Exercise 6 - Veronica Bayani

Use the psei_ret (Philippine Stock Exchange Index Log Returns) as derived from PSEI.PS_AdjClose.csv through the preamble code.

Loading the data in R

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.2.2

## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2

## Warning: package 'ggplot2' was built under R version 4.2.2

## Warning: package 'readr' was built under R version 4.2.2

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(fpp2)
```

```
## Warning: package 'fpp2' was built under R version 4.2.2

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
## -- Attaching packages ---------------------------------------------- fpp2 2.4 --
## v forecast  8.18     v expsmooth 2.3
## v fma       2.4

## Warning: package 'forecast' was built under R version 4.2.2

##
```

```
library(tinytex)
```

```
## Warning: package 'tinytex' was built under R version 4.2.2
```

```r
library(forecast)
library(MASS)
```

```
## Warning: package 'MASS' was built under R version 4.2.2

##
## Attaching package: 'MASS'
##
## The following objects are masked from 'package:fma':
##
##     cement, housing, petrol
##
## The following object is masked from 'package:dplyr':
##
##     select
```

```r
library(fGarch)
```

```
## Warning: package 'fGarch' was built under R version 4.2.2

## NOTE: Packages 'fBasics', 'timeDate', and 'timeSeries' are no longer
## attached to the search() path when 'fGarch' is attached.
##
## If needed attach them yourself in your R script by e.g.,
##         require("timeSeries")
```

```r
library(rugarch)
```

```
## Warning: package 'rugarch' was built under R version 4.2.2

## Loading required package: parallel
##
## Attaching package: 'rugarch'
##
## The following object is masked from 'package:purrr':
##
##     reduce
##
## The following object is masked from 'package:stats':
##
##     sigma
```

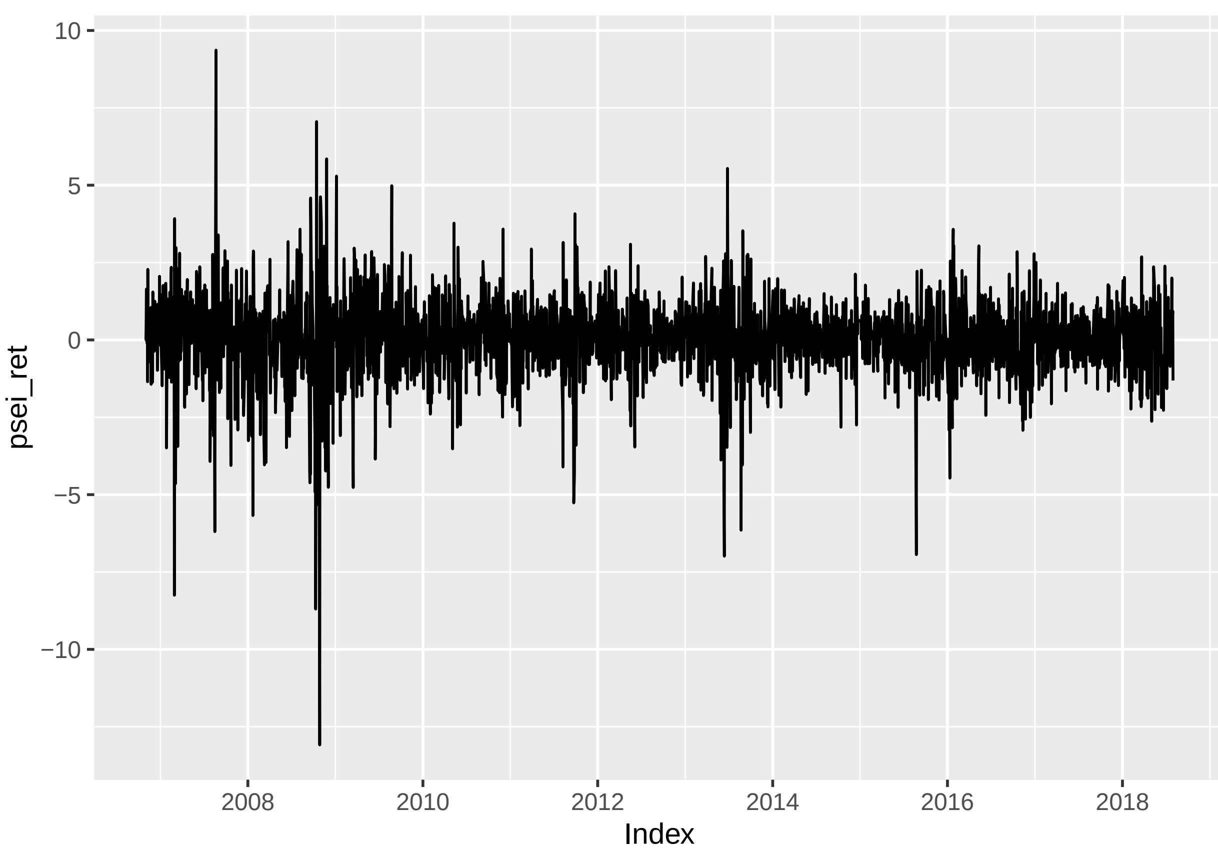
```r
library(zoo)
```

```
## Warning: package 'zoo' was built under R version 4.2.2

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

```
library(xts)
```

```
## Warning: package 'xts' was built under R version 4.2.2

##
## Attaching package: 'xts'
##
## The following objects are masked from 'package:dplyr':
##
##     first, last
```

```
#loading the monthly data
psei <- read.csv("PSEI.PS_AdjClose.csv", stringsAsFactors = FALSE, na.strings = c("NA"))

### Remove NAs
psei <- na.omit(psei)

### Forming the xts structure in PSEI
psei <- xts(psei$PSEI, as.Date(psei$Date, format = "%Y-%m-%d"))

### Generating the log returns of the PSEI with NAs removed, converted to Percent
psei_ret <- na.omit(diff(log(psei)))*100
autoplot(psei_ret)
```



Split the data into training and test data set: Test dataset = 250 most recent returns, and Training dataset = remaining observations.

```r
#for the train set
psei_train <- psei_ret[1:2631]


#for the test set

psei_test <- tail(psei_ret,250)
```

1) [3pts] Select an appropriate ARMA-GARCH model with the best-possible distribution fit for the return series. An appropriate model should have the lowest information criteria values and should fulfill GARCH assumptions and residual diagnostic tests.

