Descriptive and Predictive Analysis on Shiba-Inu

Ritwik Bahukhandi (Jamia Hamdard University)

*Abstract*— Cryptocurrency and Blockchain is one of the most trending technologies out there , Just like Stock Market which is actually Centralised , these Cryptocurrency market (Decentralised) has also gained popularity and plays an important role in Economy. This project is proposed to give insights and visualizations on how Shiba-Inu has performed in past and how will it going to perform in future by analysing it’s time series components.This papers aims to provide Historical(Descriptive) and Future(Predictve) insights using Statisical Analysis

Keywords— Time Series , Shiba Inu, Seasonality, Trend, Closing Price , Moving Averages , Statistical Tests , ARIMA

# **Introduction**

Blockchain is a revolutionary technology that allows people to record transactions on a digital, decentralised, distributed ledger, without any central authority. Some consider this technology as “the trust machine” and cryptocurrency is without doubt the most notable by-product of the blockchain revolution. Initial Coin Offering (ICO) is a new way to raise entrepreneurial finance, newly created cryptocurrencies are being sold to the public by startups in exchange of capital. This book chapter intends to clarify this phenomenon by explaining the concepts of blockchain technology, cryptocurrency and ICO, in order to provide valuable insights into this new trend of entrepreneurial finance.

A cryptocurrency is a digital or virtual currency that is secured by cryptography, which makes it nearly impossible to counterfeit or double-spend. Many cryptocurrencies are decentralized networks based on blockchain technology—a distributed ledger enforced by a disparate network of computers. A defining feature of cryptocurrencies is that they are generally not issued by any central authority, rendering them theoretically immune to government interference or manipulation.

The Shiba Inu coin launched on August 1, 2020 and at the time was priced at less than $0.00000001. With 1 quadrillion tokens minted it remained at that level for some time, but more recently thanks to the increasing interest in Dogecoin, Shiba Inu has also come to life. Created by an anonymous developer going by the name Ryoshi, the Woofpaper for Shiba Inu says the project was begun from the simple question of:

“What would happen is a cryptocurrency project was 100% run by its community?”



*Fig 1 :Shiba-inu Logo*

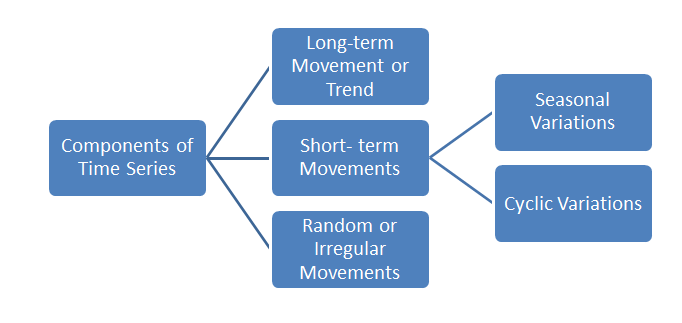
A time series analysis is a method of analysing a series of data points gathered over a period of time. Analysts use time series analysis to record data points at consistent intervals over a set period of time rather than intermittently or randomly. However, this type of analysis is more than just collecting data over time.

Time series analysis helps organizations understand the underlying causes of trends or systemic patterns over time. Using data visualizations, business users can see seasonal trends and dig deeper into why these trends occur. With modern analytics platforms, these visualizations can go far beyond line graphs.

**Models of time series analysis include:**

* **Classification:** Categorize the Data
* **Descriptive analysis:** Time series analysis identifies patterns such as patterns, periods, and seasonal fluctuations. Patterns in time series data, such as trends, cycles, or seasonal variation, are discovered.
* **Exploratory analysis:**Showcase the main characteristics of the time series data, usually in a visual format in the form of graph
* **Forecasting:** It foresees data in the future. This classification is based on historical patterns. It forecasts events that could occur along upcoming plot points utilizing past data as a model.\

**NOTE : *Time Series Analysis of Shiba Inu can be done through various Mathematical and Statistical Analysis but there are many more things which can effect the prediction so we can’t directly say its 100% accurate to perform time series analysis on cryptocurrency***



*Fig 2 :Component of Time Series*

# **PRIOR WORK**

## Univariate Time Series Forecasting of Temperature and Precipitation with a Focus on Machine Learning Algorithms

[1]We provide contingent empirical evidence on the solutions to three problems associated with univariate time series forecasting using machine learning (ML) algorithms by conducting an extensive multiple-case study. These problems are: (a) lagged variable selection, (b) hyperparameter handling, and (c) comparison between ML and classical algorithms. The multiple-case study is composed by 50 single-case studies, which use time series of mean monthly temperature and total monthly precipitation observed in Greece. We focus on two ML algorithms, i.e. neural networks and support vector machines, while we also include four classical algorithms and a naïve benchmark in the comparisons. We apply a fixed methodology to each individual case and, subsequently, we perform a cross-case synthesis to facilitate the detection of systematic patterns. We fit the models to the deseasonalized time series. We compare the one- and multi-step ahead forecasting performance of the algorithms. Regarding the one-step ahead forecasting performance, the assessment is based on the absolute error of the forecast of the last monthly observation. For the quantification of the multi-step ahead forecasting performance we compute five metrics on the test set (last year’s monthly observations), i.e. the root mean square error, the Nash-Sutcliffe efficiency, the ratio of standard deviations, the coefficient of correlation and the index of agreement. The evidence derived by the experiments can be summarized as follows: (a) the results mostly favour using less recent lagged variables, (b) hyperparameter optimization does not necessarily lead to better forecasts, (c) the ML and classical algorithms seem to be equally competitive.

## ARIMA and Exponential Smoothening

The stock market is very unpredictable, any geopolitical change can impact the share trend of stocks in the share market, recently we have seen how covid-19 has impacted the stock prices, which is why on financial data doing a  reliable trend analysis is very difficult. The most efficient way to solve this kind of issue is with the help of Machine learning and Deep learning. For time series forecasting, the ARIMA model is a well-known and frequently used statistical technique. This is one of the most commonly used models for predicting linear time series data. Since it was discovered to be dependable, efficient, and capable of predicting short-term share market changes, this model has been frequently used in banking and economics. Exponential smoothing and ARIMA models are the two most commonly used approaches to time series forecasting, and they are complementary approaches to the problem. ARIMA models attempt to describe the autocorrelation, whereas exponential smoothing techniques attempt to reflect the trend and seasonality in the data (autocorrelation is the degree of resemblance between a given time series and a lagged version of itself over subsequent time periods).

## AD-Fuller Test

The author proposed the idea of unit root or AD-Fuller test A unit root test tests whether a time series is not stationary and consists of a unit root in time series analysis. The presence of a unit root in time series defines the null hypothesis, and the alternative hypothesis defines time series as stationary.

## Exponential Smoothening (Brown,Holt,Winters)

[2]This paper is a critical review of exponential smoothing since the original work by Brown and Holt in the 1950s. Exponential smoothing is based on a pragmatic approach to forecasting which is shared in this review. The aim is to develop state-of-the-art guidelines for application of the exponential smoothing methodology. The first part of the paper discusses the class of relatively simple models which rely on the Holt-Winters procedure for seasonal adjustment of the data. Next, we review general exponential smoothing (GES), which uses Fourier functions of time to model seasonality. The research is reviewed according to the following questions. What are the useful properties of these models? What parameters should be used? How should the models be initialized? After the review of model-building, we turn to problems in the maintenance of forecasting systems based on exponential smoothing. Topics in the maintenance area include the use of quality control models to detect bias in the forecast errors, adaptive parameters to improve the response to structural changes in the time series, and two-stage forecasting, whereby we use a model of the errors or some other model of the data to improve our initial forecasts. Some of the major conclusions: the parameter ranges and starting values typically used in practice are arbitrary and may detract from accuracy. The empirical evidence favours Holt's model for trends over that of Brown. A linear trend should be damped at long horizons. The empirical evidence favours the Holt-Winters approach to seasonal data over GES. It is difficult to justify GES in standard form–the equivalent ARIMA model is simpler and more efficient. The cumulative sum of the errors appears to be the most practical forecast monitoring device. There is no evidence that adaptive parameters improve forecast accuracy. In fact, the reverse may be true.

## Forecasting Bitcoin Prices Using N-BEATS Deep Learning Architecture (Alikhan Bulatov)

[3]In the world of finance, the usage of computationally intensive systems that apply machine learning techniques is becoming more frequent. Each year, new state-of-the-art deep learning architectures for time series forecasting are created, making them more accurate than ever before. This research compares the prediction power of the N-BEATS deep learning architecture trained on Bitcoin daily, hourly, and real-time data to other popular time series forecasting approaches like LSTM and ARIMA. The Mean Average Percentage Error (MAPE) and Root Mean Squared Error (RMSE) are two methods for calculating prediction errors (RMSE). In comparison to LSTM and ARIMA models, the results show that the created N-BEATS model has potential predictive power.

## Covid-19 forecast using Holt-Winters exponential smoothing

[4]Covid-19 has spread throughout the world, including in Indonesia and it is known that the virus is reported to be highly contagious. Indonesia is the fourth-largest population for Covid-19 cases in Asia, while in the world the top three for covid-19 cases are United States, Brazil, and India, so it is inevitable, Indonesia will be greatly felt the spread of the pandemic coronavirus, even estimated to be freed from the coronavirus in a shorter period when compared to other countries with less exposure to COVID-19. In this paper, we study the covid-19 prediction model using Holt-Winters exponential smoothing, for a certain period. This study uses the Covid-19 pandemic data area of Gorontalo, Indonesia, from April 10 to October 13, 2020 (especially total cases). It was found that using Holt-Winters exponential smoothing, the best forecasting model is the one with smoothing parameters α = 0.1 and γ = δ = 0.5 for trend and seasonality respectively, which gives the smallest MAPE value of 6.14.

# PROBLEM STATEMENT

To perform Time Series Analysis on Shiba Inu historical data to gain descriptive and Predictive insights on Closing Price using Statistical and Machine Learning Models

# METHODOLOGIES

For performing Time Series Analysis , One can use many methods to gain insights , but the ultimate goal is to come up with good models to make the insights gain Simple . In this paper I majorly focused of Providing great visuals and providing excellent predictions . So to do that I’ve used many stats-model and analytical tools.

****

*Fig 3:Workflow of Methodologies*

## **Dataset and Features**

The dataset was taken from Yahoo Finance (https://finance.yahoo.com/quote/SHIB-INR/history/) It’s often considered to be the ﬁrst step towards analysis . The Data is decribed in a such way that it has record of more that 600 days of daily shiba inu price since August 2020.

