

RTSM Project on Time Series Data.

Name: Kartik Yadav | **Roll No:** 15A

Stock Considered: Apple.inc or AAPL (herein after referred as “apple_stock”) on daily basis from 2020-01-01 to 2023-12-31

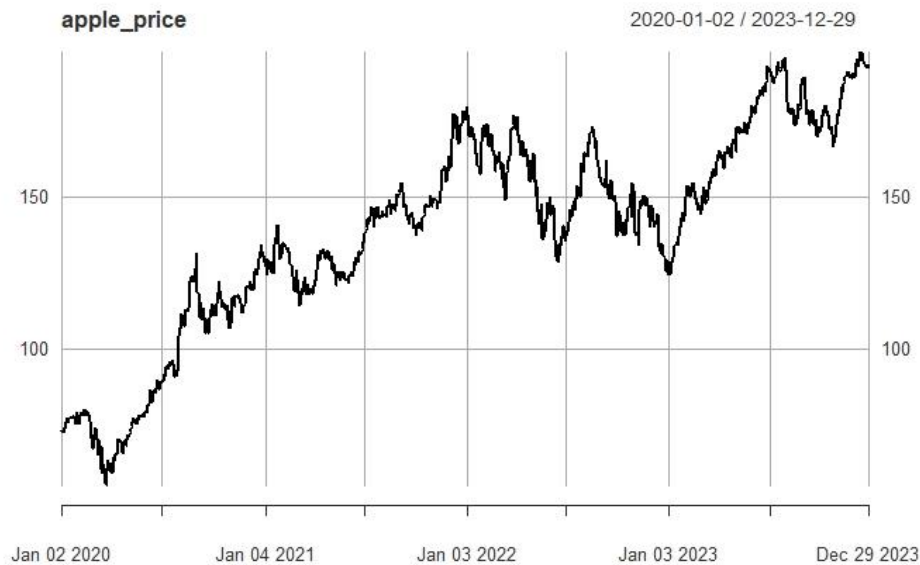
Required packages: quantmod, car ,forecast, tseries, FinTS, rugarch, utf8, ggplot2

```
packages = c('quantmod','car','forecast','tseries','FinTS', 'rugarch','utf8','ggplot2')
install.packages(packages, dependencies = TRUE)
lapply(packages, require, character.only = TRUE)
```

Loading the dataset of Apple inc. from Yahoo finance:

```
getSymbols(Symbols = 'AAPL',
           src = 'yahoo',
           from = as.Date('2020-01-01'),
           to = as.Date('2023-12-31'),
           periodicity = 'daily')

apple_price = na.omit(AAPL$AAPL.Adjusted) # Adjusted Closing Price
class(apple_price) # xts (Time-Series) Object
plot(apple_price)
```



To analyse the stock price of Apple from 2020-01-01 to 2023-12-29, extracted daily adjusted closing prices of Apple.inc from Yahoo finance. Performed ADF test on stock prices and returns.

Augmented Dickey-Fuller (ADF) Test for Stationarity with Stock Data:

```
adf_test_aapl = adf.test(apple_price); adf_test_aapl
```

Augmented Dickey-Fuller Test

```
data: apple_price
Dickey-Fuller = -2.4768, Lag order = 10, p-value = 0.3765
alternative hypothesis: stationary
```

Objective: To conduct an Augmented Dickey-Fuller (ADF) test for stationarity on the daily Stock Price.

Analysis: Performed the ADF test using the 'adf.test' function and obtained results.

Result: The Augmented Dickey-Fuller test for stationarity on stock price yields the following results: -

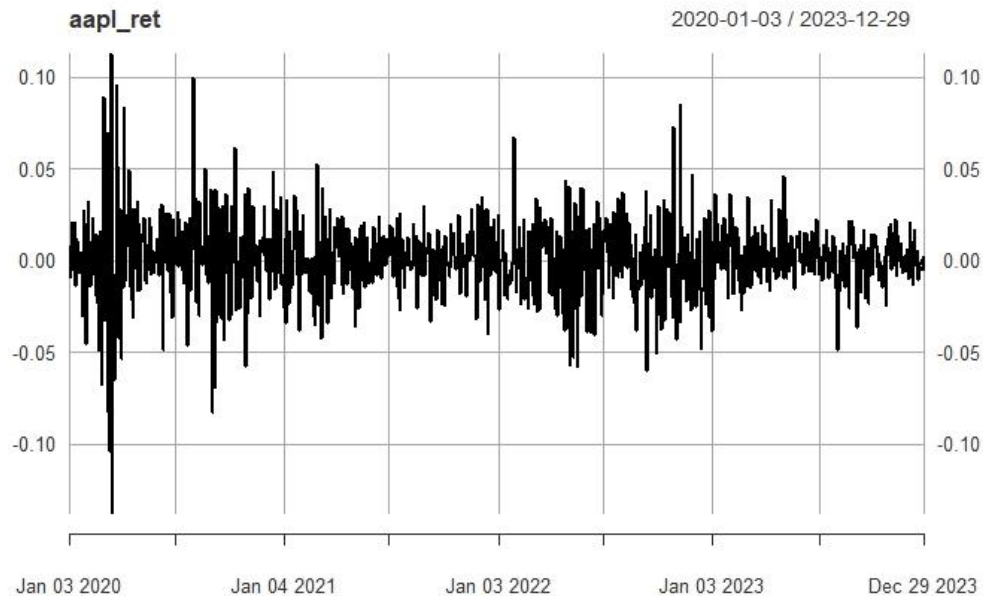
Dickey-Fuller = -2.4768, Lag order = 10, p-value = 0.3765 alternative hypothesis: stationary

Implication: The ADF test suggests that the stock price is non stationary as with a p value of 0.3765, we fail to reject the null hypothesis of non stationary.

