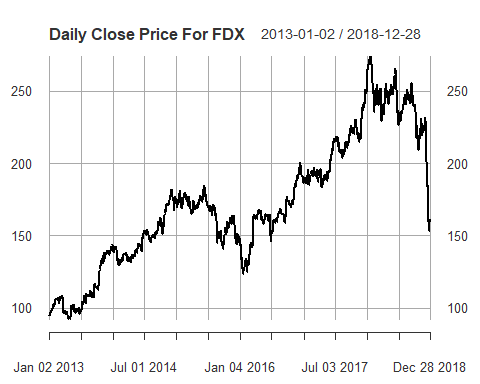
Modeling the Daily Close Price for FDX

Bilal Rehman

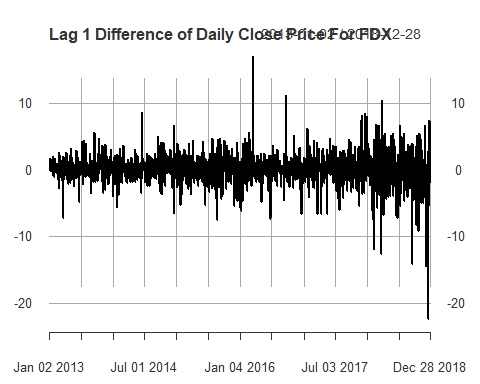
The goal of this report is to test the effectiveness of a log transformation on the closing price of FedEx stock. The daily closing prices for FedEx were acquired from Yahoo Finance and everything was coded in R.

The information included in the data set was the date, the opening price for that day, the absolute high and low prices, the closing and adj closing figure, and the volume of the entire price entries for that respective date. The date range used was January 1, 2013 to December 31, 2018 and the closing prices from the data was isolated.

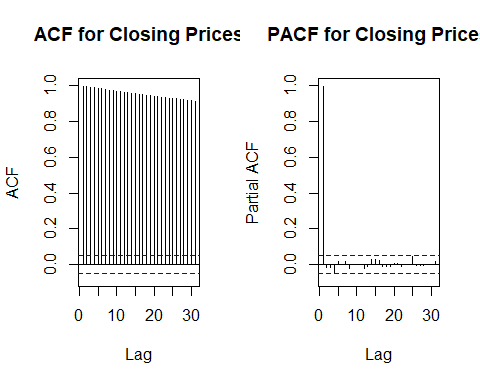
The analysis began by creating a model based off the untransformed closing prices from the dataset. The date vs closing price graph is as follows:



And the lag-1 closing price vs date graph is as follows:



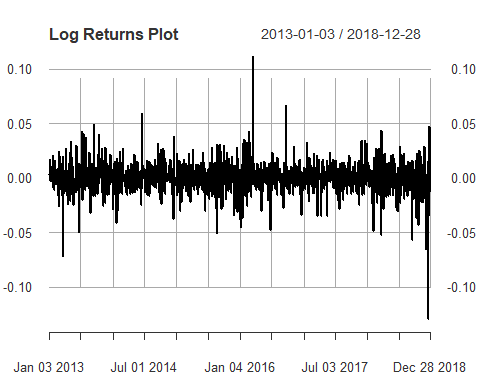
The mean can be seen centered around 0. To create an ARIMA model the ACF and PACF were graphed which is as follows:



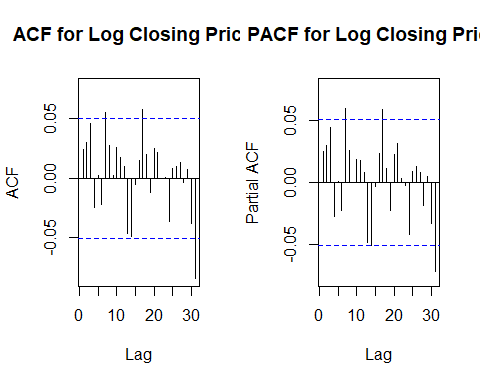
From the graphs it was concluded that the model ARI(1,1) will be used because the ACF is geometric and the PACF’s lag is only significant for the first lag. The model is as follows:

## Call:  
## arima(x = FDXClose, order = c(1, 1, 0))  
##   
## Coefficients:  
## ar1  
## 0.0305  
## s.e. 0.0257  
##   
## sigma^2 estimated as 6.164: log likelihood = -3511.09, aic = 7026.19

Next, the dataset was transformed the with a log transformation with lag-1 to see if it was a better fit. The log returns vs date plot is as follows:

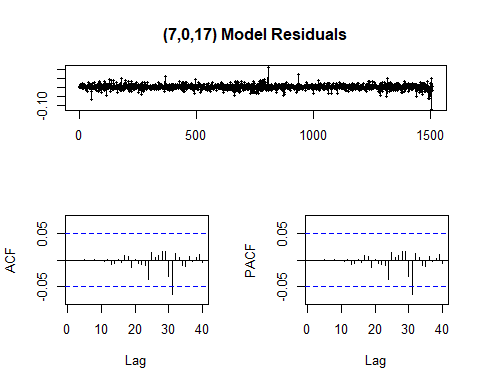


The mean is centered around 0. To figure out the best model for the transformed data the ACF and PACF were graphed, which are as follows:

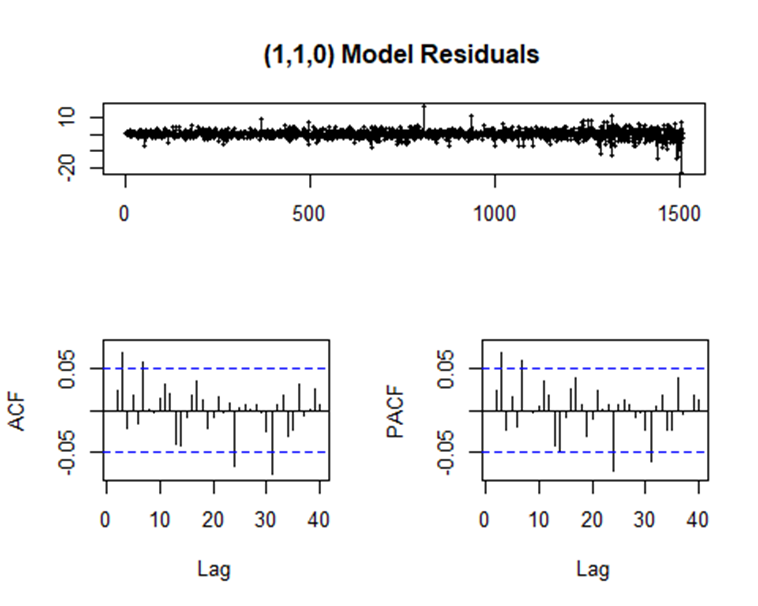


From the graphs it can be concluded that the ideal model will be ARMA(7,17) because the 7th and 17th lags are significant. The model is as follows:

## Call:  
## arima(x = logs, order = c(7, 0, 17))  
##   
## Coefficients:  
## ar1 ar2 ar3 ar4 ar5 ar6 ar7 ma1  
## 0.3823 -0.0671 0.4507 0.1647 -0.3285 0.3998 -0.4807 -0.3621  
## s.e. 0.3871 0.1953 0.2564 0.2439 0.1779 0.3330 0.1951 0.3868  
## ma2 ma3 ma4 ma5 ma6 ma7 ma8 ma9  
## 0.0931 -0.4134 -0.2150 0.3233 -0.4466 0.5596 0.0237 -0.0123  
## s.e. 0.1933 0.2490 0.2351 0.1716 0.3183 0.1950 0.0531 0.0496  
## ma10 ma11 ma12 ma13 ma14 ma15 ma16 ma17  
## 0.0407 -0.0252 0.0310 -0.0761 -0.0247 0.0211 0.0326 0.078  
## s.e. 0.0483 0.0401 0.0388 0.0350 0.0405 0.0361 0.0305 0.045  
## intercept  
## 3e-04  
## s.e. 5e-04  
##   
## sigma^2 estimated as 0.0001944: log likelihood = 4303.33, aic = -8554.66

To see which of the 2 models is better, AIC and BIC values must be compared. The AIC and BIC values of the ARMA(7,17) model are -8554.66 and -8416.38 respectively and for the ARI(1,1) model the AIC and BIC values are 7026.189 and 7036.826 respectively. Since ARMA(7,17) model’s values are so much smaller, it can be concluded that model is better. We can also compare the ACF and PACF of the models’ residuals. The model with the least significant number of lags is the better model. ARMA(7,17)’s residual graph is as follows:

Only one lag from both the ACF and PACF are significant. ARI(1,1)’s residual graph is as follows:



There are 4 significant lags in both the ACF and PACF.

In conclusion, the ARMA(7,17) model from the log transformation with lag-1 is better than the ARI(1,1). This means, the log transformation was favorable to the untransformed model. Also, in the closing price vs date that there was an odd trend at the end of the data set. There was a dip in the closing prices that could have possibly compromised the accuracy of the models. If the goal was to create the most accurate model possible, that portion of the data would have to be removed because it does not seem to be part of the overall trend.

