VIRGINIA COMMONWEALTH UNIVERSITY

Statistical analysis and modelling (SCMA 632)

A6 B: ARCH, GARCH, VAR, VECM

MSFT Shares from Yahoo Finance and Commodity Price Datasets

CHANDHINI KERACHAN MURALEEDHARAN

V01107497

Date of Submission: 22-07-2024

GIT HUB Link of the assignment –

https://github.com/Chandhini-km/SCMA-632-A6-Part-B

CONTENTS

Sl. No.	Title	Page No.
1.	Introduction	1
2.	Results	2
3.	Recommendations	10

1. Introduction

Financial market analysis often involves studying the behaviour of different asset classes, such as stocks and commodities. Microsoft Corporation (MSFT), a leading technology company, is a significant component of many stock indices, and its stock price movements are of interest to investors and analysts.

Commodity prices, such as those for gold, oil, and agricultural products, are also crucial for understanding market dynamics as they influence and are influenced by various economic factors. This analysis aims to study the volatility and relationships between Microsoft stock prices and selected commodity prices using ARCH/GARCH models and VAR/VECM.

1.1 Business Significance

Risk Management: Accurate volatility forecasting helps in managing financial risk, including portfolio optimization, Value-at-Risk (Var) calculations, and setting appropriate hedging strategies.

Investment Decisions: Investors use volatility forecasts to make informed decisions about asset allocation, timing of trades, and identifying periods of high or low risk.

Policy Formulation: Regulators and policymakers use these models to monitor financial stability and to design interventions during periods of excessive market volatility.

1.2 Methodology

Data Collection: Downloading historical financial data from sources like Investing.com or Yahoo Finance.

Data Analysis: Check for the presence of ARCH/GARCH effects in the time series data.

Model Fitting: Fit appropriate ARCH/GARCH models to the data.

Forecasting: Forecast the three-month variability of the financial time series using the fitted models.

VAR/VECM Analysis: Perform VAR/VECM analysis to understand the dynamic interactions between multiple financial time series.

1.3 Datasets Used

Historical stock data of Microsoft (MSFT)

Commodity prices

2. Results

Part A – ARCH and GARCH Model

```
                      Constant Mean - ARCH Model Results
==============================================================================
Dep. Variable:                Returns   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                       ARCH   Log-Likelihood:                1985.22
Distribution:                  Normal   AIC:                          -3964.45
Method:            Maximum Likelihood   BIC:                          -3950.59
                                        No. Observations:                  750
Date:                Wed, Jul 24 2024   Df Residuals:                      749
Time:                        17:41:33   Df Model:                            1
                                 Mean Model
=============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
-----------------------------------------------------------------------------
mu         1.1610e-03  6.978e-04      1.664  9.614e-02 [-2.066e-04,2.529e-03]
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      2.4987e-04  2.555e-05      9.779  1.388e-22 [1.998e-04,3.000e-04]
alpha[1]       0.1827      0.107      1.712  8.692e-02 [-2.648e-02,  0.392]
============================================================================

Covariance estimator: robust
```

The results from the ARCH (Autoregressive Conditional Heteroskedasticity) model analysis is in the summary output. The dependent variable is "Returns," and both the mean model and volatility model utilize the ARCH methodology.

