

# PROJECT III, Analysis of the STOCK dataset

*Group Member:*

*Ethan Edwards*

*Jie Hu*

*Xun Sun*

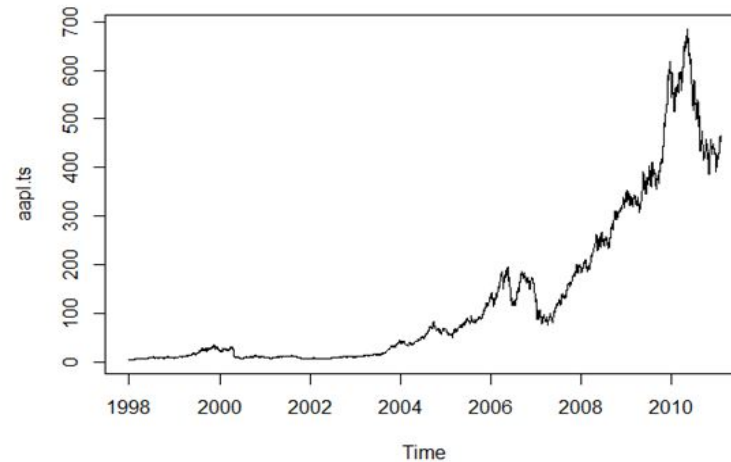
*Yijun Yang*

# Introduction

Instead of analysis the overall stock market, our study today is focus on two famous and well representative companies, APPLE and GOOGLE.

# Case I, APPLE Company

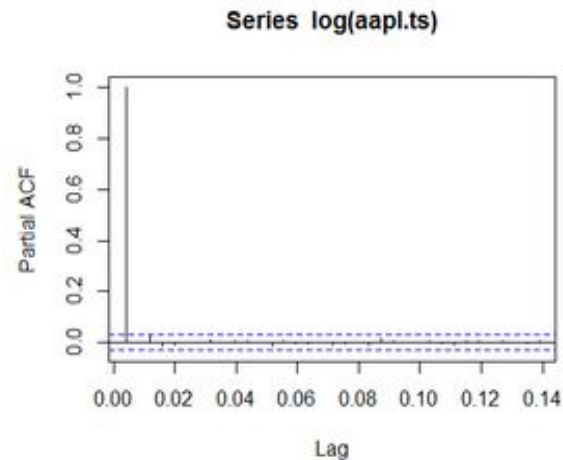
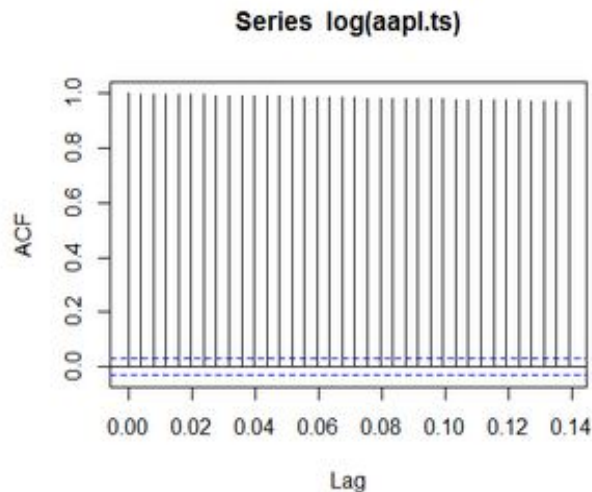
Firstly, We transformed the data to the time series, and plot the trend.



This is the time trend for Apple stock closing price from 1998 to 2013.

