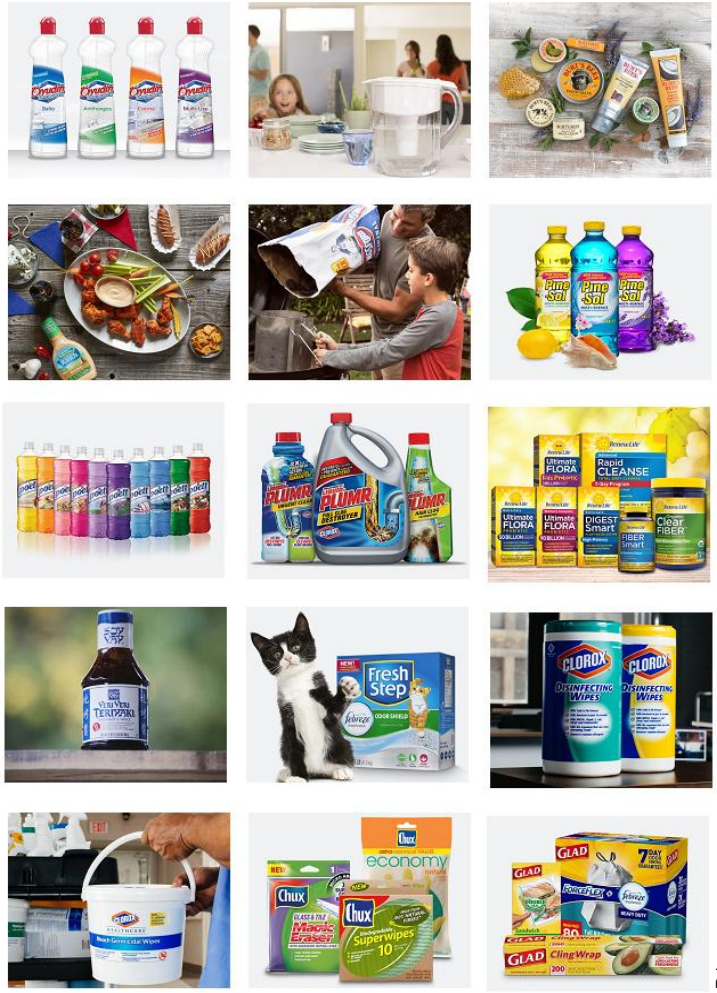
Time Series Group Project

By Michael Corcoran, Susan Hsieh,
Eric Nechayev, Chixiang Qiu, Yujia Zhou

Outline

1. Introduction to Clorox
2. ARMA model
3. GARCH model
4. SARIMA model
5. Forecasting



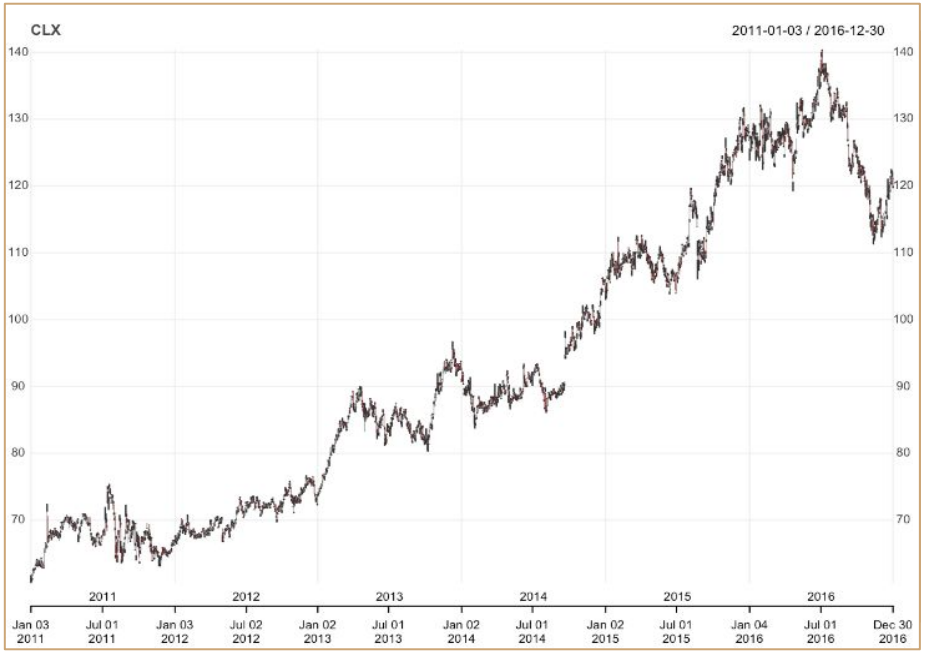
Why Clorox Stock (CLX)

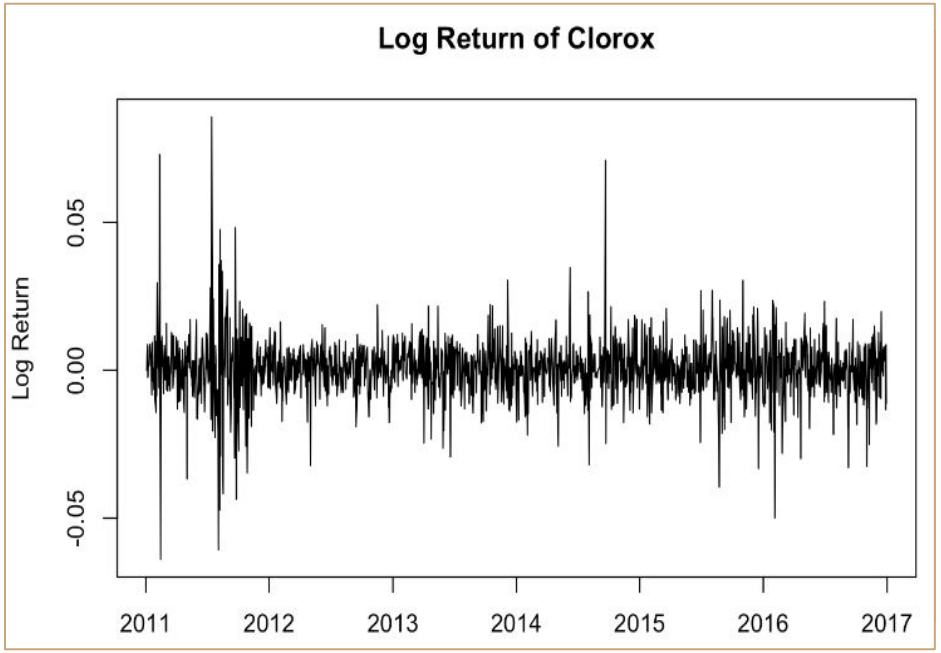
- **The background of Clorox :**
  A leading multinational manufacturer and professional products around the world
- **Portfolio of Clorox :**
  Sales of Brands hold the No. 1 or No. 2 market share positions in their categories
- **The recent performance of Clorox :**
  The company’s continued investments in product innovation and efficient marketing have led to volume expansions.

The packages of R

- **quantmod :** uantitative Financial Modelling Framework
- **urca :** Unit Root and Cointegration Tests for Time Series Data
- **tseries :**Time Series Analysis and Computational Finance
- **forecast :** Forecasting Functions for Time Series and Linear Models
- **rugarch :** Univariate GARCH Models
- **ggplot2 :** Create Elegant Data Visualisations Using the Grammar of Graphics

