ARTIFICAL INTELLIGENCE AND MACHINE LEARNING FUNDAMENTALS WITH CLOUD COMPUTING AND GEN AI BY MICROSOFT

"Agricultural Raw Material Analysis"

By

AKASH A (810021114005)

akashabraham3955@gmail.com

UNIVERSITY COLLEGE OF ENGINEERING TRICHY

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.

ABSTRACT OF THE PROJECT

This project aims to develop an efficient, sustainable system for managing agricultural raw materials, focusing on quality, availability, and ecological impact. As global demand for food and agricultural products rises, there is a critical need to enhance the production, storage, and distribution of raw materials such as grains, fruits, vegetables, and other essential crops. This project will explore sustainable practices, including optimized resource use, reduction of waste, and environmental conservation, to improve raw material quality and availability.

We will employ innovative technologies like precision agriculture, IoT-enabled monitoring, and data-driven supply chain management to enhance productivity and ensure that raw materials meet high standards. Additionally, the project will assess alternative raw materials that may be more resilient and sustainable in the long term. Our goal is to create a model that supports economic growth in the agriculture sector while promoting environmental responsibility and sustainable resource use. This project has the potential to benefit farmers, industry stakeholders, and consumers by ensuring a stable, high-quality supply of agricultural raw materials, essential for food security and economic development.

TABLE OF CONTENTS

CHAPTER 1

Introduction

1.1 Background on Agricultural Raw Materials

Agricultural raw materials, which include products like grains, oils, and fibers, are fundamental components of the global economy. They serve as essential inputs for various industries, including food production, biofuels, and textiles. Understanding their price trends is critical for stakeholders, such as farmers, investors, and policymakers, who rely on these insights to make informed decisions. The agricultural commodity market is known for its volatility, driven by factors like seasonal weather changes, trade policies, and global supply-demand shifts.

1.2. Importance of Price Analysis

Analyzing the prices of agricultural raw materials over time offers insight into their market behavior, including stability, trends, and potential correlations with other materials. Price fluctuations can affect the cost of production, consumer prices, and ultimately, the global economy. By examining historical price data, we can better understand which materials are prone to price volatility and identify stable commodities, thereby providing guidance for industry participants.

1.3. Objectives of the Analysis

The primary goals of this analysis are to:

- Identify high and low price ranges of different raw materials over the years.
- Calculate the percentage changes in prices to understand which materials are most and least volatile
- Explore how price ranges have shifted over time, identifying any trends or anomalies.

1.4. Scope of Exploratory Data Analysis (EDA)

This analysis involves applying exploratory data analysis (EDA) techniques to uncover insights from the agricultural raw material prices dataset. EDA will help:

- Detect patterns or trends in price changes.
- Quantify volatility across materials to distinguish between stable and fluctuating commodities.
- Understand interdependencies between materials, which may help forecast future price behaviors.

1.5. Significance of Findings

Insights from this analysis can guide decision-making in areas like resource allocation, risk management, and market forecasting. For instance, identifying stable materials could be beneficial for manufacturers seeking predictable costs, while understanding volatile commodities may be valuable for traders and investors looking to capitalize on price swings. Additionally, policymakers may use these findings to shape policies that stabilize food prices, thus contributing to food security

CHAPTER 2

Analyze agricultural raw material prices

For an Agricultural Raw Material Analysis using an exploratory data analysis (EDA) approach, the goal is to thoroughly explore and understand patterns, trends, and relationships within the dataset containing prices of various raw materials over time. Here's an outline of the main steps and key insights you would seek in this analysis:

1. Data Preparation and Cleaning

- Load and Inspect the Dataset: Begin by loading the data and taking an initial look at its structure. Understand what each column represents, such as raw material name, year, and price. Ensure that all necessary fields are present.
- Data Cleaning: Check for and handle any missing values or outliers, which might distort results. Data types should be verified and converted if necessary, especially date-related fields. Ensure that the data is in a tidy format, with each row representing a unique material and year combination.

2. Descriptive Analysis of Price Ranges

- High and Low Price Ranges: Identify the highest and lowest prices observed for each raw material over the years. This step reveals the relative value of different raw materials, showing which materials are typically high-priced or low-priced in the market.
- High Range Materials: These are materials that generally maintain a high market value, possibly due to high demand or limited supply.
- Low Range Materials: Materials with consistently lower prices, which might indicate an abundant supply or lesser demand.
- Insights: Describe the variations in price ranges across materials. Are there materials with very high or low prices compared to the average? Are some materials stable while others show significant fluctuations?

3. Analysis of Price Volatility (Percentage Change)

- Calculate Yearly % Change: For each raw material, compute the percentage change in price from one year to the next. This metric reveals how much a material's price fluctuates year-over-year.

- High % Change Materials: Materials with frequent or large price swings, indicating high volatility. High volatility may point to sensitivity to economic conditions, seasonality, or other external factors.
- Low % Change Materials: Materials with minimal price changes over time, showing more stability. These materials could be staples in the market with steady supply and demand.

- Insights: Discuss the economic or market factors that might drive these changes. For instance, weather events, trade policies, or shifts in demand can significantly impact certain materials.

4. Trend Analysis Over the Years

- Time Series Analysis: Plot price trends for each material over the years to understand the trajectory of prices. Line plots can reveal seasonal patterns, periodic peaks, or troughs, and highlight years with notable changes.
- Range of Price Changes Over Time: Calculate the range (difference between maximum and minimum) of prices over the years for each material, helping to identify which materials have stable prices versus those with wide variability.
- Insights: Determine if specific periods correlate with significant changes in price. For example, a particular material may show sharp increases during certain years due to economic conditions, climate impacts, or production changes.

5. Correlation Analysis

- Correlation Matrix: To examine if prices of different raw materials move together, create a correlation matrix. This matrix quantifies the relationship between prices of different materials, showing which pairs of materials have positive or negative price correlations.
- Heatmap Visualization: Use a heatmap to represent the correlation matrix visually. This helps quickly identify clusters of materials with similar price trends or contrasting movements.

- Positive Correlation: Materials with high positive correlations may be substitutes or complements. For example, if two crops compete for land, a price increase in one might drive up the price of the other.
- Negative Correlation: Materials with negative correlations might indicate complementary demand dynamics or production trade-offs.
- Interpretation: Discuss possible reasons behind observed correlations, such as cross-demand dependencies, supply chain linkages, or regional production patterns.

Chapter 3

High range and low range materials

Data Loading and Inspection

- Load the Dataset: Import the dataset into a DataFrame.
- Preview the Data: Display the first few rows to understand its structure and columns.
- Check Data Types: Ensure Date is in datetime format, and Price is a numeric type.
- Basic Summary: Get an overview with basic statistics (mean, median, min, max) of prices to understand initial price ranges.

Data Cleaning

- Handle Missing Values: Identify any missing values, especially in Price, and decide on a strategy (e.g., imputation or removal).
- Outlier Detection: Detect outliers in price data using statistical techniques (e.g., Z-score, IQR).
- Data Type Conversion: Ensure all columns are in the correct format for analysis.
- Filtering Irrelevant Data: Remove any entries that might skew the analysis, such as erroneous or incomplete records.

Exploratory Data Analysis (EDA)

1. Yearly Price Range Calculation

- Aggregate by Year: Group the dataset by Year and Material Type.
- Calculate Price Range: For each year and material type, calculate the price range:

Range = Maximum Price - Minimum Price

- Summarize Range Data: Create a summary table showing yearly price ranges for each raw material.

2. Trend Analysis

- Visualize Yearly Price Ranges: Use line charts to show the yearly price range for each raw material over time.

- Identify Trends: Note periods of high or low volatility and analyze whether certain raw materials consistently show wide price ranges compared to others.

4. Comparison Across Materials

- Compare Price Ranges: Analyze and compare yearly price ranges across different raw materials to determine which materials have the most stable or volatile prices.
- Box Plots: Use box plots to visualize the distribution of yearly price ranges for each material, showing variations over time.

Chapter 4
The range of prices changed over the years

To identify the range of prices changed over the years for each agricultural raw material, we'll need to calculate the price range (i.e., the difference between the highest and lowest prices) for each material over the available years. Here's how to approach this:

Steps to Identify Price Range Over the Years

1. Load and Preprocess the Data

- Load the dataset and check for any missing values or data inconsistencies.
- Ensure the date or year column is correctly formatted, so we can analyze prices over time.

2. Calculate Price Range for Each Material

- Group the data by each raw material.
- For each material, identify the maximum and minimum price over the years.
- Calculate the price range as the difference between the maximum and minimum prices.

3. Summarize the Results

- Create a summary table with each material's name, maximum price, minimum price, and price range.
- Sort the materials based on their price range to identify the materials with the highest and lowest range of price changes.

