8

MULTI-FACTOR LINEAR MODELS: FF3, FFC4 AND FF5

Multi-factor models are a simple extension of the single-factor model such as the CAPM (Capital Asset Pricing Model). Using Excel took, there is no difference between running a single-factor model or running a 3-factor model. Thus, we would focus on the meaning of various multi-factor model such as Fama-French 3-factor model, Fama-French-Carhart 4-factor model. In addition, we explain how to compare the performance of a single-factor model with a multi-factor model. In this chapter, the following topics will be covered:

- Review of CAPM
- A general format for a multi-factor linear model
- Fama-French 3-factor model
- Fama-French-Charhart 4-factor model
- Fama-French 5-factor model
- Adding your own factor
- Performance measure of multi-factor model

8.1 REVIEW OF A SINGLE-FACTOR MODEL AND CAPM

The general one-factor linear model is shown below.

$$y = \alpha + \beta * x \tag{1}$$

CAPM can be viewed as one-factor linear model. It has the following form.

$$E(R_i) = \alpha + \beta[E(R_{mkt}) - R_f], \tag{2}$$

where E() is the expectation, Ri is the return for security i, β is the slop or the market risk, R_{mkt} is the market return, such as S&P500, and $E(R_{mkt})$ is the expected market return. In terms of the logic, the follow formula offers a better picture.

$$R_{i,t} - R_{f,t} = \alpha + \beta\left(R_{m,t} - R_{f,t}\right) + \varepsilon_t, \tag{3}$$

where $R_{i,t}$ is the return on stock i at time t, $R_{f,t}$ is the risk-free rate at time t, and $R_{m,t}$ is the market return (such as the S&P500) at time t. The left-hand side is the risk-premium of an underlying security i, while the difference between the market and risk free rate is the market risk premium, i.e., be consistent with the left-hand side.

8.2 A GENERAL MODEL FOR MULTI-FACTOR LINEAR MODEL

A generate formula for n-factor linear model is given below.

$$y = \alpha + \sum_{i=1}^{n} \beta_i x_i + \epsilon_t \quad (4)$$

For example, for a 3-factor model, we have the following formula.

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon_t \quad (5)$$

The R codes to generate three input variables and run the above model are in the left panel below. The result, in the right panel, indicates that the intercept is 4.900 and three coefficients are 3.027, -1.281 and 5.005, respectively.

The logic behind the so-called 52-week high and low trading strategy is that if today's price is close to 50-week high (low*)*, there is little chance that the price will go higher (lower). To view 52-week's range for IBM, see below.

Since today's price is the 52-week's high (174.54 vs. 174.54), we should not purchase IBM's stock, i.e., we should sell it. Actually, we could use just one line R codes to get a 52-week's range.

8.3 FAMA-FRENCH 3-FACTOR MODEL

The famous Fama-French 3 factor model is formulated below.

$$R_i = R_f + \beta_{mkt}(R_m - R_f) + \beta_{SMB} * SMB + \beta_{HML} * HML\,, \quad (6)$$

where *SMB* (Small Minus Big) is defined as the returns of the small portfolios minus returns of the big portfolios, *HML* (High Minus Low) is the difference of returns of high book-to-market portfolios minus the returns of low book-to-market portfolios. See Ken French data library at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html for more detailed definitions.

8.4 *SMB* (SMALL MINUS BIG)

Since small stocks are more risky, investors would demand high returns for bearing extra risks for holding them compared with big stocks which are usually stocks for more mature companies. Fama and French