Stock”CLX” from 2011 to 2016

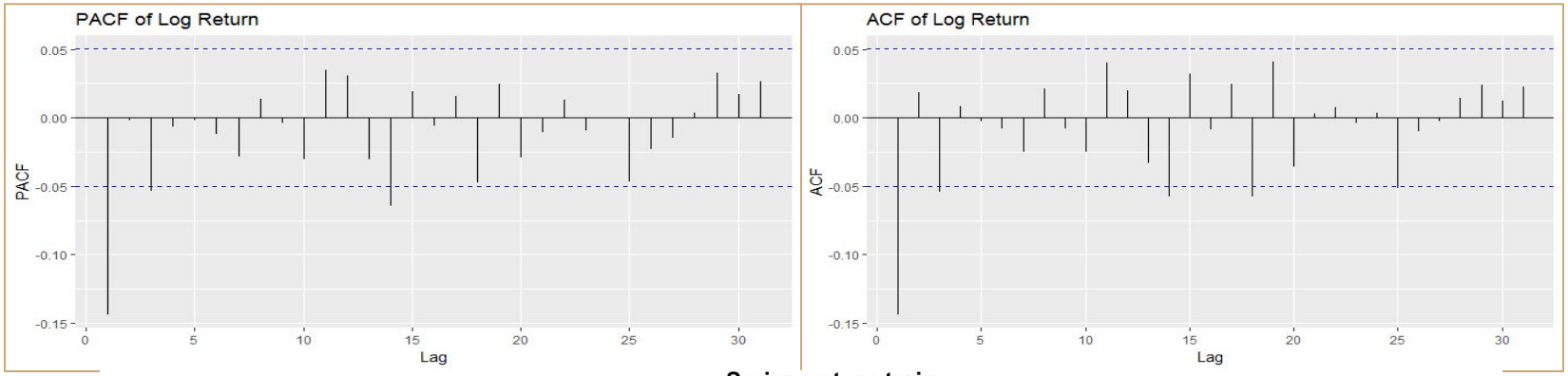


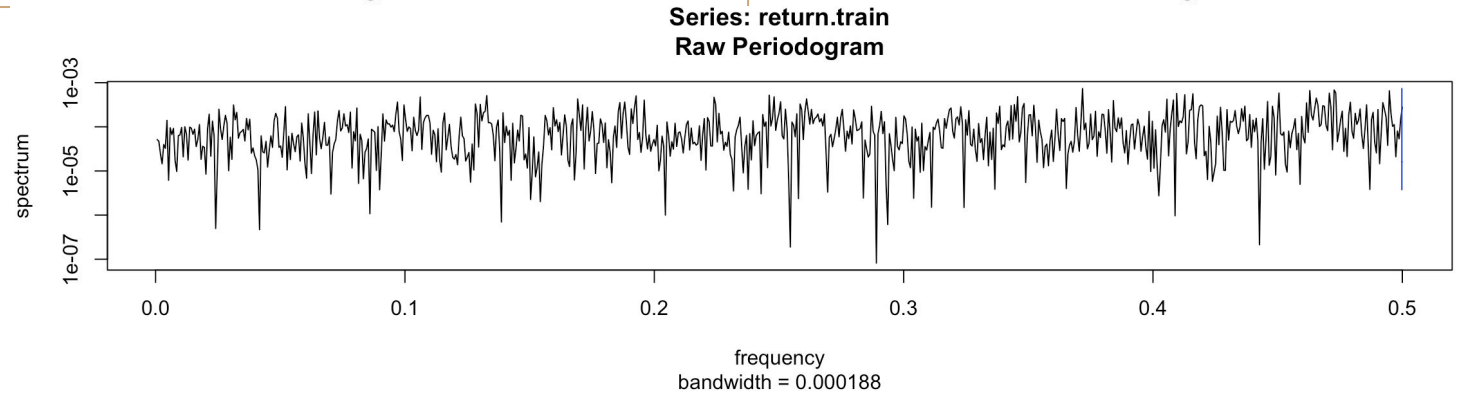


Closing price

Daily Log return

Seasonality: ACF, PACF and the Periodogram





ADF Test of Log Returns

The model of ADF test

$$\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \delta_1 \Delta X_{t-1} + \cdots + \delta_{p-1} \Delta X_{t-p-1} + \varepsilon_t$$

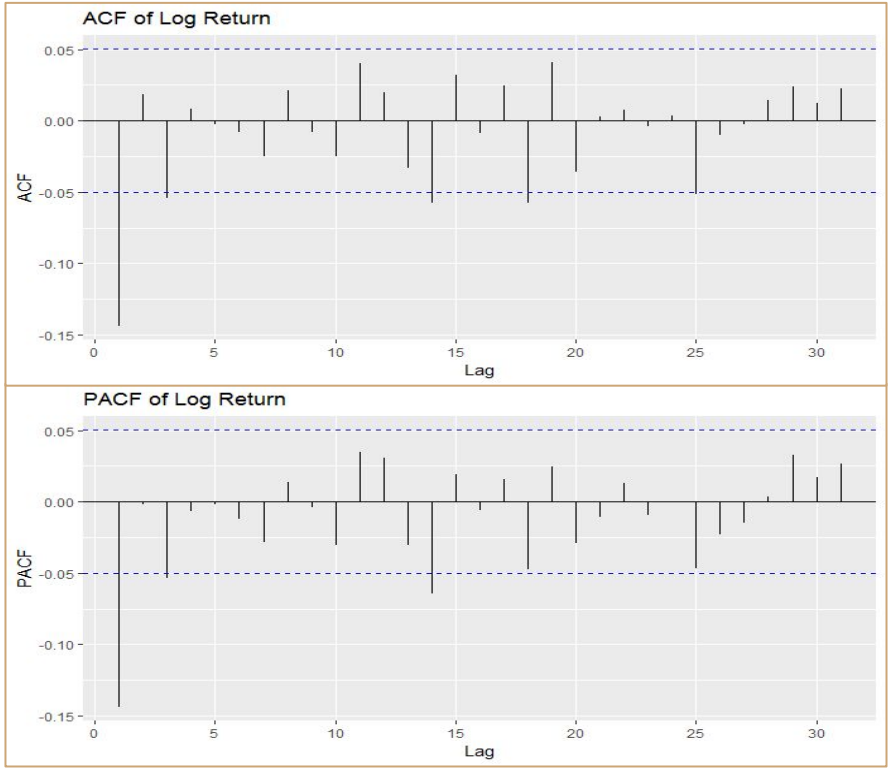
$$\Delta X_t = \gamma X_{t-1} + \varepsilon_t$$

```
Test regression trend


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

Residuals:
      Min        1Q    Median        3Q       Max
-0.064267 -0.005083  0.000210  0.005572  0.083904

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.310e-04  5.534e-04   0.959    0.337
z.lag.1     -1.149e+00  3.934e-02 -29.206   <2e-16 ***
tt          -9.369e-08  6.299e-07  -0.149    0.882
z.diff.lag   2.594e-03  2.600e-02   0.100    0.921
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0104 on 1480 degrees of freedom
Multiple R-squared:  0.5731, Adjusted R-squared:  0.5722
F-statistic: 662.2 on 3 and 1480 DF,  p-value: < 2.2e-16


Value of test-statistic is: -29.2057 284.3245 426.4856

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

Both of ACF and PACF cut off at lag1.

We will fit the ARMA (0,1), ARMA (1,0), and ARMA(1,1) models.

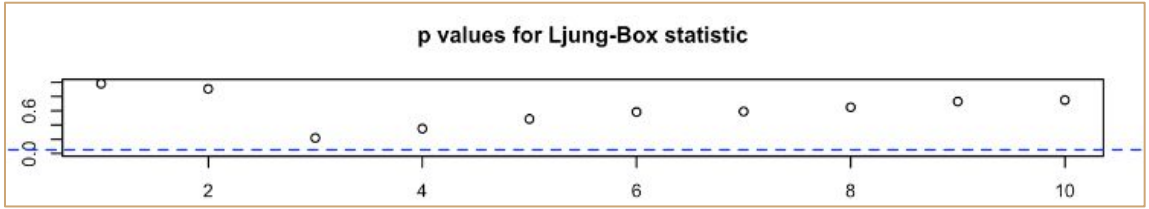


ARMA(0,1)

$$y_t = 0.0004\varepsilon_t - 0.1447\varepsilon_{t-1}$$

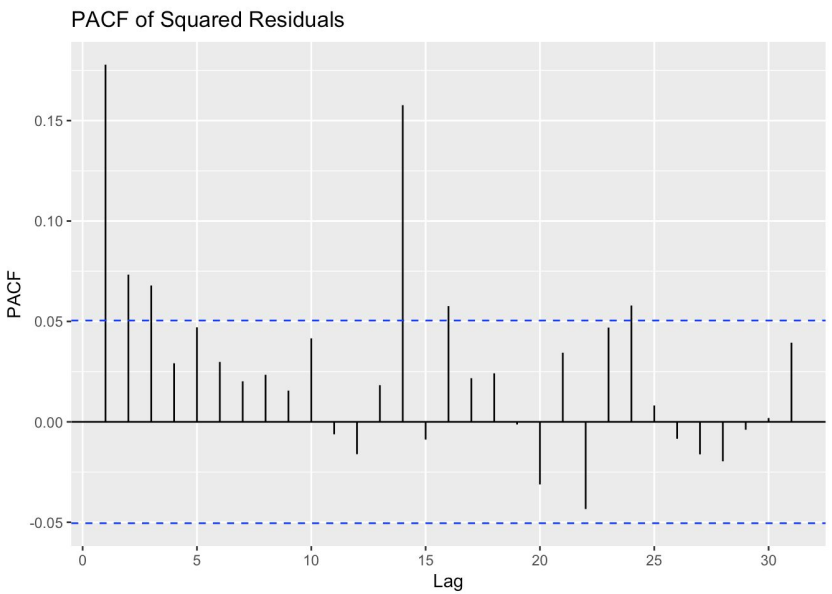
```
arima(x = return.train, order = c(i, 0, j))

Coefficients:
          ma1  intercept
      -0.1447      4e-04
s.e.   0.0256      2e-04
```

- Log likelihood = 4753.49
- AIC = -9500.98
- BIC = -9485.022




