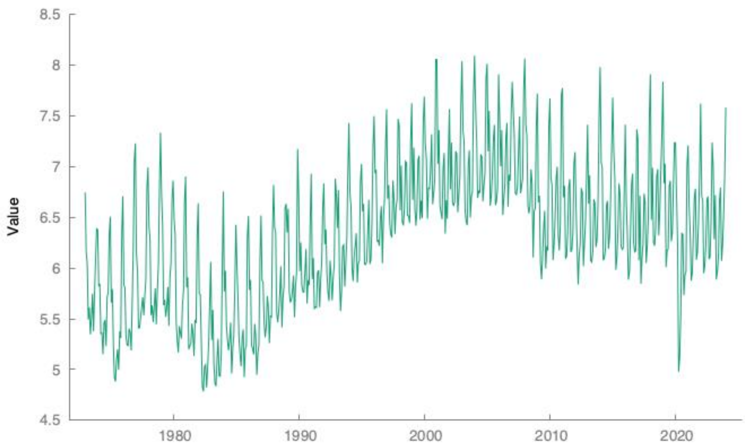
Problem 1- Forecasting energy variables

Time Series Analysis of Total Fossil Fuels Consumption

"Total Fossil Fuels Consumption" represents the aggregated usage of coal, natural gas, and petroleum products. The time series analysis, shown in Figure 1, exhibits a general upward trend interspersed with seasonal fluctuations. These variations likely reflect seasonal changes in energy demand.

Figure 1: Time Series Plot of Total Fossil Fuels Consumption

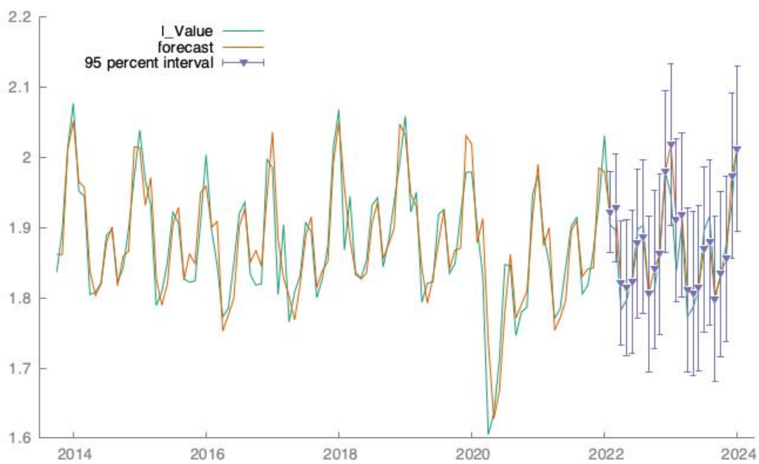


ARIMA Model Selection and Forecasting

To stabilize the variance and achieve stationarity, we first transformed the data using natural logs and differencing. Using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), we determined the most appropriate ARIMA model parameters. The diagnostic checks, such as the Ljung-Box test, affirmed the adequacy of our chosen ARIMA model.

The residuals of the ARIMA model, which confirm the absence of significant autocorrelation, are depicted in Figure 2. The forecast for the subsequent 24 months indicates a sustained increase in fossil fuel consumption, accounting for seasonal variations.

Figure 2: Residuals from ARIMA Model Forecasting



VAR Model Analysis and Comparison

To augment our analysis, we utilized a Vector Autoregression (VAR) model, integrating two additional economic indicators: GDP growth rate and Industrial Production Index. Figure 3 shows the forecast comparison between the ARIMA model and the VAR model, illustrating the VAR model's enhanced capability to account for the interactions between economic factors and energy consumption.

Figure 3: Forecast Comparison between ARIMA and VAR Models

Mean Error	-0.013518
Root Mean Squared Error	0.031991
Mean Absolute Error	0.024407
Mean Percentage Error	-0.73284
Mean Absolute Percentage Error	1.3061
Theil's U2	0.43833
Bias proportion, UM	0.17856
Regression proportion, UR	0.068264
Disturbance proportion, UD	0.75318

Additional Insights from Related Time Series

Further, we analyzed two related time series, incorporating their dynamics into our VAR model to better understand the economic influences on fossil fuel consumption. Figure 4 and Figure 5 display the time series and forecast plots for GDP growth rate and Industrial Production Index, respectively.

