

Day 3 - 2 June 2015(3hrs)

VBA Programming for MIF

Siraprapa Watakit
Technical Instructor,
BSc(Computer), MSc(Finance), PhD Student(Finance)

Day 1,2 - RECAP

Day 1: Module 1, Module 2 Part 1

- Record macros
- Implement SUBROUTINE vs. FUNCTION
- Datatype in VBA
 - Primitive: Integer, Single, Double and etc..
 - Object: Workbook, Worksheet and etc...
- VBA Integrated <u>Development</u> Environment, aka IDE(ALT+F8)
 - Project Explorer(Sheets, Modules, Class, Forms and etc..)
 - Properties Windows
 - Immediate Windows
 - Watch Windows
- Debugging code with F8(Step into), breakpoint, watch window and etc...

Day 2: Module 2 Part 2.1 and 2.2

- 2.1 Structural Programming if then else, for loop, error handling
- 2.2 Working with Excel Object Open /close multiple files, insert/delete multiple sheets

Day 3 – Agenda

- Module 3 Data Processing & Cleaning for Financial Research
 - Aggregate Data from multiple files
 - Insert a new sheet for each data file
 - Aggregate all data from all file in 1 sheet
 - Eliminate un-wanted data
 - Data transformation, from Time-Series Data to Panel Data
 - Data Processing tips when the data is from
 - Datastream
 - Bloomberg
 - Thomson Reuters Eikon Excel

Module 3 Data Processing & Cleaning for Financial Research

Workshop 1

- Suppose you HAVE 12 DATAFILES. You need to import them into a single workbook(create a new sheet for every imported file; hence there will be 12 new sheets)
 - o Source files: Workshop1.zip
- Think! Write down your idea here..
 - 1. First, I need to get the namelist of those file..(recall Day2, last program?)
 - 2. Then,..

Workshop 2

- Now, suppose you have many TimeSeries datafiles. You need to aggregate them into 1 big table, so that you can later import them into STATA, SAS or MatLab or Etc...; hence there will be only 1 sheet
 - Source files: Workshop2.zip
- Think! Write down your idea here...
 - 1. First, I need to get the namelist of those file..(recall Day2, last program?)
 - 2. Then,..

Workshop 2.1 – Delete BLANK data

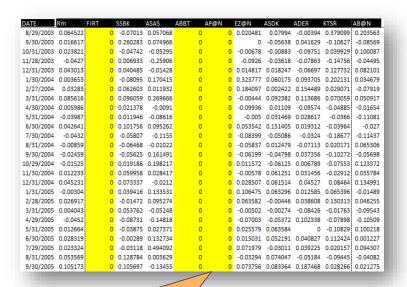
 Form previous problem, you'll see that some stocks are totally BLANK. It means that these stock are not EXISTED during the downloading period. DELETE them..