Example Code for Calculating Price Range

Here is an example using Python, assuming the data is in a DataFrame named data with columns like Material, Year, and Price.

```
import pandas as pd

# Load the dataset
data =
pd.read_csv('agricultural_raw_material_prices.csv') #
Adjust filename as needed

# Calculate max, min, and range for each material
price_range = data.groupby('Material')['Price'].agg(
   max_price='max',
   min_price='min'
)
price_range['price_range'] = price_range['max_price'] -
price_range['min_price']

# Sort by price range to see which materials had the
highest and lowest fluctuations
price_range_sorted =
price_range.sort_values(by='price_range',
ascending=False)

# Display the top materials with highest and lowest
price ranges
print("Materials with the highest price range:")
print(price_range_sorted.head(10))
print("\nMaterials with the lowest price range:")
print(price_range_sorted.tail(10))
```

Output Example

The output table will include columns like:

Material	Max Price	Min Price	Price Range
Wheat	200	100	100
Corn	180	90	90
Coffee Beans	300	150	150
...	...	...	...

Interpreting the Results

- High Price Range Materials indicate materials with significant price volatility, potentially influenced by supply, demand, or external factors.

- Low Price Range Materials are generally more stable in pricing, which might suggest less sensitivity to market changes or steady demand.

Preview Code Blame Raw

Agricultural Raw material prices Case Study (1990-2020)

The basic purpose of my case study is to have a basic understanding of the data and also identify any reason for the price change. A better understanding of these prices will allow us to predict price hikes

Load the Libraries Required

In [1]:
```
# Import required packages
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.impute import SimpleImputer
sns.set(rc={'figure.figsize':(11, 4)})
```

Exploring the Dataset

In [13]:
```
def overview():
    df = pd.read_csv("agricultural_raw.csv")
    print("The first 5 rows of data are:\n")
    print(df.head)
    print("\n\n\nDataset has {} rows and {} columns".format(df.shape[0], df.shape[1]))
    print("\n\n\nDatatype: \n")
    print(df.dtypes)
    print("\n\n\nThe number of null values for each column are: \n")
    print(df.isnull().sum())
    print("\n\n\nData summary: \n")
    print(df.describe())
    return df
```

In [14]:
```
# assigning a variable to overview()
df = overview()
```

```
The first 5 rows of data are:

<bound method NDFrame.head of       Month Coarse
wool Price Coarse wool price % Change Copra Price
\
0     Apr-90            482.34
-          236
1     May-90            447.26
-7.27%          234
2     Jun-90            440.99
-1.40%          216
3     Jul-90            418.44
-5.11%          205
4     Aug-90            418.44
0.00%          198
..        ...              ...
...       ...
356   Dec-19               NaN
NaN          NaN
357   Jan-20               NaN
NaN          NaN
358   Feb-20               NaN
NaN          NaN
359   Mar-20               NaN
NaN          NaN
360   Apr-20               NaN
NaN          NaN
```

```
    Copra price % Change  Cotton Price Cotton pri
ce % Change Fine wool Price  \
0                     -         1.83
-          1,071.63
1                -0.85%         1.89
3.28%          1,057.18
2                -7.69%         1.99
5.29%            898.24
3                -5.09%         2.01
1.01%            895.83
4                -3.41%         1.79
-10.95%            951.22
..                  ...          ...
...              ...
356                 NaN         1.67
1.21%              NaN
357                 NaN         1.74
4.19%              NaN
358                 NaN         1.69
-2.87%              NaN
359                 NaN         1.49
-11.83%              NaN
360                 NaN         1.40
-6.04%              NaN

    Fine wool price % Change  Hard log Price  ...
Plywood Price  \
0                          -          161.20  ...
312.36
1                     -1.35%          172.86  ...
350.12
2                    -15.03%          181.67  ...
373.94
3                     -0.27%          187.96  ...
378.48
4                      6.18%          186.13  ...
364.60
..                       ...             ...  ...
...
356                      NaN          272.80  ...
500.37
357                      NaN          272.40  ...
499.64
358                      NaN          270.56  ...
496.28
359                      NaN          276.93  ...
507.96
360                      NaN          276.24  ...
506.68
```

```
     Plywood price % Change Rubber Price  Rubber
price % Change Softlog Price   \
0                             -             0.84
-            120.66
1                        12.09%             0.85
1.19%            124.28
2                         6.80%             0.85
0.00%            129.45
3                         1.21%             0.86
1.18%            124.23
4                        -3.67%             0.88
2.33%            129.70
..                          ...              ...
...            ...
356                      -0.22%             1.66
7.79%               NaN
357                      -0.15%             1.68
1.20%               NaN
358                      -0.67%             1.61
-4.17%               NaN
359                       2.35%             1.50
-6.83%               NaN
360                      -0.25%             1.33
-11.33%               NaN

     Softlog price % Change Soft sawnwood Price
Soft sawnwood price % Change   \
0                             -                  218.76
-
1                         3.00%                  213.00
-2.63%
2                         4.16%                  200.00
-6.10%
3                        -4.03%                  210.05
5.03%
4                         4.40%                  208.30
-0.83%
..                          ...                     ...
...
356                         NaN                     NaN
NaN
357                         NaN                     NaN
NaN
358                         NaN                     NaN
NaN
359                         NaN                     NaN
NaN
360                         NaN                     NaN
NaN
```

```
     Wood pulp Price  Wood pulp price % Change
0             829.29                         -
1             842.51                     1.59%
2             831.35                    -1.32%
3             798.83                    -3.91%
4             818.74                     2.49%
..               ...                       ...
356           875.00                     0.00%
357           875.00                     0.00%
358           875.00                     0.00%
359           875.00                     0.00%
360              NaN                       NaN

[361 rows x 25 columns]>


Dataset has 361 rows and 25 columns


Datatype:

Month                                object
Coarse wool Price                    object
Coarse wool price % Change           object
Copra Price                          object
Copra price % Change                 object
Cotton Price                        float64
Cotton price % Change                object
Fine wool Price                      object
Fine wool price % Change             object
Hard log Price                      float64
Hard log price % Change              object
Hard sawnwood Price                 float64
Hard sawnwood price % Change         object
Hide Price                          float64
Hide price % change                  object
Plywood Price                       float64
Plywood price % Change               object
Rubber Price                        float64
Rubber price % Change                object
Softlog Price                       float64
Softlog price % Change               object
Soft sawnwood Price                 float64
Soft sawnwood price % Change         object
Wood pulp Price                     float64
Wood pulp price % Change             object
dtype: object
```

```
Data summary:

       Cotton Price  Hard log Price  Hard sawnwoo
d Price  Hide Price  \
count    361.000000      361.000000            32
7.000000  327.000000
mean       1.640000      251.034072            70
7.950367   78.566667
std        0.513319       65.628406            14
4.563241   13.690623
min        0.820000      133.280000            41
3.370000   28.590000
25%        1.290000      197.960000            57
3.470000   69.495000
50%        1.600000      253.010000            72
8.710000   77.250000
75%        1.850000      282.970000            83
1.635000   86.000000
max        5.060000      520.810000            97
3.600000  114.630000

       Plywood Price  Rubber Price  Softlog Price
Soft sawnwood Price  \
count     361.000000    361.000000     327.000000
327.000000
mean      508.216122      1.656427     164.527462
291.061713
std        89.274718      1.017086      25.596308
34.113959
min       312.360000      0.490000     119.350000
183.610000
25%       442.540000      0.860000     145.970000
277.590000
50%       505.040000      1.440000     160.370000
294.960000
75%       570.790000      2.060000     180.210000
310.865000
max       751.810000      6.260000     259.970000
372.600000
```

```
           Wood pulp Price
count           360.000000
mean            696.670889
std             161.156936
min             384.000000
25%             549.777500
50%             693.580000
75%             875.000000
max             966.490000
```