Figure 4: Time Series Plot of GDP Growth Rate

```
Function evaluations: 36
Evaluations of gradient: 15

Model 6: ARMAX, using observations 1973:01-2022:01 (T = 589)
Estimated using BHHH method (conditional ML)
Dependent variable: l_Value

             coefficient    std. error      z       p-value
  ------------------------------------------------------------
  const       1.76651       0.00932970    189.3     0.0000    ***
  theta_1     0.688268      0.0321336      21.42    8.90e-102 ***
  time        0.00101887    6.65815e-05    15.30    7.36e-53  ***
  sq_time    -1.09775e-06   1.10735e-07    -9.913   3.65e-23  ***
  dm1         0.0407288     0.00742409      5.486   4.11e-08  ***
  dm2        -0.0670603     0.00971079     -6.906   4.99e-12  ***
  dm3        -0.0603407     0.0101096      -5.969   2.39e-09  ***
  dm4        -0.166151      0.0101205     -16.42    1.44e-60  ***
  dm5        -0.171613      0.00991508    -17.31    4.07e-67  ***
  dm6        -0.163064      0.0105285     -15.49    4.19e-54  ***
  dm7        -0.107398      0.0103102     -10.42    2.08e-25  ***
  dm8        -0.0969190     0.0100827      -9.612   7.09e-22  ***
  dm9        -0.177440      0.0106334     -16.69    1.63e-62  ***
  dm10       -0.141476      0.0103873     -13.62    3.04e-42  ***
  dm11       -0.119660      0.00720352    -16.61    5.77e-62  ***

Mean dependent var   1.837277   S.D. dependent var   0.113755
Mean of innovations -0.000111   S.D. of innovations  0.041696
R-squared            0.867566   Adjusted R-squared   0.864572
Log-likelihood       1035.698   Akaike criterion    -2039.396
Schwarz criterion   -1969.341   Hannan-Quinn        -2012.102

                        Real  Imaginary    Modulus  Frequency
  -----------------------------------------------------------
  MA
    Root  1          -1.4529     0.0000     1.4529     0.5000
  -----------------------------------------------------------
```

Figure 5: Forecast Plot of Industrial Production Index

```
VAR system, maximum lag order 8

The asterisks below indicate the best (that is, minimized) values
of the respective information criteria, AIC = Akaike criterion,
BIC = Schwarz Bayesian criterion and HQC = Hannan-Quinn criterion.
```

lags	loglik	p(LR)	AIC	BIC	HQC
1	2404.21585		-7.898234	-7.789014	-7.855733
2	2538.67140	0.00000	-8.312963	-8.138210	-8.244961
3	2627.39347	0.00000	-8.576507	-8.336222	-8.483004
4	2709.38722	0.00000	-8.817809	-8.511991	-8.698805
5	2806.50891	0.00000	-9.109120	-8.737771	-8.964615
6	2876.53898	0.00000	-9.310873	-8.873991	-9.140866
7	2933.48943	0.00000	-9.469387	-8.966972	-9.273879
8	2989.65422	0.00000	-9.625303*	-9.057356*	-9.404295*