(1992, 1993) construct a risk factor called SML (Small Minus Big), http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, which is the average return on 3 small stock portfolios minus the average return on 3 big stock portfolios,

$$SMB = \frac{1}{3}(SH + SM + SL) - \frac{1}{3}(BH + BM + BL), \quad (3)$$

where *S* stands for Small, *B* for Big, *H* for High, *M* for Median and *L* for Low. Those 6 portfolios are constructed in the following way. First, at the end of each year, the market capitalizations (stock price time shares outstanding) for the stocks listed on NYSE are ranked. We calculate a median. A stock is classified as small if its market capitalization is below this median, otherwise it is a big stock. Second, the ratio of Book-value of Equity over its market values (BE/ME) is estimated for each stock. Third, all stocks are broken into 3 book-to-market equity groups based on the breakpoints for the bottom 30% (Low), middle 40% (Median) and top 30% (High) of the ranked values of BE/ME for the NYSE listed stocks. They define book value of equity (BE) as the Compustat book value of stockholders equity, plus balance sheet deferred taxed and investment tax credit, minus the book value of preferred stocks. If a stock is small and belong to high book-to-market group, it will be in the portfolio of SH. Other stocks will be classified accordingly. Eventually, each stock will be classified into one of those 6 portfolios: SH, *SM*, SL, BH, BM and BL. To estimate the returns of those 6 portfolios, the value-weighted method is applied. The weight of individual stocks is its market capitalization over the total market capitalization at the end of previous year.

8.5 *HML* (HIGH MINUS LOW)

HML is the difference in portfolio returns between the portfolio with high book-to-market values and the portfolio with low book-to-market values.

$$HML = \frac{1}{2}(SH + BH) - \frac{1}{2}(SL + BL), \quad (4)$$

where, SH, BH, SL and BL are defined above.

8.6 DOWNLOADING FAMA-FRENH FACTORS

The related web page is http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.The above web page contains both daily and monthly Fama-French factors, two most widely used frequencies. For the monthly factors, click "Fama-French Factors" to download a zip file called F-F_Research_Data_Factors.zip. After unzip it, we will have monthly and annual Fama-French factors. A few top lines are shown below.

```
This file was created by CMPT_ME_BEME_RETS using the 201012 CRSP
database.
The 1-month TBill return is from Ibbotson and Associates, Inc.

        Mkt-RF     SMB     HML      RF
192607    2.62   -2.16   -2.92    0.22
```

```
192608    2.56   -1.49    4.88    0.25
192609    0.36   -1.38   -0.01    0.23
192610   -3.43    0.04    0.71    0.32
192611    2.44   -0.24   -0.31    0.31
```

We can use following R codes to check the first several records. “skip=4” is used to skip the first 4 lines, i.e., skip the header. "*nrow=10*" imports 10 lines.

```
> x<-read.table("F-F_Research_Data_factors.txt",skip=4,nrow=10)
> colnames(x)<-c("yyyymm","mkt_rf","SMB","HML","Rf")
> x
   yyyymm mkt_rf   SMB   HML   Rf
1  192607   2.62 -2.16 -2.92 0.22
2  192608   2.56 -1.49  4.88 0.25
3  192609   0.36 -1.38 -0.01 0.23
4  192610  -3.43  0.04  0.71 0.32
5  192611   2.44 -0.24 -0.31 0.31
6  192612   2.77 -0.01 -0.10 0.28
7  192701  -0.11 -0.30  4.79 0.25
8  192702   4.32 -0.24  3.35 0.26
9  192703   0.32 -1.87 -2.58 0.30
10 192704   0.41  0.29  0.95 0.25
```

Through a trial-and-error, we could use “nrow=1014” to get all monthly factors. If omit “nrow=1014”, we could find an error message. The reason is that the original text file contains both monthly and annual factors. We could manually remove the annual factors. Then, we don’t have to specify the number of rows we would import.

```
> x<-read.table("F-F_Research_Data_factors.txt",skip=4,nrow=1014)
> tail(x)
         V1    V2    V3    V4   V5
1009 201007  7.24 -0.08  0.13 0.01
1010 201008 -4.40 -2.92 -1.71 0.01
1011 201009  9.24  3.97 -3.14 0.01
1012 201010  3.89  0.91 -2.14 0.01
1013 201011  0.58  3.72 -0.61 0.01
1014 201012  6.77  0.81  3.55 0.01
```

Below are the complete codes. Please note that the last 4 columns divided by 100.

