# Portfolio Investment Using Time Series Data

Group Member:

Menglin Liu

Sung Ik Kim

Wei Jiang

Lu Zhao

# Analysis of Time Series Data

- **Outline**
  - Introduction
  - Pfizer Inc. (PFE)
  - British Petroleum (BP)
  - Oracle Corporation (ORCL)
  - JP Morgan Chase & Co. (JPM)
  - Portfolio performance
  - Conclusion

# Introduction

- Goal
  - Analyze and model four stocks from four different industries
  - Form a tiny portfolio using these four stocks
  - Measure the performance of this portfolio based on ten days return
- Data Summary
  - Source: Yahoo Finance
  - Time period: 01-04-2010 to 03-31-2014
  - Size: 1067

# Pfizer, Inc.

Pfizer, Inc. is an American multinational pharmaceutical corporation headquartered in New York City, and with its research headquarters in Groton,Connecticut, United States. It is one of the world's largest pharmaceutical companies by revenues.

Pfizer develops and produces medicines and vaccines for a wide range of medical disciplines, including immunology, oncology, cardiology, diabetology/endocrinology, and neurology.


| | |
|---|---|
| **Type** | Public |
| **Traded as** | NYSE: PFE (component of Dow 30 and S&P 500) |
| **Industry** | Pharmaceutical |
| **Founded** | 1849 |
| **Headquarters** | New York, NY |
| **Area served** | Worldwide |
| **Key people** | Ian Read (President & CEO) |

# Pfizer, Inc.

| | |
|---|---|
| Revenue | $58.98 billion (2013) |
| Profit | $12.08 billion (2013) |
| Total assets | $172.101 billion (2013) |
| Total equity | $81.26 billion (2012) |
| Employees | 91,500 (2012) |

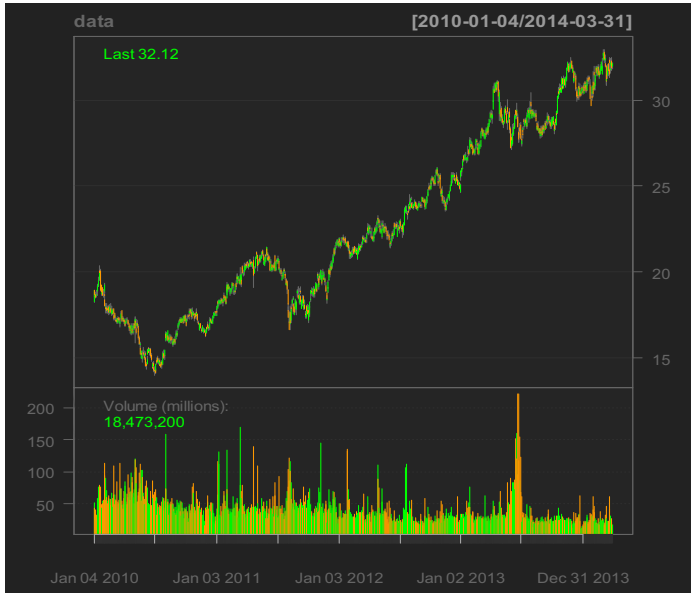
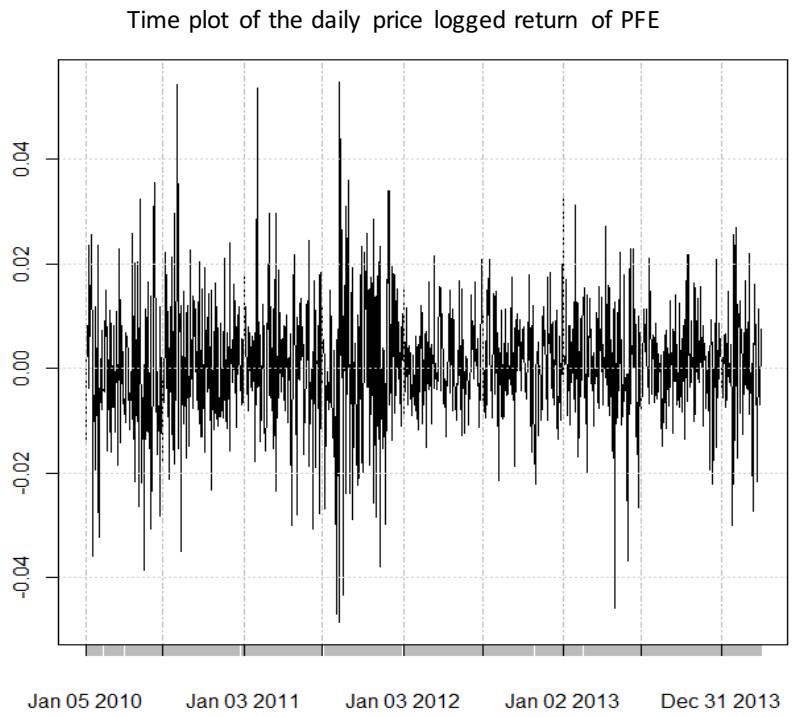


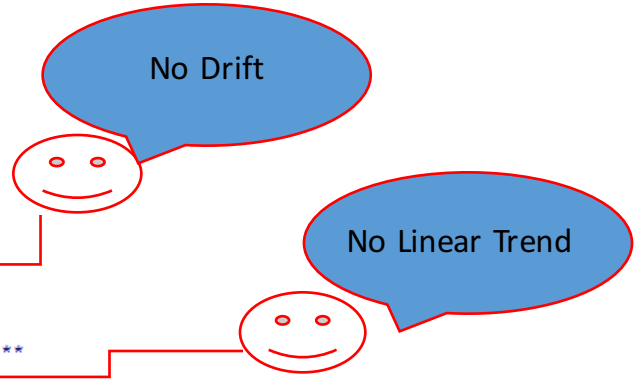

**Fig1. Line chart for the close price of Pfizer(PFE) from 2010-01-04 to 2014-03-31**

# Calculate the Logged Return of the Stock Price

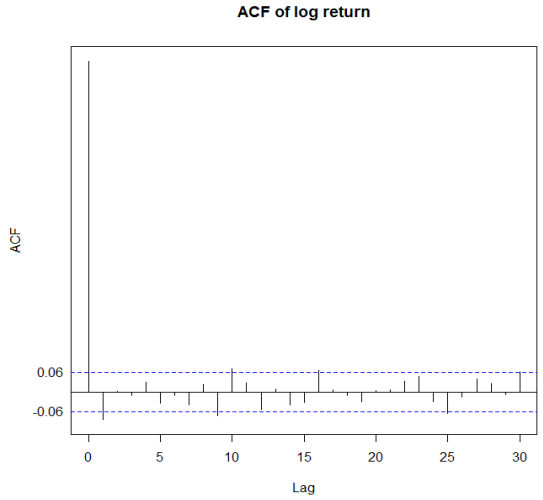


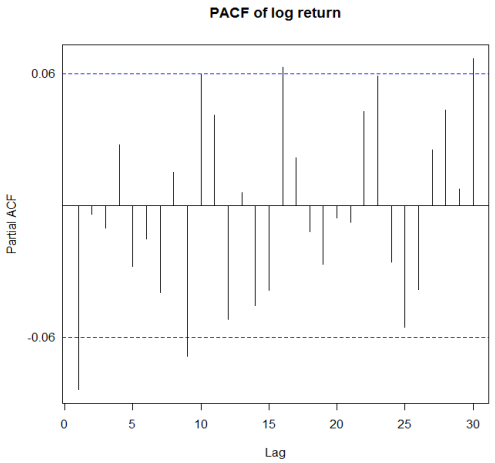
# Test the Trend and Drift

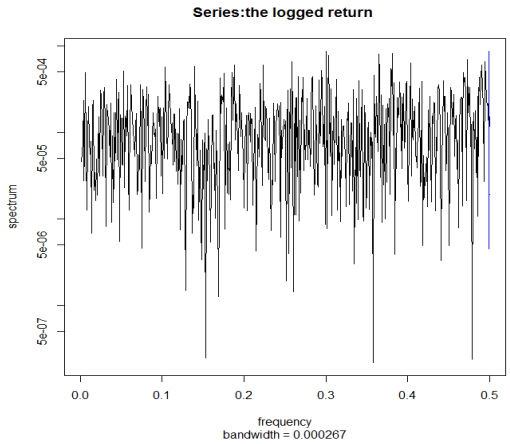
```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
      Min        1Q    Median        3Q       Max
-0.048723 -0.006723 -0.000409  0.006601  0.056771

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -8.750e-05  7.607e-04  -0.115    0.908
z.lag.1     -1.101e+00  4.567e-02 -24.098   <2e-16 ***
tt           1.212e-06  1.226e-06   0.989    0.323
z.diff.lag   1.006e-02  3.085e-02   0.326    0.744
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01194 on 1041 degrees of freedom
Multiple R-squared:  0.5466,    Adjusted R-squared:  0.5453
F-statistic: 418.3 on 3 and 1041 DF,  p-value: < 2.2e-16


Value of test-statistic is: -24.0978 193.5706 290.354

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.96 -3.41 -3.12
phi2  6.09  4.68  4.03
phi3  8.27  6.25  5.34
```