		PIRT STHO			DCO ABBT ITCP	TAKS APOH ACES	ACDG ACCR EZOH ACOP		O ADVA ARIC		ABGH
1/21/1991 2/29/1991				-0.15208 0.25062	-0.01726 0.05064	-0.09149 0.2041	-0.09051 0.26025	0.01182 -0.22285 0.05716 0.29152			0.01722
2/29/1991		0.01695		0.05411	0.12211	-0.01055	-0.00429	0.12768 0.07578		0.0201	
4/20/1991				-0.01406	0.0095	-0.01055	-0.00429	0.01942 -0.06965		-0.0201	
					0.0095			0.01947 -0.00865			
5/21/1991				0.00946		0.07146	-0.02227				0.05022
6/29/1991			-0.08798		-0.02229	-0.14942	-0.10066	-0.07281 -0.10758			-0.0622
7/21/1991			-0.02506		0.02441	0.00996	0.02401	-0.01255 0.05545			-0.04442
9/20/1991				-0.16049	0.02116	-0.20222	-0.10725	0.00722 -0.25225			-0.18141
9/20/1991				0.01106	-0.04797	0.11981	0.17441	0.07205 -0.02124			0.05215
			0.02276		0.07258	0.10212	0.07512	0.09266 -0.1652			-0.06695
				-0.09924	-0.02517	-0.19522	-0.14606	-0.14651 -0.06226			-0.12999
				0.09644	0.11991	0.05018	-0.02597	0.09195 -0.02921		-0.00669	
1/21/1992		0.095	-0.05226		-0.07555	-0.10425	-0.06227	-0.08582 -0.02785			0.11884
2/29/1992		0.1615	-0.10181		0.05404	-0.04159	-0.01129	0.01816 -0.14604			-0.14765
3/21/1992				-0.04015	0.00241	-0.05995	-0.07959	-0.04772 0.22597			-0.05835
4/20/1992			-0.09217		0.06744	-0.15092	-0.12914	-0.22272 0.02022			-0.09494
5/29/1992				0.22092	-0.0197	0.11274	0.14225	0.07297 0.04126			0.0774
6/20/1992				-0.09992	-0.92076	-0.20562	-0.15759	-0.19016 -0.04999			-0.11247
7/21/1992				-0.21952	0.05729	-0.07154	-0.09265	0.02965 -0.11993			-0.0927
9/21/1992				0.21022	-0.02622	0.22451	0.11575	0.05742 -0.01241			-0.0296
9/20/1992				-0.05292	-0.11164	-0.07916	-0.05297	0.09929 -0.01521			-0.11587
		0.042		-0.02196	0.07099	-0.05515	0.00249	0 -0.04453			-0.02077
				-0.02476	0.05992	0.02046	0.00248	0.10276 0.02251			0.02064
			-0.05824		0.01039	-0.10252	-0.02022	-0.12252 0.29749			0.14693
1/29/1992			-0.00574		-0.12085	-0.01864	0.01018	0.02641 0.02202			-0.02121
2/26/1992				0.02259	-0.07909	-0.02297	-0.09724	0.11952 0.0006			-0.02702
2/21/1992	0.10994	-0.02922		0.07514	-0.1057	0.16194	0.14291	0.01669 0.2706			0.22209
4/20/1992	0.12292	-0.09657	0.11779	0.10091	0.0492	0.09089	0.15589	0.08552 0.21587		0.15916	0.09726
5/21/1992	0.00971	-0.05915	0.04818	0.41872	0.00224	0.00759	0.07542	-0.02684 0.26941		0.12416	0.05965
6/20/1992	-0.02504	0.11988	0.05274	0.16229	-0.05997	-0.09516	-0.11225	-0.07207 -0.14994		-0.05576	0.01908
7/20/1992	0.04918	0	0.09577	0.05549	-0.02699	0.05504	0.04298	0.02576 0.15827		-0.01822	0.0279
9/21/1992	0.01979	0	0.09072	0.05657	0.00266	0.04516	-0.00516	0.0551 0.22526		-0.06722	-0.09518
9/20/1992	-0.04029	0	-0.09829	-0.15214	0.05774	-0.03988	-0.06821	-0.05972 0.29154		-0.05505	-0.04903

Hint: COUNTBLANK

Workshop 2.2 – Delete DEAD stocks

 Next, notice that some stock has a stream of 0, this PROBABLY means that these stock are DEAD or UNLISTED.
DELETE them as well...



o Hint: COUNIF

You'll need to count how many zero at **the last 60 trading days**...if there is more than **30** of zero, may be we can safely assume this stock is DEAD.... (Note: 30 of 60 or 50% of zeros number? This is subject to your judgment)

3 Data Type in Econometric

Don't get confused, this is basically a timeseries of price/volume data of <u>1 company</u>, sitting side by side with another company

- Time Series Data, the t subscript
 - Data of 1 company in many time(t) period

Date	AAPL	GOOG	BAC
8/28/2014	102.25	569.2	16.01
8/27/2014	102.13	571	16.2
8/26/2014	100.89	577.86	16.33
8/25/2014	101.54	580.2	16.29

