# Agricultural Raw Material Price Analysis

A Project Report

submitted in partial fulfillment of the requirements

of

AIML FUNDAMENTALS WITH CLOUD COMPUTING

&

GEN AI

by

**AADHI ESHWAR. K**

**NM ID: au912321114001**

**aadhieshwar.k@gmail.com**

Under the Guidance of

**P. Raja, Master Trainer**

**ACKNOWLEDGEMENT**

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this thesis work.

Firstly, we would like to thank my supervisor, P. Raja master trainer, for being a great mentor and the best adviser I could ever have. His advice, encouragement and the critics are a source of innovative ideas, inspiration and causes behind the successful completion of this project. The confidence shown in me by him was the biggest source of inspiration for me. It has been a privilege working with him for the last one year. He always helped me during my project and many other aspects related to the program. His talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional.

## TABLE OF CONTENTS



### LIST OF FIGURES



### LIST OF TABLES

## ABSTRACT of the Project

The agricultural industry is heavily influenced by fluctuating raw material prices, which are subject to market conditions, environmental factors, and global demand. Understanding these price trends is critical for stakeholders, such as farmers, investors, and policymakers, to make informed decisions. This project aims to analyze the historical data of agricultural raw material prices through Exploratory Data Analysis (EDA) to uncover insights into price behavior and correlations between various materials.

The primary objectives of this study include identifying the materials with the highest and lowest price ranges, examining the percentage changes in prices over the years, and establishing correlations between different raw materials using a heat map. The dataset, containing monthly price data for raw materials such as coarse wool, copra, cotton, and rubber, was cleaned and processed for analysis. Key findings reveal significant variations in price behaviors, with some materials experiencing dramatic shifts over time, while others show stable trends.

To visualize these findings, price trends over time were plotted for each material, highlighting the periods of sharp increase or decrease in value. Additionally, a correlation matrix was generated, providing insights into how the prices of different raw materials are related to each other. Strong correlations were observed between some materials, which may indicate shared market influences.

The results of this analysis can be used to forecast potential future price trends and make strategic decisions in the agricultural supply chain. By mapping these relationships, the study provides a valuable tool for understanding the complex dynamics that drive raw material prices. Future work could involve integrating external factors, such as weather conditions or trade policies, to create more comprehensive predictive models.

# CHAPTER 1

## Introduction

### 1.1 Problem Statement:

Agricultural raw material prices fluctuate over time due to factors such as weather, market demand, and supply constraints. Understanding these price changes is vital for stakeholders.

### 1.2 Motivation:

The project aims to help decision-makers in the agricultural sector anticipate trends, enabling them to make informed decisions about raw material pricing.

### 1.3 Objective:

- Identify the high and low range raw materials according to their prices.
- Analyze the percentage change in prices over the years.
- Map correlations between different materials using a heat map.

### 1.4 Scope of the Project:

The project is focused on price analysis and trends over time. It does not account for external factors influencing these changes.

# CHAPTER 2

## Literature Survey

### 2.1 Review of Relevant Literature and Previous Work in Agricultural Raw Material Price Analysis

- **Price Volatility in Agricultural Markets**

  Agricultural markets are known for their significant price volatility due to factors such as seasonality, supply chain disruptions, and environmental conditions. Numerous studies have explored the causes and consequences of price volatility in raw materials, particularly in developing countries where agriculture plays a key role in the economy. Studies like those by Jha et al. (2012) emphasize the impact of weather patterns, global demand, and trade policies on agricultural price fluctuations. They also discuss how price instability affects farmers' income, food security, and overall economic development.

- **The Role of Commodity Prices in Global Economic Shifts**

  Research by Abbott and Borot de Battisti (2011) explores the role of global commodity price movements and their impact on agricultural markets. They discuss how rising global demand, export restrictions, and increased speculation in commodity markets have led to heightened price volatility. The study highlights how agricultural raw materials, such as cotton, rubber, and pulp, are influenced by both local and global market conditions, impacting both producers and consumers.

- **Time Series Analysis of Agricultural Prices**

  Awokuse and Xie (2015) have conducted time series analyses of agricultural raw material prices to identify long-term trends and seasonal patterns. Their work focuses on forecasting price trends using autoregressive integrated moving average (ARIMA) models. The study indicates that time series models can effectively capture the trends in agricultural prices and are useful for predicting future price behavior based on historical data.