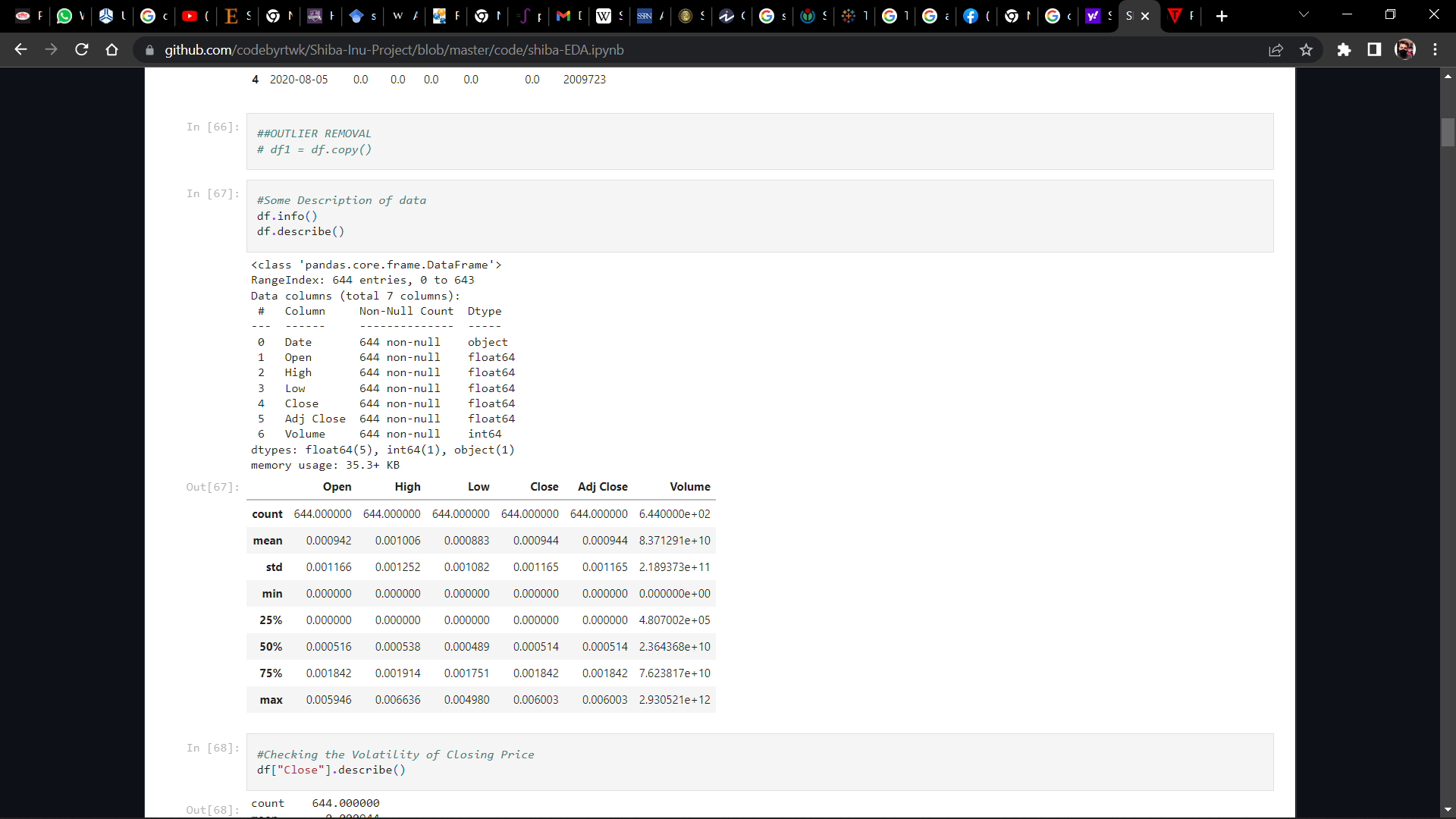
Data ser has contain multiple Attribute which are quite important for our Analysis :

* *Date*
* *Open*
* *High*
* *Low*
* *Close*
* *Adjacent Close*
* *Volume*

## **Describing the DataSet**

* ***.info()*** The info() method prints information about the DataFrame. The information contains the number of columns, column labels, column data types, memory usage, range index, and the number of cells in each column (non-null values).
* .***describe()*** The describe() method is used for calculating some statistical data like percentile, mean and std of the numerical values of the Series or DataFrame. It analyzes both numeric and object series and also the DataFrame column sets of mixed data types.

The beauty of Describe Function is that it provides the basic information such as Mean , Count , Standard Deviation , Quartile etc.

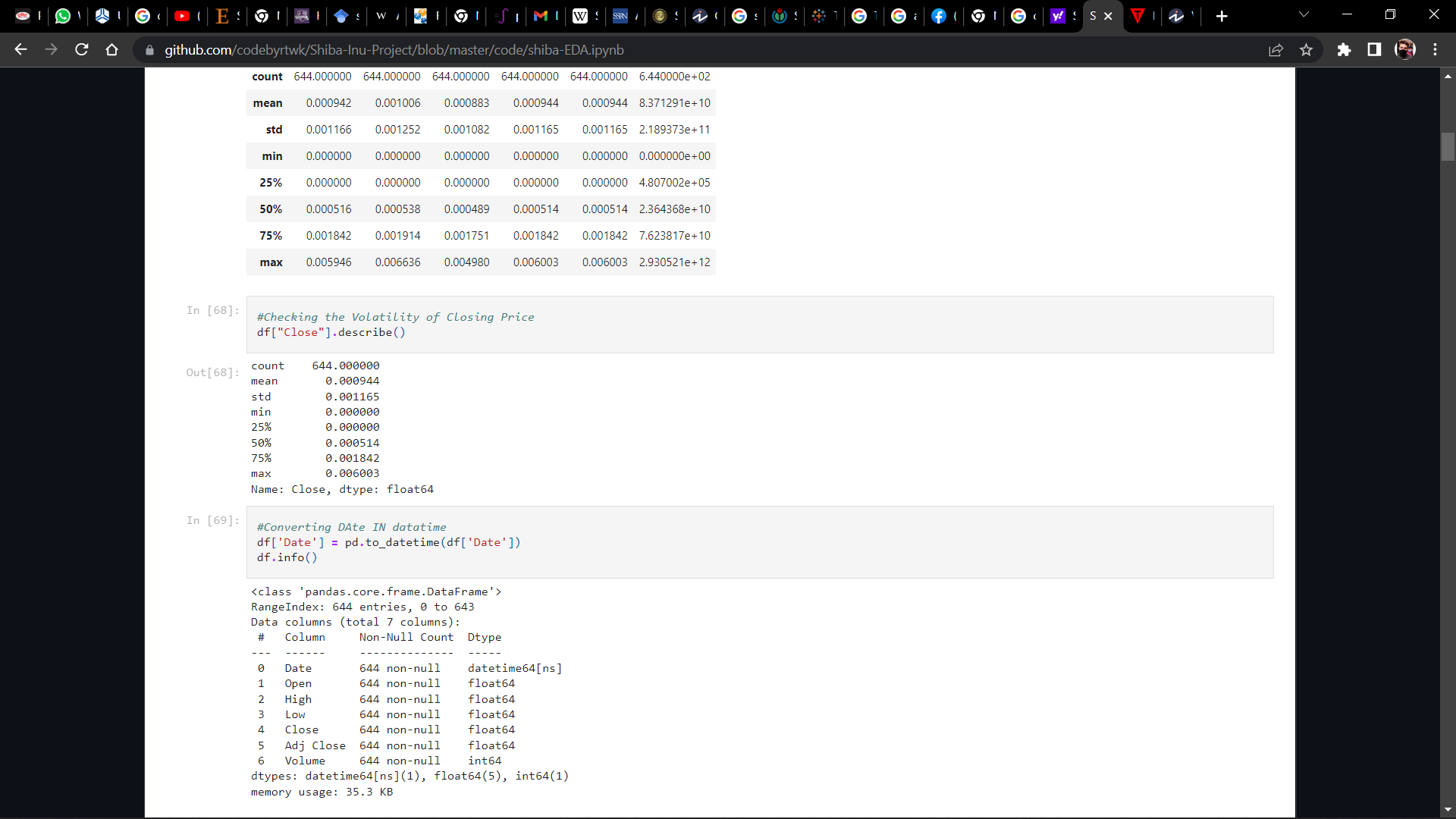


*Fig 4 Describing the Dataset*

* ***Volatility Test***

*Volatility is a statistical measure of the dispersion of returns for a given security or market index. In most cases, the higher the volatility, the riskier the security. Volatility is often measured as either the* ***standard deviation*** *or variance between returns from that same security or market index.*

*In the securities markets, volatility is often associated with big swings in either direction. For example, when the stock market rises and falls more than one percent over a sustained period of time, it is called a "volatile" market. An asset's volatility is a key factor when pricing options contracts.*

****

*Fig 5 : Volatility Check*

*Note :* ***After Analysis it turned out to be Quite a Volatile Asset but not much (Standard Deviation > 0)*. *In comparison with Coins like Bitcoin it’s much less volatile***

## **DESCRIPTIVE VISUALIZATION / ANALYSIS**

It summarizes and organizes all of the collected data into something manageable and simple to understand. The descriptions can include the entire data set or just a part of the data set.

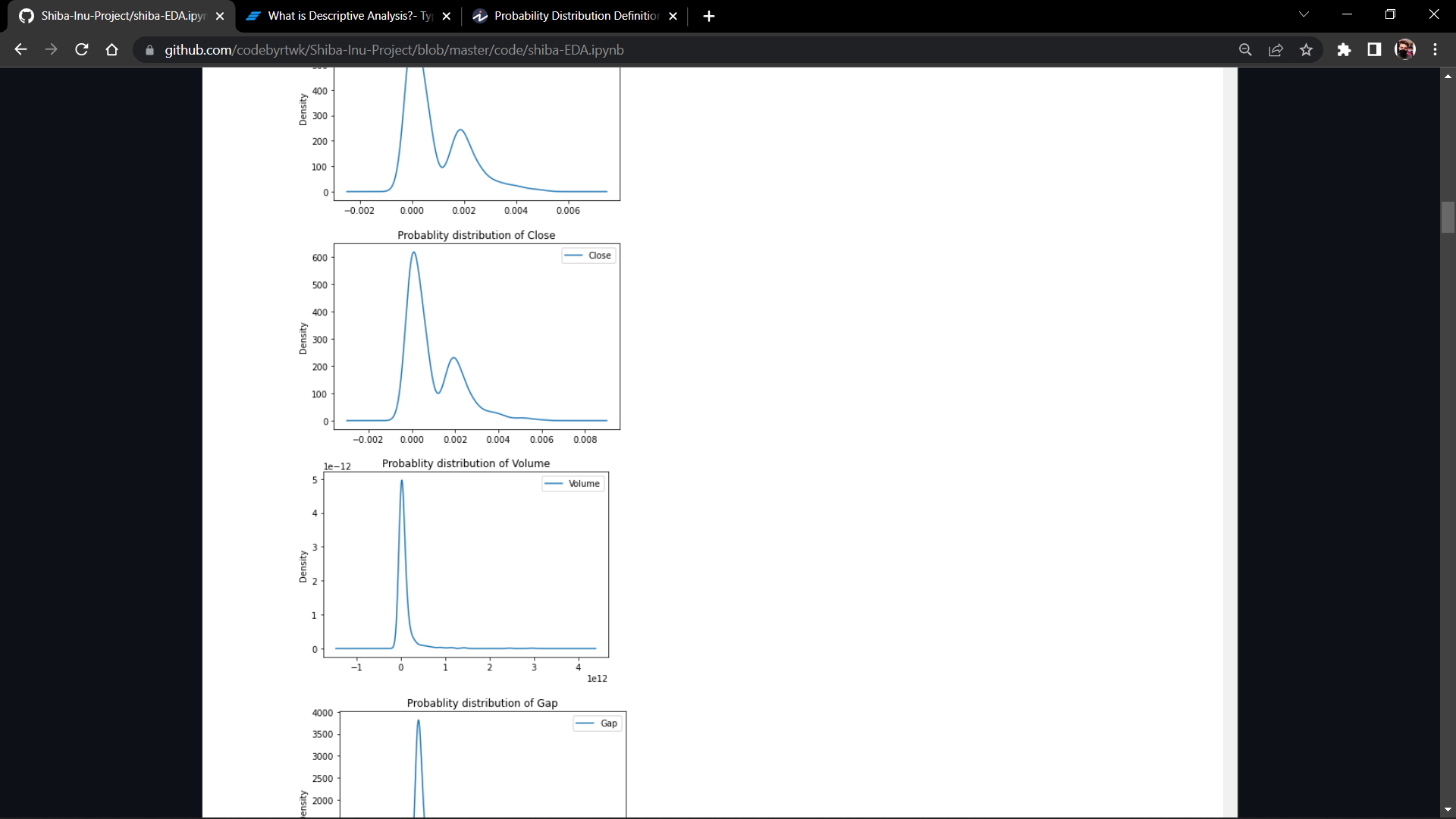
One of the most important things to know about descriptive data analysis is that it focuses on the data instead of on the implication that can be far reaching and go beyond the represented data.

