GROUP WORK PROJECT # 2 MScFE 610: FINANCIAL ECONOMETRICS
GROUP NUMBER: 5338

FULL LEGAL NAME	**LOCATION (COUNTRY)**	**EMAIL ADDRESS**	**MARK X FOR ANY NON-CONTRIBUTING MEMBER**
Shivansh Kumar	India	shivansh.business23@gmail.com	
Vaibhav Sandeep Somani	Ireland	vaibhavssomani@gmail.com	
Kshitiz Sharma	India	drkshitizsharma@gmail.com	

Statement of integrity: By typing the names of all group members in the text boxes below, you confirm that the assignment submitted is original work produced by the group (excluding any non-contributing members identified with an "X" above).	
Team member 1	**Shivansh Kumar**
Team member 2	**Vaibhav Sandeep Somani**
Team member 3	**Kshitiz Sharma**

Use the box below to explain any attempts to reach out to a non-contributing member. Type (N/A) if all members contributed. **Note:** You may be required to provide proof of your outreach to non-contributing members upon request.
(N/A)

Data Collection & Processing

For our handbook we are going to use 5 years of daily return data from META, We are using Yahoo Finance API to to collect our data. After that, we calculate the daily returns, drop missing data, and properly format it.

Step 2:

The Four challenges that we have picked are:

- **Modeling randomness**
- **Autocorrelation**
- **Heteroskedasticity**
- **Non-stationarity**

Throughout this document we are going to address:-

Definition: Technical definition using formulas or equations
Description: Written explanation (1–2 sentences)
Demonstration: Numerical example using real-world data (or simulated data if not found)
Diagram: Visual example using real-world data (using same data as above)
Diagnosis: How to recognize or test that the problem exists
Damage: Clear statement of the damaged caused by the problem
Directions: Suggested models that can address this

For More detailed report please refer to the handbook submitted or you can also go to this colab link:

Group_5338_Group_Project_Work_2.ipynb

Modeling randomness

Definition

We will be understanding about the challenges of modeling of randomness but let's first understand what is randomness actually.

In time series data the unpredictable variation in the data points over time is called randomness. it is often represented as stochastic processes or random variables. we can mathematically define it using probability distribution like normal distribution.

Now let's understand it using a mathematical approach:

Let Randomness in time series X_t represented by random variable X with a normal distribution $N(\mu,\sigma^2)$, where:

- μ is the mean of normal distribution.
- σ^2 is the variance of the distribution.

So Now the PDF of the normal density function is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where,

x is the random variable., e is the base of the natural log., π is the mathematical constant.

This PDF formula states that the probability of observing a specific value x in the distribution, have the higher probability to be around the mean μ and decreasing probability as x shift away from mean according to the variance σ^2 .

Description

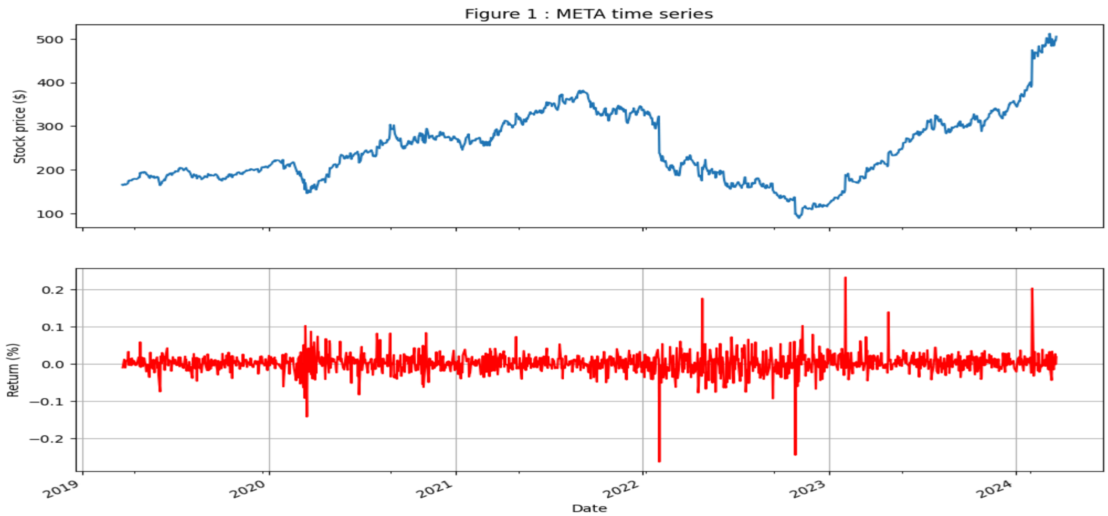
As famously said by "Nassim Nicholas Taleb" in his book (Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets) (Taleb):