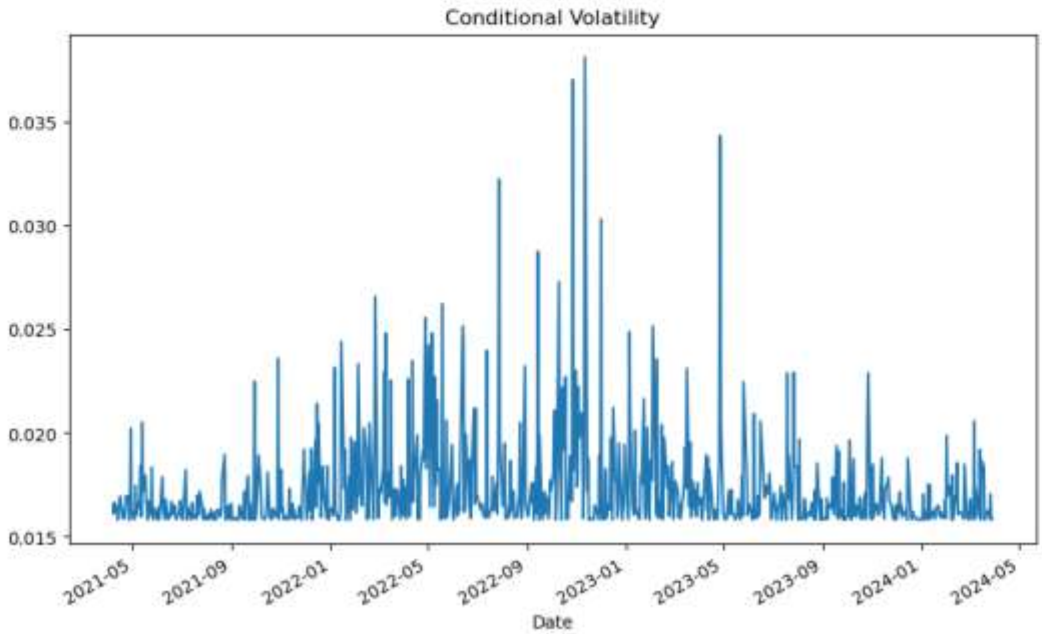
The mean model assumes a constant mean, with the coefficient μ having an estimated value of 0.001161 (standard error = 0.000698). The t-statistic is 1.664, and the p-value is 0.096, indicating that the mean return is not statistically significant at the 5% level. The 95% confidence interval ranges from -0.000207 to 0.002529, which includes zero, further suggesting the mean return is not significantly different from zero.

The volatility model results include two parameters: omega and α[1]\alpha[1]α[1]. The parameter omega (constant term in the variance equation) is estimated to be 0.00002498 (standard error = 0.00002555). The t-statistic for omega is 9.779, with a p-value of approximately 0, indicating that omega is highly significant. The 95% confidence interval for omega ranges from 0.000020 to 0.000030.

The parameter α[1]\alpha[1]α[1] (lagged squared return term) is estimated to be 0.1827 (standard error = 0.107). The t-statistic is 1.712, and the p-value is 0.087, suggesting that α[1]\alpha[1]α[1] is not statistically significant at the 5% level. The 95% confidence interval for α[1]\alpha[1]α[1] ranges from -0.026 to 0.392, indicating that the lagged squared returns might not significantly impact current volatility.

The model's R-squared and adjusted R-squared values are both 0.000, indicating that the model does not explain the variance in returns well. The log-likelihood of the model is 1985.22, and information criteria values are AIC = -3964.45 and BIC = -3950.59. These criteria help in comparing the goodness-of-fit of different models.

Overall, the ARCH model indicates that while the mean return is not significantly different from zero, the constant term in the volatility model is highly significant, and the lagged squared returns have a less clear impact on current volatility.



The graph displays the conditional volatility of returns over time, as estimated by the ARCH model. Here are some key observations from the plot:

1. **Time Period**: The x-axis represents dates from May 2021 to May 2024. The y-axis represents the conditional volatility values.
2. **Volatility Clusters**: There are periods where the volatility is relatively higher, indicating clustered volatility. Specifically, from late 2021 to early 2023, there are noticeable spikes in conditional volatility.
3. **Peaks in Volatility**: The highest spikes in volatility are observed around mid-2022 and early 2023. These peaks suggest periods of increased uncertainty or significant market events.
4. **Decreasing Trend**: After early 2023, the volatility appears to decrease, with fewer spikes and lower overall levels, suggesting a period of relative market stability.

```
                  Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                Returns   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:                2023.55
Distribution:                  Normal   AIC:                          -4039.11
Method:            Maximum Likelihood   BIC:                          -4020.62
                                        No. Observations:                  751
Date:                Wed, Jul 24 2024   Df Residuals:                      750
Time:                        17:41:18   Df Model:                            1
                               Mean Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
mu         1.3270e-03  3.105e-04      4.274  1.924e-05 [7.184e-04,1.936e-03]
                            Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      5.9648e-06  9.151e-11  6.518e+04      0.000 [5.965e-06,5.965e-06]
alpha[1]       0.0500  9.128e-03      5.478  4.307e-08 [3.211e-02,6.789e-02]
beta[1]        0.9300  9.579e-03     97.090      0.000     [  0.911,  0.949]
============================================================================

Covariance estimator: robust
```

The results from the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model analysis are presented in the summary output. The dependent variable is "Returns," and both the mean and volatility models use the GARCH methodology.