Problem 2 – Modelling volatility

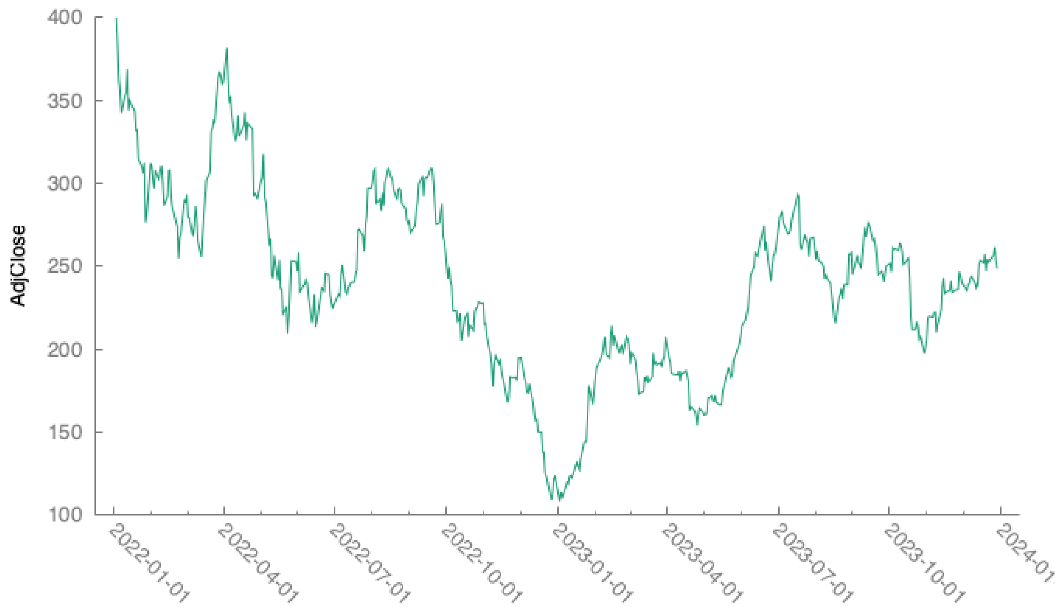
TSLA (01/01/2022 – 01/01/2024): The Product hereby described is the Tesla,Inc.

SUMMARY AFTER UPLOADING THE DATA

Summary statistics, using the observations 2022-01-03 - 2023-12-29 for the variable 'ld_AdjClose' (500 valid observations)

Mean	-0.00095184
Median	0.00099742
Minimum	-0.13059
Maximum	0.10436
Standard deviation	0.038102
C.V.	40.030
Skewness	-0.29333
Ex. kurtosis	0.74510
5% percentile	-0.068778
95% percentile	0.060470
Interquartile range	0.041742
Missing obs.	1

TIME PLOT GRAPH



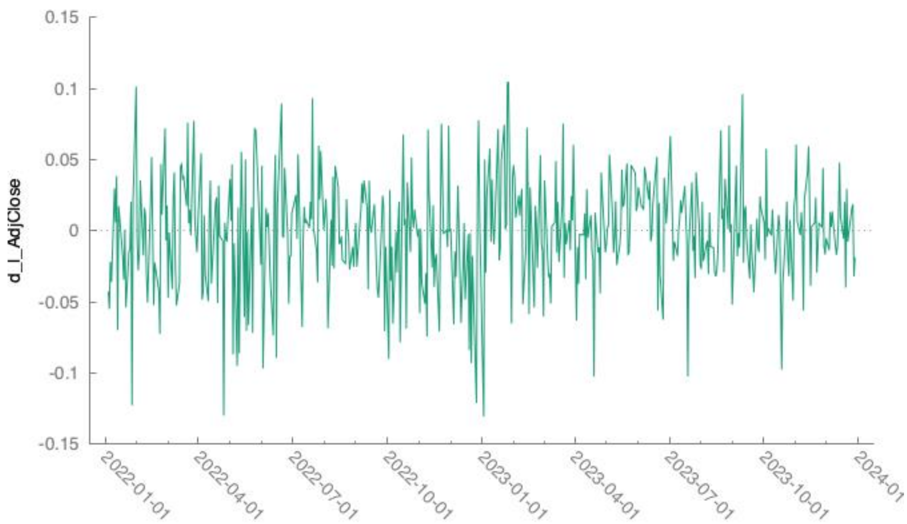
Model the Logarithm of AdjClose as an AR-1.

```
Model 1: ARMA, using observations 2022-01-04:2023-12-29 (T = 500)
Estimated using least squares (conditional ML)
Dependent variable: l_AdjClose

             coefficient   std. error      z      p-value
  --------------------------------------------------------
  const       0.0902113    0.0376501      2.396   0.0166  **
  phi_1       0.983283     0.00689713   142.6     0.0000  ***

Mean dependent var   5.452320   S.D. dependent var   0.244935
Mean of innovations  0.000000   S.D. of innovations  0.037917
R-squared            0.976084   Adjusted R-squared   0.976084
Log-likelihood       927.7092   Akaike criterion    -1851.418
Schwarz criterion   -1842.989   Hannan-Quinn        -1848.111

                        Real  Imaginary    Modulus  Frequency
  -----------------------------------------------------------
  AR
    Root  1           1.0170     0.0000     1.0170     0.0000
  -----------------------------------------------------------
```