# ARMA-GARCH Model



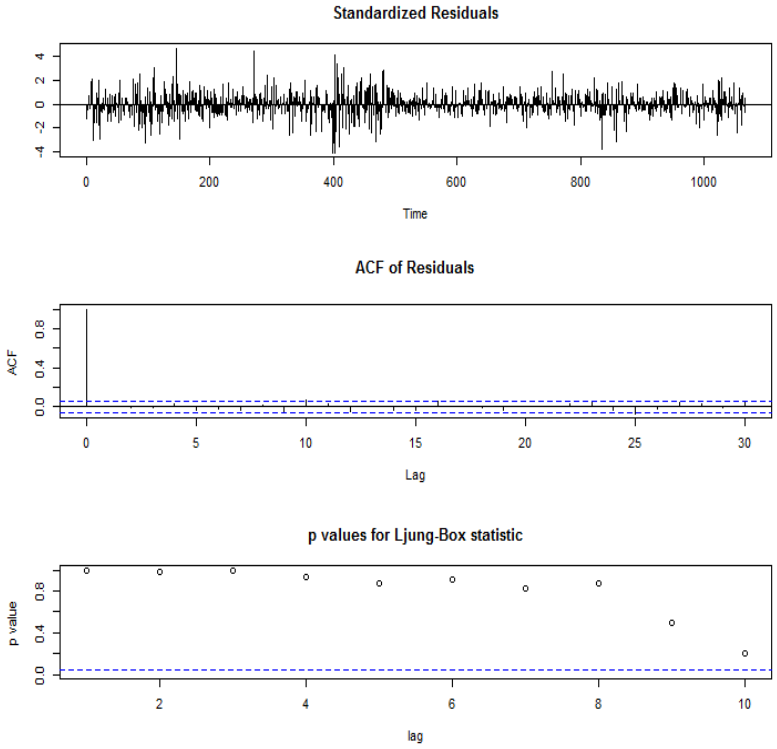



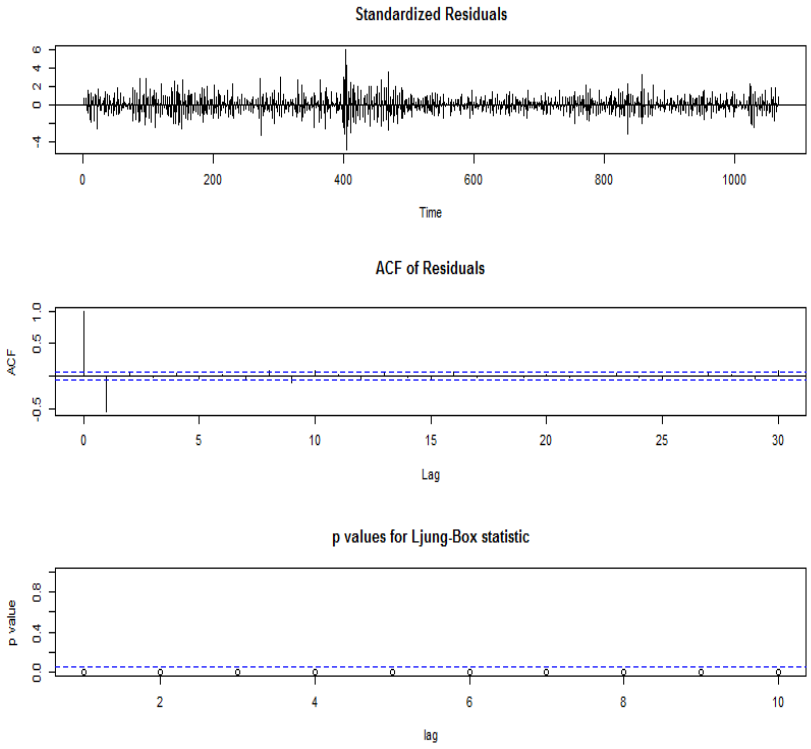
# ARIMA Model Comparison

| Model | AIC | AICc | BIC |
|---|---|---|---|
| ARIMA(1,0,0) | -6404.66 | -6404.64 | -6389.74 |
| ARIMA(1,1,0) | -5935.93 | -5935.91 | -5925.98 |
| ARIMA(1,1,1) | -6390.51 | -6390.49 | -6375.6 |
| ARIMA(0,1,1) | -6385.17 | -6385.16 | -6375.23 |
| ARIMA(0,0,1) | -6404.63 | -6404.62 | -6389.70 |
| ARIMA(1,0,1) | -6402.67 | -6402.63 | -6382.78 |
| ARIMA(0,1,0) | -5571.15 | -5571.14 | -5566.18 |

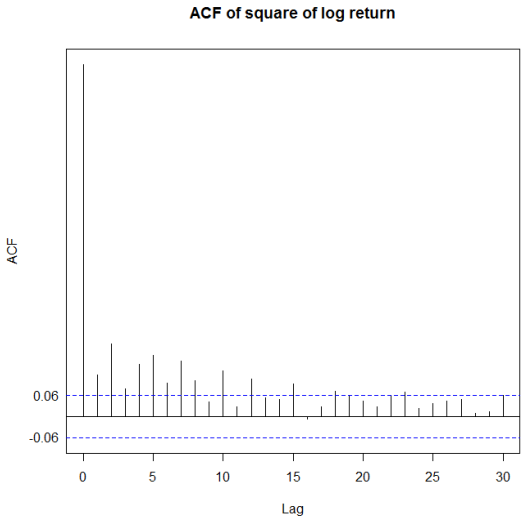
AR(1) Model

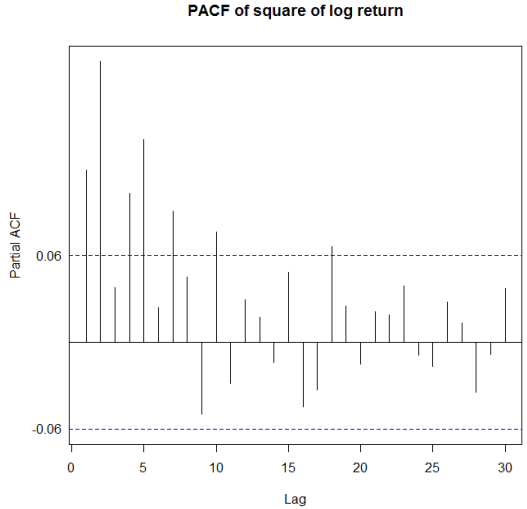



MA(1)Model

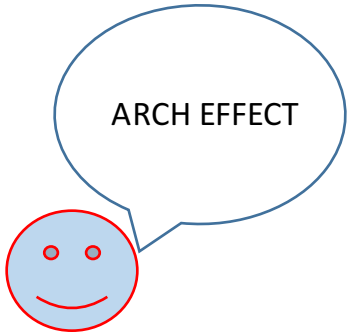


# ACF and PACF of the Squared Residuals




```
> r1=r-mean(r)
> Box.test(r1^2,lag=10,type='Ljung')

        Box-Ljung test

data:  r1^2
X-squared = 221.6592, df = 10, p-value < 2.2e-16
```



# ARMA-GARCH Model Comparison

| Model | AIC | Mean squared Prediction error |
|---|---|---|
| AR(1)-Garch(1,1) with Normal conditional distribution | -6.107765 | 0.25553708 |
| AR(1)-Garch(1,1) with T conditional distribution | -6.151196 | 0.2550218 |
| AR(1)-Garch(1,1) with generalized error distribution | - 6.144887 | 0.254877 |
| AR(1)-Garch(1,1) with other distributions | >-6.10598 | >0.2553843 |

# Fitted Model

```
> m2=garchFit(~arma(1,0) +garch(1,1),data=r,cond.dist="std",trace=F)
> summary(m2)

Title:
 GARCH Modelling

Call:
 garchFit(formula = ~arma(1, 0) + garch(1, 1), data = r, cond.dist = "std",
    trace = F)

Mean and Variance Equation:
 data ~ arma(1, 0) + garch(1, 1)
<environment: 0x000000000c90e708>
 [data = r]

Conditional Distribution:
 std

Coefficient(s):
        mu          ar1        omega       alpha1        beta1        shape
 7.6204e-04  -8.4115e-02   3.0906e-06   5.6240e-02   9.2249e-01   6.3633e+00

Std. Errors:
 based on Hessian

Error Analysis:
         Estimate  Std. Error  t value Pr(>|t|)
mu      7.620e-04   3.111e-04    2.450  0.01430 *
ar1    -8.412e-02   3.032e-02   -2.775  0.00553 **
omega   3.091e-06   1.574e-06    1.963  0.04963 *
alpha1  5.624e-02   1.661e-02    3.386  0.00071 ***
beta1   9.225e-01   2.342e-02   39.384  < 2e-16 ***
shape   6.363e+00   1.216e+00    5.232 1.68e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 3284.588    normalized:  3.081227
```

The fitted AR(1)-GARCH(1,1) model can be expressed as:

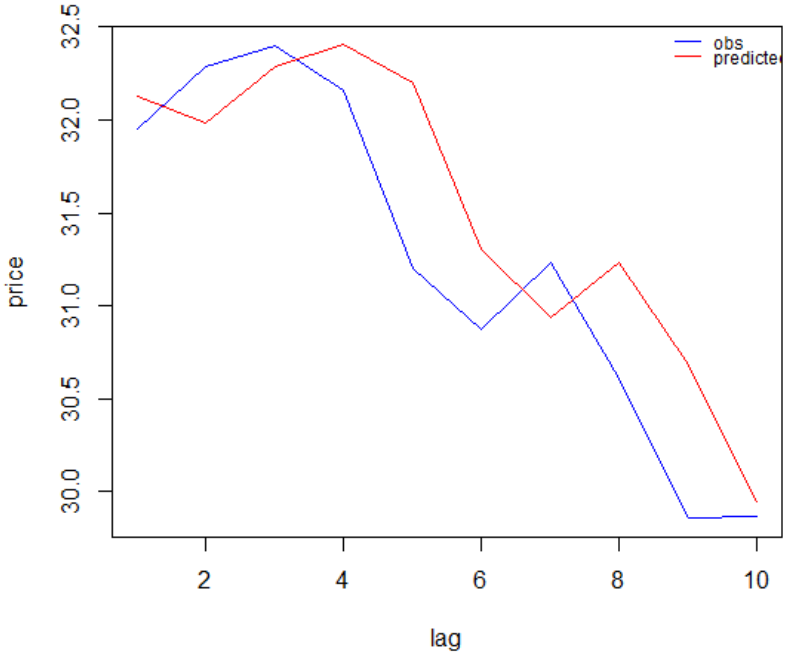
$$X_t = -8.412 \times 10^{-2} X_{t-1} + r_t$$

$$r_t = 7.620 \times 10^{-4} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{6.36}$$

$$\sigma_t^2 = 3.091 \times 10^{-6} + 5.624 \times 10^{-2} a_t^2 + 9.225 \times 10^{-1} \sigma_{t-1}^2$$

