**A REVIEW OF GLOBAL METAL SCENARIO AND METAL PRICE FORECASTING**

Submitted for the Requirement of

MINE DESIGN PROJECT-II END-TERM EVALUATION

by

**ANANYA CHAVALI (201MN009)**

**ANSHIKA KHANDELWAL (201MN010)**

Under the guidance of

**DR. ANUP KUMAR TRIPATHI**

Associate Professor

Department of Mining Engineering

NITK, Surathkal



NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL - 575025

APRIL – 2024

CERTIFICATE

This is to certify that the major project report - 2 entitled “A REVIEW OF GLOBAL METAL SCENARIO AND METAL PRICE FORECASTING” submitted by Ananya Chavali (201MN009) and Anshika Khandelwal (201MN010) for the evaluation of partial requirements for the award of Bachelor of Technology degree in Mining Engineering at the National Institute of Technology Karnataka, Surathkal is an authentic work carried out by them under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University/Institute for the award of any Degree or Diploma.

Date:

Dr. Anup Kumar Tripathi

Associate Professor

Department of Mining Engineering

NITK, Surathkal

ACKNOWLEDGEMENT

First and foremost, we would like to extend our gratitude and indebtedness towards our guide Dr. Anup Kumar Tripathi, for his kindness in allowing us for introducing the present topic and for his inspiring guidance, constructive criticism and valuable suggestions. We are sincerely thankful to him for his able guidance and pain taking effort in improving our understanding of this project.

We would also like to thank our HOD Dr. Harsha Vardhan for giving us the opportunity to

work on the project. An assemblage of this nature could never have been attempted without reference to and inspiration from the works of others whose details are mentioned in reference section. We acknowledge our gratitude to all of them.

Last but not the least we would like to thank each and every one who helped us directly or

indirectly.

Thank you.

Ananya Chavali

(201MN009)

Anshika Khandelwal

(201MN010)

**ABSTRACT**

This report explores the multifaceted landscape of metal price forecasting, combining economic analysis with advanced predictive modeling techniques. The demand for accurate metal price predictions has grown exponentially in a dynamic global market characterized by intricate supply chains, geopolitical uncertainties, and rapid technological advancements. This review synthesizes insights from diverse sources, encompassing economic indicators, technological trends, and regulatory frameworks.

The results of the study demonstrate the effectiveness of the proposed predictive modeling approach in forecasting metal prices. The model provides valuable insights into the factors influencing metal price movements and offers reliable forecasts for strategic decision-making and risk management in industries reliant on metal commodities.

**Table of Contents**

[**Introduction** 10](#_Toc163515849)

[1.1 Overview 10](#_Toc163515850)

[1.2 Objective 11](#_Toc163515851)

[**Literature review** 13](#_Toc163515852)

[2.1 Historical Trends in Metal Demand 13](#_Toc163515853)

[2.2 Current Supply and Future Demand for Metals 15](#_Toc163515854)

[Population Growth and Urbanization: 15](#_Toc163515855)

[Technological Advancements: 16](#_Toc163515856)

[Renewable Energy Transition: 16](#_Toc163515857)

[Electric Vehicles (EVs): 16](#_Toc163515858)

[Infrastructure Development: 16](#_Toc163515859)

[Consumer Electronics: 16](#_Toc163515860)

[Environmental and Sustainability Initiatives: 16](#_Toc163515861)

[2.3 Critical and Rare earth Metals - Trends and Prospects 16](#_Toc163515862)

[Clean Energy Transition 17](#_Toc163515863)

[Electric Vehicles (EVs) 17](#_Toc163515864)

[Technology and Electronics 17](#_Toc163515865)

[Strategic Importance 17](#_Toc163515866)

[Supply Chain Concerns 17](#_Toc163515867)

[Recycling Initiatives 18](#_Toc163515868)

[Research and Exploration 18](#_Toc163515869)

[Policy and Regulation 18](#_Toc163515870)

[Price Dynamics 18](#_Toc163515871)

[Environmental and Social Considerations 18](#_Toc163515872)

[2.4 Rare earth Metals – The Shift 18](#_Toc163515873)

[2.5 The Key Metals 20](#_Toc163515874)

[2.6 Metal Price Forecasting 22](#_Toc163515875)

[2.6.1 Time Series regression model 24](#_Toc163515876)

[2.6.2 Naïve Forecasting 27](#_Toc163515877)

[2.6.3 Moving Average Forecasting 28](#_Toc163515878)

[2.6.4 LSTM Method 31](#_Toc163515879)

[2.7 Technological Innovation 39](#_Toc163515880)

[2.8 Socio-Economic Dynamics 41](#_Toc163515881)

[Labor Conditions and Employment Opportunities: 41](#_Toc163515882)

[Community Impacts and Social Development: 41](#_Toc163515883)

[Resource Dependency and Economic Diversification: 42](#_Toc163515884)

[Environmental Sustainability and Social Responsibility: 42](#_Toc163515885)

[Inclusive Growth and Stakeholder Collaboration: 42](#_Toc163515886)

[2.9 Emergence of Renewable energy 42](#_Toc163515887)

[**Methodology** 48](#_Toc163515888)

[3.1 Data Collection 48](#_Toc163515889)

[3.2 Data Preprocessing 49](#_Toc163515890)

[3.3 Model Selection 51](#_Toc163515891)

[**Data Analysis and Modelling** 54](#_Toc163515892)

[4.1 Data pre-processing 54](#_Toc163515893)

[4.1.1. Importing all datasets 54](#_Toc163515894)

[4.1.2 Copper price trends 54](#_Toc163515895)

[4.2.3 Silver price trends 54](#_Toc163515896)

[4.2.4 Gold price trends 54](#_Toc163515897)

[4.2 Model Implementation 54](#_Toc163515898)

[4.2.1 Simple Linear Regression 54](#_Toc163515899)

[4.2.2 Multiple Linear Regression 54](#_Toc163515900)

[4.2.3 Moving average forecasting 54](#_Toc163515901)

[4.2.4 Exponential Moving average Forecasting 58](#_Toc163515902)

[4.2.5 Naïve Forecasting 58](#_Toc163515903)

[4.2.6 Long Short Term Memory 60](#_Toc163515904)

[4.2.7 Auto Regressive Integrated Moving Average (ARIMA) 63](#_Toc163515905)

[4.2.8 Seasonal Autoregressive Integrated Moving Average (SARIMA) 64](#_Toc163515906)

[4.2.9 Model Performance Comparison of ARIMA, SARIMA and LSTM 65](#_Toc163515907)

[**Dashboard Integration** 68](#_Toc163515908)

[5.1 Selecting the right Dashboarding solution 69](#_Toc163515909)

[5.2 Comparing dashboarding tools 71](#_Toc163515910)

[5.3 Streamlit as a tool 72](#_Toc163515911)

[5.4 Making predictions using dashboard 73](#_Toc163515912)

[**Conclusion** 75](#_Toc163515913)

[**Bibliography** 76](#_Toc163515914)

**LIST OF FIGURES**

Figure 1: 10 year historical price of LME Index 13

Figure 2: 10 year global gold price (in US dollars adjusted as per current inflation) 14

Figure 3: Global Uranium production (in tonnes) in 2015 19

Figure 4: Real commodity price index, metals, 1850 to 2020 21

Figure 5: Graphical Representation of Simple Linear Regression 25

Figure 6: Graphical Representation of Simple Moving Average 29

Figure 7: Graphical Representation of Exponential Moving Average 30

Figure 8: Illustration of recurrent connections in RNNs 32

Figure 9: Graphical Representation of trained ,tested and predicted values for LSTM network. 36

Figure 10: Representation of LSTM architecture 36

Figure 11: Long term and short-term memory representation 38

Figure 12: A more intuitive representation of LSTM network 38

Figure 13: Current Global share of renewable energy 44

Figure 14: Representation of steps involved in data pre processing 49

Figure 15: Representation of model selection considerations 53

Figure 16: Graphical representation of copper prices between 2020-2024 56

Figure 17: Graphical Representation of copper opening prices between 2020-2024 56

Figure 18: Graphical representation of copper high prices between 2020-2024 56

Figure 19: Graphical representation of volume of copper futures vs date 59

Figure 20: Graphical representation of silver price trends between 2020-2024 61

Figure 21: Representation of gold high prices between 2020-2024 62

Figure 22:Graphical Representation of open, high and low gold prices over 2019-2024 64

Figure 23: Pictorial Representation of Train and Test split 65

Figure 24:Graphical comparison of actual vs naive vs EMA forecast 71

Figure 25: Plot obtained for actual vs predicted values in Naive forecast 73

Figure 26:Reprsenatation of actual vs forecasted prices using LSTM 76

Figure 27: Making prediction using streamlit 86

**Chapter 1 - Introduction**

# **Introduction**

## 1.1 Overview

The global metal scenario represents a dynamic interplay of economic, technological, and geopolitical forces, influencing the supply, demand, and pricing of essential metals. This literary overview delves into the intricate fabric of the metal industry, exploring both historical perspectives and contemporary challenges. Emphasizing the crucial need for accurate metal price forecasting, the review integrates insights from economic analyses, technological trends, and predictive modeling methodologies.

The historical evolution of the global metal scenario reveals a complex tapestry shaped by industrialization, technological revolutions, and geopolitical shifts. From the industrial boom of the 19th century to the current era of advanced manufacturing and renewable energy, metals have been at the forefront of global economic development. This section examines key trends in metal production, consumption patterns, and the emergence of critical metals in strategic industries.

Accurate metal price forecasting is paramount for informed decision-making in the volatile metal markets. This section provides a comprehensive overview of forecasting methodologies, encompassing economic analyses, statistical modeling, and machine learning applications. The literature explores the role of supply and demand dynamics, macroeconomic factors, and technological advancements in shaping metal prices. Regression analyses, time series modeling, and machine learning algorithms are scrutinized for their effectiveness in predicting future price trends.

Technological innovation stands as a cornerstone in shaping the future of the metal industry. This section delves into the transformative impact of advancements such as automation, artificial intelligence, and sustainable mining practices. The integration of technology not only enhances operational efficiency but also introduces new variables into forecasting models. Case studies highlight the influence of breakthrough technologies on metal prices and the adaptive strategies adopted by industry players.

## 1.2 Objective

The aim of this project is to apply descriptive, predictive, and prescriptive analytics in the field of metal pricing and to elevate understanding of metal prices and econometrics which will help to gain strategic insights about global metal pricing forecasts and allied areas.

The project has the following main objectives:

1. Literature review of Global Metal scenario, Current and Future Demand and Metal Price Forecasting.

2. Data collection from historical prices and trends.

3. Application of different techniques such as exponential forecasting, regression analysis, LSTM and other algorithms for metal price prediction.

4. Use of Python software in metal price forecasting.

**Chapter 2 – Literature Review**

# **Literature review**

### 2.1 Historical Trends in Metal Demand

Historical trends in metal demand have been shaped by various factors, including industrialization, technological advancements, population growth, and global economic conditions.

Movements in metals prices have long been a focus of empirical study and have leant themselves as examples for numerous economic studies. Many mineral economists have dedicated their academic careers to studying various issues relating to primary commodities and their price movements [1].

Metals play a crucial role not only for producing nations but also for consumer countries, serving as a vital input in numerous industries. Consequently, the dynamics of commodity prices hold significant importance for global economic activities. For many developing nations, revenue generated from metal exports often constitutes a primary income source. Consequently, fluctuations in commodity prices can exert a substantial impact on the overall macroeconomic performance [2].

Figure 1: 10 year historical price of LME Index

Moreover, companies engaged in raw material processing face potential adverse effects due to abrupt increases in commodity prices, leading to elevated input costs. The recent emergence of markets in developing regions has contributed to an upward push on commodity prices. Simultaneously, technological advancements and notable improvements in mineral extraction methods exert downward pressure on commodity prices. This complex interplay of factors underscores the intricate relationship between commodity prices, global economic dynamics, and the well-being of nations and industries alike [3].

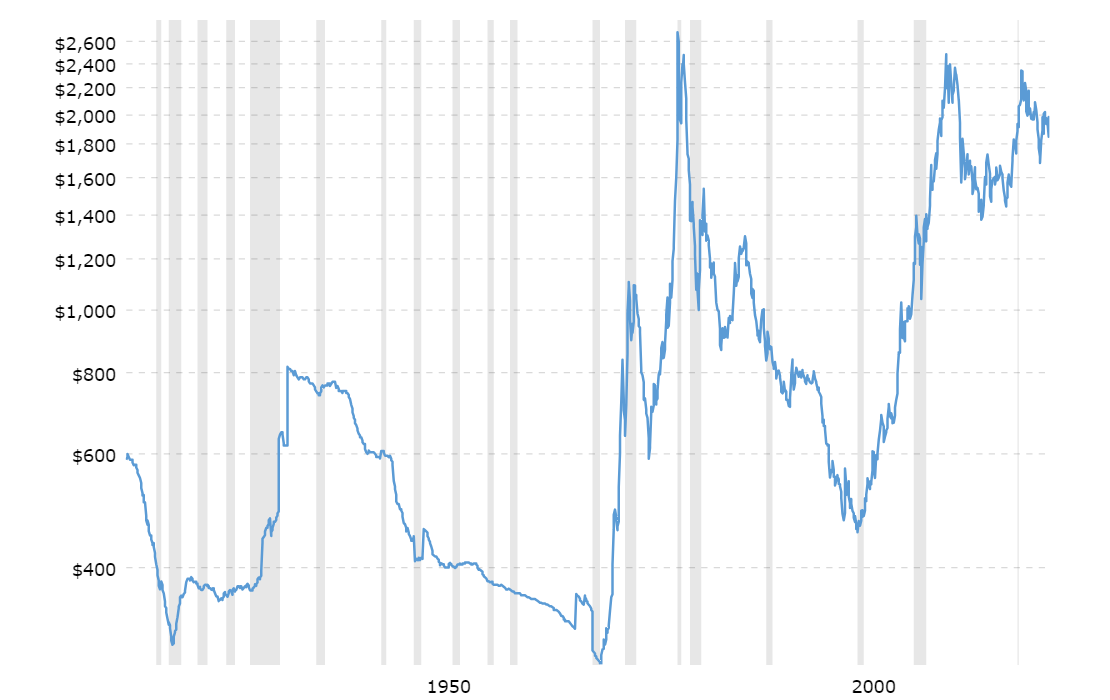


Figure 2: 10 year global gold price (in US dollars adjusted as per current inflation)

For the majority of metals, it is reasonable to posit that there has been no significant overall change in the quantity of unused raw stock over the past approximately 100 years. This assumption implies that the rate of metal extraction has closely aligned with industry demand. Notably, precious metals may deviate from this general trend, a point I'll delve into shortly. Assuming that extraction rates have historically mirrored demand, one can extrapolate historical data on mineral extraction to infer anticipated future demand. This assumption serves as a foundation for constructing an empirical function, leveraging historical per-capita production trends to estimate forthcoming per-capita demand.

## 2.2 Current Supply and Future Demand for Metals

Examining the sustainable balance between mineral resource consumption and reserves is a critical aspect of long-term supply and demand analysis. To assess the long-term outlook accurately, there has been a concerted effort to compare consumption rates with the Earth's crustal abundance. The depletion rate, notably highest for Pb, serves as a benchmark. Depletion rates for elements such as Au, Sb, Cu, Sn, Ag, and Zn fall within the 40-10 times the rate of Ni, while Hg, Cd, Mo, Se, Cr, Bi, W, and As register depletion rates in the 1-10 times the rate of Ni range. As mentioned earlier, mineral resources with a shorter lifespan, including Pb, Au, Cu, Zn, Sn, and Ag (excluding Ge and Ta), exhibit high depletion rates for these elements [4].

Over the observed period, global per capita metal consumption has increased, yet the intensity of use, measured by the tonnage of metal consumed relative to Gross Domestic Product (GDP), has shown a decline. While per capita consumption has risen nearly universally, the intensity of use has increased only in developing nations and in Japan, specifically for nonferrous metals.

A distinctive shift characterizes the trend in global metal consumption over the last decade, diverging significantly from the preceding decade. This shift can be attributed to a substantial decrease in the rate of growth in worldwide GDP per capita during the last decade. Additionally, a notable contributing factor is the steep upward trajectory in real metal prices that transpired in the post-1973 period [5].

The future demand for metals is influenced by various factors, including population growth, urbanization, technological advancements, and shifts in global economic activities. The following factors summarise the overall effect of the intangible factors [6].

### Population Growth and Urbanization:

As the global population continues to grow, especially in developing countries, there is an increasing demand for infrastructure development, housing, and consumer goods. Urbanization drives the need for metals in construction, transportation, and manufacturing.

### Technological Advancements:

The rapid pace of technological innovation is driving the demand for metals such as rare earth elements, lithium, cobalt, and others used in the production of electronic devices, batteries, renewable energy technologies, and electric vehicles (EVs).

### Renewable Energy Transition:

The global push towards renewable energy sources, such as solar and wind power, has led to a growing demand for metals like copper, aluminum, and rare earth elements for the manufacturing of solar panels, wind turbines, and energy storage systems.

### Electric Vehicles (EVs):

The increasing adoption of electric vehicles is a major driver for the demand for metals like lithium, cobalt, nickel, and copper. These metals are essential components in batteries and electric motor systems.

### Infrastructure Development:

Ongoing and future infrastructure projects, including those related to transportation (roads, bridges, railways), energy (power grids), and communication (5G networks), contribute to the demand for various metals.

### Consumer Electronics:

The continued proliferation of smartphones, laptops, and other electronic devices sustains the demand for metals like gold, silver, copper, and rare earth elements.

### Environmental and Sustainability Initiatives:

The emphasis on sustainable practices and environmental initiatives may lead to increased demand for metals used in energy-efficient technologies and green infrastructure.

It's important to note that the demand for specific metals can vary based on technological developments and market dynamics. Additionally, recycling efforts and the development of alternative materials can influence the overall demand for traditional metals. Future trends and developments in technology, geopolitics, and global economic conditions will play a significant role in shaping the demand for metals.

## 2.3 Critical and Rare earth Metals - Trends and Prospects

Critical minerals, also known as strategic minerals or critical materials, are a group of raw materials that are considered essential for the economic and industrial well-being of a country, and whose supply may be at risk due to geological scarcity, geopolitical issues, trade policies, or other factors. These minerals play a crucial role in various industries, including technology, defense, energy, and manufacturing .