Since the stock prices are non stationary, we will test the return on stocks.

Augmented Dickey-Fuller (ADF) Test for Stationarity with Return on stocks:

```
aapl_ret = na.omit(diff(log(apple_price)))  
adf_test_aapl_ret = adf.test(aapl_ret); adf_test_aapl_ret  
plot(aapl_ret)
```



Augmented Dickey-Fuller Test

```
data: aapl_ret  
Dickey-Fuller = -9.2926, Lag order = 10, p-value = 0.01  
alternative hypothesis: stationary
```

Objective: To conduct an Augmented Dickey-Fuller (ADF) test for stationarity on the daily return on stock.

Analysis: Performed the ADF test using the 'adf.test' function and obtained results.

Result: The Augmented Dickey-Fuller test for stationarity on return on stock yields the following results: -

Dickey-Fuller = -9.2926, Lag order = 10, p-value = 0.01 alternative hypothesis: stationary

Implication: The ADF test suggests that the return on stock is **stationary** as with a p value of 0.01, we can reject the null hypothesis of non stationary.

Ljung-Box Test for Autocorrelation - Stock Data (H0: No Autocorrelation):

```
lb_test_aapl_ret = Box.test(aapl_ret); lb_test_aapl_ret
```

Box-Pierce test

data: aapl_ret
X-squared = 17.547, df = 1, p-value = 2.803e-05

Objective: To perform a Ljung-Box test for autocorrelation on the stock returns.

Analysis: Conducted the Ljung-Box test using the 'Box.test' function and obtained results.

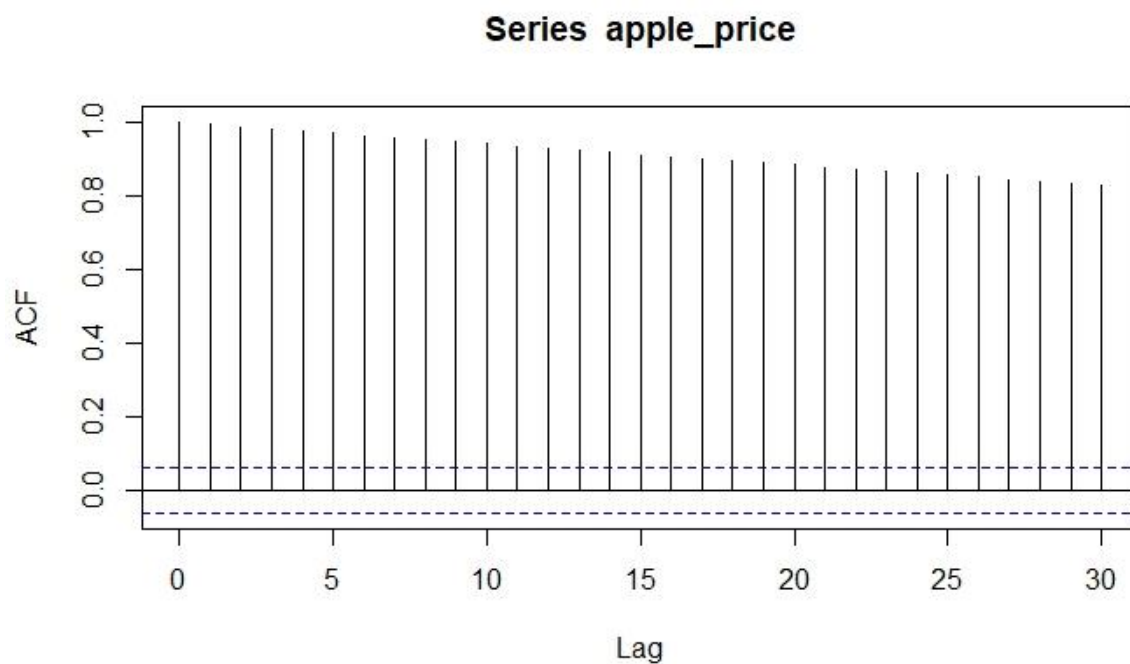
Result: The Ljung-Box test for autocorrelation on stock returns yields the following results:
X-squared = 17.547, df = 1, p-value = 2.803e-05

Implication: The Ljung-Box test indicates significant autocorrelation in the stock returns. The small p-value ($< 2.803e-05$) suggests evidence against the null hypothesis of no autocorrelation.

Action Step: Given the presence of autocorrelation, it may be advisable to consider an autoARIMA model for time series forecasting. AutoARIMA can help in automatically selecting an appropriate ARIMA model with differencing to account for the observed autocorrelation.