# Prediction for Ten Days

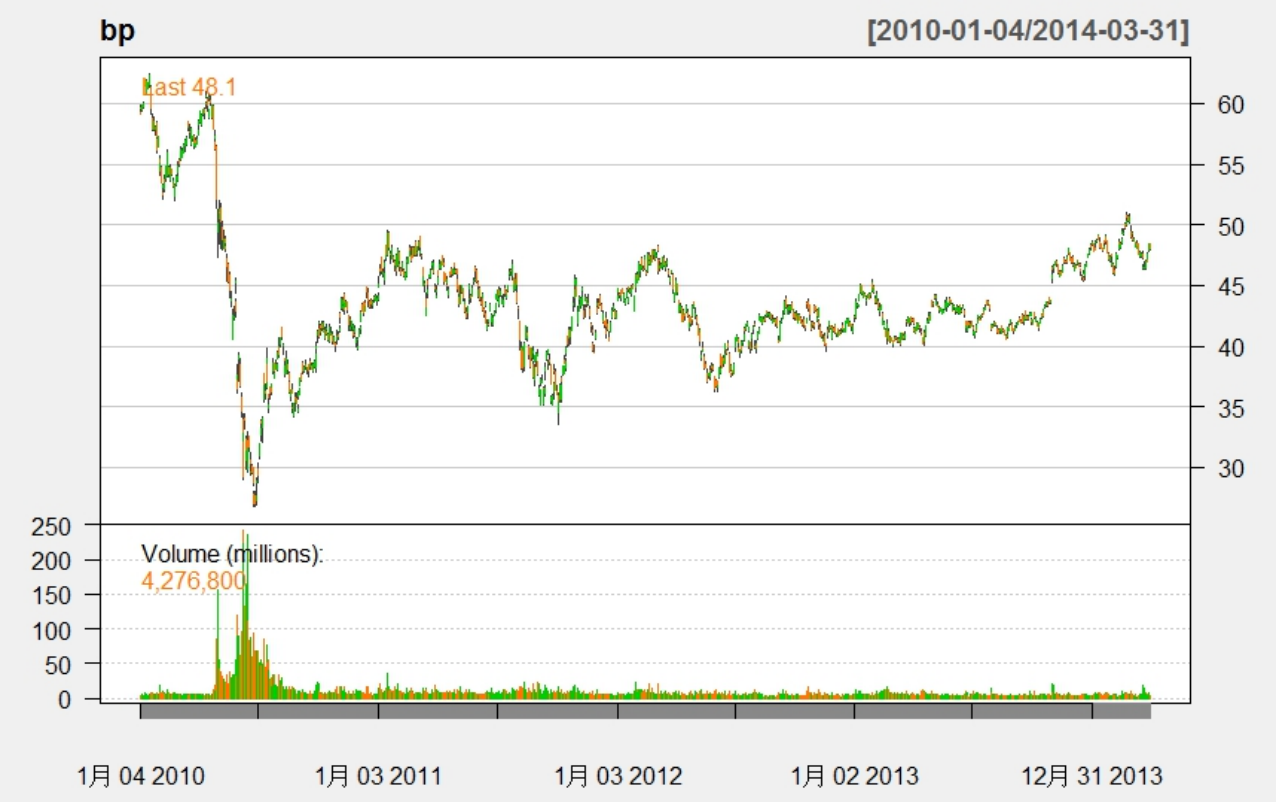
| Date | PFE.Adjusted | predict.value |
|---|---|---|
| 4/1/2014 | 31.95 | 32.12421352 |
| 4/2/2014 | 32.29 | 31.98606034 |
| 4/3/2014 | 32.4 | 32.28838973 |
| 4/4/2014 | 32.16 | 32.41125681 |
| 4/7/2014 | 31.2 | 32.20284297 |
| 4/8/2014 | 30.87 | 31.30678378 |
| 4/9/2014 | 31.23 | 30.93500168 |
| 4/10/2014 | 30.61 | 31.23401354 |
| 4/11/2014 | 29.86 | 30.68006209 |
| 4/14/2014 | 29.87 | 29.9506736 |

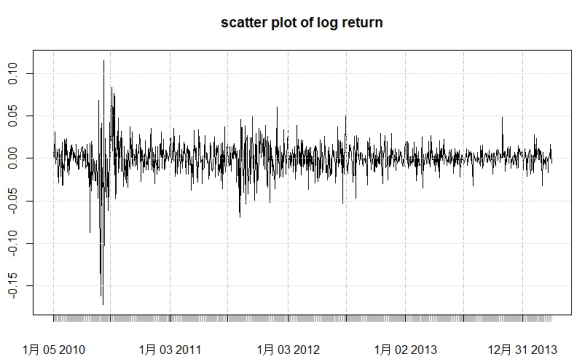


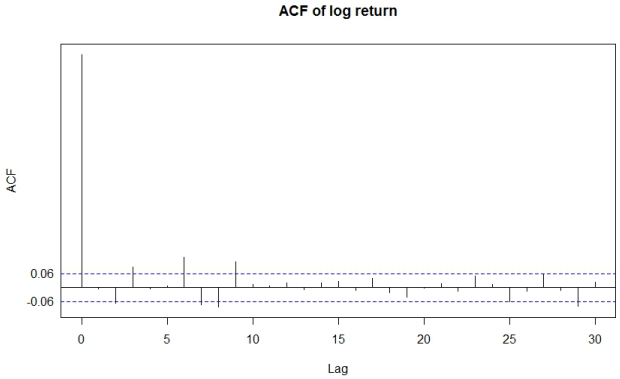
# British Petroleum

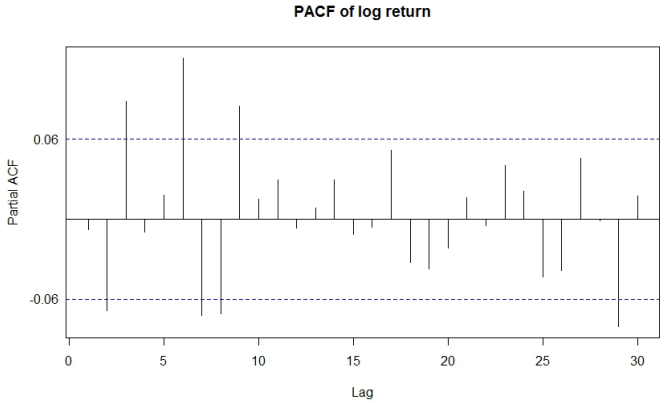
**Bp plc** is a British multinational oil and gas company headquartered in London, England. It is one of the six oil and gas "supermajors".
BP is vertically integrated and operates in all areas of the oil and gas industry, including exploration and production, refining, distribution and marketing, petrochemicals, power generation and trading.
It also has renewable energy activities in biofuels and wind power.

# Data


bp
[2010-01-04/2014-03-31]
last 48.1
60
55
50
45
40
35
30
250
200
150
100
50
0
Volume (millions):
4,276,800
1月 04 2010
1月 03 2011
1月 03 2012
1月 02 2013
12月 31 2013


scatter plot of log return
0.10
0.05
0.00
-0.05
-0.10
-0.15
1月 05 2010
1月 03 2011
1月 03 2012
1月 02 2013
12月 31 2013


ACF of log return
ACF
0.06
-0.06
0
5
10
15
20
25
30
Lag


PACF of log return
Partial ACF
0.06
-0.06
0
5
10
15
20
25
30
Lag

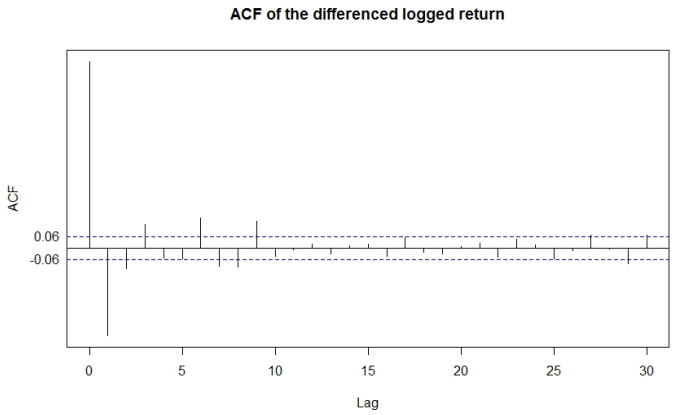
# Check for Trend

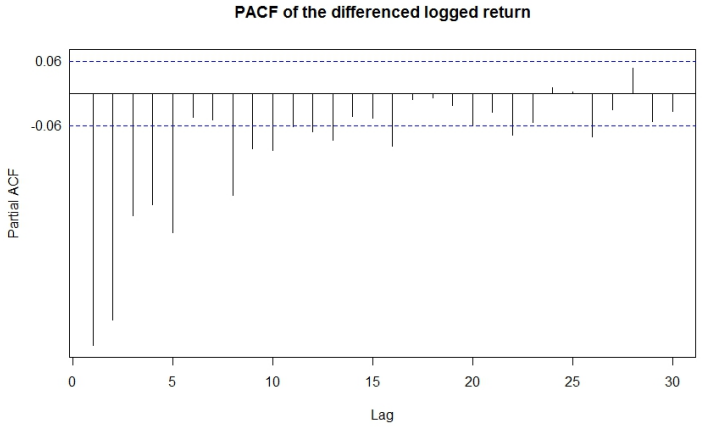
```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
      Min        1Q    Median        3Q       Max
-0.167446 -0.007640  0.000440  0.008375  0.112686
Coefficients:
              Estimate   Std. Error  t value  Pr(>|t|)
(Intercept) -1.025e-03  1.179e-03   -0.869   0.38518
z.lag.1     -9.863e-01  5.503e-02  -17.922   < 2e-16 ***
tt           1.958e-06  1.899e-06    1.031   0.30286
z.diff.lag1 -1.654e-02  4.381e-02   -0.378   0.70579
z.diff.lag2 -8.793e-02  3.085e-02   -2.850   0.00445 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01849 on 1040 degrees of freedom
Multiple R-squared:  0.5107,    Adjusted R-squared:  0.5088
F-statistic: 271.4 on 4 and 1040 DF,  p-value: < 2.2e-16
```

No drift

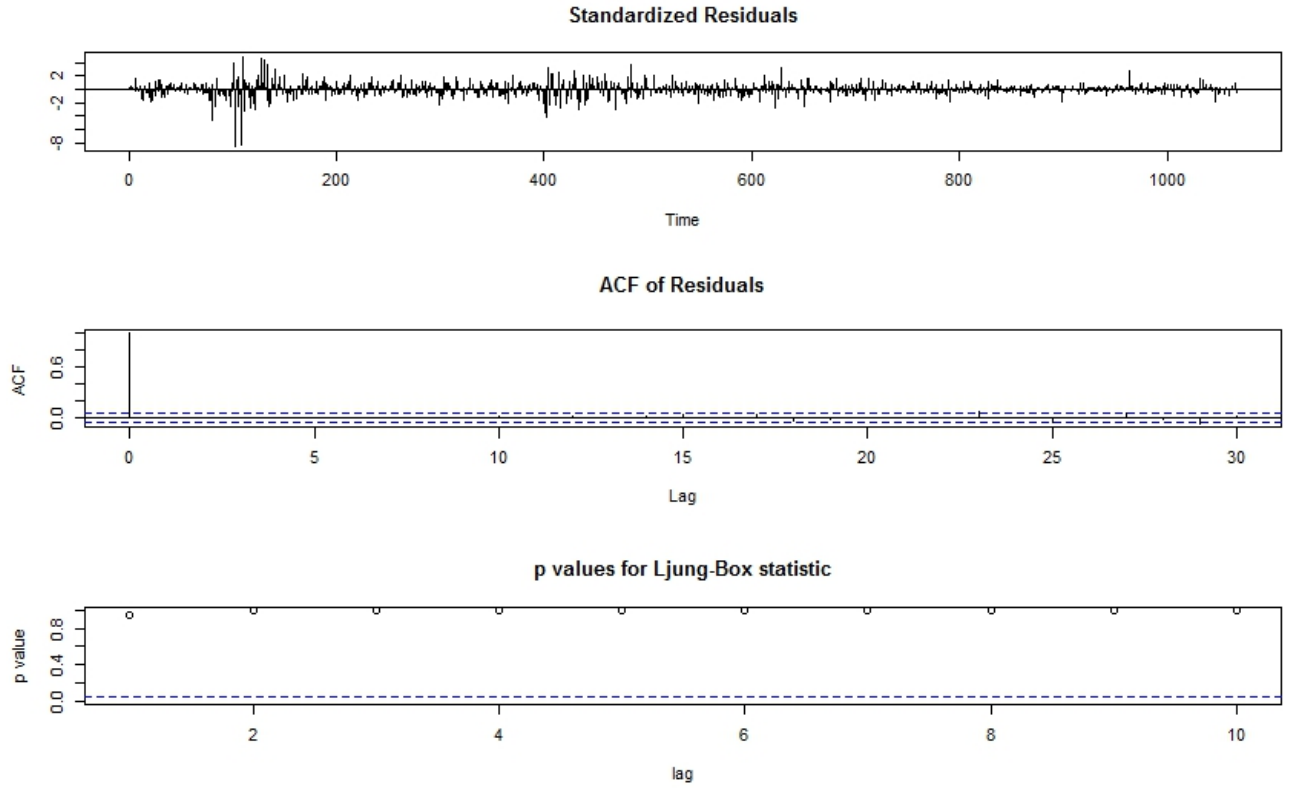
No linear trend

# Difference


ACF of the differenced logged return
ACF
0.06
-0.06
0
5
10
15
20
25
30
Lag


PACF of the differenced logged return
Partial ACF
0.06
-0.06
0
5
10
15
20
25
30
Lag