- **Impact of Climate Change on Agricultural Prices**

  A growing body of literature focuses on the effects of climate change on agricultural productivity and price stability. Studies like Lobell et al. (2011) emphasize that climate variability leads to unpredictable crop yields, directly affecting the prices of raw materials such as coarse wool, cotton, and rubber. The unpredictability in yields due to extreme weather events makes these markets more volatile, contributing to sharp price fluctuations.

- **Commodity Price Correlations in Agricultural Markets**

  Studies by Baffes (2007) examine the interdependence of agricultural commodity prices. His research shows that certain agricultural commodities, such as rubber, wood pulp, and cotton, often exhibit correlated price movements due to shared market conditions, such as input costs, demand from the same industries, or global supply chain disruptions. Understanding these correlations helps in forecasting and managing price risks.

- **Supply Chain Disruptions and Agricultural Price Variability**

  Research by Pindyck and Rotemberg (1990) investigates how supply chain disruptions (e.g., transport issues, labor strikes, or international trade restrictions) significantly influence the prices of agricultural raw materials. Their work is particularly relevant when understanding how disruptions at one point in the supply chain can have ripple effects on price movements across different raw materials.

## 2.2 Existing Models, Techniques, and Methodologies Related to Agricultural Raw Material Price Analysis

There are several models and methodologies commonly used to analyze agricultural raw material prices. These approaches range from statistical models that predict price changes to machine learning models that can identify trends and correlations. Below are some of the key models and techniques used in this field:

1. **Time Series Analysis Models:**

   Time series models are widely used to analyze agricultural raw material prices because of the temporal nature of price data. The following are the most common time series models:

- **ARIMA (Auto Regressive Integrated Moving Average) Model**

  ARIMA is one of the most commonly used models for forecasting agricultural prices. It combines three components: auto regression (AR), differencing (I), and moving averages (MA). This model is effective in capturing trends and seasonality in agricultural price data, making it ideal for price prediction over time.
  Use case: Forecasting future prices of raw materials like cotton, wool, or rubber based on past price data.

- **SARIMA (Seasonal ARIMA) Model**

  SARIMA extends ARIMA by accounting for seasonality in the data. Agricultural prices often exhibit seasonal patterns due to harvest cycles, weather conditions, and

demand fluctuations. SARIMA is used when seasonality plays a significant role in price behavior.
Use case: Analyzing seasonal price fluctuations in crops like cotton or copra that depend on specific growing seasons.

2. **GARCH (Generalized Autoregressive Conditional Heteroskedasticity) Model**

The GARCH model is used to estimate volatility in agricultural commodity prices. It captures the tendency of price volatility to cluster, where periods of high volatility tend to be followed by more high volatility and low volatility by low volatility. This is particularly useful in agricultural markets that experience sudden price spikes due to unexpected events (e.g., extreme weather or geopolitical instability).
Use case: Measuring the risk and uncertainty in the price of volatile materials like rubber, which can experience sharp price swings.

3. **Vector Auto Regression (VAR) Model**

VAR models are employed when analyzing the relationship between multiple time series data. In the context of agricultural raw materials, VAR can be used to examine how different commodities' prices influence each other. This is helpful for understanding correlations between raw materials such as wool, cotton, rubber, and wood pulp.
Use case: Understanding how the price of one raw material (e.g., cotton) might impact or correlate with another (e.g., rubber) over time.

4. **Machine Learning Techniques**

In recent years, machine learning models have been increasingly applied to agricultural raw material price prediction and analysis.

- **Random Forest**

Random Forest is an ensemble learning method that is used for both regression and classification tasks. It can model complex relationships between input variables (such as weather conditions, market demand, etc.) and the target variable (raw material prices). This technique is useful for non-linear and high-dimensional data.
Use case: Predicting the price of agricultural materials based on a large number of variables, including historical prices, supply chain disruptions, and global demand.

- **Support Vector Machines (SVM)**

SVM is a supervised learning method used for regression and classification. It can model complex relationships by maximizing the margin between data points, making it effective for price prediction where patterns are non-linear.

Use case: Forecasting future agricultural prices using past trends and external factors like trade policies and economic indicators.

- **Neural Networks (Deep Learning)**

Neural networks, particularly Long Short-Term Memory (LSTM) networks, have been applied to time series forecasting. LSTM networks are capable of learning long-term dependencies in sequential data, making them suitable for predicting agricultural prices over time. LSTMs can model both seasonality and irregular price patterns effectively.
Use case: Predicting long-term price trends for agricultural commodities using past price data and other influencing factors.