In [15]:

```
# checking the Data shape
df.shape
```

Out[15]:

```
(361, 25)
```

In [16]:

```
df.info
```

Out[16]:

```
<bound method DataFrame.info of         Month Coars
e wool Price Coarse wool price % Change Copra Pri
ce  \
0     Apr-90            482.34
-          236
1     May-90            447.26
-7.27%          234
2     Jun-90            440.99
-1.40%          216
3     Jul-90            418.44
-5.11%          205
4     Aug-90            418.44
0.00%          198
..        ...               ...
...           ...
356   Dec-19               NaN
NaN           NaN
357   Jan-20               NaN
NaN           NaN
358   Feb-20               NaN
NaN           NaN
359   Mar-20               NaN
NaN           NaN
360   Apr-20               NaN
NaN           NaN
```

```
     Copra price % Change  Cotton Price Cotton pri
ce % Change Fine wool Price  \
0                       -          1.83
-         1,071.63
1                  -0.85%          1.89
3.28%         1,057.18
2                  -7.69%          1.99
5.29%           898.24
3                  -5.09%          2.01
1.01%           895.83
4                  -3.41%          1.79
-10.95%           951.22
..                    ...           ...
...              ...
356                   NaN          1.67
1.21%              NaN
357                   NaN          1.74
4.19%              NaN
358                   NaN          1.69
-2.87%             NaN
359                   NaN          1.49
-11.83%             NaN
360                   NaN          1.40
-6.04%             NaN

    Fine wool price % Change  Hard log Price  ...
Plywood Price  \
0                         -          161.20  ...
312.36
1                    -1.35%          172.86  ...
350.12
2                   -15.03%          181.67  ...
373.94
3                    -0.27%          187.96  ...
378.48
4                     6.18%          186.13  ...
364.60
..                      ...             ...  ...
...
356                     NaN          272.80  ...
500.37
357                     NaN          272.40  ...
499.64
358                     NaN          270.56  ...
496.28
359                     NaN          276.93  ...
507.96
360                     NaN          276.24  ...
506.68
```

```
      Plywood price % Change Rubber Price  Rubber
price % Change Softlog Price   \
0                           -            0.84
-         120.66
1                      12.09%            0.85
1.19%          124.28
2                       6.80%            0.85
0.00%          129.45
3                       1.21%            0.86
1.18%          124.23
4                      -3.67%            0.88
2.33%          129.70
..                        ...             ...
...            ...
356                    -0.22%            1.66
7.79%             NaN
357                    -0.15%            1.68
1.20%             NaN
358                    -0.67%            1.61
-4.17%            NaN
359                     2.35%            1.50
-6.83%            NaN
360                    -0.25%            1.33
-11.33%            NaN

      Softlog price % Change Soft sawnwood Price
Soft sawnwood price % Change   \
0                           -                 218.76
-
1                       3.00%                 213.00
-2.63%
2                       4.16%                 200.00
-6.10%
3                      -4.03%                 210.05
5.03%
4                       4.40%                 208.30
-0.83%
..                        ...                    ...
...
356                       NaN                    NaN
NaN
357                       NaN                    NaN
NaN
358                       NaN                    NaN
NaN
359                       NaN                    NaN
NaN
360                       NaN                    NaN
NaN
```

```
     Wood pulp Price  Wood pulp price % Change
0             829.29                         -
1             842.51                     1.59%
2             831.35                    -1.32%
3             798.83                    -3.91%
4             818.74                     2.49%
..               ...                       ...
356           875.00                     0.00%
357           875.00                     0.00%
358           875.00                     0.00%
359           875.00                     0.00%
360              NaN                       NaN

[361 rows x 25 columns]>
```

looking for the Null values

In [17]:
```
#Checking Null Values of each column
df.isnull().sum()
```

Out[17]:
```
Month                             0
Coarse wool Price                34
Coarse wool price % Change       34
Copra Price                      22
Copra price % Change             22
Cotton Price                      0
Cotton price % Change             0
Fine wool Price                  34
Fine wool price % Change         34
Hard log Price                    0
Hard log price % Change           0
Hard sawnwood Price              34
Hard sawnwood price % Change     34
Hide Price                       34
Hide price % change              34
Plywood Price                     0
Plywood price % Change            0
Rubber Price                      0
Rubber price % Change             0
Softlog Price                    34
Softlog price % Change           34
Soft sawnwood Price              34
Soft sawnwood price % Change     34
Wood pulp Price                   1
Wood pulp price % Change          1
dtype: int64
```

replacing Null, NaN values

In [18]:

```
# Replacing %, "," and "-"
df = df.replace('%', '', regex=True)
df = df.replace(',', '', regex=True)
df = df.replace('-', '', regex=True)
df = df.replace('', np.nan)
df = df.replace('MAY90', np.nan)
```

In [19]:

```
# Dropping rows with NaN values
df = df.dropna()
```

In [20]:

```
# Check to see if all NaN values are resolved
df.isnull().sum()
```

Out[20]:

```
Month                              0
Coarse wool Price                  0
Coarse wool price % Change         0
Copra Price                        0
Copra price % Change               0
Cotton Price                       0
Cotton price % Change              0
Fine wool Price                    0
Fine wool price % Change           0
Hard log Price                     0
Hard log price % Change            0
Hard sawnwood Price                0
Hard sawnwood price % Change       0
Hide Price                         0
Hide price % change                0
Plywood Price                      0
Plywood price % Change             0
Rubber Price                       0
Rubber price % Change              0
Softlog Price                      0
Softlog price % Change             0
Soft sawnwood Price                0
Soft sawnwood price % Change       0
Wood pulp Price                    0
Wood pulp price % Change           0
dtype: int64
```

Converting data types

In [21]:
```
# Converting data type to float
lst = ["Coarse wool Price", "Coarse wool price % Change", "Copra Price", "Copra price % Change", "Cotto
df[lst] = df[lst].astype("float")
df.dtypes
```

Out[21]:
```
Month                              object
Coarse wool Price                 float64
Coarse wool price % Change        float64
Copra Price                       float64
Copra price % Change              float64
Cotton Price                      float64
Cotton price % Change             float64
Fine wool Price                   float64
Fine wool price % Change          float64
Hard log Price                    float64
Hard log price % Change           float64
Hard sawnwood Price               float64
Hard sawnwood price % Change      float64
Hide Price                        float64
Hide price % change               float64
Plywood Price                     float64
Plywood price % Change            float64
Rubber Price                      float64
Rubber price % Change             float64
Softlog Price                     float64
Softlog price % Change            float64
Soft sawnwood Price               float64
Soft sawnwood price % Change      float64
Wood pulp Price                   float64
Wood pulp price % Change          float64
dtype: object
```

Date time col Formating

Setting up this col as an index for the entire dataset

In [22]:
```
df.Month  = pd.to_datetime(df.Month.str.upper(), format='%b%y', yearfirst=False)
# Indexing month
df = df.set_index('Month')
```

In [23]:
```
df.head()
```

Out[23]:

Month	Coarse wool Price	Coarse wool price % Change	Copra Price	Copra price % Change	Cotton Price	Cotton price % Change	Fine wool Price	Fine wool price % Change	Hard log Price	Hard log price % Change	...	Plywood Price	Plywood price % Change	Rubber Price
1990-05-01	447.26	7.27	234.0	0.85	1.89	3.28	1057.18	1.35	172.86	7.23	...	350.12	12.09	0.85
1990-06-01	440.99	1.40	216.0	7.69	1.99	5.29	898.24	15.03	181.67	5.10	...	373.94	6.80	0.85
1990-07-01	418.44	5.11	205.0	5.09	2.01	1.01	895.83	0.27	187.96	3.46	...	378.48	1.21	0.86
1990-08-01	418.44	0.00	198.0	3.41	1.79	10.95	951.22	6.18	186.13	0.97	...	364.60	3.67	0.88
1990-09-01	412.18	1.50	196.0	1.01	1.79	0.00	936.77	1.52	185.33	0.43	...	384.92	5.57	0.90

5 rows × 24 columns

Exploratory Analysis and Visualization

First, To ready the required enviroments

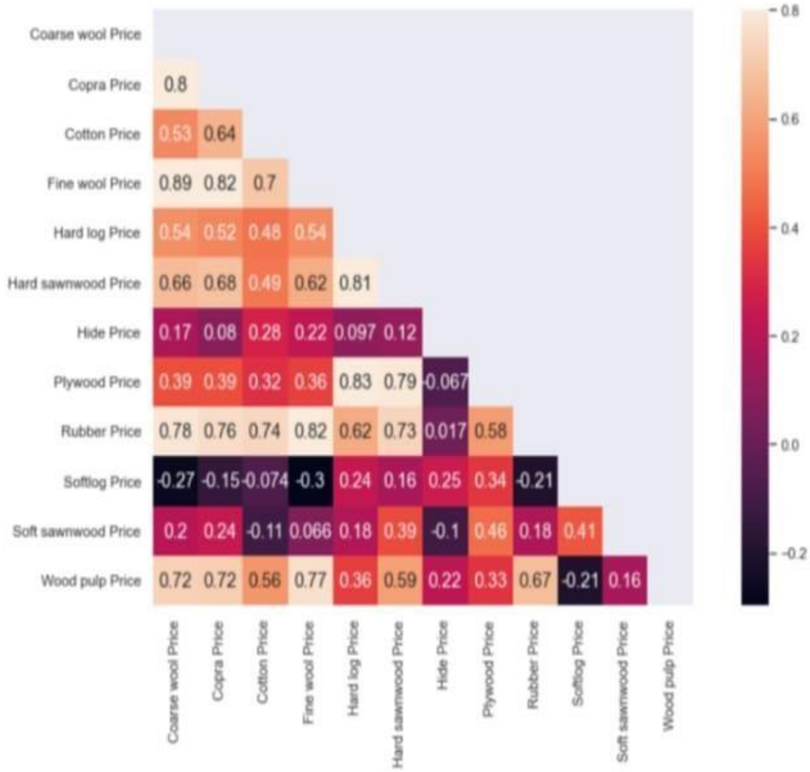
In [24]:

```
import seaborn as sns
import matplotlib
import matplotlib.pyplot as plt
%matplotlib inline
sns.set_style('darkgrid')
matplotlib.rcParams['font.size'] = 14
matplotlib.rcParams['figure.figsize'] = (9, 5)
matplotlib.rcParams['figure.facecolor'] = '#00000000'
```