Model 1: model=sGarch, garchOrder(1,1), armaOrder(1,1), distribution model=jsu

```r
### 1. SPECIFICATION (Mean, Variance, and Distribution)

library(rugarch)

spec1 <- ugarchspec(
    variance.model = list(model = "sGARCH",
                          garchOrder = c(1, 1),
                          external.regressors = NULL),
    mean.model = list(armaOrder = c(1, 1),
                      include.mean = TRUE,
                      archm = FALSE, archpow = 1,
                      external.regressors = NULL),
    distribution.model = "jsu")

### 2. FIT
fit1 <- ugarchfit(spec1,
                  psei_train ,
                  out.sample = 250)

print(fit1)
```

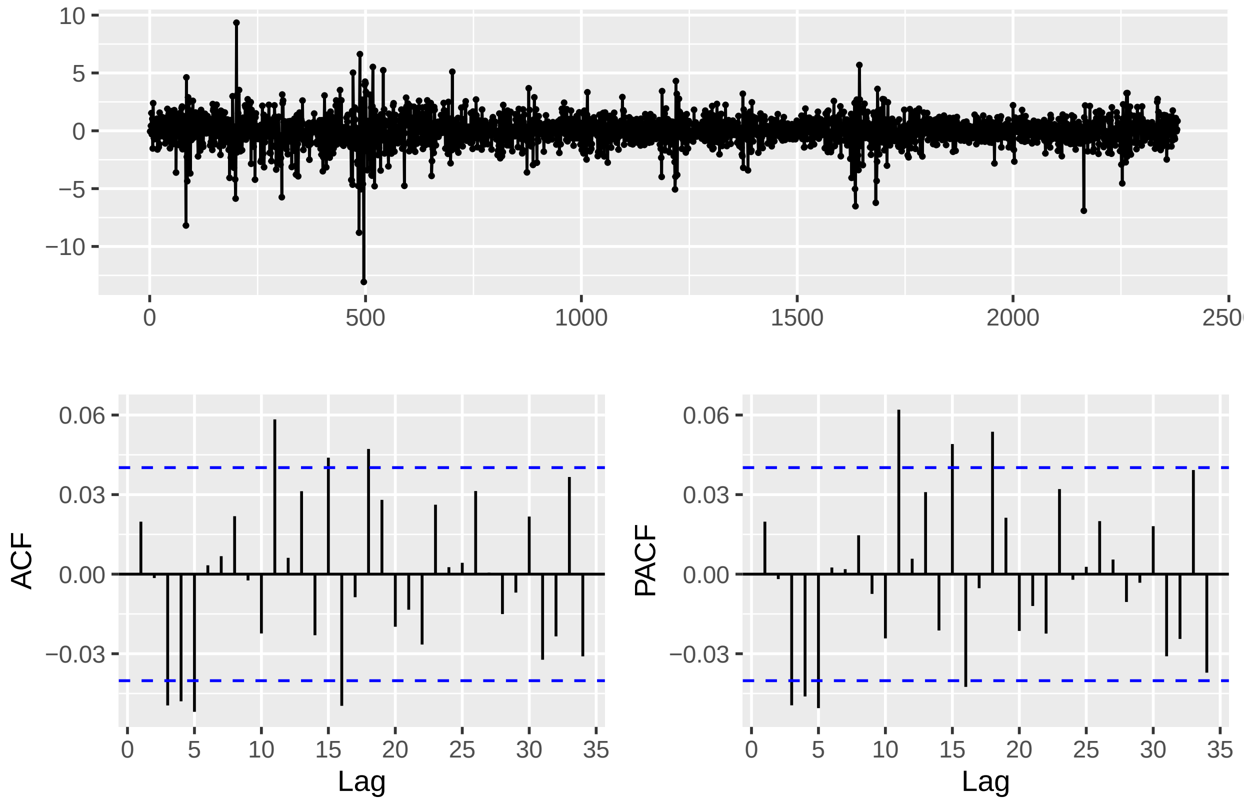
```
##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(1,1)
## Mean Model   : ARFIMA(1,0,1)
## Distribution : jsu
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.065426    0.022176   2.9503 0.003174
## ar1    -0.222570    0.146850  -1.5156 0.129615
## ma1     0.319224    0.142080   2.2468 0.024653
```

```
## omega   0.062643    0.015977   3.9209 0.000088
## alpha1  0.141246    0.020306   6.9558 0.000000
## beta1   0.822280    0.023742  34.6336 0.000000
## skew   -0.417359    0.109251  -3.8202 0.000133
## shape   1.945509    0.147876  13.1563 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.065426    0.021173   3.0900 0.002001
## ar1    -0.222570    0.101708  -2.1883 0.028646
## ma1     0.319224    0.100210   3.1856 0.001445
## omega   0.062643    0.019423   3.2252 0.001259
## alpha1  0.141246    0.021551   6.5540 0.000000
## beta1   0.822280    0.026929  30.5353 0.000000
## skew   -0.417359    0.107665  -3.8765 0.000106
## shape   1.945509    0.167020  11.6484 0.000000
##
## LogLikelihood : -3625.028
##
## Information Criteria
## ------------------------------------
##
## Akaike       3.0517
## Bayes        3.0711
## Shibata      3.0517
## Hannan-Quinn 3.0587
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                      2.729 0.09857
## Lag[2*(p+q)+(p+q)-1][5]     4.050 0.05853
## Lag[4*(p+q)+(p+q)-1][9]     5.207 0.40774
## d.o.f=2
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                    0.01393  0.9060
## Lag[2*(p+q)+(p+q)-1][5]   0.73139  0.9168
## Lag[4*(p+q)+(p+q)-1][9]   1.89085  0.9176
## d.o.f=2
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[3]   0.06246 0.500 2.000  0.8027
## ARCH Lag[5]   1.04277 1.440 1.667  0.7203
## ARCH Lag[7]   1.77734 2.315 1.543  0.7643
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  1.9632
```