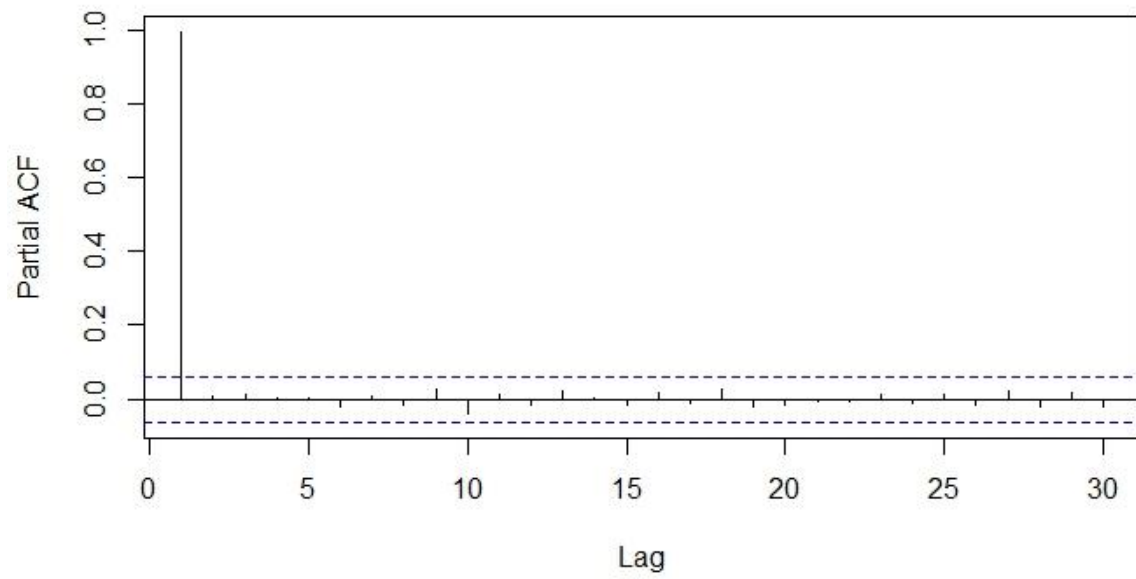
Autocorrelation Function (ACF) | Partial Autocorrelation Function (PACF)

acf(apple_price)



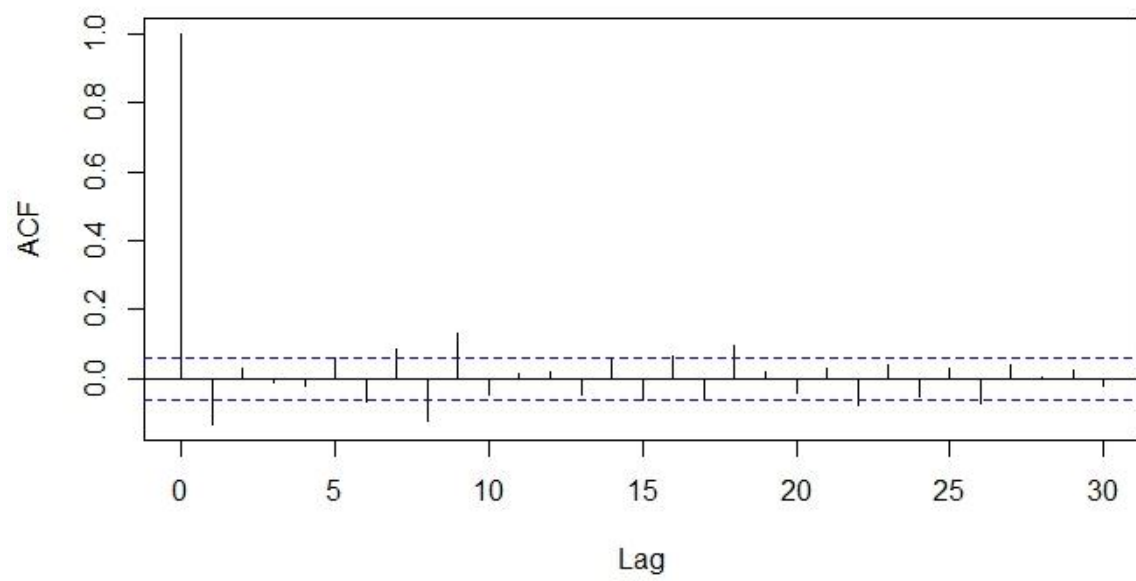
pacf(apple_price)

