

```

# Load Data and Set 'climate' as the Index
import pandas as pd

def load_timeseries_data(file_path: str):
    df = pd.read_csv(file_path)
    df['Year'] = pd.to_datetime(df['Year'], format='%Y')

    df.set_index('Year', inplace=True)

    return df

df =
load_timeseries_data('C:/Users/Lenovo/Downloads/Climate_Change_Indicators.csv
')
print(df.head())

```

Year	Global Average Temperature (°C)	CO2 Concentration (ppm)
1948-01-01	13.17	397.04
1996-01-01	13.10	313.17
2015-01-01	14.67	311.95
1966-01-01	14.79	304.25
1992-01-01	13.15	354.52

Year	Sea Level Rise (mm)	Arctic Ice Area (million km ²)
1948-01-01	116.25	5.97
1996-01-01	277.92	9.66
2015-01-01	290.32	8.40
1966-01-01	189.71	11.83
1992-01-01	14.84	11.23

In [7]:

```

# Clean and Preprocess the Data
# Clean missing values (e.g., forward fill missing values)
def clean_missing_values(df):
    df = df.fillna(method='ffill') # Forward fill missing values
    return df

# Remove duplicates if any
def remove_duplicates(df):
    df = df.drop_duplicates()
    return df

# Example usage
df_cleaned = clean_missing_values(df)
df_cleaned = remove_duplicates(df_cleaned)

# Visualize the time series data
import matplotlib.pyplot as plt

def plot_timeseries(df):
    df.plot(figsize=(10, 6))

```

```
plt.title("Climate Change Indicators")
plt.xlabel('Year')
plt.ylabel('Values')
plt.show()
```

```
# Plot the cleaned data
plot_timeseries(df_cleaned)
```

In [8]:

```
# Create lag features (e.g., lag of 1 year for temperature)
def create_lag_features(df, lag=1):
    df['lag_1'] = df['Global Average Temperature (°C)'].shift(lag)
    return df

# Example usage
df_with_lags = create_lag_features(df_cleaned)
plot_timeseries(df_with_lags)
```

In [9]:

```
def preprocess_timeseries(file_path):
    # Step 1: Load the data
    df = load_timeseries_data(file_path)

    # Step 2: Clean the data
    df = clean_missing_values(df)
    df = remove_duplicates(df)

    # Step 3: Feature Engineering (Create lag features, etc.)
    df = create_lag_features(df)

    # Step 4: Visualize the cleaned data
    plot_timeseries(df)

    return df

# Example usage
df_preprocessed =
preprocess_timeseries('C:/Users/Lenovo/Downloads/Climate_Change_Indicators.csv')
```

Global Average Temperature (°C)	CO2 Concentration (ppm)	\
Year		
1948-01-01	13.17	397.04
1996-01-01	13.10	313.17
2015-01-01	14.67	311.95
1966-01-01	14.79	304.25
1992-01-01	13.15	354.52

	Sea Level Rise (mm)	Arctic Ice Area (million km ²)
Year		
1948-01-01	116.25	5.97
1996-01-01	277.92	9.66
2015-01-01	290.32	8.40
1966-01-01	189.71	11.83
1992-01-01	14.84	11.23