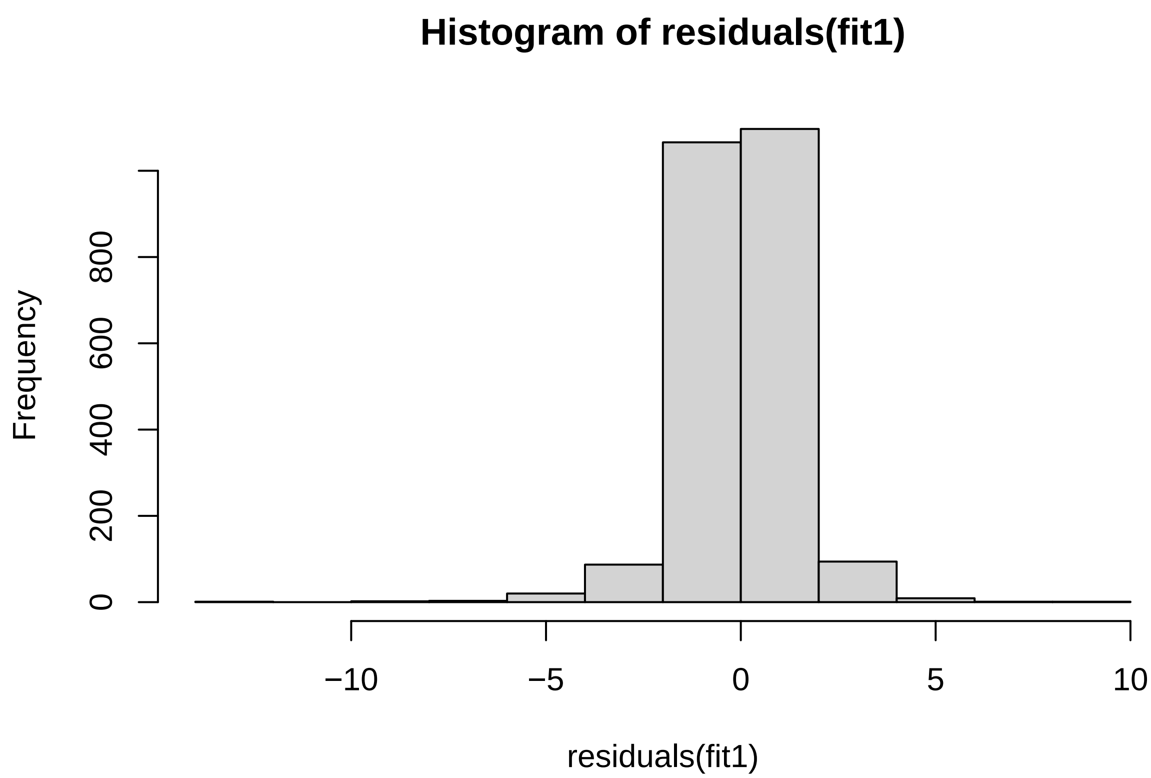
```
## Individual Statistics:
## mu     0.15028
## ar1    0.06930
## ma1    0.06287
## omega  1.27737
## alpha1 0.92660
## beta1  1.06054
## skew   0.15723
## shape  0.14539
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          1.89 2.11 2.59
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                      t-value     prob sig
## Sign Bias           2.23398 0.025577  **
## Negative Sign Bias  0.06182 0.950715
## Positive Sign Bias  0.86171 0.388934
## Joint Effect       15.33648 0.001551 ***
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##   group statistic p-value(g-1)
## 1    20     8.996       0.9735
## 2    30    29.823       0.4229
## 3    40    34.976       0.6540
## 4    50    45.102       0.6319
##
##
## Elapsed time : 0.4825718
```

Comments: AICc is 3.0517. Weighted Ljung-Box Test on Standardized Residuals and Weighted Ljung-Box Test on Standardized Squared Residuals indicate that there is no serial correlation. Adjusted Pearson Goodness-of-Fit Test indicate that the distribution model follows a jsu distribution. Weighted ARCH LM Tests indicate that no additional ARCH terms are needed.

```
ggtsdisplay(residuals(fit1))
```




```
hist(residuals(fit1))
```



Model 2: gjrARCH, garchOrder(1,1), armaOrder(2,2), distribution model=jsu

```
spec2 <- ugarchspec(
    variance.model = list(model = "gjrGARCH",
                          garchOrder = c(1, 1),
                          external.regressors = NULL),
    mean.model = list(armaOrder = c(2, 2),
                      include.mean = TRUE,
                      archm = FALSE, archpow = 1,
                      external.regressors = NULL),
    distribution.model = "jsu")

fit2 <- ugarchfit(spec2,
                  psei_train,
                  out.sample = 250, solver = "hybrid")
fit2
```