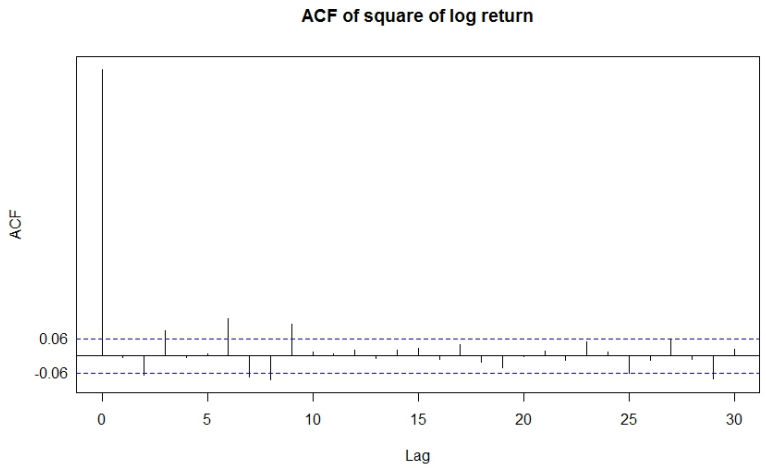
# ARIMA(0,1,9)

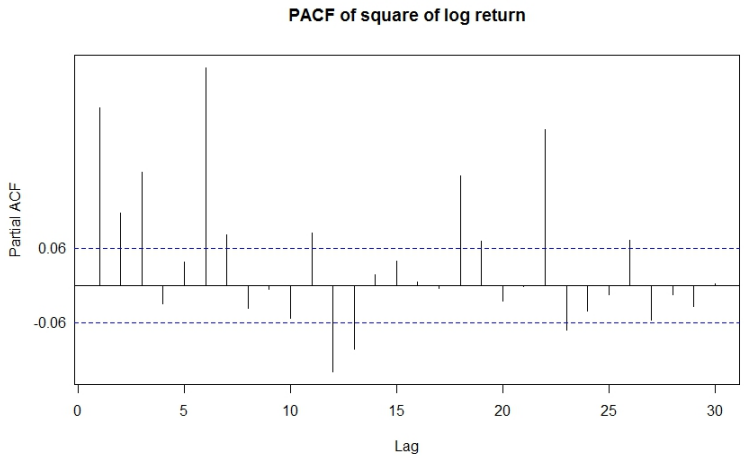

Standardized Residuals
-8
-2
2
0
200
400
600
800
1000
Time
ACF of Residuals
ACF
0.0
0.6
0
5
10
15
20
25
30
Lag
p values for Ljung-Box statistic
p value
0.0
0.4
0.8
2
4
6
8
10
lag

# Comparison of Different ARIMA Model

| Model | AIC | AICc | BIC |
|---|---|---|---|
| ARIMA(0,1,9) | -5508.05 | -5507.84 | -5458.33 |
| ARIMA(0,1,8) | -5496.7 | -5496.53 | -5451.95 |
| ARIMA(0,1,10) | -5506.28 | -5506.03 | -5451.59 |
| ARIMA(0,1,1) | -5474.3 | -5474.28 | -5464.35 |
| ARIMA(1,1,1) | -5472.61 | -5472.59 | -5457.7 |
| ARIMA(9,1,0) | -5506.45 | -5506.24 | -5456.73 |
| ARIMA(10,1,0) | -5504.7 | -5504.45 | -5450.01 |

# ACF and PACF of the Squared Log return




Box-Ljung test data:

bp.sq X-squared = 487.7167, df = 10, p-value < 2.2e-16

ARCH effect

# Comparison of Different GARCH Model

| Model | AIC | Mean squared Prediction error |
|---|---|---|
| MA(9)-Garch(1,1) with Normal conditional distribution | -5.529191 | 0.1952921 |
| MA(9)-Garch(1,1) with T conditional distribution | -5.597735 | 0.1946812 |
| Other possible Garch models | >-5.529191 | >0.1947 |

The fitted GARCH(1,1) model is

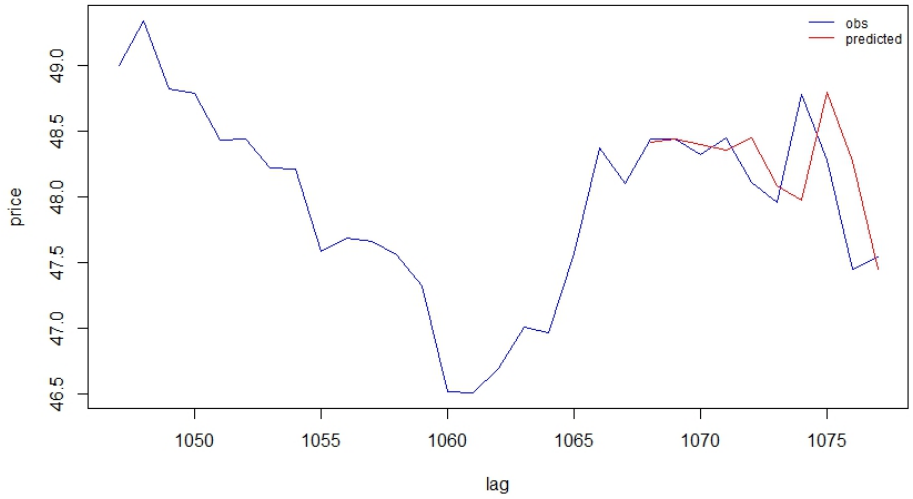
$$X_t = -9.602\times10^{-1}X_{t-1} + r_t$$

$$r_t = 8.732\times10^{-7} + a_t\,, \quad a_t = \sigma_t\epsilon_t, \quad \epsilon_t \sim t_{5.07}$$

$$\sigma_t^2 = 1.18\times10^{-6} + 7.403\times10^{-2}a_t^2 + .918\sigma_{t-1}^2$$

# Prediction

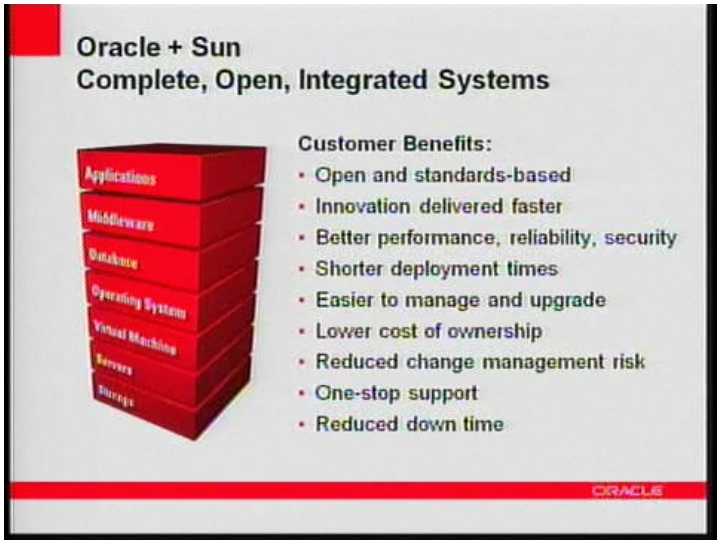
| Date | Adjusted Close | Predict.value |
|---|---|---|
| 2014/4/14 | 47.54 | 47.45004 |
| 2014/4/11 | 47.45 | 48.27468 |
| 2014/4/10 | 48.28 | 48.79573 |
| 2014/4/9 | 48.78 | 47.97152 |
| 2014/4/8 | 47.96 | 48.08075 |
| 2014/4/7 | 48.11 | 48.44901 |
| 2014/4/4 | 48.45 | 48.35662 |
| 2014/4/3 | 48.32 | 48.39810 |
| 2014/4/2 | 48.44 | 48.44193 |
| 2014/4/1 | 48.44 | 48.41262 |



# Oracle Corporation

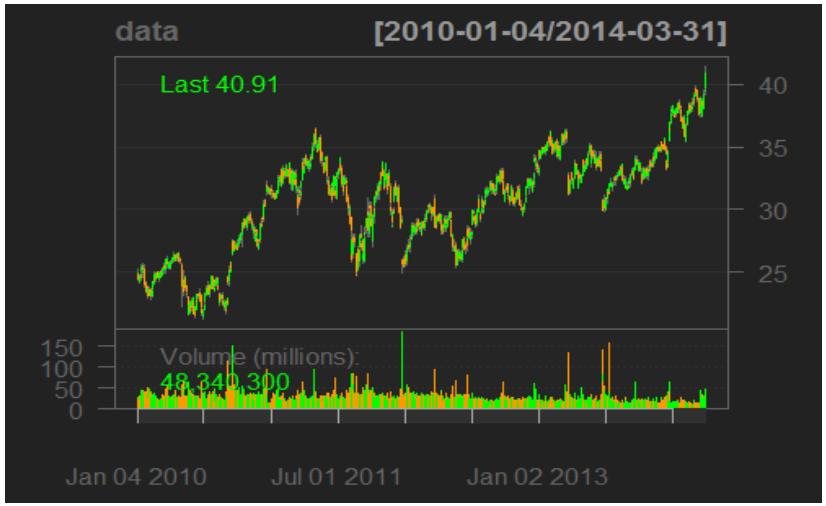
- It is an American multinational computer technology corporation headquartered in Redwood City, California, United States. The company specializes in developing and marketing compute hardware systems and enterprise software products – particularly its own brands of database management systems.