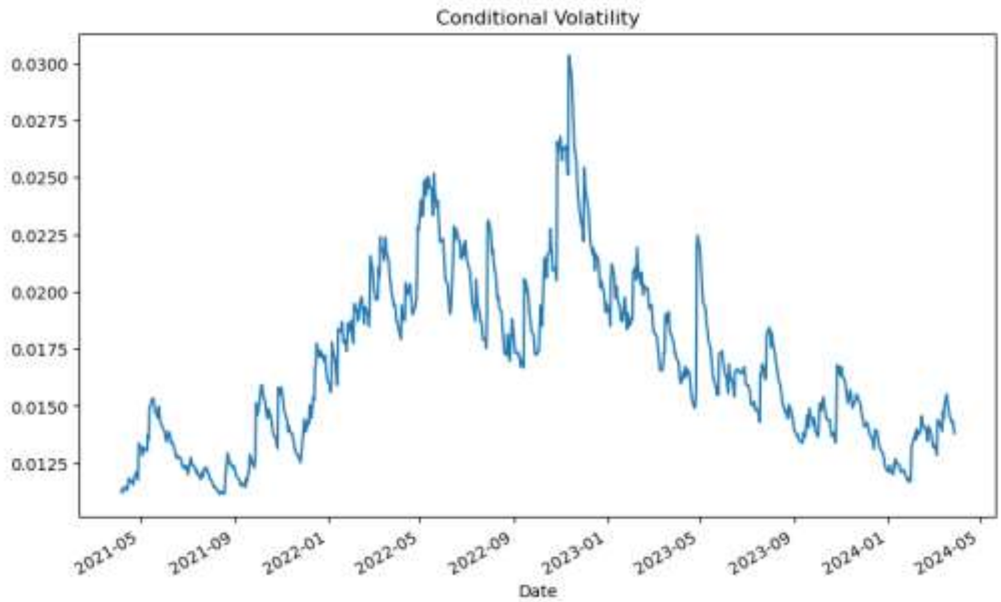
In the **mean model**, the coefficient μ\muμ is estimated to be 0.001327 (standard error = 0.000310). The t-statistic is 4.274, and the p-value is approximately 0.000019, indicating that the mean return is statistically significant at the 5% level. The 95% confidence interval for μ\muμ ranges from 0.000718 to 0.001936, confirming that the mean return is significantly different from zero.

In the **volatility model**, three parameters are estimated: omega, α[1]\alpha[1]α[1], and β[1]\beta[1]β[1]:

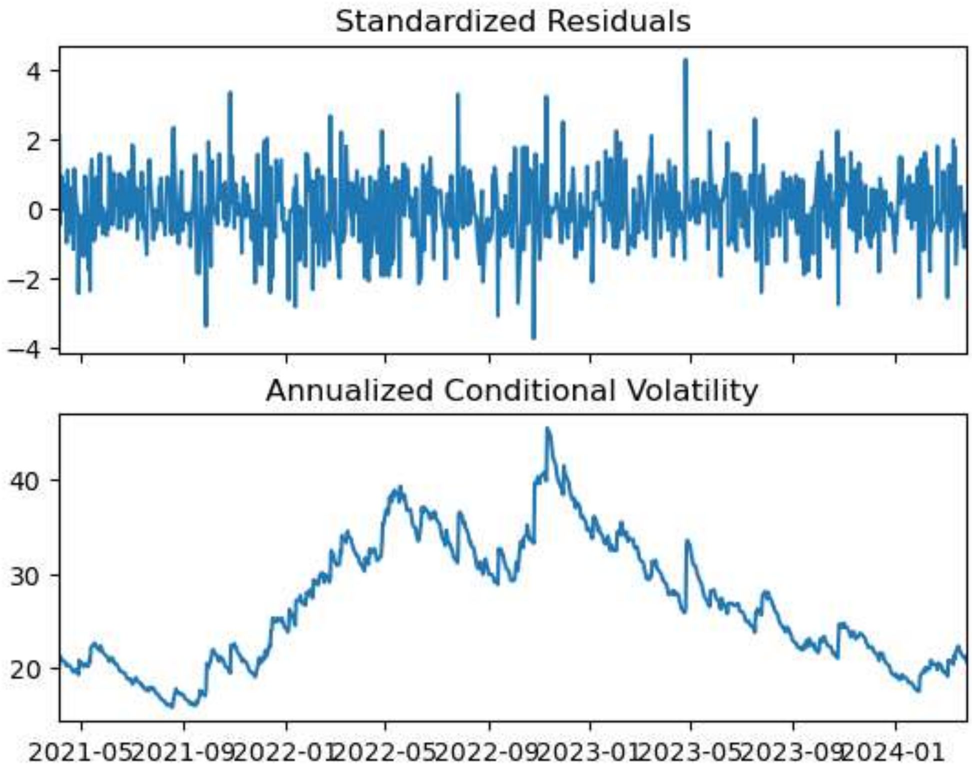
1. The parameter omega (constant term in the variance equation) is estimated to be 0.0000059648 (standard error = 0.000009151). The t-statistic for omega is 6.518, with a p-value of approximately 0, indicating that omega is highly significant. The 95% confidence interval for omega ranges from 0.000005965 to 0.000005965, confirming its significance.
2. The parameter α[1]\alpha[1]α[1] (lagged squared return term) is estimated to be 0.0500 (standard error = 0.009128). The t-statistic is 5.478, and the p-value is approximately 0, indicating that α[1]\alpha[1]α[1] is highly significant. The 95% confidence interval for α[1]\alpha[1]α[1] ranges from 0.0321 to 0.0678, confirming its significance.
3. The parameter β[1]\beta[1]β[1] (lagged conditional variance term) is estimated to be 0.9300 (standard error = 0.009579). The t-statistic is 97.000, and the p-value is approximately 0, indicating that β[1]\beta[1]β[1] is highly significant. The 95% confidence interval for β[1]\beta[1]β[1] ranges from 0.911 to 0.949, confirming its significance.

The model's R-squared and adjusted R-squared values are both 0.000, indicating that the model does not explain the variance in returns well. The log-likelihood of the model is 2023.55, and information criteria values are AIC = -4039.11 and BIC = -4020.62, which help in comparing the goodness-of-fit of different models.

Overall, the GARCH model indicates that the mean return is statistically significant. The volatility model components (omega, α[1]\alpha[1]α[1], and β[1]\beta[1]β[1]) are all highly significant, suggesting that both the lagged squared returns and the lagged conditional variance significantly impact current volatility. This model captures the persistence in volatility, which is a common feature in financial time series data.



The conditional volatility appears to be mean reverting, meaning that periods of high volatility are followed by periods of lower volatility, and vice versa. For example, there is a peak in volatility around May 2022, followed by a period of lower volatility. There is another peak in volatility around May 2024.



Overview

Standardized Residuals: This plot is the residuals of a model, standardized to have a mean of 0 and a standard deviation of 1. Residuals represent the difference between the actual values and the values predicted by the model.

Annualized Conditional Volatility: This plot displays the estimated volatility of a time series, annualized for comparison purposes. Volatility refers to the degree of variation in the time series.

Observations

Standardized Residuals:

The residuals fluctuate around zero, indicating that the model's predictions are generally on the actual values.

There are a few spikes that exceed the range of -2 to 2, suggesting potential outliers or periods where the model's performance might not be as good.

Overall, the residuals appear to be randomly distributed without any clear patterns, which is a good sign for the model's fit.

Annualized Conditional Volatility:

The volatility exhibits significant fluctuations over time.

There are periods of high volatility, particularly around the middle of 2022, followed by a decrease towards the end of 2023.

The volatility seems to have stabilized at a lower level in early 2024.