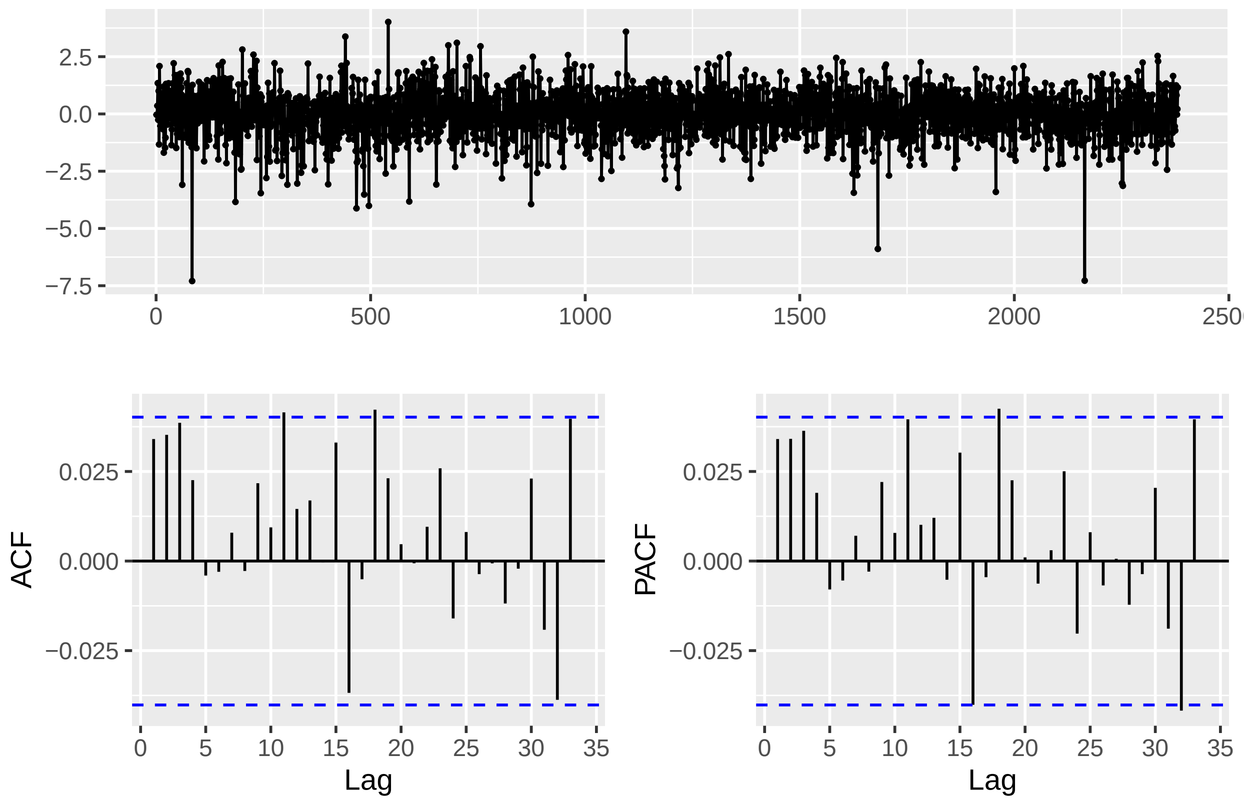
```
##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : gjrGARCH(1,1)
## Mean Model   : ARFIMA(2,0,2)
```

```
## Distribution : jsu
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.055809    0.019784  2.82096 0.004788
## ar1     0.855422    0.212177  4.03165 0.000055
## ar2    -0.200581    0.186200 -1.07723 0.281376
## ma1    -0.760345    0.217569 -3.49473 0.000475
## ma2     0.079481    0.189259  0.41996 0.674515
## omega   0.069384    0.016981  4.08608 0.000044
## alpha1  0.082133    0.019840  4.13972 0.000035
## beta1   0.823156    0.024138 34.10218 0.000000
## gamma1  0.099344    0.029380  3.38132 0.000721
## skew   -0.412506    0.111140 -3.71160 0.000206
## shape   1.973300    0.156743 12.58937 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.055809    0.022061  2.52977 0.011414
## ar1     0.855422    0.268644  3.18421 0.001451
## ar2    -0.200581    0.229208 -0.87510 0.381517
## ma1    -0.760345    0.277669 -2.73831 0.006176
## ma2     0.079481    0.233215  0.34081 0.733250
## omega   0.069384    0.022543  3.07785 0.002085
## alpha1  0.082133    0.019222  4.27288 0.000019
## beta1   0.823156    0.028134 29.25828 0.000000
## gamma1  0.099344    0.037640  2.63931 0.008308
## skew   -0.412506    0.110654 -3.72788 0.000193
## shape   1.973300    0.174078 11.33575 0.000000
##
## LogLikelihood : -3609.504
##
## Information Criteria
## ------------------------------------
##
## Akaike       3.0412
## Bayes        3.0678
## Shibata      3.0411
## Hannan-Quinn 3.0509
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                         statistic   p-value
## Lag[1]                      2.765 9.634e-02
## Lag[2*(p+q)+(p+q)-1][11]   10.080 6.653e-09
## Lag[4*(p+q)+(p+q)-1][19]   15.019 2.828e-02
## d.o.f=4
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                        statistic p-value
## Lag[1]                   0.02825  0.8665
```