Heat Map

In [25]:

```
#raw-materials list
raw_data=['Coarse wool Price',  'Copra Price','Cotton Price', 'Fine wool Price',  'Hard log Price', 'Ha
 'Hide Price', 'Plywood Price', 'Rubber Price', 'Softlog Price', 'Soft sawnwood Price', 'Wood pulp Pric

#getting the correlation matrix
corrmat = df[raw_data].corr()

#setting the size of plot
fig = plt.figure(figsize = (14, 8))

#masking the upper traingle part since matrix is symmetric(repetitive)
mask = np.triu(np.ones_like(corrmat, dtype=bool))
sns.heatmap(corrmat, vmax = .8,mask=mask, square = True, annot = True)
plt.show()
```

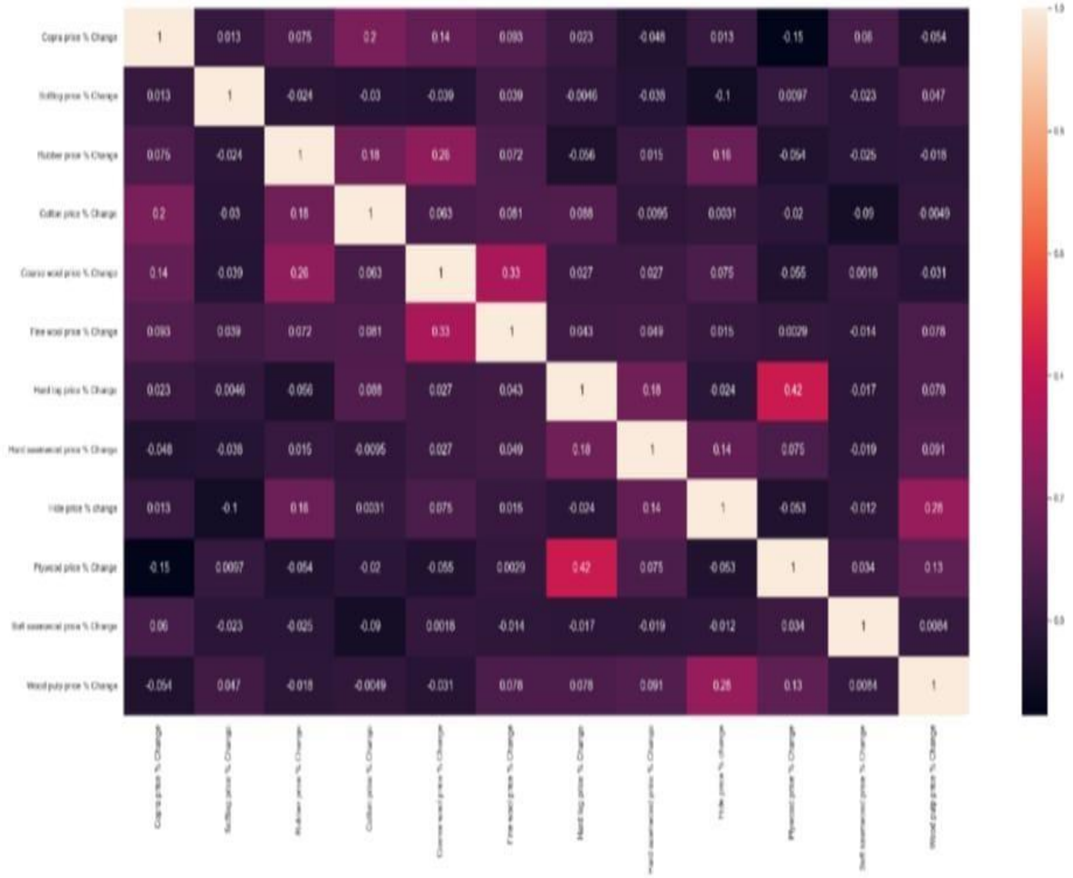


The Heatmap depicts correlation between the raw-materials:

higher the correlated value higher chance of being two raw-materials related but not necessarily.

In [26]:

```
plt.figure(figsize=(28,14))
changelist=['Copra price % Change','Softlog price % Change','Rubber price % Change','Cotton price % Cha

#generate a correlation matrix for the whole dataset
corrMatrix = df[changelist].corr()
sns.heatmap(corrMatrix, annot=True)
plt.show()
```



From the above plots we can say that there is almost no relation between % change of raw-material prices

- The negative value implies two variables are negatively correlated (one increase,other decrease)

- Zero implies no relation

- other wise higher the value higher the chance of relation.

prices and their % change plots

Coarse wool

In [27]:
```
axes=df[["Coarse wool Price", "Coarse wool price % Change"]].plot(figsize=(11, 9), subplots=True, linew
```

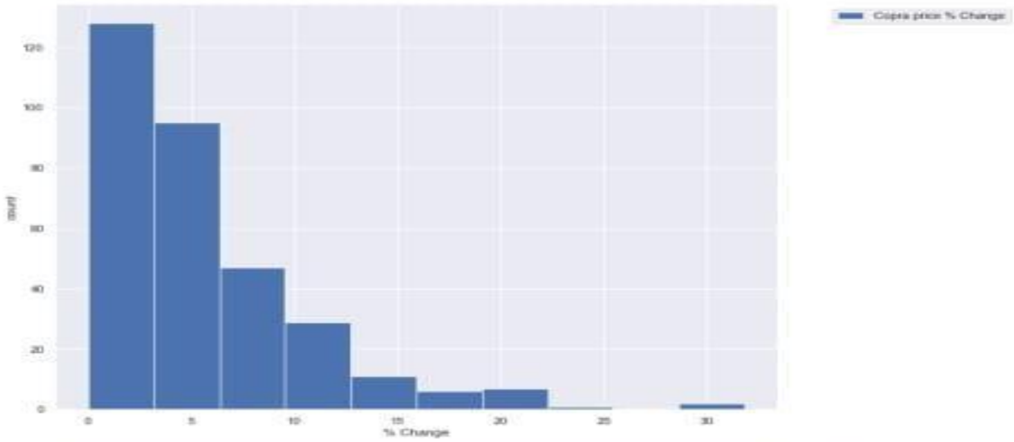
In [29]:
```
axes=df[["Fine wool Price", "Fine wool price % Change"]].plot(figsize=(11, 9), subplots=True, linewidth
```

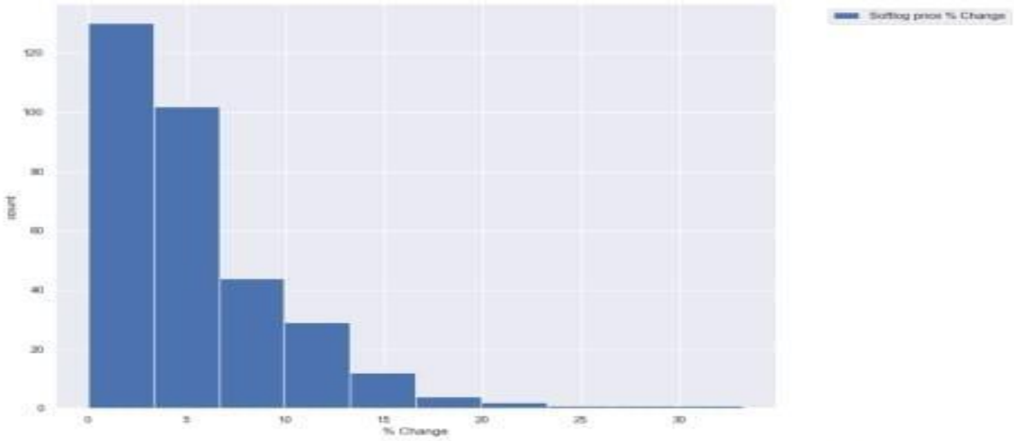
Similarly we could find the visualization of all other features given in dataset.