This is the main difference between inferential statistics and descriptive statistics. Inferential statistics uses complicated calculations to make predictions while descriptive statistics does not.

High degree of objectivity and neutrality of the researchers are one of the main advantages of Descriptive Analysis. The reason why researchers need to be extra vigilant is because descriptive analysis shows different characteristics of the data extracted and if the data doesn’t match with the trends then it will lead to major dumping of data ***Descriptive analysis is considered to be more vast than other quantitative methods and provide a broader picture of an event or phenomenon***. It can use any number of variables or even a single number of variables to conduct a descriptive research

* ***Probability Distribution***

A probability distribution is a statistical function that describes all the possible values and likelihoods that a [random variable](https://www.investopedia.com/terms/r/random-variable.asp) can take within a given range. This range will be bounded between the minimum and maximum possible values, but precisely where the possible value is likely to be plotted on the probability distribution depends on a number of factors. These factors include the distribution's mean (average), [standard deviation](https://www.investopedia.com/terms/s/standarddeviation.asp)

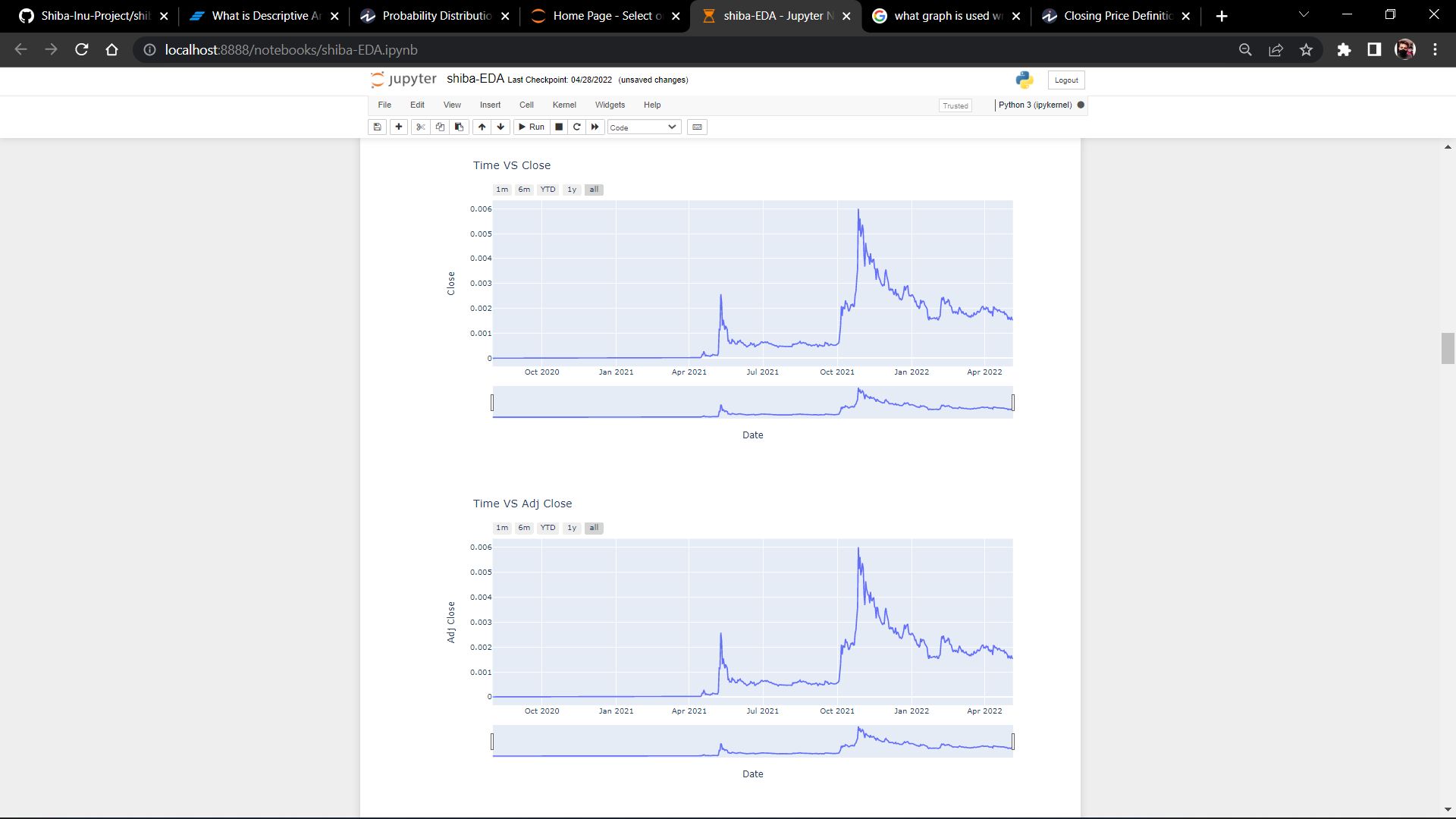


*Fig 6 : Probability Density (Closing Price & Volume)*

*Note :* ***We can see the PDF came out to be left skewed which shows that most of the value revolves nears are Zero(0)***

* ***Line Graph vs Time***

The closing price is the raw price or cash value of the last transacted price in a security before the market officially closes for normal trading. It is often the reference point used by investors to compare a stock's performance since the previous day—and closing prices are frequently used to construct line graphs depicting historical price changes over time. The adjusted closing price factors in anything that might affect the stock price after the market closes, such as dividends or splits. Most stocks and other financial instruments are traded after-hours, although in far smaller volumes. Therefore, the closing price of any security is often different from its after-hours trading price



*Fig 7 : Line Graph (Close and Adj Close)*

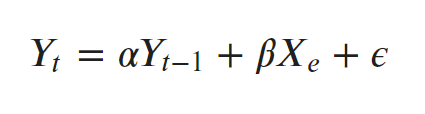
* ***Hypothesis Testing (AD-Fuller Test)***

Our initial Hypothesis was that data is non stationary to check that I’ve use AD-Fuller Test.

[5]The ADF Test (Augmented Dickey Fuller Test) is a typical statistical test used to determine whether or not a time series is stationary. When it comes to analysing the stability of a series, it is one of the most widely employed statistical tests.

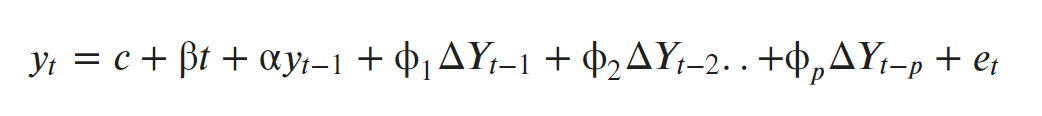
[6]Since testing the stationarity of a time series is a frequently performed activity in autoregressive models, the ADF test along with KPSS test is something that you need to be fluent in when performing time series analysisThe null hypothesis of the Augmented Dickey-Fuller is that there is a unit root, with the alternative that there is no unit root. If the pvalue is above a critical size, then we cannot reject that there is a unit root p-values are obtained through regression surface approximation from MacKinnon 1994, but using the updated 2010 tables. If the p-value is close to significant, then the critical values should be used to judge whether to reject the null.

The ADF test relates to the 'Unit Root Test' group of tests, which is the proper method for evaluating a time series' stationarity. A time series' unit root is a feature that renders it non-stationary. A unit root is claimed exist in a time series of the value of alpha = 1 in the equation below..



where, Yt is the value of the time series at time ‘t’ and Xe is an exogenous variable (a separate explanatory variable, which is also a time series).

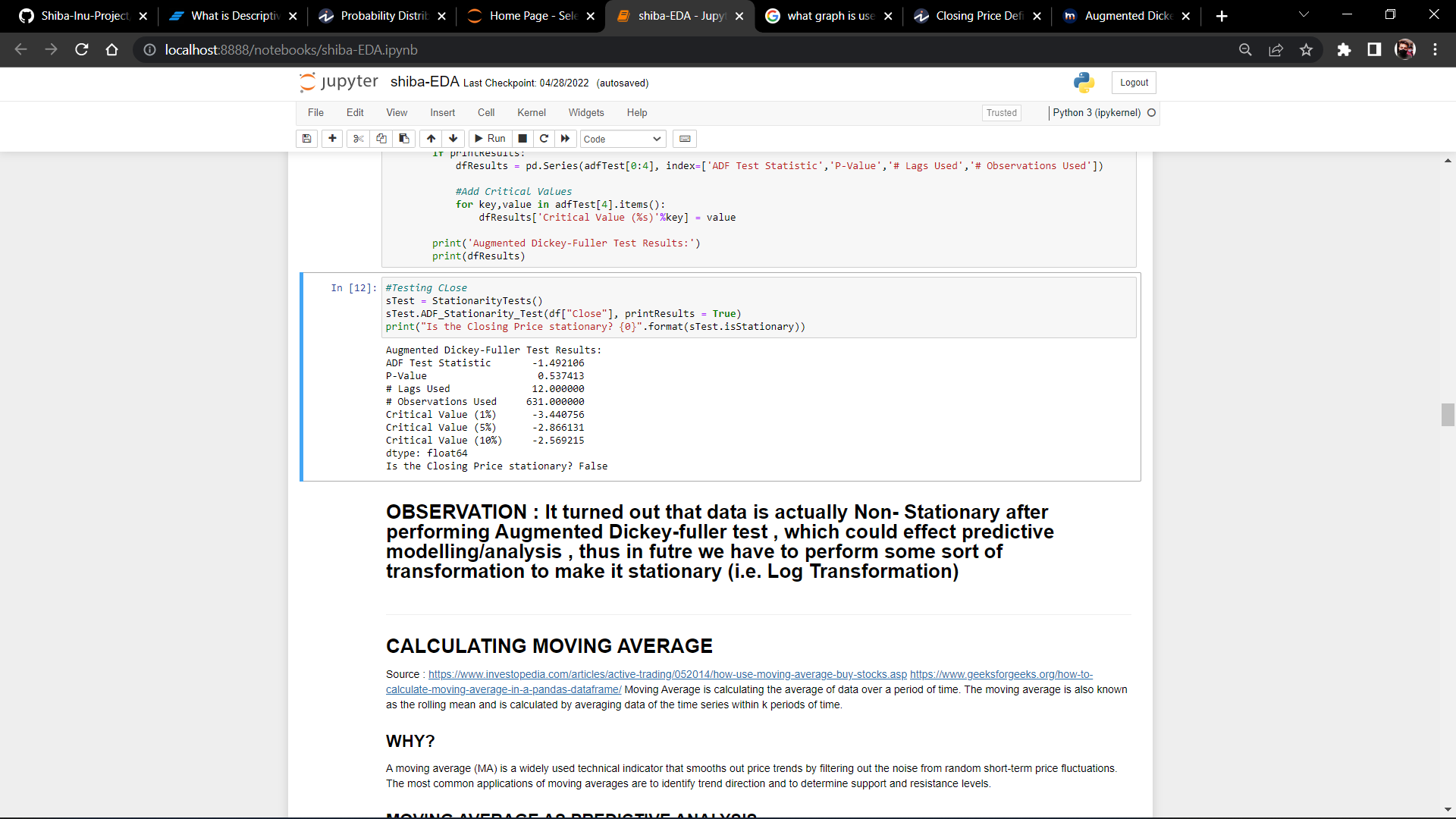
The ADF test expands the Dickey-Fuller test equation to include high order regressive process in the model.