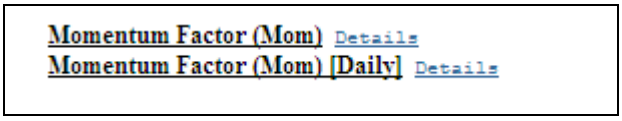
```
> x<-read.table("F-F_Research_Data_factors.txt",skip=4,nrow=1014)
> colnames(x)<-c("date","mkt_rf","SMB","HML","Rf")
> x[,2:5]<-x[,2:5]/100            # from percent to decimal
> ff_monthly_factors<-x           # use a better name
> save(ff_monthly_factors,file="ff_monthly_factors.RData")
```

8.7 MOMENTUM STRATEGY

Momentum is defined as a stock moves in the same direction for consecutive periods of time. Accordingly, a trading strategy, called "buy winners and sell losers", could be desinged. Jagadeesh and Titman (1993) test this strategy. For example, they use 6 months as an evaluation period to classify winners and losers and the next 6 months as a holding period. First, they rank all stocks into 10 portfolios (deciles) based on their past 6 months total returns. The highest return decile is called 'winner' while the bottom decile is called losers. The trading strategy is to take a long position with winners and a short position with lowers, i.e., long the 1st portfolio and short the 10th portfolio. For each month, they re-classify both winners and losers, and rebalance their long-short portfolio. They find that this strategy is indeed a profitable trading strategy. Note that Jagadeesh and Titman (1993) do not consider transaction costs.

8.8 DOWNLOADING THE MOMENTUM FACTORS

If add the momentum factor, we could have a 4-fatcor model. To download the momentum factor, go to Prof. Ken French data library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html



Below is the Fama-French-Carhart's 4-factor model.

$$R_i = R_f + \beta_{mkt}(R_m - R_f) + \beta_{SMB} * SMB + \beta_{HML} * HML + \beta_{Mom} * MOM \quad (5)$$

To download those 4-factors (Market, SMB, HML and Momentum), we go to Prof. Ken French's data library to download two zipped files for those 4 factors.

8.9 HOW TO RUN A MULTI-FACTOR LINEAR MODEL

First, let's look at the general formula for a single-linear regression.

$$y = \alpha + \beta * x \quad (5)$$

For the following example, we generate x and y columns with known alpha and beta.

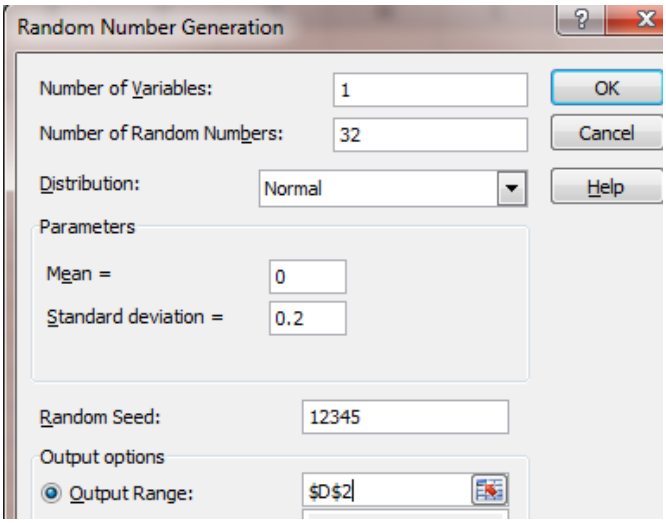
$$y = 0.134 + 1.2x + \varepsilon_t \quad (6)$$

The alpha is 0.134, beta is 1.2 while y is the function of x plus a random variable. First, we assume that x takes values from -0.5 with an incremental value of 0.025, see below.

