Project 6 Using Arch Model to Predict Apple Stock

Introduction:

In the stock market, many scholars are concerned about the changes in stock prices and the prediction of stock prices. Uncertainty changes in stock prices often manifest themselves as market fluctuations. The ARCH model can accurately simulate the change of volatility of time series variables. It is widely used in financial engineering and empirical research, so that people can grasp the risk (volatility) more accurately, especially in the value at risk theory. In our project, we used the arch model for volatility prediction of the apple stocks, we tested our model and we also forecasted the apple stocks.

Required R Packages:

- Quantmod: Quantitative financial modeling and trading framework in R. It was used in building trading strategies in our project.
- Xts: Used to handle time-based dataset and helping create time series model
- PerformanceAnalytics: Collection of econometrics functions for model performance and risk evaluation
- Rugarch: It's a premier open source software for univariate GARCH modeling.

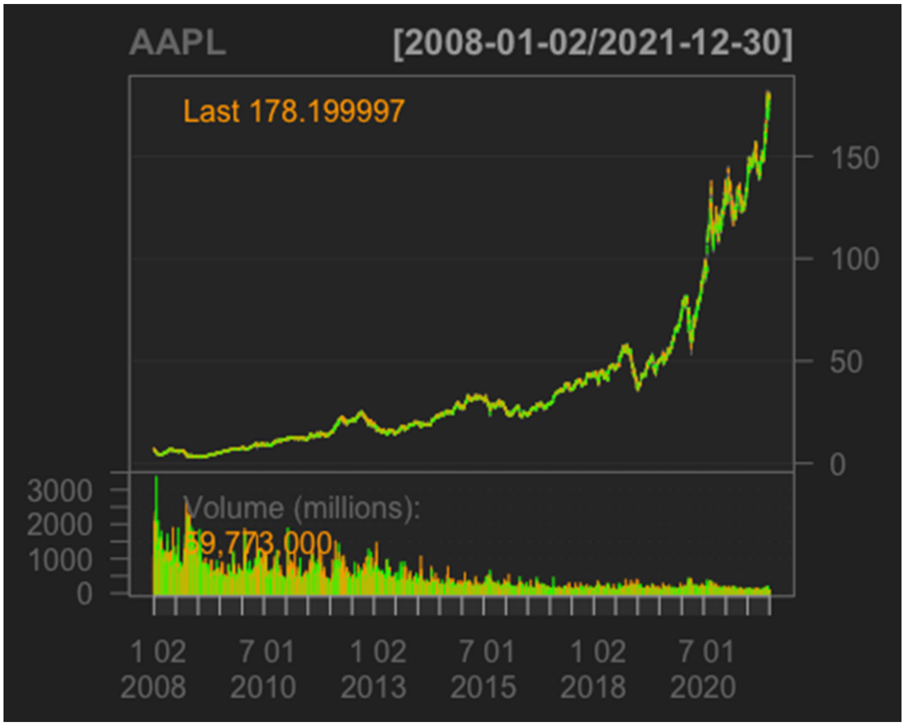
Data Description:

In this project, the data set we used is the stock price of APPLE from 2008 to 2021

	AAPL.Open	AAPL.High	AAPL.Low	AAPL.Close	AAPL.Volume	AAPL.Adjusted
2008-01-02	7.116786	7.152143	6.876786	6.958571	1079178800	5.958444
2008-01-03	6.978929	7.049643	6.881786	6.961786	842066400	5.961198
2008-01-04	6.837500	6.892857	6.388929	6.430357	1455832000	5.506149
2008-01-07	6.473214	6.557143	6.079643	6.344286	2072193200	5.432449
2008-01-08	6.433571	6.516429	6.100000	6.116071	1523816000	5.237033
2008-01-09	6.117857	6.410714	6.010714	6.407143	1813882000	5.486271
2008-01-10	6.342143	6.464286	6.264643	6.357857	1482975200	5.444069
2008-01-11	6.285714	6.351786	6.071429	6.167500	1232285600	5.281072

There are 6 columns in the data set:
Opening price
High of the day
Low of the day
Closing price
Volume
Adjusted prices

For example, this is the chart of information on stock price for December, 2021. We can find the High of the day and Low of the day from the top and lower line. We can also know the opening and closing price depending on the color. For example, if the color is orange, then the opening price is lower than the closing price.