The p-value obtained should be less than the significance level (say 0.05) in order to reject the null hypothesis, because the null hypothesis assumes the presence of unit root, that really is, 1. As an outcome, the series is also said to be stationary.

[7]A **null hypothesis** is a type of statistical hypothesis that proposes that no statistical significance exists in a set of given observations. Hypothesis testing is used to assess the credibility of a hypothesis by using sample data. The null hypothesis, also known as the conjecture, is used in quantitative analysis to test theories about markets, investing strategies, or economies to decide if an idea is true or false.

[8]An important point to note is that we are testing the null hypothesis because there is an element of doubt about its validity. Whatever information that is against the stated null hypothesis is captured in the alternative hypothesis (H1).



*Fig 8 : ADF Test Result*

*Note :* ***It turned out that our Initial Hypothesis is failed cause the Closing Price is a non Staionary , as the p-value > 0.05***

* ***Moving Average (Trend Analysis)***

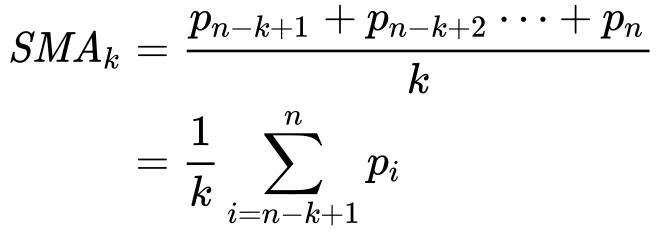
[9]The commonly used time series method is Moving Average. This method is slick with random short-term variations. Relatively associated with the components of time series. **The Moving Average (MA) (Or) Rolling Mean:**In which MA has calculated by taking averaging data of the time-series, within k periods.

The concept of a moving average is straightforward. You generally apply a moving average to a series of observations taken over time. The average moves forward with the data series. So, the first moving average might include Days 1 through 3, the second moving average might include Days 2 through 4, and so on. This chapter discusses some of the reasons why you might want to make those calculations.

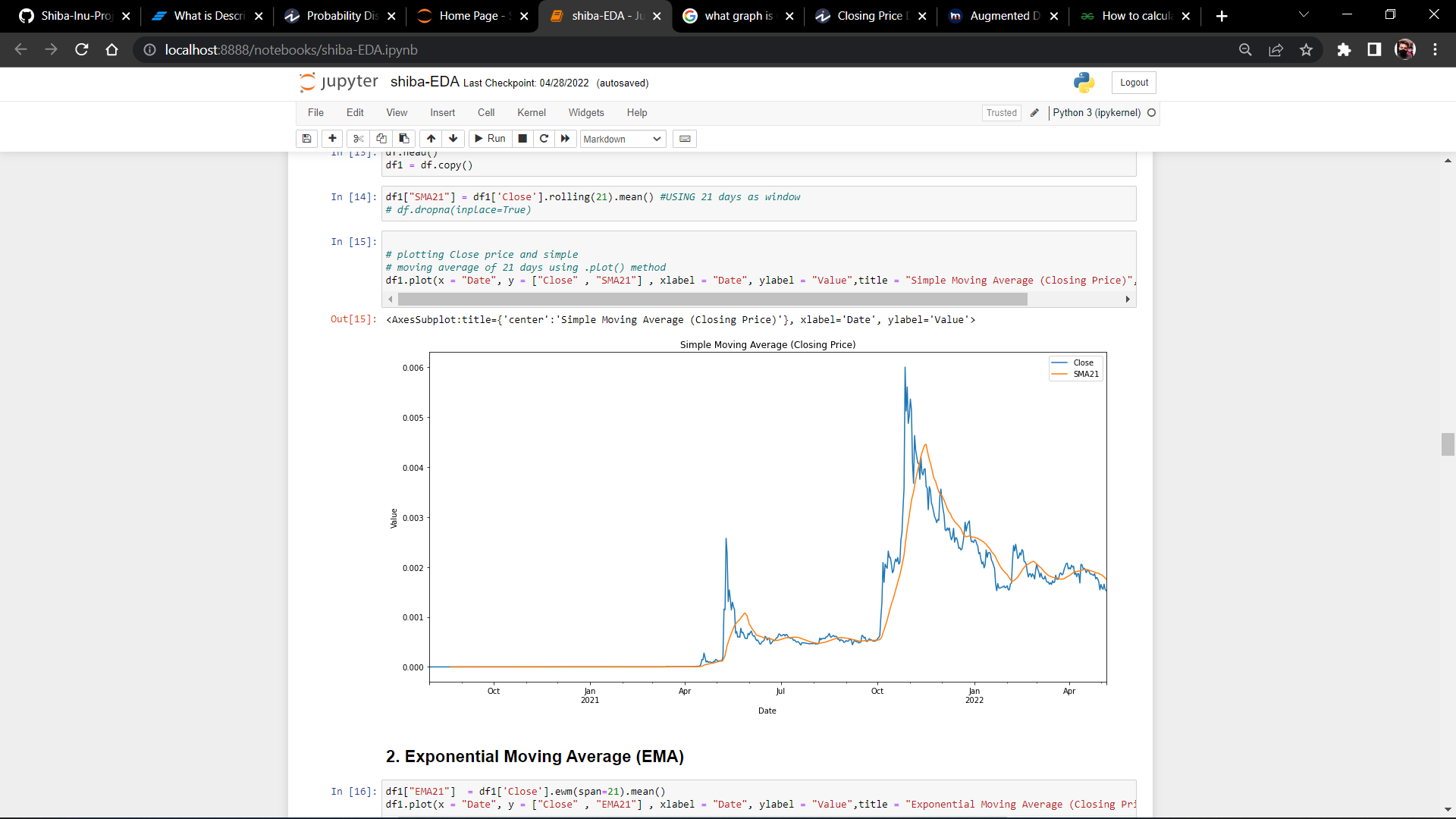
*It is the first step towards the predictive/Inferential Analysis*

#### **Simple Moving Average**

A simple moving average tells us the unweighted mean of the previous K data points. The more the value of K the more smooth is the curve, but increasing K decreases accuracy. If the data points are p1, p2, . . . , pn then we calculate the simple moving average



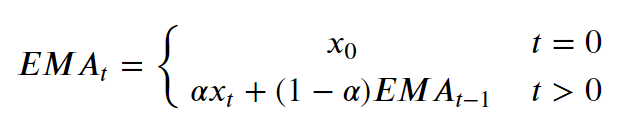
In Python, we can calculate the moving average using .rolling() method. This method provides rolling windows over the data, and we can use the mean function over these windows to calculate moving averages. The size of the window is passed as a parameter in the function .rolling(window)



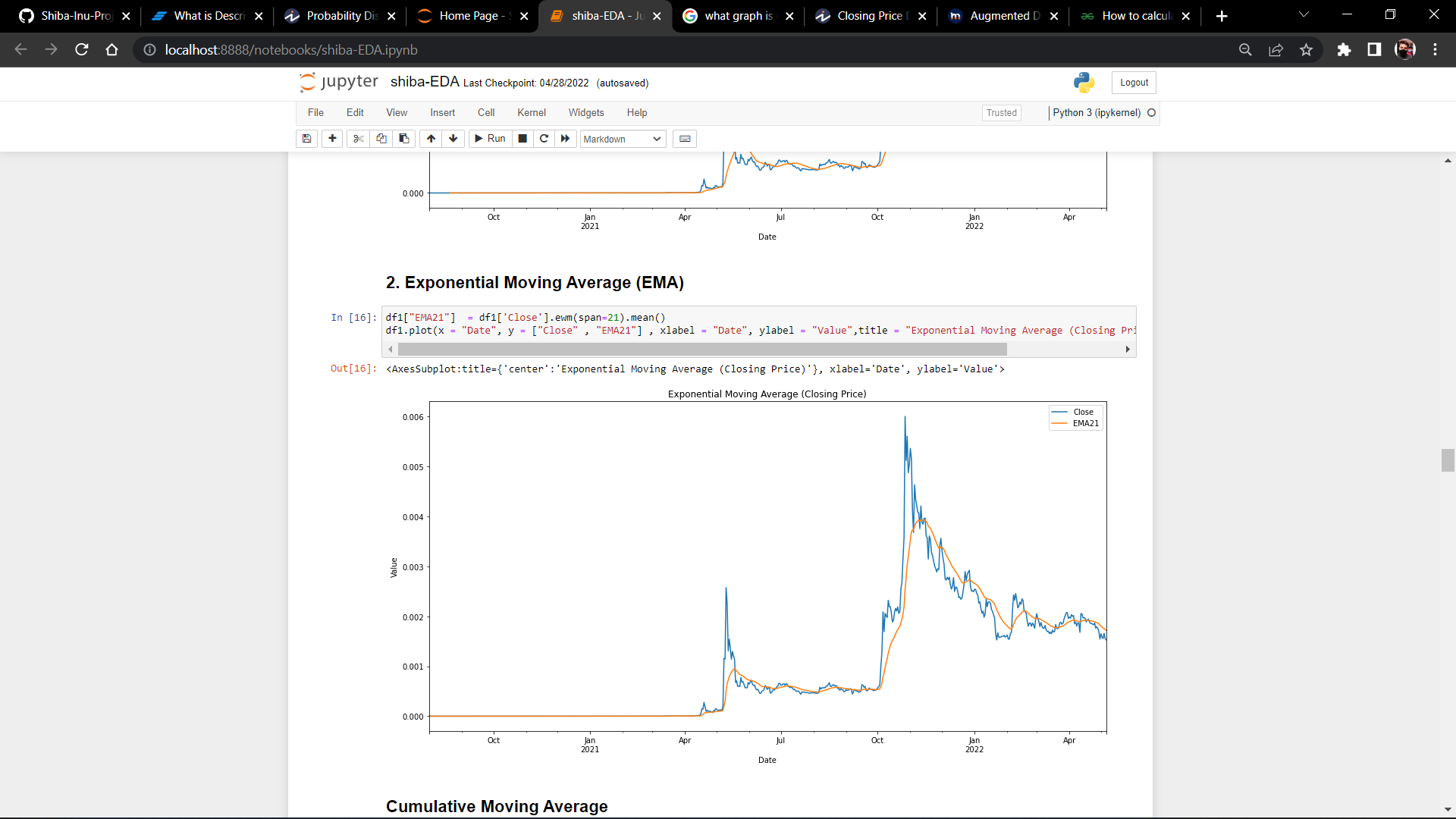
*Fig 9 : Simple Moving Average ( SMA21)*

#### **Exponential Moving Average**

#### Exponential moving average (EMA) tells us the weighted mean of the previous K data points. EMA places a greater weight and significance on the most recent data points. The formula to calculate EMA at the time period t is:



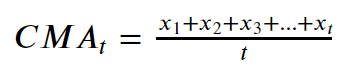
where xt is the value of observation at time t & α is the smoothing factor. In Python, EMA is calculated using .ewm() method. We can pass span or window as a parameter to .ewm(span = ) method.