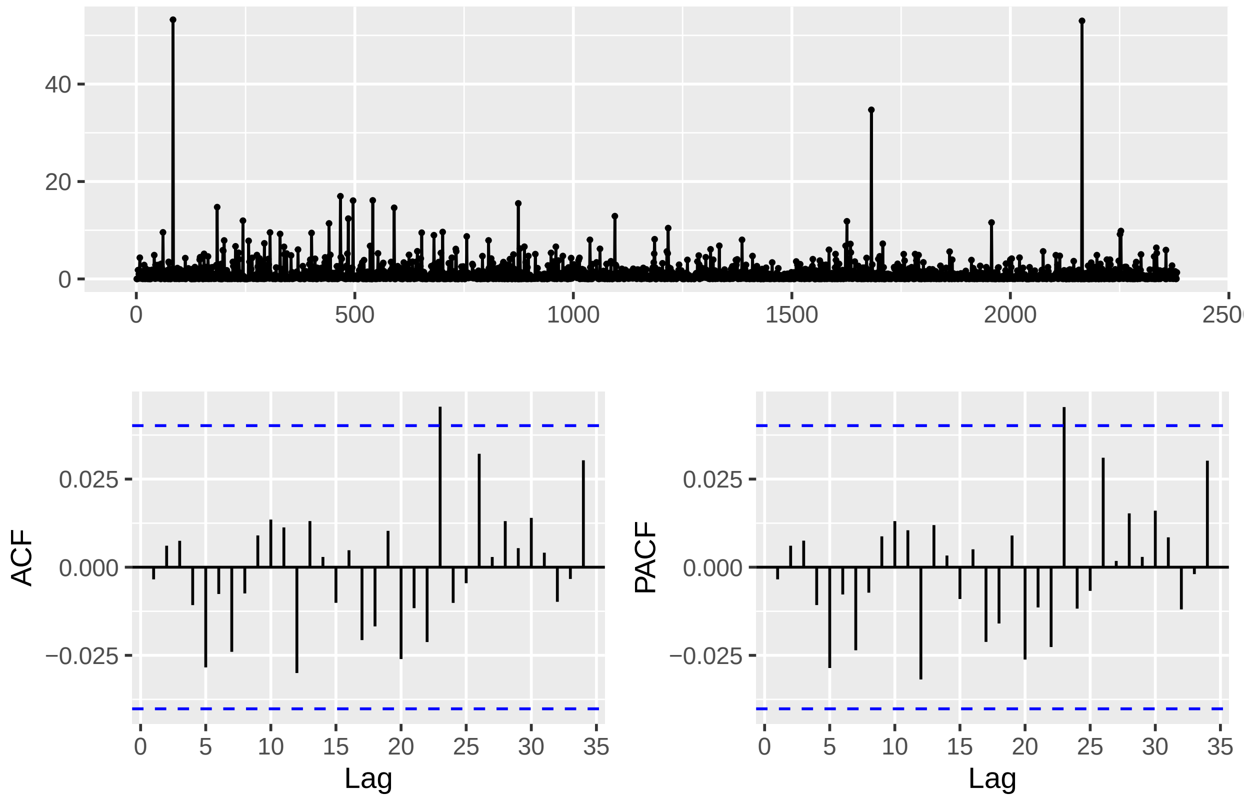
```
## Lag[2*(p+q)+(p+q)-1][5]   0.67573  0.9275
## Lag[4*(p+q)+(p+q)-1][9]   2.03792  0.9004
## d.o.f=2
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[3]    0.1337 0.500 2.000  0.7147
## ARCH Lag[5]    1.5077 1.440 1.667  0.5903
## ARCH Lag[7]    2.4094 2.315 1.543  0.6311
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  3.1681
## Individual Statistics:
## mu     0.3112
## ar1    0.2524
## ar2    0.2365
## ma1    0.2672
## ma2    0.2320
## omega  1.9338
## alpha1 1.3848
## beta1  1.4335
## gamma1 0.7407
## skew   0.3459
## shape  0.2592
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          2.49 2.75 3.27
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                    t-value    prob sig
## Sign Bias           0.8258 0.40899
## Negative Sign Bias  0.1100 0.91242
## Positive Sign Bias  1.3271 0.18460
## Joint Effect        6.7501 0.08031   *
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##   group statistic p-value(g-1)
## 1    20     7.366       0.9920
## 2    30    21.684       0.8330
## 3    40    31.012       0.8154
## 4    50    43.003       0.7137
##
##
## Elapsed time : 0.9003391
```

Comments: AICc is 3.0411 which is lower than Model 1 (model=sGarch, garchOrder(1,1), armaOrder(1,1), distribution model=jsu) but the Weighted Ljung-Box Test on Standardized Residuals indicates that there is serial correlation so model 1 is still preferable.

```r
fit2 %>% residuals(standardize = TRUE) %>% ggtsdisplay()
```




```r
residuals(fit2, standardize = TRUE)^2 %>% ggtsdisplay()
```