```
Shapiro-Wilk normality test

data:  res_0
W = 0.91838, p-value < 2.2e-16
```

Squared Residuals

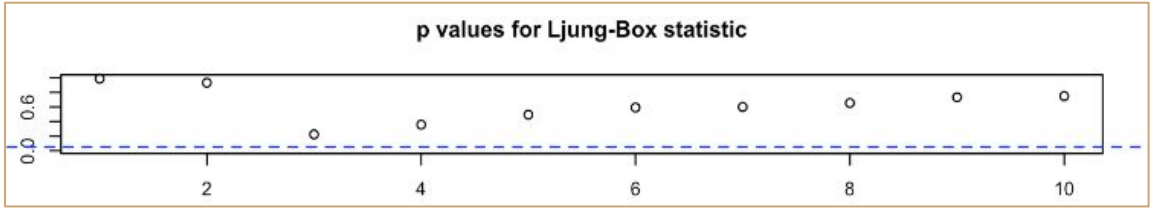


ARMA(1,0)

$$y_t = 0.0004 - 0.1439 y_{t-1} + \varepsilon_t$$

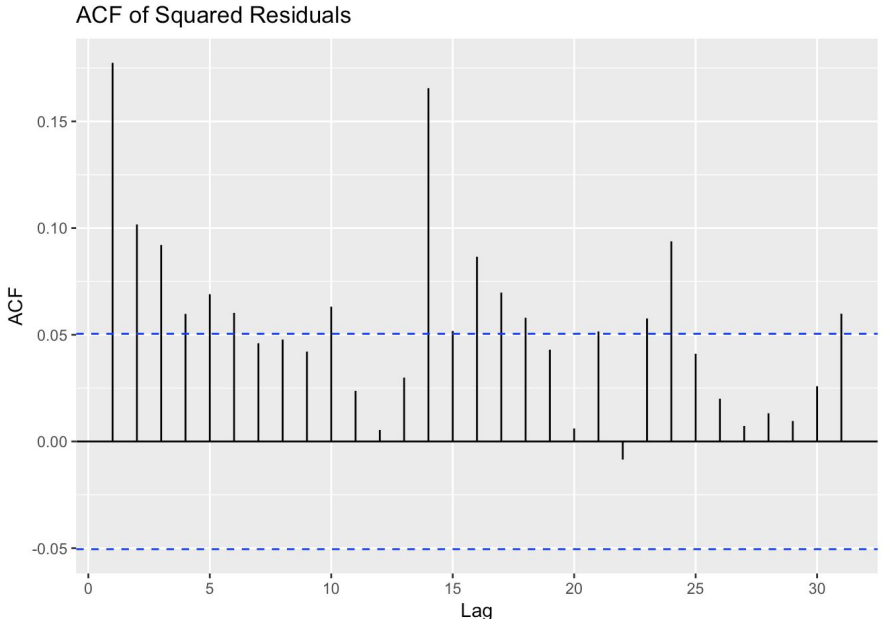
```
arima(x = return.train, order = c(p, 0, q))

Coefficients:
          ar1  intercept
      -0.1439      4e-04
s.e.   0.0255      2e-04
```

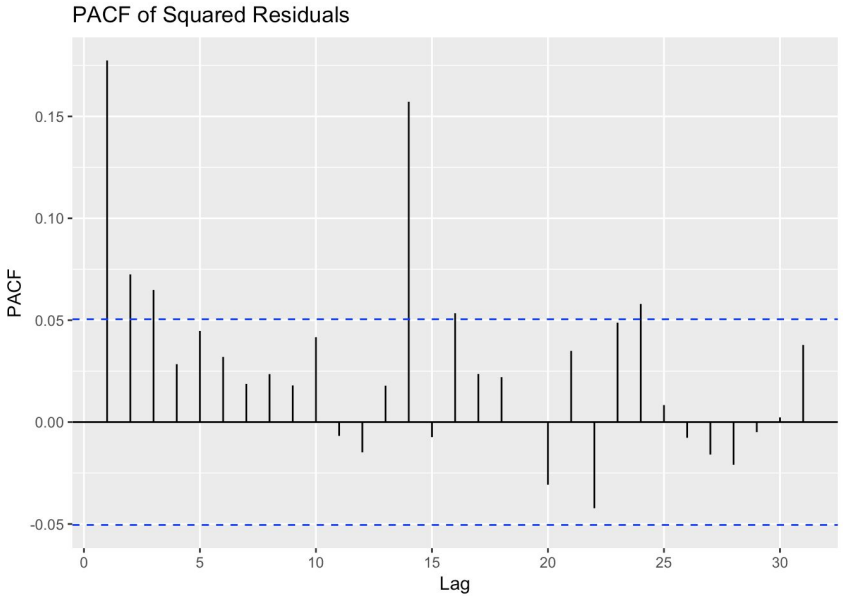
- Loglikelihood = 4753.51
- AIC = -9501.03
- BIC = -9485.071




```
 Shapiro-Wilk normality test

data:  res_0
W = 0.91892, p-value < 2.2e-16
```

Squared Residuals



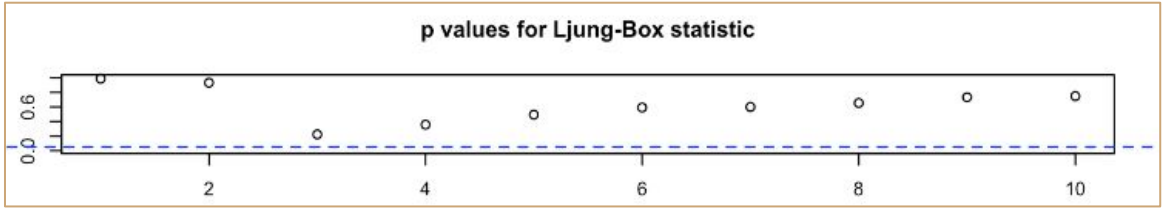


ARMA(1,1)

$$y_t = 0.0004 - 0.1439 y_{t-1} + \varepsilon_t - 0.0719 \varepsilon_{t-1}$$

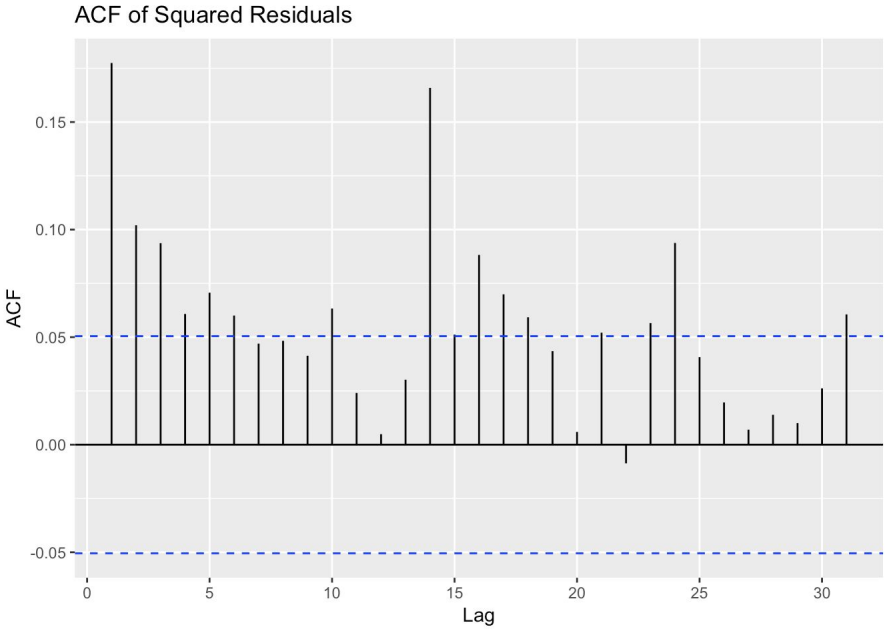
```
arima(x = return.train, order = c(p, 0, q))

Coefficients:
          ar1      ma1  intercept
      -0.0729  -0.0719      4e-04
s.e.   0.3493   0.3518      2e-04
```

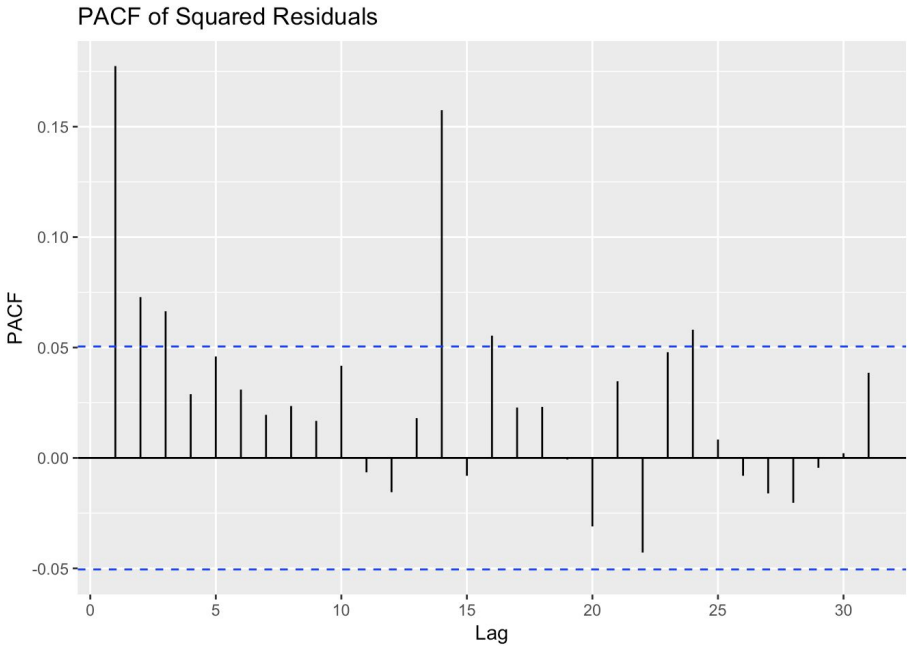
- Loglikelihood = 4753.52
- AIC = -9499.05
- BIC = -9477.771