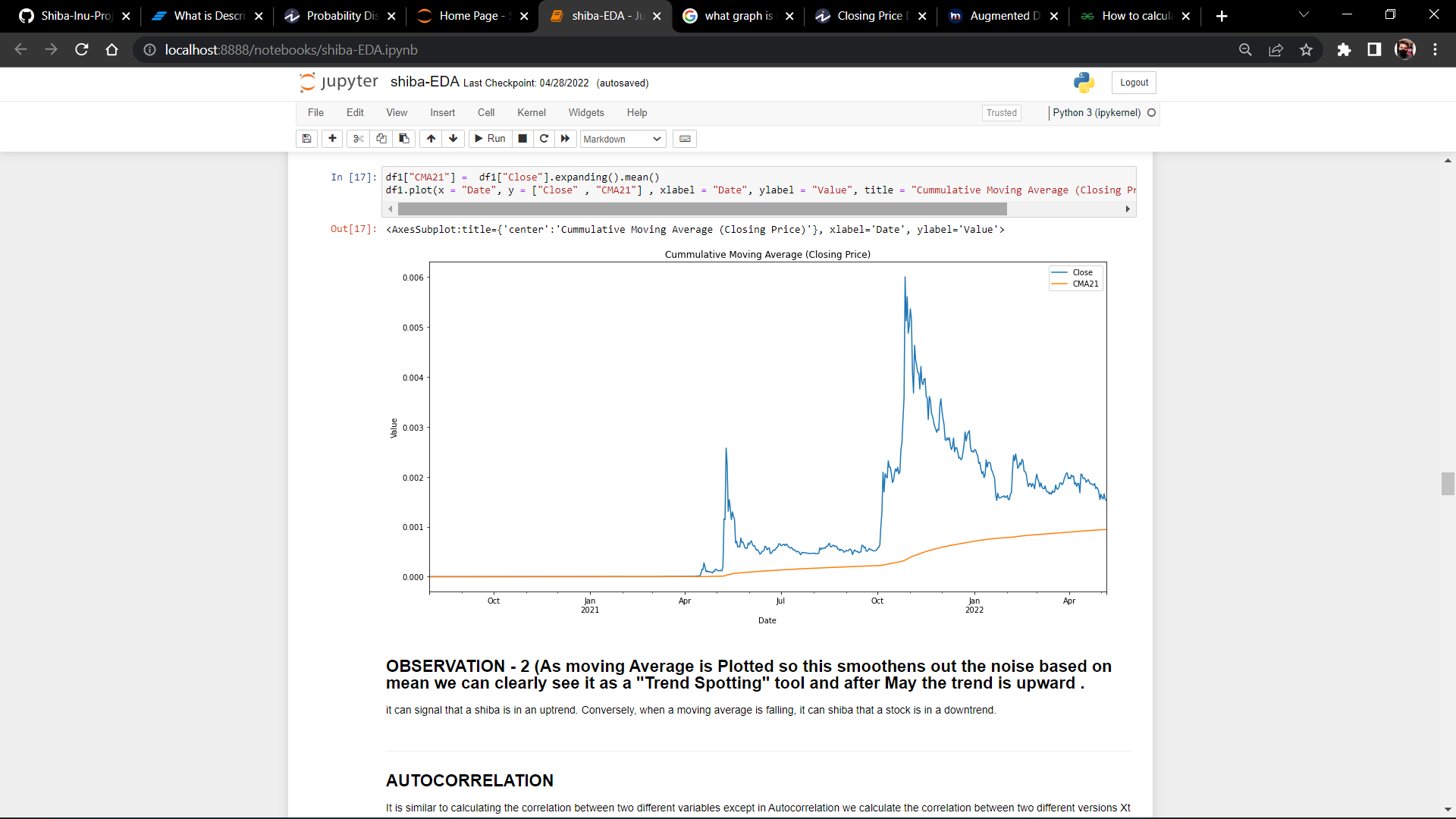
*Fig 10 : Exponential Moving Average(EMA21)*

#### **Cummulative Moving Average**

#### The Cumulative Moving Average is the mean of all the previous values up to the current value. CMA of dataPoints x1, x2 ….. at time t can be calculated as,



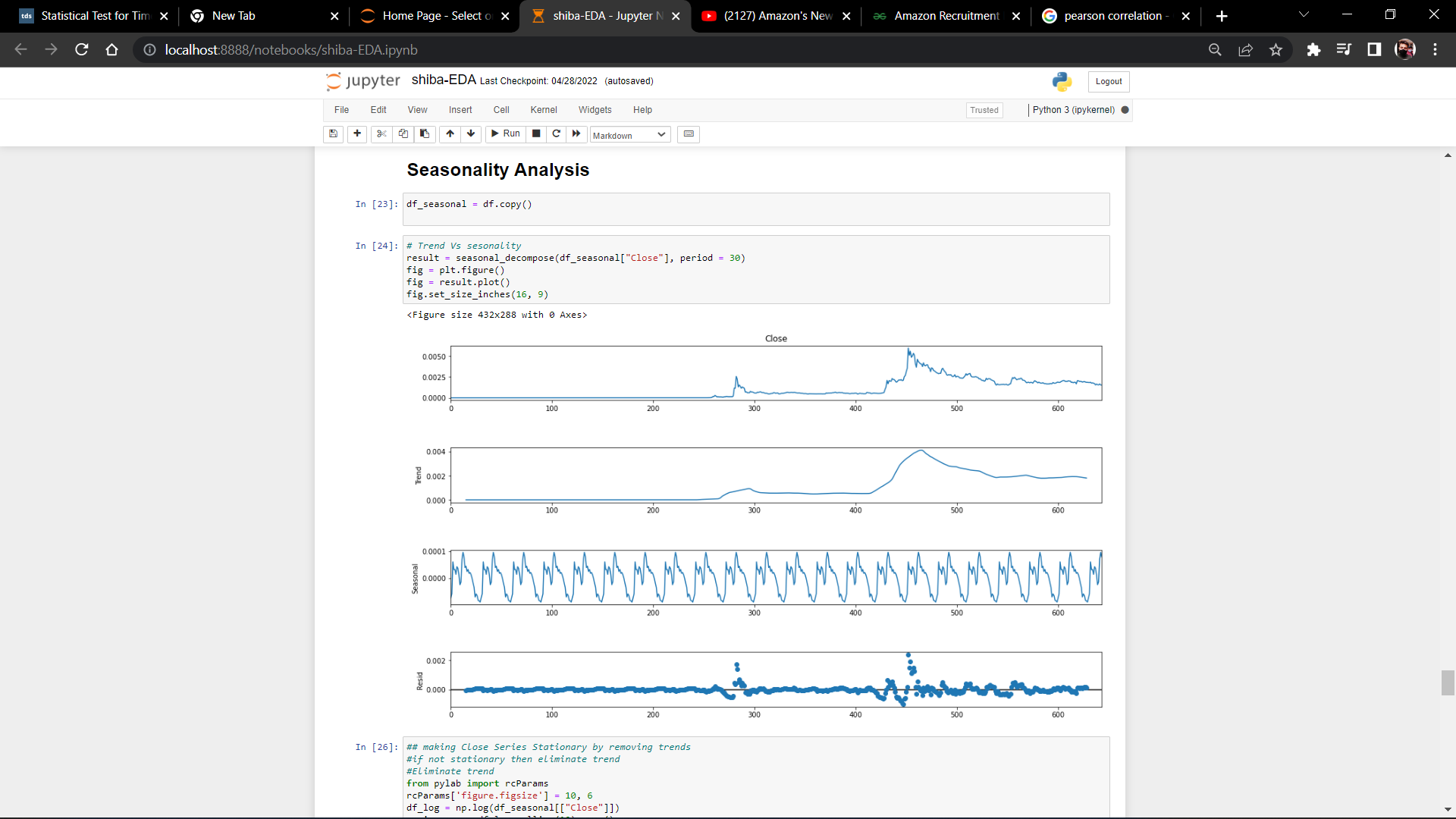
While calculating CMA we don’t have any fixed size of the window. The size of the window keeps on increasing as time passes. In Python, we can calculate CMA using .expanding() method.



*Fig 11 : Cummulative Moving Average*

* ***Seasonality Analysis***

[10]Time series decomposition involves thinking of a series as a combination of level, trend, seasonality, and noise components.Decomposition provides a useful abstract model for thinking about time series generally and for better understanding problems during time series analysis and forecasting.