Model 3: gjrGarch, garchOrder(2,2),armaOrder(1,1), distribution model=jsu

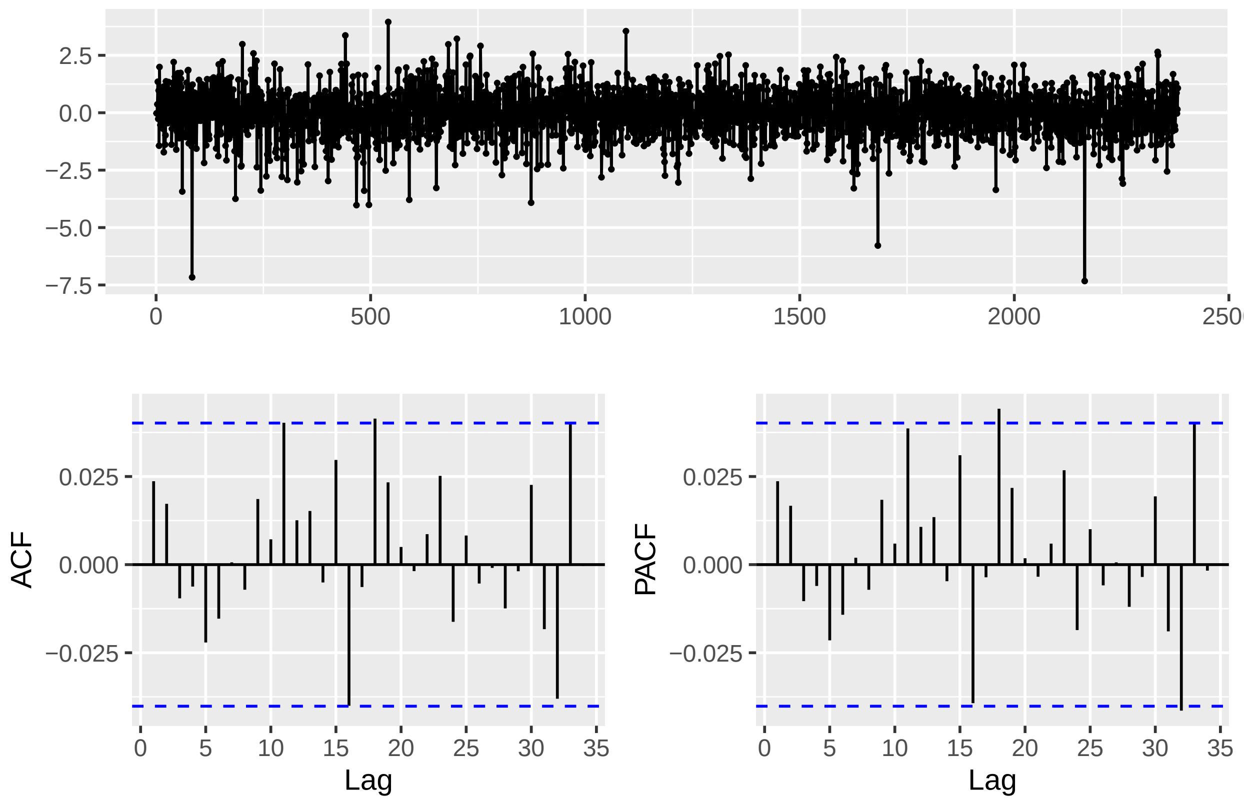
```
spec3 <- ugarchspec(
    variance.model = list(model = "gjrGARCH",
                          garchOrder = c(2,2),
                          external.regressors = NULL),
    mean.model = list(armaOrder = c(1,1),
                      include.mean = TRUE,
                      archm = FALSE, archpow = 1,
                      external.regressors = NULL),
    distribution.model = "jsu")

fit3 <- ugarchfit(spec3,
                  psei_train,
                  out.sample = 250, solver = "hybrid")
fit3
```

```
##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : gjrGARCH(2,2)
## Mean Model   : ARFIMA(1,0,1)
## Distribution : jsu
```

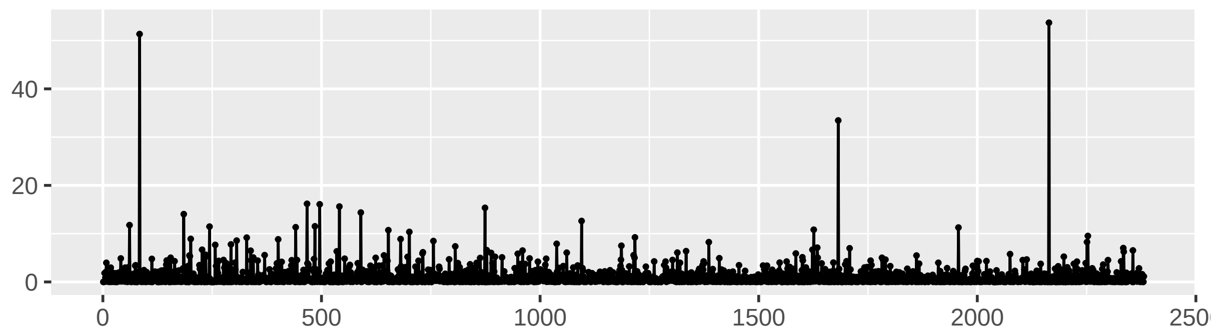
```
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.043584    0.022891  1.90396 0.056916
## ar1    -0.183776    0.143771 -1.27825 0.201161
## ma1     0.289954    0.139326  2.08112 0.037423
## omega   0.067885    0.018851  3.60114 0.000317
## alpha1  0.060333    0.035790  1.68574 0.091845
## alpha2  0.010317    0.031682  0.32566 0.744684
## beta1   0.826200    0.284355  2.90552 0.003666
## beta2   0.000000    0.223477  0.00000 1.000000
## gamma1  0.146482    0.038563  3.79847 0.000146
## gamma2 -0.026437    0.065026 -0.40655 0.684337
## skew   -0.438870    0.116081 -3.78073 0.000156
## shape   2.033143    0.164829 12.33488 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.043584    0.023733  1.83642 0.066296
## ar1    -0.183776    0.105568 -1.74082 0.081715
## ma1     0.289954    0.103856  2.79188 0.005240
## omega   0.067885    0.028515  2.38067 0.017281
## alpha1  0.060333    0.033315  1.81101 0.070140
## alpha2  0.010317    0.038297  0.26940 0.787620
## beta1   0.826200    0.327463  2.52303 0.011635
## beta2   0.000000    0.273022  0.00000 1.000000
## gamma1  0.146482    0.043441  3.37200 0.000746
## gamma2 -0.026437    0.045995 -0.57477 0.565446
## skew   -0.438870    0.114136 -3.84516 0.000120
## shape   2.033143    0.182831 11.12033 0.000000
##
## LogLikelihood : -3614.326
##
## Information Criteria
## ------------------------------------
##
## Akaike       3.0461
## Bayes        3.0752
## Shibata      3.0460
## Hannan-Quinn 3.0566
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                      1.335  0.2479
## Lag[2*(p+q)+(p+q)-1][5]     2.303  0.8706
## Lag[4*(p+q)+(p+q)-1][9]     3.212  0.8559
## d.o.f=2
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                          statistic p-value
```