ORACLE®



# Oracle Corporation

| | |
|---|---|
| Revenue | $37.18 billion (2013) |
| Profit | $10.92 billion (2013) |
| Total assets | $81.81 billion (2013) |
| Total equity | $45.14 billion (2013) |
| Employees | 122,458 (2013) |

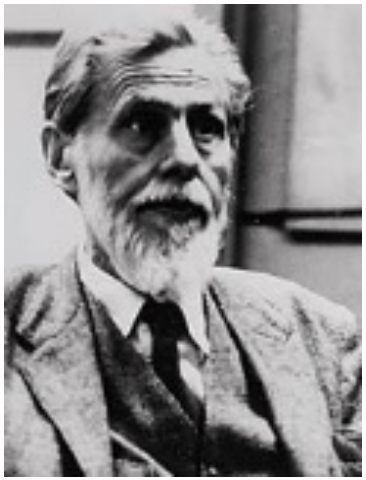


# Oracle Corporation

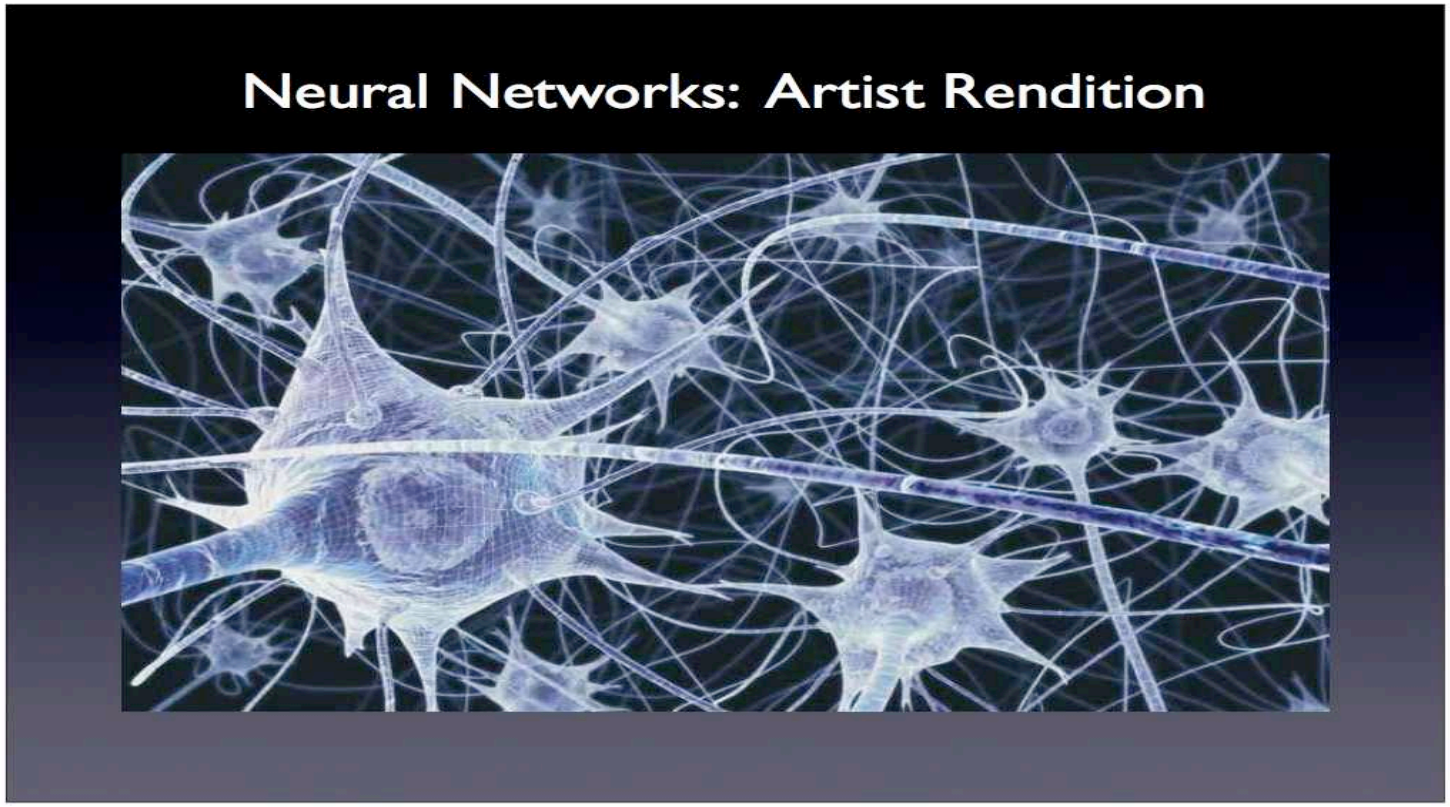
- Outline
  - Artificial Neural network(ANN)
  - State Space Model(SSM)
  - ARMA + GARCH Model

# Artificial Neural Network

- Warren Sturgis McCulloch (November 16, 1898 – September 24, 1969)
- Walter Harry Pitts, Jr (23 April 1923 – 14 May 1969)

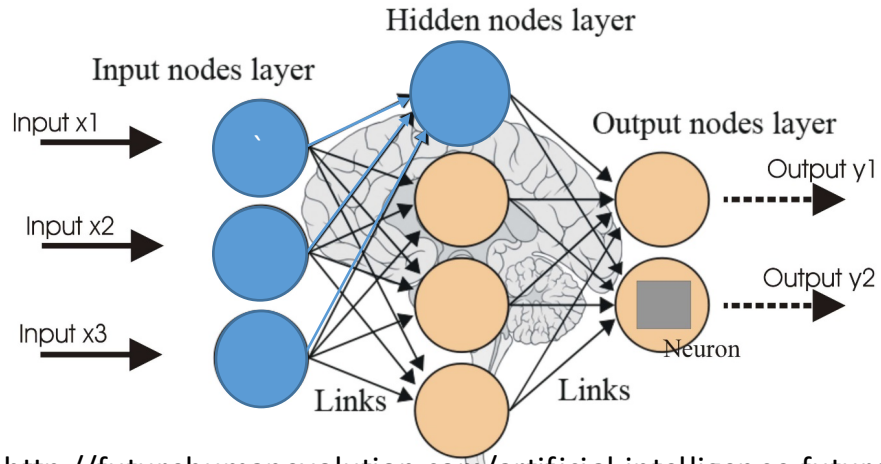


# Artificial Neural Network

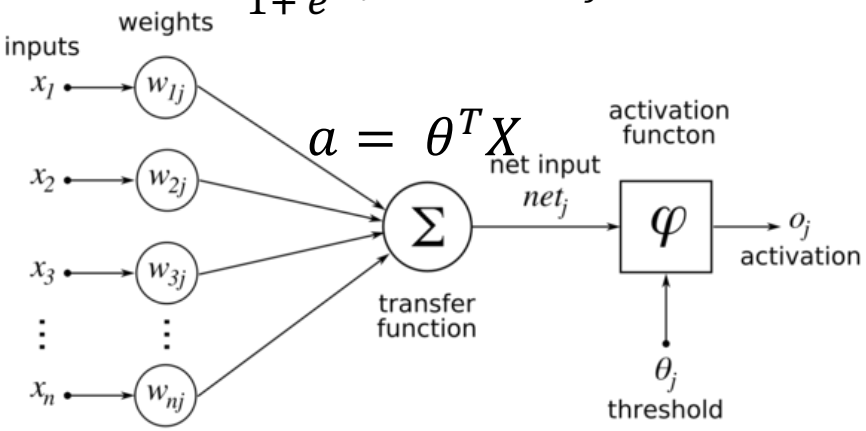


# Artificial Neural Network

- Single hidden layer back-propagation network(Vanilla neural network))



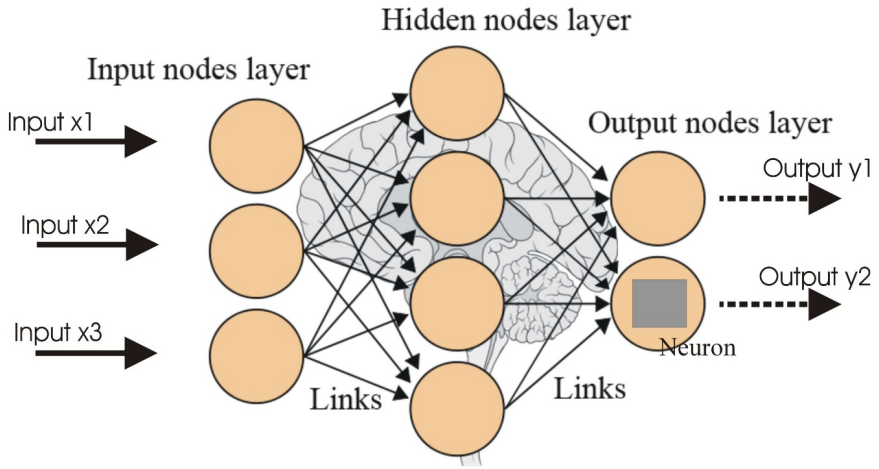

$$f(x) = \frac{1}{1+e^{-v}} \longleftarrow o_j = \varphi(a)$$




http://futurehumanevolution.com/artificial-intelligence-future-human-evolution/artificial-neural-networks
http://en.wikibooks.org/wiki/Artificial_Neural_Networks/Print_Version

# Artificial Neural Network

- To train ANN
  - Network structural
  - Activation function
  - Input features
- Problem
  - Computational
  - Over-fitting
  - Hard to Interpret