C3 =C2+B3

	A	B	C	D	E
1	alpha	0.134	x		
2	beta	1.2	-0.5		
3	delta	0.025	-0.475		
4			-0.45		
5	mean	0	-0.425		
6	std	0.2	-0.4		
7	seed	12345	-0.375		
8			-0.35		
9	32		-0.325		

Assume that we have 32 x values, the count in cell A9. Next we generate a set of ransom values drawn from a normal distribution with a mean of zero and standard deviation of 0.2, see below.

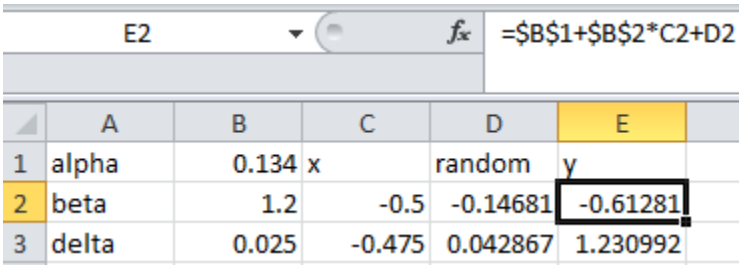


Finally, we generate a y column, see the formula in E2.

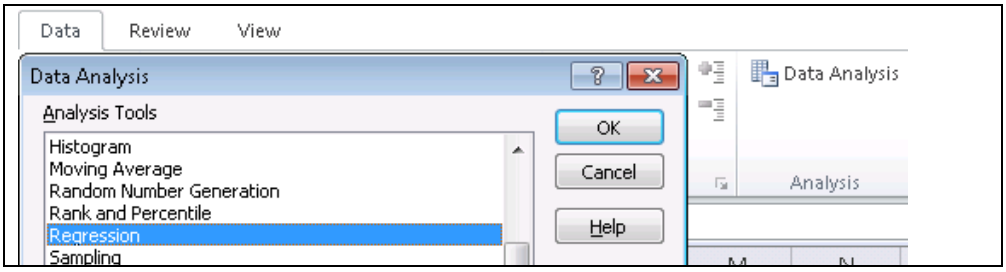
E2 =B1+B2*C2+D2

	A	B	C	D	E
1	alpha	0.134	x	random	y
2	beta	1.2	-0.5	-0.14681	-0.61281
3	delta	0.025	-0.475	0.042867	

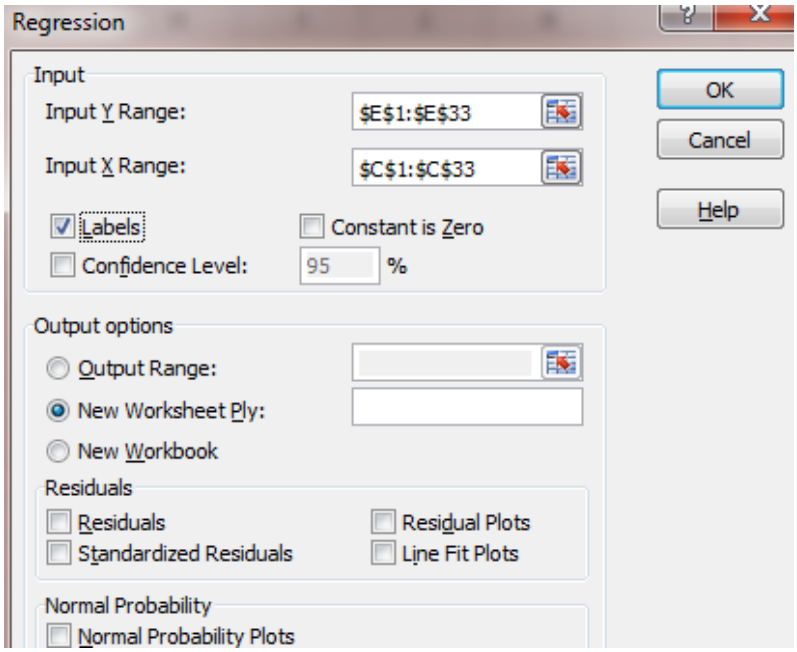
Before copy and paste, we have to fix the alpha and beta, see below.