```
 Shapiro-Wilk normality test

data:  res_0
W = 0.91867, p-value < 2.2e-16
```

Squared Residuals





Summary of ARMA Models

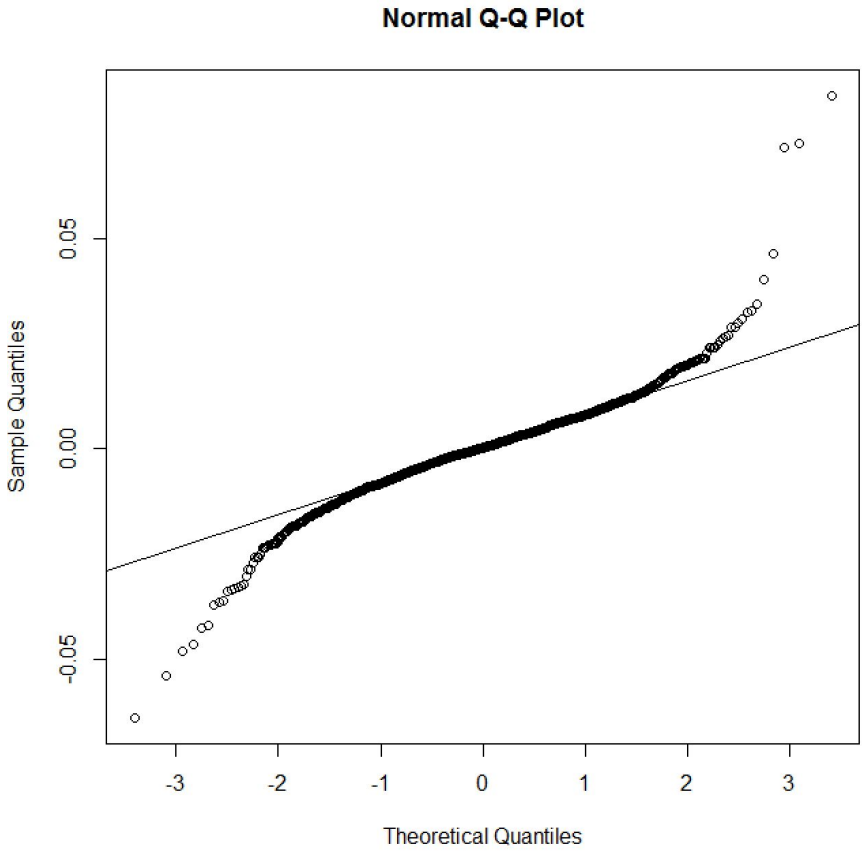
N=1509	AIC	BIC	Log likelihood
MA(1)	-9500.98	-9485.02	4753.49
AR(1)	**-9501.03**	**-9485.07**	4753.51
ARMA(1,1)	-9499.05	-9477.77	4753.52

$$X_t = -0.1439X_{t-1} + \varepsilon_t$$

Testing Residuals

```
        Shapiro-Wilk normality test

data:  res
W = 0.91998, p-value < 2.2e-16
```

This means the residuals are not normally distributed



GARCH

Above we found the best fit model assuming that the variance is a Gaussian White Noise but the variance could follow a different form of White Noise. We will look at the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model to model the variance. As shown before, residuals are not normally distributed. Therefore, we use t-distribution.

We will use the ARMA (0,1), ARMA (1,0), and ARMA(1,1) models to fit the mean and the GARCH model to fit the variance.

ARMA(0,1)+GARCH(1,1)

AIC

	Order ARCH1	Order ARCH2	Order ARCH3	Order ARCH4	Order ARCH5
Order GARCH0	-9824.072	-9822.567	-9820.578	-9818.441	-9816.592
Order GARCH1	-9914.639	-9912.635	-9911.615	-9909.686	-9908.989
Order GARCH2	-9912.432	-9910.635	-9910.093	-9908.615	-9907.582
Order GARCH3	-9910.280	-9908.499	-9908.093	-9906.612	-9905.582
Order GARCH4	-9908.187	-9906.406	-9905.972	-9904.728	-9903.582
Order GARCH5	-9905.986	-9904.209	-9903.786	-9902.568	-9901.582

BIC

	Order ARCH1	Order ARCH2	Order ARCH3	Order ARCH4	Order ARCH5
Order GARCH0	-9797.476	-9790.652	-9783.343	-9775.888	-9768.719
Order GARCH1	-9882.723	-9875.400	-9869.062	-9861.814	-9855.797
Order GARCH2	-9875.197	-9868.082	-9862.221	-9855.423	-9849.070
Order GARCH3	-9867.726	-9860.627	-9854.901	-9848.100	-9841.751
Order GARCH4	-9860.315	-9853.214	-9847.461	-9840.897	-9834.432
Order GARCH5	-9852.794	-9845.698	-9839.956	-9833.418	-9827.113

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-6.429171	-6.411546	-6.429193	-6.422607

```
Title:
 GARCH Modelling

Call:
 garchFit(formula = ~arma(0, 1) + garch(1, 1), data = r, trace = F)

Mean and Variance Equation:
 data ~ arma(0, 1) + garch(1, 1)
<environment: 0x1044b7ac>
 [data = r]

Conditional Distribution:
 norm

Coefficient(s):
        mu          ma1        omega       alpha1        beta1
 5.7059e-04  -8.5999e-02   1.3333e-05   1.5095e-01   7.3027e-01

Std. Errors:
 based on Hessian

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu     5.706e-04   2.161e-04    2.641  0.00827 **
ma1   -8.600e-02   2.996e-02   -2.870  0.00410 **
omega  1.333e-05   2.957e-06    4.508 6.53e-06 ***
alpha1 1.510e-01   2.711e-02    5.569 2.56e-08 ***
beta1  7.303e-01   4.418e-02   16.529  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 4855.809    normalized:  3.217899

Description:
 Sun Mar 26 19:31:23 2017 by user: mcorc


Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  3681.125  0
 Shapiro-Wilk Test  R    W      0.9403947 0
 Ljung-Box Test     R    Q(10)  3.622732  0.9627659
 Ljung-Box Test     R    Q(15)  10.96386  0.755152
 Ljung-Box Test     R    Q(20)  14.75533  0.790232
 Ljung-Box Test     R^2  Q(10)  2.169872  0.9948646
 Ljung-Box Test     R^2  Q(15)  6.43044   0.9715756
```

ARMA(1,0)+GARCH(1,1)

AIC

	Order ARCH1	Order ARCH2	Order ARCH3	Order ARCH4	Order ARCH5
Order GARCH0	-9822.900	-9820.882	-9818.908	-9816.862	-9814.418
Order GARCH1	-9914.260	-9912.258	-9911.275	-9909.349	-9908.694
Order GARCH2	-9912.049	-9910.259	-9909.764	-9908.302	-9907.291
Order GARCH3	-9909.898	-9908.122	-9907.764	-9906.302	-9905.291
Order GARCH4	-9907.801	-9906.027	-9905.641	-9904.424	-9903.291
Order GARCH5	-9905.597	-9903.828	-9903.453	-9902.356	-9901.291

BIC

	Order ARCH1	Order ARCH2	Order ARCH3	Order ARCH4	Order ARCH5
Order GARCH0	-9796.304	-9788.967	-9781.674	-9774.308	-9766.545
Order GARCH1	-9882.345	-9875.024	-9868.722	-9861.476	-9855.502
Order GARCH2	-9874.815	-9867.705	-9861.891	-9855.110	-9848.780
Order GARCH3	-9867.344	-9860.249	-9854.572	-9847.791	-9841.461
Order GARCH4	-9859.928	-9852.835	-9847.130	-9840.594	-9834.142
Order GARCH5	-9852.405	-9845.316	-9839.622	-9833.206	-9826.822

```
Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
-6.429315 -6.411690 -6.429337 -6.422751
```

```
Title:
 GARCH Modelling

Call:
 garchFit(formula = ~arma(1, 0) + garch(1, 1), data = r, trace = F)

Mean and Variance Equation:
 data ~ arma(1, 0) + garch(1, 1)
<environment: 0x0c58b670>
 [data = r]

Conditional Distribution:
 norm

Coefficient(s):
         mu           ar1         omega        alpha1         beta1
 6.2175e-04  -8.8023e-02   1.3228e-05   1.5107e-01   7.3127e-01