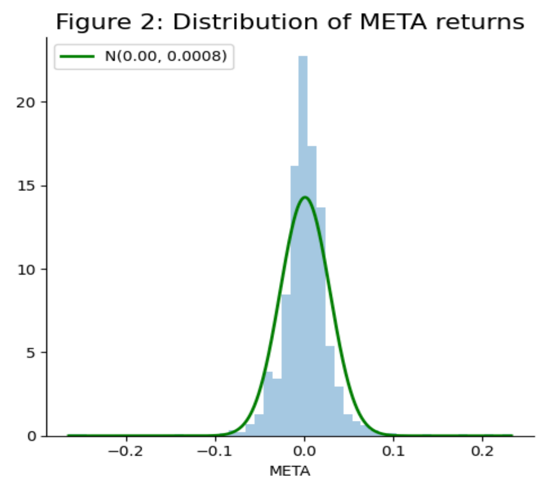
"No matter how sophisticated our choices, how good we are at dominating the odds, randomness will have the last word."

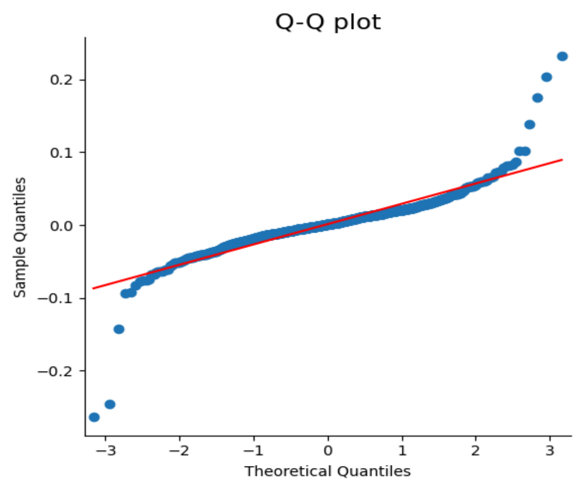
We also face a similar situation while dealing with financial time series data, Uncertainty is one of the most Prevailing factors while dealing with financial time series data which is random. And this makes it one the most important challenges in Quantitative finance as it's challenging to predict future values. That's why we need to consider randomness effectively which accounts for this uncertainty in forecasting processes.

Demonstration & Diagram

Now we will understand randomness in our financial time series data. To demonstrate this we are going to plot a time series plot of stock price as well as stock return to understand the variation in the time series data. We are also going to plot the histogram and calculate the descriptive statistics such as mean, median, std dev. etc to understand the variation and potential randomness.








```
---------- Descriptive Statistics ----------
Range of dates: 2019-03-21 - 2024-03-20
Number of observations: 1259
Mean: 0.0013
Median: 0.0011
Min: -0.2639
Max: 0.2328
Standard Deviation: 0.0279
Skewness: -0.2812
Kurtosis: 19.0142
```

Figure 3: Descriptive Statistics

As we can see in the above diagrams, in The first figure we can see the plot of stock prices of META as well as stock returns of META.

In the next diagram, we can see the histogram of stock return in a Normal distribution for Gaussian randomness with its QQ plot. At last, we can see the Descriptive Statistics of the stock return.

Now we are going to do a diagnosis of the problem in the next section.

Diagnosis

To diagnose the issues we are going to do visual inspection and try to understand the randomness behavior. In the first diagram as we can see the time series plot of stock prices we can say that it is following a random walk (Random walk suggest that the price changes in an assets are random) (Smith). and from the **Figure 1** we can also conclude that the stocks prices are random and unpredictable which shows potential randomness. Now we look into the stock return diagram in **Figure 1** we can see that it also shows a random pattern of returns over time as sometimes the returns are high sometimes they are low, it is moving with an unpredictable movement which shows randomness in stock returns.

After that if we see our histogram in **Figure 2** we can see negative skewness as skewness < 0 is equal to left skewed (Team Member 2: GWP1: Skewness) (Team member 2) also it has high kurtosis which is also confirmed by our QQ plot in **Figure 2** which shows heavy tails and from our descriptive statistic in **Figure 3** results we can confirm where skewness -0.2812 which is negative skewed and kurtosis is 19.0142 the major reason behind this is outlier which makes analysis challenging and also for test reasons we assumes the data distributions to normal but skewness violates these assumptions (Team Member 2: GWP1: Skewness) (Team member 2) and it signifies potential randomness in the data.

Also Visual inspection as the first layer of analysis is enough for a start of the analysis process but it's not tangible. The better approach to test randomness in a times series data is autocorrelation function (ACF) and partial autocorrelation function (PACF) analysis. If there are no significant autocorrelation or partial autocorrelation patterns beyond a few lags, it suggests randomness. We can also do Cross-validation, Residual analysis, Unit root tests etc.