We have the following steps to run a linear aggression.
Step 1: Highlight data set, see below. (note that the labels of y and x are included)
Step 2: Click “Data”, then “Data Analysis”, then choose “Regression”



Step 3: choose y and x data ranges, see below.



If you choose label in Step 1, then you have to click "Labels" in Step 3. You could choose to include the output within the same spreadsheet (output options). Below is the final result. The beta will be 0.96 while the alpha is 0.078. Since the T-value for Beta estimation is 8.2, we could claim that the estimate beta of 9.96 is statistically significant different from zero. This is true for the alpha estimate which has a T-value of 2.58.

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple R	0.831543							
5	R Square	0.691464							
6	Adjusted R Square	0.681179							
7	Standard Error	0.152952							
8	Observations	32							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*gnificance F*			
12	Regression	1	1.572888	1.572888	67.23339	3.75E-09			
13	Residual	30	0.701833	0.023394					
14	Total	31	2.274722						
15									
16		*Coefficients*	*ndard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
17	Intercept	0.077841	0.030079	2.587902	0.014743	0.016412	0.13927	0.016412	0.139270074
18	x	0.960476	0.117137	8.199597	3.75E-09	0.721251	1.199702	0.721251	1.199702221

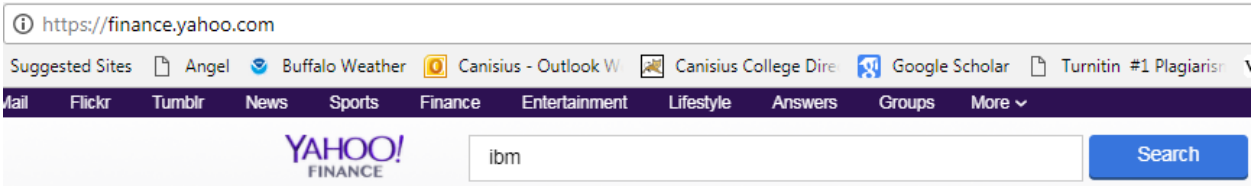
Alternatively, we could look at the P-value. Since our critical value is 5% and both p-values for beta (0.0147) and alpha (3.75E-9) are less than 5%, we conclude that the intercept (alpha) and slope (beta) are statistically different from zero.

8.5 HOW TO DOWNLOAD DATA

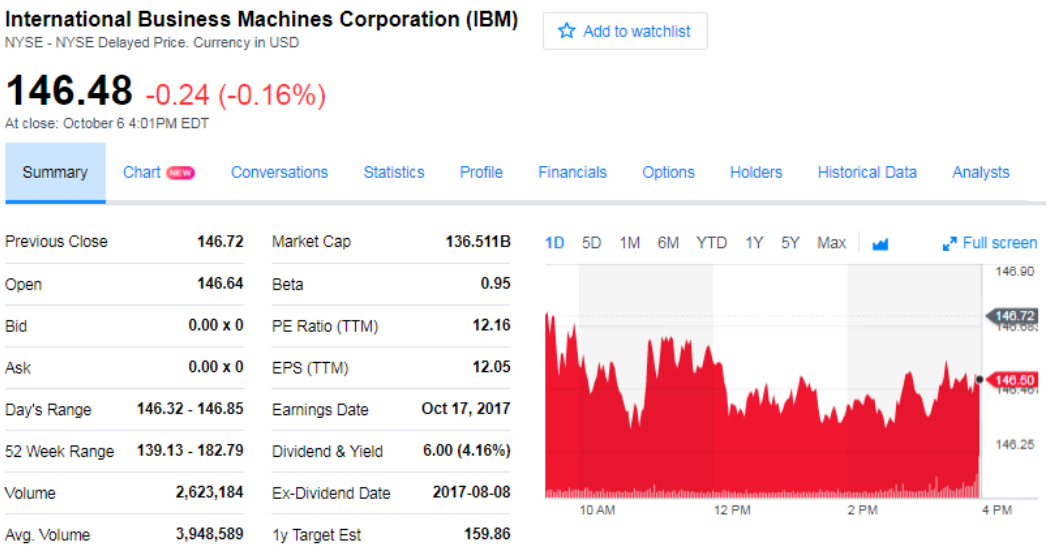
From Chapter 5, we know that monthly, weekly and daily price information is available from Yahoo Finance. We can convert price into return. To manually download data from Yahoo!finace, we have the following procedure:

Step 1: go to http://finance.yahoo.com

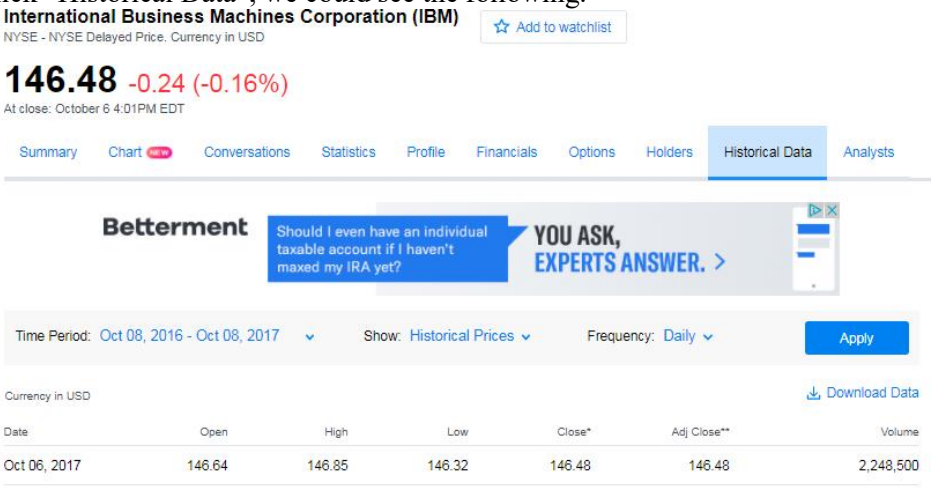
Step 2: enter a ticker, such as IBM, see the following image.



After click "Search", we would see the following image (on 10/7/2017).



Step 3: Click "Historical Data", we could see the following.



Step 5: Choose "Time Period" to fix the beginning and ending dates (remember to click "Done).
Step 6: Choose appropriate "Frequency" (default is Daily).

Step 7: After click “Apply”, Download Data.

See the first several observations downloaded on 10/8/2018.

	A	B	C	D	E	F	G
1	Date	Open	High	Low	Close	Adj Close	Volume
2	1/2/1962	7.713333	7.713333	7.626667	7.626667	2.192523	387200
3	1/3/1962	7.626667	7.693333	7.626667	7.693333	2.211689	288000
4	1/4/1962	7.693333	7.693333	7.613333	7.616667	2.189648	256000
5	1/5/1962	7.606667	7.606667	7.453333	7.466667	2.146526	363200
6	1/8/1962	7.46	7.46	7.266667	7.326667	2.10628	544000
7	1/9/1962	7.36	7.506667	7.36	7.413333	2.131195	491200

To get a risk-free rate, we have several sources. The easiest one is from Prof. French’s Data Library. The risk-free rates are included in his daily and monthly factor time series.

Step 1: go to Prof. French’s Data Library website at
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
Step 2: Click "Fama/French Factors" for monthly (daily) factors
Step 3: Save it to your local computer and unzip it.