If we make this plot for the entire data, we can see that there are a lot of fluctuations in the price. This kind of data will be difficult to model, so we convert this data to daily return data instead of use it directly.

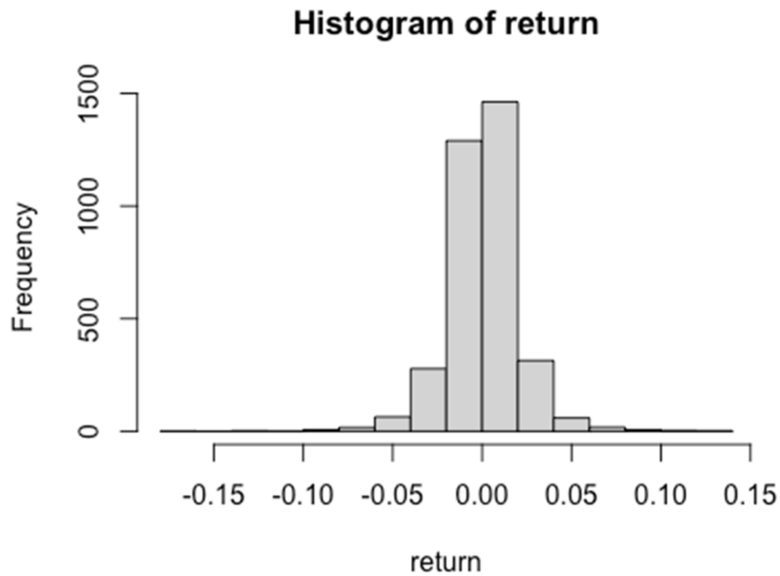
Since the closing price is important, so we calculate the daily return of the closing price.

Return value = (current closing price - previous closing price)/ previous closing price

AAPL.Close
6.958571
6.961786
6.430357
6.344286
6.116071
6.407143

→

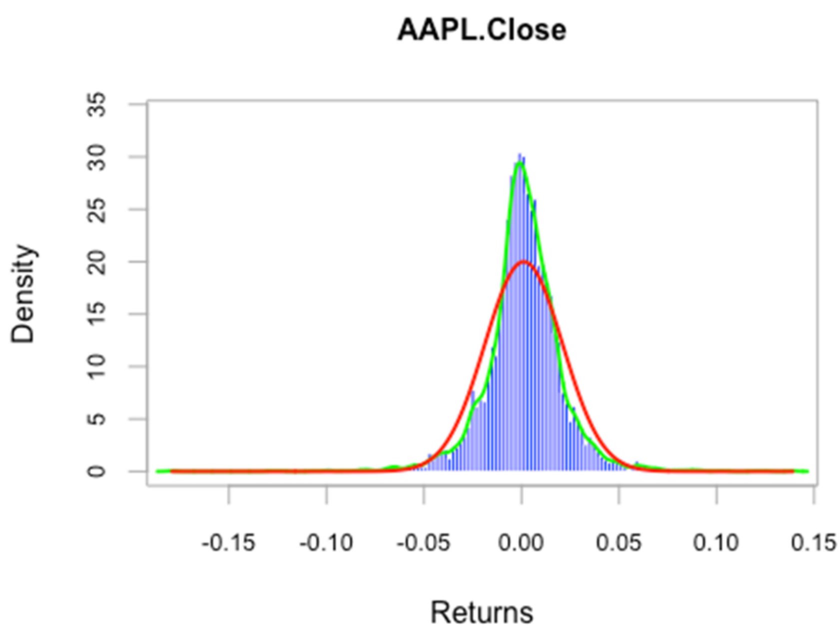
2008-01-03	0.0004620201
2008-01-04	-0.0763351531
2008-01-07	-0.0133851044
2008-01-08	-0.0359717390
2008-01-09	0.0475913376
2008-01-10	-0.0076923521
2008-01-11	-0.0299404343
2008-01-14	0.0352655047