The criticality of these minerals arises from their importance in key technologies and industries, as well as concerns about the concentration of their production in a small number of countries. Governments and industries often work to diversify sources, invest in exploration and mining technologies, and promote recycling to ensure a stable and secure supply of these critical minerals. As technologies evolve and new industries emerge, the list of critical minerals may be subject to change [7].

the prospects of critical metals continue to be significant, driven by various factors including technological advancements, the transition to clean energy, and the increasing demand for high-tech products. Here are some key aspects influencing the prospects of critical metals:

### Clean Energy Transition

Critical metals, such as lithium, cobalt, rare earth elements, and others, are essential components in renewable energy technologies like solar panels and wind turbines. Additionally, these metals play a crucial role in energy storage systems, such as lithium-ion batteries, which are integral to the widespread adoption of electric vehicles and grid storage.

### Electric Vehicles (EVs)

The global shift toward electric vehicles is a major driver for certain critical metals like lithium, cobalt, and nickel. These metals are crucial for the production of batteries that power electric cars [8].

### Technology and Electronics

Critical metals are essential in the manufacturing of a wide range of electronic devices, including smartphones, laptops, and other high-tech products. This demand is expected to grow as technology continues to advance [9].

### Strategic Importance

Some critical metals, including rare earth elements, are strategically important for national security and defense applications. They are used in the production of advanced materials, electronics, and military technologies.

### Supply Chain Concerns

Many critical metals have concentrated production in a few countries, leading to concerns about supply chain vulnerabilities. Diversifying sources and developing new mining projects are considerations to address these concerns.

### Recycling Initiatives

As the awareness of environmental sustainability increases, there is a growing emphasis on recycling critical metals from end-of-life products. Recycling efforts can help mitigate supply risks and reduce the environmental impact of mining.

### Research and Exploration

Ongoing research and exploration efforts are aimed at discovering new sources of critical metals and developing more efficient extraction and processing technologies. This includes exploration in unconventional sources such as deep-sea mining for certain minerals.

### Policy and Regulation

Governments and international bodies are increasingly recognizing the importance of critical metals and implementing policies to secure their supply. This may include investments in domestic mining projects, trade agreements, and environmental regulations.

### Price Dynamics

The prices of critical metals can be influenced by market dynamics, geopolitical events, and changes in demand. As technology evolves and new applications emerge, the demand for certain critical metals may experience price fluctuations.

### Environmental and Social Considerations

The mining and processing of critical metals can raise environmental and social concerns. Sustainable and responsible mining practices are becoming increasingly important in shaping the prospects of critical metals.

## 2.4 Rare earth Metals – The Shift

The rare earth elements (REEs) represent a class of critical minerals that include the fifteen lanthanide elements (Z = 57 through 71) and yttrium (Z = 39) in the periodic table. They are called REEs because most of them were originally isolated in the 18th and 19th centuries as oxides from rare minerals. Because of their reactivity, the REEs were set up to be delicate to upgrade to pure essence. likewise, effective separation processes weren't developed until the 20th century because of REEs’ chemical similarity. Current global primary REE product is about 130,000 metric tons of rare- earth oxide (REO) original content per time, with the 2014 REE request worth ∼ US$2.05 billion [10].

Global demand for these rudiments has steadily increased. still, prices have remained steady after a significant shaft caused by the operation and posterior junking of Chinese import restrictions and a posterior surfeit of some of the REE. utmost REE consumption is by mature requests like catalysts, glassmaking, lighting, and metallurgy. Newer high- growth requests similar as attractions, pottery, and batteries enthrall the remaining 41 [11].

The increase in operation of the REE has led to a coincident increase in demand for these rudiments. still, utmost of the demand is still met by primary product from mines, primarily within China, which dominated the global consumption of 119,000 Ton of rare earth oxides (REO) in 2014. REEs could be substituted by some other rudiments, but overall, some REEs are considered more critical than others. still, one of the main issues with REE mine product is the so- called balance problem, where the vast maturity of REE product is dominated by La and Ce but the maturity of REE demand is for Nd or Dy. This could lead to a situation where La and Ce are overproduced, and the demand for Nd or Dy for use in end- products analogous as lodestones and batteries may not be met by primary product alone still, this issue can partly be overcome by recovering, as the products that would be reclaimed generally contain the potentially undersupplied Nd or Dy rather than the potentially overproduced lighter REE similar as La or Ce. Although, the small quantum of recycling of these rudiments means it's presently unclear how recycling could affect the REE request [12].

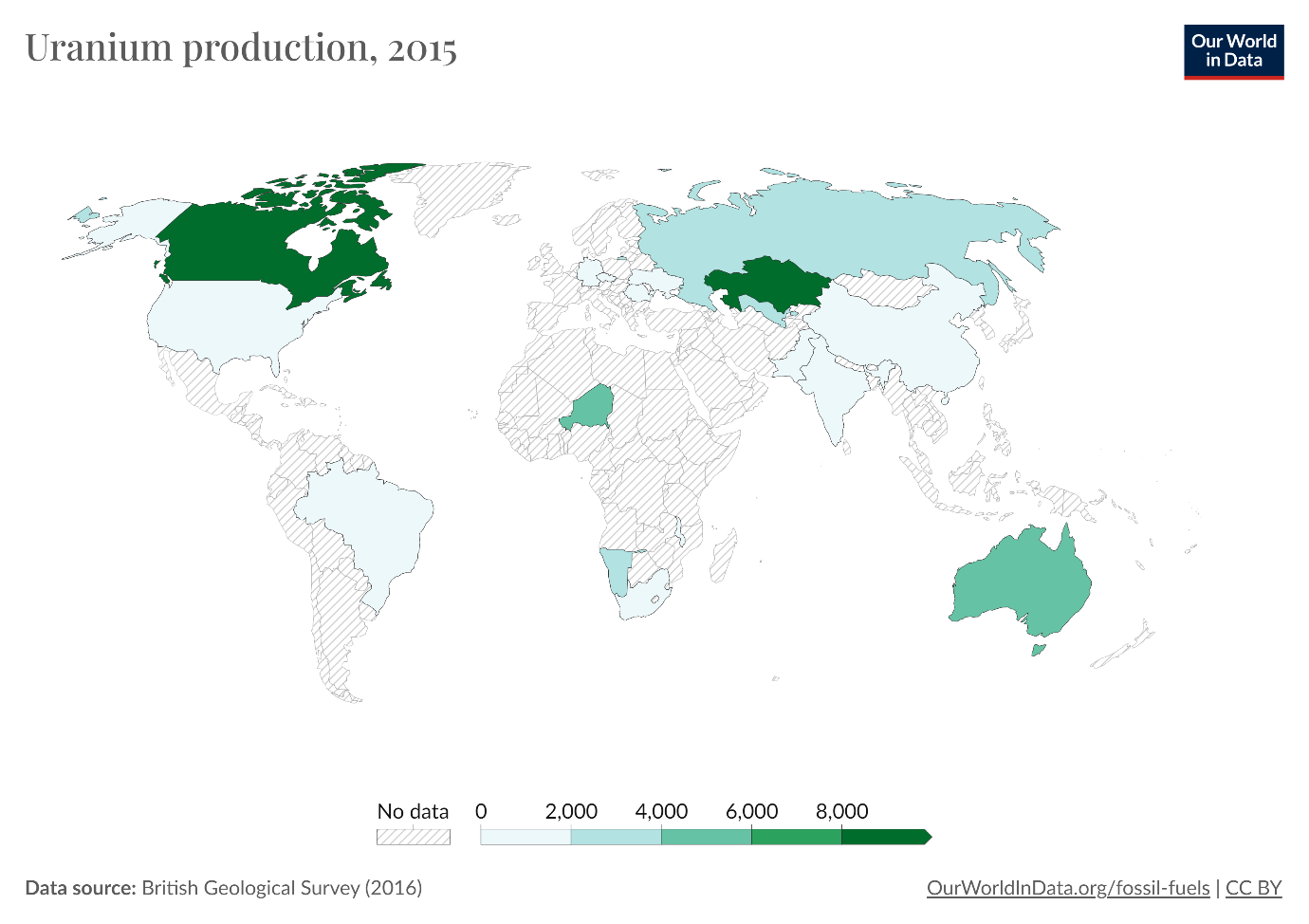


Figure 3: Global Uranium production (in tonnes) in 2015

One of the main obstacles to the recycling of these rudiments is the fact that the quantum of REE used in end- products ranges from <mg to several kg. This, combined with the complexity of their uses, the difficulty essential in separating the individual REE from each other to yield pure single rudiments, the occasionally-long life of certain uses (e.g., endless attractions in electrical technologies), and a variety of other further general reasons mean that < 1% of the REE used moment are reclaimed. The recycling that takes place is generally in the form of REE within endless attractions, fluorescent lighting, batteries, and the REE that are used as catalysts within the chemical assiduity, including in petroleum refining [13].

REE minerals can be concentrated in various ways based on physical, chemical, and mechanical properties that differ from the other constituents of an ore. Efforts are being made to develop cleaner and more sustainable processing technologies and improve waste management practices to mitigate these impacts [14].

Efforts are being made to mitigate the environmental impacts of rare earth elements. For example, research is ongoing to develop cleaner extraction and processing technologies, improve waste management practices, and promote recycling of rare earth materials to reduce the need for new mining operations. Additionally, stricter regulations and environmental standards can help minimize the negative effects of rare earth element production.

The recycling of rare earth elements (REEs) plays a vital role in reducing the demand for primary mining and minimizing the environmental impacts associated with their extraction [15].

Promoting and expanding rare earth element recycling requires collaboration among various stakeholders, including manufacturers, recyclers, policymakers, and consumers. Developing efficient recycling technologies, improving collection systems, and creating economic incentives for recycling can all contribute to the growth of a sustainable rare earth element recycling industry [16].

## 2.5 The Key Metals

Several metals play crucial roles in various industries and have significant importance on the global stage. These metals are essential for manufacturing, construction, technology, and other sectors. Given below are few of the key metals globally:

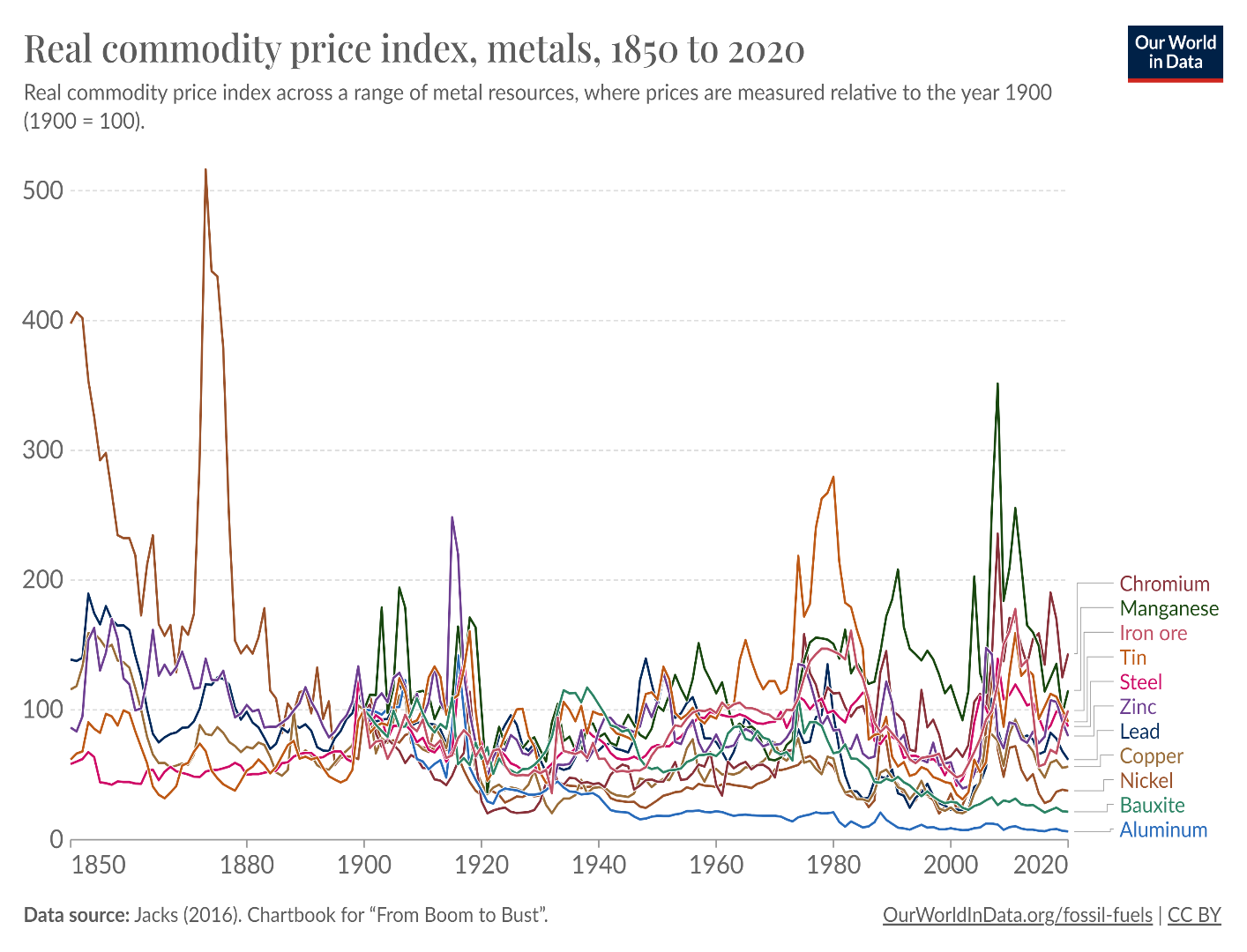


Figure 4: Real commodity price index, metals, 1850 to 2020

1. **Aluminum:** Lightweight and corrosion-resistant, aluminum is widely used in transportation (including automobiles and aircraft), construction, packaging, and electrical applications.
2. **Copper:** A highly conductive metal, copper is essential in electrical wiring, electronics, construction, and transportation. It is a critical component in renewable energy systems and electric vehicles.
3. **Iron and Steel:** Iron is a fundamental component in the production of steel, which is a major construction material and is used in manufacturing, infrastructure, and transportation.
4. **Gold and Silver:** Precious metals like gold and silver are valued for their use in jewelry, investment, and electronic components. They are also used in certain medical and industrial applications.
5. **Nickel:** Valued for its corrosion resistance and ability to withstand high temperatures, nickel is used in stainless steel production, batteries (especially in electric vehicles), and various other industrial applications.
6. **Lead:** Lead is used in batteries, ammunition, construction, and as a radiation shield in medical imaging and nuclear facilities.
7. **Zinc:** Zinc is primarily used in galvanizing steel to prevent corrosion. It also has applications in the production of alloys, batteries, and as a dietary supplement.
8. **Tin:** Tin is used in the production of solder, as well as in coatings for food packaging. It is also used in the manufacturing of electronic devices.
9. **Platinum Group Metals (PGMs):** Platinum, palladium, and rhodium are crucial in catalytic converters for automobiles. They also find applications in the production of jewelry, electronics, and fuel cells.
10. **Titanium:** Known for its strength-to-weight ratio and corrosion resistance, titanium is used in aerospace applications, medical implants, and various industrial processes.
11. **Lithium:** Essential for lithium-ion batteries used in electric vehicles, portable electronics, and energy storage systems.
12. **Cobalt:** Important in the production of rechargeable batteries, particularly in electric vehicles and electronic devices.

These metals are integral to modern industrial processes and technologies. The demand for them is influenced by economic trends, technological advancements, and shifts in consumer preferences. Additionally, as the world transitions toward renewable energy and electric vehicles, certain metals like lithium, cobalt, and rare earth elements are gaining increased attention due to their role in energy storage and clean technologies.

## 2.6 Metal Price Forecasting

Metal prices hold a central role in guiding investment decisions within the mineral industry, exerting a significant impact on the financial performance of companies operating in this sector. Two primary reasons underscore the influence of mineral prices on changes in mineral expenditure. Firstly, both current and historical price movements serve to mould expectations concerning future prices and the prospective profits derived from exploration and mining activities. Secondly, these prices play a pivotal role in shaping mining revenues and influencing the cost of capital essential for financing exploration endeavours. Given these considerations, mining companies would greatly benefit from the ability to forecast metal prices, providing invaluable assistance in their strategic planning [17] .

Metal price forecasting is crucial for various stakeholders, including industries, investors, governments, and consumers. Here are several reasons highlighting the importance of metal price forecasting:

1. **Strategic Planning for Industries:**

Industries heavily reliant on metals, such as automotive, construction, and electronics, depend on accurate price forecasts to plan their supply chains efficiently. This includes decisions related to procurement, inventory management, and production schedules.

1. **Risk Mitigation for Investors:**

Investors in metal markets need reliable forecasts to make informed decisions. Accurate price predictions help investors identify potential risks and opportunities, enabling them to optimize their portfolios and manage investment risks effectively.

1. **Government Policy Formulation:**

Governments use metal price forecasts to formulate economic policies, especially in countries where metal production is a significant contributor to the GDP. This includes planning for export-import policies, taxation, and infrastructure development.

1. **Impact on Inflation and Consumer Goods:**

Fluctuations in metal prices can impact inflation rates. Central banks and policymakers use metal price forecasts to implement measures that control inflation and ensure economic stability.

Many consumer goods, such as appliances and vehicles, incorporate metals in their production. Accurate metal price forecasts help manufacturers set appropriate prices for their products, preventing sudden cost increases that could affect consumer affordability.

1. **Supply Chain Resilience:**

Metal price forecasting assists companies in identifying potential risks in the supply chain, such as shortages or price spikes. This enables them to develop strategies to mitigate risks and ensure a resilient supply chain.

1. **Mining Industry Investment:**

Mining companies use metal price forecasts to plan exploration activities and decide on whether to invest in new extraction projects. These forecasts influence decisions related to opening new mines or expanding existing ones.

1. **Financial Planning for Metal Producers:**

Metal producers need accurate price forecasts to budget effectively and plan their financial strategies. This includes decisions on capital expenditure, operational expenses, and debt management.