Series apple_price

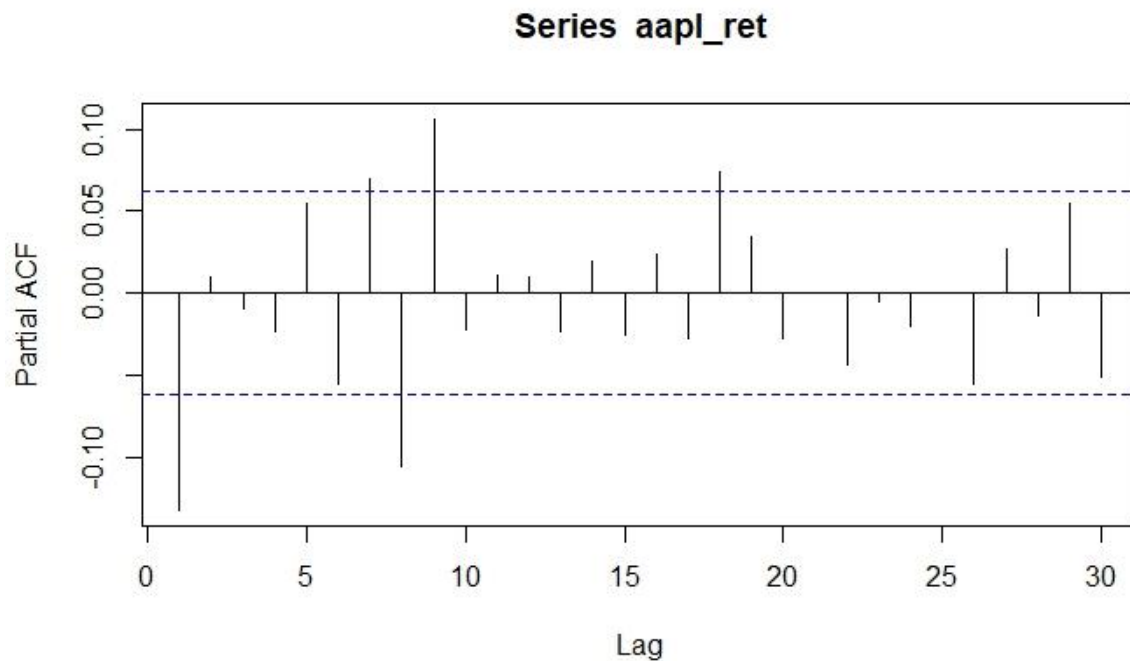


`acf(aapl_ret)`

Series aapl_ret



```
pacf(aapl_ret)
```



Auto ARIMA on Stock Return and Stock Price:

```
arma_pq_aapl_ret = auto.arima(aapl_ret); arma_pq_aapl_ret
```

```
Series: aapl_ret  
ARIMA(0,0,1) with non-zero mean
```

```
Coefficients:  
          ma1      mean  
      -0.1280  1e-03  
s.e.    0.0306  6e-04
```

```
sigma^2 = 0.0004396: log likelihood = 2459.06  
AIC=-4912.11  AICc=-4912.09  BIC=-4897.37
```

```
arma_pq_aapl = auto.arima(apple_price); arma_pq_aapl
```

```
Series: apple_price  
ARIMA(0,1,0) with drift
```

```
Coefficients:  
      drift  
      0.1186  
s.e.    0.0824
```

```
sigma^2 = 6.827: log likelihood = -2390.75  
AIC=4785.5  AICc=4785.51  BIC=4795.33
```

Objective: To perform autoARIMA modeling on the stock returns and stock price.

Analysis: Used the 'auto.arima' function to automatically select the ARIMA model for both returns and prices.

Results:

For stock returns ('stock_ret'): The autoARIMA model suggests an ARIMA(0,0,1) with non zero mean. Coefficients: MA: ma1 - sigma^2 = 0.0004396: log likelihood = 2459.06
AIC=-4912.11 AICc=-4912.09 BIC=-4897.37

For Adjusted Closing Prices ('stock_price'): The autoARIMA model suggests an ARIMA(0,1,0) with drift. Coefficients: drift- sigma^2 = 6.827: log likelihood = -2390.75
AIC=4785.5 AICc=4785.51 BIC=4795.33

$$y(t) = c + (-0.4205) * y(t-1) - (-0.4677) * e(t-1) - (-0.3902) * e(t-2) - (-0.1162) * e(t-3) - (-0.0908) * e(t-4) + 0.0878 * e(t-5) + e(t)$$

Implication: The autoARIMA models provide a statistical framework to capture the underlying patterns in both stock returns and stock price. These models can be used for forecasting future values, and the AIC, AICc, and BIC values help in model comparison.

Note: The log likelihood is positive, indicating a good fit of the model to the data. The information criteria (AIC, AICc, BIC) are relatively low, suggesting a good balance between model fit and complexity.

Ljung-Box Test for Autocorrelation - Model Residuals(H0: No Autocorrelation)

```
lb_test_arma_pq_aapl_ret = Box.test(arma_pq_aapl_ret$residuals); lb_test_arma_pq_aapl_ret
```

Box-Pierce test

```
data: arma_pq_aapl_ret$residuals  
X-squared = 0.0096919, df = 1, p-value = 0.9216
```

Objective: To perform a Ljung-Box test for autocorrelation on the arima residuals.

Analysis: Conducted the Ljung-Box test using the 'Box.test' function and obtained results.