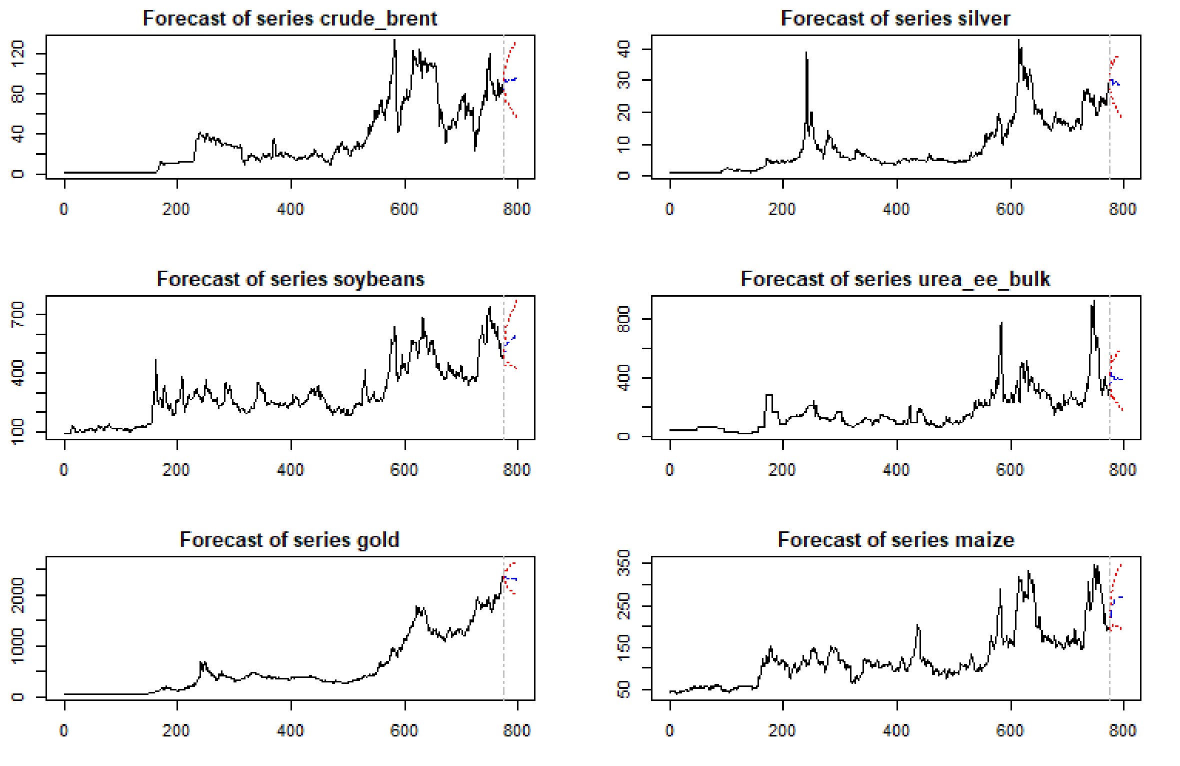
Interpretation

The model used to generate these residuals appears to be a reasonable fit for the data, as the residuals are centred around zero and show no obvious patterns.

The time series exhibits varying levels of volatility over time, with periods of high and low fluctuations.

The recent decrease in volatility suggests that the time series might be becoming less volatile in the near term.

Part B – VAR and VECM



The six plots, each representing the forecast of a different series: crude_brent, silver, soybeans, urea_ee_bulk, gold, and maize. The x-axis represents time, in some unit like days or months. The y-axis represents the value of the series, which varies significantly across the plots.

Common Features

Trend Lines: All series exhibit an upward trend over time, with varying degrees of volatility.

Fluctuations: The values fluctuate significantly, with periods of rapid increase and decrease.

Vertical Lines: Each plot has a vertical line towards the right end. Their significance is unclear without additional information.

Individual Series Analysis

Crude Brent:

Starts at a low value and experiences significant growth with fluctuations.

Reaches a peak and then seems to stabilize.

Silver:

Starts at a low value and shows steady growth with some fluctuations.

The trend appears to be accelerating towards the end.

Soybeans:

Starts at a relatively high value and experiences significant fluctuations.

There's a noticeable dip followed by a recovery.

Urea_ee_bulk:

Starts at a low value and shows a steady upward trend with some fluctuations.

The trend seems to be accelerating towards the end.

