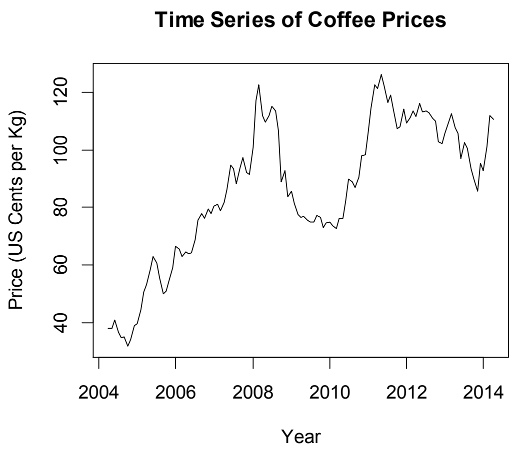
Time Series
Student Project

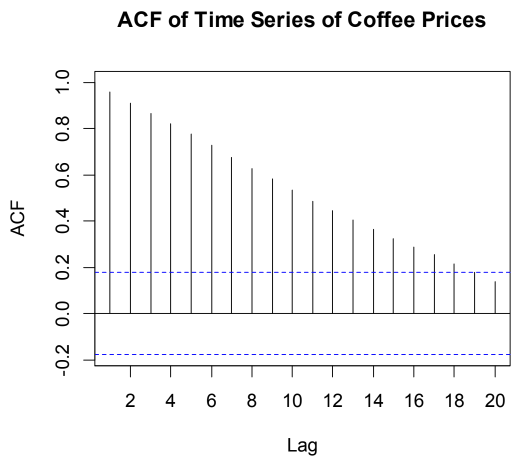
Yen-Chun Lin
peterlin003@gmail.com

There are more and more people enjoy coffee. Currently, coffee is one of the most widely consumed beverages in the world. This project is about the development of ARIMA model for monthly coffee prices over the past ten years.

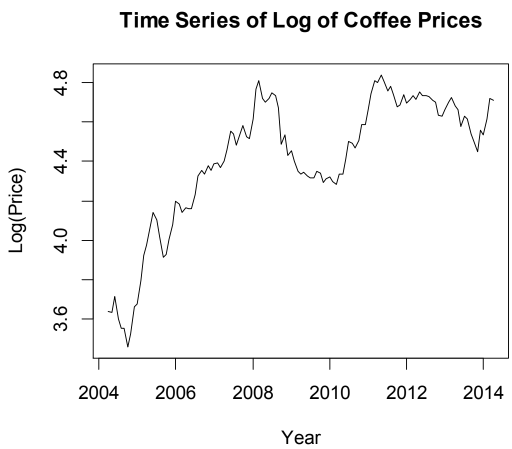
I obtain the data regarding the monthly coffee prices at the following website. http://www.indexmundi.com/commodities/?commodity=robusta-coffee&months=120. In this project, I would ignore inflation impact within the coffee prices over the past years and use the nominal coffee prices in my analysis.

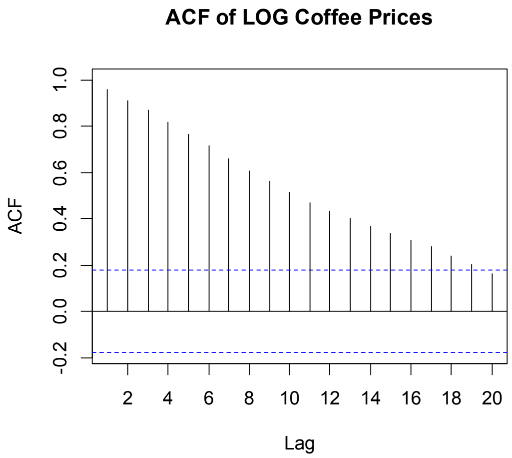
In the graph of time series, coffee prices show an upward trend before 2008 and then downward to 2010 and upward in 2011. There is fairly large fluctuation in the coffee price with exponential path. And autocorrelation of time series of coffee prices will not quickly decay toward zero. There are many evidences to show that the data is non-stationary.