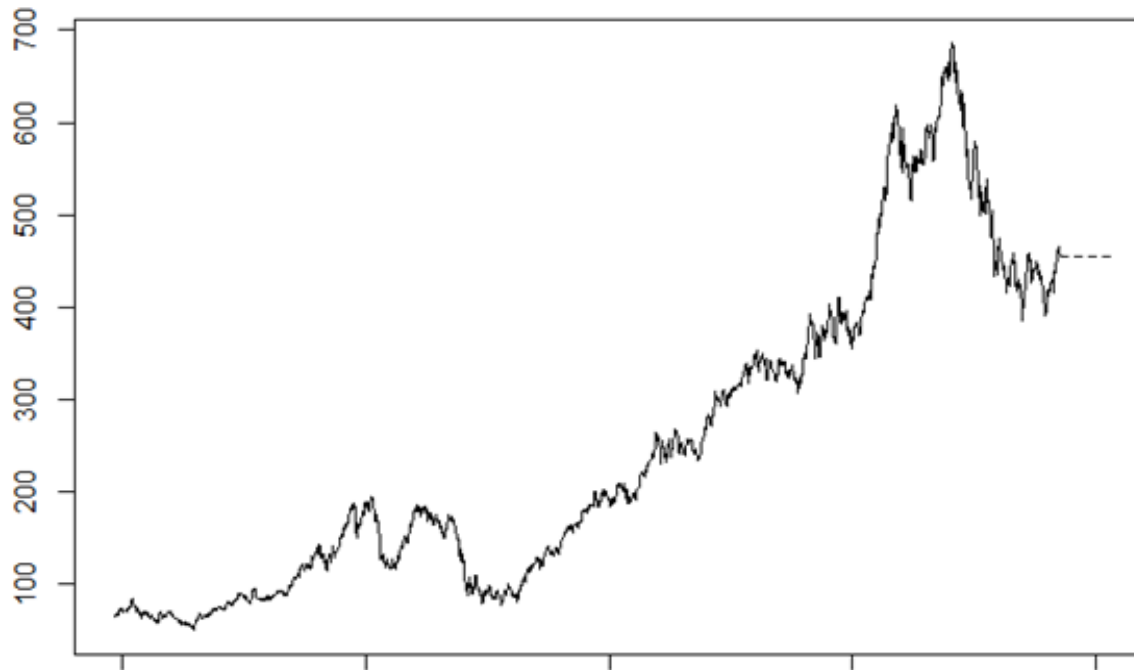
# ACF

Then, we plot the ACF for Apple stock closing price. It suggests there is strong autocorrelation in the time series. Also, use the log transformation to adjust the scale.



# Model

We use the ARIMA model for the time series in order to adjust for the autocorrelation.



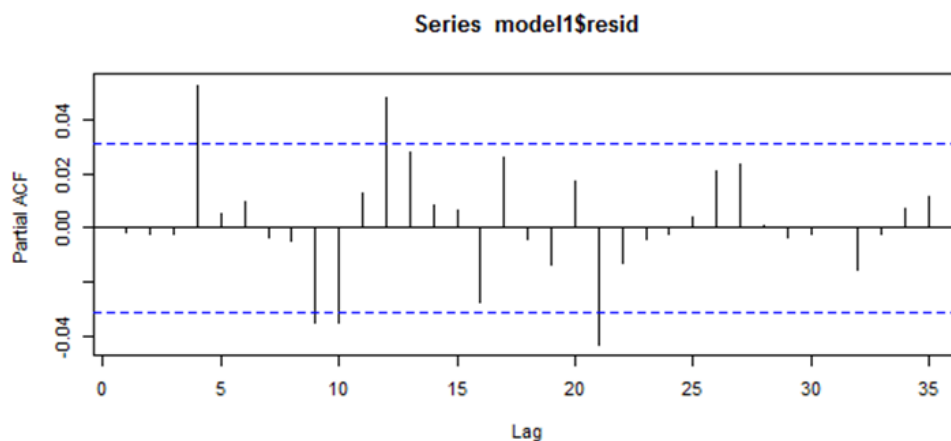
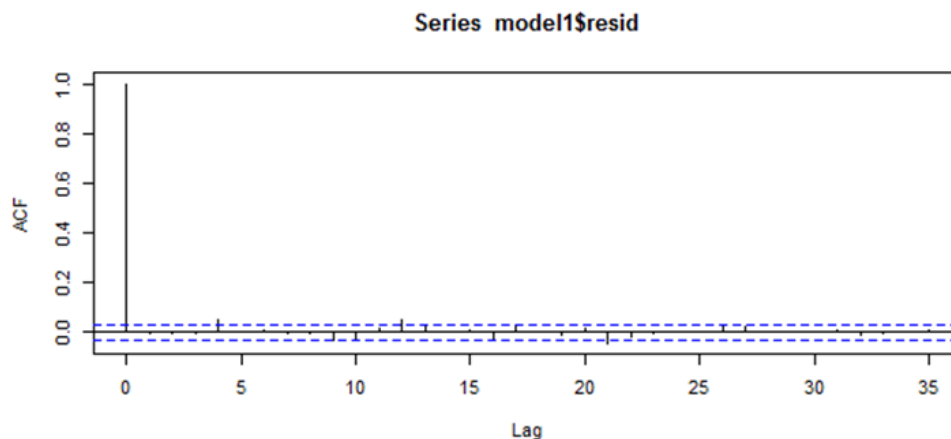
# Prediction

The predicted and the true values of the Apple stocks are compared in the table. The prediction was a bit conservative, that is lower than the true prices.