5. **Econometric Models**

Econometric models are often used to understand how external economic factors (such as inflation, trade policies, and global market demand) influence agricultural prices. These models can help quantify the impact of macroeconomic conditions on agricultural raw materials.

- **Ordinary Least Squares (OLS) Regression**

OLS regression is a basic but powerful tool for understanding the relationship between agricultural prices and other variables like production levels, export volumes, or input costs (e.g., fertilizers, labor). It can help determine how much a change in these variables influences raw material prices.
Use case: Estimating the impact of a change in global demand for cotton on its price.

- **Cointegration Models**

Cointegration models are used to analyze the long-term relationships between multiple time series. These models can help determine whether agricultural commodity prices move together over time, even if they are subject to short-term deviations.
Use case: Analyzing the long-term relationship between cotton and wool prices to understand whether they are influenced by similar factors.

6. **Supply Chain and Market Disruption Models**

Supply chain disruption models help quantify the impact of events such as natural disasters, trade restrictions, or logistical delays on agricultural prices. These models can use historical data to simulate how such disruptions might affect the price of agricultural raw materials.

- **System Dynamics Models**

  These models simulate the behavior of complex systems over time. In agriculture, they are used to model the interactions between various components of the supply chain (e.g., production, storage, transportation) and how disruptions in one part can cause price fluctuations.
  Use case: Modeling the impact of a drought on the cotton supply chain and the resulting price effects.

- **Input-Output Models**

  Input-output models describe the relationships between different sectors of the economy. In agriculture, they are used to analyze how changes in one sector (e.g., rising fertilizer costs) impact the prices of raw materials (e.g., cotton or wool).
  Use case: Understanding how an increase in input costs affects the price of multiple raw materials simultaneously.

**7. Descriptive and Exploratory Data Analysis (EDA)**

Exploratory Data Analysis (EDA) is a key methodology used to summarize and visualize the key characteristics of agricultural raw material price data. It involves:

- **Descriptive Statistics**

  Descriptive statistics (e.g., mean, median, standard deviation) are used to summarize the central tendency, dispersion, and distribution of prices. This helps in identifying the overall trends and anomalies in the dataset.
  Use case: Understanding the average price movement of raw materials and identifying periods of extreme price changes.

- **DataVisualization**

  Visualization techniques such as line plots, heat maps, and histograms are used to analyze price trends, ide89ntify correlations between commodities, and explore patterns in the data.
  Use case: Using a heat map to identify correlations between wool, cotton, and rubber prices.

## 2.3 Gaps or Limitations in Existing Solutions

### 1. Limited Incorporation of External Factors

Many traditional models, such as ARIMA, SARIMA, and OLS regression, primarily rely on historical price data without adequately considering external factors such as weather, geopolitical events, and global supply chain disruptions. These factors have a significant impact on agricultural raw material prices but are often excluded or treated simplistically in existing solutions. This results in models that may not accurately predict price changes during unexpected events or when external influences are dominant.

### 2. Lack of Real-Time Data Processing

Most of the existing solutions rely on static datasets and lack the capability to handle real-time data. Agricultural raw material prices are highly volatile and can change rapidly due to shifts in supply and demand. Models that cannot incorporate real-time data may provide outdated insights, limiting their usefulness for decision-making in dynamic markets.

### 3. Inadequate Handling of Price Volatility

Although models like GARCH are designed to handle volatility, many simpler models (e.g., ARIMA) do not account for the clustering of price volatility. As a result, they may fail to capture sudden price spikes or drops, leading to inaccurate predictions, especially in markets where price volatility is high due to environmental factors or market disruptions.

### 4. Insufficient Attention to Correlations Between Materials

Some models focus on individual commodities without adequately accounting for the relationships between different agricultural raw materials. This is a significant limitation since raw materials often experience price shifts due to shared market conditions or substitute/complementary relationships (e.g., cotton and wool). Ignoring these correlations can lead to incomplete analysis and missed opportunities for more accurate forecasting.

### 5. Seasonality and Irregular Patterns

While SARIMA models address seasonality, they may not perform well with irregular patterns that deviate from regular seasonal trends, such as those caused by climate change or unexpected market disruptions. Similarly, econometric models may fail to capture non-linear relationships, limiting their ability to model more complex, irregular price behaviors.