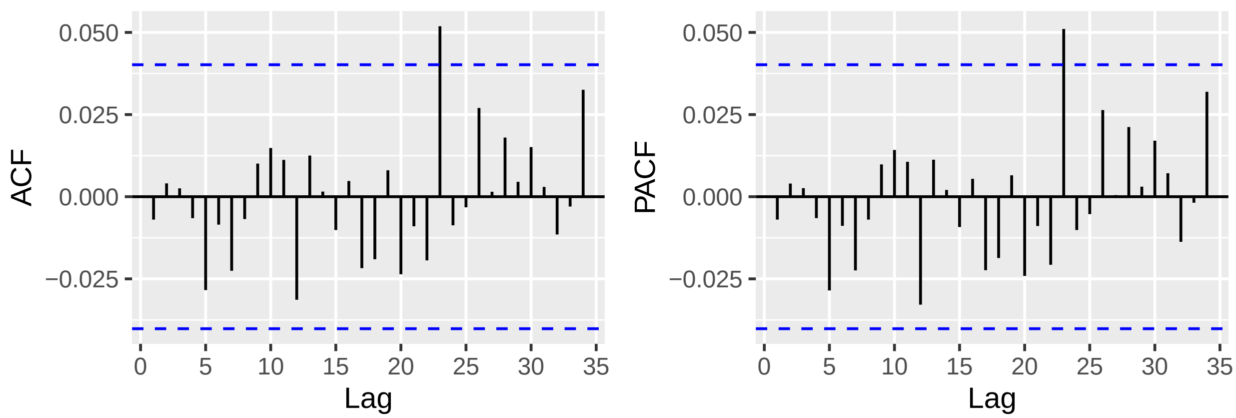
```
## Lag[1]                         0.1156  0.7339
## Lag[2*(p+q)+(p+q)-1][11]    2.3405  0.9355
## Lag[4*(p+q)+(p+q)-1][19]    4.8497  0.9470
## d.o.f=4
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[5]     1.923 0.500 2.000  0.1656
## ARCH Lag[7]     2.934 1.473 1.746  0.3305
## ARCH Lag[9]     3.224 2.402 1.619  0.5247
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  6.2518
## Individual Statistics:
## mu     0.30367
## ar1    0.08314
## ma1    0.07272
## omega  1.94950
## alpha1 1.37240
## alpha2 1.46081
## beta1  1.45709
## beta2  1.40623
## gamma1 0.77150
## gamma2 0.73026
## skew   0.34486
## shape  0.26515
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          2.69 2.96 3.51
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                    t-value   prob sig
## Sign Bias           0.7630 0.4456
## Negative Sign Bias  0.2539 0.7996
## Positive Sign Bias  1.2104 0.2263
## Joint Effect        4.7517 0.1909
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##   group statistic p-value(g-1)
## 1    20     12.09       0.8819
## 2    30     21.96       0.8218
## 3    40     24.29       0.9685
## 4    50     32.96       0.9617
##
##
## Elapsed time : 1.168362
```

```r
fit3 %>% residuals(standardize = TRUE) %>% ggtsdisplay()
```




```r
residuals(fit3, standardize = TRUE)^2 %>% ggtsdisplay()
```





Comments: AICc is lower versus model 1. Weighted Ljung-Box Test on Standardized Residuals and Weighted Ljung-Box Test on Standardized Squared Residuals indicate that there is no serial correlation. Weighted ARCH LM Tests results indicate that no additional ARCH terms are needed and Adjusted Pearson Goodness-of-Fit Test indicate that a jsu model is appropriate.

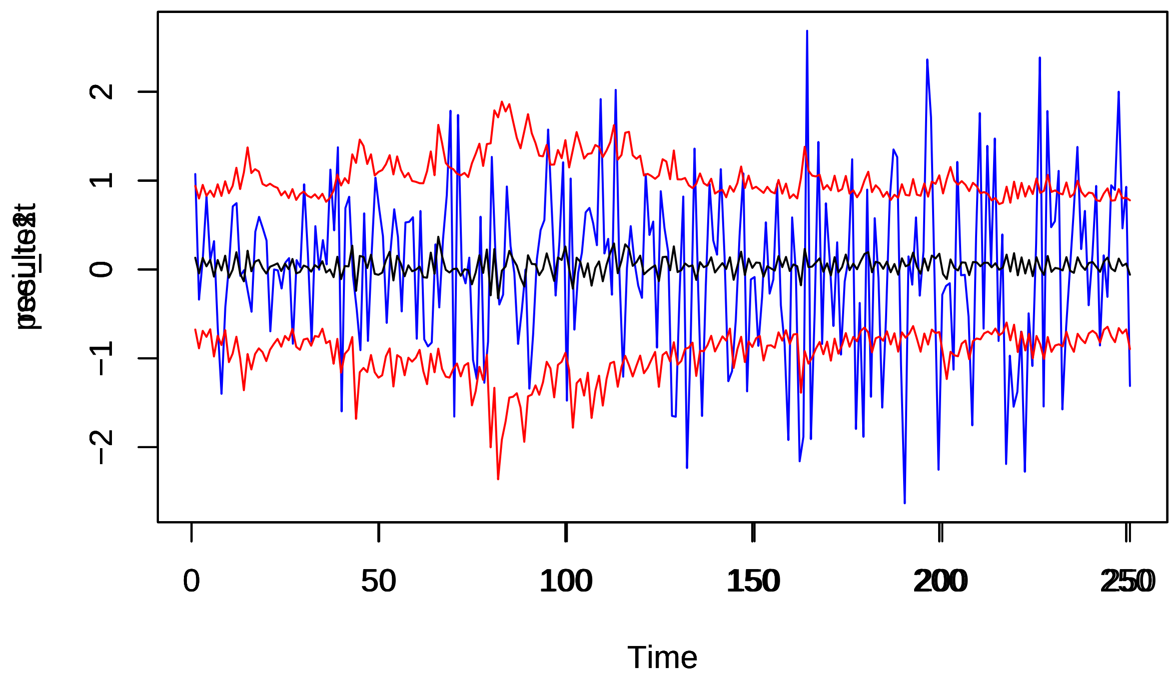
Since Model 3 (gjrGarch, garchOrder(2,2),armaOrder(1,1), distribution model=jsu) has the lowest information criteria (AICc) values, fulfills the GARCH assumptions and passes the residual diagnostic tests then it is recommended as an appropriate ARMA-GARCH model with the best possible distribution fit for the return series.