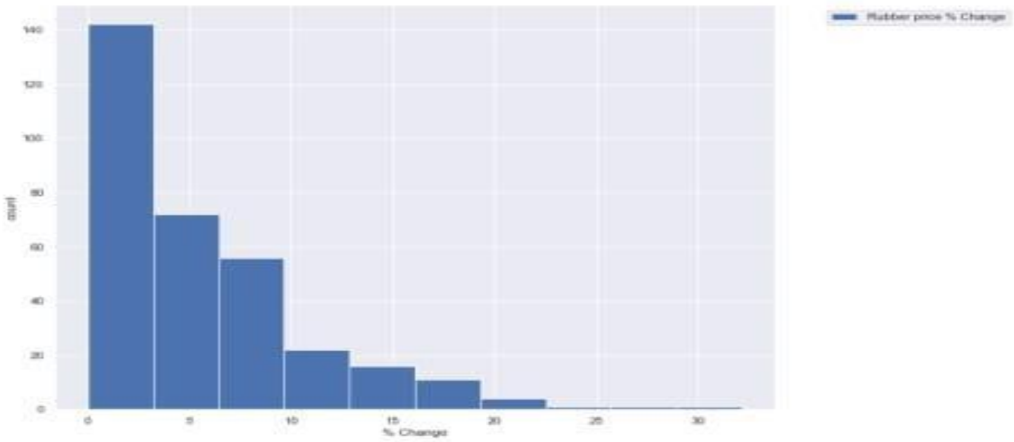
Asking and Answering Questions

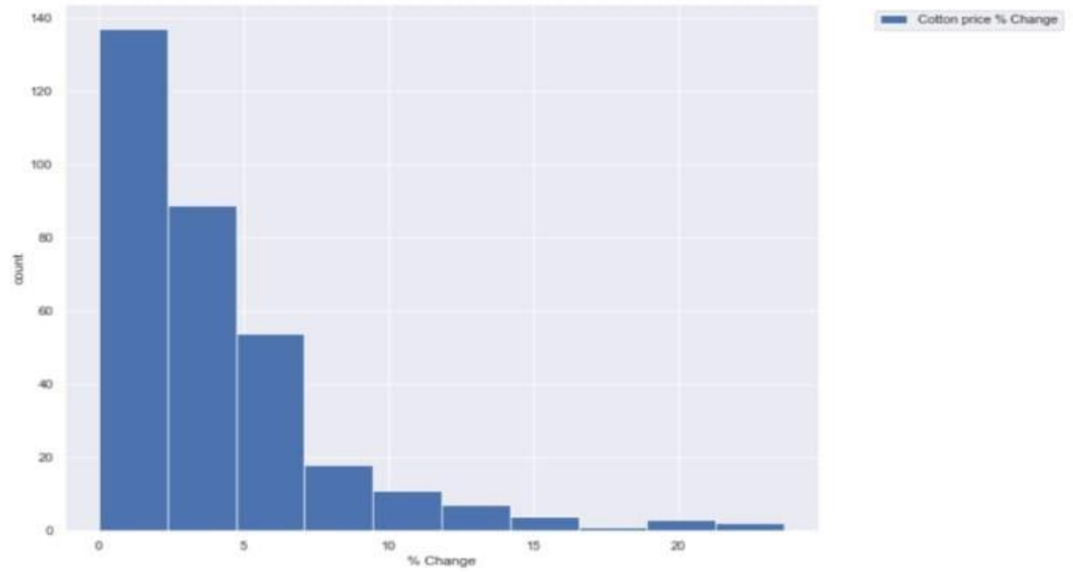
Q1: Find out the normal price change for each raw material

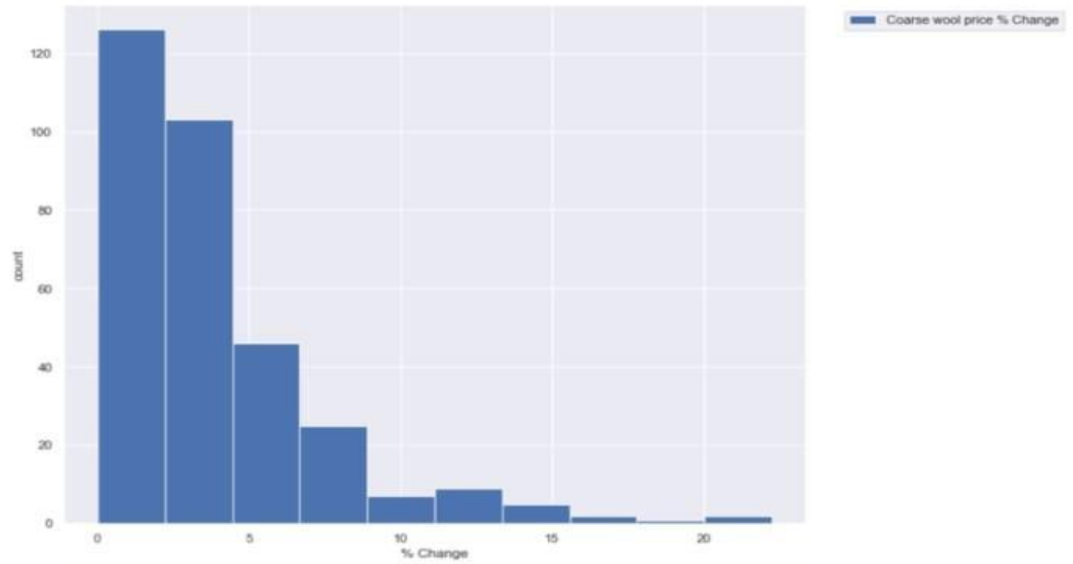
In [30]:
```python
#df[['Wood pulp Price', 'Wood pulp price % Change']].hist(figsize=(11, 9), linewidth=1)

changelist=['Copra price % Change','Softlog price % Change','Rubber price % Change','Cotton price % Cha
for i in range(len(changelist)):
    plt.figure(figsize=(12,12))
    df[changelist[i]].hist(figsize=(11, 9), linewidth=1)
    plt.xlabel('% Change')
    plt.ylabel('count')
    plt.legend(changelist[i:],loc='upper center',bbox_to_anchor=(1.2,1))
```