Appendix

**Project-Code.R**

2019-12-09

**library**(forecast)

## Warning: package 'forecast' was built under R version 3.5.3

**library**(tseries)

## Warning: package 'tseries' was built under R version 3.5.3

**library**(timeSeries)

## Warning: package 'timeSeries' was built under R version 3.5.3

## Loading required package: timeDate

## Warning: package 'timeDate' was built under R version 3.5.3

**library**(quantmod)

## Warning: package 'quantmod' was built under R version 3.5.3

## Loading required package: xts

## Warning: package 'xts' was built under R version 3.5.3

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 3.5.3

##   
## Attaching package: 'zoo'

## The following object is masked from 'package:timeSeries':  
##   
## time<-

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

## Loading required package: TTR

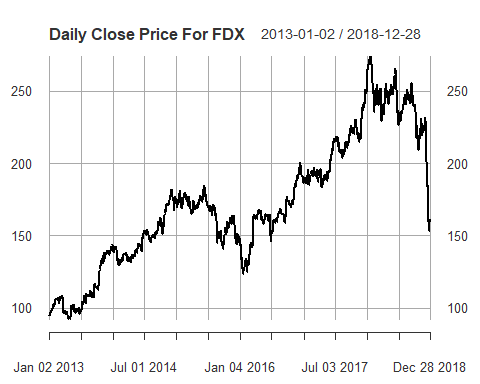
## Warning: package 'TTR' was built under R version 3.5.3

## Version 0.4-0 included new data defaults. See ?getSymbols.

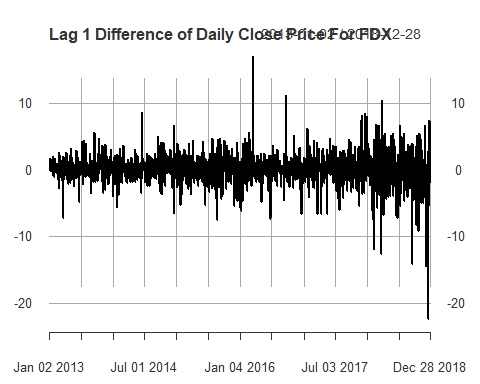
##Setting up a model without any transformation  
  
**options**("getSymbols.warning4.0"=FALSE)  
**getSymbols**('FDX', from='2013-01-01', to='2018-12-31')

## [1] "FDX"

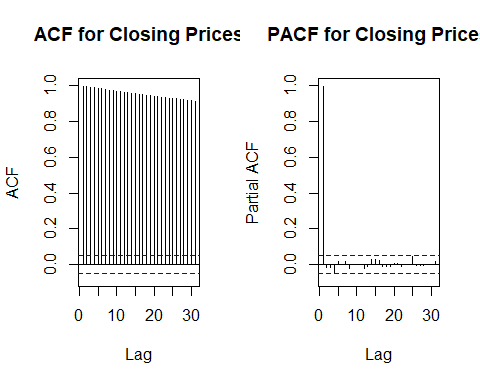
*#imports data set from Yahoo Finance*  
FDXClose=FDX[,4]  
*#Isolates closing Price*  
  
**plot**(FDXClose, main="Daily Close Price For FDX")



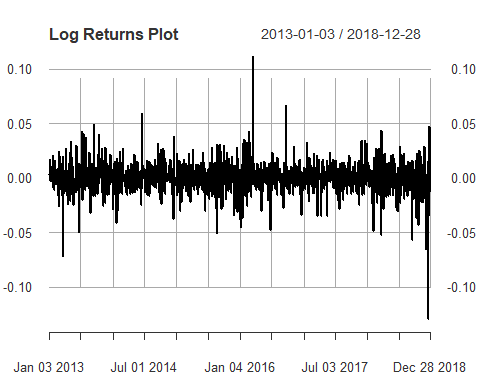
*#Plot of closing price*   
**plot**(**diff**(FDXClose, lag=1), main="Lag 1 Difference of Daily Close Price For FDX")



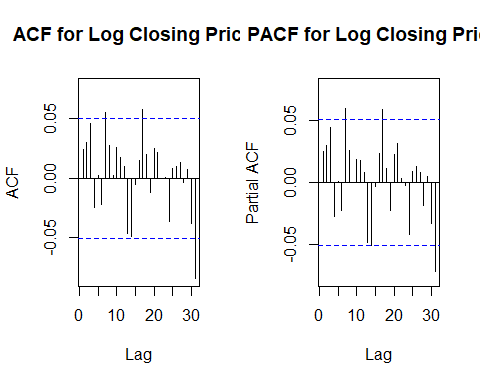
*#plot of the difference of closing prices with lag=1*  
  
**par**(mfrow=**c**(1,2))  
**Acf**(FDXClose, main="ACF for Closing Prices")  
**Pacf**(FDXClose, main="PACF for Closing Prices")