2) [1pt] Based on the model in (1), show a plot of the forecasted mean and variance of psei_ret for the test data added into the plot of the full dataset. Analyze the plots. (1)

Plot of the forecasted mean and variance of psei_ret for the test data:

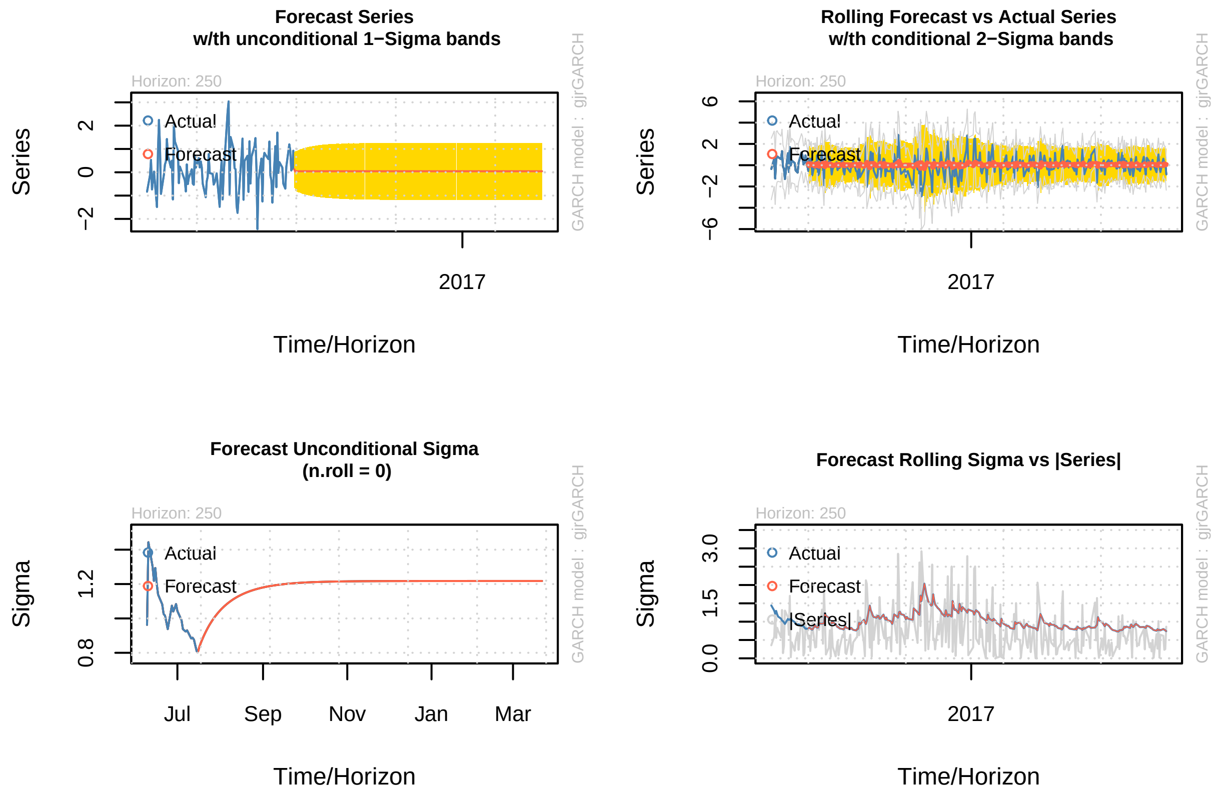
```
fcast3 <- ugarchforecast(fit3,
                         data = NULL,
                         n.ahead = 250,
                         n.roll = 250,
                         out.sample = 0,
                         external.forecasts = list(mregfor = NULL, vregfor = NULL))

results1 <- fcast3@forecast$seriesFor[1,] + fcast3@forecast$sigmaFor[1,]
results2 <- fcast3@forecast$seriesFor[1,] - fcast3@forecast$sigmaFor[1,]
```

```r
ylim <- c(min(psei_test), max(psei_test))
plot.ts(psei_test , col="blue", ylim=ylim)
par(new=TRUE)
fcast3@forecast$seriesFor[1,] %>% plot.ts(ylim=ylim)
par(new=TRUE)
plot.ts(results1, col="red", ylim=ylim)
par(new=TRUE)
plot.ts(results2, col="red", ylim=ylim)
```



Plot of the dataset with the forecasts:

```r
fcast3 <- ugarchforecast(fit3,
                         data = NULL,
                         n.ahead = 250,
                         n.roll = 250,
                         out.sample = 0,
                         external.forecasts = list(mregfor = NULL, vregfor = NULL))

plot(fcast3, which="all")
```



The forecast values are flat as compared to the actual values but the volatility seems to be captured well when the conditional 2 sigma bands are used.

3) [1pt] Generate the accuracy measures of the selected model in (1) with respect to the testing dataset. Write a short analysis based on the accuracy measures.

```
fit3 <- ugarchfit(spec3,
                  psei_train,
                  out.sample = 250, solver = "hybrid")

fcast3 <- ugarchforecast(fit3,
                         data = NULL,
                         n.ahead = 250,
                         n.roll = 0,
                         out.sample = 0,
                         external.forecasts = list(mregfor = NULL, vregfor = NULL))
```

Calculating for the MSE

```
MSE <- mean((psei_test[,1]-fitted(fcast3))^2)
MSE
```

```
## [1] 0.9179808
```

Calculating for the MAE

```
MAE <- mean(abs((psei_test[,1]-fitted(fcast3))))
MAE
```

```
## [1] 0.744124
```

The MSE and MAE of the test data set are low since especially since we are looking at a range of values from -13.09 to 9.36 of the PSEI log returns.This means that even if we visually saw that the forecast is lower and more flat than the actual test data set in item 2, the errors are still acceptable.