Damage

The major damage due to failure in modeling time series data is misleading forecasts and as we have previously discussed uncertainty is the most Prevailing factor while dealing with time series data misleading forecasts can lead to Illusionary risk assessments which will lead to suboptimal decision making and due to all of this mess our castle of glass will be shattered into pieces with financial losses and missed opportunities.

Ignoring randomness or not correctly modeling randomness can result in understating or overstating risk associated with the assets that we are analyzing which can lead to inefficient resource allocation to a particular asset. At last, I would like to quote one more quote from Nassim Taleb's book (Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets) (Taleb):

“We favor the visible, the embedded, the personal, the narrated, and the tangible; we scorn the abstract.”

We humans have no control in our emotions, we only favor that which we can see, analyze or understand and we always ignore that which doesn't have any concrete existence. That's why we need to work with a more tangible way of modeling randomness which we will discuss in our next section.

Autocorrelation

Definition

Autocorrelation is a statistical concept that measures the degree of similarity between a given time series and its lagged version over successive time intervals (Smith and Murry).

The range of autocorrelation is -1 to 1.

We can calculate autocorrelation using the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF).

Description

Autocorrelation ranges from -1 to +1:

- +1: Perfect Positive Correlation
- 0: No Correlation
- -1: Perfect Negative Correlation

Autocorrelation can help reveal hidden patterns in the data, and analyze stationarity and seasonality.

For calculation of Autocorrelation:

We can use ACF and PACF. The ACF measures the linear relationship between an observation at time (t) and the observations at previous times, while the PACF measures the association between (y(t)) and (y(t-k)) while filtering out the linear influence of the random variables that lie in between.

Demonstration and Diagram

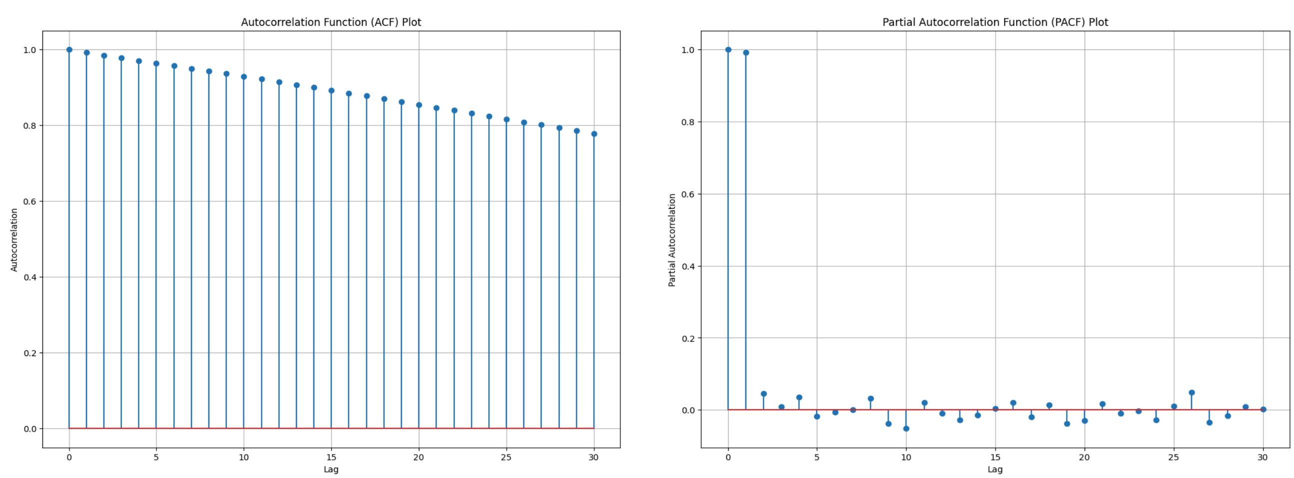


Figure: 4

As from the above two plots in **Figure 4**, we can see that :

In the ACF Plot, the autocorrelation values gradually decrease and will become insignificant after some lags, will indicate Tail- off.

In the PACF Plot, the autocorrelation value becomes insignificant after the first lag, indicating a Cut-off after one lag.

Thus, from the above two conditions we can conclude that the autoregressive model - AR(1) might be suitable for modeling this time series.

Diagnosis

There are various ways to diagnose the Autocorrelation in the time series. Here are a few of the most used ways:

1. **Visual Inspection of ACF Plot**: ACF Plot displays the autocorrelation between the series and its lagged version for different lags.
2. **Visual Inspection of PACF Plot**: The PACF plot focuses on the unique contribution of a specific lag, excluding the influence of earlier lags.
3. **Ljung-Box Test**: It is a statistical test to check if there is any overall autocorrelation.