Std. Errors:
 based on Hessian

Error Analysis:
         Estimate  Std. Error  t value Pr(>|t|)
mu       6.217e-04   2.369e-04    2.624  0.00868 **
ar1     -8.802e-02   3.027e-02   -2.908  0.00364 **
omega    1.323e-05   2.930e-06    4.514 6.36e-06 ***
alpha1   1.511e-01   2.708e-02    5.579 2.41e-08 ***
beta1    7.313e-01   4.388e-02   16.666  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 4855.918    normalized:  3.217971

Description:
 Sun Mar 26 19:23:28 2017 by user: mcorc


Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  3667.235  0
 Shapiro-Wilk Test  R    W      0.9405165 0
 Ljung-Box Test     R    Q(10)  3.855672  0.953623
 Ljung-Box Test     R    Q(15)  11.2564   0.7342155
 Ljung-Box Test     R    Q(20)  15.02967  0.7747073
 Ljung-Box Test     R^2  Q(10)  2.184716  0.9947183
 Ljung-Box Test     R^2  Q(15)  6.366755  0.9729069
```

ARMA(1,1)+GARCH(1,1)

AIC

	Order ARCH1	Order ARCH2	Order ARCH3	Order ARCH4	Order ARCH5
Order GARCH0	-9831.547	-9829.861	-9828.357	-9826.189	-9823.700
Order GARCH1	-9915.632	-9913.608	-9912.664	-9910.787	-9909.962
Order GARCH2	-9913.448	-9911.607	-9911.381	-9909.994	-9908.676
Order GARCH3	-9911.319	-9909.490	-9909.381	-9907.994	-9906.676
Order GARCH4	-9909.232	-9907.405	-9907.266	-9906.004	-9904.676
Order GARCH5	-9907.066	-9905.242	-9905.116	-9903.598	-9902.676

BIC

	Order ARCH1	Order ARCH2	Order ARCH3	Order ARCH4	Order ARCH5
Order GARCH0	-9799.632	-9792.627	-9785.803	-9778.316	-9770.508
Order GARCH1	-9878.397	-9871.054	-9864.791	-9857.595	-9851.451
Order GARCH2	-9870.895	-9863.734	-9858.189	-9851.482	-9844.845
Order GARCH3	-9863.446	-9856.297	-9850.869	-9844.163	-9837.526
Order GARCH4	-9856.040	-9848.893	-9843.436	-9836.854	-9830.207
Order GARCH5	-9848.554	-9841.412	-9835.966	-9829.129	-9822.888

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-6.428143	-6.406994	-6.428175	-6.420267

```
 GARCH Modelling

Call:
 garchFit(formula = ~arma(1, 1) + garch(1, 1), data = r, trace = F)

Mean and Variance Equation:
 data ~ arma(1, 1) + garch(1, 1)
<environment: 0x0ffe1a64>
 [data = r]

Conditional Distribution:
 norm

Coefficient(s):
         mu          ar1          ma1        omega       alpha1        beta1
7.2816e-04  -2.7190e-01   1.8587e-01   1.2981e-05   1.5126e-01   7.3373e-01

td. Errors:
based on Hessian

rror Analysis:
         Estimate  Std. Error  t value Pr(>|t|)
u       7.282e-04   3.395e-04    2.145    0.032 *
r1     -2.719e-01   3.344e-01   -0.813    0.416
a1      1.859e-01   3.420e-01    0.544    0.587
omega   1.298e-05   2.905e-06    4.468 7.89e-06 ***
alpha1  1.513e-01   2.698e-02    5.606 2.06e-08 ***
beta1   7.337e-01   4.340e-02   16.908  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 4856.034    normalized:  3.218048

Description:
 Sun Mar 26 19:36:40 2017 by user: mcorc


Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  3628.739  0
 Shapiro-Wilk Test  R    W      0.9407641 0
 Ljung-Box Test     R    Q(10)  4.384542  0.9283369
 Ljung-Box Test     R    Q(15)  11.87416  0.6885278
 Ljung-Box Test     R    Q(20)  15.67123  0.7367995
 Ljung-Box Test     R^2  Q(10)  2.218191  0.9943778
 Ljung-Box Test     R^2  Q(15)  6.23237   0.9755745
```

Summary of GARCH Models

N=1509	AIC	BIC	Log Likelihood
ARMA(0,1) + GARCH(1,1)	-9701.62	-9675.02	4855.81
ARMA(1,0) + GARCH(1,1)	**-9701.84**	**-9675.24**	4855.92
ARMA(1,1) + GARCH(1,1)	-9700.07	-9668.15	4856.03

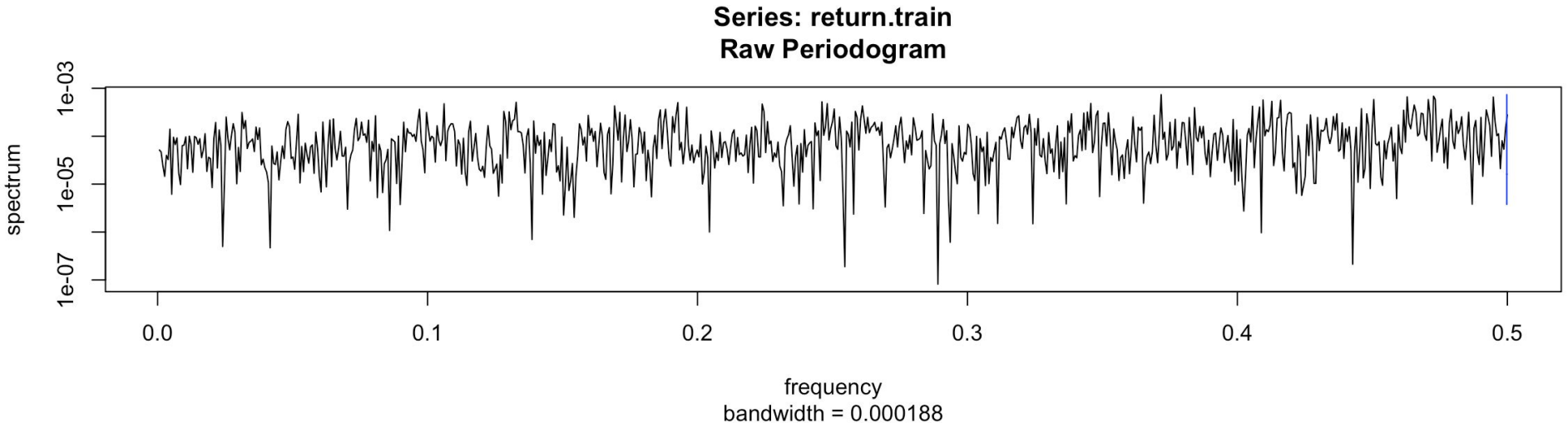
$$X_t = .0006217 - .08802X_{t-1} + o_t e_t$$
$$o^2_t = .1511x^2_{t-1} + .7313o^2_{t-1}$$

SARIMA

So far, we have restricted our attention to non-seasonal ARIMA models. However, ARIMA models are also capable of modelling a wide range of seasonal data. A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA models we have seen so far.

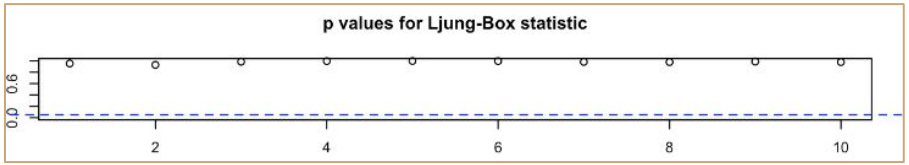
We will analyze the AR(1), MA(1), and ARMA(1,1) models in three different scenarios separately.

Period


Series: return.train
Raw Periodogram
spectrum
1e-03
1e-05
1e-07
0.0
0.1
0.2
0.3
0.4
0.5
frequency
bandwidth = 0.000188

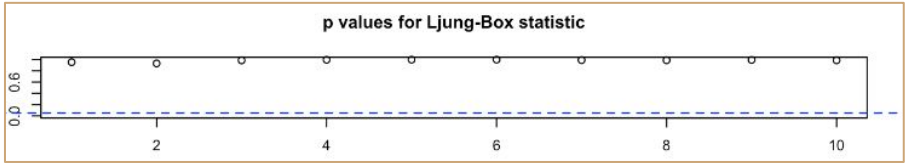
SARIMA(1,0,0)

1. SARIMA(1,0,0)(1,0,0)[3] $(1-\alpha B^3)(1-\beta B)y_t = \varepsilon_t$



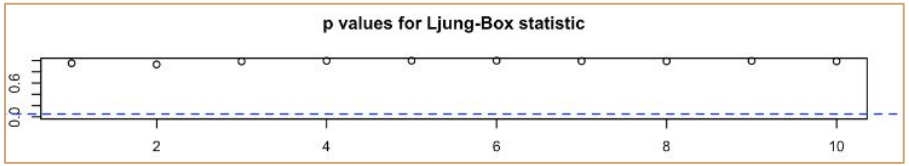

```
 Shapiro-Wilk normality test

data:  res_2
W = 0.9185, p-value < 2.2e-16
```

2. SARIMA(1,0,0)(0,0,1)[3] $(1-\beta B)y_t = (1-\theta B^3)\varepsilon_t$