	1	2	3	4	5
Predicted	454.5740	454.5192	454.5225	454.5220	454.5220
True	454.45	467.36	489.57	498.50	497.91

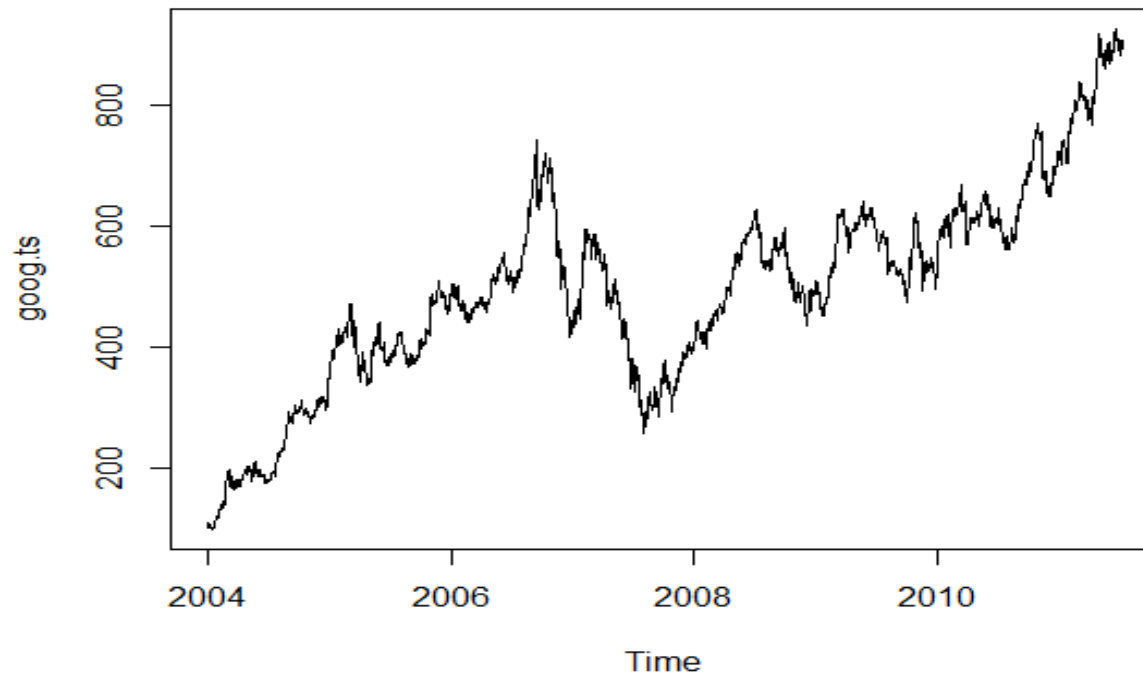
# Diagnostics

In order to check whether the model is fit or not, we need to see the ACF and PACF of the model, it shows the autocorrelation was adjusted. That is good.