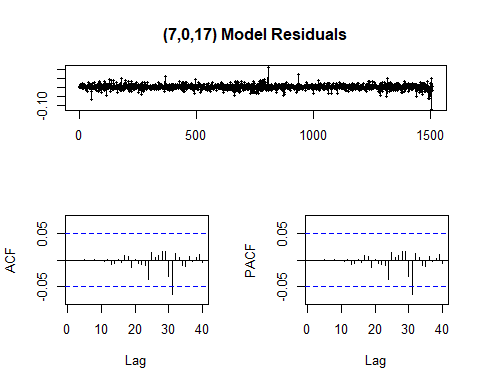
*#creates ACF and PACF for closing prices*  
*#from the graph we can see that the ACF plot is geometric and PACF is significat untill the second lag, this means we will*  
*#use the AR(1) model*  
  
  
##Setting up a model for log transformation  
  
logs=**diff**(**log**(FDXClose),lag=1)  
logs=logs[**!is.na**(logs)]  
*#log transformation with lag=1*  
  
**par**(mfrow=**c**(1,1))  
**plot**(logs,main="Log Returns Plot")



*#plot of the transformed data with lag=1*  
  
**par**(mfrow=**c**(1,2))  
**Acf**(logs, main="ACF for Log Closing Prices")  
**Pacf**(logs, main="PACF for Log Closing Prices")



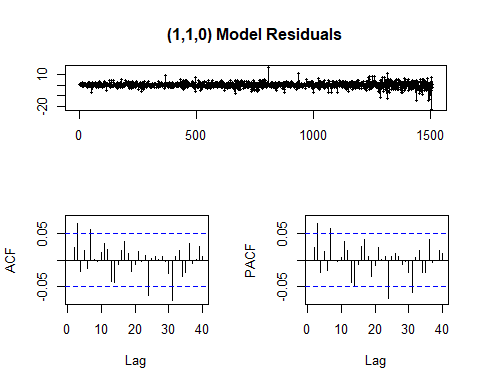
*#creates ACF and PACF for closing prices*  
*#from the 2 graphs we can see that our arima model will be (7,0,17), d=0 becaude that data is already transformed at lag=1*  
  
  
fit<-**arima**(logs, order = **c**(7,0,17))  
**tsdisplay**(**residuals**(fit), lag.max=40, main="(7,0,17) Model Residuals")



*#this is the arima model from the log transformation*   
fit

##   
## Call:  
## arima(x = logs, order = c(7, 0, 17))  
##   
## Coefficients:  
## ar1 ar2 ar3 ar4 ar5 ar6 ar7 ma1  
## 0.3823 -0.0671 0.4507 0.1647 -0.3285 0.3998 -0.4807 -0.3621  
## s.e. 0.3871 0.1953 0.2564 0.2439 0.1779 0.3330 0.1951 0.3868  
## ma2 ma3 ma4 ma5 ma6 ma7 ma8 ma9  
## 0.0931 -0.4134 -0.2150 0.3233 -0.4466 0.5596 0.0237 -0.0123  
## s.e. 0.1933 0.2490 0.2351 0.1716 0.3183 0.1950 0.0531 0.0496  
## ma10 ma11 ma12 ma13 ma14 ma15 ma16 ma17  
## 0.0407 -0.0252 0.0310 -0.0761 -0.0247 0.0211 0.0326 0.078  
## s.e. 0.0483 0.0401 0.0388 0.0350 0.0405 0.0361 0.0305 0.045  
## intercept  
## 3e-04  
## s.e. 5e-04  
##   
## sigma^2 estimated as 0.0001944: log likelihood = 4303.33, aic = -8554.66

fit2<-**arima**(FDXClose, order = **c**(1,1,0))  
**tsdisplay**(**residuals**(fit2), lag.max=40, main="(1,1,0) Model Residuals")



*#this is the arima model from before the log transformation*   
fit2

##   
## Call:  
## arima(x = FDXClose, order = c(1, 1, 0))  
##   
## Coefficients:  
## ar1  
## 0.0305  
## s.e. 0.0257  
##   
## sigma^2 estimated as 6.164: log likelihood = -3511.09, aic = 7026.19

*#from the graphs we can see that the arima model from the log transformation has only one lag that passes the CI in the ACF*   
*#and PACF graph while the original model has 4 lags that pass the CI in the ACF and PACF graphs. We can use this to judge the*  
*#accuracy of the models. Since the model with the log transformation has fewer lags that leave the CI, it is more accurate.*   
  
  
**AIC**(fit,k=2)

## [1] -8554.657

**AIC**(fit2,k=2)

## [1] 7026.189

*#the log based model has a smaller AIC value than the other model. This means it is a better fit*  
  
  
**BIC**(fit)

## [1] -8416.375

**BIC**(fit2)

## [1] 7036.826

*#the log based model has a smaller BIC value than the other model. This means it is a better fit*