   The Ljung–Box test may be defined as:

   - **H0**: The data are independently distributed
   - **Ha**: The data are not independently distributed, they exhibit autocorrelation.
4. Another test commonly used is the **Durbin-Watson test. (Rehal)**

From the above ways, we have already seen how to analyze the ACF and PACF plots in the above section. Here is how we can use the Ljung-Box Test:

```
lb_test = sm.stats.acorr_ljungbox(main_df['Adj Close'], lags=[10])
print(lb_test)
```

```
        lb_stat  lb_pvalue
10  11678.171192        0.0
```

A p-value is 0, typically less than 0.05, suggests that the null hypothesis of no autocorrelation can be rejected. Based on these results, there is very strong evidence of autocorrelation.

Damage

Autocorrelation, if not addressed efficiently, can cause the following issues in time series analysis:

- **Increased Variance of Estimates:** In the case of the autocorrelation, the variance of estimates becomes higher.
- **Incorrect Parameter Estimates:** Autocorrelation implies a relation between the observations, thus, it can lead to biased parameter estimates in the models.
- **Impact on Tests of Significance:** Due to the autocorrelation, the results of significance tests become unreliable, due to incorrect parameter estimates.
- Due to all of the above reasons, it becomes difficult for prediction and forecast based on the model.

Heteroskedasticity

Definition

One of the basic assumptions of the classical regression model while applying OLS is that the variance of the regression disturbance e is constant for all observations Var(e) = $\sigma 2$. The violation of this assumption is called heteroskedasticity (kmenta). In other terms variance of the residuals is not IID (independent and identically distributed) over a range of measured values.

e~N(0,σ2)

Multiple regression mean function as follows:

E(Y|X)=βX

Because the variance function is for multiple regression, Var(e) is actually a covariance matrix. When the error term variances are not constant, the covariance matrix for **n** observations will be like the following:

{Var}(e) = $$\begin{pmatrix} \delta_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \delta_2^2 & 0 & \cdots & 0 \\ 0 & 0 & \delta_3^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \delta_n^2 \end{pmatrix}$$

where $\delta 21 \neq \delta 22 \neq \delta 23 \neq \cdots \neq \delta 2n$

Description

Heteroskedasticity is known as heterogeneity of variance. when a variable is assumed to be homoscedastic when it is actually heteroscedastic, it will result in biased estimates of standard errors and overestimating the goodness of fit as measured by the Pearson coefficient. Heteroskedasticity can be plotted in a scatter plot where it results in an unequal scatter of the residuals, a cone or megaphone shape distribution. The heteroskedasticity is caused by large differences in smallest and largest data points (CFI).

Demonstration

Now we will focus on heteroskedasticity issues.

For this we need to create a time-series model. For time-series Model 1 the 'Adj Close' values of the main-df file are focused. For time-series Model 2 the 'META' value (Stock Return) of the main-df file are focused.

The Hypothesis for Heteroskedasticity:

$H0$:Homoscedasticity exists

$H1$:Heteroskedasticity exists

We run a Breusch-Pagan test to the hypothesis.

Model 1: Modeling Stock Price

Detrend, Model Fit and **Figure 5** - Residuals from Simple Regression of META Stock Price on Trend

OLS Regression Results

Dep. Variable:	Adj Close	**R-squared:**	0.111
Model:	OLS	**Adj. R-squared:**	0.110
Method:	Least Squares	**F-statistic:**	156.6
Date:	Mon, 25 Mar 2024	**Prob (F-statistic):**	5.95e-34
Time:	05:40:15	**Log-Likelihood:**	-7218.5
No. Observations:	1258	**AIC:**	1.444e+04
Df Residuals:	1256	**BIC:**	1.445e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	201.8838	4.242	47.591	0.000	193.561	210.206
x1	0.0731	0.006	12.515	0.000	0.062	0.085

Omnibus:	0.221	**Durbin-Watson:**	0.008
Prob(Omnibus):	0.895	**Jarque-Bera (JB):**	0.177
Skew:	0.027	**Prob(JB):**	0.915
Kurtosis:	3.022	**Cond. No.**	1.45e+03