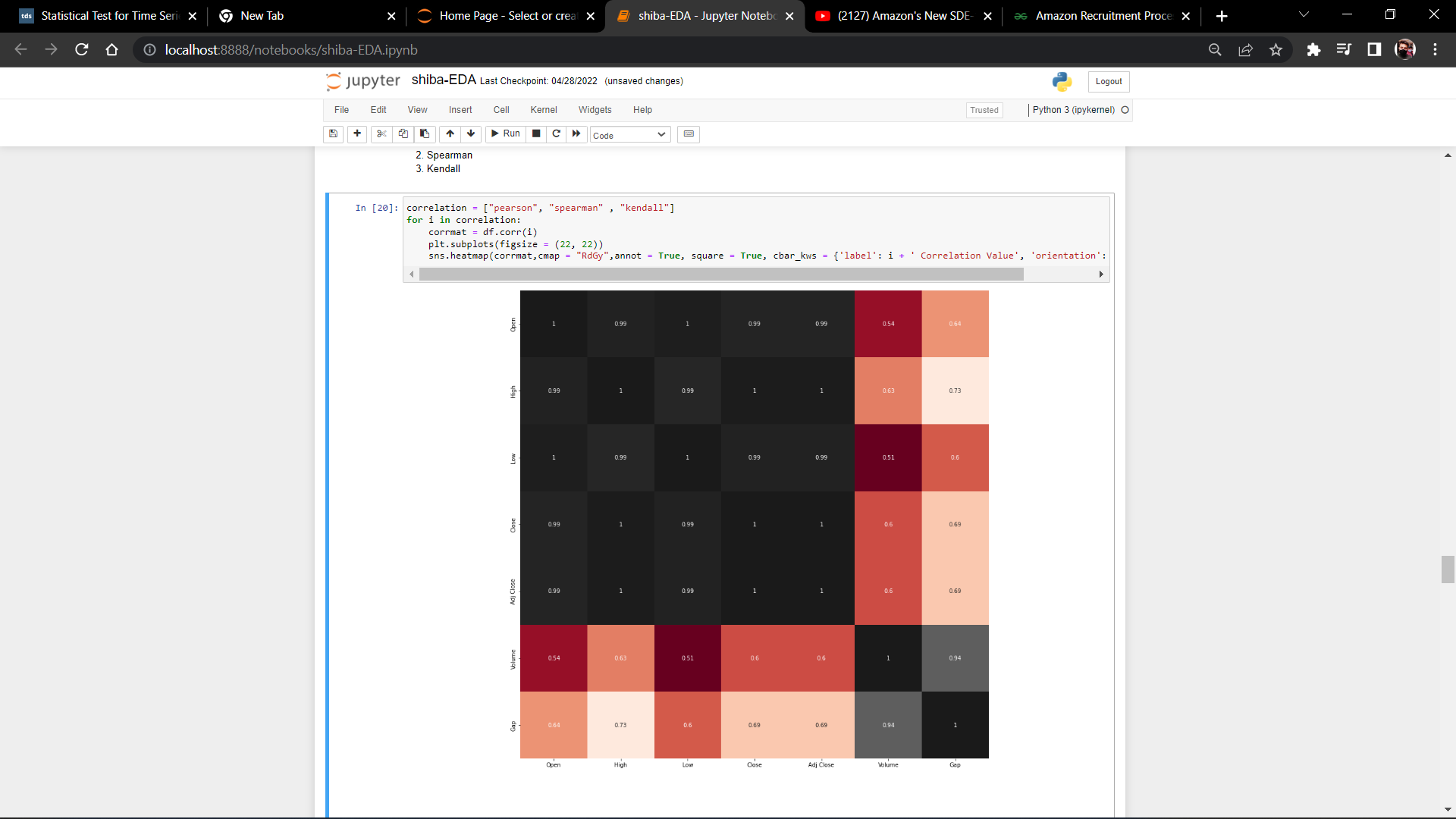


*Fig 12 : Seasonal Decoposition (Closing Price)*

* ***Correlation Analysis***

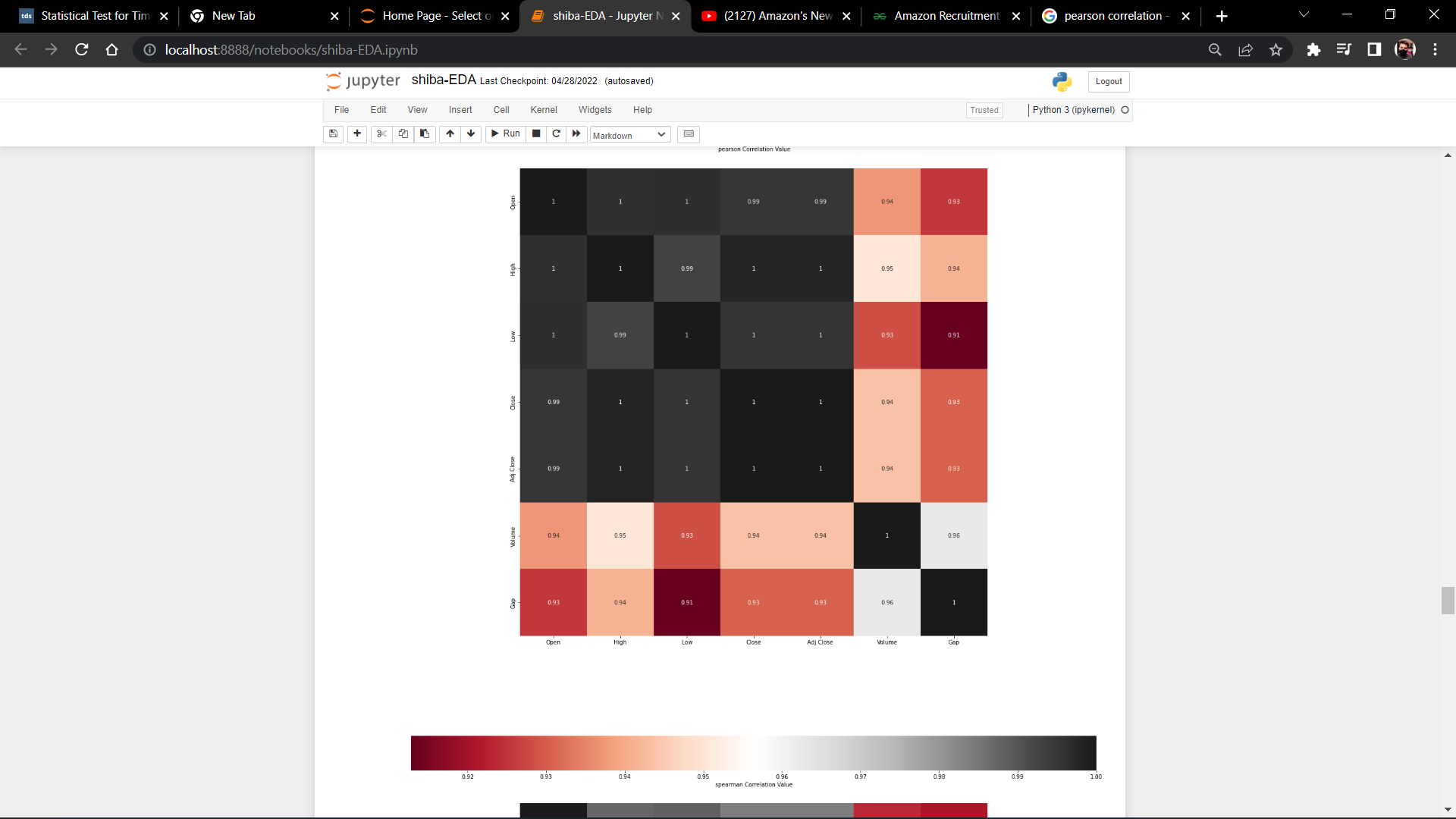
In statistics, correlation or dependence is any statistical relationship, whether causal or not, between two random variables or bivariate data. In the broadest sense correlation is any statistical association, though it commonly refers to the degree to which a pair of variables are linearly related. One of the statistical concepts that is most related to this type of analysis is the correlation coefficient.The correlation coefficient is the unit of measurement used to calculate the intensity in the linear relationship between the variables involved in a correlation analysis, this is easily identifiable since it is represented with the symbol r and is usually a value without units which is located between 1 and -1