# Case III, GOOGLE Company

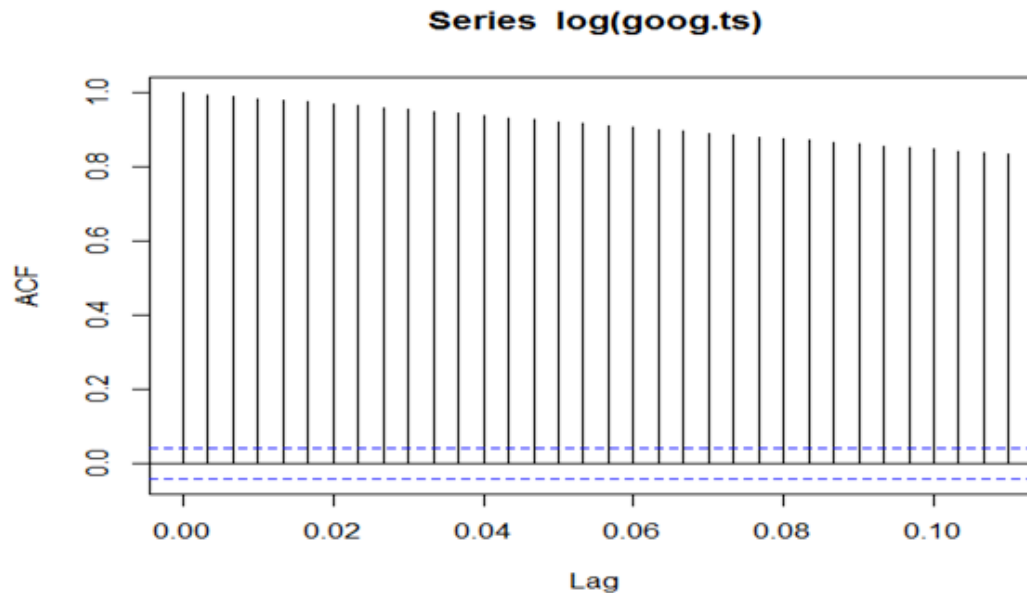
Here is the time trend for Google stock closing price from 2004 to 2013.





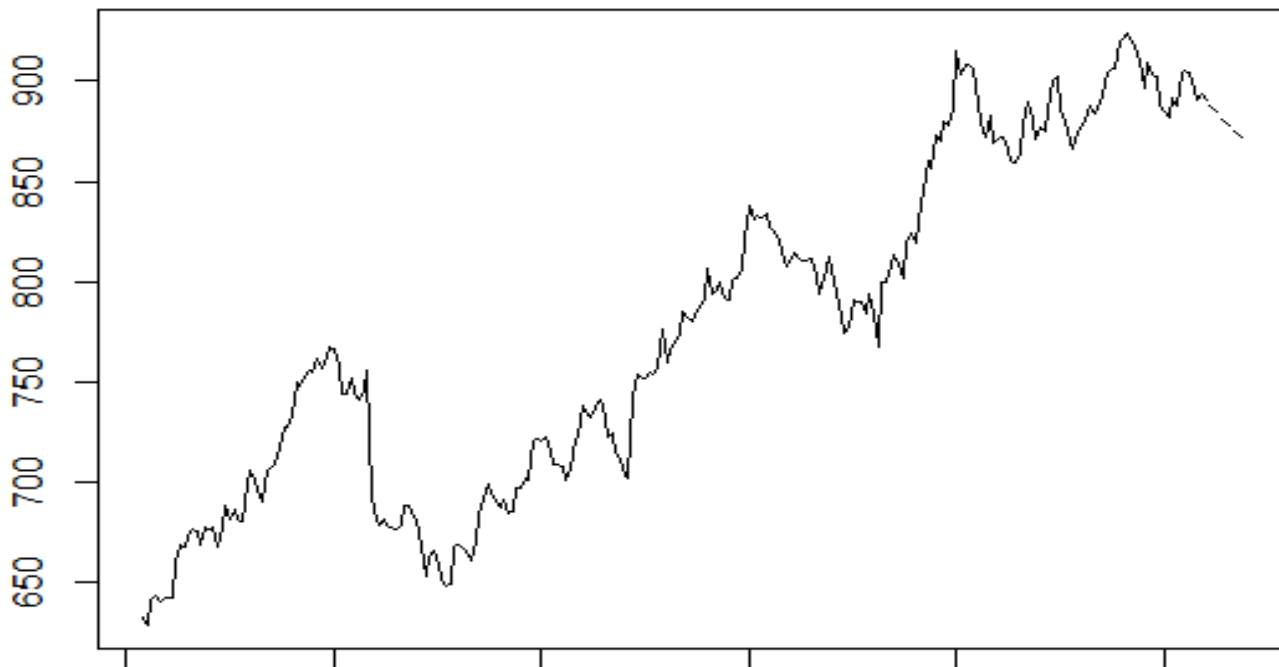
# ACF

The plot below is the ACF for Google stock closing price. It suggests there is strong autocorrelation in the time series. Also, use the log transformation to adjust the scale.



# Model

We use the ARIMA model for the time series in order to adjust for the autocorrelation.



# Prediction

The predicted and the true values of the Google stocks are compared in the table. The prediction was still a bit conservative, that is lower than the true prices.