Histogram of return

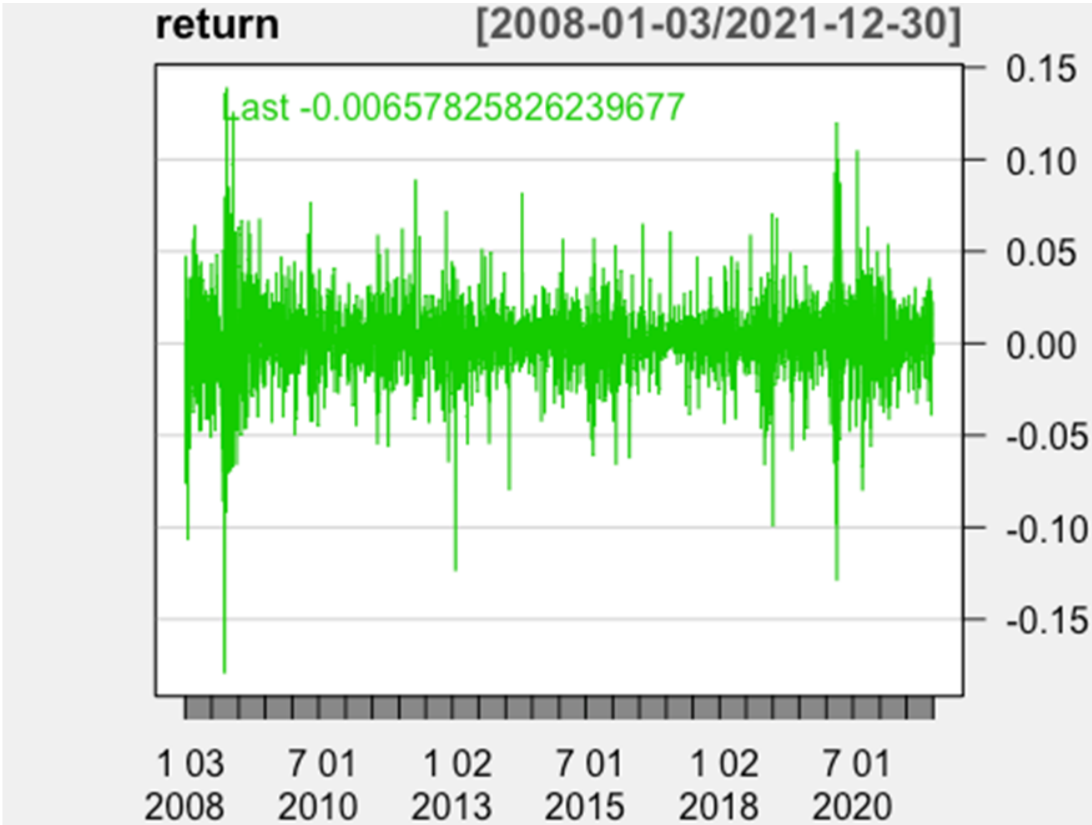
From this histogram we can see that the average return is approximately 0, but there are some with high or low return.

We can also compare how close this histogram is to the normal distribution:



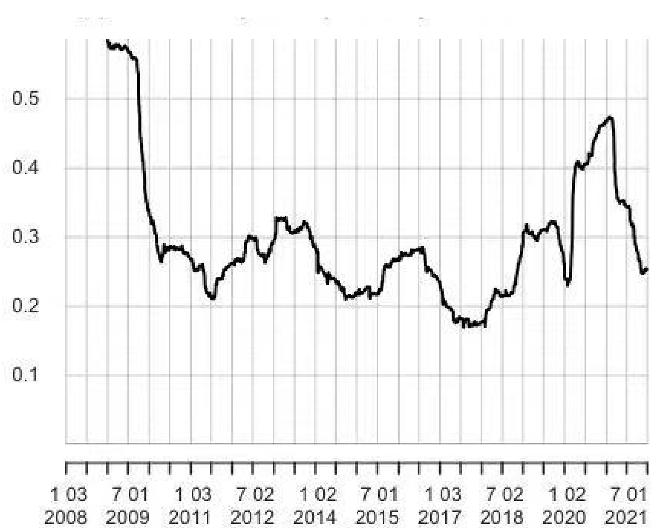
From the histogram, we can see that the curve of return is higher than normal distribution. And there are some areas on the left and right side of the histogram that green curve is larger than the normal distribution. This means that we will see more very high and low returns compared to normal distribution.