1. **Environmental and Social Impact:**

**:** Metal price forecasts can influence the adoption of sustainable practices in mining and metal production. As prices fluctuate, companies may adjust their strategies to align with environmental and social responsibilities.

Time series analysis is a powerful tool in forecast metal price, allowing analysts to identify trends, patterns, and potential future movements based on historical data. Various statistical methods are widely employed in this context to model the relationship between time and metal prices.

With the above context in mind, the primary objective of this paper is to delve into the exploration of five distinct time series forecasting methods designed for predicting metal prices. These methods include:

### 2.6.1 Time Series regression model

This section explores the application of regression models in time series analysis for accurate and insightful metal price forecasting. Linear regression is a statistical technique that models the linear relationship between a dependent variable and one or more independent variables. In the context of time series analysis for metal prices, time is typically the independent variable, and the metal price is the dependent variable. The goal is to establish a linear equation that best represents the historical price movements over time.

The basic concept is that we forecast the time series of metal prices **y** assuming that it has a linear relationship with other parameters ***x*** as predictor [18].

The **forecast variable** *y* is sometimes called the regressand, dependant or explained variable. The **predictor variables** x are sometimes also called the regressors, independent or explanatory variables. In this section, we will always refer to them as “forecast” variable and “predictor” variables.

##### **2.6.1a Simple Linear Regression**

In the simplest cased, the regression model allows for a linear relationship between the forecast variable y and the single predictor variable x:

An artificial example of data from such a model is shown in Fig.1. The coefficients and denote the intercept and the slope of the line respectively. The intercept represents the predicted value of y when x=0. The slope represents the average predicted change in y resulting from a one unit increase in x.

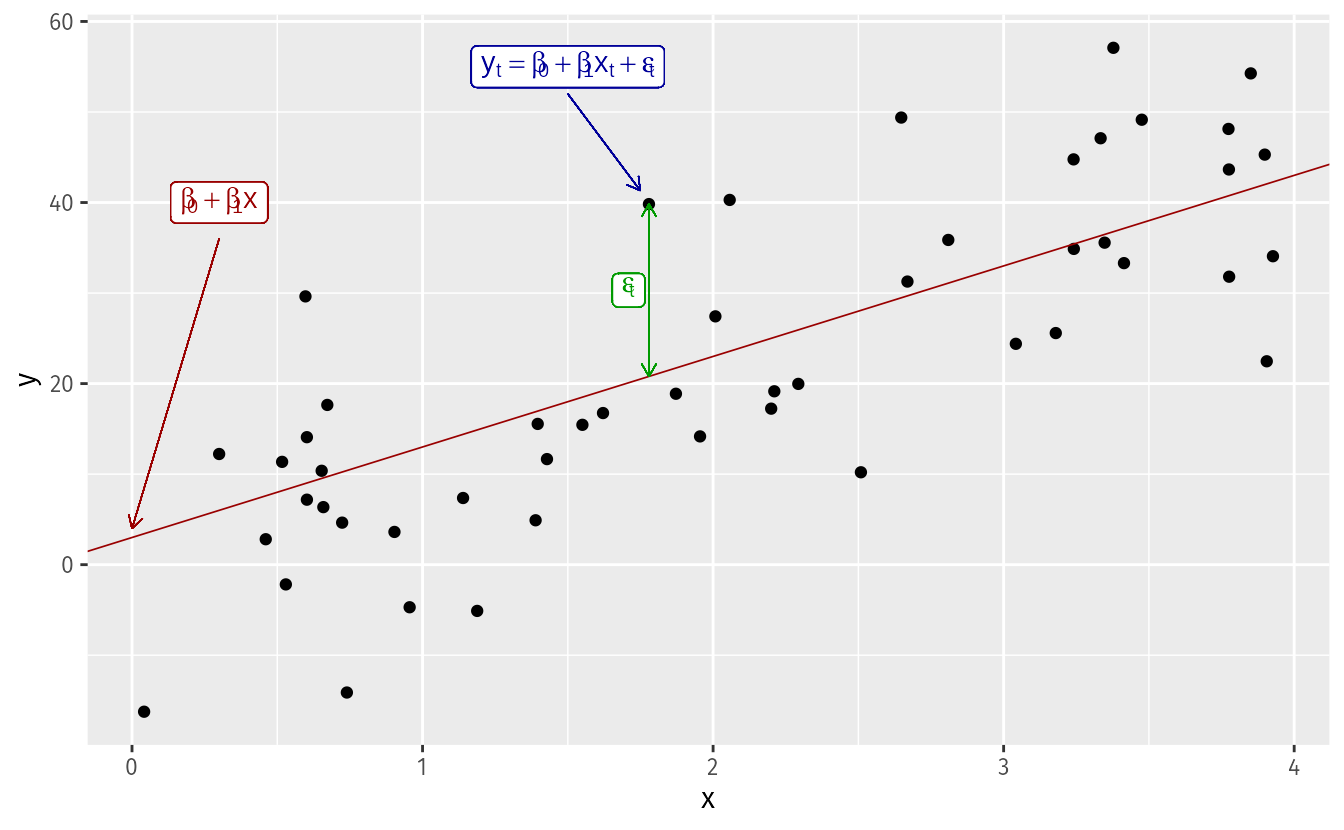


Figure 5: Graphical Representation of Simple Linear Regression

Notice that the observations do not lie on the straight line but are scattered around it. We can think of each observation as consisting of the systematic or explained part of the model, , and the random “error”, . The error term does not imply a mistake, but a deviation from the underlying straight line model. It captures anything that may affect other than .

##### **2.6.1b Multiple Linear Regression**

When there are two or more predictor variables, the model is called a multiple regression model. The general form of a multiple regression model is:

Where is the variable to be forecasted and , , ,…, are the predictor variables. Each of the predictor variables must be numerical. The coefficients ,…,measure the effect of each predictor after taking into account the effects of all the other predictors in the model. Thus, the coefficients measure the marginal effects of the predictor variables.

**Assumptions**

When we use a linear regression model, we are implicitly making some assumptions about the variables.

First, we assume that the model is a reasonable approximation to reality; that is, the relationship between the forecast variables and the predictor variables satisfy this linear equation. Second, we make the following assumptions about the errors (,…,:

* They have mean zero; otherwise the forecasts will be systematically biased.
* They are not autocorrelated; otherwise the forecasts will be inefficient, as there is more information in the data that can be exploited.
* They are unrelated to the predictor variables; otherwise there would be more information that should be included in the systematic part of the model.

It is also useful to have the errors being normally distributed with a constant variance in order to easily produce prediction intervals. Another important assumption in the linear regression model is that each predictor is not a random variable. If we were performing a controlled experiment in a laboratory, we could control the values of each (so they would not be random) and observe the resulting values of . With observational data (including most data in business and economics), it is not possible to control the value of , we simply observe it. Hence, we make this an assumption.

Linear regression is a valuable tool in time series analysis for metal price forecasting. By capturing the linear relationship between time and metal prices, this method provides a foundation for more advanced modelling techniques. While it has its limitations, particularly in capturing complex non-linear trends, when appropriately applied and interpreted, linear regression contributes valuable insights into the temporal dynamics of metal prices. An approach that involves modelling the relationship between metal prices and other relevant variables, facilitating a linear prediction [19].

### 2.6.2 Naïve Forecasting

A straightforward method that relies on past observations without incorporating complex algorithms, offering a baseline for comparison with more sophisticated techniques. Naive Forecasting technique can be used for predicting metal prices using time series analysis, especially as a baseline or benchmark model. The Naive Forecasting technique is a simple method where the prediction for the next time period is based on the observed value from the current time period. There are different variants of Naive Forecasting, such as the Naive Approach and the Seasonal Naive Approach [20].

#### 2.6.2.a Naïve Approach:

The Naïve Approach assumes that the future value of the time series is equal to the last observed value. The mathematical representation is:

Where:

* is the forecasted value for the next time period
* is the observed value at time t.

#### 2.6.2.b Seasonal Naïve Approach:

The seasonal Naïve Approach is an extension of the Naïve Approach, particularly useful when there is a clear seasonality pattern in the data. In this approach, the forecasted value for the next time period is the observed value from the same season in the previous year (or season of a fixed length). The mathematical representation is:

Where:

* is the forecasted value for the next time period
* is the observed value at time t.
* m is the length of the season.

While Naive Forecasting can be a useful starting point, especially for its simplicity and interpretability, it is often advisable to explore more sophisticated time series models like ARIMA, SARIMA, or machine learning techniques for better accuracy and capturing the nuances of metal price movements.

##### **Application to Metal Price Forecasting**

While Naive Forecasting is straightforward and easy to implement, it is important to note that it assumes no underlying pattern or trend in the data other than persistence from the previous observation. This makes it a basic model that serves as a reference point for evaluating the performance of more sophisticated forecasting methods.

#### Advantages:

Simplicity: Naive Forecasting requires minimal computational effort.

Benchmark: It provides a baseline for comparison with more complex models.

#### Limitations:

Ignores patterns and trends: Naive Forecasting does not capture any underlying patterns or trends in the data.

Limited accuracy: It may not perform well when the time series exhibits complexity beyond simple persistence.

### 2.6.3 Moving Average Forecasting

This technique involves calculating the average of a set of past data points, providing a smoothed projection of future metal prices. In the realm of time series analysis for metal price forecasting, moving averages stand out as a widely employed and effective technique. This method, rooted in smoothing historical data, provides a clearer picture of trends and patterns by reducing noise and fluctuations. This section explores the application of moving averages, its variations, and its utility in predicting metal prices over time [21].

It is a statistical calculation used to analyse data points by creating a series of averages of different subsets of the full dataset. In the context of time series analysis for metal prices, a moving average is computed by taking the average of metal prices over a specified time window, or “moving” period. The resulting smoothed line provides a clearer representation of the underlying trend in the data.

Types of Moving Averages:

#### 2.6.3.a Simple Moving Average (SMA):

The SMA is calculated by averaging a set number of past prices. It is easy to compute and provides a smooth curve, making it suitable for capturing long-term trends.

Where :

* is the simple moving average at time t,
* is the metal price at the time t,
* is the number of time periods in the moving average

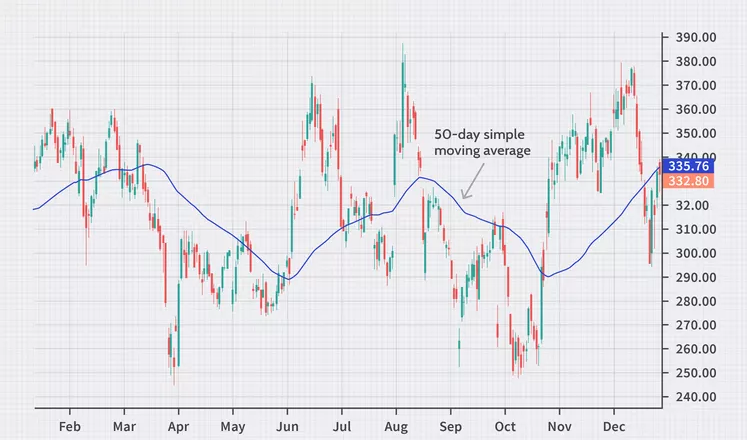


Figure 6: Graphical Representation of Simple Moving Average

#### 2.6.3.b Exponential Moving Average (EMA):

The EMA gives more weight to recent prices, reacting more quickly to changes in the metal price. It is calculated using a smoothing factor (.

EMA is mainly used to identify trends and filter out noise. The weight of elements is decreased gradually over time. This means it gives weights to recent data points, not historical ones. Compared with SMA, the EMA is faster to change and more sensitive.

Where:

* is the exponential moving average at time t,
* is the metal price at time t,
* It has a value between [0,1]
* Represents the weighting applied to the very recent period.



Figure 7: Graphical Representation of Exponential Moving Average

Moving averages offer a valuable tool in the time series analysis toolkit for metal price forecasting. Whether through the simplicity of SMA or the responsiveness of EMA, this technique aids analysts in discerning trends and making informed predictions about the future direction of metal prices. While recognizing its limitations, particularly in terms of lag, the judicious application of moving averages enhances the ability to navigate the complexities of the dynamic metal market.

##### Special Considerations:

##### Analytical Significance:

Moving averages are an important analytical tool to identify current price trends and the potential for a change in an established trend. The simplest use of an SMA in technical analysis is using it to quickly determine if an asset is in an uptrend or downtrend.

##### SMA vs. EMA :

The major difference between an exponential moving average (EMA) and a simple moving average is the sensitivity each one shows to changes in the data used in its calculation. More specifically, the EMA gives a higher weighting to recent prices, while the SMA assigns an equal weighting to all values.

The two averages are similar because they are interpreted in the same manner and are both commonly used by technical traders to smooth out price fluctuations. Since EMAs place a higher weighting on recent data than on older data, they are more reactive to the latest price changes than SMAs are, which makes the results from EMAs more timely and explains why the EMA is the preferred average among many forecasters.

### 2.6.4 LSTM Method

Leveraging Long Short-Term Memory networks, a type of recurrent neural network, this approach is designed to capture complex patterns and dependencies within metal price time series data [22].

Each of these forecasting techniques is thoroughly elucidated in the methodology section, outlining their operational principles and distinctive features. This comprehensive investigation aims to equip mining companies with the analytical tools necessary for informed decision-making, offering insights into the potential trajectories of metal prices and aiding in strategic planning for exploration and mining activities.

#### Basic Structure of RNN

Recurrent Neural Networks (RNNs) are a class of artificial neural networks designed for processing sequential data. Unlike traditional feedforward neural networks, RNNs have connections that form directed cycles, allowing them to maintain a hidden state or memory of previous inputs. This unique architecture enables RNNs to capture and learn dependencies in sequential data, making them well-suited for time series forecasting, including metal price prediction [23].

The basic structure of an RNN involves three main components:

Input Layer: Receives input sequences at each time step.

Hidden Layer: Maintains a hidden state that captures information from previous time steps. The hidden state is updated at each time step based on the input and the previous hidden state.

Output Layer: Produces the output for the current time step based on the information in the hidden state.

The recurrent connections in the hidden layer allow RNNs to consider the entire sequence of inputs, enabling them to capture long-term dependencies and temporal patterns in the data [24].

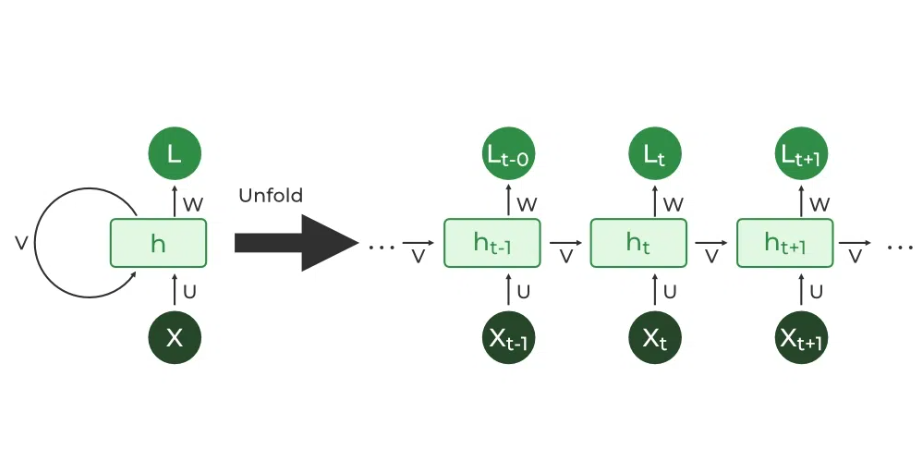


Figure 8: Illustration of recurrent connections in RNNs

#### Mathematics behind RNN Architecture

To understand the mathematics behind Recurrent Neural Networks (RNNs) used for metal price forecasting, let's break down the fundamental equations involved in the forward pass and backpropagation through time [25].

The hidden state at each time step is computed as a function of the input and the previous hidden state :

Where:

* is the hidden state at time t.
* is the input at time t.
* is the weight matrix for the recurrent connections.
* is the weight matrix for the input connections.
* is the bias term.
* is the activation function, often the hyperbolic tangent (tanh) or rectified linear unit (ReLU).

The output at each time step is then computed based on the hidden state:

Where:

* is the weight matrix for the output connections.
* is the bias term.

1. Backpropagation Through Time (BPTT):

The objective is to minimize the error between the predicted output and the actual target . The loss function at each time step is often the mean squared error:

The total loss is the sum of the losses across all time steps:

To update the weights during training, gradients with respect to the parameters are calculated through backpropagation through time. The gradients for the output layer are computed as:

Similarly, gradients for the hidden layer and input layer are computed:

The update rules for the weights are then applied using an optimization algorithm such as stochastic gradient descent (SGD).

1. Challenges and Enhancements:

Despite their ability to capture sequential dependencies, traditional RNNs may suffer from the vanishing gradient problem, hindering their performance in long term dependency tasks. This limitation has led to the development of more advanced architectures such as Long Short-Term Memory (LSTM) networks, which introduce memory cells and gating mechanism to better handle long-range dependencies [26].

In the next section, we can explore how LSTM networks address these challenges and provide an effective solution for metal price forecasting.

##### **Challenges of Vanishing and Exploding Gradients in Traditional RNN**

While RNNs offer a powerful mechanism for capturing sequential dependencies, they are susceptible to two main challenges during training: vanishing gradients and exploding gradients.

Vanishing Gradients:

In the training process, when backpropagating errors through time, gradients can become extremely small, leading to the vanishing gradient problem.

This occurs because the gradients are multiplied during each time step, and if the weights are not properly adjusted, they may approach zero, hindering the learning of long-term dependencies.

For metal price forecasting, where historical trends may be critical, the vanishing gradient problem can impede the model's ability to capture and learn from distant past information.

Exploding Gradients:

Conversely, exploding gradients occur when the gradients become excessively large during training.

This can lead to numerical instability, causing the weights to grow exponentially and making the model difficult to train.

The exploding gradient problem can disrupt the learning process, especially when dealing with large and complex RNN architectures [27].

##### **Mitigating Strategies**

Researchers have developed strategies to address these challenges in traditional RNNs. One common approach is the use of gated recurrent units (GRUs) and Long Short-Term Memory (LSTM) networks, which are specialized RNN architectures designed to better handle long-range dependencies and mitigate the vanishing gradient problem.