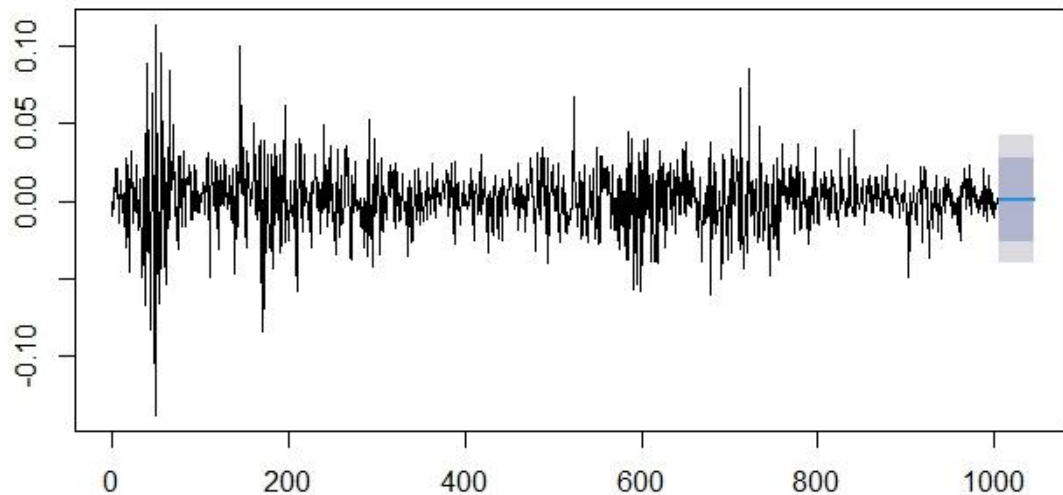
Result: The Ljung-Box test for autocorrelation on arima residual yields the following results:
X-squared = 0.0096919, df = 1, p-value = 0.9216

It indicates no significant autocorrelation in the residuals of the ARIMA(0,0,1) model residual. The high p-value (0.9216) suggests that there is no evidence against the null hypothesis of no autocorrelation.

Forecasting with ARIMA Models

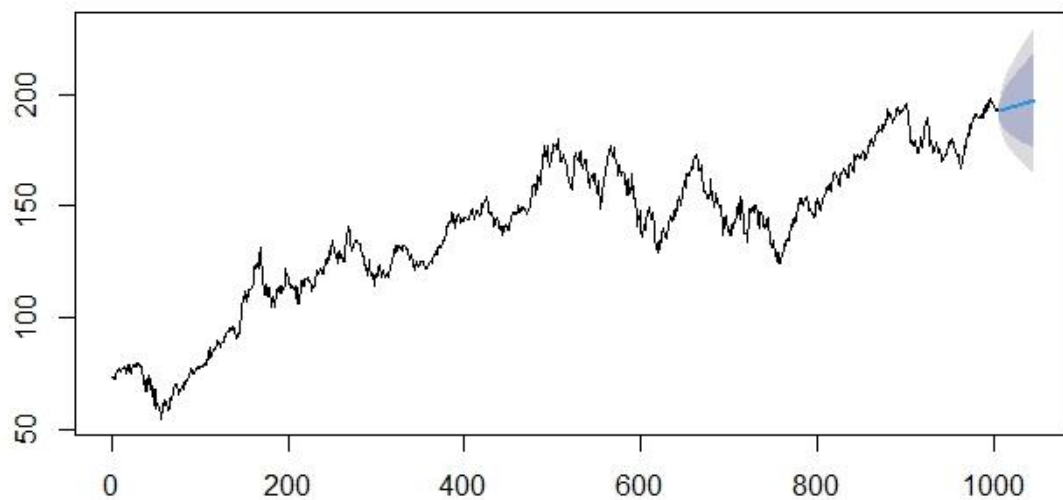
```
aapl_ret_fpq = forecast(arma_pq_aapl_ret, h = 40)  
plot(aapl_ret_fpq)
```

Forecasts from ARIMA(0,0,1) with non-zero mean



```
aapl_fpq = forecast(arma_pq_aapl, h = 40)  
plot(aapl_fpq)
```

Forecasts from ARIMA(0,1,0) with drift



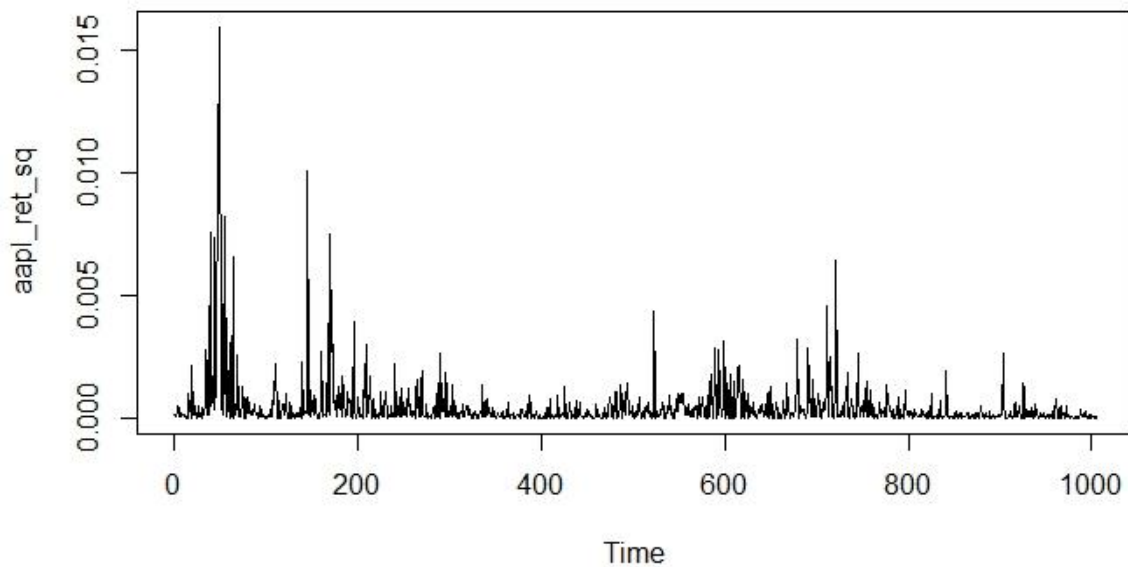
Objective: To fit an ARIMA(0,0,1) model to the stock returns & price and generate forecasts. Analysis: Used the 'arima' function to fit the ARIMA model and the '**forecast**' function to generate forecasts.