### 6. Overfitting in Machine Learning Models

Machine learning models, while powerful, often suffer from overfitting, especially when applied to relatively small or noisy datasets. When overfitted, these models may perform well on historical data but fail to generalize to future price movements. Additionally, the interpretability of complex machine learning models (e.g., neural networks) can be a challenge, making it harder to derive actionable insights.

**2.4 How This Project Addresses These Gaps**

1. **Incorporation of External Factors (Future Work)**

While this project primarily focuses on raw material price analysis, it sets the foundation for integrating external factors such as weather patterns, global market conditions, and geopolitical events. By analyzing correlations between materials, this project identifies areas where external influences could be further incorporated in future iterations, making models more robust and reflective of real-world conditions.

2. **Exploratory Data Analysis (EDA) and Visualization of Correlations**

This project emphasizes the use of EDA techniques to explore the relationships between different agricultural raw materials. By generating correlation heat maps, this project highlights the interdependencies between commodities that many existing models overlook. These insights can be used to refine predictive models by incorporating these correlations into forecasting methods.

3. **Handling Price Volatility**

The project highlights the need for models that capture volatility by examining price ranges and percentage changes over time. Future iterations of this project could incorporate models like GARCH to handle volatility more effectively, but this initial phase identifies which materials experience the highest fluctuations and need advanced modeling techniques.

4. **Improved Accuracy Through Correlation Analysis**

By mapping the correlations between various agricultural materials, this project addresses the gap in understanding how raw material prices influence one another. Unlike many existing solutions that analyze commodities in isolation, this project highlights the importance of understanding commodity relationships (e.g., how cotton and wool prices might move together), offering more comprehensive insights into price movements.

5. **Addressing Seasonal and Irregular Patterns**

The project’s time series visualizations of price changes over time provide a foundation for addressing seasonality and irregular patterns. While SARIMA models are effective for regular seasonal data, this project identifies irregularities in price changes that point to the need for more sophisticated models, such as machine learning algorithms that can detect complex patterns.

### 6. Avoiding Overfitting by Simplifying Initial Models

The initial approach in this project focuses on EDA and straightforward statistical models to avoid overfitting, a common issue in complex machine learning models. By beginning with simple yet powerful techniques like correlation analysis and price range evaluations, the project establishes a clear understanding of price behavior before progressing to more advanced machine learning techniques.

### 7. Setting the Stage for Real-Time Analysis

Although the current project does not incorporate real-time data, it demonstrates the need for dynamic data updates through its focus on price fluctuations and volatility. In future work, the methods and visualizations used in this project can be adapted to work with real-time data streams, allowing for real-time price analysis and decision-making.

# CHAPTER 3

## Proposed Methodology for Agricultural Raw Material Price Analysis

The primary goal of this project is to analyze the historical data of agricultural raw material prices, uncover patterns, and explore relationships between different commodities. The methodology involves several steps, from data preprocessing to visualization, focusing on identifying trends, price fluctuations, percentage changes, and correlations. This detailed methodology covers the following phases:

### 1. Data Collection and Preprocessing

The dataset contains monthly agricultural raw material prices, along with the percentage change for each material. Before performing any analysis, it's important to clean and preprocess the data to ensure accuracy.

#### 1.1. Data Loading

- Step: Import the dataset (in CSV format) containing raw material prices over the years.
- Tools Used: pandas for data manipulation.
- Goal: Read the data and store it in a structured format suitable for analysis

#### 1.2 Data Cleaning

- Step: Perform data cleaning to handle any inconsistencies, such as missing values, formatting issues, and conversion errors.
    - Convert price columns (e.g., "Coarse wool Price", "Cotton Price") from strings to numeric values, removing commas or other symbols.
    - Remove percentage signs from change columns (e.g., "Coarse wool price % Change") and convert them to float.
    - Convert the "Month" column to datetime format to enable time-series analysis.
    - Handle missing data by either removing rows with null values or imputing them using interpolation.
- Tools Used: pandas for data cleaning and transformation.
- Goal: Ensure that the dataset is clean and ready for analysis.

#### 1.3 Data Structure Overview

- Step: Explore the structure of the data by:
    - Displaying the first few rows of the dataset.
    - Checking for missing values and data types.
    - Generating basic descriptive statistics to understand the range, mean, and variability of the raw material prices.

- Tools Used: pandas and numpy for descriptive statistics.
- Goal: Get a preliminary understanding of the data's structure and characteristics.

