1. Implement programs for time series data cleaning, loading and handling times series data and pre-processing techniques.

EX.N0 : 1	**Implement programs for time series data cleaning, loading and handling times series data and pre-processing techniques.**
DATE : 25/01/2025	

AIM:

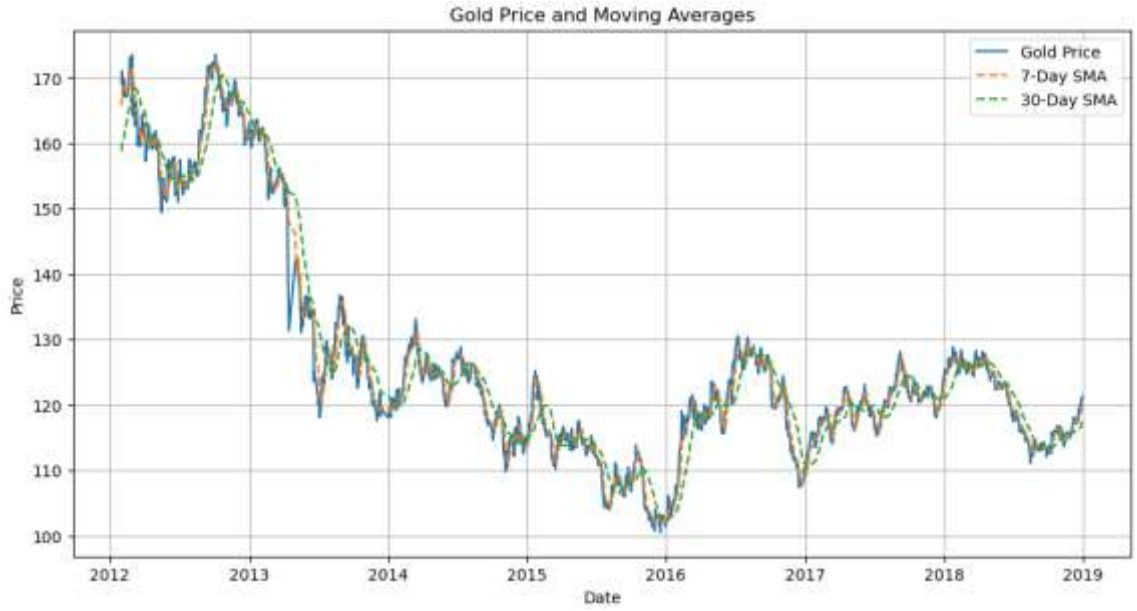
PROGRAM:

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

def load_data(filepath):

    try:

        data = pd.read_csv(filepath, parse_dates=True, index_col='Date')

        print("Dataset loaded successfully.")

        return data

    except Exception as e:

        print(f"Error loading dataset: {e}")

        return None

def clean_data(data):

    print("Initial dataset shape:", data.shape)

    data = data.drop_duplicates()

    data = data.fillna(method='ffill')  # Forward fill missing values

    data = data.fillna(method='bfill')  # Backward fill for remaining missing values
```

```
    data = data.dropna()
    print("Dataset shape after cleaning:", data.shape)
    return data
def preprocess_time_series(data):
    print("Index type:", type(data.index))
    if not isinstance(data.index, pd.DatetimeIndex):
        data.index = pd.to_datetime(data.index)
    data = data.sort_index()
    return data
def feature_engineering(data):
    data['SMA_7'] = data['Close'].rolling(window=7).mean()  # 7-day Simple Moving Average
    data['SMA_30'] = data['Close'].rolling(window=30).mean()  # 30-day Simple Moving
Average
    data['Lag_1'] = data['Close'].shift(1)  # Previous day's price
    data['Lag_7'] = data['Close'].shift(7)  # Price a week ago
    data = data.dropna()
    return data
def visualize_data(data):
    plt.figure(figsize=(12, 6))
    plt.plot(data['Close'], label='Gold Price')
    plt.plot(data['SMA_7'], label='7-Day SMA', linestyle='--')
    plt.plot(data['SMA_30'], label='30-Day SMA', linestyle='--')
    plt.title('Gold Price and Moving Averages')
    plt.xlabel('Date')
    plt.ylabel('Price')
```

```
    plt.legend()
    plt.grid()
    plt.show()
def main():
    filepath = "C:\\Users\\jaya karthick\\Downloads\\archive (1) (1)\\FINAL_USO.csv"
    data = load_data(filepath)
    if data is None:
        return
    data = clean_data(data)
    data = preprocess_time_series(data)
    data = feature_engineering(data)
    visualize_data(data)
    print("Processed dataset preview:\n", data.head())
if __name__ == "__main__":
    main()
```

OUTPUT:



RESULT:

Thus, the program for Implement programs for time series data cleaning, loading and handling times series data and pre-processing techniques is executed successfully.