In the next section, we will delve into LSTM networks, a particularly effective solution for capturing sequential dependencies in time series data, including metal prices [28].

##### **Basic Structure of LSTM**

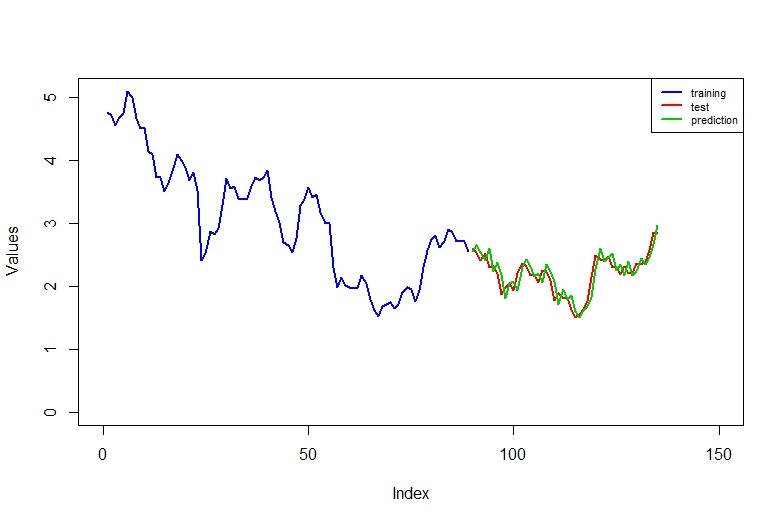
LSTM networks represent an advanced form of Recurrent Neural Network (RNNs) specifically designed to address the vanishing gradient problem encountered in traditional RNNs. In the context of forecasting metal prices, where capturing both short-term fluctuations and long-term trends is crucial , 

Figure 9: Graphical Representation of trained ,tested and predicted values for LSTM network.

##### **Architecture of LSTM**

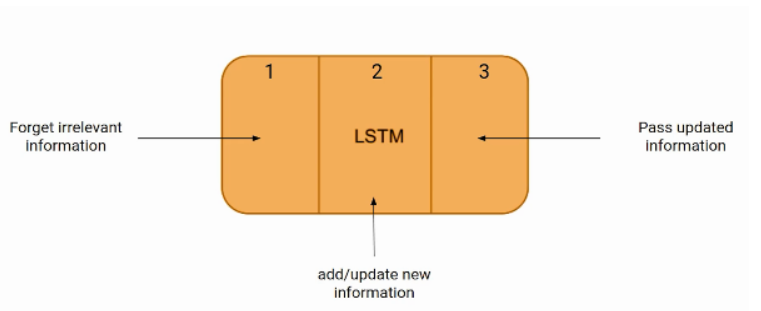


Figure 10: Representation of LSTM architecture

* Memory Cell :

The memory cell in an LSTM allows the network to maintain a long-term memory state. In the context of metal price forecasting, this long-term memory is essential for capturing trends and patterns that unfold over extended periods.

* Input Gate (:

The input gate determines how much of the new information should be stored in the memory cell. It considers the current input and the previous hidden state . In metal price forecasting, this gate enables the model to decide the relevance of the latest market information for predicting future prices.

* Forget Gate (:

The forget gate decides what information from the previous memory cell state (​) should be discarded. It, too, considers the current input ) and the previous hidden state ). For metal prices, this gate allows the model to selectively forget or remember past market conditions based on their significance.

* Cell State Update:

The memory cell state is updated based on the input gate, forget gate, and a candidate new state . The candidate new state is determined by the hyperbolic tangent activation function (tanh) applied to a combination of the current input and the previous hidden state .

* Output Gate (

The output gate determines the next hidden state and the output of the LSTM cell. It considers the current input( and the previous hidden state . The new hidden state is then determined by multiplying the output gate with the hyperbolic tangent of the current memory cell state .

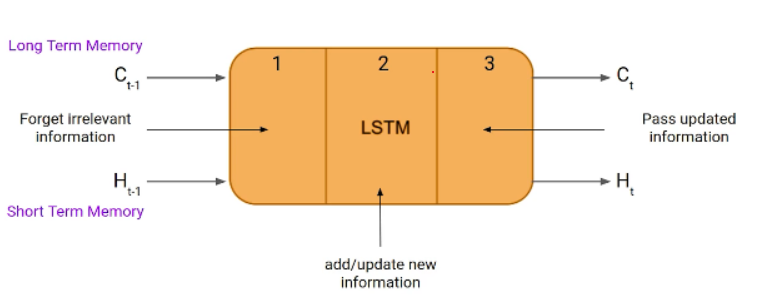


Figure 11: Long term and short-term memory representation

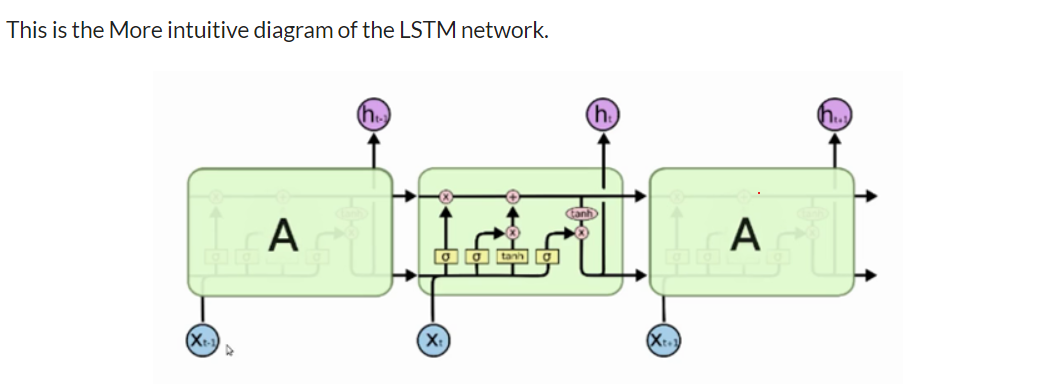


Figure 12: A more intuitive representation of LSTM network

##### **Application in metal Price Forecasting**

In the context of forecasting metal prices, the LSTM architecture allows the model to capture both short-term fluctuations and long-term trends. The memory cell retains important information over extended periods, enabling the network to learn and adapt to the dynamic nature of metal markets. The gates provide the flexibility to selectively store or discard information, ensuring that the model focuses on the most relevant aspects of the input sequence.

By incorporating LSTM networks into time series analysis for metal price forecasting, the model gains the ability to effectively learn and remember patterns over various time scales, contributing to more accurate and adaptive predictions [29].

## 2.7 Technological Innovation

The metal mining sector, a cornerstone of global industrialization, has historically been viewed as a traditional and conservative industry. However, the tide is turning as technological innovation sweeps through its operations, revolutionizing everything from exploration and extraction to processing and environmental management. This essay explores the profound impact of technological innovation on the metal mining sector, shedding light on how advancements are reshaping practices, enhancing efficiency, and addressing environmental concerns.

One of the primary areas witnessing transformation is exploration. Traditional prospecting methods are being replaced by sophisticated technologies such as remote sensing, geospatial analytics, and artificial intelligence. These innovations enable more precise identification of mineral deposits, reducing exploration costs and minimizing environmental impact. Drones and autonomous vehicles equipped with advanced sensors have made it possible to survey vast and challenging terrains, providing a more comprehensive understanding of potential mining sites.

Technological advancements in extraction methods have significantly increased efficiency and safety in mining operations. Automated machinery and robotics have replaced manual labor in hazardous environments, reducing the risk to human life. Intelligent drilling and blasting technologies, guided by real-time data analytics, optimize ore recovery while minimizing waste. Moreover, sensor-equipped equipment enables continuous monitoring of operational parameters, enhancing predictive maintenance and minimizing downtime.

In the processing stage, innovation has led to the development of more energy-efficient and environmentally sustainable methods. Advanced sorting technologies, such as sensor-based ore sorting, allow for the selective extraction of valuable minerals, reducing the volume of material that needs to be processed. Breakthroughs in hydrometallurgy and bioleaching offer eco-friendly alternatives to traditional extraction methods, addressing concerns about the environmental impact of chemical processes.

Technological innovation has also become a crucial ally in addressing environmental challenges associated with mining activities. Advanced monitoring systems track air and water quality, enabling rapid responses to mitigate potential environmental harm. Reclamation efforts are benefitting from the use of drones and satellite imagery, facilitating the restoration of mined areas to their natural state. Additionally, the integration of renewable energy sources into mining operations is reducing the industry's carbon footprint.

Despite the positive impact of technological innovation, the mining sector faces challenges in adopting these advancements universally. High initial investment costs, the need for skilled personnel to operate new technologies, and the potential displacement of traditional jobs are among the hurdles. Moreover, cybersecurity threats pose risks to data-driven mining systems, necessitating robust cybersecurity measures.

The metal mining sector is undergoing a profound transformation propelled by technological innovation. From exploration to environmental management, advancements are driving efficiency, sustainability, and safety. While challenges persist, the industry's embrace of innovation heralds a future where metal mining is not only economically viable but also environmentally responsible and socially sustainable. As technology continues to evolve, the metal mining sector stands on the precipice of a new era, characterized by smarter, safer, and more sustainable practices.

The mining industry's disposition towards innovation has been a subject of debate, often characterized as conservative by critics or heralded as sophisticated by advocates. By examining productivity statistics, a parallel can be drawn between the innovation rates of metal mining companies and those associated with general manufacturing over the past five decades. Notably, the mining and general manufacturing sectors have exhibited comparable annual rates of productivity increases, standing at 2.3% and 2.5%, respectively. However, a noteworthy deviation arises in the high-tech manufacturing sector, which has maintained a substantially higher productivity increase of 9.5% per year for the past approximately twelve years.

In terms of groundbreaking technological advancements, the mining industry has witnessed a frequency of one to three revolutionary technologies per century, aligning with the innovation rates of other mature industries like cement and glass-making. This rate sharply contrasts with the micro-computing sector, which has experienced a revolutionary technology development rate four times higher. In the realm of innovation, mining seems to share more similarities with well-established manufacturing industries than with the high-tech sector.

Despite this, investment in basic research and development (R&D) within the mining industry remains relatively low compared to other sectors and shows signs of decreasing. Presently, much of the industry's productivity strides stem from external sources, including equipment manufacturers and suppliers. As the mining sector and government programs witness declines in basic R&D investments, the sustainability of the industry's productivity advancements at comparable rates becomes a looming question.

## 2.8 Socio-Economic Dynamics

The metal mining sector plays a pivotal role in the global economy, contributing significantly to industrial development and infrastructure growth. However, the socio-economic conditions within this sector are complex and multifaceted, encompassing both opportunities for economic advancement and challenges related to labor, community well-being, and sustainable development. This essay explores the intricate socio-economic dynamics in the metal mining sector, examining the impact on local communities, labor forces, and broader economic landscapes.

### Labor Conditions and Employment Opportunities:

The metal mining sector, known for its labor-intensive nature, provides employment opportunities to millions worldwide. However, the socio-economic conditions of labor in this sector are marked by challenges such as hazardous working environments, job insecurity, and sometimes inadequate wages. The need for skilled labor in modern mining, driven by technological advancements, poses a challenge for traditional mining communities where the skillsets required may not align with existing capabilities. Balancing the creation of quality jobs with the adoption of advanced technologies becomes crucial for ensuring positive socio-economic outcomes.

### Community Impacts and Social Development:

The presence of mining activities can have profound impacts on local communities, influencing their socio-economic conditions. While mining projects can inject capital into local economies, fostering infrastructural development and creating business opportunities, they also bring challenges. Environmental degradation, disruption of traditional livelihoods, and social inequalities can arise, affecting the well-being of communities. Effective community engagement, transparent revenue-sharing mechanisms, and sustainable development initiatives are essential for mitigating negative social impacts and promoting inclusive growth.

### Resource Dependency and Economic Diversification:

Countries heavily reliant on metal mining often face challenges related to economic diversification. The susceptibility of such economies to fluctuations in metal prices and global demand can lead to economic volatility. Governments and industry stakeholders must strategize to diversify their economies, investing in sectors beyond mining to reduce vulnerability and promote long-term socio-economic stability. This shift requires proactive policies that encourage innovation, entrepreneurship, and the development of alternative industries.

### Environmental Sustainability and Social Responsibility:

As concerns about environmental sustainability grow, the metal mining sector faces increasing pressure to adopt responsible practices. Sustainable mining practices not only contribute to environmental conservation but also have socio-economic implications. Community health, access to clean water, and the protection of indigenous rights become central issues. Mining companies that prioritize social responsibility, engage in transparent communication, and invest in community development initiatives are more likely to foster positive socio-economic conditions.

### Inclusive Growth and Stakeholder Collaboration:

To address the socio-economic challenges in the metal mining sector, a collaborative approach involving governments, industry players, local communities, and non-governmental organizations is imperative. Inclusive growth strategies should prioritize the empowerment of local communities through education, skills development, and entrepreneurship initiatives. Engaging stakeholders in decision-making processes and recognizing the rights of indigenous communities are crucial steps towards fostering socio-economic conditions that benefit all.

The metal mining sector stands at a crossroads where socio-economic conditions are deeply intertwined with global economic trends, technological advancements, and environmental imperatives. While challenges persist, there are opportunities to shape the sector in a way that promotes inclusive growth, environmental sustainability, and social responsibility. By acknowledging the complexities of the socio-economic landscape, stakeholders can work collaboratively to navigate these challenges and unlock the full potential of the metal mining sector for the benefit of communities, labour forces, and economies at large.

## 2.9 Emergence of Renewable energy

the global transition to renewable energy was well underway, with numerous countries and industries making significant strides.

From 2010 to 2019, the global investment in renewable energy capacity has soared to $2.6 trillion, a staggering threefold increase from the previous decade. The solar sector has been the primary beneficiary, attracting a substantial $1.3 trillion, while wind secured $1 trillion, and biomass and waste-to-energy collectively garnered $115 billion.

China has emerged as the undisputed leader in renewable energy investments during this period, committing an impressive $758 billion. The United States follows at a distance with $356 billion, and Japan takes the third spot with $202 billion. In Europe, a collective investment of $698 billion was made, with Germany leading at $179 billion and the U.K. contributing $122 billion. India has also become a notable player, having committed $90 billion by the midpoint of 2019.

A remarkable aspect of this decade has been the substantial improvement in the cost-competitiveness of renewable technologies. The levelized cost of electricity for solar photovoltaics plummeted by 81%, onshore wind by 46%, and offshore wind by 44%. Consequently, renewable energy has become the most economical option for new power generation in many countries globally.

The driving forces behind these cost reductions are a synergy of factors, including economies of scale in manufacturing, intense competition along the supply chain (fueled by auctions in many nations), historically low costs of finance, and enhancements in the efficiency of generating equipment.

Solar capacity installations have outpaced all other generating technologies, reaching an impressive 638GW during 2010-2019, a remarkable feat considering that global solar power capacity was only 25GW at the end of 2009. The overall power capacity installed during this decade is projected to reach a net 2.4 terawatts, with solar leading, followed by coal, and wind narrowly surpassing gas for third place. However, the continued reliance on fossil fuels has led to an estimated increase of at least 10% in global power sector emissions between 2009 and 2019.

In 2018, global investment in renewable energy capacity reached $272.9 billion, marking the fifth consecutive year surpassing $250 billion. Although down 12% from 2017, this figure was achieved despite falling capital costs of solar and wind projects. Solar retained its dominance, attracting $133.5 billion, a 22% decrease from 2017, while wind secured $129.7 billion, reflecting a 3% increase.

Renewable energy capacity investment was more widespread globally than ever before, with 29 countries each investing over $1 billion in 2018. China, despite a 38% decrease, led with $88.5 billion, followed by Europe at $59.9 billion (a 45% increase), and the U.S. at $42.8 billion (a 6% decrease).

Several countries, including Spain, Vietnam, Ukraine, and South Africa, witnessed a more than fivefold increase in capacity investment in 2018. The world added a record 167GW of new renewable capacity, contributing to a share of 12.9% in global electricity generation in 2018, up from 11.6% in 2017. This progress helped avoid an estimated 2 gigatonnes of carbon dioxide emissions.

In addition to capacity investment, other forms of investment in renewables experienced increases in 2018. Government and corporate research and development rose by 10% to $13.1 billion, equity raising by specialist companies on public markets increased by 6% to $6 billion, and venture capital and private equity investment surged by 35% to $2 billion. Despite an 11% decline in overall renewable energy investment, including these categories, it amounted to $288.3 billion in 2018.