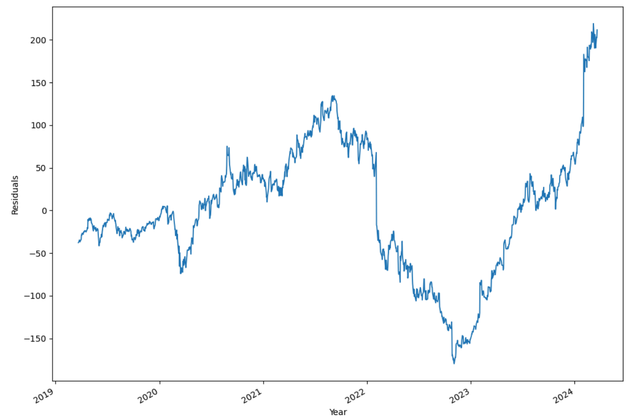


Figure: 5

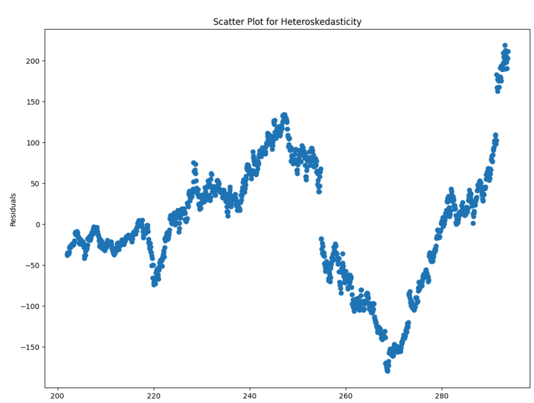
Heteroskedasticity test

Lagrange multiplier statistic	2.596865e+02
p-value	2.008427e-58
f-value	3.267172e+02
f p-value	4.303885e-65

Diagram

For deep analysis we will plot scatter plot for OLS Fitted Value and OLS Residuals to check presence of heteroskedasticity issue

Figure 6: Scatter Plot for OLS Fitted Values and OLS Residuals for Stock Price.



Diagnosis

Now let's analyze the Model and Diagrams :
To diagnose the issue of heteroskedasticity we will first analyze the OLS model result. The time-series model is built with one lagged time stamp of META stock price and META Return. The condition number is large which indicates that there could be problems of multicollinearity and heteroskedasticity.
Further, the Breusch-Pagan (BP) test is a formal statistical test to detect the issue. BP test checks the above hypothesis.This is a Chi-square test.The p -value (1.637852e-58) is less than the significant level of 0.05, we reject the null hypothesis. The residuals are not constant, so there is a heteroskedasticity issue in stock price data. The stock return does not show this issue as P-value (0.086) is more than the significance value(0.05).

In the scatter plot Now we look into the stock price diagram. It shows the increasing distribution from left to right, which means the error term is not constant in stock price data, and the presence of heteroskedasticity.
The scatter plot of stock return data is uniform from left to right. The error term is constant, no presence of heteroskedasticity.
Adding a trend variable in the model We can see the residuals from the simple regression model (Model 1)with a trend variable does not appear to exhibit stationarity. The time series of residuals shows what resembles a "V pattern," and therefore, it does not appear that the mean is constant over time. Thus, the residuals from the regression do not form a stationary time series.
Adding a trend variable does not seem to be enough to make META stock price into a stationary time series.

Damage

Lets understand the damage of heteroskedasticity in a time-series model :
The key assumptions of homoscedasticity are vital for OLS regression. OLS estimators are still unbiased, but they are no longer efficient. Inefficient estimators mean that the parameter estimates have higher variances, which reduces the precision of the estimates.
Heteroskedasticity can lead to incorrect inference, including incorrect standard errors, confidence intervals, and hypothesis tests. This can cause incorrect conclusions about the significance of coefficients and the overall goodness-of-fit of the model.
Models that do not account for heteroskedasticity may produce forecasts with underestimated or overestimated uncertainty, leading to inaccurate predictions of future values.
Overall the reliability of the model is at stake.

Figure: META Stock Price vs. Fitted META Stock Price Using First Differencing

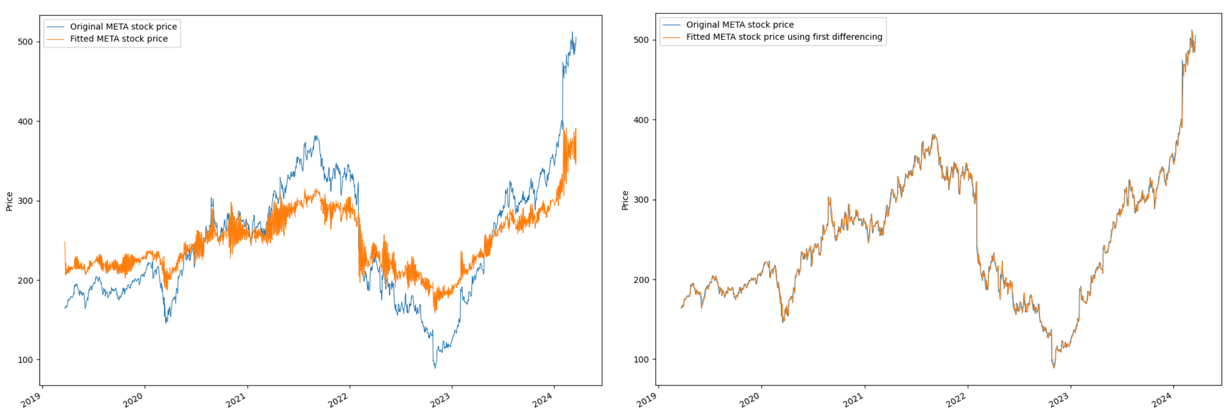


Figure: 7

Directions

<u>Direction for modeling randomness:</u>

The direction in which we should move forward to properly model randomness includes:

Direction 1: Test the Autocorrelation function (ACF) and partial autocorrelation function (PACF) to check whether there is any significant autocorrelation after a few lags.