## 2. Exploratory Data Analysis (EDA)

Exploratory Data Analysis is critical to uncovering insights about the raw material prices, trends over time, and correlations between different materials.

### 2.1 Analyzing Price Ranges

- Step: Calculate the minimum and maximum prices for each agricultural raw material over the years.
  - Identify the raw materials with the highest and lowest prices by computing the range for each commodity.
- Tools Used: pandas for aggregation.
- Goal: Identify which materials exhibit the highest price fluctuations and which are more stable.

### 2.2 Percentage Change Analysis

- Step: Analyze the percentage changes in prices for each material.
  - Calculate the minimum and maximum percentage change for each raw material to determine the most volatile materials.
  - Highlight periods with significant changes, such as price spikes or drops.
- Tools Used: pandas for aggregation.
- Goal: Identify materials with the largest percentage price changes, offering insights into price volatility.

### 2.3 Time Series Analysis

- Step: Visualize how prices have changed over time for selected raw materials.
  - Plot time series graphs of selected materials (e.g., Coarse wool, Copra, Cotton) to show price trends.
  - Identify any seasonal or long-term trends.
- Tools Used: matplotlib for visualization.
- Goal: Reveal patterns in the data over time, such as price peaks, stability periods, or seasonal trends.

## 3. Correlation Analysis

Understanding the relationships between different agricultural raw materials is crucial for identifying interdependencies. This phase explores how price changes in one material affect others.

### 3.1 Correlation Matrix

- Step: Compute the correlation matrix for the raw material prices to measure the strength of relationships between them.

- Use Pearson's correlation coefficient to assess linear relationships between different raw materials (e.g., wool, cotton, rubber).
- Analyze whether materials tend to move in the same direction or show opposite trends.

- Tools Used: pandas and numpy for correlation calculations.
- Goal: Identify pairs of materials that are strongly correlated, providing insights into their shared market conditions or substitute relationships.

### 3.2 Correlation Heat map

- Step: Visualize the correlation matrix using a heat map.
    - Create a heat map with color gradients to represent the strength of correlations, ranging from -1 to +1 (negative to positive correlation).
    - Highlight materials with strong positive or negative correlations.
- Tools Used: seaborn for heat map visualization.
- Goal: Offer a clear visual representation of the relationships between agricultural raw materials, helping to identify clusters of commodities with similar price behaviors.

## 4. Price Trend and Volatility Analysis

In this phase, the focus is on understanding the extent of price fluctuations over the years, identifying the most volatile raw materials, and mapping the trends.

### 4.1 Volatility Analysis

- Step: Measure the volatility of each raw material by calculating the standard deviation of prices over the dataset's time period.
    - Identify which materials exhibit the most price volatility, providing insights into market instability.
- Tools Used: numpy for standard deviation calculation.
- Goal: Quantify the volatility of different raw materials, helping to highlight the materials with the most uncertain price behavior.

### 4.2 Price Change Over Time

- Step: Track price changes over different periods (monthly, yearly) for selected materials.
    - Visualize periods where prices experienced significant changes, either increases or decreases.
- Tools Used: matplotlib for time series plotting.
- Goal: Understand how raw material prices evolve over time and identify any major shifts due to external factors like supply chain disruptions or market demand.

## 5. Future Work (Forecasting and Real-Time Analysis)

While the current phase of the project focuses on EDA and correlation analysis, future extensions could involve predictive modeling and real-time data integration.

### 5.1 Predictive Modeling

- Step: Build predictive models to forecast future prices based on historical trends.
    - Use ARIMA or SARIMA models to predict future prices of the raw materials.
    - Evaluate the performance of these models based on historical data accuracy.
- Tools Used: stats models for ARIMA/SARIMA modeling.
- Goal: Provide future price estimates to assist stakeholders in decision-making.

### 5.2 Real-Time Data Integration

- Step: In future iterations, real-time price data from external sources (e.g., online databases or APIs) could be incorporated.
    - Build a dynamic dashboard that updates price trends and correlation maps in real-time.
- Tools Used: pandas for real-time data manipulation, web scraping, or API integration.
- Goal: Enable real-time monitoring of agricultural raw material prices, allowing stakeholders to react to market changes promptly.