The first several lines (Fama-French monthly factors) are given below. Note that every values (returns) shown below should be divided by 100. Whenever, we offer a data case by using Fama-French data, there always some students use those values without dividing them by 100.

```
This file was created by CMPT_ME_BEME_RETS using the 201006 CRSP
database.
The 1-month TBill return is from Ibbotson and Associates, Inc.

         Mkt-RF      SMB      HML       RF
192607     2.62    -2.16    -2.92     0.22
192608     2.56    -1.49     4.88     0.25
192609     0.36    -1.38    -0.01     0.23
192610    -3.43     0.04     0.71     0.32
```

There is a much easy and efficient way to retrieve the risk-free rate is to issue the command of `.getdata` (note there is a period before `getdata`), see the image below.

```
> .getdata
function(){
"
 *-------------------------------------------------------------------------*
 * .getdata                                                                *
 *-------------------------------------------------------------------------*
 *  Economics                 Finance                Accounting            *
 *-------------------------------------------------------------------------*
 * .show_usGDPannual         .showff3Monthly        .getBSannual          *
 * .show_usGDPquarterly      .showff3Daily          .getBSquarterly       *
 * .show_usUnemployRate      .showffc4Monthly       .getISannual          *
 * .show_usDebt_annual       .showffc4Daily         .getISquarterly       *
 * .show_usCPI_annual        .showff5Monthly        .getCFannual          *
 * .show_usCPI_monthly       .showff5Daily          .getCFquarterly       *
 * .show_euroDollar_1m       .showAaaYieldDaily                           *
 * .show_dollarIndexMonthly  .showBaaYieldMonthly   .getCIKgivenTicker     *
 * .show_goldPrice           .showBaaYieldDaily                           *
 * .show_fedFundRate         .showTradingDaysDaily                        *
 * .show_vix                 .showTradingDaysMonthly                      *
 *                           .show_ct1day                                 *
 *                           .show_cq1day            .copy2clipboard      *
 *                           .show_ct3month          .saveYan             *
 *                           .show_cq3month          .saveFinStatement    *
 *                           .show_sp500monthly      .inputCSV            *
 *                           .show_sp500daily        .mergeTwo            *
 *                           .show_bondSpread                             *
 *                           .getDailyPrice                               *
 *-------------------------------------------------------------------------*
 * >.show_usGDPannual        # find the usage of this function            *
 * >.getdata                 # back to this menu                          *
 * >.fm                      # back to the main menu                      *
 *-------------------------------------------------------------------------*
```

For the Fama-French there are 4 data sets, see above. For the usage of each function, just type its name, such as .ff3Monthly, see below.

```
> .showff3Monthly
function(n=2){
"Objective: show Fama-French 3 monthly factors
        n  : n > 0 for the first n obs (default is 2)
             n < 0 for the last  n obs
             n = 0 for all observations

   source    : http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
   frequency : monthly

 Example 1:> .showff3Monthly()
                      DATE MKT_RF    SMB     HML     RF
              1 1926-07-01 0.0296 -0.023 -0.0287 0.0022
              2 1926-08-01 0.0264 -0.014  0.0419 0.0025

 Example 2: > .showff3Monthly(-5)
                    DATE  MKT_RF     SMB     HML    RF
          1075 2016-01-01 -0.0577 -0.0366  0.0313 1e-04
          1076 2016-02-01 -0.0007  0.0073 -0.0003 2e-04
          1077 2016-03-01  0.0696  0.0087  0.0129 2e-04
          1078 2016-04-01  0.0091  0.0084  0.0312 1e-04
          1079 2016-05-01  0.0178 -0.0028 -0.0185 1e-04

 Example 3: > x=.showff3Monthly(0)
            > .saveYan(x,'c:/temp/ff.csv')
               [1] 'Your saved file is ==>c:/temp/ff.csv'
```

Another advantage of using the method is the all the returns data are divided by 100 if compared with the original data downloaded from Prof. French's Data Library.