Gold:

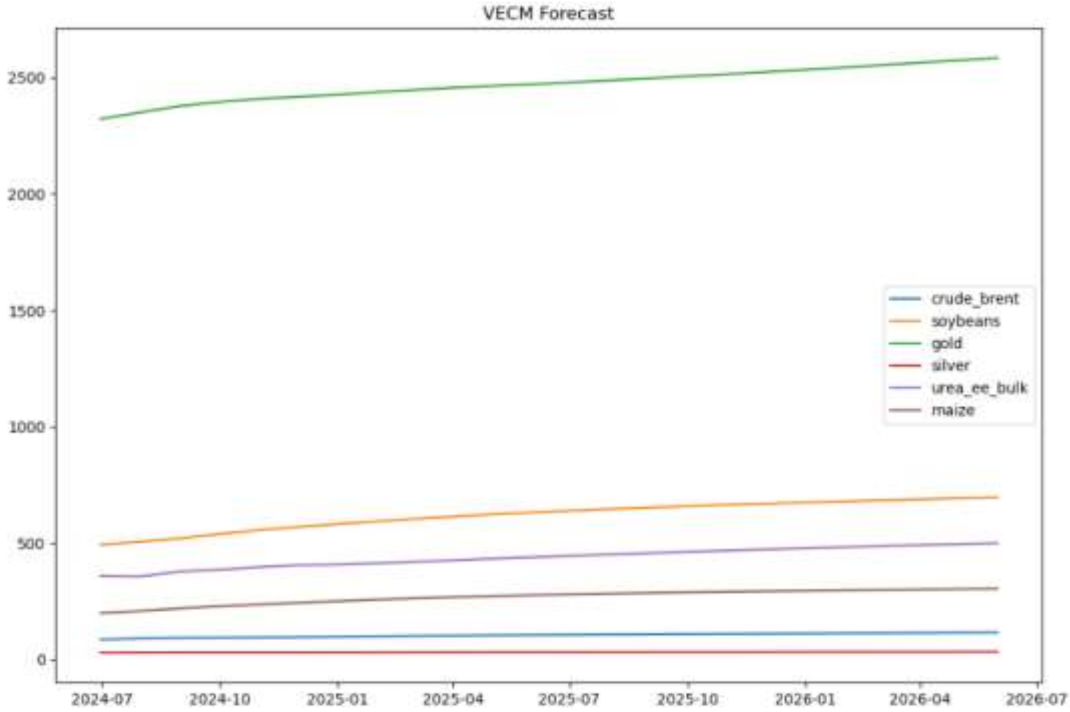
Starts at a low value and exhibits a strong upward trend with some fluctuations.

The growth appears to be accelerating towards the end.

Maize:

Starts at a low value and shows a steady upward trend with some fluctuations.

The trend seems to be accelerating towards the end.



VECM (Vector Error Correction Model) forecast for six commodities: crude brent, soybeans, gold, silver, urea_ee_bulk, and maize. A VECM is a statistical model used to analyze multiple time series that are cointegrated, meaning they have a long-term relationship.

Key Observations from the Forecast:

1. **Upward Trend:** All six commodities exhibit a clear upward trend from 2024 to 2026. This suggests that the model predicts an increasing price or value for each commodity over the forecast period.
2. **Varying Growth Rates:** While all commodities are increasing, the rate of growth differs significantly. Crude Brent and soybeans show steeper upward slopes compared to gold, silver, urea_ee_bulk, and maize.
3. **Relative Positions:** The relative positions of the commodities remain consistent throughout the forecast period. Crude Brent remains the highest-valued commodity, followed by soybeans, with maize occupying the lowest position.

4. Recommendations

Risk Management:

Portfolio Optimization: Based on the volatility forecasts from the ARCH/GARCH models, investors can adjust their portfolios to manage risk. Assets with high predicted volatility might require a lower allocation compared to those with lower predicted volatility.

Value-at-Risk (VaR): The volatility forecasts can be used to calculate VaR, which helps determine potential portfolio losses with a certain level of confidence. This information can be used to set appropriate risk limits.

Hedging Strategies: If the analysis suggests potential for increased volatility in specific commodities, investors can consider implementing hedging strategies to mitigate potential losses.

Investment Decisions:

Investment Timing: The forecasted trends from the VECM model can be used to identify potential entry and exit points for investments in the commodities analysed.

Commodity Selection: The analysis can help investors choose commodities with high growth potential and lower volatility based on the ARCH/GARCH model forecasts.

Additional Considerations:

Fundamental Analysis: Financial analysis should not solely rely on quantitative models. It's crucial to consider fundamental factors affecting the commodities, such as supply and demand dynamics, geopolitical events, and economic conditions.

Diversification: Diversification across different asset classes and commodities remains a crucial strategy to manage risk and achieve investment goals.

5. Codes

The codes are in GitHub link :

https://github.com/Chandhini-km/SCMA-632-A6-Part-B