Results: The plot displays the original time series of stock returns along with the forecast values.

Implication: The ARIMA(0,0,1) model is fitted to the historical stock returns, providing insights into the underlying patterns. The generated forecast can be used for future predictions, and the plot visually represents the model's performance.

Test for Volatility Clustering or Heteroskedasticity: Box Test:

```
aapl_ret_sq = arma_pq_aapl_ret$residuals^2 # Return Variance (Since Mean Returns is approx. 0)
plot(aapl_ret_sq)
```



```
aapl_ret_sq_box_test = Box.test(aapl_ret_sq, lag = 10) # H0: Return Variance Series is Not Serially Correlated
aapl_ret_sq_box_test # Inference : Return Variance Series is Heteroskedastic (Has Volatility Clustering)
```

Box-Pierce test

```
data: aapl_ret_sq
x-squared = 507.38, df = 10, p-value < 2.2e-16
```

Test for Volatility Clustering or Heteroskedasticity: ARCH Test:

```
aapl_ret_arch_test = ArchTest(arma_pq_aapl_ret$residuals^2, lags = 10) # H0: No ARCH Effects
```

```
aapl_ret_arch_test # Inference : Return Series is Heteroskedastic (Has Volatility Clustering)
```

ARCH LM-test; Null hypothesis: no ARCH effects

```
data: arma_pq_aapl_ret$residuals^2  
Chi-squared = 213.05, df = 10, p-value < 2.2e-16
```

Objective: To test for volatility clustering or heteroskedasticity in the arima residuals.

Analysis: Conducted Box test and ARCH test on the squared residuals to assess the presence of volatility clustering.

Results:

1. Box Test for Volatility Clustering:

X-squared statistic: 507.38

Degrees of freedom: 10

p-value: < 2.2e-16

Thus, the test indicates significant evidence against the null hypothesis, suggesting that the return variance series exhibits volatility clustering or heteroskedasticity

2. ARCH Test for Volatility Clustering:

Chi-squared statistic: 213.05

Degrees of freedom: 10

p-value: < 2.2e-16

Thus, The ARCH test provides significant evidence against the null hypothesis. It confirms the presence of ARCH effects in the return series.

Implication: The results from above test suggest the variation and the presence of volatility clustering or heteroskedasticity in the residual. Hence, we proceed with Residual modelling assuming Heteroskedasticity.

GARCH Model

```
garch_model1 = ugarchspec(variance.model = list(model = 'sGARCH', garchOrder = c(1,1)),  
mean.model = list(armaOrder = c(1,10), include.mean = TRUE))  
aapl_ret_garch1 = ugarchfit(garch_model1, data = aapl_ret); aapl_ret_garch1
```

```
*-----*  
*          GARCH Model Fit          *  
*-----*
```

Conditional Variance Dynamics

```
-----  
GARCH Model      : sGARCH(1,1)  
Mean Model       : ARFIMA(1,0,10)  
Distribution      : norm
```

Optimal Parameters

```
-----
```

	Estimate	Std. Error	t value	Pr(> t)
mu	0.001629	0.000483	3.37609	0.000735
ar1	-0.850800	0.158378	-5.37194	0.000000
ma1	0.805218	0.162075	4.96817	0.000001
ma2	-0.061504	0.042434	-1.44941	0.147224
ma3	-0.048708	0.042663	-1.14167	0.253590
ma4	-0.022547	0.042802	-0.52678	0.598349
ma5	0.016659	0.041980	0.39683	0.691495
ma6	0.006016	0.041383	0.14538	0.884413
ma7	0.011800	0.043742	0.26975	0.787349
ma8	-0.026587	0.041384	-0.64244	0.520586
ma9	-0.004569	0.041580	-0.10988	0.912506
ma10	0.041400	0.031017	1.33476	0.181955
omega	0.000009	0.000005	1.85442	0.063679
alpha1	0.100560	0.007315	13.74806	0.000000
beta1	0.879553	0.015039	58.48443	0.000000

Robust Standard Errors:

	Estimate	Std. Error	t value	Pr(> t)
mu	0.001629	0.000554	2.939004	0.003293
ar1	-0.850800	0.107414	-7.920721	0.000000
ma1	0.805218	0.108781	7.402179	0.000000
ma2	-0.061504	0.040998	-1.500167	0.133571
ma3	-0.048708	0.040213	-1.211249	0.225800
ma4	-0.022547	0.037859	-0.595552	0.551475
ma5	0.016659	0.042378	0.393100	0.694245
ma6	0.006016	0.048009	0.125312	0.900277
ma7	0.011800	0.049301	0.239337	0.810844
ma8	-0.026587	0.042590	-0.624253	0.532461
ma9	-0.004569	0.048314	-0.094564	0.924661
ma10	0.041400	0.034243	1.208998	0.226664
omega	0.000009	0.000015	0.607830	0.543300
alpha1	0.100560	0.035885	2.802271	0.005074
beta1	0.879553	0.031194	28.196524	0.000000