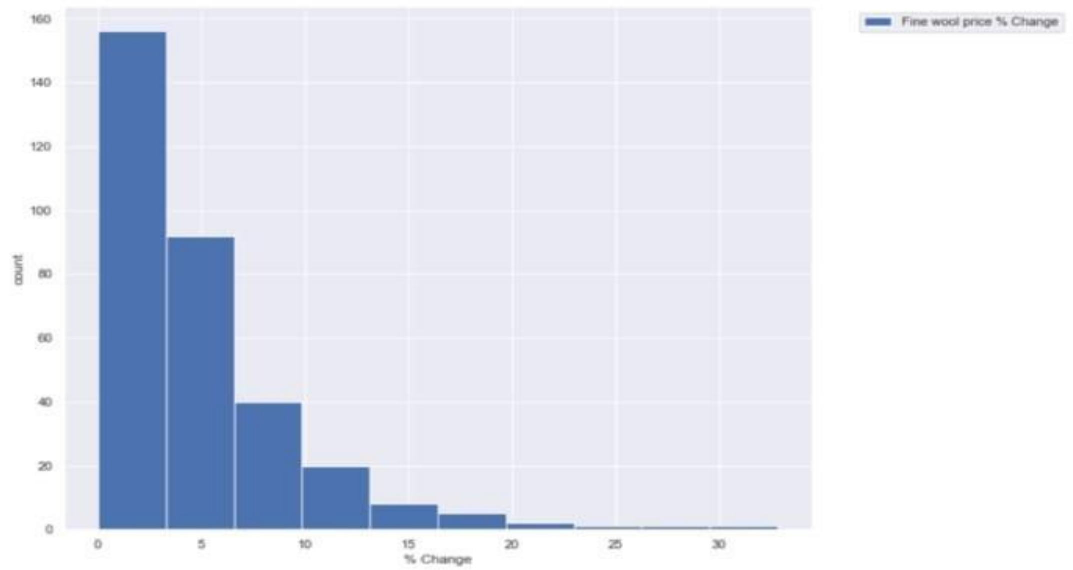


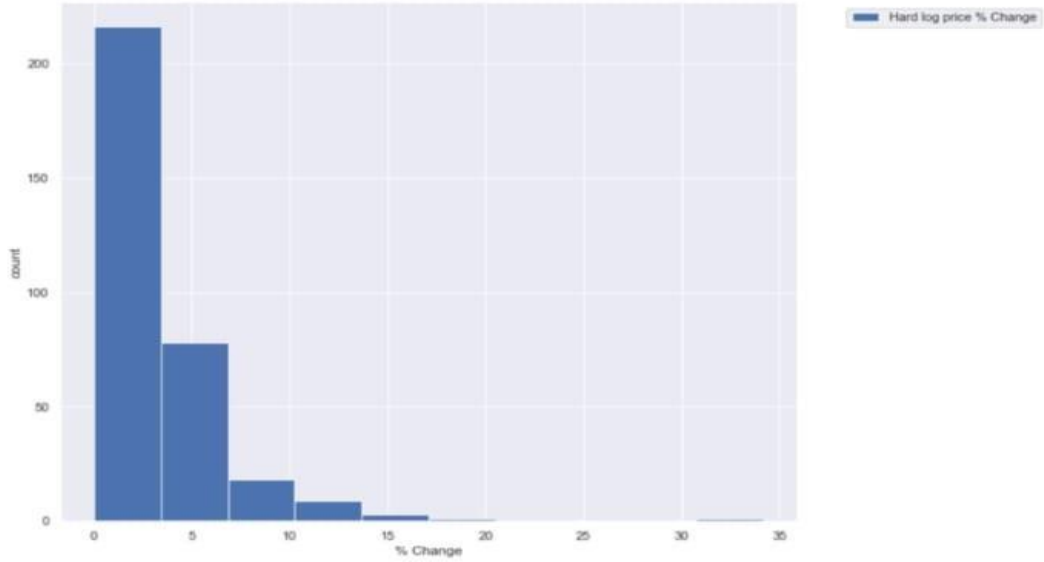


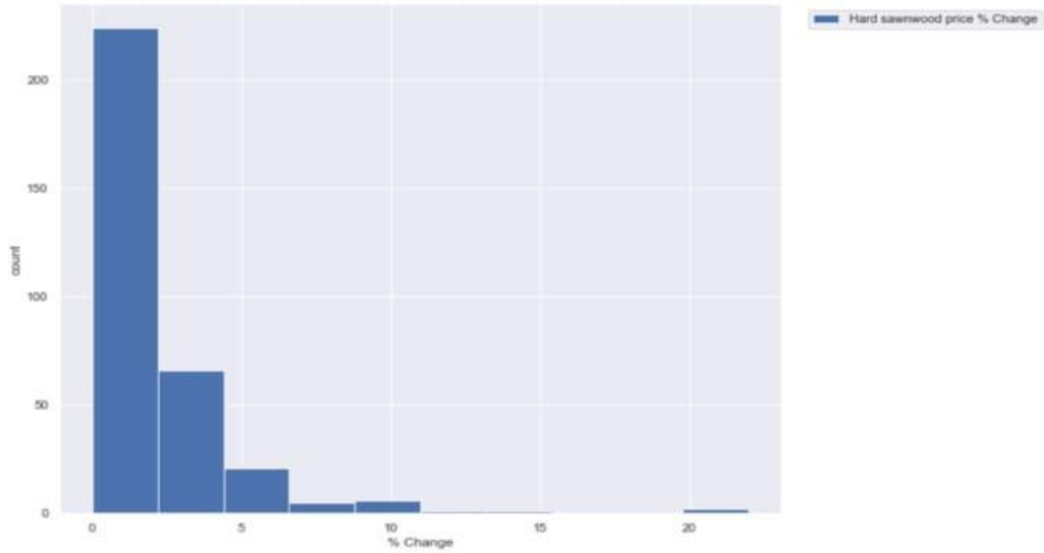


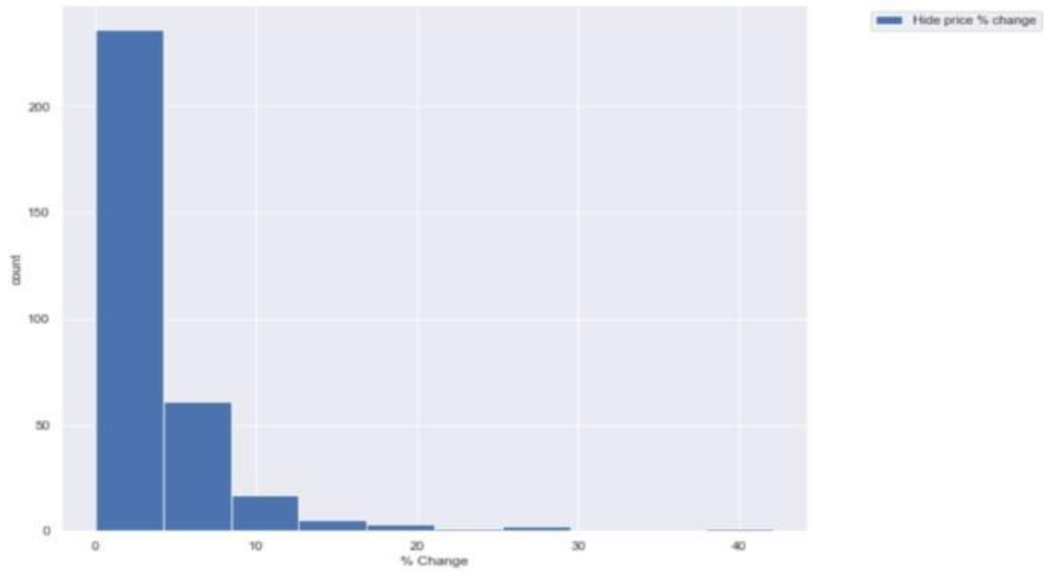

Cotton price % Change
count
% Change
0
20
40
60
80
100
120
140
5
10
15
20

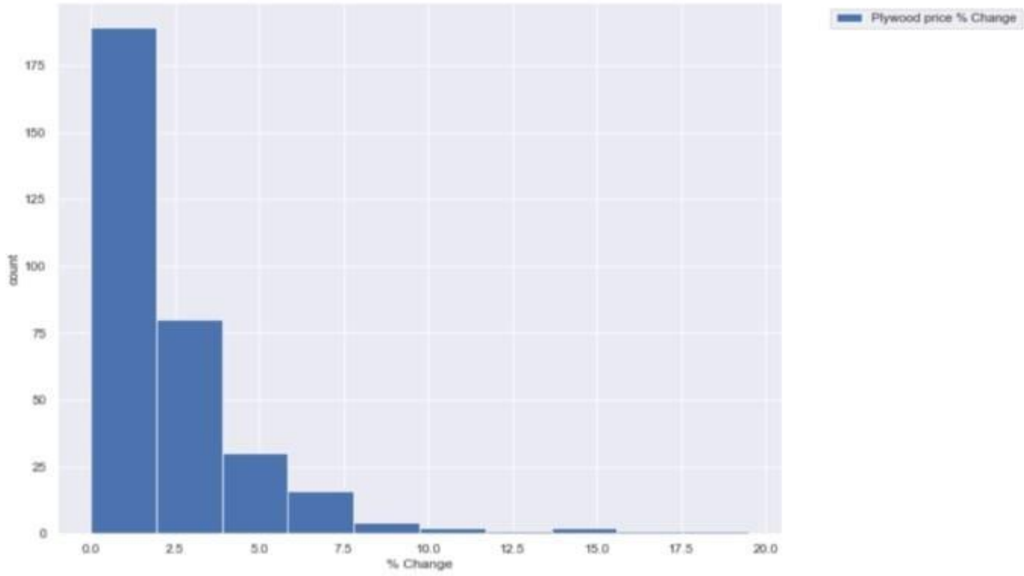

Coarse wool price % Change
count
% Change
0
20
40
60
80
100
120
5
10
15
20

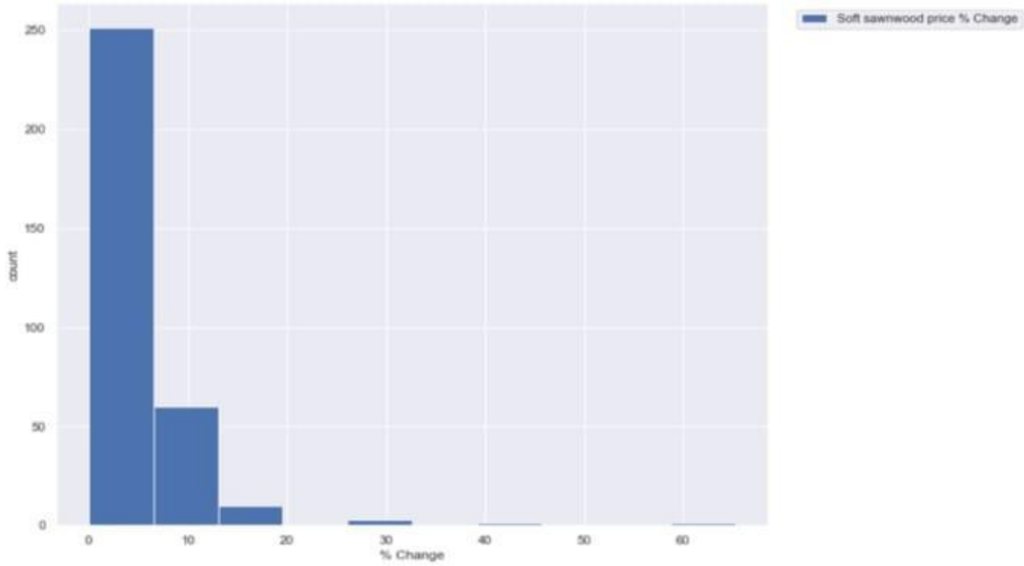

Fine wool price % Change
count
% Change
0
20
40
60
80
100
120
140
160
5
10
15
20
25
30