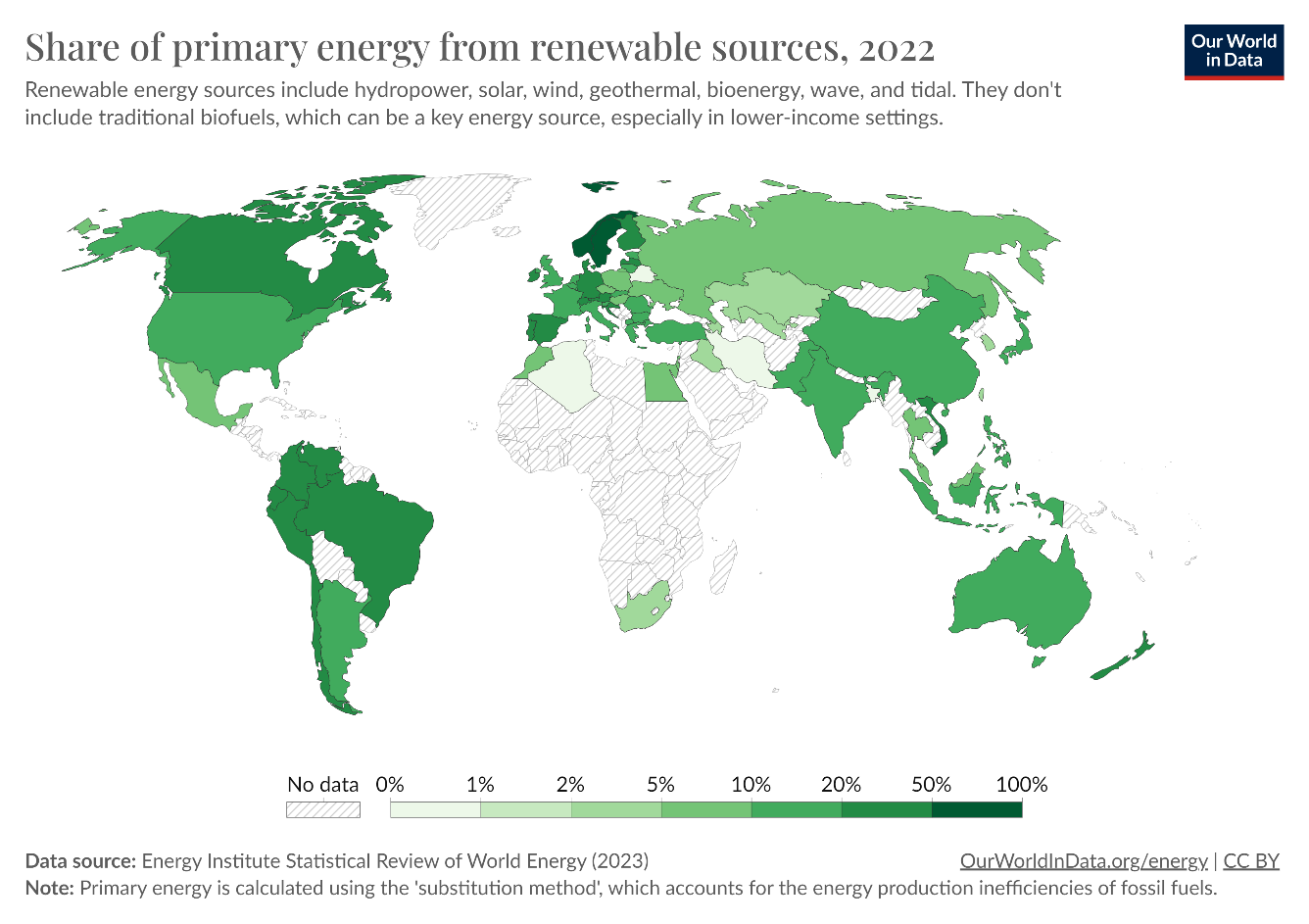


Figure 13: Current Global share of renewable energy

**Increasing Renewable Energy Capacity:**

Many countries have been investing in and expanding their renewable energy capacity, particularly in solar and wind power. Governments and businesses are increasingly adopting renewable energy sources to meet electricity demand and reduce reliance on fossil fuels.

**Policy Support:**

Government policies and incentives have played a crucial role in promoting the adoption of renewable energy. Incentives such as subsidies, tax credits, and feed-in tariffs have been used to encourage investment in renewable technologies.

**Technological Advances:**

Ongoing technological advancements in solar, wind, and energy storage technologies have led to increased efficiency and cost reductions. Innovations in battery storage have been particularly significant for enabling the integration of intermittent renewable sources into the energy grid.

**Decentralization and Distributed Energy Resources:**

The transition to renewable energy is characterized by a shift toward decentralized energy systems. Distributed energy resources, including rooftop solar panels and community-based renewable projects, are becoming more common, contributing to a more resilient and flexible energy infrastructure.

**Electric Vehicles and Electrification:**

The electrification of transportation, including the widespread adoption of electric vehicles (EVs), has become a key aspect of the transition to renewable energy. Governments and automakers are investing in EV infrastructure and incentivizing consumers to make the switch from traditional combustion engine vehicles.

**International Commitments:**

Various international agreements and commitments, such as the Paris Agreement, have spurred countries to set ambitious targets for reducing greenhouse gas emissions and increasing the share of renewable energy in their energy mix.

**Challenges and Barriers:**

Despite progress, challenges remain. These include the intermittency of renewable sources, the need for energy storage solutions, the phasing out of existing fossil fuel infrastructure, and the economic considerations associated with the transition.

**Corporate Sustainability Initiatives**

Many corporations are incorporating renewable energy into their sustainability strategies. This includes commitments to using renewable energy for their operations, investing in renewable projects, and setting carbon reduction targets.

**Public Awareness and Support:**

Public awareness of environmental issues and the importance of transitioning to renewable energy has increased. This has led to greater support for policies and initiatives that promote sustainability and combat climate change.

**Chapter 3 - Methodology**

# **Methodology**

## 3.1 Data Collection

Sources of Historical Metal Price Data:

Collecting accurate and relevant historical metal price data is crucial for building a robust forecasting model. The sources chosen should provide comprehensive and reliable information on metal prices over time. Common sources for historical metal price data include:

1. Financial Markets Data Providers: Utilize financial market data platforms and providers that offer historical price data for various metals. Examples include Bloomberg, Reuters, and Yahoo Finance.
2. Commodity Exchanges: Access data from major commodity exchanges such as the London Metal Exchange (LME) or the Chicago Mercantile Exchange (CME), where metal prices are traded.
3. Government Agencies: Explore datasets provided by government agencies that track and publish economic indicators, including metal prices. National statistical offices or the U.S. Geological Survey are potential sources.
4. Industry Reports and Publications: Refer to industry reports, publications, and research papers that provide historical metal price trends. Industry associations and research organizations may offer valuable datasets.
5. Open Data Platforms: Explore open data platforms that provide free access to historical datasets. Websites like Quandl, Kaggle, or data.gov may offer relevant datasets on metal prices.

Selection Criteria for the Dataset:

Following are the common selection criteria for choosing the dataset:

1. Comprehensive Coverage: Choosing a dataset that covers a comprehensive range of metals, including but not limited to gold, silver, copper, aluminium, and others. A diverse set of metals allows the model to capture broader market dynamics.
2. Granularity of Data: Opting for datasets with a suitable level of granularity, such as daily, hourly, or minute-by-minute prices. The granularity should align with the forecast horizon and the desired level of detail for analysis.
3. Historical Time Period: Ensuring the dataset spans a sufficiently long historical time period. This allows the model to capture different market conditions, economic cycles, and trends over time.
4. Data Quality and Consistency: Prioritizing datasets with high data quality and consistency. Checking for missing values, outliers, and inconsistencies in the data, and apply appropriate preprocessing techniques to address any issues.
5. Inclusion of Relevant Features: In addition to metal prices, considering datasets that include relevant features such as trading volumes, economic indicators, geopolitical events, or other factors that may impact metal prices. This enhances the model's ability to capture multifactorial influences
6. Currency and Units: Ensuring consistency in currency and units across the dataset. This is crucial for accurate analysis and model training. If prices are in multiple currencies, consider converting them to a standardized currency.
7. Accessibility and Updates: Choosing datasets that are easily accessible and, if possible, regularly updated. Timely updates ensure that the model is trained on the most recent information, enhancing its forecasting accuracy.

By adhering to these selection criteria, the historical metal price dataset becomes a reliable foundation for training and evaluating time series forecasting models. It enables the development of a model that can capture the complexities and nuances of metal markets over time.

## 3.2 Data Preprocessing

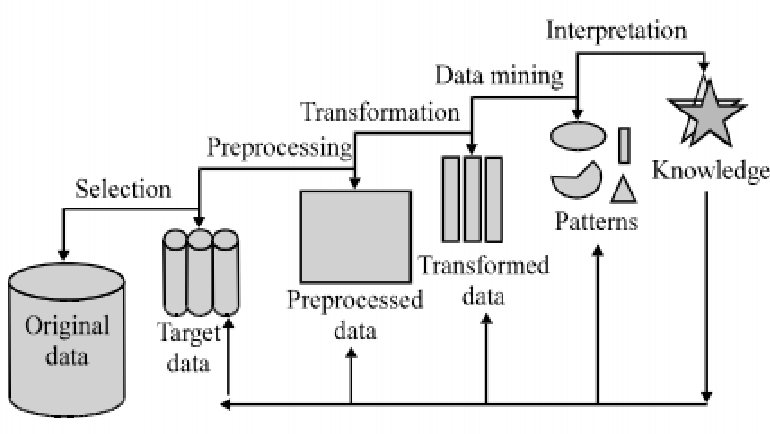


Figure 14: Representation of steps involved in data pre processing

Data preprocessing is a crucial step in preparing historical metal price data for time series analysis. This stage involves cleaning and transforming the raw data to ensure its suitability for training and evaluating forecasting models.

Steps Taken in Data Preprocessing:

1. Handling Missing Values: Identify and handle missing values in the dataset. This may involve imputation techniques such as forward filling, backward filling, or using interpolation to estimate missing values based on neighbouring observations.

# importing pandas package

import pandas as pd

# making data frame from csv file

data = pd.read\_csv("sample.csv")

# creating bool series True for NaN values

bool\_series = pd.isnull(data["Column"])

# filtering data

# displaying data only with Gender = NaN

data[bool\_series]

1. Outlier Detection and Removal: Identify outliers in the metal price data that may disrupt the analysis. Outliers can be detected using statistical methods or visualization techniques, and appropriate measures such as trimming or imputation can be applied to address their impact.
2. Data Transformation: Convert the time series data into a suitable format for analysis. Ensure that the data is in a chronological order and, if needed, resample it to the desired frequency (e.g., daily to monthly) to match the forecast horizon.
3. Handling Non-stationarity: Check for non-stationarity in the time series data. If trends or seasonality are present, apply differencing or other transformation techniques to make the data stationary. This is essential for certain time series models like ARIMA.
4. Feature Engineering: Augment the dataset with additional relevant features that may impact metal prices, such as economic indicators, geopolitical events, or external factors. Ensure that these features are synchronized with the time series data.
5. Normalization or Standardization: Normalize or standardize the data if the values are on different scales. This step ensures that all features contribute proportionately to the model and prevents certain features from dominating others during training.
6. Handling Categorical Variables: If the dataset includes categorical variables (e.g., metal type), encode them appropriately for model compatibility. Techniques such as one-hot encoding can be applied to convert categorical variables into numerical representations.
7. Train-Test Split: Split the dataset into training and testing sets. The training set is used to train the model, while the testing set is reserved for evaluating the model's performance on unseen data.

## 3.3 Model Selection

Choosing the right model for metal price forecasting depends on the characteristics of your data and the nature of the task. Here are several models commonly used for time series forecasting, including metal prices:

1. **ARIMA (AutoRegressive Integrated Moving Average)**

Effective for capturing linear trends and seasonality in time series data.

May not handle complex relationships or nonlinear patterns well.

1. **LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit):**

Well-suited for capturing long-term dependencies and nonlinear patterns in sequential data.

Computationally expensive and may require more data for training.

1. **Prophet**

Developed for time series forecasting, handles missing data and outliers well, and incorporates seasonality and holiday effects.

May not capture complex relationships as effectively as some other models.

1. **XGBoost (eXtreme Gradient Boosting):**

A powerful ensemble model that can handle complex relationships, nonlinear patterns, and feature interactions.

May require more tuning of hyperparameters, and its interpretability may be lower than simpler models.

1. **Random Forest:**

Robust to overfitting, handles nonlinear relationships well, and provides feature importance.

May require careful tuning, and its predictions may be less interpretable than simpler models.

1. **SARIMA (Seasonal AutoRegressive Integrated Moving Average):**

An extension of ARIMA that accounts for seasonality in time series data.

May be sensitive to the choice of parameters and may not handle complex patterns as effectively as other models.

1. **Neural Networks (Feedforward or Deep Learning Models):**

Can capture complex, nonlinear relationships in data, and handle large amounts of data.

Computationally expensive, may require extensive tuning, and can be prone to overfitting with limited data.

**Model Selection Considerations:**

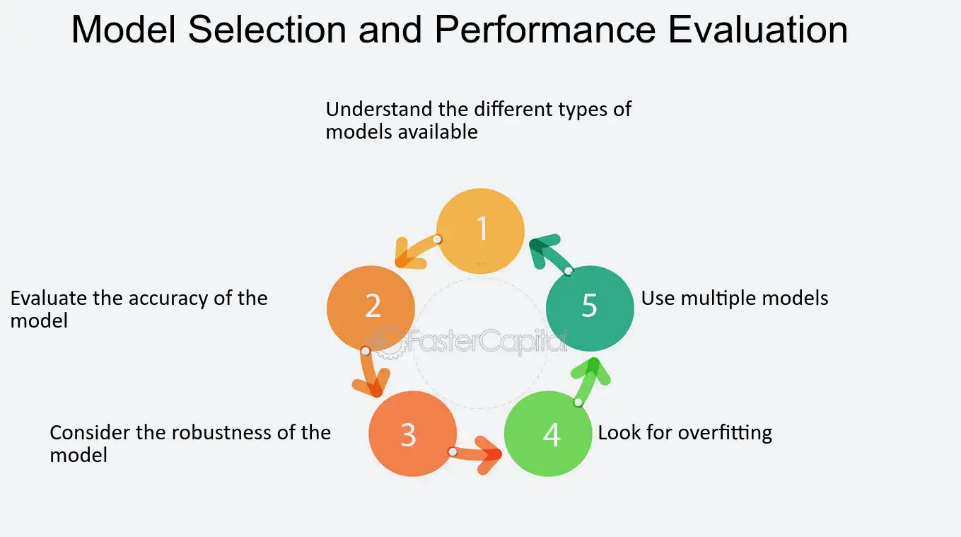
****

Figure 15: Representation of model selection considerations

* **Data Characteristics:** Consider the characteristics of your metal price data, such as seasonality, trends, and potential nonlinearity.
* **Amount of Data:** Some models, particularly deep learning models, may require a large amount of data to perform well.
* **Interpretability:** Consider whether interpretability is crucial. Simpler models like ARIMA or SARIMA may offer more straightforward interpretation of results.
* **Computational Resources:** Some models, particularly deep learning models, may require significant computational resources for training.
* **Ensemble Approaches:** Combining predictions from multiple models using ensemble methods (e.g., stacking or bagging) can sometimes lead to better overall performance.
* **Hyperparameter Tuning:** Be prepared to fine-tune hyperparameters for the selected model to achieve optimal performance.

**Chapter 4 – Data Analysis and Modelling**

# **Data Analysis and Modelling**

## 4.1 Data pre-processing

### 4.1.1. Importing all datasets

We import datasets of all the metals we intend on analysing. Namely gold, copper and silver.

data\_copper=pd.read\_csv("lme\_copper.csv")

data\_silver=pd.read\_excel("comex\_silver.xlsx")

data\_gold=pd.read\_excel("comex\_gold.xlsx")

Each of the datasets have been collected for a duration of 5 years (Feb 2020- Feb 2024)

### 4.1.2 Copper price trends

We perform basic exploratory analysis on the copper dataset to obtain the variations in copper price trends over years.

# Plot the histogram

data\_copper.plot(x="Date", y="Price", color="r",figsize=(20, 6))

data\_copper.plot(x="Date", y="Open", color="b",figsize=(20, 6))

data\_copper.plot(x="Date", y="High", color="g",figsize=(20, 6))

#figsize=(10, 6)

# Set labels and title

plt.xlabel("Date")

plt.ylabel("Price")

plt.title("Copper Price Trend")

# Show the plot

plt.show()

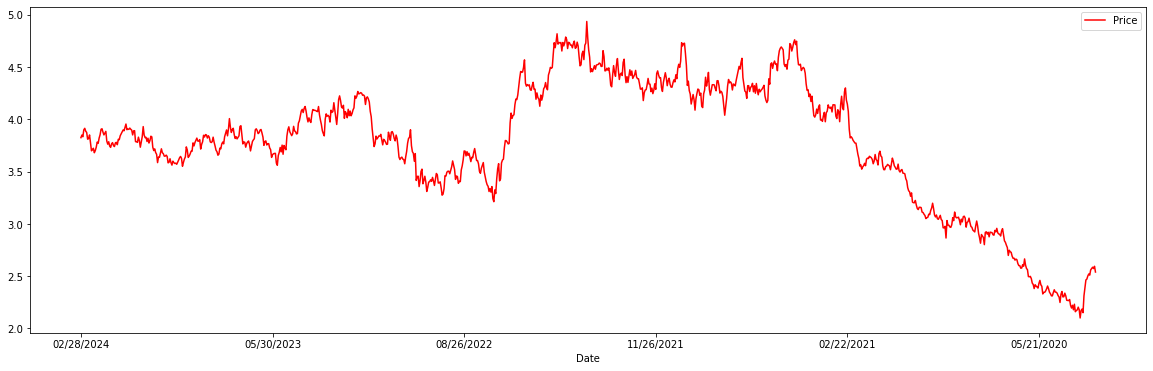


Figure 16: Graphical representation of copper prices between 2020-2024

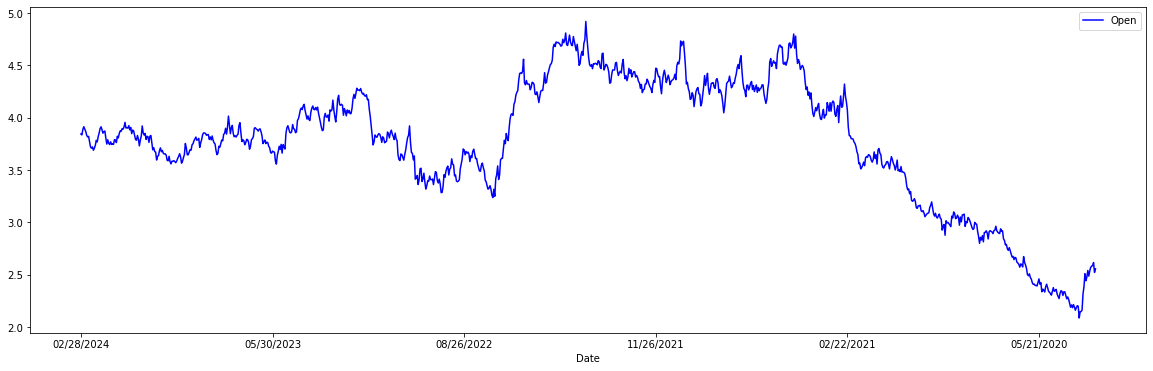


Figure 17: Graphical Representation of copper opening prices between 2020-2024

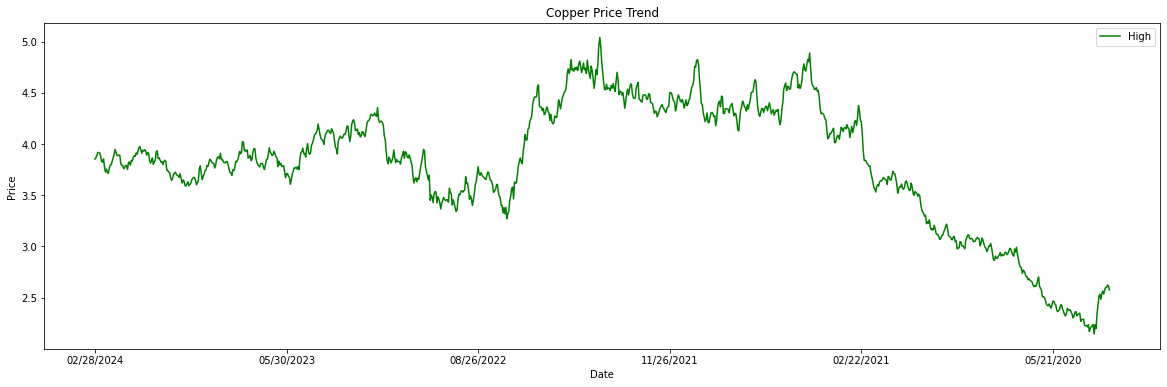


Figure 18: Graphical representation of copper high prices between 2020-2024

**Interpretation of underlying copper trends through above visualization**

The provided visualization illustrates the trend of copper prices over time, represented by three key metrics: price, opening price, and high price. Each line plot displays the respective metric against dates, allowing for a comparative analysis of how copper prices have evolved over the specified time period.