LogLikelihood : 2581.116

Information Criteria

Akaike	-5.1067
Bayes	-5.0334
Shibata	-5.1071
Hannan-Quinn	-5.0788

Weighted Ljung-Box Test on Standardized Residuals

	statistic	p-value
Lag[1]	0.2168	0.6415
Lag[2*(p+q)+(p+q)-1][32]	10.5817	1.0000
Lag[4*(p+q)+(p+q)-1][54]	23.7921	0.8306

d.o.f=11
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals

	statistic	p-value
Lag[1]	0.397	0.5286
Lag[2*(p+q)+(p+q)-1][5]	1.293	0.7905
Lag[4*(p+q)+(p+q)-1][9]	2.580	0.8258

d.o.f=2

Weighted ARCH LM Tests

	Statistic	Shape	Scale	P-Value
ARCH Lag[3]	0.5320	0.500	2.000	0.4658
ARCH Lag[5]	0.8216	1.440	1.667	0.7865
ARCH Lag[7]	2.1047	2.315	1.543	0.6951

Nyblom stability test

Joint Statistic: 4.8789

Individual Statistics:

mu	0.07357
ar1	0.32658
ma1	0.31180
ma2	0.19616
ma3	0.19373
ma4	0.24128
ma5	0.34670
ma6	0.34163
ma7	0.19204
ma8	0.15536
ma9	0.15109
ma10	0.12448
omega	0.91431
alpha1	0.35049
beta1	0.24359

Asymptotic Critical values (10% 5% 1%)

Joint Statistic: 3.26 3.54 4.07

Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

Sign Bias t-value prob sig
Negative Sign Bias 0.84147 0.4003
Positive Sign Bias 0.36873 0.7124
Joint Effect 0.03922 0.9687
1.11029 0.7746

Adjusted Pearson Goodness-of-Fit Test:

group statistic p-value(g-1)
1 20 42.18 0.0016737
2 30 61.66 0.0003826
3 40 71.82 0.0010628
4 50 90.47 0.0002870

```
garch_model2 = ugarchspec(variance.model = list(model = 'sGARCH', garchOrder = c(1,1)),  
mean.model = list(armaOrder = c(1,5), include.mean = FALSE))  
aapl_ret_garch2 = ugarchfit(garch_model2, data = aapl_ret); aapl_ret_garch2
```

```
*-----*  
*          GARCH Model Fit          *  
*-----*
```

Conditional Variance Dynamics

GARCH Model : sGARCH(1,1)
Mean Model : ARFIMA(1,0,5)
Distribution : norm

Optimal Parameters

Estimate Std. Error t value Pr(>|t|)
ar1 -0.749896 0.210726 -3.558623 0.000373
ma1 0.717150 0.212757 3.370751 0.000750
ma2 -0.041408 0.040885 -1.012799 0.311156
ma3 -0.035621 0.040539 -0.878685 0.379572
ma4 0.000682 0.040731 0.016736 0.986647
ma5 0.038736 0.032252 1.201061 0.229728
omega 0.000009 0.000003 2.807170 0.004998

Sign Bias Test

	t-value	prob	sig
Sign Bias	1.1724	0.2413	
Negative Sign Bias	0.2548	0.7989	
Positive Sign Bias	0.1650	0.8690	
Joint Effect	1.9601	0.5807	

Adjusted Pearson Goodness-of-Fit Test:

	group	statistic	p-value(g-1)
1	20	41.39	0.002142
2	30	51.09	0.006875
3	40	62.98	0.008844
4	50	73.96	0.012150

Objective: The objective was to fit GARCH models to the residuals of the ARIMA(1, 1, 5) mean model and evaluate volatility clustering.

Analysis: Two GARCH models ('garch_model1' and 'garch_model2') were fitted to the residuals, and an ARCH test was conducted on squared residuals to assess volatility clustering.

Results:

Model Comparison: garch_model2 (sGARCH(1,1) with ARFIMA(1,0,5) mean model) appears to be superior to garch_model1 (sGARCH(1,1) with ARFIMA(1,0,10) mean model) based on the following:

Higher LogLikelihood: garch_model1 has a higher LogLikelihood value (2581.116) compared to garch_model2 (2574.294) indicating a better fit to the data.

Information Criteria: Two information criteria are lower for model 1, two are lower for model 2.

Ljung-Box Tests: Both models exhibit significant serial correlation in the standardized residuals and squared residuals, as indicated by the p-values of the Ljung-Box tests.

ARCH LM Tests: There are no significant effect of conditional heteroskedasticity in either of the models.

Nyblom Stability Test: Both models pass the Nyblom stability test, with all individual statistics below the critical values.

Adjusted Pearson Goodness-of-Fit Test: Both models exhibit a small p-value for the test statistic, indicating good goodness-of-fit.

Therefore, considering the higher LogLikelihood, similar information criteria, and the absence of significant ARCH effects in both models, garch_model1 (sGARCH(1,1) with ARFIMA(1,0,10) mean model) can be considered a better fit for the data compared to garch_model2. Further analysis will be required to address the significant serial correlation observed in the standardized residuals and squared residuals.