***Pearson Correlation***



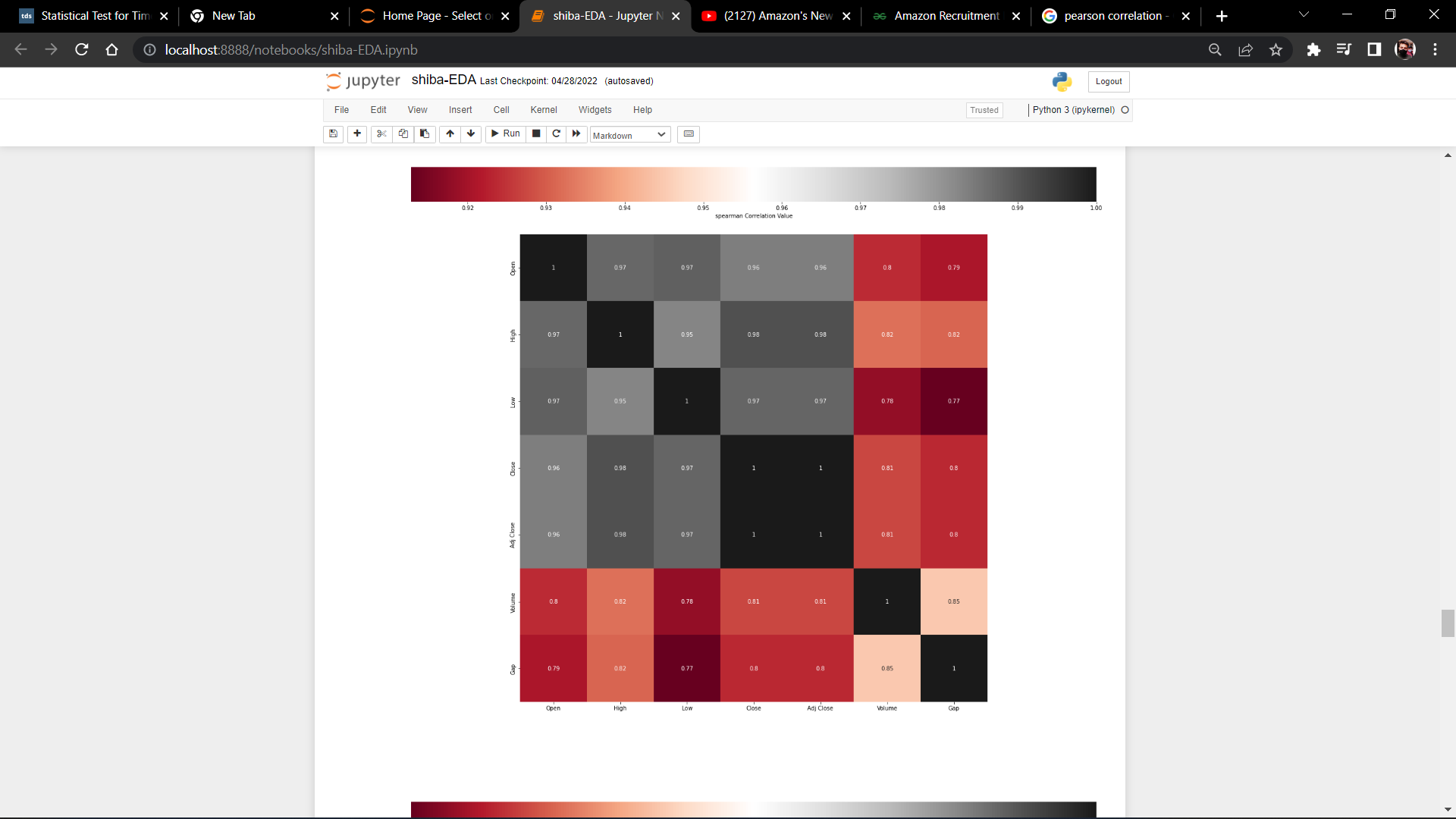
*Fig 13 : Pearson Correlation*

***Spearman Correlation***



*Fig 13 : Spearman Correlation*

***Kendall Correlation***



*Fig 14 : Kendall Correlation*

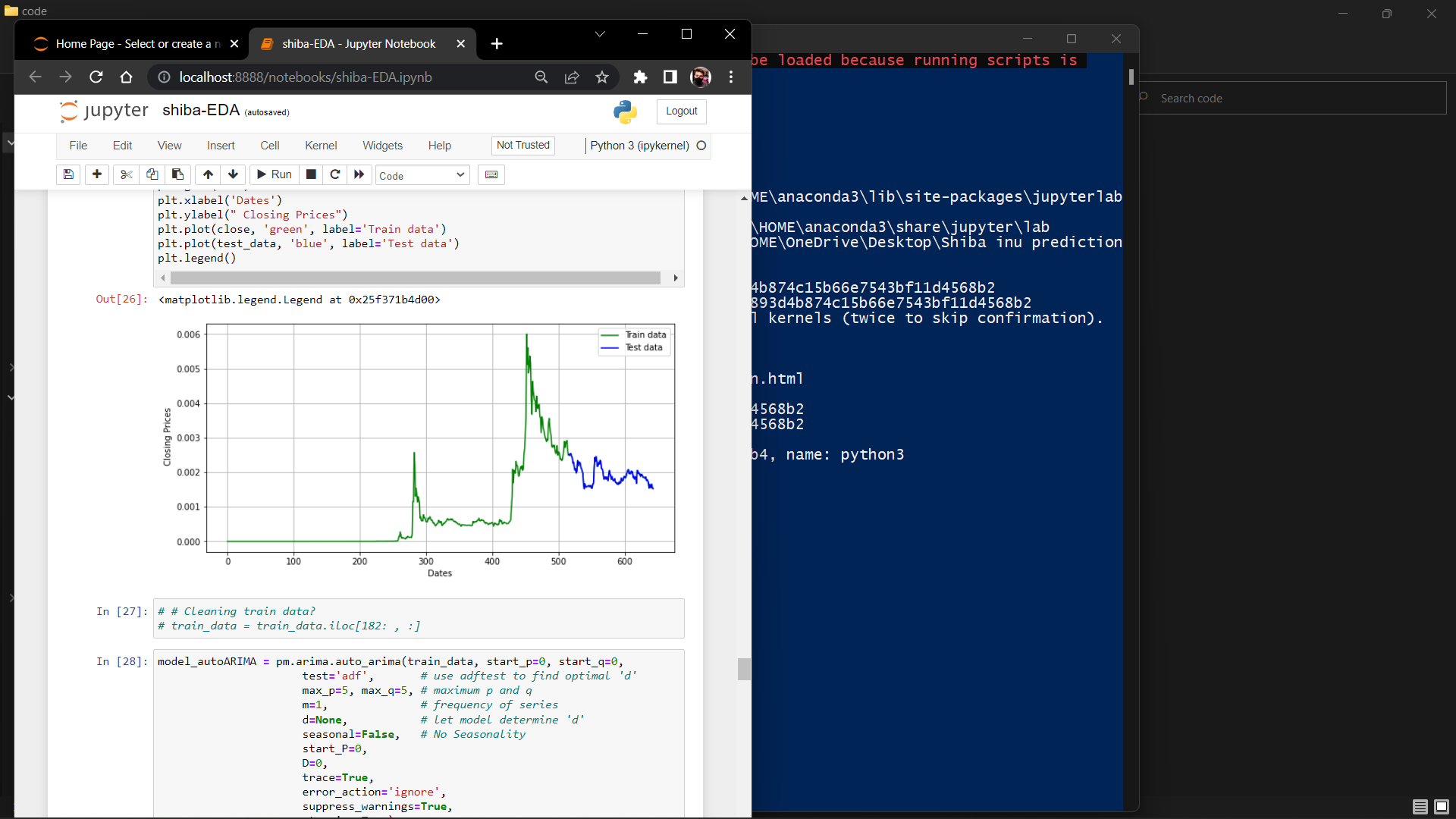
*Note :* ***We can see pearson and spearman are roughly the same, but kendall is very much different. That’s because Kendall is a test of strength of dependece (one could be written as a linear function of the other), whereas Pearson and Spearman are nearly equivalent in the way they correlate normally distributed data. All of these correlations are correct in their result, it’s just that Pearson/Spearman are looking at the data in one way, and Kendall in another***

## **PREDICTIVE ANALYSIS (ARIMA)**

Predictive Analysis/Modeling is a common statistical technique for predicting future behaviour . Predictive solutions are a form of data technology that analyses historical and current data to build a model which can be used to forecast future results.

Now we hace come to the stage where we prepare our data for prediction/forecasting Purpose .The data gets divided into Train (80%) and Test(20%) set

The training data is used for model to learn and Testing Data is use to predict and evaluating Model



*Fig 14 : Train-Test Split*

#### **ARIMA**

An autoregressive integrated moving average(ARIMA), is a statistical analysis model/tool that uses [time series data](https://www.investopedia.com/terms/t/timeseries.asp) to either better understand the data set or to predict future trends.

A statistical model is autoregressive if it predicts future values based on past values. For example, an ARIMA model might seek to predict a crypto's future closing prices based on its past performance or forecast a company's earnings based on past periods

An ARIMA model can be understood by outlining each of its components as follows:

[***Autoregression (AR****)*](https://www.investopedia.com/terms/a/autoregressive.asp): is a model in which a changing variable regresses on its own lagged, or previous, value.

***Integrated (I):****denotes the separating of raw observations in order for the time series to become stationary (i.e., data values are replaced by the difference between the data values and the previous values).*

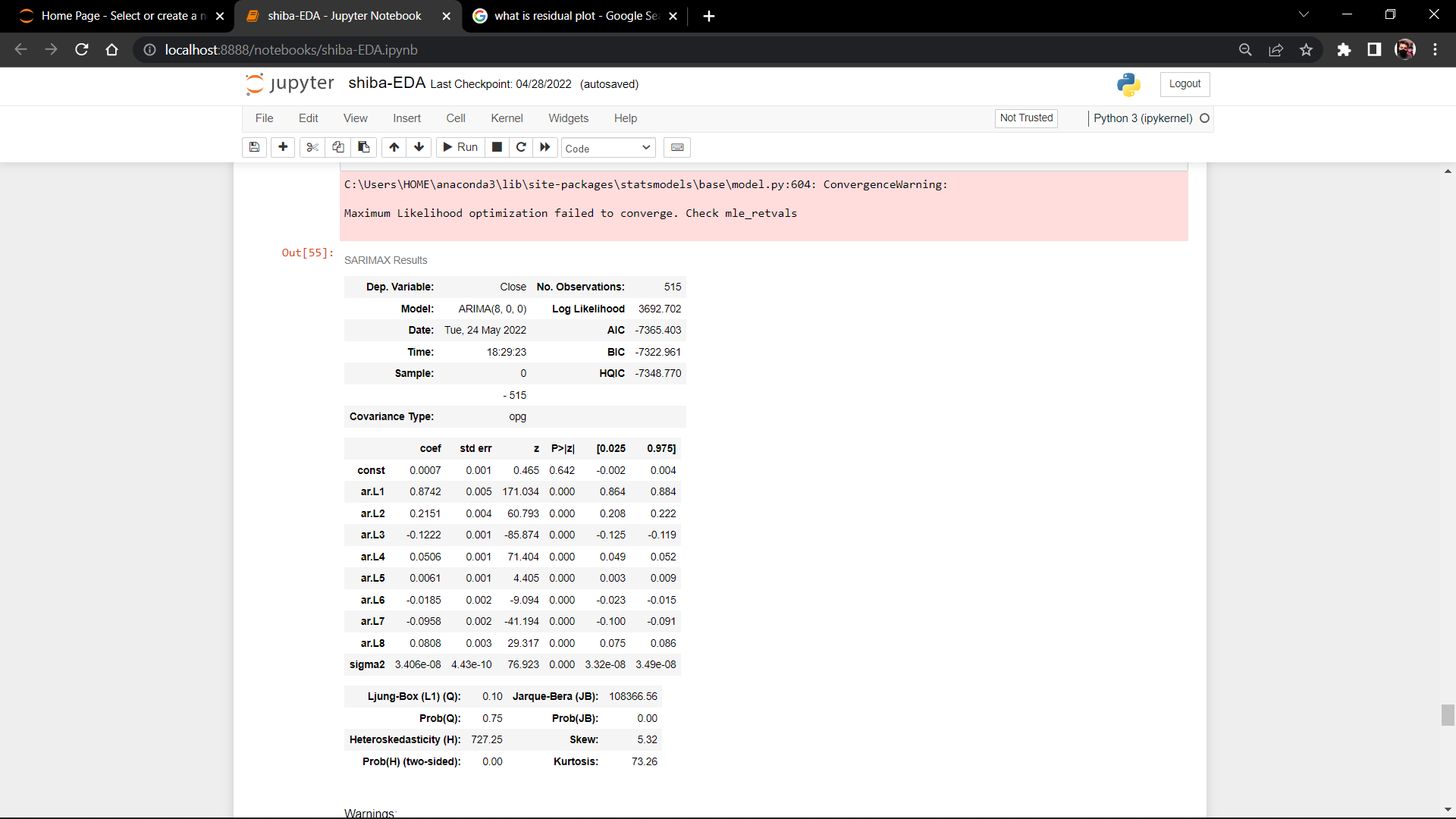
[***Moving average (MA)***](https://www.investopedia.com/terms/m/movingaverage.asp): *a rolling average model applied to delayed observations that considers the reliance between an observation and a leftover error*. ARIMA treats each component as a parameter with a consistent nomenclature. ARIMA with p, d, and q is a standard protocol for ARIMA models, where integer values substitute the parameters to indicate the type of ARIMA model used. The parameters are as continues to follow

Each component in ARIMA functions as a parameter with a standard notation. For ARIMA models, a standard notation would be ARIMA with p, d, and q, where integer values substitute for the parameters to indicate the type of ARIMA model used. The parameters can be defined as:

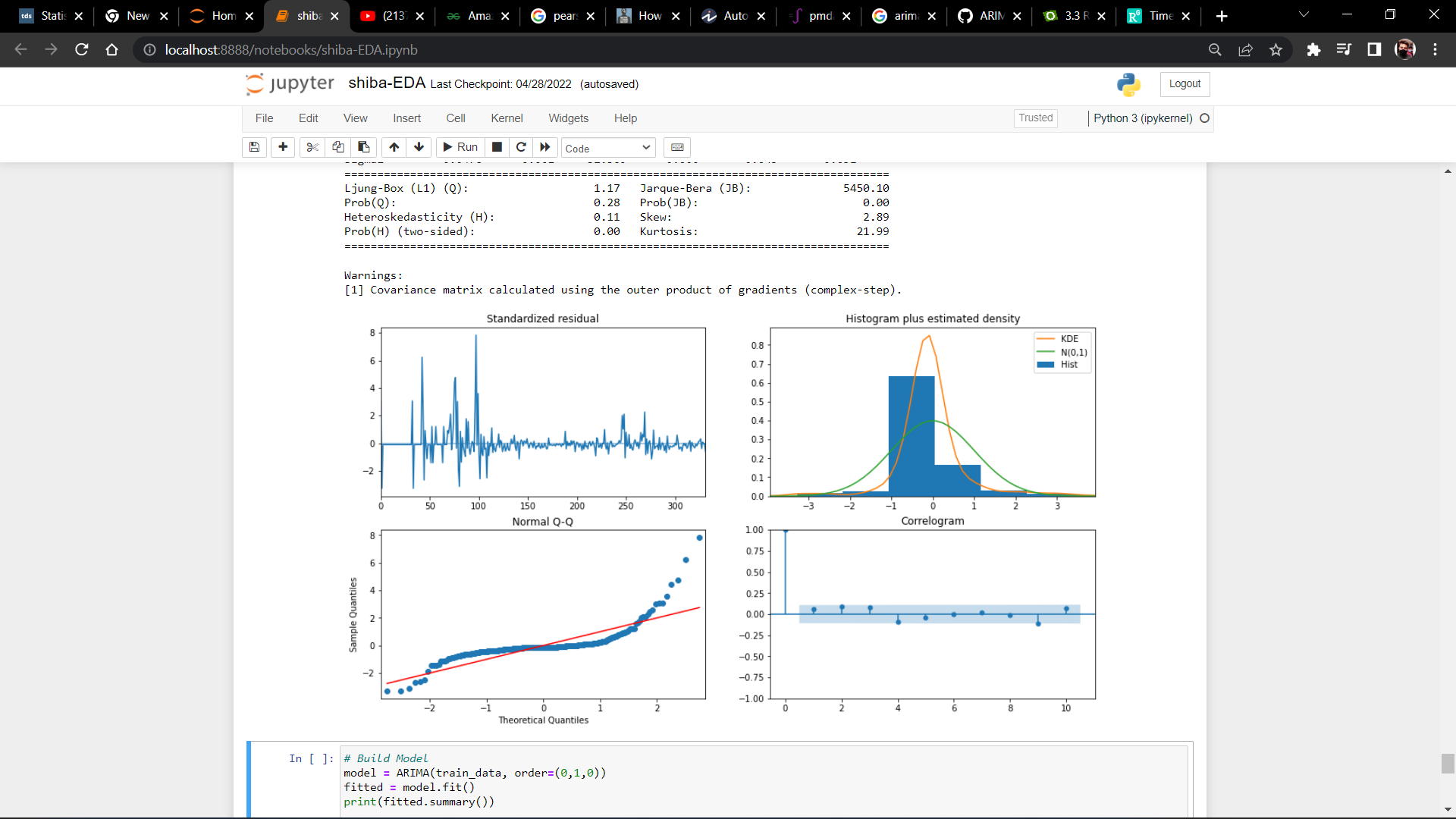
*p: the number of lag observations in the model; also known as the lag order.*