# Artificial Neural Network

- Feature extraction
  - The stock price of the nearest five lags

    $x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}$

  - The price difference between each lag and current stock price

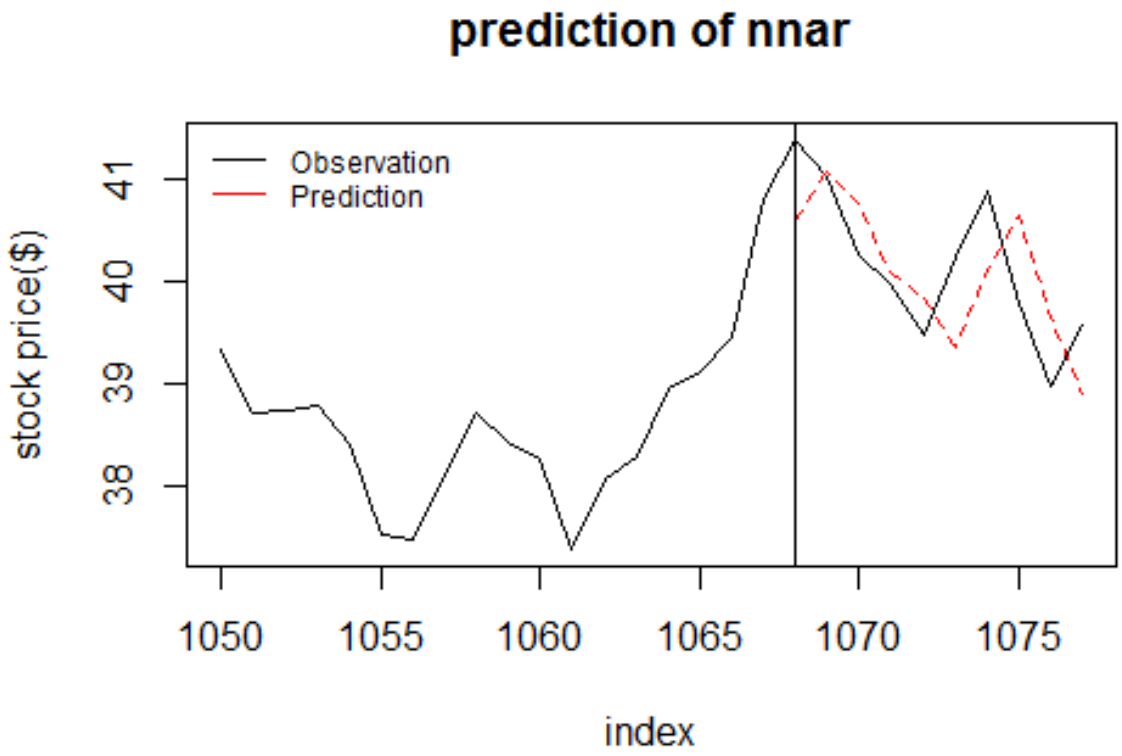
    $\Delta x_1, \Delta x_2, \Delta x_3, \Delta x_4, \Delta x_5$

- Activation function
  - Hidden layer: Sigmoid function
  - Output layer: Identity link
- Structural
  - Ten input nodes
  - One output nodes
  - One hidden layer

# Artificial Neural Network

- Fitted model
  - 10-10-1 structure
  - There are totally 121(11*10 + 11*1) weights.

# Artificial Neural Network



# State Space Model

- Specification of initial state distribution

$$a_0 \sim N(m_0, C_0)$$

  - Here $m_0$ is set as the mean of the series
  - $C_0$ is set to be zero since the series is stationary.

- Conditional Maximum Likelihood method was used to estimate system matrices.

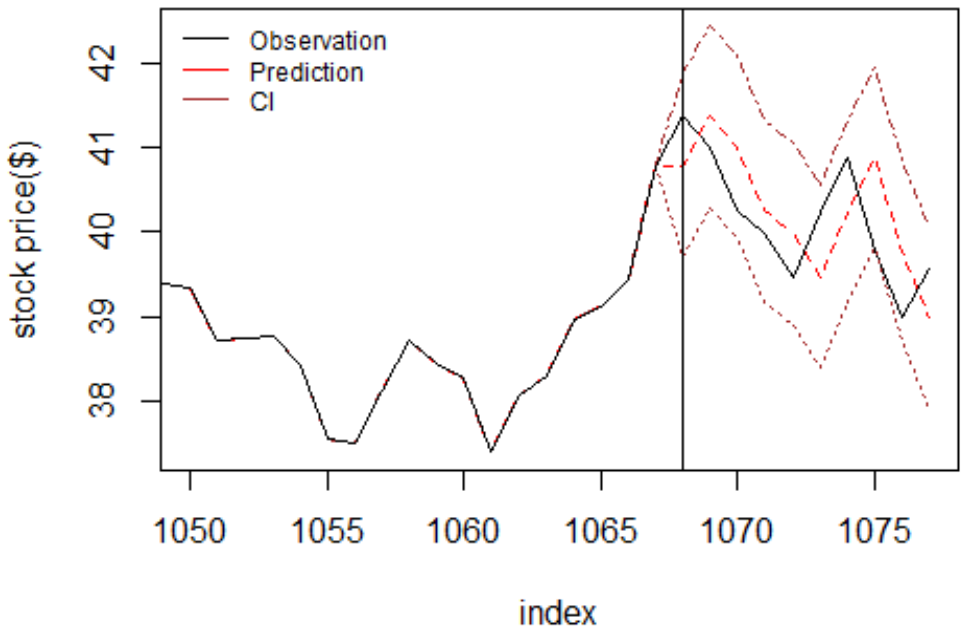
# State Space Model

- The fitted model

$$r_t = 6.15\times10^{-13} a_t + \varepsilon_t, \quad \varepsilon_t \sim N(0,1.234),$$
$$a_t = 0.8242 a_{t-1} + \eta_t, \ \eta_t \sim N(0,0.645)$$

- Different initial values seem do not affect the model too much. The prediction is almost same.

# State Space Model



# ARMA + GARCH Model

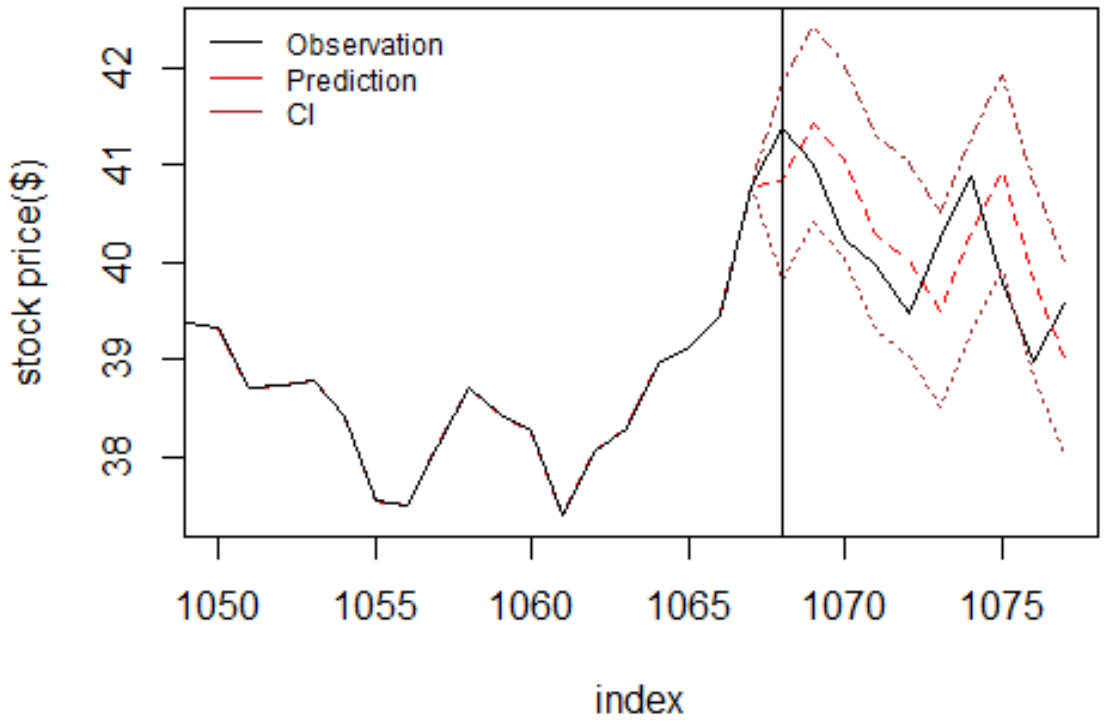
| Model | AIC | Mean squared Prediction error |
|---|---|---|
| Garch(1,1) with Normal conditional distribution | -5.445494 | 0.4675814 |
| Garch(1,1) with T conditional distribution | -5.562868 | 0.4703446 |
| Other possible Garch models | > - 5.4454 | > 0.47 |

The fitted GARCH(1,1) model is

$$r_t = 1.096 \times 10^{-3} + \ a_t \ , \quad a_t = \ \sigma_t \epsilon_t, \ \ \epsilon_t \sim t_{5.12}$$

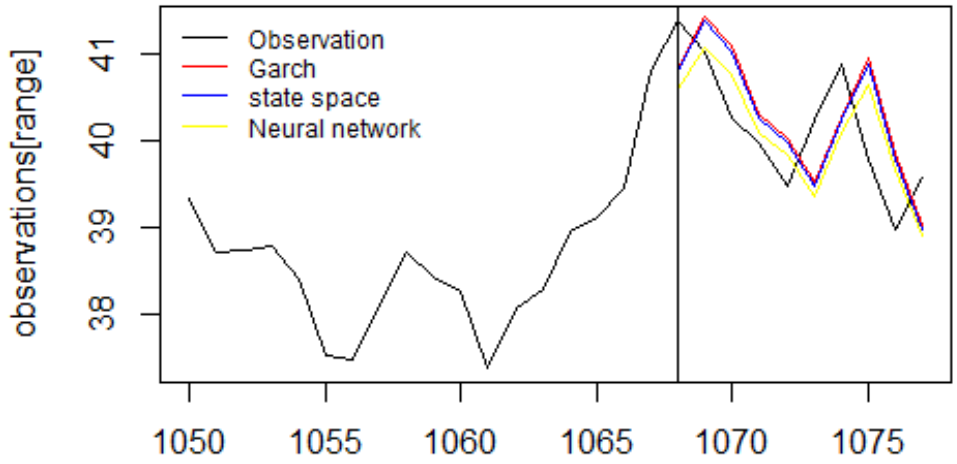
$$\sigma_t^2 = 1.18 \times 10^{-5} + 0.076 a_t^2 + \ .088 \sigma_{t-1}^2$$

# ARMA + GARCH Model



# Predict accuracy comparison

| Method | Mean square prediction error |
|---|---|
| ARMA + GARCH | 0.4703446 |
| STATE SPACE MODEL | 0.45714 |
| ARTIFICIAL NEURAL NETWORK | 0.4047023 |