Direction 2: Augmented Dickey-Fuller Test (ADF) test to check if stock price follows a random walk that is if the time series data is non-stationary or not.

Direction 3: Stochastic Models such as ARIMA and GARCH to capture and simulate randomness.

Direction 4: Bayesian Methods can be applied to use prior knowledge and update probability distribution i.e. posterior knowledge, which can enhance the model from uncertainty and randomness.

<u>Directions for autocorrelation</u>

The following models can address the autocorrelation:

Direction 1: ARIMA Model (Autoregressive Integrated Moving Average): ARIMA Models can capture the autocorrelation in time series data. By appropriately specifying the order of the AR, differencing, and MA terms, ARIMA models can effectively address autocorrelation (Cole and Richards).

Direction 2: GARCH Model (Generalized Autoregressive Conditional Heteroskedasticity): GARCH models are used to model the conditional variance of a time series. GARCH can model autocorrelation in the error terms alongside the heteroscedasticity.

Direction 3: Applying robust standard errors: It adjusts the standard errors to address autocorrelation in the residuals.

Directions for heteroskedasticity

The direction in which we should move forward to properly model heteroskedasticity includes:

Direction 1: Visual Inspection: Begin by visually inspecting the time series plot of the residuals obtained from the model. Look for patterns or trends in the variance of the residuals over time. Heteroskedasticity often manifests as increasing or decreasing variance over time. Use residual plot and scatter Plot for visual inspection.

Direction 2: Perform the Breusch-Pagan test and white test diagnostic tests for heteroskedasticity.

Direction 3: Remove Trend from Non-Stationary Time Series by using trend variable removal, differencing method, white noise removal, etc.

Direction 4: Use MA, ARIMA, and GARCH methods for statistical modeling time-series data.
As we can see above by using MA(1) an efficient model has been generated.

Non-stationarity

Definition

To understand Non-stationary let's first define stationary, A time series is considered as stationary if has no trends and it does not exhibit constant variance over time also it has a constant autocorrelation structure over time.

Now let's understand non-stationary A time series Xt is to be non-stationary if its statistical properties such as mean μt , variance σ2t , and autocorrelation structure over time, which violates the assumption of stationarity (Iordanova).

A time series is non-stationary if:

- μt is a function of t, i.e μt is not equal to μ for all t.
- σ2t function of t, i.e σ2t is not equal to σ2 for all t.
- non-constant Autocorrelation structure for all t.

Description

Non-stationary in time series leads to challenges in modeling, forecasting, and statistical analysis, because in non-stationary data changes over time. non-stationary requires special handling and techniques to make the data stationary for accurate modeling and analysis.

Demonstration & Diagram

Before special handling of non-stationary, we need to identify if the data is non-stationary or not.

one of the way to find if the data is non-stationary or not is to perform an augmented dickey-fuller test, which uses null and alternative hypotheses, let's first understand that:

- H0: The time series is non-stationary means it has some time-dependent structure and is not constant over time.
- HA: The time series is stationary.

In the test, if the P-value is less than the threshold i.e. **(α=.05)**, we can then reject the null hypothesis and conclude that the time series is stationary (Zach).

Now we will test the stock price and stock return of META to test if these time series data are non-stationary or not :