**Key Observations**

* Price Dynamics: The red line plot represents the overall trend of copper prices over time. By examining this plot, we can observe fluctuations and trends in the market value of copper. Notably, [mention any observed trends, such as increasing, decreasing, or volatile price movements].
* Opening Prices: The blue line plot depicts the opening prices of copper for each corresponding date. This metric provides insights into the initial trading values at the beginning of each trading session.
* High Prices: The green line plot illustrates the highest recorded prices of copper during the specified time period. This metric helps identify peak price levels reached during the observation period.

**Implications and Insights**

* Market Volatility: Fluctuations in copper prices, as evidenced by the red line plot, reflect the inherent volatility of the commodity market. Factors such as supply and demand dynamics, economic indicators, geopolitical events, and global trade policies can influence price movements.
* Opening vs. High Prices: Comparing opening prices (blue) with high prices (green) provides insights into intraday price movements and the potential range of price fluctuations within each trading session.
* Trend Analysis: Analyzing the trend of copper prices over time enables stakeholders to assess market performance, identify potential patterns or cycles, and make informed decisions regarding investment strategies, risk management, and procurement decisions.

In conclusion, the visualization of copper price trends offers valuable insights into the dynamics of the copper market, highlighting fluctuations, trends, and intraday price movements. By leveraging this information, stakeholders can enhance their understanding of market dynamics, mitigate risks, and capitalize on opportunities in the copper market.

**Feature Engineering**

As part of the data preprocessing phase, feature engineering plays a crucial role in preparing the dataset for analysis and modeling. One common preprocessing task involves handling non-numeric characters within numeric columns, ensuring consistency and compatibility for further analysis. In this report, we discuss the application of feature engineering techniques to address non-numeric characters within the "Vol." (Volume) column of the copper dataset.

Background: The copper dataset contains a column labeled "Vol." representing the volume of copper transactions. Upon inspection, it was discovered that the volume values contain non-numeric characters, particularly the letter 'K', denoting thousands. For instance, volume values might appear as "1.2K" indicating 1.2 thousand units of copper traded.

To address the presence of non-numeric characters and convert the volume values to numerical format, the following feature engineering approach was applied

1. Remove non-numeric characters and convert to numerical format

data\_copper['Vol.'] = data\_copper['Vol.'].str.replace('K', '')

This code snippet utilizes the .str.replace() method in pandas to remove the letter 'K' from volume values, effectively eliminating the non-numeric character. Subsequently, the cleaned volume values are converted to numerical format, enabling quantitative analysis and modeling tasks.

By applying the feature engineering technique outlined above, the "Vol." column now contains numeric values representing the volume of copper transactions without non-numeric characters. This enhances the consistency and integrity of the dataset, facilitating accurate analysis and interpretation of volume-related metrics.

In conclusion, Feature engineering, including the handling of non-numeric characters, is essential for preparing datasets for analysis and modeling tasks. The application of appropriate techniques ensures data consistency, quality, and suitability for downstream analytical processes, ultimately contributing to more robust insights and decision-making.

import matplotlib.pyplot as plt

# Set the default figure size

plt.figure(figsize=(10, 6)) # Adjust the width and height as needed

# Plot histogram

data\_copper["Vol."].plot(kind="hist", color="g", bins=20) # Adjust the number of bins as needed

plt.xlabel("Volume")

plt.ylabel("Frequency")

plt.title("Volume Histogram")

# Show the plot

plt.show()

data\_copper.plot(x="Date", y="Vol.", color="g", figsize=(10, 6))

plt.xlabel("Date")

plt.ylabel("Volume")

plt.title("Volume vs. Date")

plt.show()

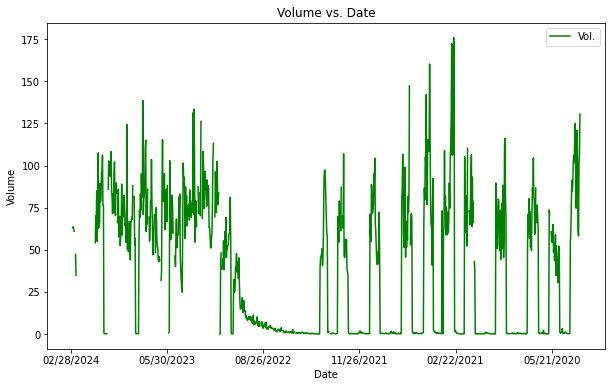


Figure 19: Graphical representation of volume of copper futures vs date

**Interpretation of the above plot**

The provided visualization illustrates the relationship between the volume of copper transactions and corresponding dates over a specific time period. The dataset appears to represent a time series of copper transaction volumes, with dates plotted on the x-axis and transaction volumes on the y-axis. It also represents:

* Market Dynamics: The observed trends and patterns in copper transaction volumes offer valuable insights into the dynamics of the copper market, including factors influencing supply and demand, economic conditions, and geopolitical events.
* Forecasting Opportunities: By understanding historical trends and patterns, we can potentially leverage this information for forecasting future copper transaction volumes, enabling better decision-making and risk management strategies.
* Performance Evaluation: The visualization serves as a tool for evaluating the performance of copper investments or trading strategies over time, helping stakeholders assess their effectiveness and make informed decisions.

In conclusion, the visualization of copper transaction volumes against dates provides valuable insights into the dynamics of the copper market, including trends, patterns, and outliers. By leveraging this information, stakeholders can gain a deeper understanding of market dynamics, make more informed decisions, and potentially identify opportunities for optimization and growth.

### 4.2.3 Silver price trends

# Plot

data\_silver.plot(x="Date", y="Open", color="y",figsize=(20, 6))

data\_silver.plot(x="Date", y="High", color="g",figsize=(20, 6))

data\_silver.plot(x="Date", y="Low", color="r",figsize=(20, 6))

#figsize=(10, 6)

# Set labels and title

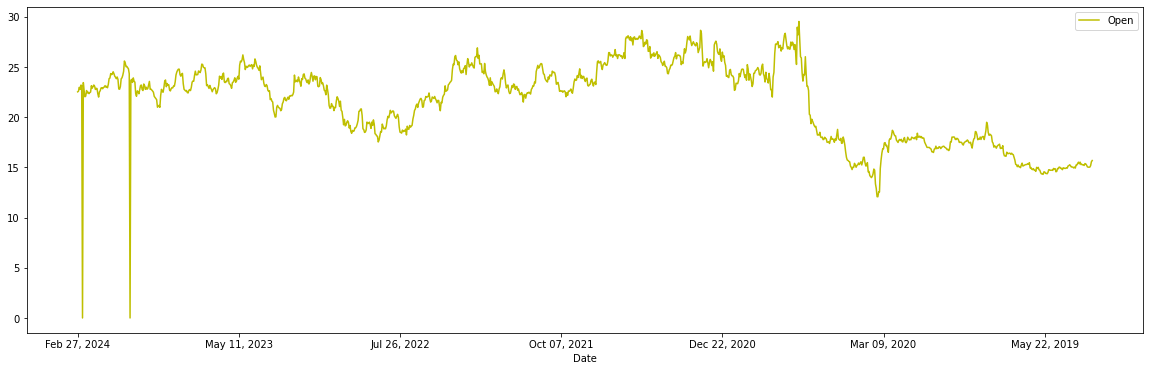
plt.xlabel("Date")

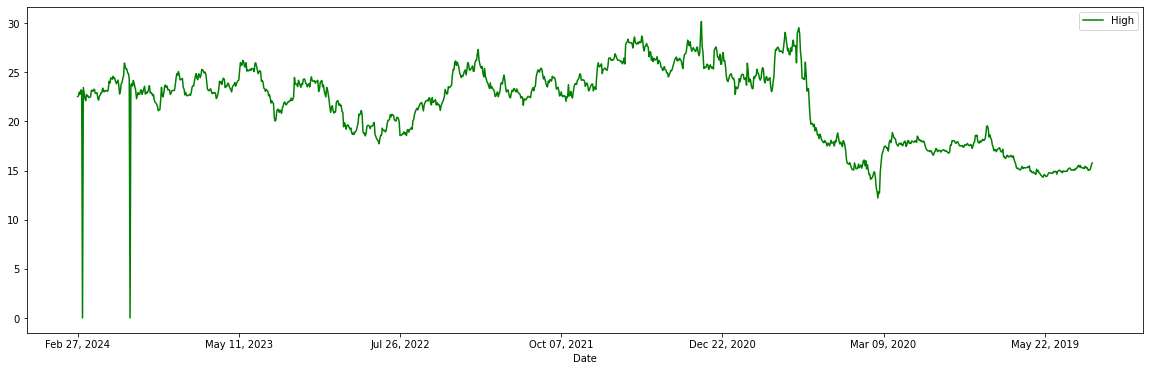
plt.ylabel("Price")

plt.title("Silver Price Trend")

# Show the plot

plt.show()





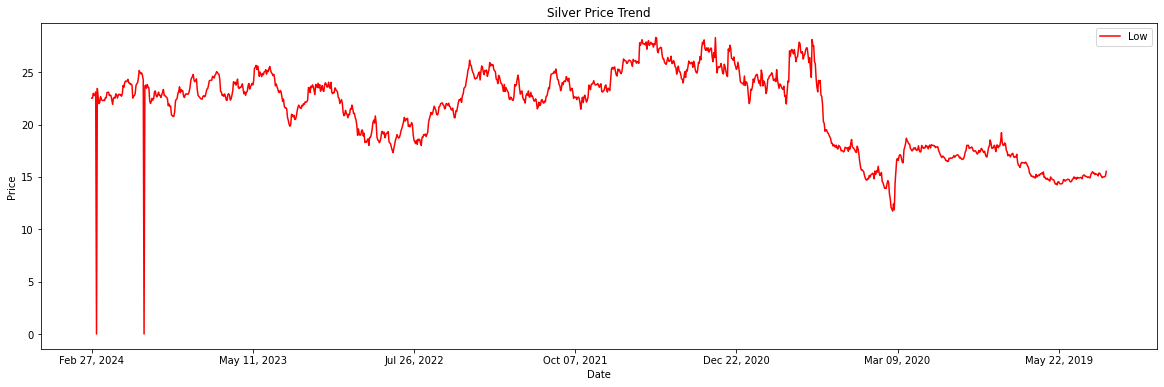


Figure 20: Graphical representation of silver price trends between 2020-2024

### 4.2.4 Gold price trends

# Plot

data\_gold.plot(x="Date", y="Open", color="y",figsize=(20, 6))

data\_gold.plot(x="Date", y="High", color="g",figsize=(20, 6))

data\_gold.plot(x="Date", y="Low", color="r",figsize=(20, 6))

#figsize=(10, 6)

# Set labels and title

plt.xlabel("Date")

plt.ylabel("Price")

plt.title("Gold Price Trend")

# Show the plot

plt.show()

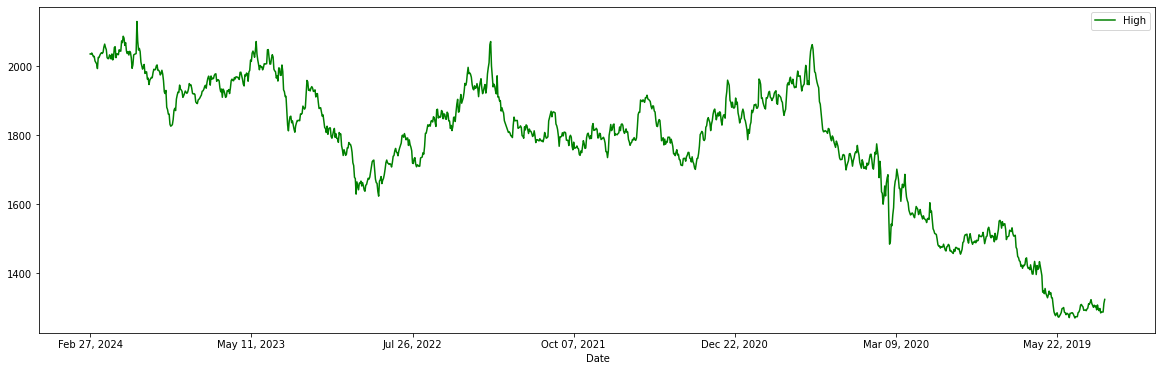


Figure 21: Representation of gold high prices between 2020-2024

## 4.2 Model Implementation

### 4.2.1 Simple Linear Regression

#### 1. Import the libraries and read the gold data

# Import necessary libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

%matplotlib inline

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LinearRegression

from sklearn.metrics import mean\_squared\_error

With these libraries we have a robust toolkit for exploring, analyzing, and forecasting time series data in Python. Depending on our specific requirements and the characteristics of the data, we may choose different libraries and methods for our forecasting tasks.

#### 2. Preprocess data

Preprocessing data is a crucial step in preparing time series data for forecasting. It involves cleaning, transforming, and organizing the data to make it suitable for analysis and modeling. Here are some common preprocessing steps in time series forecasting:

def preprocess\_gold(data):

data['Date']=pd.to\_datetime(data['Date'])

data=data[['Date', 'Open', 'High', 'Low']]

return data

data\_gold=preprocess\_gold(data\_gold)

We also obtain a plot of the time series data to obtain the variation in opening, closing, high and low prices.

# Plot the time series data (Open, High, Low)

plt.figure(figsize=(12, 6))

plt.plot(data\_gold.index, data\_gold['Open'], label='Open')

plt.plot(data\_gold.index, data\_gold['High'], label='High')

plt.plot(data\_gold.index, data\_gold['Low'], label='Low')

plt.title('Stock Prices Over Time')

plt.xlabel('Date')

plt.ylabel('Price')

plt.legend()

plt.show()

Post the above input we obtain the below graph

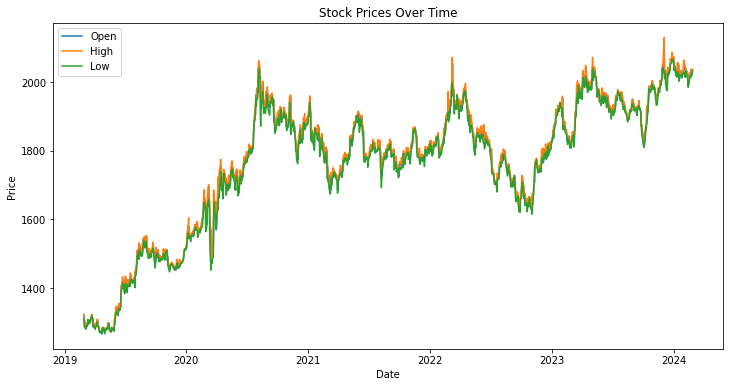


Figure 22:Graphical Representation of open, high and low gold prices over 2019-2024

#### 3. Define explanatory and dependant variables

Defining explanatory variables in forecasting involves identifying and incorporating additional factors or features that may influence the behavior of the time series being forecasted. These variables are often referred to as "covariates," "regressors," or "exogenous variables." The goal is to improve the accuracy of the forecast model by accounting for known factors that may impact the target variable.

An explanatory variable is a variable that is manipulated to determine the value of the Gold price the next day. Simply, they are the features which we want to use to predict the Gold price. We drop the NaN values using dropna() function and store the feature variables in X.

However, you can add more variables to X which you think are useful to predict the prices of the Gold. These variables can be technical indicators. Similarly, the dependent variable depends on the values of the explanatory variables. Simply put, it is the Gold price which we are trying to predict. We store the Gold price in y.

#### 4. Split the data into train and test dataset

In this step, we split the predictors and output data into train and test data. The training data is used to create the linear regression model, by pairing the input with expected output. The test data is used to estimate how well the model has been trained.