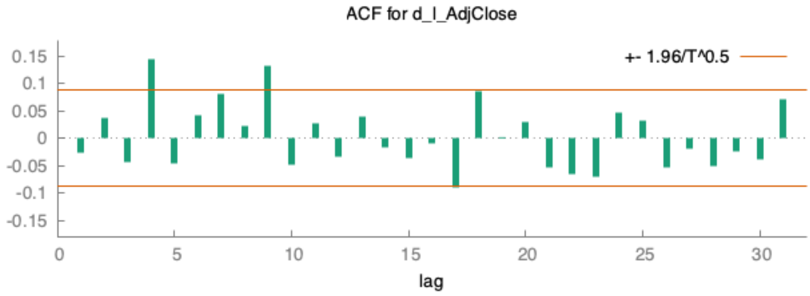
Our estimated coefficient is very close to 1.

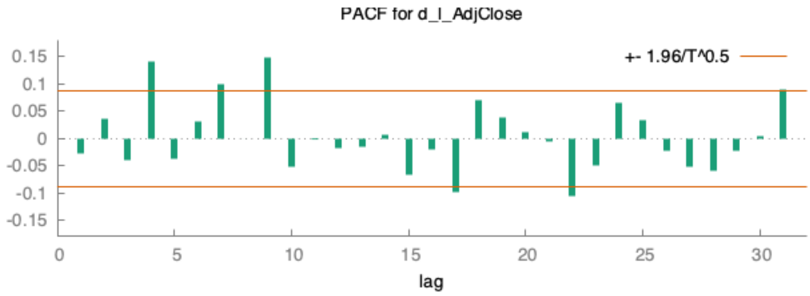
Time Series Plot of the Difference in Log-Transformed Returns:

- The Mean is 0 but we see a lot of variability around it.
- We do not see any visible pattern creating predictability of this variable.

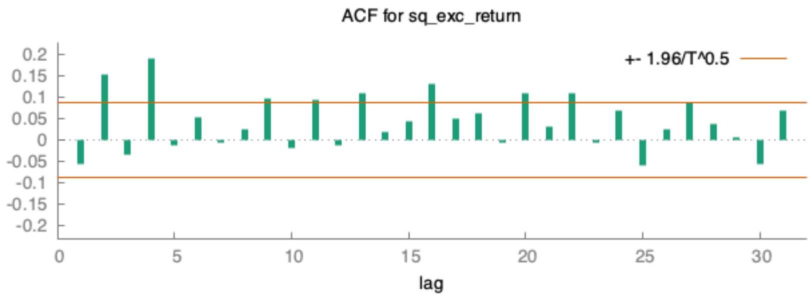


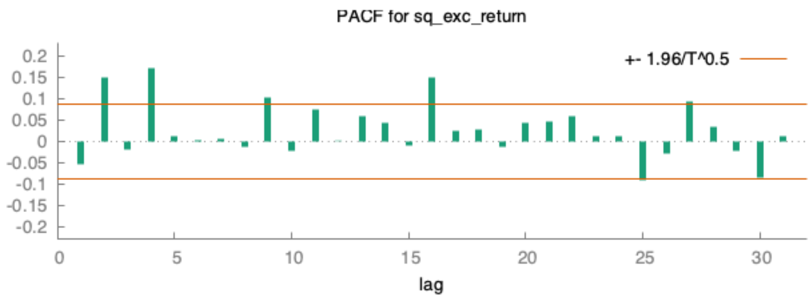
Try to show that the returns are not correlated with their own past.