Finally, we plot the series:



From the graph, we can see that this is stationary time series, and there are a lot of fluctuations. We has especially a lot of volatility in 2008 and 2020.

This is because the market become volatile due to the financial crisis in 2008 and COVID-19 in 2020



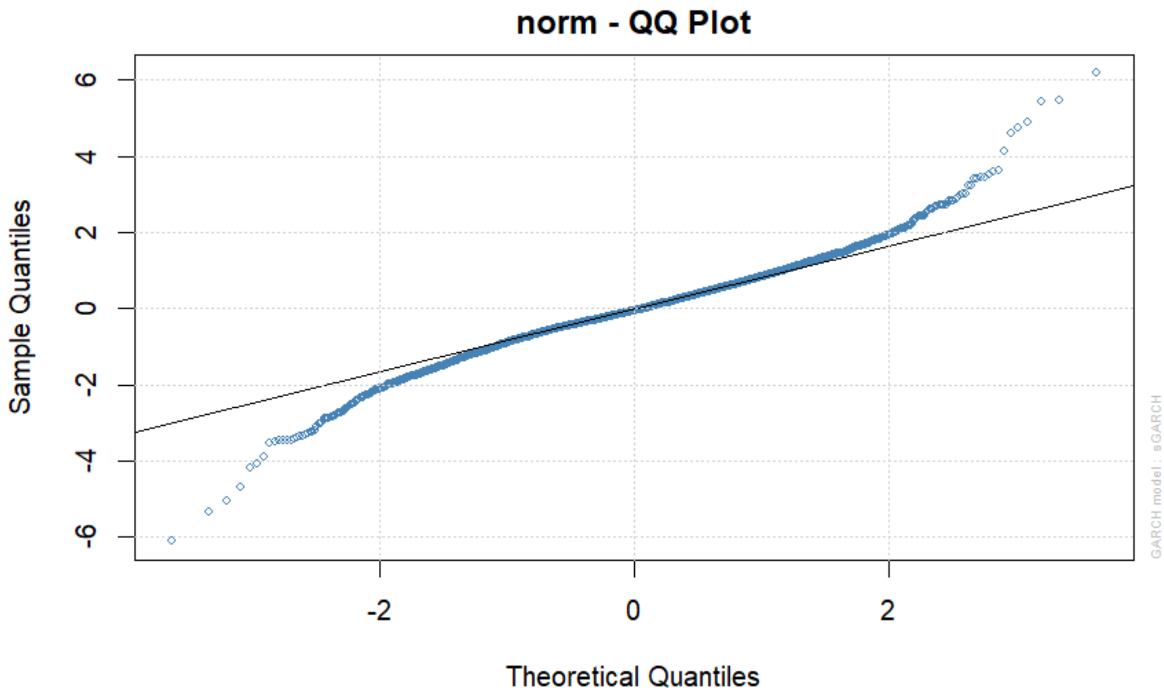
Analysis:

sGARCH model with constant mean

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

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(1,0,1)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.001917    0.000270   7.1067 0.000000
ar1     0.849663    0.243760   3.4857 0.000491
ma1    -0.855458    0.239052  -3.5785 0.000346
omega   0.000015    0.000003   4.5018 0.000007
alpha1  0.116317    0.017722   6.5634 0.000000
beta1   0.844153    0.011571  72.9545 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.001917    0.000454  4.22192 0.000024
ar1     0.849663    0.141903  5.98764 0.000000
ma1    -0.855458    0.140102 -6.10598 0.000000
omega   0.000015    0.000015  0.99724 0.318649
alpha1  0.116317    0.075711  1.53634 0.124456
beta1   0.844153    0.023573 35.80983 0.000000
```

The model parameters are shown above. The p-values of the parameters are less than 0.05, which means that all of the parameters are statistically significant. So the equations of our model are: R_T= 0.001917+ e_T and σ^2=0.000015+0.1116317e2+0.844153σ^2

Then we plot a QQ-Plot of Standardized Residuals.

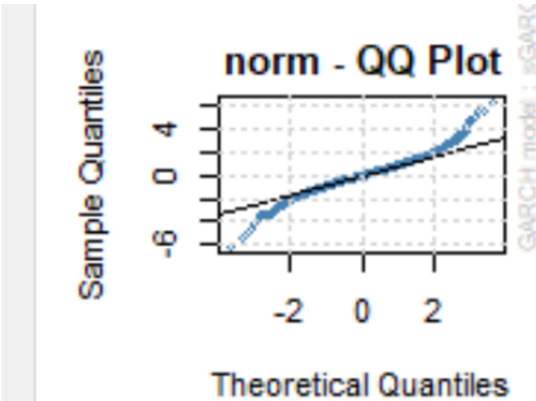


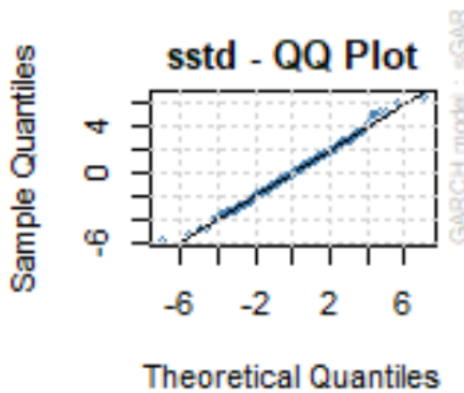
From the plot, we found that the tails deviate from the normal distribution significantly both for lower and higher values.

Model Evaluation:

After finding out normal distribution may not fit our data well. And the residual in the QQ plot is not distributed as a normal distribution. This residual looks like a skewed distribution from our perspective. Therefore, we decided to use skewed student T distribution on the data.

The residual in the QQ Plot of skewed student T distribution is much more fitted than normal distribution by looking at two plots below. Despite that, we have a Ljung-Box test on our standardized residuals. It shows that the p values of all residuals are all larger than 0.05, which means they are not in the 95% confidence interval. Therefore, we conclude we fail to reject that the residuals are correlated. In the adjusted Pearson Goodness-of-fit, our null hypothesis is the Skewed Student T Distribution is better than the normal distribution. It also emphasizes the failed rejection of our new distribution.

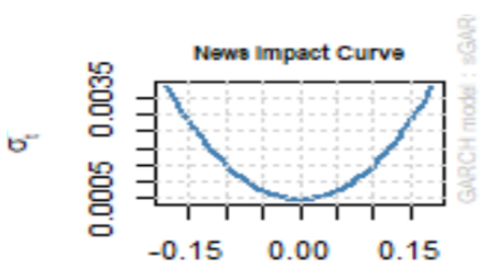


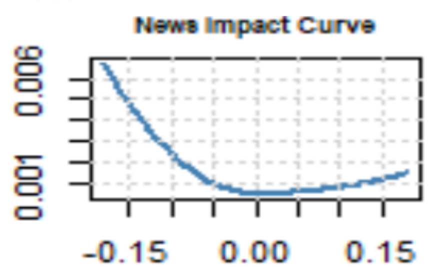



```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.555 0.21238
Lag[2*(p+q)+(p+q)-1][5]     4.208 0.03848
Lag[4*(p+q)+(p+q)-1][9]     7.946 0.06037
d.o.f=2
H0 : No serial correlation
```

```
Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     21.06       0.3334
2    30     30.10       0.4091
3    40     39.61       0.4427
4    50     51.14       0.3896
```

In previous data analysis, the Bear market seems to have a larger impact on volatility of the market. Instead of the standard Garch model, our group also tried the GJR-Garch model; it offers all the functions of the Garch model but includes leverage effects. The higher the fluctuation the bear market brings to volatility, the model is more levered. Considering how negative news affects volatility is necessary in our Garch model.