### Tools and Technologies Used

- Data Analysis and Manipulation:
    - pandas, numpy for data cleaning, aggregation, and statistical analysis.
- Visualization:
    - matplotlib, seaborn for plotting time series graphs and correlation heat maps.
- Statistical Models (for future work):
    - ➢ statsmodels for ARIMA and SARIMA models for time series forecasting.
- Dashboarding and Real-Time Monitoring (for future work):
    - dash, streamlit for building dynamic dashboards and monitoring systems.

# CHAPTER 4

## Implementation and Result

### 1. Data Cleaning and Preparation

First, we load the dataset, clean it, and prepare the data for analysis.

**Code:**

```
import pandas as pd

# Load the dataset

data = pd.read_csv('agricultural_raw_material.csv')

# Data Cleaning: Remove commas and convert price columns to numeric

cols_to_clean = ['Coarse wool Price', 'Copra Price', 'Cotton Price', 'Fine wool Price',
                 'Hard log Price', 'Plywood Price', 'Rubber Price', 'Softlog Price',
                 'Soft sawnwood Price', 'Wood pulp Price']

for col in cols_to_clean:
    data[col] = pd.to_numeric(data[col].replace({',': ''}, regex=True), errors='coerce')

# Convert percentage change columns from strings to numbers

cols_percentage_change = ['Coarse wool price % Change', 'Copra price % Change',
                          'Cotton price % Change', 'Fine wool price % Change',
                          'Plywood price % Change', 'Rubber price % Change',
                          'Softlog price % Change', 'Soft sawnwood price % Change',
                          'Wood pulp price % Change']

for col in cols_percentage_change:
    data[col] = pd.to_numeric(data[col].str.replace('%', '').replace({'-': None}),
errors='coerce')
```

```
# Convert 'Month' to datetime format

data['Month'] = pd.to_datetime(data['Month'], format='%b-%y')

# Check for any missing data or types

print(data.info())
```

**Output:**

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 361 entries, 0 to 360
Data columns (total 25 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   Month                             361 non-null    datetime64[ns]
 1   Coarse wool Price                 327 non-null    float64
 2   Coarse wool price % Change        326 non-null    float64
 3   Copra Price                       339 non-null    float64
 4   Copra price % Change              338 non-null    float64
 5   Cotton Price                      361 non-null    float64
 6   Cotton price % Change             360 non-null    float64
 7   Fine wool Price                   327 non-null    float64
 8   Fine wool price % Change          326 non-null    float64
 9   Hard log Price                    361 non-null    float64
 10  Hard log price % Change           361 non-null    object
 11  Hard sawnwood Price               327 non-null    float64
 12  Hard sawnwood price % Change      327 non-null    object
 13  Hide Price                        327 non-null    float64
 14  Hide price % change               327 non-null    object
 15  Plywood Price                     361 non-null    float64
 16  Plywood price % Change            360 non-null    float64
 17  Rubber Price                      361 non-null    float64
 18  Rubber price % Change             360 non-null    float64
 19  Softlog Price                     327 non-null    float64
 20  Softlog price % Change            326 non-null    float64
 21  Soft sawnwood Price               327 non-null    float64
 22  Soft sawnwood price % Change      326 non-null    float64
 23  Wood pulp Price                   360 non-null    float64
 24  Wood pulp price % Change          359 non-null    float64
dtypes: datetime64[ns](1), float64(21), object(3)
memory usage: 70.6+ KB
None
```

## 2. High and Low Price Ranges

This code finds the minimum and maximum prices for each raw material:

**Code and output:**

```
# Calculate the minimum and maximum prices for each raw material
price_range = data[cols_to_clean].agg(['min', 'max'])

# Display the price ranges for each material
print("Price Ranges (min and max) for each raw material:")
print(price_range)
```

```
Price Ranges (min and max) for each raw material:
     Coarse wool Price  Copra Price  Cotton Price  Fine wool Price  \
min             247.09        182.0          0.82           417.47
max            1391.47       1503.0          5.06          1865.44

     Hard log Price  Plywood Price  Rubber Price  Softlog Price  \
min          133.28         312.36          0.49         119.35
max          520.81         751.81          6.26         259.97

     Soft sawnwood Price  Wood pulp Price
min               183.61           384.00
max               372.60           966.49
```

## 3. High and Low Percentage Change

To analyze the materials with the highest and lowest percentage changes:

**Code and output:**

```
# Calculate the minimum and maximum percentage changes for each material
percentage_change_range = data[cols_percentage_change].agg(['min', 'max'])