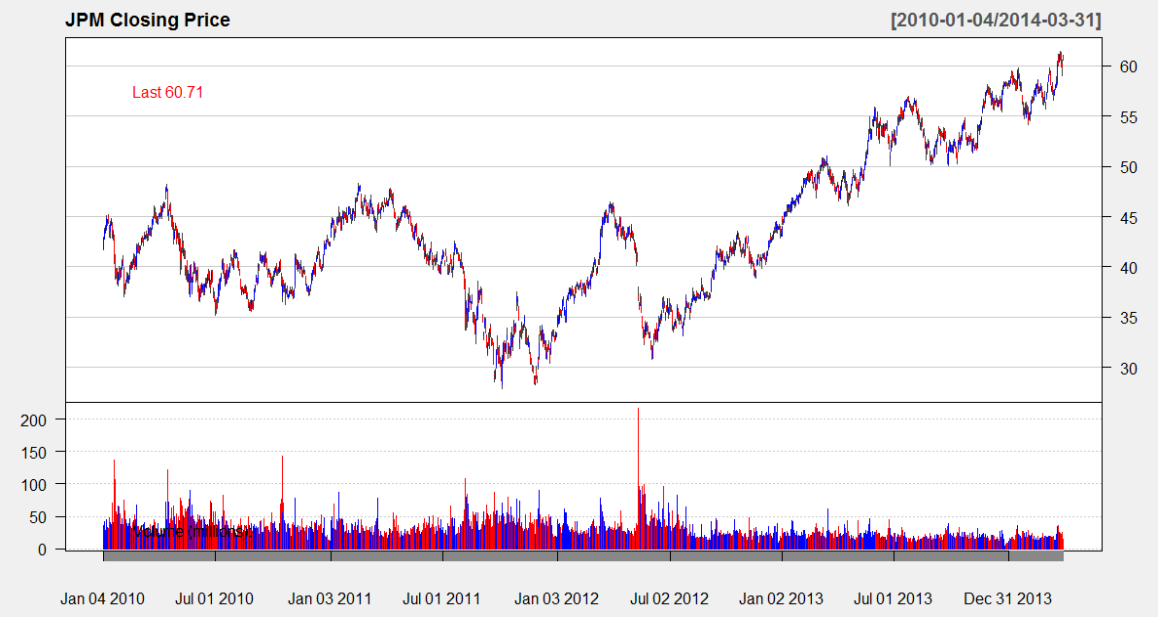


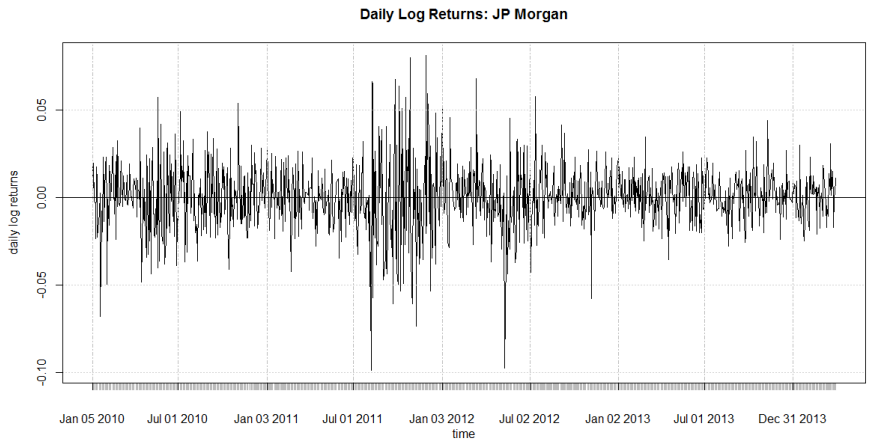
# JP Morgan Chase & Co. (JPM)

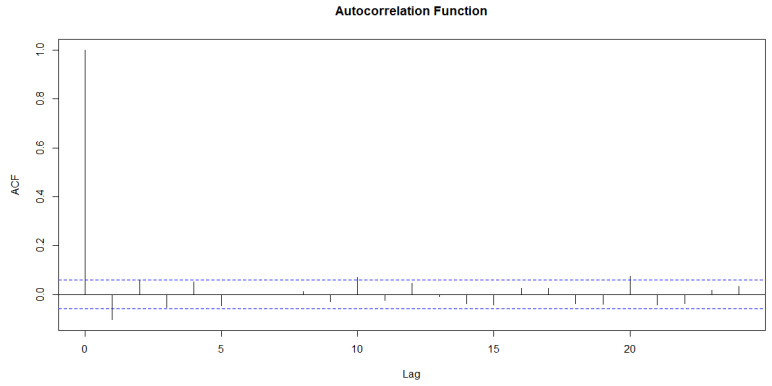
**JPMorgan Chase & Co.** is an American multinational banking and financial services holding company. It is the largest bank in the United States, with total assets of US$2.415 trillion. It is a major provider of financial services, and according to *Forbes* magazine is the world's third largest public company based on a composite ranking. The hedge fund unit of JPMorgan Chase is the second largest hedge fund in the United States. The company was formed in 2000, when Chase Manhattan Corporation merged with J.P. Morgan & Co.

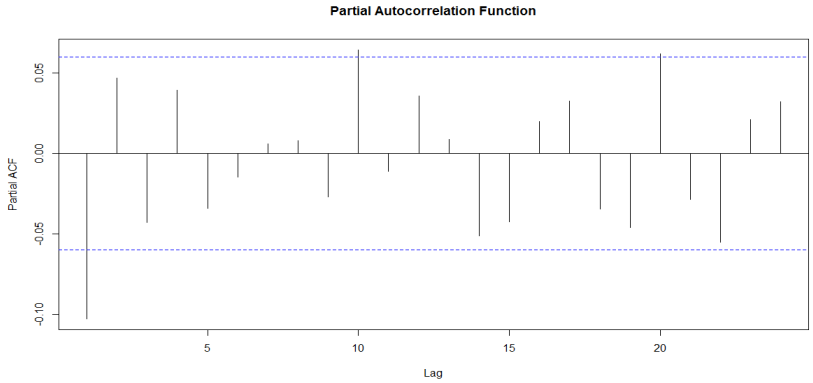
JPMorgan Chase is one of the Big Four banks of the United States with Bank of America, Citigroup and Wells Fargo. According to Bloomberg, as of October 2011, JPMorgan Chase surpassed Bank of America as the largest U.S. bank by assets. Its predecessor, the Bank of the Manhattan Company, was the 22nd oldest bank in the world.

# 1. DATA


JPM Closing Price
[2010-01-04/2014-03-31]
Last 60.71
60
55
50
45
40
35
30
200
150
100
50
0
Volume (millions)
Jan 04 2010
Jul 01 2010
Jan 03 2011
Jul 01 2011
Jan 03 2012
Jul 02 2012
Jan 02 2013
Jul 01 2013
Dec 31 2013


Daily Log Returns: JP Morgan
daily log returns
0.05
0.00
-0.05
-0.10
Jan 05 2010
Jul 01 2010
Jan 03 2011
Jul 01 2011
Jan 03 2012
Jul 02 2012
Jan 02 2013
Jul 01 2013
Dec 31 2013
time


Autocorrelation Function
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
Lag


Partial Autocorrelation Function
Partial ACF
0.05
0.00
-0.05
-0.10
5
10
15
20
Lag

## - Descriptive Statistics : Log Returns

| | |
|---|---|
| Sample size | 1066 |
| Min | -0.0987973 |
| Max | 0.0811109 |
| Mean | 0.0004102 |
| SD | 0.01899152 |
| Skewness | -0.1550941 |
| Kurtosis | 5.819095 |
| Jarque-Bera | 360.0213 |

## - ARCH Effect

```
var = (rtn-mean(rtn))^2
Box.test(var, lag=20, type='Ljung')
```

Box-Ljung test

data: var X-squared = 517.6004, df = 20, p-value < 2.2e-16

⇨ Log-returns have ARCH effect !

# 2. Selection of Best Model

## - Check for the trend

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula  = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-0.097601 -0.009571 -0.000011  0.010281  0.083128

Coefficients:
              Estimate   Std. Error  t value  Pr(>|t|)
(Intercept) -4.363e-04  1.198e-03   -0.364    0.716
z.lag.1     -1.060e+00  4.615e-02  -22.972   <2e-16 ***
tt           1.732e-06  1.930e-06    0.897    0.370
z.diff.lag  -4.588e-02  3.094e-02   -1.483    0.138
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01881 on 1041 degrees of freedom
Multiple R-squared:  0.5565,    Adjusted R-squared:  0.5552
F-statistic: 435.3 on 3 and 1041 DF,  p-value: < 2.2e-16
```

## - Finding the best model

| Model | AIC | BIC | SIC |
|---|---|---|---|
| ARMA(1, 0)-ARCH(1) | -5.228032 | -5.204713 | -5.228076 |
| ARMA(0, 1)-ARCH(1) | -5.227969 | -5.204650 | -5.228013 |
| ARMA(1, 1)-ARCH(1) | -5.227839 | -5.227902 | -5.217237 |
| ARMA(1, 0)-GARCH(1, 1) | -5.337447 | -5.309463 | -5.337510 |
| ARMA(0, 1)-GARCH(1, 1) | -5.337454 | -5.309471 | -5.337517 |
| ARMA(1, 1)-GARCH(1, 1) | -5.335606 | -5.302959 | -5.335691 |
| ARCH(1) | -5.224538 | -5.205883 | -5.224566 |
| GARCH(1, 1) | **-5.334971** | **-5.311651** | **-5.326135** |

**Best Model : GARCH(1, 1)**

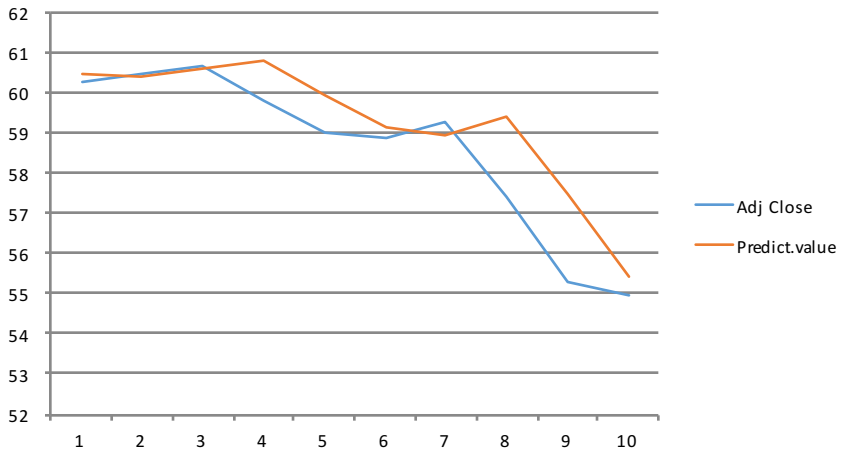
*** ARMA(1,0)-GARCH(1,1) / ARMA(0,1)-GARCH(1,1) / ARMA(1,1)-GARCH(1,1)**
**: one or both of ARMA process parameters is/are not significant**

# 3. GARCH (1, 1) Model

```
Title:
 GARCH Modelling

Call:
 garchFit(formula = ~arma(0, 0) + garch(1, 1), data = rtn, cond.dist = "std")

Mean and Variance Equation:
 data ~ arma(0, 0) + garch(1, 1)
<environment: 0x00000000168d5660>
 [data = rtn]

Conditional Distribution:
 std

Coefficient(s):
        mu       omega      alpha1       beta1       shape
9.2395e-04  4.9024e-06  6.3927e-02  9.2125e-01  7.0176e+00