```
Information Criteria
------------------------------------

Akaike       -5.3468
Bayes        -5.3310
Shibata      -5.3468
Hannan-Quinn -5.3412
```

```
Information Criteria
------------------------------------

Akaike       -5.3305
Bayes        -5.3165
Shibata      -5.3305
Hannan-Quinn -5.3255
```

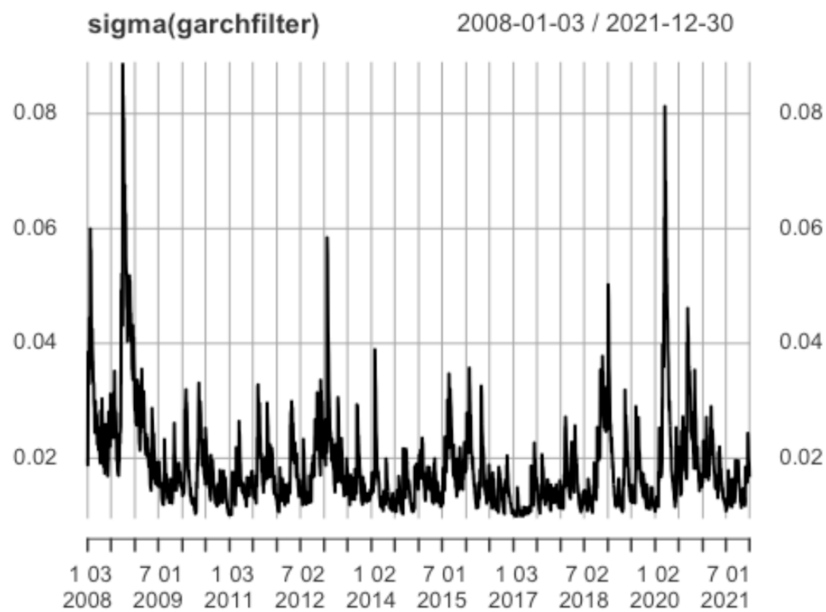
Simulating Stock prices

Approach 1:

Use the validated GARCH model in production to analyze the recent dynamics in the mean and volatility, so we can make predictions about the future mean and volatility

Step 1: Define the final model specs

Step 2: Analysis of past mean and volatility dynamics



2008: Global financial crisis 2020: Covid-19

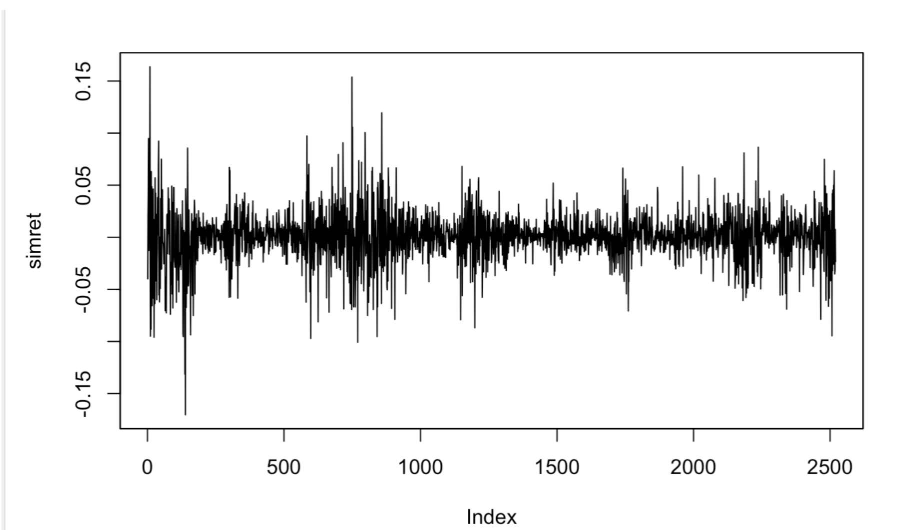
Step 3: Make predictions about future returns