SUMMARY

In this chapter, we have learnt the CAPM (Capital Asset Pricing Model) which is a single-factor linear model. In finance, CAPM is one of the well know model. The CAPM theory was developed in the 1960s. It is quite popular among academics and practitioners due to its simplicity and usefulness. CAPM can be viewed as a single-factor linear model the simplest asset pricing model. In the previous chapters, we have learnt that we could use standard deviation of returns to represent the total risk. Based on CAPM, the beta (β) represents the market risk of an underlying security.

There are some shortcomings of CAPM since it is quite basic. In the next chapter, Chapter 8: Multiple-factor models: FF3, FFC4 and FF5 models, we would learn several import models such as Fama-French 3-factor model, Fama-French–Carhart 4-factor model and Fama-French 5-factor model. More importantly, we learnt how to construct our own multi-factor linear models and how to evaluate whether our extra factor is significantly or more and how to evaluate the model overall performance.

REFERENCES

Damodaran, Aswath, Estimating beta

http://people.stern.nyu.edu/adamodar/pdfiles/eqnotes/discrate2.pdf
Fernandez, Pablo, CAPM: an absurd model,
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2505597
Qin, Jie, 2002, Human-Capital-Adjusted Capital Asset Pricing Model, Japanese Economy Review,
http://onlinelibrary.wiley.com/doi/10.1111/1468-5876.00222/abstract
YouTube: CAPM Capital Asset Pricing Model in 4 Easy Steps - What is Capital Asset Pricing Model Explained, https://www.youtube.com/watch?v=LWsEJYPSw0k
Calculating stock beta using Excel (4m28s)
https://www.youtube.com/watch?v=zlClflcSrM8
What is Beta? - MoneyWeek Investment Tutorials (5m29s)
https://www.youtube.com/watch?v=etlv7qTQUSY
Beta Calculation on Excel (9m48s)
https://www.youtube.com/watch?v=7LiK-qbmPsw
How to Calculate Beta with Excel, Calculation of Beta (18m43s)
https://www.youtube.com/watch?v=LRyFn_T94IU

EXERCISES

8.1 What is the Fama-French 3 factor model?
8.2 What is the definition of SMB?
8.3 What is the definition of HML?
8.4 What is the momentum strategy?
8.5 From where we could download Fama-French monthly factors?
8.6 From where we could download Fama-French daily factors?
8.7 Is the SMB factor a portfolio?
8.8 If hold SMB portfolio from January 1, 2000 to December 31, 2010, what is the total return? Estimate total returns based on both daily and monthly SMB.
8.9 Why you get different results in 8.8?
8.10 Trying the same thing (same as questions of 8.8 and 8.9) for HML.
8.11 Trying the same thing (same as questions of 8.8 and 8.9) form mkt_rf.
8.12 Estimating IBM's beta based on the Fama-French 3-factor model by using the monthly data from Yahoo finance and from Prof. French' data library.
8.13 What is the difference between Sharpe ratio and Treynor's ratio?
8.14 Can we use just one period to estimate Jensen's alpha?
8.15 What is the Sharpe ratio for IBM based on the last 5-year's monthly data?
8.16 Based on the Sharpe ratio and the past 3-year daily data, which stock you should choose, IBM or Google?
8.17 Will your decision, based on the previous question, changes if applying Treynor's ratio instead of Sharpe ratio?
8.18 Constructing a portfolio of 5 stocks, such as IBM,C, GE, GOOG and WMT, and estimating their monthly portfolio returns from 2001 to 2010. What is their portfolio's beta?
8.19 What is the alpha of the above portfolio is we use the 1st 5-year to estimate beta of the portfolio?

8.20 Writing a program to estimate 52-week's high and low for DELL.

8.21 Downloading F-F_Research_Data_Factor.zip from French Data library. Unzip it and the name of a text file is F-F_Research_Data_Factors.txt. Write an R program to process annual Fama-French factors and save your final output to an R data set