Std. Errors:
 based on Hessian

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu     9.240e-04   4.520e-04    2.044   0.0409 *
omega  4.902e-06   2.331e-06    2.103   0.0354 *
alpha1 6.393e-02   1.530e-02    4.179 2.92e-05 ***
beta1  9.213e-01   1.816e-02   50.741  < 2e-16 ***
shape  7.018e+00   1.307e+00    5.371 7.82e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 2848.539    normalized:  2.672176

Description:
 Wed Apr 30 12:50:06 2014 by user: SUNG


Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  369.7358  0
 Shapiro-Wilk Test  R    W      0.9770234 5.903182e-12
 Ljung-Box Test     R    Q(10)  9.727982  0.4646738
 Ljung-Box Test     R    Q(15)  12.99365  0.6027881
 Ljung-Box Test     R    Q(20)  17.67099  0.6090716
 Ljung-Box Test     R^2  Q(10)  9.448032  0.4901756
 Ljung-Box Test     R^2  Q(15)  11.54841  0.7128391
 Ljung-Box Test     R^2  Q(20)  15.30473  0.7587132
 LM Arch Test       R    TR^2   10.68526  0.5560673

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
-5.334971 -5.311651 -5.335014 -5.326135
```

# 4. 10 – Day Predictions : 4-1-2014 to 4-14-2014

| Date | Adj Close | Predict.value |
|---|---|---|
| 4/1/2014 | 60.29 | 60.4585 |
| 4/2/2014 | 60.48 | 60.4138 |
| 4/3/2014 | 60.66 | 60.6048 |
| 4/4/2014 | 59.81 | 60.7858 |
| 4/7/2014 | 59 | 59.9313 |
| 4/8/2014 | 58.85 | 59.1174 |
| 4/9/2014 | 59.27 | 58.9661 |
| 4/10/2014 | 57.4 | 59.3883 |
| 4/11/2014 | 55.3 | 57.5109 |
| 4/14/2014 | 54.96 | 55.4057 |



# Portfolio Performance

## 1. Assumptions

- Can short our stocks and hold any fraction of the stocks.
- No transaction costs / No taxes / No commissions
- Can borrow and lend at the one risk-free rate
- Can sell / buy the stocks at their closing prices

## 2. Investment Strategy

- CAPM (Capital Asset Pricing Model) : Mean-Variance Portfolio Theory
- Simple (Naïve) Strategy

## 3. CAPM ( Capital Asset Pricing Model )

$$\boldsymbol{w}_{eff} = \underset{\boldsymbol{w}}{argmin}\ \boldsymbol{w}^T \Sigma\, \boldsymbol{w}$$

$$subject\ to \qquad \boldsymbol{w}^T \boldsymbol{\mu} + (1 - \boldsymbol{w}^T \mathbf{1}) r_f = \boldsymbol{\mu}_*$$

$$L(\boldsymbol{w}, \boldsymbol{\alpha}) = \boldsymbol{w}^T \Sigma\, \boldsymbol{w} - \boldsymbol{\alpha}(\boldsymbol{w}^T \boldsymbol{\mu} + (1 - \boldsymbol{w}^T \mathbf{1}) r_f - \boldsymbol{\mu}_*)$$

$$\boldsymbol{\partial L / \partial w} = \boldsymbol{w}^T = \Sigma\, \boldsymbol{w} - \boldsymbol{\alpha}(\boldsymbol{\mu} - \mathbf{1}) = 0$$

$$\boldsymbol{\partial L / \partial \alpha} = \boldsymbol{w}^T \boldsymbol{\mu} + (1 - \boldsymbol{w}^T \mathbf{1}) r_f - \boldsymbol{\mu}_* = 0$$

$$\Rightarrow \boldsymbol{w}_{eff} = \frac{\boldsymbol{\mu}_* - r_f}{(\boldsymbol{\mu} - \mathbf{1} r_f)^T \Sigma^{-1} (\boldsymbol{\mu} - \mathbf{1} r_f)} \Sigma^{-1} (\boldsymbol{\mu} - \mathbf{1} r_f)$$

## 4. Simple (Naïve) Strategy

: Based on today' our expected returns, we determined the next day's optimal weights

Example.

Expected Returns :

| Date | ORCL | BP | PFE | JPM |
|---|---|---|---|---|
| 4/9/14 | 0.001492 | 0.001843 | 0.002105 | 0.008101 |

Optimal Weights :

| Date | ORCL | BP | PFE | JPM | Sum |
|---|---|---|---|---|---|
| 4/9/14 | 0.1102 | 0.1361 | 0.1555 | 0.5982 | 1 |

# 5. Portfolio Performance

## - Optimal weights of portfolio

### 1) CAPM ( $\mu_* = 0.01$ , $r_f$ = 0.0000001 )

| | [Day 1] | [Day 2] | [Day 3] | [Day 4] | [Day 5] |
|---|---|---|---|---|---|
| ORCL | 11.222466 | -0.4143350 | 0.9778525 | 3.3618611 | 0.4913769 |
| BP | -4.311396 | -1.0064645 | -3.1386777 | 1.1550289 | -3.1565856 |
| PFE | -7.340797 | -0.1807968 | -1.9150774 | 1.6956641 | 3.8875300 |
| JPM | 5.613608 | 6.1125807 | 4.0573297 | -5.0706057 | 6.3722198 |
| RF | -4.183881 | -3.5109845 | 1.0185728 | -0.1419484 | -6.5945410 |

| | [Day 6] | [Day 7] | [Day 8] | [Day 9] | [Day 10] |
|---|---|---|---|---|---|
| ORCL | -0.6936181 | 1.502197 | 0.7061865 | 0.47880719 | 0.5155507 |
| BP | -1.8138328 | 1.698035 | 0.3225542 | 0.09656084 | 0.9687703 |
| PFE | 2.1579918 | 6.159588 | 0.5684975 | 0.81123969 | 2.2932737 |
| JPM | 2.1468143 | -4.375536 | -1.5481485 | -1.21418886 | -2.3562321 |
| RF | -0.7973552 | -3.984283 | 0.9509103 | 0.82758114 | -0.4213627 |

### 2) Simple Strategy

| | [Day 1] | [Day 2] | [Day 3] | [Day 4] | [Day 5] |
|---|---|---|---|---|---|
| ORCL | 0.843923 | 0.194688 | 0.369851 | 0.572129 | 0.796590 |
| BP | 0.081405 | 0.157276 | -0.264659 | 0.183538 | 0.270413 |
| PFE | 0.100913 | 0.139203 | -0.006808 | 0.132697 | 0.718255 |
| JPM | -0.026242 | 0.508831 | 0.901616 | 0.111634 | -0.785259 |

| | [Day 6] | [Day 7] | [Day 8] | [Day 9] | [Day 10] |
|---|---|---|---|---|---|
| ORCL | 0.414255 | 0.110202 | 0.095740 | 0.714873 | 0.256088 |
| BP | -0.204098 | 0.136098 | -0.017721 | -0.102914 | 0.591131 |
| PFE | 0.984640 | 0.155477 | 1.800503 | 1.122383 | 0.555898 |
| JPM | -0.194796 | 0.598220 | -0.878522 | -0.734342 | -0.403119 |

# - 10 Days Realized Returns ( 4/1/14 to 4/14/14)

## 1) CAPM

| | ORCL | BP | PFE | JPM | Portf. Ret. |
|---|---|---|---|---|---|
| Day 1 | 0.370840693 | -0.070495865 | 0.089152490 | -0.008566371 | **0.38093094** |
| Day 2 | 0.008255019 | 0.000000000 | -0.004460846 | 0.044446685 | **0.04824085** |
| Day 3 | -0.041223902 | 0.017874686 | -0.015055303 | 0.027858516 | **-0.01054600** |
| Day 4 | -0.051700300 | 0.007167794 | -0.028782137 | 0.162112325 | **0.08879768** |
| Day 5 | -0.014313282 | 0.050772708 | -0.262025871 | -0.196958762 | **-0.42252521** |
| Day 6 | -0.031553946 | 0.012995300 | -0.052194649 | -0.012546711 | **-0.08330001** |
| Day 7 | 0.055583655 | 0.067594985 | 0.166656852 | -0.072237917 | **0.21759757** |
| Day 8 | -0.042605274 | -0.007562069 | -0.025652128 | 0.110165858 | **0.03434639** |
| Day 9 | -0.022146390 | -0.003779602 | -0.045039546 | 0.099856226 | **0.02889069** |
| Day 10 | 0.018145294 | 0.004236226 | 0.001768791 | 0.033223553 | **0.05737386** |
| | | | | | **0.131107** |

## 2) Simple Strategy

| Date | ORCL | BP | PFE | JPM | Portf. Ret. |
|---|---|---|---|---|---|
| Day 1 | 0.011999894 | 0.000575422 | -0.0005341 | 1.73992E-05 | **0.012058615** |
| Day 2 | -0.001694169 | 0 | 0.001481351 | 0.001603549 | **0.001390731** |
| Day 3 | -0.006854108 | 0.000655639 | -2.31943E-05 | 0.002683383 | **-0.00353828** |
| Day 4 | -0.003837884 | 0.000493792 | -0.000982945 | -0.001564283 | **-0.00589132** |
| Day 5 | -0.010161606 | -0.001897637 | -0.021440472 | 0.010634679 | **-0.02286538** |
| Day 6 | 0.008081495 | 0.000636351 | -0.010414467 | 0.000495247 | **-0.00120137** |
| Day 7 | 0.001752731 | 0.002326957 | 0.001813153 | 0.004269374 | **0.01016221** |
| Day 8 | -0.00255277 | 0.000181649 | -0.035744859 | 0.027717836 | **-0.01039814** |
| Day 9 | -0.014552588 | 0.00176924 | -0.027500402 | 0.026866178 | **-0.01341757** |
| Day 10 | 0.003876154 | 0.001121218 | 0.000186168 | 0.00247849 | **0.00766203** |
| | | | | | **-0.02627549** |

# Conclusion

- We successfully formed an optimal portfolio which gives us about 13.1% investment return using the CAPM model under some assumptions.

- Possible improvements
  - Improve the prediction accuracy using more powerful models
  - Add more stocks to stabilize the portfolio
  - Give more reliable portfolio by adding some basic risk measurement

- Reference
  - http://home.agh.edu.pl/~vlsi/AI/intro/
  - http://www.stephanieharris-biofeedback.com/stephanie-harris-winf.html
  - http://cran.r-project.org/web/packages/dlm/index.html
  - http://www.rinfinance.com/agenda/2012/workshop/Zivot+Yollin.pdf
  - http://en.wikipedia.org/wiki/JPMorgan_Chase
  - Applied Predictive Modeling by Max Kuhn, Kjell Johnson
  - The Element of Statistical Learning by Trevor Hastie, Robert Tibshirani et al
  - Pattern Recognition and Machine Learning by Christopher M.Bishop
  - Statistical Models and Methods for Financial Markets by Tze Lai and Haipeng Xing

# Thank You