```
##         2021-12-30 2021-12-30
## T+1  0.001325347 0.01623213
## T+2  0.001370615 0.01638451
## T+3  0.001370874 0.01653195
## T+4  0.001370875 0.01667468
## T+5  0.001370875 0.01681290
## T+6  0.001370875 0.01694679
## T+7  0.001370875 0.01707652
## T+8  0.001370875 0.01720227
## T+9  0.001370875 0.01732420
## T+10 0.001370875 0.01744244
```

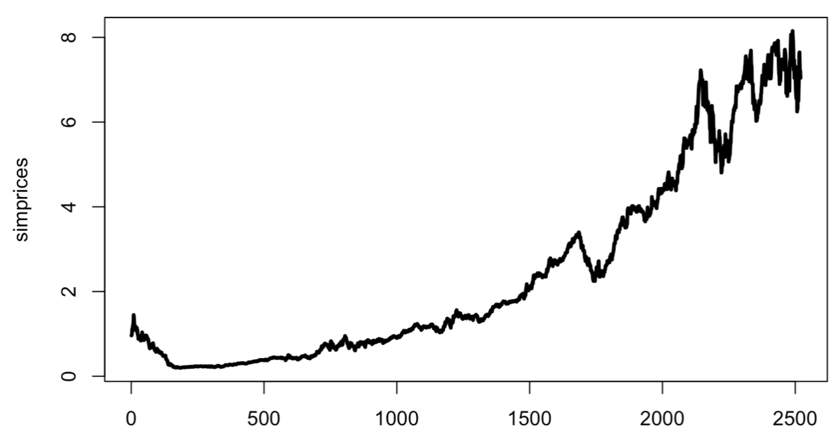
Approach2 :

We also used the complete model spec to simulate artificial log-returns, defined as the difference between the current log-price and the past log-price.

Simulated returns:



We can then plot the simulated stock prices



We observed that stock returns tend to have an asymmetric distribution instead of normal distribution. Even though it's unlikely that we can find an optimal model to fit an environment with high volatility and randomness like the stock market, we can still make predictions about the time when the next financial crisis is going to occur or the general trend of stock prices in the long run etc.

Date	Open	High	Low	Close*	Adj Close**	Volume
Jan 28, 2022	165.71	170.35	162.80	170.33	170.11	179,935,700
Jan 27, 2022	162.45	163.84	158.28	159.22	159.02	121,954,600
Jan 26, 2022	163.50	164.39	157.82	159.69	159.49	108,275,300
Jan 25, 2022	158.98	162.76	157.02	159.78	159.58	115,798,400
Jan 24, 2022	160.02	162.30	154.70	161.62	161.41	162,294,600
Jan 21, 2022	164.42	166.33	162.30	162.41	162.20	122,848,900
Jan 20, 2022	166.98	169.68	164.18	164.51	164.30	91,420,500
Jan 19, 2022	170.00	171.08	165.94	166.23	166.02	94,815,000
Jan 18, 2022	171.51	172.54	169.41	169.80	169.58	90,956,700
Jan 14, 2022	171.34	173.78	171.09	173.07	172.85	80,440,800
Jan 13, 2022	175.78	176.62	171.79	172.19	171.97	84,505,800
Jan 12, 2022	176.12	177.18	174.82	175.53	175.31	74,805,200
Jan 11, 2022	172.32	175.18	170.82	175.08	174.86	76,138,300
Jan 10, 2022	169.08	172.50	168.17	172.19	171.97	106,765,600
Jan 07, 2022	172.89	174.14	171.03	172.17	171.95	86,709,100
Jan 06, 2022	172.70	175.30	171.64	172.00	171.78	96,904,000
Jan 05, 2022	179.61	180.17	174.64	174.92	174.70	94,537,600
Jan 04, 2022	182.63	182.94	179.12	179.70	179.47	99,310,400
Jan 03, 2022	177.83	182.88	177.71	182.01	181.78	104,487,900
Dec 31, 2021	178.09	179.23	177.26	177.57	177.34	64,062,300

Conclusion:

Overall, we compared normal distribution and skewed student T distribution. We also tried both the Garch model and GJR-Garch model. Based on the stock price of APPLE from 2008 to 2021, our project simulates the change of volatility of time series variables. It shows high volatility in 2008 and 2021.

Volatility is not always a bad thing, as it can sometimes provide entry points from which investors can take advantage. Hence, this model can help an investor to estimate the fluctuations that may happen in the future which help them make decisions.

Reference:

https://bookdown.org/ccolonescu/RPoE4/time-varying-volatility-and-arch-models.html

https://rpubs.com/cyobero/arch

https://finance.yahoo.com/quote/AAPL?p=AAPL&.tsrc=fin-srch

https://www.princeton.edu/~yacine/leverage.pdf