Date	AAPL	GOOG	BAC
8/28/2014	68,389,800	1,289,400	61,862,400
8/27/2014	46,827,400	1,698,700	62,651,300
8/26/2014	33,119,800	1,635,200	73,164,900
8/25/2014	40,144,700	1,357,700	88,818,400

- Cross-sectional Data, the i subscript
 - o Data of many company(i) in 1 time period

Time	Stock	Price	Volume
8/28/2014	AAPL	102.25	68,389,800.00
8/28/2014	GOOG	569.2	1,289,400.00
8/28/2014	BAC	16.01	61,862,400.00

Time	JUCK	FIICE	Volume
8/26/2014	AAPL	100.89	33,119,800.00
8/26/2014	GOOG	577.86	1,635,200.00
8/26/2014	BAC	16.33	73,164,900.00

Time	Stock	Price	Volume
8/27/2014	AAPL	102.13	46,827,400.00
8/27/2014	GOOG	571	1,698,700.00
8/27/2014	BAC	16.2	62,651,300.00

Time	Stock	Price	Volume
8/25/2014	AAPL	101.54	40,144,700.00
8/25/2014	GOOG	580.2	1,357,700.00
8/25/2014	BAC	16.29	88,818,400.00

3 Data Type in Econometric

- Panel Data, the it subscript
 - Cross-Sectional+Time Series altogether

Time	Stock	Price	Volume
8/28/2014	AAPL	102.25	68,389,800.00
8/28/2014	GOOG	569.2	1,289,400.00
8/28/2014	BAC	16.01	61,862,400.00

Time	Stock	Price	Volume
8/27/2014	AAPL	102.13	46,827,400.00
8/27/2014	GOOG	571	1,698,700.00
8/27/2014	BAC	16.2	62,651,300.00

Time	Stock	Price	Volume
8/26/2014	AAPL	100.89	33,119,800.00
8/26/2014	GOOG	577.86	1,635,200.00
8/26/2014	BAC	16.33	73,164,900.00

Time	Stock	Price	Volume
8/25/2014	AAPL	101.54	40,144,700.00
8/25/2014	GOOG	580.2	1,357,700.00
8/25/2014	BAC	16.29	88,818,400.00

Sort By Time

	Time	Stock	Price	Volume
Ī	8/28/2014	AAPL	102.25	68,389,800.00
	8/28/2014	GOOG	569.2	1,289,400.00
	8/28/2014	BAC	16.01	61,862,400.00
ı	8/27/2014	AAPL	102.13	46,827,400.00
ı	8/27/2014	GOOG	571	1,698,700.00
	8/27/2014	BAC	16.2	62,651,300.00
	8/26/2014	AAPL	100.89	33,119,800.00
ı	8/26/2014	GOOG	577.86	1,635,200.00
	8/26/2014	BAC	16.33	73,164,900.00
ı	8/25/2014	AAPL	101.54	40,144,700.00
ı	8/25/2014	GOOG	580.2	1,357,700.00
	8/25/2014	BAC	16.29	88,818,400.00

Sort By Stock

Time	Stock	Price	Volume		
8/28/2014	AAPL	102.25	68,389,800.00		
8/27/2014	AAPL	102.13	46,827,400.00		
8/26/2014	AAPL	100.89	33,119,800.00		
8/25/2014	AAPL	101.54	40,144,700.00		
8/28/2014	BAC	16.01	61,862,400.00		
8/27/2014	BAC	16.2	62,651,300.00		
8/26/2014	BAC	16.33	73,164,900.00		
8/25/2014	BAC	16.29	88,818,400.00		
8/28/2014	GOOG	569.2	1,289,400.00		
8/27/2014	GOOG	571	1,698,700.00		
8/26/2014	GOOG	577.86	1,635,200.00		
8/25/2014	GOOG	580.2	1,357,700.00		