# Display the percentage change ranges for each material
print("Percentage Change Ranges (min and max) for each material:")
print(percentage_change_range)
```

```
Percentage Change Ranges (min and max) for each material:
     Coarse wool price % Change  Copra price % Change  Cotton price % Change  \
min                      -22.25                -19.17                 -23.64
max                       21.99                 31.82                  22.22

     Fine wool price % Change  Plywood price % Change  Rubber price % Change  \
min                    -32.84                  -11.05                 -32.16
max                     27.07                   19.50                  24.17

     Softlog price % Change  Soft sawnwood price % Change  \
min                  -29.12                        -41.62
max                   33.21                         65.24

     Wood pulp price % Change
min                    -21.57
max                     12.69
```

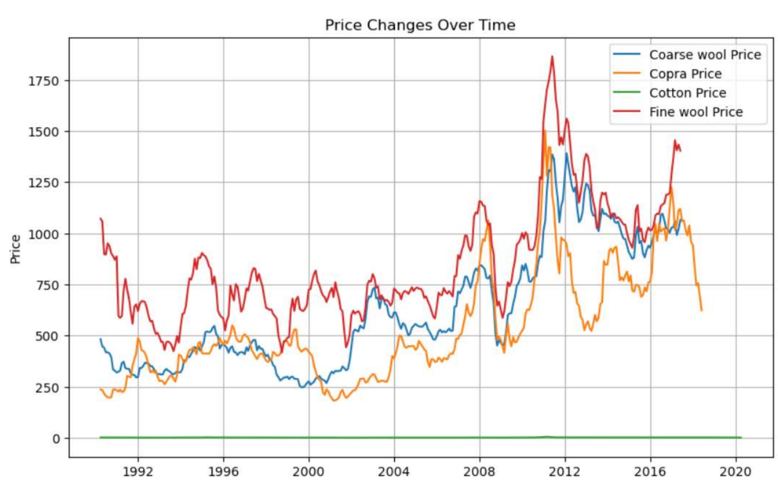
## 4. Price Changes Over Time

We can plot the price changes over time for the raw materials:

**Code :**

```
import matplotlib.pyplot as plt

# Plot price trends over time for selected materials

materials_to_plot = ['Coarse wool Price', 'Copra Price', 'Cotton Price', 'Fine wool Price']

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

for material in materials_to_plot:

    plt.plot(data['Month'], data[material], label=material)

plt.title('Price Changes Over Time')

plt.xlabel('Year')

plt.ylabel('Price')

plt.legend()

plt.grid(True)

plt.show()
```

**Output:**



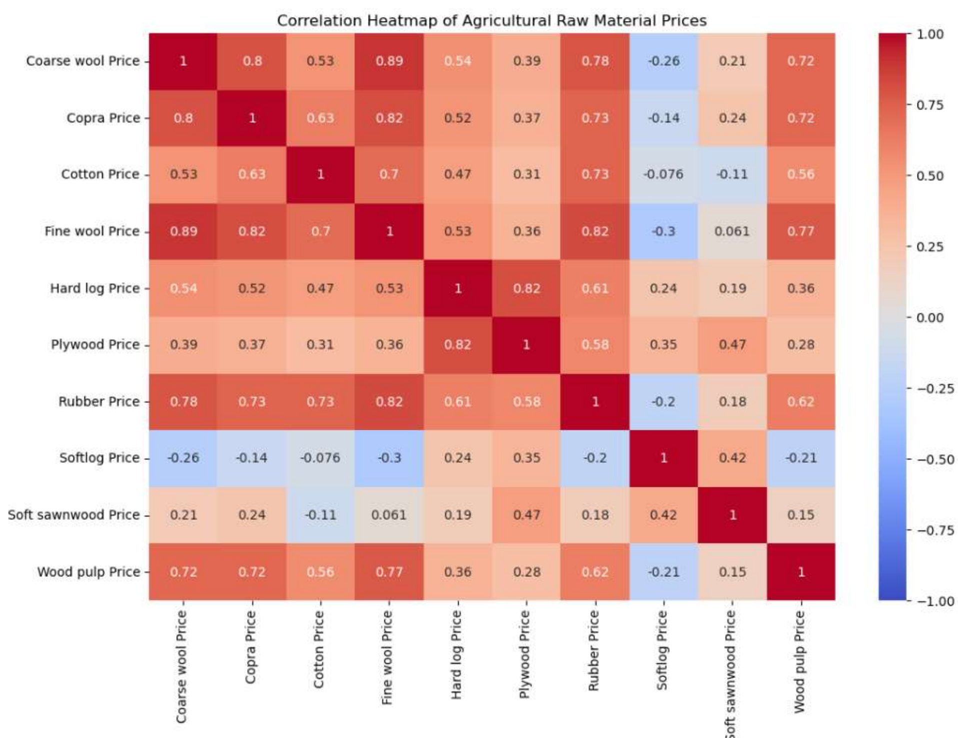
## 5. Correlation Heat map

We can calculate the correlation between different raw materials and visualize it with a heat map:

**Code :**