```
 Shapiro-Wilk normality test

data:  res_3
W = 0.91849, p-value < 2.2e-16
```

3. SARIMA(1,0,0)(1,0,1)[3] $(1-\alpha B^3)(1-\beta B)y_t = (1-\theta B^3)\varepsilon_t$



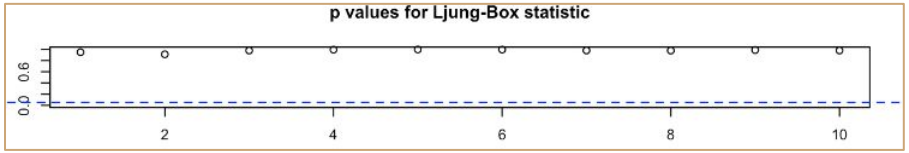

```
 Shapiro-Wilk normality test

data:  res_4
W = 0.91852, p-value < 2.2e-16
```

Summary of SARIMA(1,0,0)

n=1509	AIC	BIC	Log Likelihood
SARIMA(1,0,0)(1,0,0)[3]	-9503.27	-9481.99	4755.63
SARIMA(1,0,0)(0,0,1)[3]	**-9503.39**	**-9482.11**	4755.69
SARIMA(1,0,0)(1,0,1)[3]	-9501.62	-9475.02	4755.81

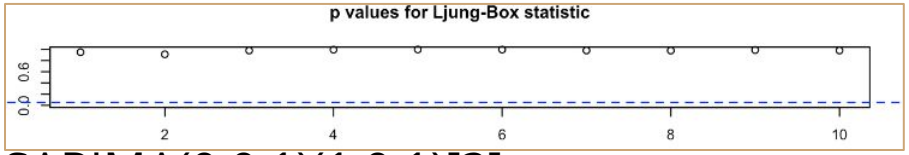
SARIMA(0,0,1)

1. SARIMA(0,0,1)(1,0,0)[3] $(1-\alpha B^3)y_t = (1-\delta B)\varepsilon_t$



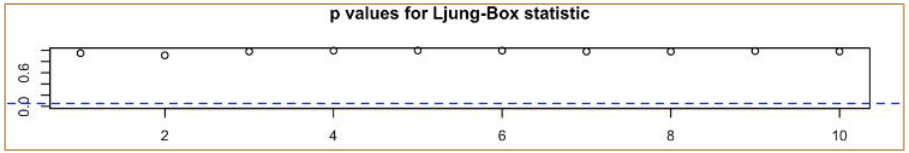

```
 Shapiro-Wilk normality test

data:  res_2
W = 0.918, p-value < 2.2e-16
```

2. SARIMA(0,0,1)(0,0,1)[3] $y_t = (1-\theta B^3)(1-\delta B)\varepsilon_t$




```
 Shapiro-Wilk normality test

data:  res_3
W = 0.91799, p-value < 2.2e-16
```

3. SARIMA(0,0,1)(1,0,1)[3] $(1-\alpha B^3)y_t = (1-\theta B^3)(1-\delta B)\varepsilon_t$



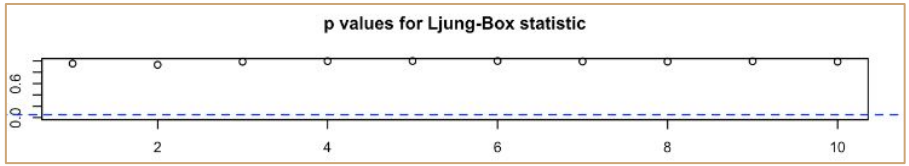

```
 Shapiro-Wilk normality test

data:  res_4
W = 0.91804, p-value < 2.2e-16
```

Summary of SARIMA(0,0,1)

n=1509	AIC	BIC	Log Likelihood
SARIMA(0,0,1)(1,0,0)[3]	-9503.23	-9481.95	4755.61
SARIMA(0,0,1)(0,0,1)[3]	**-9503.35**	**-9482.07**	4755.67
SARIMA(0,01)(1,0,1)[3]	-9501.58	-9474.98	4755.79

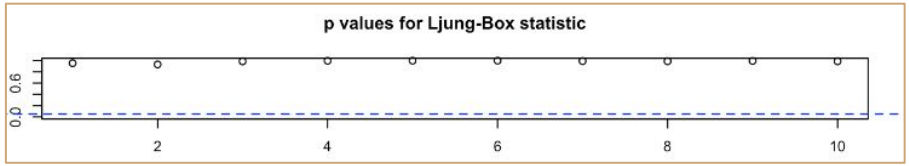
SARIMA(1,0,1)

1. SARIMA(1,0,1)(1,0,0)[3] $(1-\alpha B^3)(1-\beta B)y_t = (1-\delta B)\varepsilon_t$



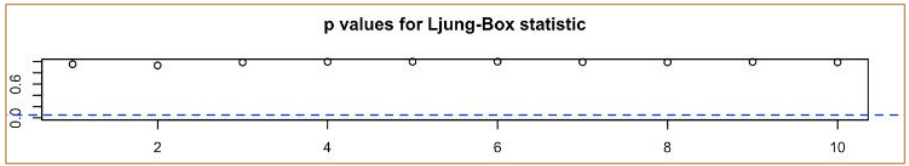

```
 Shapiro-Wilk normality test

data:  res_2
W = 0.91829, p-value < 2.2e-16
```

2. SARIMA(1,0,1)(0,0,1)[3] $(1-\beta B)y_t = (1-\theta B^3)(1-\delta B)\varepsilon_t$




```
 Shapiro-Wilk normality test

data:  res_3
W = 0.91828, p-value < 2.2e-16
```

3. SAMA(1,0,1)(1,0,1)[3] $(1-\alpha B^3)(1-\beta B)y_t = (1-\theta B^3)(1-\delta B)\varepsilon_t$