<u>Correlogram of the Squared Excess Returns :</u>





GARCH MODEL:

1. Model Parameters:

- Constant (const): The estimated value is -0. 000528564 and a p-value of 0. 7483, indicating **it is not statistically significant**.

- Alpha(0): The coefficient is 2. 72254 x 10^{-5} (p=0.2308) and is not **statistically significant** (p-value > 0.001).

- Alpha(1): The coefficient is (0. 0319516) and a p-value of 0. 0157, suggesting that it is **statistically significant**.

- Beta(1) (β1): Coefficient: 0. 948037, P-value< 0.0001, this indicates high persistence in volatility.

2. Model Fit and Statistics:

- Mean of Dependent Variable Variance: -0. 000952 showing the average change in the dependent variable.

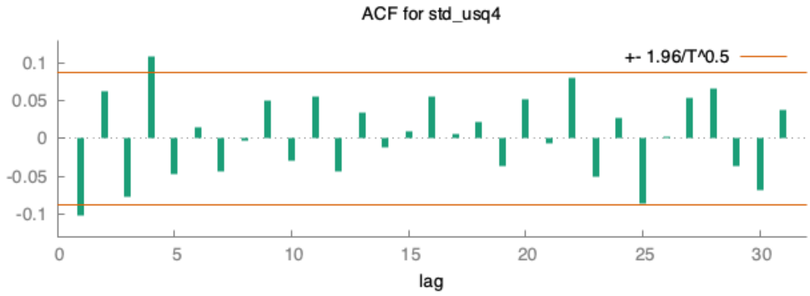
- Standard Error of Dependent Variable Variance : 0. 038102, showing the variation in the dependent variable around the mean.

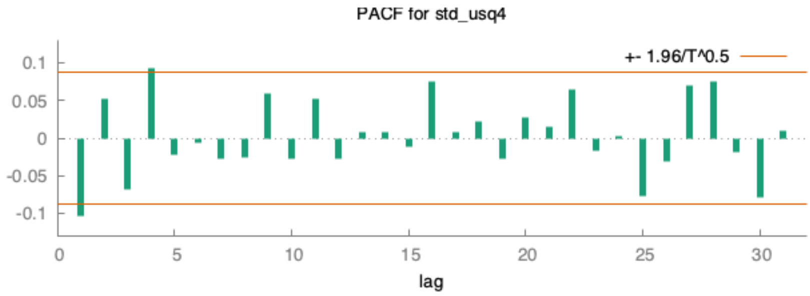
- Unconditional Variance of the Error: 0. 00136048\), with lower values indicating a better model relative to others.

- Likelihood Ratio Test for GARCH: A highly significant chi-square value of 15. 2984 with a p-value close to zero (0.000476415) suggests that including GARCH terms significantly improves the model's ability to capture the dynamics in the data.

Unconditional Standard Error : 0. 03688468516

```
Function evaluations: 114
Evaluations of gradient: 24

Model 4: GARCH, using observations 2022-01-04:2023-12-29 (T = 500)
Dependent variable: d_l_AdjClose
Standard errors based on Hessian

             coefficient    std. error       z      p-value
  ---------------------------------------------------------
  const      -0.000528564   0.00164696    -0.3209   0.7483

  alpha(0)    2.72254e-05   2.27209e-05    1.198    0.2308
  alpha(1)    0.0319516     0.0132303      2.415    0.0157   **
  beta(1)     0.948037      0.0233576     40.59     0.0000   ***

Mean dependent var  -0.000952   S.D. dependent var   0.038102
Log-likelihood       932.4265   Akaike criterion    -1854.853
Schwarz criterion   -1833.780   Hannan-Quinn        -1846.584

Unconditional error variance = 0.00136048
Likelihood ratio test for (G)ARCH terms:
  Chi-square(2) = 15.2984 [0.000476415]
```





FORECASTING:

The Forecast for the Next day (2024-02-01) has a standard deviation of 0.038102.

It is relatively close to the unconditional standard error 0.03688468516.