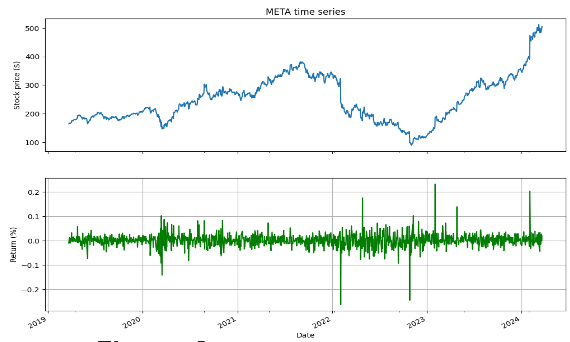


Figure: 8

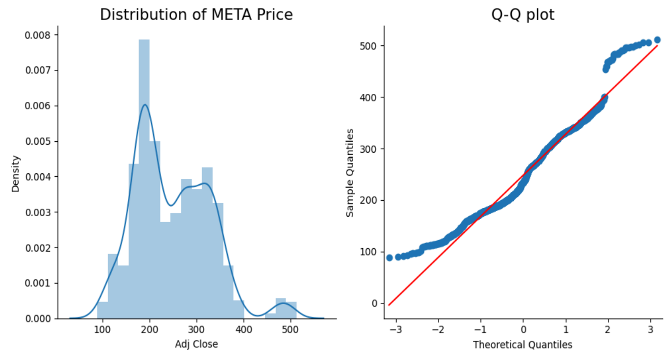


Figure: 9

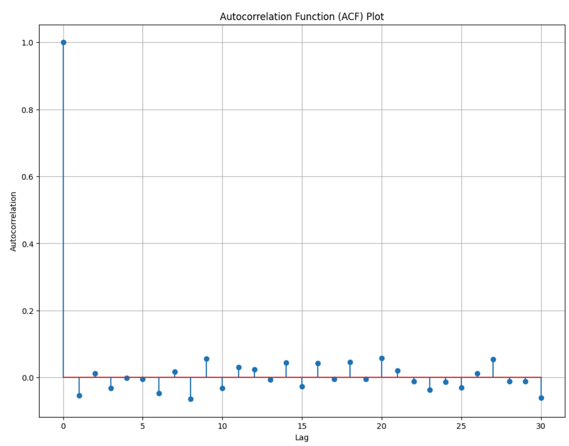


Figure: 10

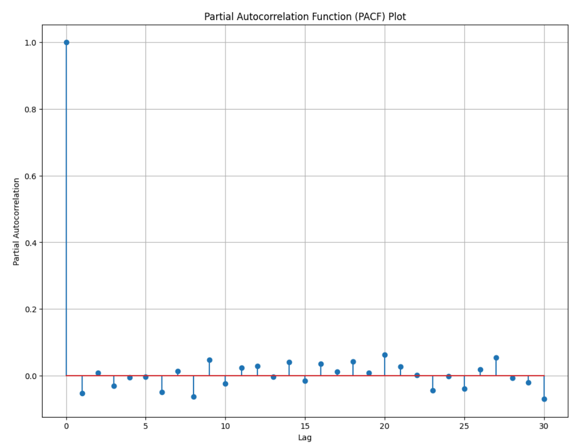


Figure: 11

```
X = check_for_stationary(main_df['Adj Close']);

 p-value= 0.976808755655692 The series Adj Close is likely non-stationary.

X = check_for_stationary(main_df['META']);
 p-value= 0.0 The series META is likely stationary.
```

Figure: 12

Diagnosis

Let's analyze the results:

In the above image we can see time series plot and the first plot is META price plot and the second plot is META return plot in **Figure 8**, and through visual inspection we can see that the META price plot does not have a constant mean and variance over time which makes the price plot non-stationary on the other have the stock return have constant mean and variance over time, appearing as a horizontal line or band around a constant value.

In modeling randomness, the histogram plot by (Team member 1: Modeling randomness GWP2) (Team Member 1) in **Figure 9** shows that the stock prices are random and do not exhibit constant mean and variance over time while the returns have constant mean and variance over time.

one more inspection that we have done is ACF plots by (Team member 2: Autocorrelation GWP2) (Team member 2) in which we can see the stock price ACF & PACF plot the autocorrelation values is not rapidly decaying to zero as we can see in **Figure 10** and **Figure 11** which indicated significant correlation which means the data is non-stationary while on the other hand, we can see for stock return ACF and PACF plot the autocorrelation values are decaying rapidly to zero making it stationary.

But these are just visual analyses we need something more tangible than statistical test because sometimes we can't comprehend everything through visual inspection. So we have some ADF(Augmented Dickey fuller) test in **Figure 12** and the p-value that we got for stock price is p-value= 0.976808755655692 which doesn't reject the null hypothesis which makes it non-stationary on the other hand the p-value for stock return is p-value = 0.0 which makes the returns stationary.

Note: The stationarity of a stock return depends on the specific stock which is being analyzed and the time period that is being considered, generally stock returns exhibits stationary behavior but it may exhibit non-stationary behavior over longer time horizons.

Damage

There are several damaging effects on the analysis and interpretation of non-stationary data:

- **Inaccurate model predictions**: Building accurate models using non-stationary data is very challenging as the data changes over time and the models assume stationarity in the data while modeling which results in failure to observe the underlying patterns and trends.
- **Biased Estimates**: Estimating values from non-stationary data may not provide a true relationship between variables which can lead to biased results.
- **Unreliable forecasts:** The changing nature of non-stationary data when forecasted can be unreliable and misleading as it is very difficult to predict the future values of changing data..
- **False correlations**: Non-stationary data can create false correlations between variables that are not really related which can mislead and lead to false conclusions.

To counter these issues we are going to discuss some directions:

Directions

Direction 1: Differencing: we can apply differencing and make the data stationary by taking first difference or seasonal difference (Iordanova).

Direction 2: Trend modeling: trend modeling methods can be used to capture and model the changing mean or trend in the data.