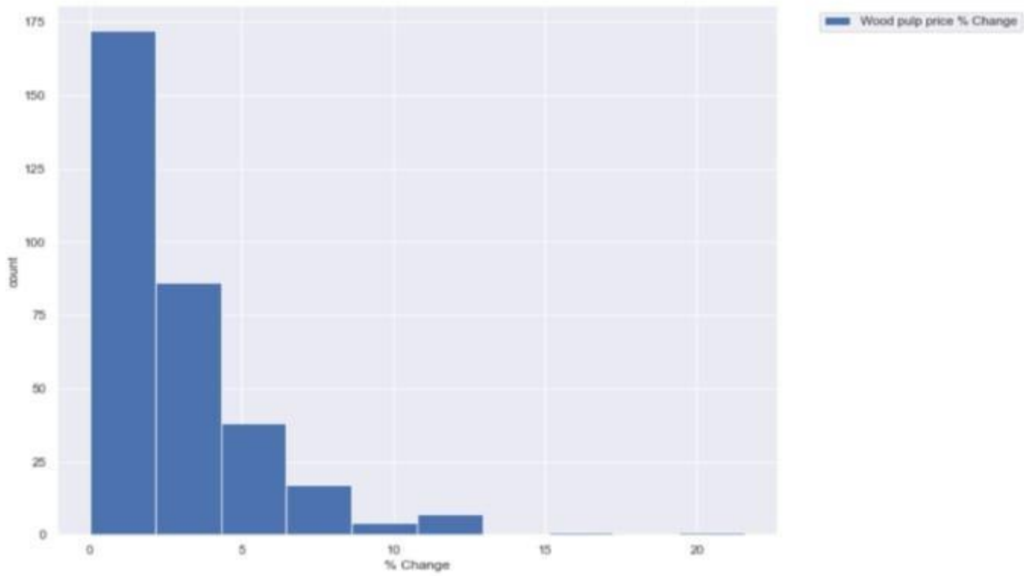

Hard log price % Change
count
% Change


Hard sawnwood price % Change
count
% Change


Hide price % change
count
% Change


Plywood price % Change
count
% Change


Soft sawnwood price % Change
count
% Change

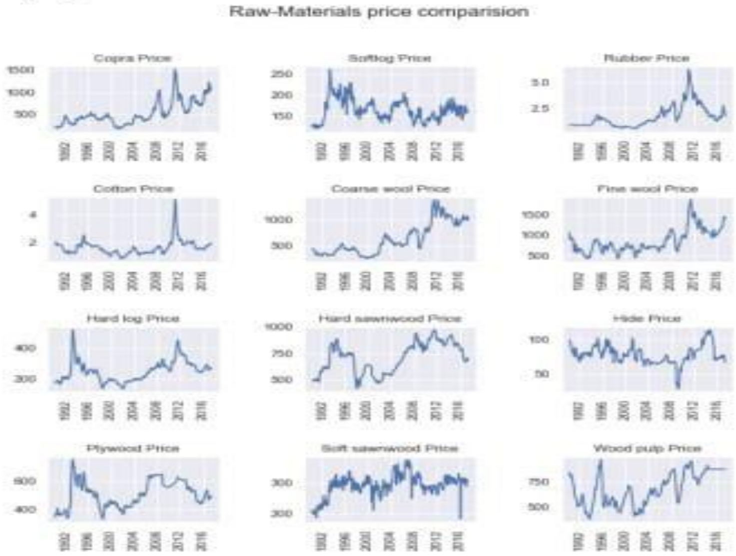

Wood pulp price % Change
count
% Change

We can observe that most raw-materials have ideal frequent %change less than 5%

Q2: Find the raw-material that has lowest price over years

In [31]:
```
plt.figure(figsize=(10, 10))
materialslist=['Copra Price','Softlog Price','Rubber Price','Cotton Price','Coarse wool Price','Fine wo
for i in range(len(materialslist)):
    plt.subplot(4,3,i+1)
    plt.subplots_adjust( hspace=1 ,wspace=0.5)
    plt.title(materialslist[i])
    plt.plot(df[materialslist[i]])
    plt.xticks(rotation=90)
plt.suptitle("Raw-Materials price comparision")
```

Out[31]:
```
Text(0.5, 0.98, 'Raw-Materials price comparision')
```

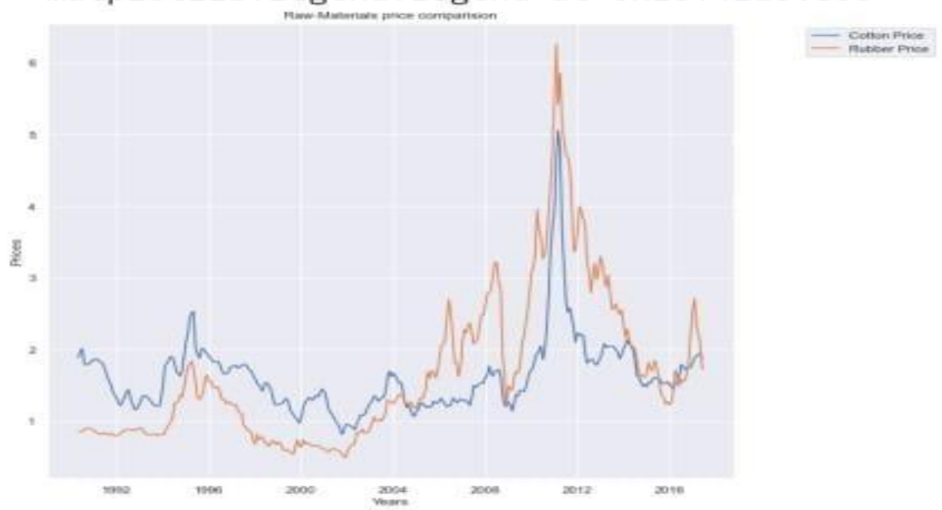


We can see cotton and rubber are of lowest prices.

lets compare prices to better understand which is lowest.

In [32]:
```
plt.figure(figsize=(10, 10))
plt.plot(df[['Cotton Price','Rubber Price']])
plt.title("Raw-Materials price comparision")
plt.xlabel('Years')
plt.ylabel('Prices')
plt.legend(['Cotton Price','Rubber Price'],loc='upper center',bbox_to_anchor=(1.2,1))
```

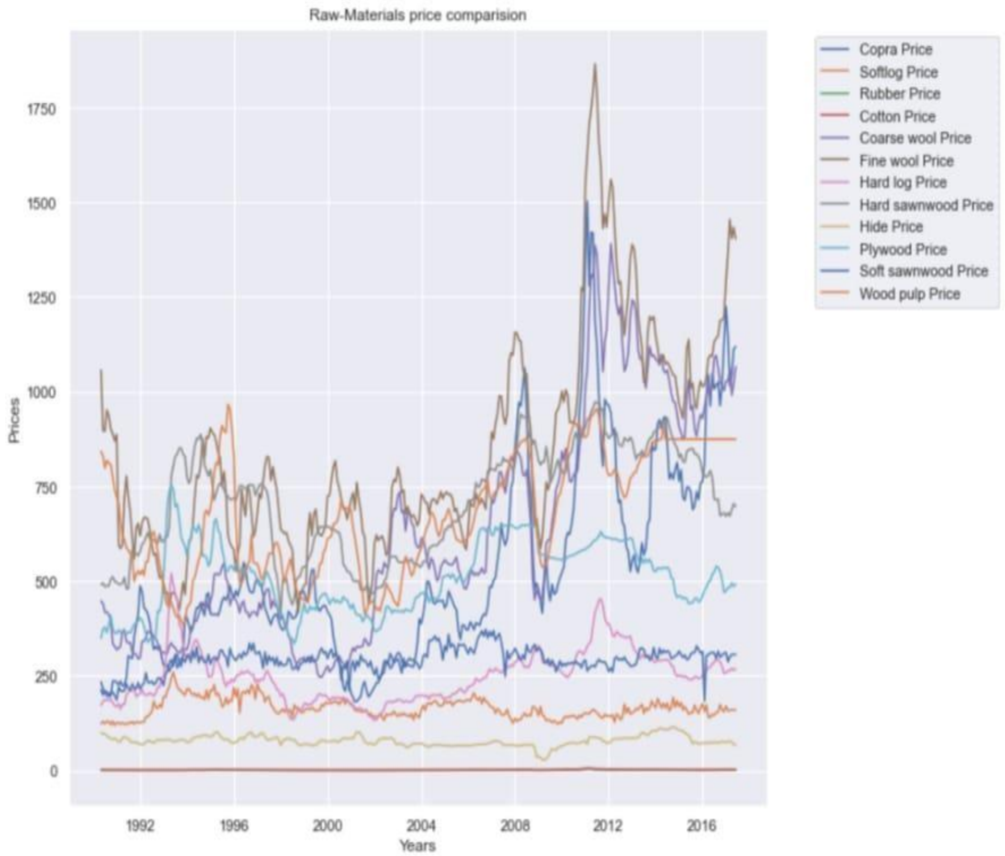
Out[32]:
```
<matplotlib.legend.Legend at 0x26142d81e80>
```



cotton is the lowest price raw-materials in recent years

In [33]:
```
plt.figure(figsize=(10, 10))
plt.plot(df[['Copra Price','Softlog Price','Rubber Price','Cotton Price','Coarse wool Price','Fine wool
plt.title("Raw-Materials price comparision")
plt.xlabel('Years')
plt.ylabel('Prices')
plt.legend(['Copra Price','Softlog Price','Rubber Price','Cotton Price','Coarse wool Price','Fine wool
```

Out[33]: <matplotlib.legend.Legend at 0x26142703730>



From the graphs we can analyze raw materials into different catagories according to their price over years.