Workshop 3

Convert Time-Series to Panel Data

Source files: Workshop3

DATE	Market	SSBK	ASAS	EZ@N	ASDK	ADER	KTSR	AB@N
1/31/1991	-0.0133	-0.00396	-0.15208	-0.09051	0.011832	-0.32285	-0.08011	0.017235
2/28/1991	0.136071	0.121068	0.350631	0.360246	0.057158	0.291533	0.24784	0.291934
3/29/1991	0.005296	0.01344	0.054105	-0.00438	0.127681	0.075782	0.020099	0.117698
4/30/1991	-0.00372	0.088963	-0.01406	-0.03467	0.01942	-0.06865	-0.03107	-0.0296
5/31/1991	0.000687	0.010572	0.008457	-0.02327	0	-0.02884	0.01148	0.050235
6/28/1991	-0.07708	-0.08798	-0.07433	-0.10066	-0.07381	-0.10758	-0.09841	-0.0632
7/31/1991	0.021811	-0.03506	-0.0417	0.024015	-0.01355	0.055448	0.01624	-0.04442
8/30/1991	-0.07077	-0.01811	-0.16049	-0.10725	0.007316	-0.25235	-0.14383	-0.18141
9/30/1991	0.056183	0.089858	0.011063	0.174413	0.073047	-0.03134	0.030531	0.052153
10/31/1991	0.029709	0.032765	0.119005	0.075131	0.092659	-0.1653	0	-0.06695
11/29/1991	-0.08635	0.010749	-0.08834	-0.14606	-0.14651	-0.06226	-0.10322	-0.12998
12/31/1991	-0.00965	0.04161	0.096436	-0.02597	0.091951	-0.03931	-0.00669	0
1/31/1992	-0.05007	-0.05236	-0.11437	-0.06337	-0.08582	-0.02785	-0.08867	0.118842
2/28/1992	-0.04801	-0.10181	-0.04059	-0.01128	0.018157	-0.14604	-0.04994	-0.14765

- o Hint
 - You should name "DATESERIES", becoz you will reuse it quite a lot for copy paste
 - You can also count how many row in a range selection for e.g. Range("DATESERIES").Rows.Count

Date	Stock	Price
1/31/1991	Market	-0.013296
2/28/1991	Market	0.1360706
3/29/1991	Market	0.0052963
4/30/1991	Market	-0.003716
5/31/1991	Market	0.0006873
1/31/1991	SSBK	-0.003956
2/28/1991	SSBK	0.1210677
3/29/1991	SSBK	0.0134395
4/30/1991	SSBK	0.0889627
5/31/1991	SSBK	0.0105723
1/31/1991	ASAS	-0.152083
2/28/1991	ASAS	0.3506313
3/29/1991	ASAS	0.0541051
4/30/1991	ASAS	-0.014056
5/31/1991	ASAS	0.0084574
6/28/1991	ASAS	-0.074326
1/31/1991	EZ@N	-0.090509
2/28/1991	EZ@N	0.3602464
3/29/1991	EZ@N	-0.004382
4/30/1991	EZ@N	-0.034668
5/31/1991	EZ@N	-0.023275

Data Processing: DataStream

- Source files: <u>Sample_Datastream_Indo.xlsm</u>
 - Data Summary: The files contains 4 sheets: price(P),total price(RI),market value(MV) and the list of ordinary equity
 - Things to clean
 - Delete all ERROR, NODATA stock
 - Delete all that IS NOT ordinary stock
 - Delete all "NA", "\$\$ERR" and etc.
 - Then
 - Rename all column to the stock name, e.g. ID:WHU(P)→WHU
 - Since each sheet represent different fundamental data, save each sheet to new CSV file...(CSV can be imported by many other program such as SAS, STATA, MatLab and etc..)
 - Macro
 - RunAll_Loop