GARCH Forecast

```
aapl_ret_garch_forecast1 = ugarchforecast(aapl_ret_garch1, n.ahead = 50); aapl_ret_garch_f  
orecast1
```

```
*-----*  
*          GARCH Model Forecast          *  
*-----*
```

```
Model: sGARCH  
Horizon: 50  
Roll Steps: 0  
Out of Sample: 0
```

0-roll forecast [T0=2023-12-29]:

	Series	Sigma
T+1	0.001469	0.01091
T+2	0.001961	0.01121
T+3	0.001848	0.01149
T+4	0.001252	0.01176
T+5	0.001768	0.01202
T+6	0.001273	0.01227
T+7	0.001652	0.01251
T+8	0.001716	0.01274
T+9	0.001599	0.01296
T+10	0.001349	0.01318
T+11	0.001868	0.01339
T+12	0.001426	0.01359
T+13	0.001802	0.01378
T+14	0.001482	0.01396
T+15	0.001754	0.01414
T+16	0.001523	0.01432
T+17	0.001720	0.01449
T+18	0.001552	0.01465
T+19	0.001695	0.01481
T+20	0.001574	0.01496
T+21	0.001677	0.01511
T+22	0.001589	0.01526
T+23	0.001664	0.01540
T+24	0.001600	0.01554
T+25	0.001654	0.01567
T+26	0.001608	0.01580
T+27	0.001647	0.01592
T+28	0.001614	0.01604
T+29	0.001642	0.01616
T+30	0.001618	0.01628
T+31	0.001639	0.01639
T+32	0.001621	0.01650
T+33	0.001636	0.01661
T+34	0.001624	0.01671
T+35	0.001634	0.01681
T+36	0.001625	0.01691
T+37	0.001633	0.01701
T+38	0.001626	0.01710
T+39	0.001632	0.01719
T+40	0.001627	0.01728
T+41	0.001631	0.01737
T+42	0.001628	0.01745
T+43	0.001631	0.01753
T+44	0.001628	0.01761
T+45	0.001630	0.01769
T+46	0.001629	0.01777
T+47	0.001630	0.01784
T+48	0.001629	0.01792
T+49	0.001630	0.01799
T+50	0.001629	0.01806

```
aapl_ret_garch_forecast2 = ugarchforecast(aapl_ret_garch2, n.ahead = 50); aapl_ret_garch_forecast2
```

```
*-----*
*      GARCH Model Forecast      *
*-----*
```

```
Model: sGARCH
Horizon: 50
Roll Steps: 0
Out of Sample: 0
```

```
0-roll forecast [T0=2023-12-29]:
```

	Series	Sigma
T+1	-3.410e-04	0.01106
T+2	2.796e-04	0.01136
T+3	1.158e-06	0.01165
T+4	9.598e-05	0.01192
T+5	-2.926e-04	0.01218
T+6	2.194e-04	0.01243
T+7	-1.646e-04	0.01267
T+8	1.234e-04	0.01290
T+9	-9.254e-05	0.01313
T+10	6.940e-05	0.01334
T+11	-5.204e-05	0.01355
T+12	3.902e-05	0.01374
T+13	-2.926e-05	0.01394
T+14	2.195e-05	0.01412
T+15	-1.646e-05	0.01430
T+16	1.234e-05	0.01447
T+17	-9.254e-06	0.01464
T+18	6.940e-06	0.01480
T+19	-5.204e-06	0.01495
T+20	3.903e-06	0.01510
T+21	-2.926e-06	0.01525
T+22	2.195e-06	0.01539
T+23	-1.646e-06	0.01553
T+24	1.234e-06	0.01566
T+25	-9.254e-07	0.01579
T+26	6.940e-07	0.01592
T+27	-5.204e-07	0.01604
T+28	3.903e-07	0.01616
T+29	-2.927e-07	0.01627
T+30	2.195e-07	0.01638
T+31	-1.646e-07	0.01649
T+32	1.234e-07	0.01660
T+33	-9.255e-08	0.01670
T+34	6.940e-08	0.01680
T+35	-5.204e-08	0.01689
T+36	3.903e-08	0.01699
T+37	-2.927e-08	0.01708
T+38	2.195e-08	0.01717
T+39	-1.646e-08	0.01726
T+40	1.234e-08	0.01734
T+41	-9.255e-09	0.01742
T+42	6.940e-09	0.01750
T+43	-5.204e-09	0.01758
T+44	3.903e-09	0.01766
T+45	-2.927e-09	0.01773
T+46	2.195e-09	0.01780
T+47	-1.646e-09	0.01787
T+48	1.234e-09	0.01794
T+49	-9.255e-10	0.01801
T+50	6.940e-10	0.01807

Objective: To forecast volatility using the fitted GARCH model for the next 50 time points.

Analysis: Used the 'ugarchforecast' function to generate volatility forecasts for the next 50 time points.

Result: GARCH Model Forecast: - Model: sGARCH - Horizon: 50 - Roll Steps: 0 - Out of Sample: 0 0-roll forecast [T0=2023-12-29]: - Forecasted Series: - T+1 to T+100: Contains forecasted values of volatility (Sigma) for each time point.

Implication: The forecasted values represent the predicted volatility for the next 50 time points based on the fitted GARCH model. These forecasts can be useful for risk management and decision-making, providing insights into the expected future volatility of the financial time series.

`plot(aapl_ret_garch_forecast1)`