	1	2	3	4	5
Predicted	888.0494	885.8044	883.5748	881.3605	879.1613
True	909.32	903.61	902.34	887.74	885.08

# Using R-package “quantmod” for APPLE Company

R package “quantmod” provides a function that does this. However, the difficulty was associated with the package since the plotting function `barChart` and `candleChart` identifies only `xts` object. The data set was transformed into the set using `as.date` function.



One classical approach to analyze stock data is to use KDJ plot.

rsv is defined as the

$$RSV = \frac{\text{Current closing price} - 9 \text{ days low}}{9 \text{ day price range}} \times 100$$

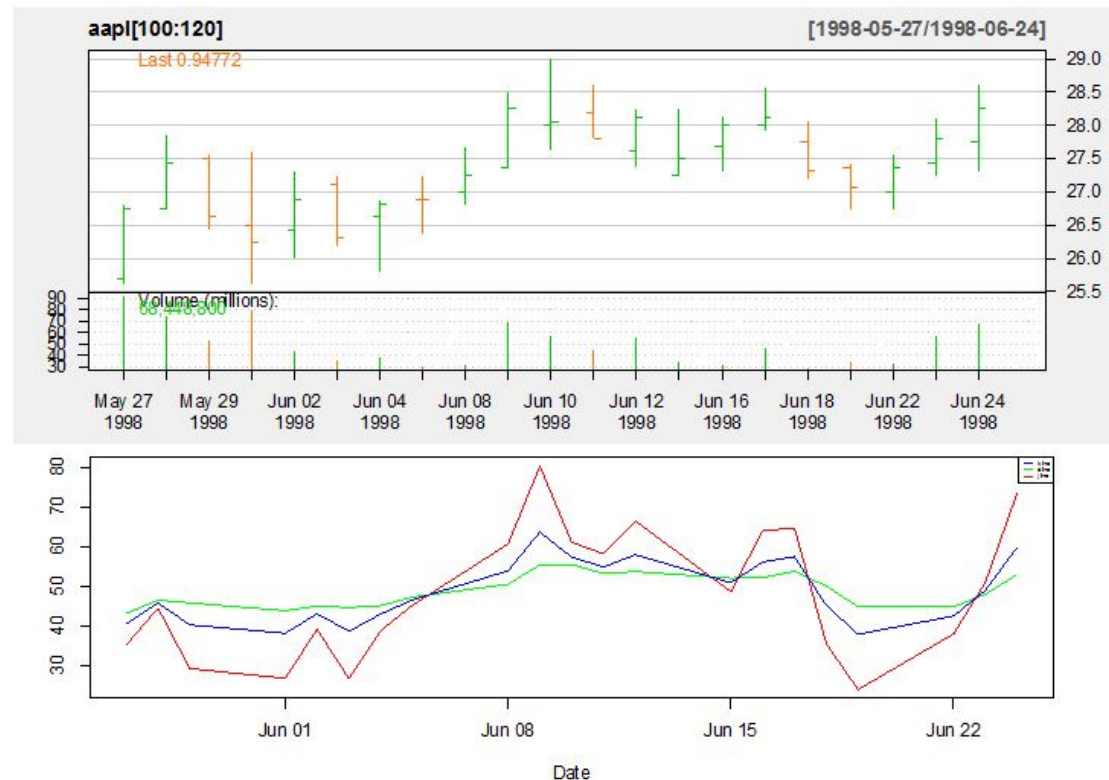
$$K = \frac{2}{3} * \text{last } K + \frac{1}{3} * \text{current } RSV$$

$$D = \frac{2}{3} * \text{last } D + \frac{1}{3} * \text{current } K$$

$$\underline{J} = 3 * \text{current } K - 2 * \text{current } D$$

By plotting KDJ line to the corresponding candle line, it is easy to see that when K line cross D line from below indicates a start of price increment. In addition, a lower K value ( $K < 30$ ) indicates a oversell in the market.

The data was screened by the decision function built on the two categories, the result was 97 out of 110 trade in point was profitable.



# Obstacles

- The prevalence of NAs was fairly sizeable in the data.
- `ts()` time series format.

# Future Works

- For error analysis, we would consider GARCH-modeled errors (General Autoregressive Conditionally Heteroskedastic).
- Not accurate prediction
- Stock Market is unpredictable!!