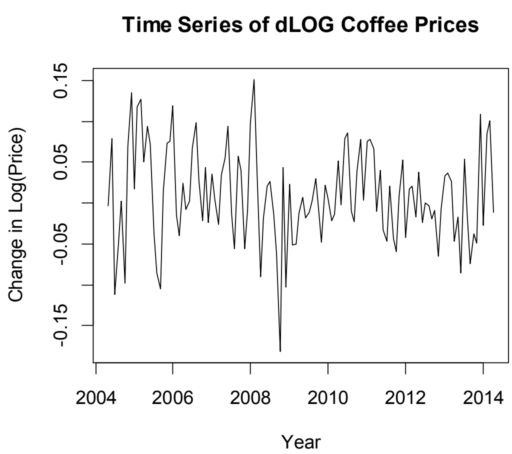


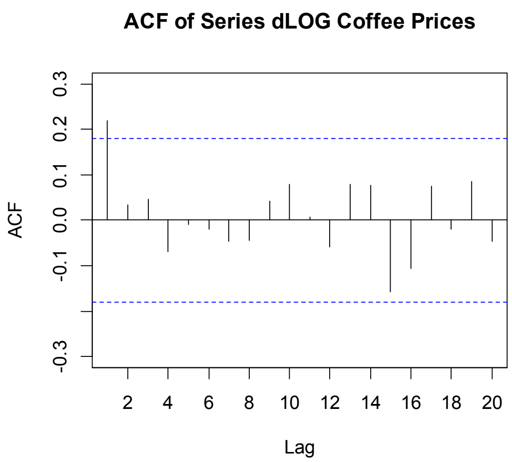
Taking the log of the coffee prices, the log of the coffee prices appears to be more linearly than the above graph. Additionally, the autocorrelation are not decaying quickly toward zero than the above.



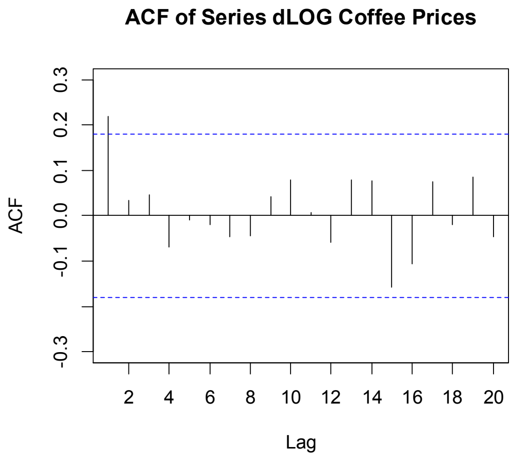


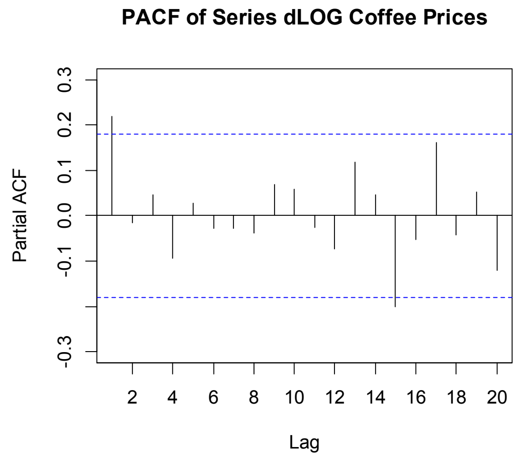
In order to get a stationary time series, I would take the first difference of the log of coffee prices. The time series is more stationary and the autocorrelation clearly decay towards zero.





Based on the above analysis, I would proceed by fitting the log-differenced price data with a stationary time series model. Using the below autocorrelation and partial autocorrelation functions of the log-differenced data, the autocorrelation exceeding the 95% confidence band at lag 1 in the autocorrelation function, and at lag 1 in the partial autocorrelation function. This suggests that MA(1) and AR(1) model would be appropriate in this case. Otherwise, the partial autocorrelation at lag 15 also exceed the 95% confidence band, which I consider that it would be due to outlier.





There are some models fitting as follows:

(I) ARIMA(0,1,1)

Coefficients:

	ma1
	0.2396
s.e.	0.0903

sigma^2 estimated as 0.003281: log likelihood = 172.87, aic = -341.74

(II) ARIMA(1,1,0)

Coefficients:

	ar1
	0.2354
s.e.	0.0883

sigma^2 estimated as 0.00328: log likelihood = 172.9, aic = -341.79

(III) ARIMA(1,1,1)

Coefficients:

	ar1	ma1
	0.1830	0.0554
s.e.	0.6153	0.6378

sigma^2 estimated as 0.00328: log likelihood = 172.9, aic = -339.8

According to measure of the log likelihood score and AIC, ARIMA(1,1,0) has higher log likelihood score and lower AIC. I would test the ARIMA(1,1,0) in my next analysis.

Using the Ljung-Box statistic for ARIMA(1,1,0), the result is shown as follows:

Box-Ljung test

data: residuals(AR1)

X-squared = 2.9699, df = 9, p-value = 0.9655

Under the 9 degrees of freedom, the p-value = 0.9655. There is no significant evidence to reject the null hypothesis that the residuals are uncorrelated at the 95% confidence level. According the my above analysis, an autoregressive model of order 1(AR(1)) is a good fit for the coffee prices in this study.