low price materials

-cotton, hide, softlog, Hard log, Soft sawnwood Price, rubber

High price materials

-coarse wool, copra, fine wool, hard sawnwood, woodpulp, plywood

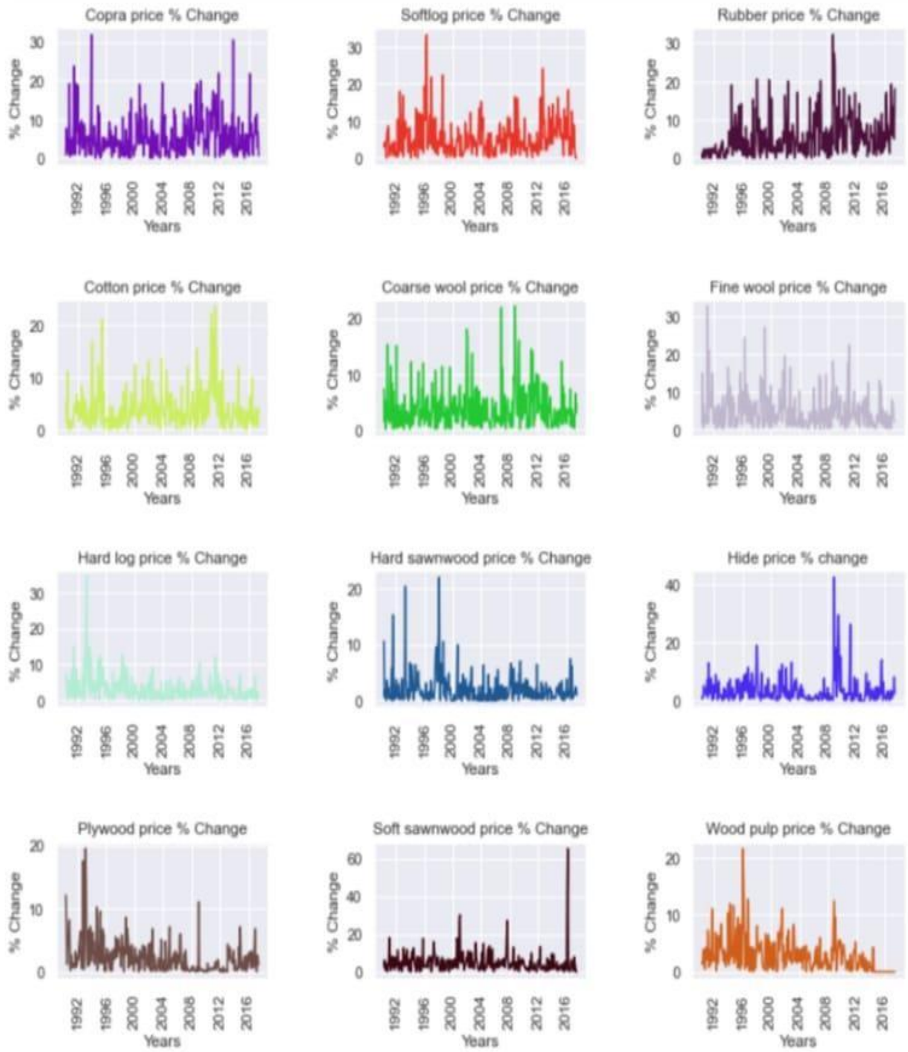
Q3: which raw material has the highest and lowest price % change

In [34]:

```
import random as random
```

In [35]:

```
#lowchangelist=['Cotton price % Change','Hide price % change','Softlog price % Change','Hard log price
plt.figure(figsize=(12,12))
for i in range(len(changelist)):

    r = random.random()
    b = random.random()
    g = random.random()
    color = (r, g, b)
    plt.subplot(4,3,i+1)
    plt.subplots_adjust( hspace=1 ,wspace=0.5)
    plt.plot(df[changelist[i]],c=color)
    plt.xticks(rotation=90)
    plt.title(changelist[i])
    plt.xlabel('Years')
    plt.ylabel('% Change')
    #plt.legend(changelist[i:],loc='upper center',bbox_to_anchor=(1.2,1))
```



From the above Graphs, We can see that, the highest % change is more than 60 for soft sawnwood and lowest % change is for plywood at less at 20.

Q4: Find the raw materials with drastic price change

In [36]:

```
#plt.figure(figsize=(12,12))
#sns.scatterplot(df['Cotton Price'],df.index,hue=df['Fine wool Price']);
lowlist=['Cotton Price','Hide Price','Softlog Price','Hard log Price','Soft sawnwood Price','Rubber Pri
plt.figure(figsize=(12,12))
plt.ylabel('Prices')
plt.xlabel('Years')
#plt.legend(lowlist,loc='upper center',bbox_to_anchor=(1.2,1))
for i in range(len(lowlist)):
    sns.scatterplot(y=df[lowlist[i]],x=df.index);
    plt.legend(lowlist,loc='upper center',bbox_to_anchor=(1.2,1))
```

In [37]:

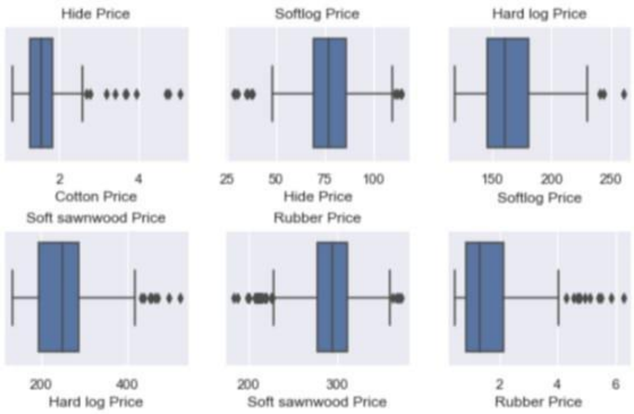
```
plt.figure(figsize=(12,12))
#sns.scatterplot(df['Cotton Price'],df.index,hue=df['Fine wool Price']);
#lowlist=['Cotton Price','Hide Price','Softlog Price','Hard log Price','Soft sawnwood Price','Plywood P
highlist=['Coarse wool Price','Copra Price','Fine wool Price','Hard sawnwood Price','Plywood Price','Wo
#lowchangelist=['Cotton price % Change','Hide price % change','Softlog price % Change','Hard log price
plt.figure(figsize=(12,12))
plt.xlabel('Prices')
plt.ylabel('Years')
for i in range(len(highlist)):
    sns.scatterplot(x=df[highlist[i]],y=df.index);
    plt.legend(highlist,loc='upper center',bbox_to_anchor=(1.2,1))
```

<Figure size 864x864 with 0 Axes>

The Price change is drastic for hard log price among low price range materials and Among high price materials it is Fine wool prices.

Q5: Figure out the price range of low priced raw-materials

In [39]:
```
#plt.figure(figsize=(10, 10))
lowlist=['Cotton Price','Hide Price','Softlog Price','Hard log Price','Soft sawnwood Price','Rubber Pri
for i in range(6):
    plt.title(lowlist[i])
    plt.subplot(2, 3, i+1)
    plt.subplots_adjust( hspace=0.5 )
    sns.boxplot(x=df[lowlist[i]])
```



Box Plot gives us the distribution of data

It includes: Inter quartile range, which is between Q3 and Q1 minimum, first quartile (Q1), median, third quartile (Q3), and maximum and also an outliers.

Inferences and Conclusion

We found out the high range and low range raw-materials according to their prices.

- High and low %Change materials

- We could identify the the range of prices change over the years.

- Correlation between them using a heatmap

Discussion

The analysis of agricultural raw material prices over the years has revealed several key insights. Firstly, the time series analysis shows notable trends and fluctuations in prices, with certain raw materials experiencing higher volatility than others. For instance, products like wheat and soybeans may exhibit seasonal price changes due to planting and harvesting cycles, which are heavily influenced by climate conditions, demand, and regional availability.

The data also indicates price variations across regions, suggesting that factors like regional climate, transportation costs, and local demand significantly impact prices. For example, raw materials sourced from regions with ideal growing conditions may have more stable prices, while those subject to extreme weather fluctuations could see higher volatility.

Correlation analysis reveals that some raw materials may be influenced by related market trends. For instance, price changes in one agricultural commodity, like corn, can have ripple effects on related products, such as livestock feed. This interconnectedness highlights the complexities within agricultural markets, where the price of one commodity can be affected by the supply and demand of another.

Conclusion

In conclusion, the analysis has highlighted the variability and complexity of agricultural raw material prices. Seasonal patterns, regional influences, and correlations among various materials all play crucial roles in price dynamics. For stakeholders in the agricultural industry, understanding these trends can provide valuable insights for strategic decision-making, such as optimizing purchasing cycles or adjusting supply chain logistics.

Future analysis could benefit from examining external factors, such as weather patterns, trade policies, and global demand, to further understand their impact on raw material prices. Additionally, using predictive models could help anticipate future price changes, providing an even greater strategic advantage for market participants.

References

- Handbook of Agricultural Analysis" by Leo M.L. Nollet
- "Postharvest Technology of Agricultural Crops" by A.A. Kader
- "Agricultural Biomass Based Potential Materials" by Khalid Rehman Hakeem

1. Project Github link,2024
2. Project Sheets & Report github link, 2024

Link

https://github.com/AU810021114005/Akash-A

Video Recording of Project Demonstration:

The recorded video is in the form of PPT and it is attached in the Git Hub

https://github.com/AU810021114005/Akash-A