*d: the number of times that the raw observations are differenced; also known as the degree of differencing.*

*q: the size of the moving average window; also known as the order of the moving average*.



*Fig 15 : ARIMA Model Summary*



*Fig 15 : Residual Plot*

*Note :* ***The optimal p,d,q value are (8,0,0)***

***PREDICTION ON TEST DATA (ARIMA)***

After Performing bunch of predictive Analysis The Predictions are as follow (ARIMA MODEL)

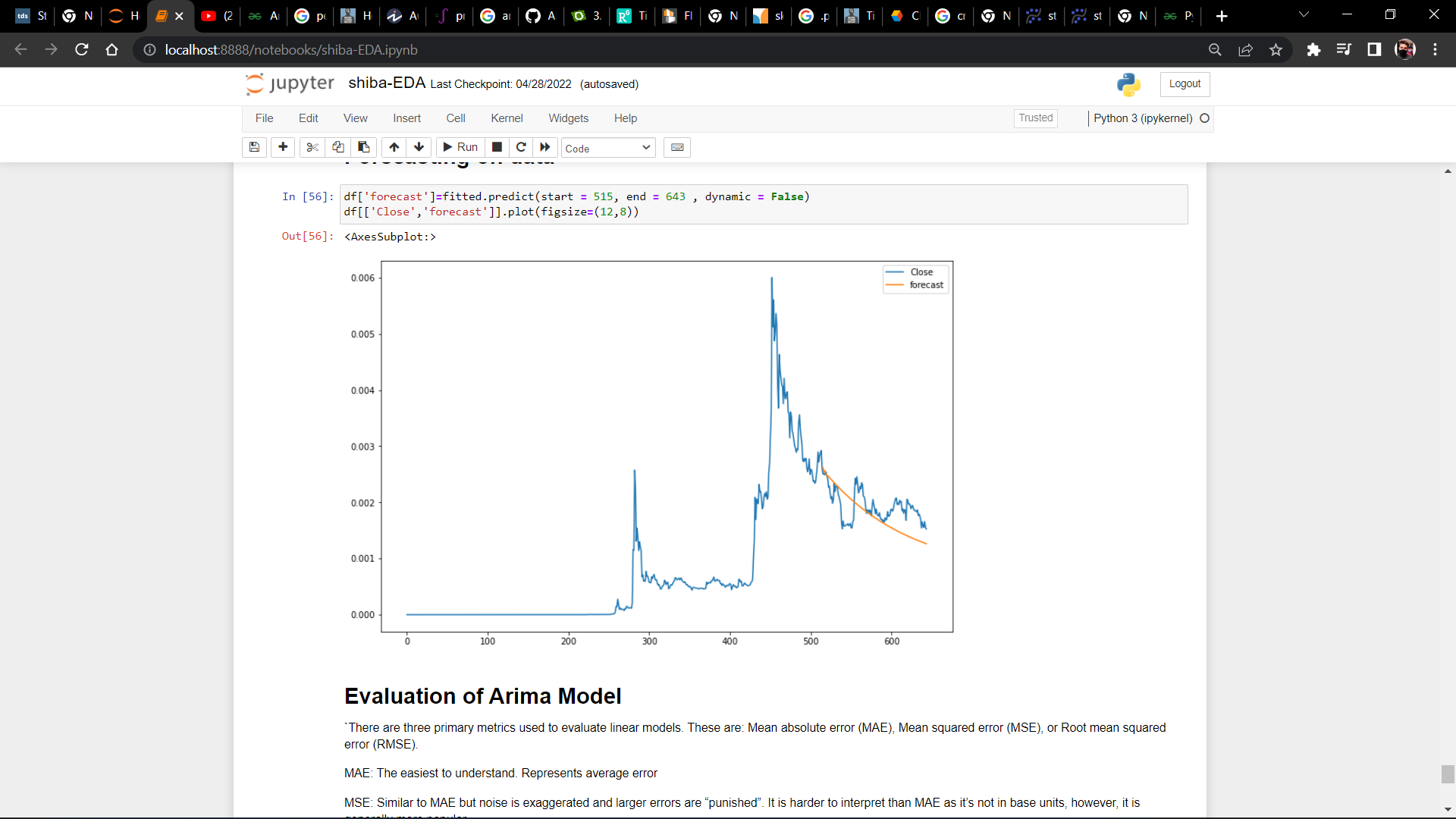


Fig 16 : Prediction on Test Data

***EVALUATION OF MODEL (ARIMA)***

|  |  |  |  |
| --- | --- | --- | --- |
| Model | MAE | MSE | RMSE |
| ARIMA | 0.000308 | 0.000000 | 0.000364 |

***Note : With the ARIMA , I was able to minimize the Error and MSE has showed 0 (Zero)***

# FUTURE SCOPE

[11]Many of the top IT businesses have made information about how they forecast their most essential services public in recent years. We're not talking about the necessity for many concurrent projections of the enormous number of measures that affect a company's bottom line here, but rather key concerns. Companies typically indicate that they are utilising deep learning with a probabilistic component in these circumstances, where the quality of the forecast is critical.

Further implementation can be done through implementing models like N-Beats , RNN , Stacked RNN etc.

[12] Empirically, it appears that accurate statistics are being used less and less in data modelling and prediction. Do not be discouraged: the field of statistics continues to thrive and answer fascinating questions about statistics. Machine learning approaches and results-oriented statistical methods, rather than sophisticated theories, closed-form solutions, or proofs of convergence, are winning out in actual deployment and real-world use cases, particularly for low-stakes forecasts that simply need to be good enough.

# OBSSERVATIONS

* OBSERVATION1 - The historical trend of Shiba inu before April2021 is kind of constant and quite dead in case of growth , Since this series contain Lots of zero value it might affect the predictive analysis hence we have to perform some Preprocessing further
* OBSERVATION 2 : It turned out that data is actually Non- Stationary after performing Augmented Dickey-fuller test , which could effect predictive modelling/analysis , thus in future we have to perform some sort of transformation to make it stationary (i.e. Log Transformation)
* OBSERVATION - 3 (As moving Average is Plotted so this smoothens out the noise based on mean we can clearly see it as a "Trend Spotting" tool and after May the trend is upward .
* OBSERVATION - 4 : Here, we can see that Durbin-Watson statistics are closer to 0 (0.05). Hence, there is some positive autocorrelation to the linear model of Closing Values
* OBSERVATION - 5 - We can see pearson and spearman are roughly the same, but kendall is very much different. That’s because Kendall is a test of strength of dependece (one could be written as a linear function of the other), whereas Pearson and Spearman are nearly equivalent in the way they correlate normally distributed data. All of these correlations are correct in their result, it’s just that Pearson/Spearman are looking at the data in one way, and Kendall in another.
* OBSERVATION – 6 – Due to recent drop in shiba inu price , the predictions turned out to be very precise.

# CONCLUSION

In this report we evaluated the insights gained from the Shiba Inu Historical Data, we come to an understanding that how the it is mathematically and statistically possible to make sense out of Time Series Data. We came up with lots of Observations to justify whether it’s a good asset or not, In particular the ARIMA and Auto-Arima has given good results on test data. However, comparing these mathematical approach to more general approach we can conclude that Shiba Inu Is not just about getting good prediction using machine learning , It’s a quite Volatile market and investing is quite risky because there are many more technical factor that effects the price of Shiba Inu . Considering it as a meme-coin we can also conclude that the Community and Social Media has also effect this market .

# REFERENCES

|  |  |
| --- | --- |
| [1] | D. K. ,. G. A. P. Hristos Tyralis. [Online]. Available: https://www.researchgate.net/publication/329279811\_Univariate\_Time\_Series\_Forecasting\_of\_Temperature\_and\_Precipitation\_with\_a\_Focus\_on\_Machine\_Learning\_Algorithms\_a\_Multiple-Case\_Study\_from\_Greece. |
| [2] | E. S. G. Jr.. [Online]. Available: https://doi.org/10.1002/for.3980040103. |
| [3] | A. Bulatov. [Online]. Available: https://source.sheridancollege.ca/fast\_sw\_mobile\_computing\_theses/5/. |
| [4] | D. a. S. E. Saleh. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1882/1/012033. |
| [5] | KSSRINIVASARAO. [Online]. Available: https://github.com/KSSRINIVASARAO/ARIMA-Time-Series-Model. |
| [6] | [Online]. Available: https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html. |
| [7] | A. Hayes. [Online]. Available: https://www.investopedia.com/terms/n/null\_hypothesis.asp#:~:text=Key%20Takeaways%201%20A%20null%20hypothesis%20is%20a,null%20hypothesis%20within%20a%20certain%20confidence%20level.%20. |
| [8] | A. Hayes. [Online]. Available: https://www.investopedia.com/terms/n/null\_hypothesis.asp#:~:text=Key%20Takeaways%201%20A%20null%20hypothesis%20is%20a,null%20hypothesis%20within%20a%20certain%20confidence%20level.%20. |
| [9] | S. Pandian. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-to-time-series-analysis/. |
| [10] | J. Brownlee, 30 January 2017. [Online]. Available: https://machinelearningmastery.com/decompose-time-series-data-trend-seasonality/. |
| [11] | A. Nielsen. [Online]. Available: https://medium.com/oreillymedia/the-future-of-time-series-forecasting-bd83c2aca9a8. |
| [12] | [Online]. Available: Empirically there seems to be less and less use of proper statistics in modeling data and generating predictions. Do not despair: the field of statistics continues to thrive and answer interesting questions that are related to statistics. And yet — partic. |