Data Processing: Bloomberg and Eikon

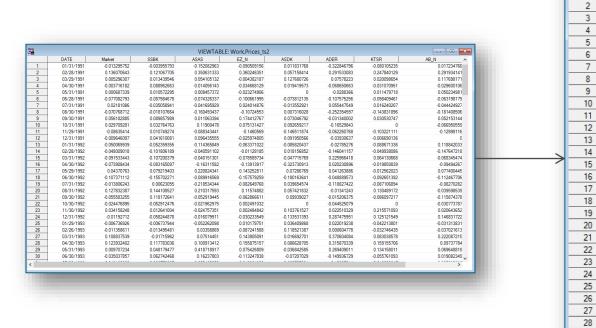
- Pretty much the same logics from the previous programs can handle the data from both Bloomberg and Eikon
 - Bloomberg : https://db.tt/QMCTacEF
 - Eikon: https://db.tt/dzGQnoxA
- The missing values... what to do with it?
 - Either 1) delete the whole line
 of observation or, 2) replace with
 the previous price, or 3) consolidate them
 - SAS PROC EXPAND/TIMESERIES can do it with 4 line of code..

```
time
            1378.50
10:00:04
            1379.50
10:00:07
10:00:08
                         1379.50
10:00:15
            1379.00
                         1379.80
            1378.25
10:00:18
10:00:19
                         1380.00
10:00:22
                                     1380.78
10:00:23
            1378.00
```

```
PROC EXPAND DATA=PRICES_TS
OUT=PRICES_TS_FULL TO=SECOND METHOD=STEP;
ID TIME;
VAR FES FSP SPY;
RUN;
```

```
PROC TIMESERIES DATA=PRICES_TS
   OUT=PRICES_TS_5MIN;
   ID TIME INTERVAL=MINUTE5 ACCUMULATE=AVERAGE;
   VAR FES FSP SPY;
RUN;
```

Convert Time-Series to Panel Data Using SAS



```
PROC TRANSPOSE DATA=PRICES_TS2

OUT=PRICE_PANEL(RENAME=(_NAME_=STOCK COL1=PRICE));

VAR AB_N ADER ASAS ASDK EZ_N KTSR MARKET SSBK;

BY DATE;

RUN;

PROC SORT DATA=PRICE_PANEL; BY STOCK; RUN;
```

DATE

01/31/1991 AB N

02/28/1991 AB N

03/29/1991 AB N

04/30/1991 AB N

05/31/1991 AB N

06/28/1991 AB N

07/31/1991 AB_N

08/30/1991 AB_N

09/30/1991 AB N

10/31/1991 AB N

11/29/1991 AB N

12/31/1991 AB N

01/31/1992 AB_N

02/28/1992 AB N

03/31/1992 AB N

04/30/1992 AB N

05/29/1992 AB N

06/30/1992 AB N

07/31/1992 AB N

08/31/1992 AB N

09/30/1992 AB N

10/30/1992 AB_N

11/30/1992 AB_N

12/31/1992 AB_N

01/29/1993 AB N

02/26/1993 AB N

03/31/1993 AB N

04/30/1993 AB N

STOCK

PRICE

0.017234766

0.291934141

0.117698171

-0.029600106

0.050234981

-0.063198179

-0.044424927

-0.181408506

0.052153144

-0.066950559

-0.12998116

0.118842033

-0.147647218

-0.068345474

-0.09494267

0.077400445

-0.112467706

-0.08270282

-0.039598539

-0.115874378

-0.030773787

0.020643652

0.146831722

-0.031313831

-0.037021613

0.222087215

0.09737784

Recommendations...

- Microsoft Online Training
 - http://office.microsoft.com
- Excel Quick Tips Videos
 - www.contextures.com
- Advanced VBA Excel Programming
 - www.cpearson.com
- Excel VBA & Macros Tips & Techniques
 - www.ozgrid.com/VBA/



End of The Show! Thank You Siraprapa Watakit