Direction 3: Detrending: As done by (Team Member 3: Heteroskedasticity GWP2) (Team member 3), it removes the trend component from the data and makes the data stationary (Iordanova).

Direction 4: Seasonal adjustment: some data have seasonal components so we can adjust the data according to that using seasonal decomposition methods.

Relation Between Challenges

For this challenge we are going to understand the relation between Modeling Randomness from GWP2 and Sensitivity to outliers from GWP1 we are going to discuss the damage and direction of these challenges and understand their relations.

When we talk about Time series analysis two major related challenges are modeling randomness and sensitivity to outliers which have a impact on the accuracy and reliability of the analysis.

Model performance: If we don't properly handle the randomness or outliers it can lead to biased model estimates and reduce our prediction accuracy which will affect our overall performance of our models.

Risk assessment: Again modeling of randomness or outliers wrong can result in overestimating and underestimating the risk which will lead to suboptimal risk management strategies.

Decision making: At last unreliable predictions due to inadequate handling of randomness or outliers can results will become an obstacle in optimal decision making which will affet the business, investment as well as resource allocation.

The challenges of modeling randomness and sensitivity are closely related, they both impact accuracy, robustness, and integrity of time series analysis. one more interesting by which they are closely related is that sensitivity to outliers suggests that there can be a potential of randomness in data, Due to this issue addressing these challenges collectively is much required to enhance the quality of our model and ensure reliable working processes.

Some Directions we can look into to tackle these issues are:

Statistical tests: We can also use some tests like ACF and PACF to check significant autocorrelation which can lead us to understand the randomness which can also explain use the outlier sensitivity, also we can use ADF test in initial phase to check if our data is non-stationary or not.

Data Preprocessing: we can use data preprocessing techniques like data cleaning using truncation, or outlier removal based on robust statistical measures.

Robust models: we can use robust statistical models such as robust regression or robust estimation methods that are less sensitive to outliers.

Transformation techniques: we can apply different transformation techniques like log transformation or Box-Cox transformations to reduce the impact of outliers on the data and improve model performance. Not only they will take care of outliers they will also smoothen the randomness.

Stochastic Models: we can use models like ARIMA and GARCH to understand the randomness accurately.

Bayesian Methods: we can use Bayesian inference techniques to understand prior knowledge and update the probability distribution, enhancing the modeling of uncertainty and randomness.

References

1. CFI. "Heteroskedasticity - Overview, Causes and Real-World Example." *Corporate Finance Institute*, http://corporatefinanceinstitute.com/resources/data-science/heteroskedasticity/. Accessed 25 March 2024.
2. Cole, Martin, and Marcus Richards. "Understanding and Addressing Autocorrelation in Regression – Data Magazine." *Data Magazine*, 6 August 2023, https://datamagazine.co.uk/understanding-and-addressing-autocorrelation-in-regression/. Accessed 26 March 2024.
3. Iordanova, Tzveta. "Introduction to Non-Stationary Processes." *Investopedia*, https://www.investopedia.com/articles/trading/07/stationary.asp. Accessed 25 March 2024.
4. kmenta, J. ""Heteroskedasticity." Palgrave Macmillan UK EBooks." 2018, pp. 5796-5797, https://doi.org/10.1057/978-1-349-95189-5_1162.
5. Rehal, Viren. "Autocorrelation: Causes and Consequences." *SPUR ECONOMICS*, 23 February 2024, https://spureconomics.com/autocorrelation-causes-and-consequences/. Accessed 26 March 2024.
6. Smith, Tim. "Random Walk Theory: Definition, How It's Used, and Example." *Investopedia*, 23 Feb 2023, https://www.investopedia.com/terms/r/randomwalktheory.asp. Accessed 25 March 2024.
7. Smith, Tim, and Cierra Murry. "Autocorrelation: What It Is, How It Works, Tests." *Investopedia*, https://www.investopedia.com/terms/a/autocorrelation.asp. Accessed 26 March 2024.
8. Taleb, Nassim Nicholas. "Fooled by Randomness Quotes by Nassim Nicholas Taleb." *Goodreads*, https://www.goodreads.com/work/quotes/3119175-fooled-by-randomness-the-hidden-role-of-chance-in-life-and-in-the-marke. Accessed 25 March 2024.
9. Team Member 1. "GWP2: Modeling Randomness." *Modeling randomness*.
10. Team member 2. "GWP1: Skewness." Skewness. *Skewness*.
11. Team member 2. "GWP2: Autocorrelation." *Autocorrelation*.

12. Team member 3. “GWP2: Heteroskedasticity.” Heteroskedasticity. *Heteroskedasticity*.
13. Zach. “Augmented Dickey-Fuller Test in Python (With Example).” *Statology*, 25 May 2021, https://www.statology.org/dickey-fuller-test-python/. Accessed 25 March 2024.