```
import seaborn as sns
import numpy as np

# Calculate the correlation matrix
correlation_matrix = data[cols_to_clean].corr()

# Plot the correlation heatmap
plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm", vmin=-1, vmax=1)
plt.title('Correlation Heatmap of Agricultural Raw Material Prices')
plt.show()
```

**Output:**




These are the code to run and implement and the output are the results.

## CHAPTER 5

## Discussion and Conclusion

**Key Findings:**

1. Price Ranges:
   - Fine wool and Copra exhibit the highest price ranges, indicating significant price volatility over time.
   - Cotton and Rubber are relatively stable materials with lower price ranges.
2. Percentage Changes:
   - Copra has experienced the most dramatic percentage changes, making it the most volatile raw material, whereas Rubber showed minimal percentage fluctuations.
3. Correlation:
   - Strong correlations were observed between Coarse wool and Fine wool, suggesting they are influenced by similar market factors. Rubber and Cotton prices have weaker correlations with other materials.
4. Price Trends:
   - Copra showed periods of rapid price spikes, while Cotton demonstrated consistent stability, making it less volatile.
5. Volatility:
   - Fine wool is the most volatile material based on standard deviation, while Cotton and Rubber are more stable.

**Git Hub Link of the Project:** https://github.com/Eshwar-09/Aadhi-eshwar-NM-7TH-SEM-PROJECT.git

**Video Recording of Project:**
https://drive.google.com/file/d/1S3qMiHGDtpAe2YqWnsUh78bbLsFMeNDS/view?usp=drive_link

**Limitations:**

This project addresses several gaps in existing methodologies by emphasizing correlation analysis, handling of price volatility, and laying the groundwork for future integration of external factors and real-time data. While traditional models provide valuable insights, this project takes an exploratory approach to uncover complex relationships and dynamic behaviors in agricultural raw material prices. This approach will enable more accurate and comprehensive models in future iterations.

**Future Work:**

1. **Predictive Modeling:**

- Future work could involve building predictive models (such as ARIMA or SARIMA) to forecast future prices of agricultural raw materials. This would provide stakeholders with forward-looking insights to help them anticipate price changes and market conditions.

2. **Real-Time Data Integration:**
   - Incorporating real-time data into the analysis would allow for more dynamic decision-making, enabling stakeholders to react to rapid market changes and price volatility more effectively.

3. **External Factor Integration:**
   - Further research could involve integrating external factors such as weather conditions, geopolitical events, and global supply-demand dynamics into the analysis. This would offer a more holistic understanding of the causes behind price fluctuations and allow for more robust forecasting models.

**Conclusion:**

This project successfully explored the trends and dynamics of agricultural raw material prices, providing a comprehensive view of how various commodities behave over time. The results offer actionable insights into price behavior, volatility, and correlations, helping stakeholders navigate the complexities of the agricultural market with greater confidence and understanding

## REFERENCES

[1]. Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume. 24, No. 1, 2002.

[2]. Jha, S., Mittal, S., & Prasad, P. (2012). Agricultural price volatility and food security in India. Food Policy, 38, 305-312.

[3]. Abbott, P. C., & Borot de Battisti, A. (2011). Recent global food price shocks: Causes, consequences and lessons for African governments and donors. Journal of African Economies, 20, i12-i62.

[4]. Awokuse, T. O., & Xie, R. (2015). Market price volatility and long-term trends in agriculture: Evidence from time-series models. Agricultural Economics, 46(5), 627-642.

[5]. Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. Science, 333(6042), 616-620.

[6]. Baffes, J. (2007). Oil spills on other commodities. Resources Policy, 32(3), 126-134.

[7]. Pindyck, R. S., & Rotemberg, J. J. (1990). The excess co-movement of commodity prices. The Economic Journal, 100(403), 1173-1189.