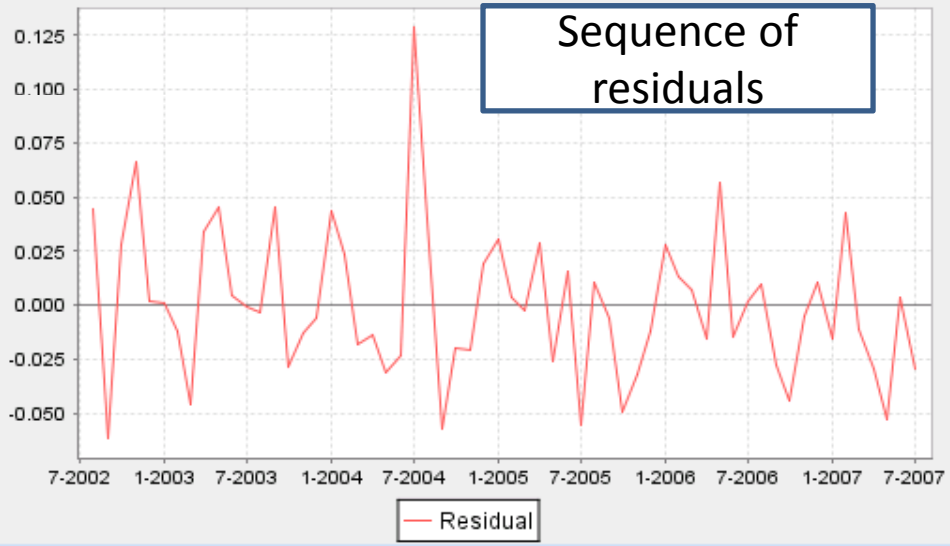
List of hedge funds

This table shows result of fitting an asset pricing model

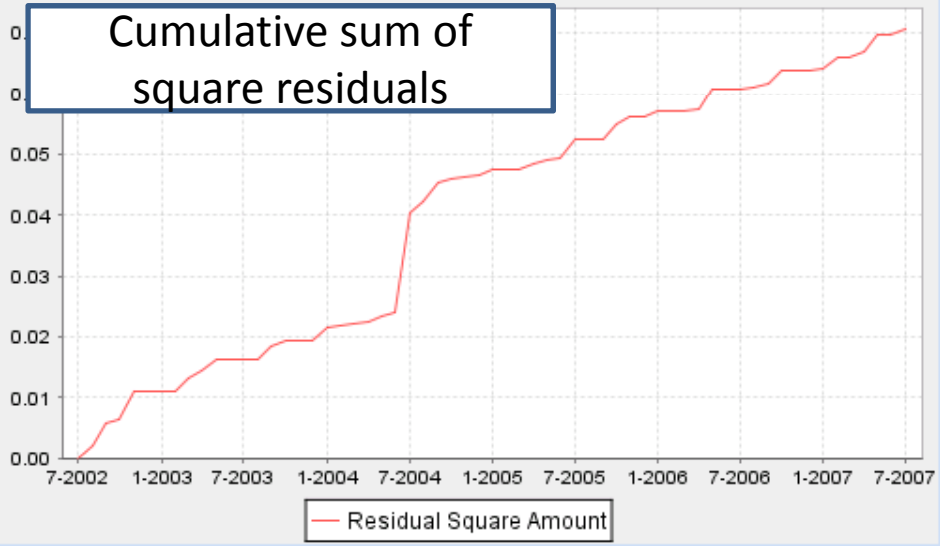
$$R_{i,t} = \alpha_t + \beta_t R_{M,t}$$

Residual Chart | Residual Data

Sequence of residuals



Cumulative sum of square residuals



スタート 98% 3:17

Conclusion

- ICSFA is a DandD client for helping interactive data analysis through a plot.
  - It may be useful for not only hedge fund return data but also other financial time series data.
- We are developing ICSFA, and analyzing hedge fund return data.
  - Our analysis of hedge fund return data is still quite primitive at the present.
  - We will present a result of our analysis for the hedge fund return data at 15th International Conference Computing in Economics and Finance held in Sydney on July 2009.

Thank you all for listening attentively.