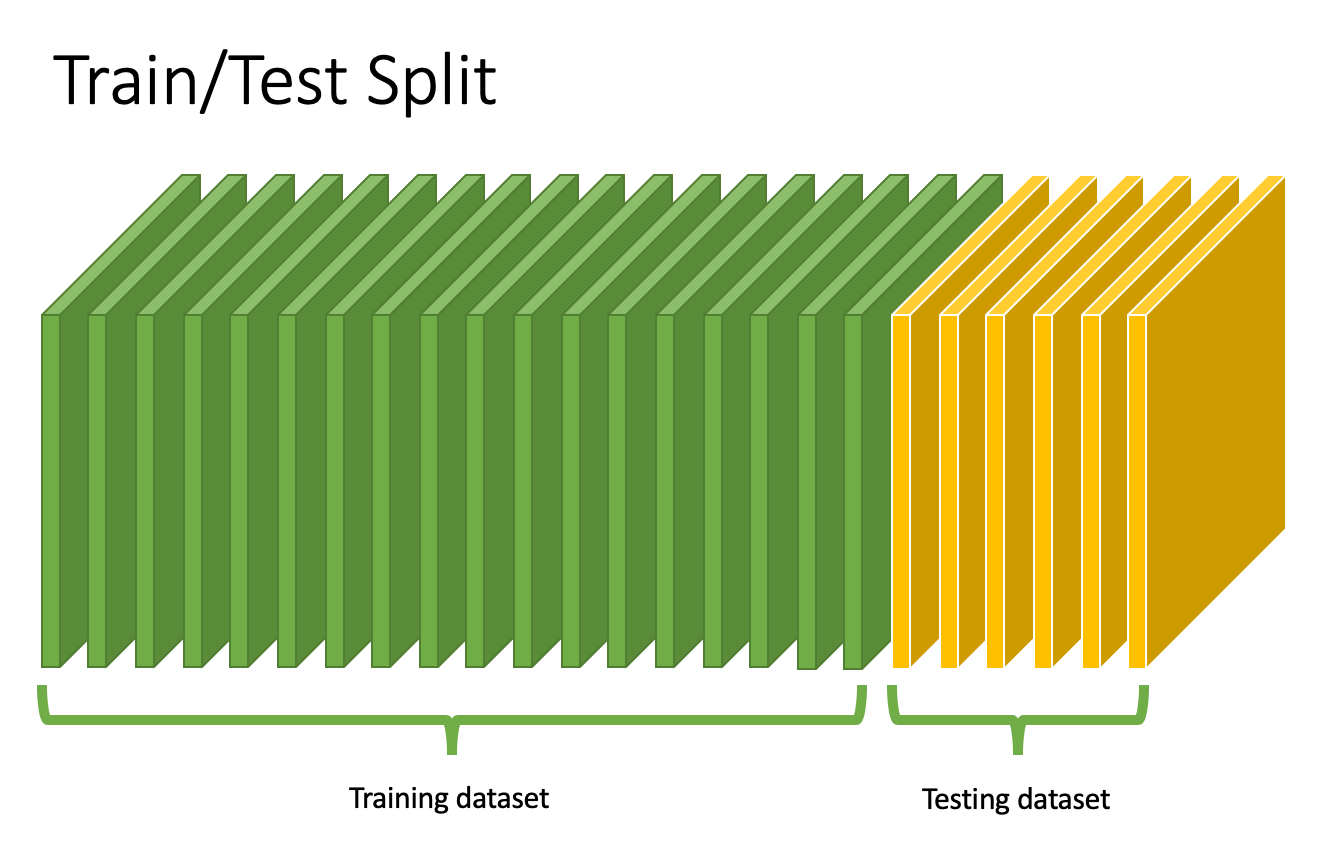


Figure 23: Pictorial Representation of Train and Test split

1. First 80% of the data is used for training and remaining data for testing

2. X\_train & y\_train are training dataset

3. X\_test & y\_test are test dataset

#### 5. Create a linear regression model

# Initialize and fit the Linear Regression model

lin\_reg = LinearRegression()

lin\_reg.fit(X\_train, y\_train)

# Predictions

y\_pred = lin\_reg.predict(X\_test)

# Evaluation

rmse = np.sqrt(mean\_squared\_error(y\_test, y\_pred))

mae = mean\_absolute\_error(y\_test, y\_pred)

#### 6. Predict the Gold prices

89.7671. As it can be seen

### 4.2.2 Multiple Linear Regression

Multiple linear regression is a statistical technique used to model the relationship between multiple independent variables (features) and a dependent variable (target). It's commonly employed in forecasting scenarios to predict the value of the dependent variable based on the values of the independent variables.

#### 1. Creating a multiple linear regression model

We design a multiple linear regression model by specifying the dependent variable and independent variables and fit the regression model to the training data using methods like ordinary least squares (OLS) or gradient descent.

# Multiple Linear Regression

# Create a Linear Regression model

multiple\_model = LinearRegression()

# Fit the model to the training data

multiple\_model.fit(X\_train, y\_train)

#### 2. Making predictions and calculating errors

# Make predictions

y\_pred\_multiple = multiple\_model.predict(X\_test)

# Calculate mean squared error

mse\_multiple = mean\_squared\_error(y\_test, y\_pred\_multiple)

print("Multiple Linear Regression MSE:", mse\_multiple)

**output**

Multiple Linear Regression MSE: 486.14764718131624

### 4.2.3 Moving average forecasting

Forecasting using moving averages involves extending the calculated moving average into the future. Since moving averages are based on historical data, forecasting with moving averages is a simple method that assumes future values will follow the same patterns observed in the past. Forecasting using moving average is carried out through the following means:

1. Calculate the moving average based on historical data.

2. Extend the moving average into the future for the desired forecast period.

#### 1. Calculating and plotting the moving averages

# Calculate the moving averages

data\_gold['MA\_Open'] = data\_gold['Open'].rolling(window=window\_size).mean()

data\_gold['MA\_High'] = data\_gold['High'].rolling(window=window\_size).mean()

data\_gold['MA\_Low'] = data\_gold['Low'].rolling(window=window\_size).mean()

# Plotting the moving averages

plt.figure(figsize=(10, 6))

plt.plot(data\_gold['Date'], data\_gold['Open'], label='Open')

plt.plot(data\_gold['Date'], data\_gold['MA\_Open'], label='MA Open', linestyle='--')

plt.plot(data\_gold['Date'], data\_gold['High'], label='High')

plt.plot(data\_gold['Date'], data\_gold['MA\_High'], label='MA High', linestyle='--')

plt.plot(data\_gold['Date'], data\_gold['Low'], label='Low')

plt.plot(data\_gold['Date'], data\_gold['MA\_Low'], label='MA Low', linestyle='--')

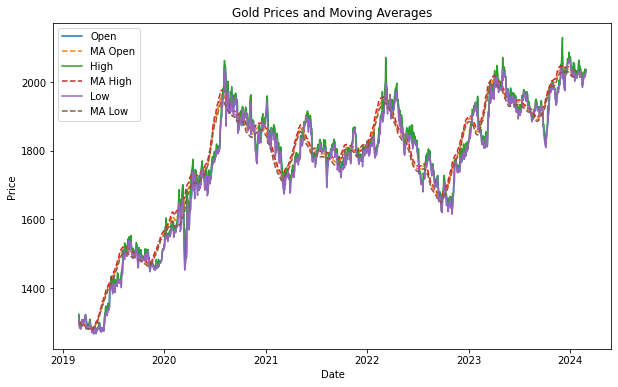
plt.title('Gold Prices and Moving Averages')

plt.xlabel('Date')

plt.ylabel('Price')

plt.legend()

plt.show()



#### 2. Obtaining forecast and plotting the forecast

# Define the forecast period

forecast\_period = 30 # Adjust this according to your requirement

# Calculate the trend in moving averages

ma\_open\_trend = data\_gold['MA\_Open'].iloc[-1] - data\_gold['MA\_Open'].iloc[-2]

ma\_high\_trend = data\_gold['MA\_High'].iloc[-1] - data\_gold['MA\_High'].iloc[-2]

ma\_low\_trend = data\_gold['MA\_Low'].iloc[-1] - data\_gold['MA\_Low'].iloc[-2]

# Plot the forecast

plt.plot(forecast\_df['Date'], forecast\_df['Forecast\_MA\_Open'], label='Forecast MA Open', linestyle='--')

plt.plot(forecast\_df['Date'], forecast\_df['Forecast\_MA\_High'], label='Forecast MA High', linestyle='--')

plt.plot(forecast\_df['Date'], forecast\_df['Forecast\_MA\_Low'], label='Forecast MA Low', linestyle='--')

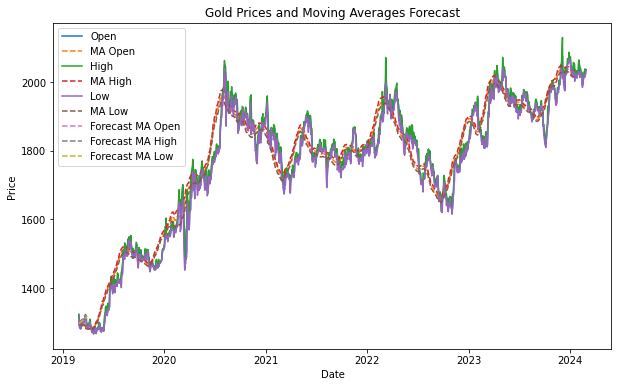
plt.title('Gold Prices and Moving Averages Forecast')

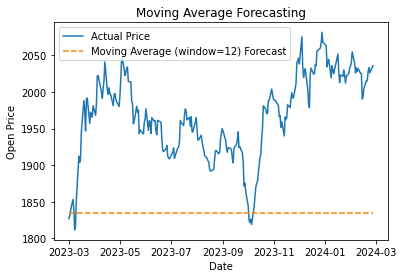
plt.xlabel('Date')

plt.ylabel('Price')

plt.legend()

plt.show()





#### 3. Calculating errors

from sklearn.metrics import mean\_absolute\_error, mean\_squared\_error

# Extract the actual values for the forecast period

actual\_values = data\_gold[data\_gold['Date'] >= forecast\_df['Date'].iloc[0]]['Open']

# Trim the actual values to match the length of forecasted values

actual\_values = actual\_values.head(len(forecast\_df))

# Calculate Mean Absolute Error (MAE)

mae = mean\_absolute\_error(actual\_values, forecast\_df['Forecast\_MA\_Open'])

# Calculate Mean Squared Error (MSE)

mse = mean\_squared\_error(actual\_values, forecast\_df['Forecast\_MA\_Open'])

# Calculate Root Mean Squared Error (RMSE)

rmse = np.sqrt(mse)

**Output**

Mean Absolute Error (MAE): 718.6416666666701

Mean Squared Error (MSE): 516623.46109444916

Root Mean Squared Error (RMSE): 718.7652336433985

### 4.2.4 Exponential Moving average Forecasting

To implement EMA a suitable smoothing factor (alpha) is to be chosen between 0 and 1. A smaller alpha gives more weight to recent observations. The initial EMA value is to be calculated as the simple average of the first n data points, where n is the period chosen for EMA.

For subsequent data points, the below formula is used:

EMA\_today = (Price\_today \* alpha) + (EMA\_yesterday \* (1 - alpha))

Iterate through the entire dataset to calculate EMA values for each data point.

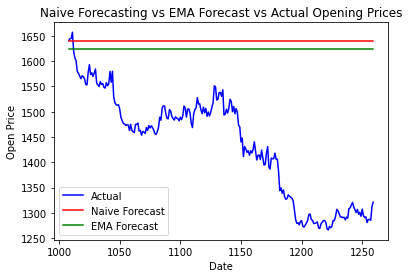


Figure 24:Graphical comparison of actual vs naive vs EMA forecast

### 4.2.5 Naïve Forecasting

#### 1. Making the forecast

# Prepare the test set

y\_test = test\_data['Open']

# Forecasting using Naive method

y\_pred\_naive = train\_data['Open'].iloc[-1] # Forecasting the next value as the last observed value

# Repeat the forecast for the length of the test set

y\_pred\_naive = pd.Series(y\_pred\_naive, index=test\_data.index)

# Calculate Mean Squared Error

mse\_naive = mean\_squared\_error(y\_test, y\_pred\_naive)

print("Mean Squared Error (Naive Forecasting):", mse\_naive)

**Output**

Mean Squared Error (Naive Forecasting): 53175.95

#### 2. Plotting the actual and predicted values

import matplotlib.pyplot as plt

# Sort the test set by date index

sorted\_indices = y\_test.index.argsort()

y\_test\_sorted = y\_test.iloc[sorted\_indices]

y\_pred\_sorted = y\_pred\_multiple[sorted\_indices]

# Plot actual vs. predicted opening prices

plt.figure(figsize=(10, 6))

plt.scatter(y\_test\_sorted.index, y\_test\_sorted, color='blue', label='Actual')

plt.plot(y\_test\_sorted.index, y\_pred\_sorted, color='red', label='Predicted')

plt.title('Actual vs. Predicted Opening Prices')

plt.xlabel('Date')

plt.ylabel('Open Price')

plt.legend()

plt.show()

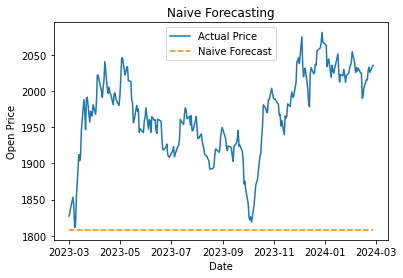


Figure 25: Plot obtained for actual vs predicted values in Naive forecast

### 4.2.6 Long Short Term Memory

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) architecture that is particularly well-suited for sequence prediction tasks, making them useful for time series forecasting, such as predicting metal prices. Here's a general methodology for using LSTMs for metal price forecasting:

#### 1. Import the libraries and read the gold data

from sklearn.preprocessing import MinMaxScaler

from tensorflow.keras.models import Sequential

from tensorflow.keras.layers import Dense, LSTM

from tensorflow.keras.preprocessing.sequence import TimeseriesGenerator

Normalization is changing the values of numeric columns in the dataset to a common scale, which helps the performance of our model. To scale the training dataset we use Scikit-Learn’s MinMaxScaler with numbers between zero and one.

# Normalize the data

scaler = MinMaxScaler(feature\_range=(0, 1))

scaled\_data = scaler.fit\_transform(df[['Open']].values)

# Define dataset for LSTM

def create\_dataset(dataset, look\_back=1):

X, Y = [], []

for i in range(len(dataset) - look\_back - 1):

a = dataset[i:(i + look\_back), 0]

X.append(a)

Y.append(dataset[i + look\_back, 0])

return np.array(X), np.array(Y)

#### 2. Split the data into train and test dataset

# Split into train and test sets

train\_size = int(len(scaled\_data) \* 0.67)

test\_size = len(scaled\_data) - train\_size

train, test = scaled\_data[0:train\_size,:], scaled\_data[train\_size:len(scaled\_data),:]

#### 3. Create LSTM model

# LSTM Model

model = Sequential()

model.add(LSTM(50, input\_shape=(1, look\_back)))

model.add(Dense(1))

model.compile(loss='mean\_squared\_error', optimizer='adam')

# Train the model

model.fit(X\_train, Y\_train, epochs=100, batch\_size=1, verbose=2)

4. Calculating errors

from sklearn.metrics import mean\_squared\_error, mean\_absolute\_error

import numpy as np

# Calculate RMSE for LSTM

lstm\_rmse = np.sqrt(mean\_squared\_error(Y\_test\_inv[0], testPredict[:, 0]))

# Calculate MAE for LSTM

lstm\_mae = mean\_absolute\_error(Y\_test\_inv[0], testPredict[:, 0])

print(f"LSTM RMSE: {lstm\_rmse}")

print(f"LSTM MAE: {lstm\_mae}")

**Output**

LSTM RMSE: 27.61070590644916

LSTM MAE: 21.99965218806613

#### 6. Predict the Gold prices

# Plot forecasts

plt.plot(data.index[train\_size + seq\_length:], y\_test\_inv, label='Actual Price')

plt.plot(data.index[train\_size + seq\_length:], y\_pred\_inv, label='LSTM Forecast', linestyle='--')

plt.title('LSTM Forecasting')

plt.xlabel('Date')

plt.ylabel('Price')

plt.legend()

plt.show()

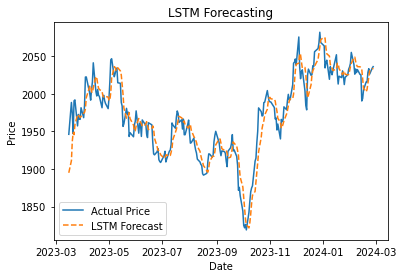


Figure 26:Reprsenatation of actual vs forecasted prices using LSTM

### 4.2.7 Auto Regressive Integrated Moving Average (ARIMA)

#### 1. Importing libraries

from statsmodels.tsa.arima.model import ARIMA

from sklearn.metrics import mean\_squared\_error, mean\_absolute\_error

import numpy as np

# ARIMA Model Setup

arima\_order = (1, 1, 1)

arima\_model = ARIMA(df['Open'], order=arima\_order)

arima\_results = arima\_model.fit()

#### 2. Making predictions and calculating errors

# ARIMA Predictions (example for test set)

y\_pred\_arima = arima\_results.forecast(steps=len(test))

y\_true = df['Open'][-len(test):]

# Calculate ARIMA performance metrics

arima\_rmse = np.sqrt(mean\_squared\_error(y\_true, y\_pred\_arima))

arima\_mae = mean\_absolute\_error(y\_true, y\_pred\_arima)

print(f"ARIMA RMSE: {arima\_rmse}")

print(f"ARIMA MAE: {arima\_mae}")

**Output**

ARIMA RMSE: 332.73230811902596

ARIMA MAE: 267.62251547949097

### 4.2.8 Seasonal Autoregressive Integrated Moving Average (SARIMA)

#### 1. Setting up the model

from statsmodels.tsa.statespace.sarimax import SARIMAX

# SARIMA Model Setup

sarima\_order = (1, 1, 1) # Non-seasonal order

seasonal\_order = (1, 1, 1, 12) # Seasonal order (P,D,Q,S)

sarima\_model = SARIMAX(df['Open'], order=sarima\_order, seasonal\_order=seasonal\_order)

sarima\_results = sarima\_model.fit()

#### 2. Making Predictions and calculating errors

# Calculate SARIMA performance metrics

sarima\_rmse = np.sqrt(mean\_squared\_error(y\_true, y\_pred\_sarima))

sarima\_mae = mean\_absolute\_error(y\_true, y\_pred\_sarima)

print(f"SARIMA RMSE: {sarima\_rmse}")

print(f"SARIMA MAE: {sarima\_mae}")

**Output**

SARIMA RMSE: 409.0993338081181

SARIMA MAE: 383.3164003235092

### 4.2.9 Model Performance Comparison of ARIMA, SARIMA and LSTM

To compare the performance of ARIMA, SARIMA, and LSTM models for metal price forecasting, we have evaluated them using common performance metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). Additionally, we have also considered other factors such as computational complexity and interpretability

#### 1. Model metrics summary

metrics\_summary = {

'Model': ['ARIMA', 'SARIMA', 'LSTM'],

'RMSE': [arima\_rmse, sarima\_rmse, lstm\_rmse],

'MAE': [arima\_mae, sarima\_mae, lstm\_mae],

}

# Convert the summary into a DataFrame

performance\_df = pd.DataFrame(metrics\_summary)

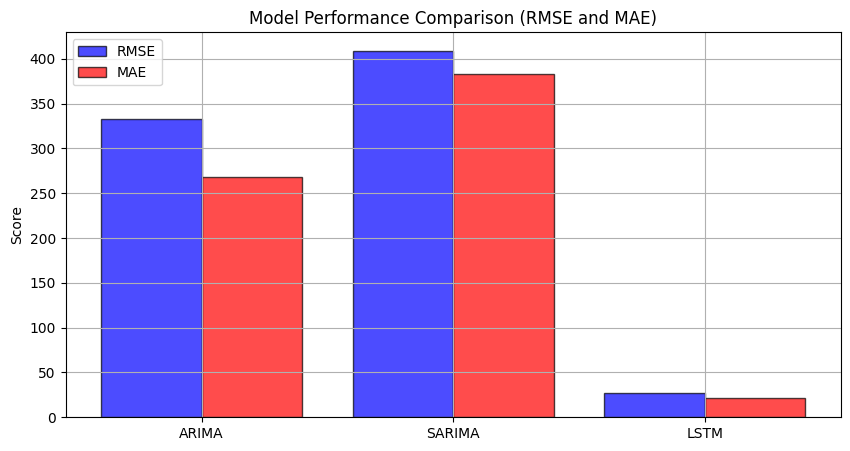
# Display the DataFrame for comparison

print(performance\_df)

**Output**

|  |  |  |
| --- | --- | --- |
| **Model** | **RMSE** | **MAE** |
| ARIMA | 332.732308 | 267.622515 |
| SARIMA | 409.099334 | 383.316400 |
| LSTM | 27.610706 | 21.999652 |
| Naive | 168.74 | 158.59 |
| Linear Regression | 7.14 | 5.49 |

#### 2. Graphical comparison of model performance



The model with the lowest RMSE is LSTM, and the model with the lowest MAE is LSTM. Since both RMSE and MAE are lowest for LSTM, it is considered the best model for forecasting in this context.