```
 Shapiro-Wilk normality test

data:  res_4
W = 0.91835, p-value < 2.2e-16
```

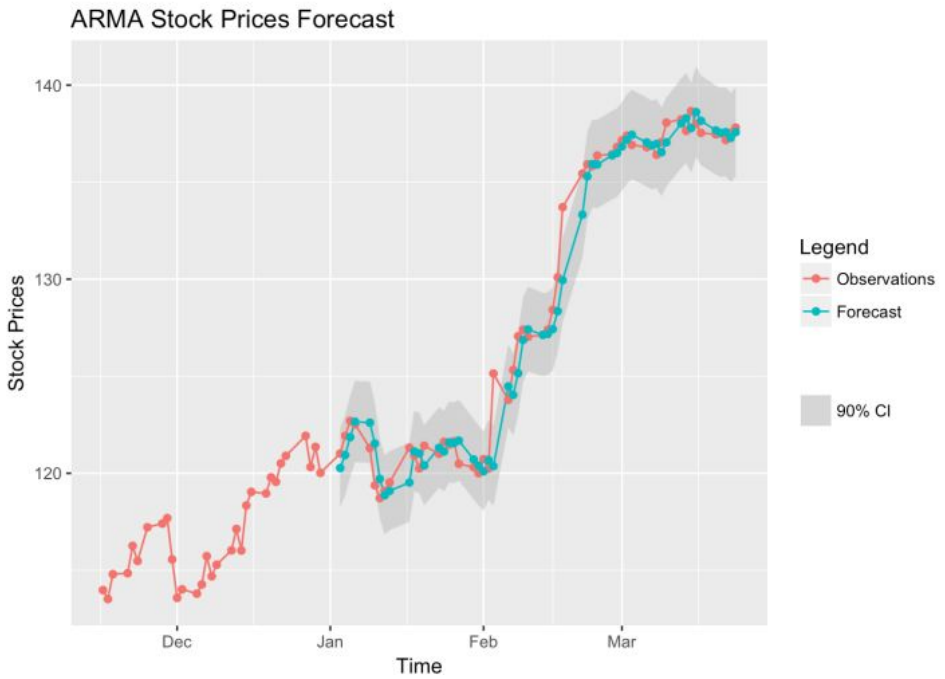
Summary of SARIMA(1,0,1)

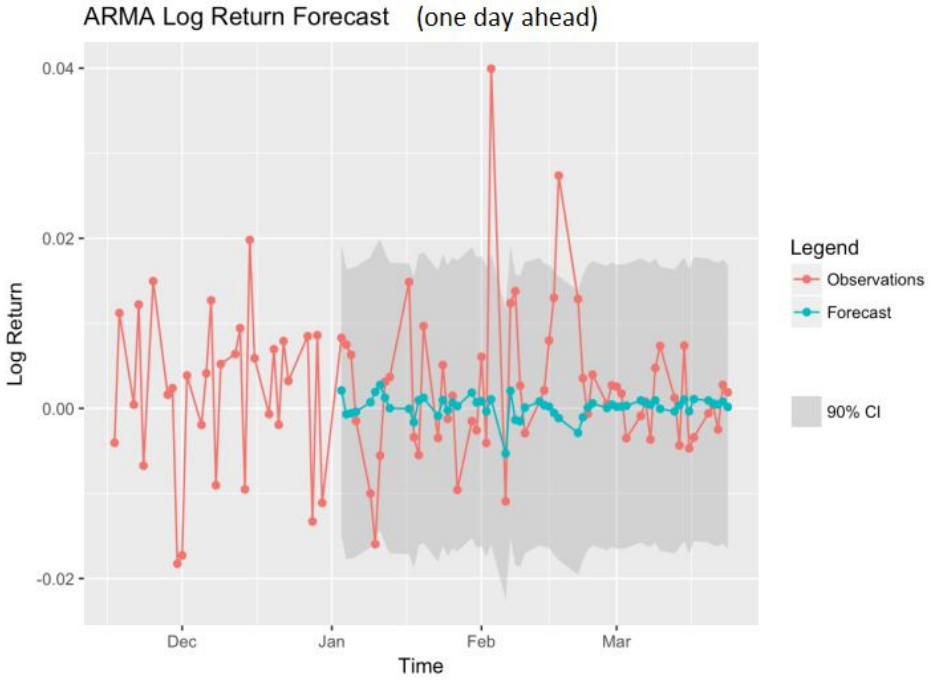
n=1509	AIC	BIC	Log Likelihood
SARIMA(1,0,1)(1,0,0)[3]	-9501.41	-9474.81	4755.70
SARIMA(1,0,1)(0,0,1)[3]	**-9501.53**	**-9474.94**	4755.77
SARIMA(1,0,1)(1,0,1)[3]	-9499.77	-9499.72	4855.89

SUMMARY OF ALL MODELS

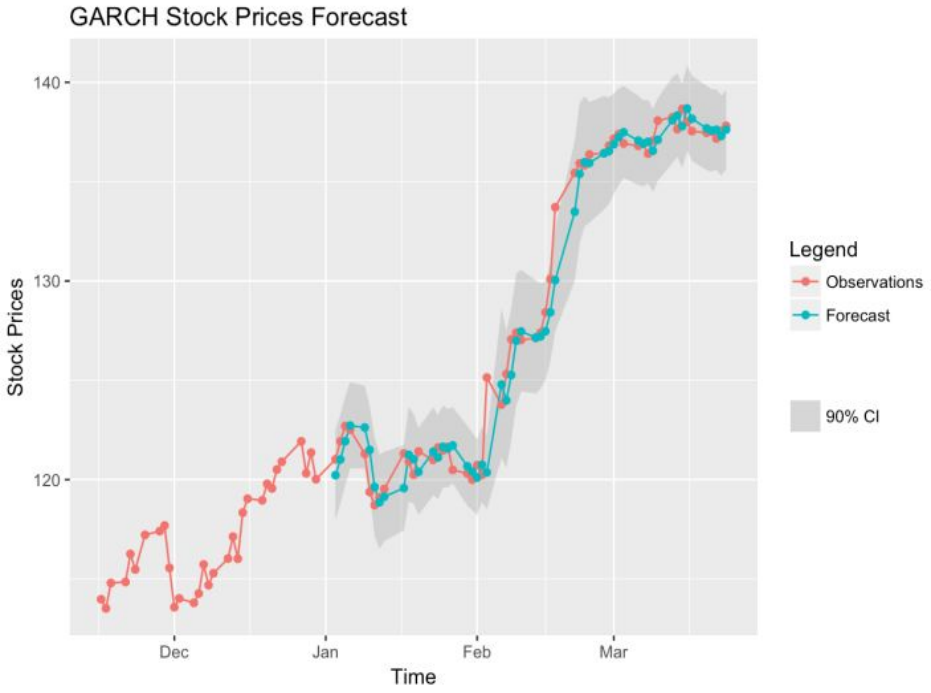
n=1509	AIC	BIC
AR(1)	-9501.03	-9485.07
AR(1)+GARCH(1,1)	-9701.84	-9675.24
SARIMA(1,0,0)(0,0,1)[3]	-9503.39	-9482.11
SARIMA(0,0,1)(0,0,1)[3]	-9501.58	-9474.98
SARIMA(1,0,1)(0,0,1)[3]	-9501.53	-9474.94

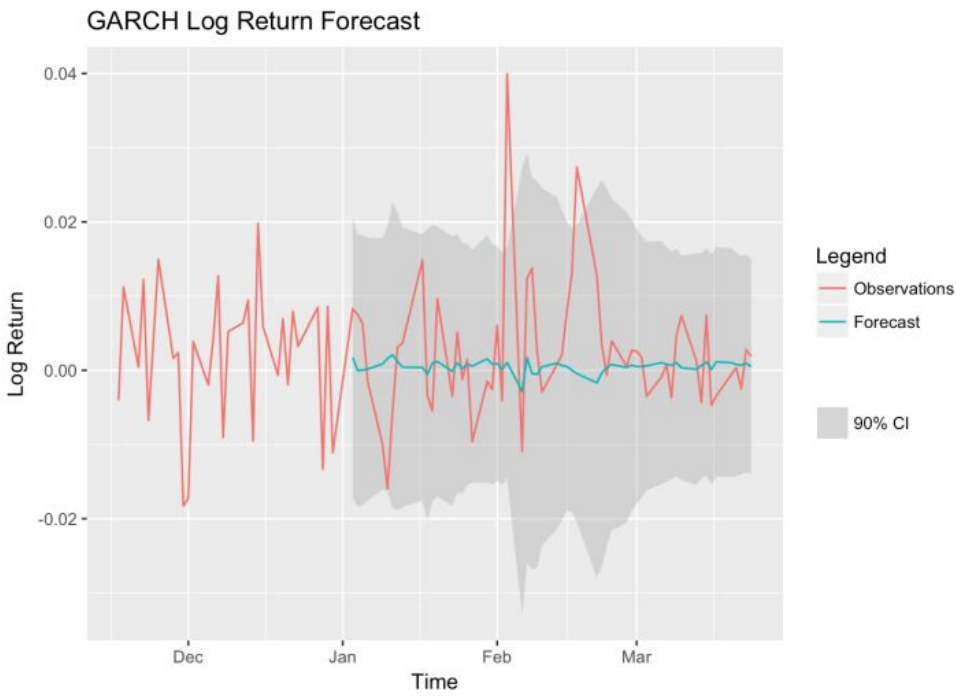
Forecasting: AR(1)



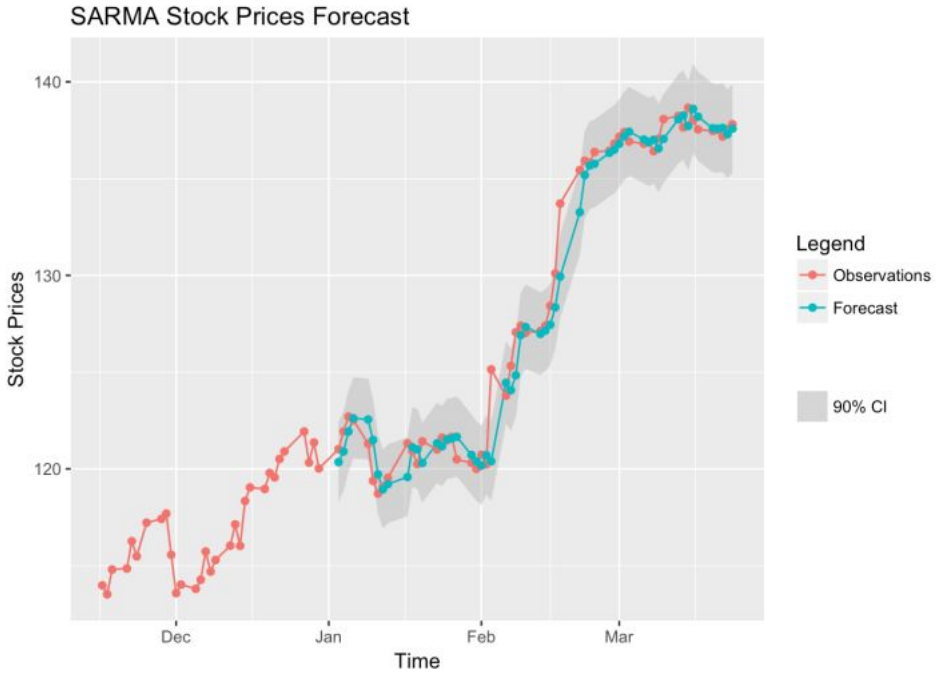


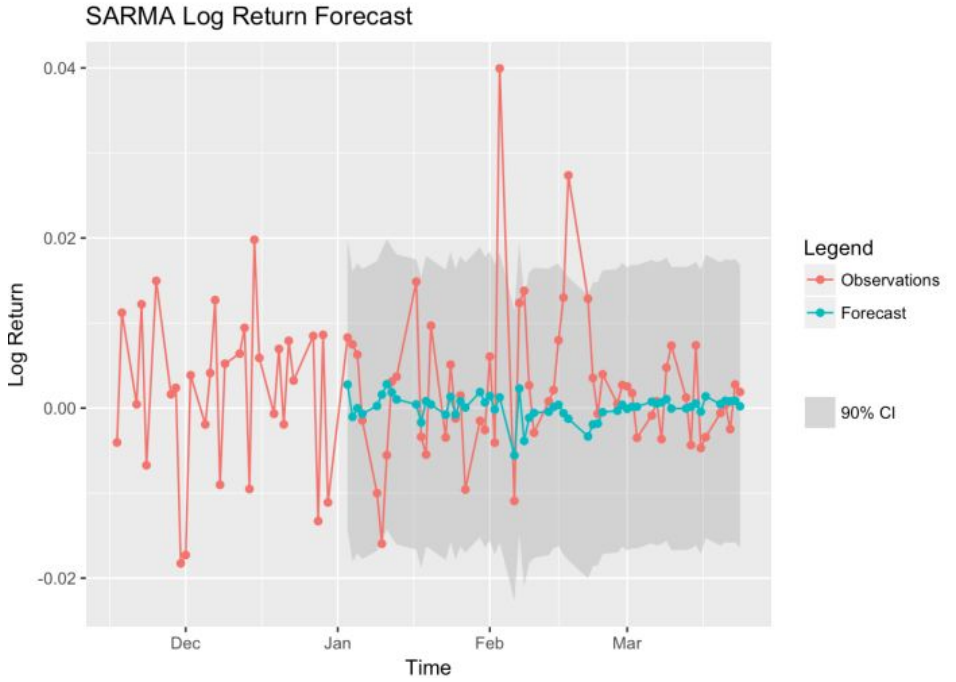
Forecasting: AR(1)+GARCH(1,1)



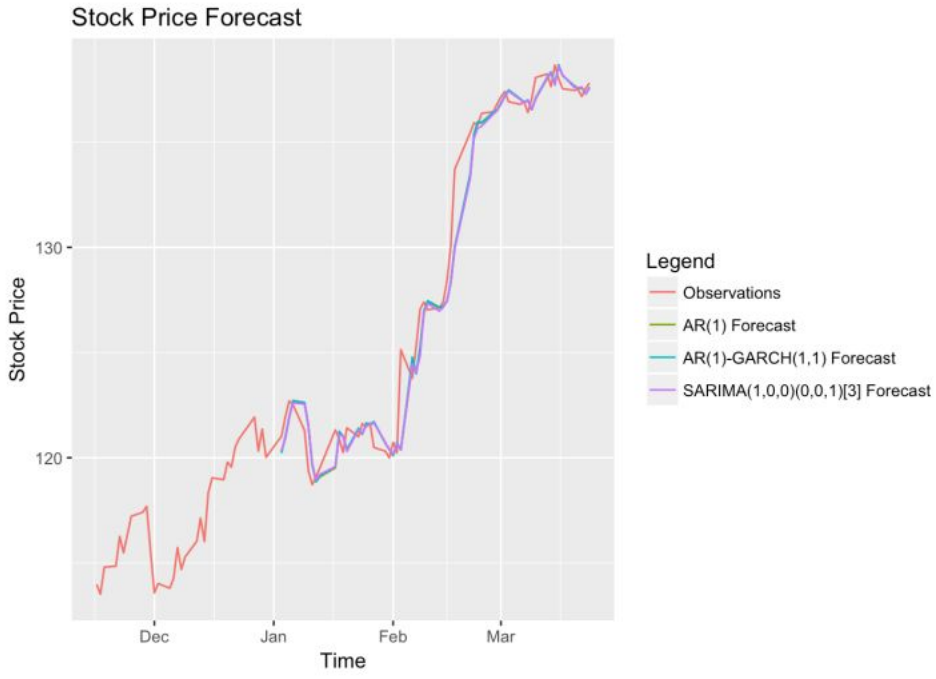


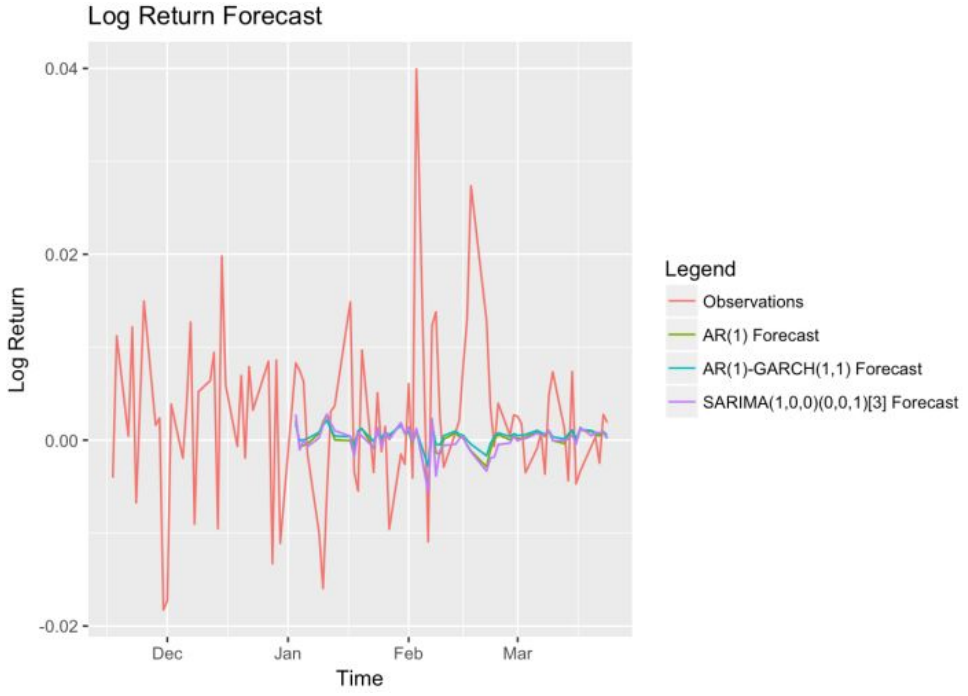
Forecasting: SARIMA(1,0,0)(0,0,1)[3]



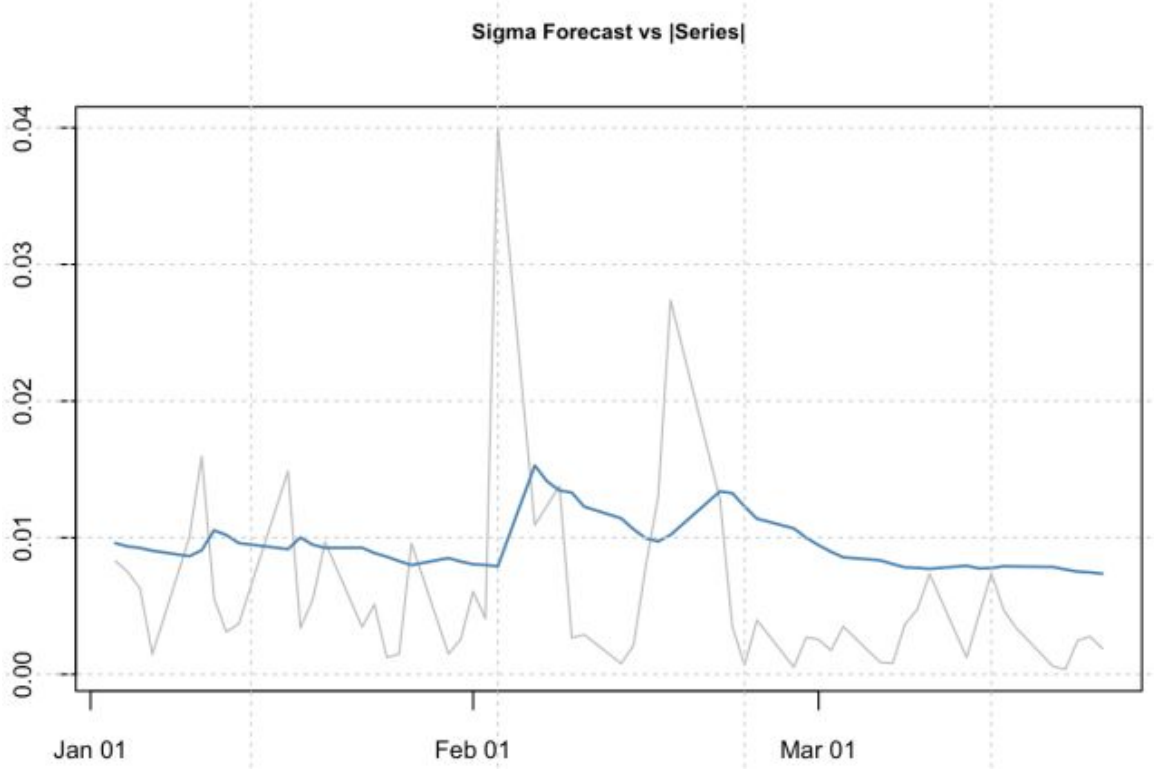


Forecasting: Comparisons





Forecasting Volatility



Conclusion

- We analyzed stock price and log returns of Clorox by using:
  - ARIMA models
  - ARIMA+GARCH models
  - SARIMA models
- We recommend fitting data in a AR(1)+GARCH(1,1) model.