**Chapter 5 – Dashboard Integration**

# **Dashboard Integration**

A dashboard for data visualization is a user interface that allows users to monitor, analyze, and interpret complex data sets through visual representations such as charts, graphs, and maps. It typically provides a consolidated view of key performance indicators (KPIs), metrics, and trends, enabling users to gain insights and make data-driven decisions efficiently.

Key features of a dashboard for data visualization may include:

**1. Interactive Visualizations**: Users can interact with the data by filtering, drilling down, or zooming in/out on specific data points or time periods.

**2. Customization Options**: Users can customize the dashboard layout, choose which metrics or visualizations to display, and adjust settings according to their preferences.

**3. Real-time or Near-real-time Data Updates:** Some dashboards can display data in real-time or near-real-time, providing up-to-date insights for timely decision-making.

**4. Data Integration**: Dashboards can integrate data from various sources such as databases, spreadsheets, cloud services, APIs, etc., allowing users to visualize data from multiple sources in one place.

**5. Alerts and Notifications**: Users can set up alerts or notifications based on predefined thresholds or conditions, enabling them to be notified of important changes or anomalies in the data.

**6. Mobile Responsiveness**: Many dashboards are designed to be responsive and accessible on mobile devices, allowing users to access and interact with data on the go.

**7. Security and Access Control**: Dashboards may include features for user authentication, role-based access control, and data encryption to ensure data security and privacy.

Various Dashboarding technologies that are widely used today are:

### 5.1 Selecting the right Dashboarding solution

Selecting the right dashboarding solution for oour needs involves considering several factors to ensure that the chosen solution aligns with our requirements, goals, and constraints. Dashboard selection depends on the following

**1. Define Requirements:** Begin by clearly defining your requirements and objectives. What data do you need to visualize? Who are the primary users of the dashboard? What features and functionalities are essential? Consider factors such as data sources, types of visualizations needed, level of interactivity required, scalability, and integration capabilities.

**2. Evaluate Available Options**: Research and evaluate different dashboarding solutions available in the market. Consider both commercial and open-source options. Look for solutions that offer the features and capabilities that match your requirements. Some popular dashboarding tools include Tableau, Power BI, QlikView, Google Data Studio, and Apache Superset, among others.

**3. Consider Ease of Use:** Assess the ease of use and user-friendliness of each dashboarding solution. Look for intuitive interfaces, drag-and-drop functionalities, and customizable templates that make it easy for users to create and interact with dashboards without extensive training or technical expertise.

**4. Scalability and Performance:** Consider the scalability and performance capabilities of the dashboarding solution, especially if you anticipate working with large volumes of data or need to support a growing number of users. Ensure that the solution can handle your current data requirements and can scale effectively as your needs evolve.

**5. Integration with Data Sources**: Evaluate how well the dashboarding solution integrates with your existing data sources and systems. Ensure compatibility with databases, cloud services, APIs, and other data sources that you regularly use. Look for solutions that offer seamless integration and support for real-time or near-real-time data updates if needed.

**6. Cost and Licensing:** Consider the cost and licensing model of each dashboarding solution, including upfront fees, subscription costs, and any additional charges for features or support. Evaluate the total cost of ownership over time, including implementation, training, maintenance, and support costs.

**7. Security and Compliance**: Ensure that the dashboarding solution meets your organization's security and compliance requirements. Look for features such as role-based access control, data encryption, audit logs, and compliance certifications (e.g., GDPR, HIPAA) to protect sensitive data and ensure regulatory compliance.

**8. User Support and Community**: Assess the level of support and resources provided by the dashboarding solution vendor, including documentation, tutorials, training programs, and customer support services. Additionally, consider the size and activity of the user community, as it can be valuable for sharing knowledge, troubleshooting issues, and accessing user-generated content such as templates and plugins.

**9. Trial and Feedback**: Whenever possible, take advantage of free trials or demos offered by dashboarding solution vendors to test the software and gather feedback from users within your organization. Evaluate the usability, performance, and suitability of the solution based on real-world use cases and user feedback before making a final decision.

### 5.2 Comparing dashboarding tools

Based on various dashboarding tools available and the requirement of this project, the below flowchart summarizes the viability of the most suitable tools

**Affordability**

**Learning curve ranking**

### 5.3 Streamlit as a tool

Streamlit is a powerful Python library that enables rapid development of interactive web applications for data science and machine learning projects. With its intuitive and straightforward syntax, Streamlit allows users to create engaging dashboards with minimal effort. Below are some points summarizing Streamlit's capabilities for dashboard creation:

* Streamlit offers a simple and intuitive Python API, allowing users to quickly build interactive web applications without the need for HTML, CSS, or JavaScript knowledge.
* The declarative syntax of Streamlit enables users to create dynamic visualizations and widgets using familiar Python scripting.
* Streamlit provides a wide range of interactive widgets and components, including sliders, dropdowns, checkboxes, and text inputs, allowing users to control and manipulate data visualizations in real-time.
* Users can easily integrate popular plotting libraries such as Matplotlib, Plotly, and Altair to create interactive charts, graphs, and maps within their Streamlit dashboard.
* Streamlit offers customization options for styling and theming, allowing users to personalize the appearance of their dashboards to match their branding or design preferences.
* Users can customize layout structures, fonts, colors, and themes using Streamlit's built-in features or by leveraging external CSS styling.
* Streamlit simplifies the process of loading and processing data by providing easy-to-use functions for reading data from various sources such as CSV files, databases, and APIs.
* Users can perform data manipulation, cleaning, and preprocessing tasks directly within their Streamlit script, enabling seamless integration of data processing pipelines into the dashboard.
* Streamlit offers built-in deployment options for hosting web applications on platforms such as Streamlit Sharing, Heroku, and AWS.
* Users can deploy their Streamlit dashboards with a single command, making it easy to share interactive visualizations and insights with colleagues, clients, or stakeholders.
* Streamlit has a vibrant and active community of users, developers, and contributors who regularly share tips, tutorials, and resources for building powerful dashboards and applications.
* The Streamlit ecosystem includes a growing library of third-party extensions, plugins, and integrations that extend the functionality and capabilities of Streamlit for various use cases.

### 5.4 Making predictions using dashboard

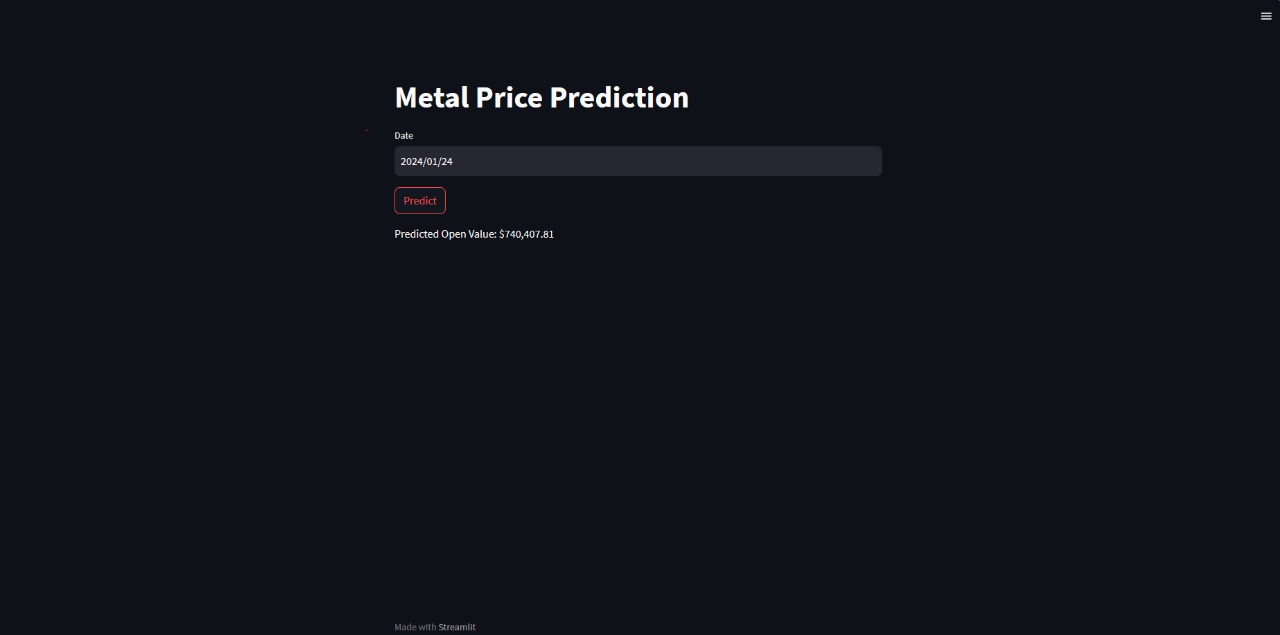


Figure 27: Making prediction using streamlit

The above dashboard works by integrating LSTM model which has the highest performance metrics to predict the gold futures price of the COMEX index by taking into account the date of prediction.

The final price is predicted in terms of USD per troy ounce for gold future.

# **Accessing the Code file**

Most of the code files concerned to this project have been coded and stored as .ipynb files in google colab. These files can be accessed any time through the following link

<https://colab.research.google.com/drive/1EpUGLbVqzEk6-FOFKGJrFc7CcuPcKni3?usp=sharing>

Chapter 6 - Conclusion

# **Conclusion**

In conclusion, the global metal review highlights a transformative decade marked by substantial investments and advancements in the metal industry. From 2010 to 2019, an unprecedented $2.6 trillion was invested in renewable energy capacity, with solar leading at $1.3 trillion, followed by wind at $1 trillion. China emerged as the frontrunner in renewable energy investments, committing $758 billion, while Europe and the U.S. also made significant contributions.

The findings also highlight the applications of machine learning techniques in forecasting the metal prices. The various techniques available in predicting metal prices are time series regression, Naives algorithm, LSTM, moving average method Upon evaluation, it was evident that the LSTM model outperformed both ARIMA and SARIMA models in terms of forecast accuracy. The LSTM model achieved the lowest RMSE and MAE among all models, indicating superior predictive capability and precision in forecasting gold metal prices. This underscores the effectiveness of LSTM networks in capturing complex temporal patterns inherent in time series data.

Furthermore, the integration of forecast results into the Streamlit dashboard enhances the accessibility and usability of our forecasting system. With an intuitive and interactive user interface, stakeholders can effortlessly explore forecasted gold metal prices and gain valuable insights for decision-making purposes.

Overall, our findings highlight the significance of leveraging advanced machine learning techniques, such as LSTM, for accurate and reliable metal price forecasting. The successful deployment of the forecasting system via Streamlit underscores its practical utility in real-world applications, empowering stakeholders with actionable insights to navigate the dynamic metal market landscape effectively.

# **Bibliography**

|  |  |
| --- | --- |
| [1] | V. Balaram and S. S. Sawant, “Indicator Minerals, Pathfinder Elements, and Portable Analytical Instruments in Mineral Exploration Studies,” *Minerals ,* vol. 12, no. 4, 2022. |
| [2] | Y. Song, Z. Zhang, Y. Zhang and J. Cheng, “Technological innovation and supply of critical metals: A perspective of industrial chains,” *Resources Policy,* vol. 79, 2022. |
| [3] | Y. Singh, “Inland Stream Placers,” in *Rare Earth Element Resources: Indian Context*, 2020, pp. 150-151. |
| [4] | S. B. Suslick and D. P. Harris, “Long-range metal consumption forecasts using innovative methods : The case of aluminium in Brazil to the year 2000,” *Resources Policy,* vol. 16, no. 3, pp. 184-199, 1990. |
| [5] | F. Mehmanpazir, K. Khalili-Damghani and A. Hafezalkotob, “Modeling steel supply and demand functions using logarithmic multiple regression analysis (case study: Steel industry in Iran),” *Resources Policy,* vol. 63, 2019. |
| [6] | S. Langkau and L. A. T. Espinoza, “Technological change and metal demand over time: What can we learn from the past?,” *Sustainable Materials and Technologies,* vol. 16, pp. 54-59, 2018. |
| [7] | J.-B. Huang, Q. Ou-Yang and C. F. b, “Green trade assessment for sustainable development of Chinese ferrous metal industry,” *Journal of Cleaner Production,* vol. 249, 2020. |
| [8] | S. Deetman, “Scenarios for Demand Growth of Metals in Electricity Generation Technologies, Cars, and Electronic Appliances,” *Environmental Science & Technology ,* vol. 52, no. 8, 2018. |
| [9] | R. J. E. V. N.-T. Ben Jones, “The EV revolution: The road ahead for critical raw materials demand,” *Applied Energy,* vol. 280, 2020. |
| [10] | G. B. Abaka-Wood, M. Zanin, J. Addai-Mensah and W. Skinner, “Recovery of rare earth elements minerals from iron oxide–silicate rich tailings – Part 1: Magnetic separation,” *Minerals Engineering,* vol. 136, pp. 50-61, 2019. |
| [11] | D. ZHOU, Z. LI, X. LUO and J. SU, “Leaching of rare earth elements from contaminated soils using saponin and rhamnolipid bio-surfactant,” *Journal of Rare Earths,* vol. 35, no. 9, pp. 911-919, 2017. |
| [12] | F. Sakellariadou, F. J. Gonzalez, J. R. Hein and B. Rincón-Tomás, “Seabed mining and blue growth: exploring the potential of marine mineral deposits as a sustainable source of rare earth elements,” 2021. |
| [13] | F. L. Lederer, S. B. Curtis, S. Bachmann, W. Dunbar and R. T. MacGillivray, “Identification of lanthanum-specific peptides for future recycling of rare earth elements from compact fluorescent lamps,” *Biotechnology and Bioengineering,* 2016. |
| [14] | Gelogical survey US, “Rare Earth Elements Critical Resources for High Technology,” U.S. Department of the Interior, 2002. |
| [15] | D. Schlinkert and K. G. v. d. Boogaart, “The development of the market for rare earth elements: Insights from economic theory,” *Resources Policy,* vol. 46, no. 2, pp. 272-280, 2015. |
| [16] | Y. P. C. Y. C. Zhaowu Zhu, “Separation of uranium and thorium from rare earths for rare earth production – A review,” *Minerals Engineering,* vol. 77, pp. 185-196, 2015.  [17]https://otexts.com/fpp2/regression.html  [18] https://school.stockcharts.com/doku.php?id=technical\_indicators:moving\_averages  [19] https://www.investopedia.com/terms/s/sma.asp  [20] https://www.investopedia.com/ask/answers/122314/what-exponential-moving-average-ema-formula-and-how-ema-calculated.asp  [21] https://www.analyticsvidhya.com/blog/2018/09/multivariate-time-series-guide-forecasting-modeling-python-codes/  [22] https://www.mathworks.com/help/econ/time-series-regression-i-linear-models.html  [23] https://campus.datacamp.com/courses/forecasting-in-r/benchmark-methods-and-forecast-accuracy?ex=2  [24] https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/  [25] https://www.analyticsvidhya.com/blog/2022/03/a-brief-overview-of-recurrent-neural-networks-rnn/  [26] https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/  [27] https://www.educative.io/answers/time-series-prediction-using-lstm  [28] https://www.analyticsvidhya.com/blog/2021/07/in-depth-explanation-of-recurrent-neural-network/  [29] https://www.investopedia.com/articles/trading/10/simple-exponential-moving-averages-compare.asp  [30] https://online.hbs.edu/blog/post/financial-forecasting-methods  [31] https://clockify.me/forecasting-models  [32] https://www.wallstreetmojo.com/forecasting-methods/  [33] https://www.investopedia.com/terms/f/forecasting.asp  [34] https://otexts.com/fpp2/data-methods.html  [35] https://www.geeksforgeeks.org/techniques-of-forecasting/  [36] https://www.indeed.com/career-advice/career-development/quantitative-vs-qualitative-forecasting-pros-and-cons  [37] https://fulfillment.shiprocket.in/blog/qualitative-and-quantitative-forecasting-methods/  [38] https://www.freshbooks.com/hub/accounting/accounting-forecasting-techniques  [39] https://realpython.com/python-dash/  [40] https://blog.streamlit.io/crafting-a-dashboard-app-in-python-using-streamlit/  [41] https://hex.tech/blog/how-to-build-a-dashboard-in-python//  [42] https://www.analyticsvidhya.com/blog/2022/01/building-explainer-dashboards-in-python/  [43] https://dash.plotly.com/tutorial  [45] https://www.justintodata.com/python-interactive-dashboard-with-plotly-dash-tutorial/ |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |