

```
In [10]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split

df = pd.read_csv(r'C:\Users\PC30\Downloads\climate_change_impact_on_agriculture_2024

print("Preview of the dataset:")
print(df.head())

print("\nDataset information:")
print(df.info())

print("\nMissing values in each column:")
print(df.isnull().sum())

# Relationship between temperature and crop yield by crop type
plt.figure(figsize=(10, 6))
sns.scatterplot(data=df, x='Average_Temperature_C', y='Crop_Yield_MT_per_HA',
hue='Crop_Type')
plt.title('Temperature vs Crop Yield by Crop Type')
plt.xlabel('Average Temperature (°C)')
plt.ylabel('Crop Yield (MT/HA)')
plt.legend(title='Crop Type')
plt.show()

# Visualize economic impact per crop type
plt.figure(figsize=(10, 6))
sns.boxplot(data=df, x='Crop_Type', y='Economic_Impact_Million_USD')
plt.title("Economic Impact by Crop Type")
plt.xlabel('Crop Type')
plt.ylabel('Economic Impact (Million USD)')
plt.xticks(rotation=45)
plt.show()
```

Preview of the dataset:

	Year	Country	Region	Crop_Type	Average_Temperature_C	\
0	2001	India	West Bengal	Corn	1.55	
1	2024	China	North	Corn	3.23	
2	2001	France	Ile-de-France	Wheat	21.11	
3	2001	Canada	Prairies	Coffee	27.85	
4	1998	India	Tamil Nadu	Sugarcane	2.19	

	Total_Precipitation_mm	CO2_Emissions_MT	Crop_Yield_MT_per_HA	\
0	447.06	15.22	1.737	
1	2913.57	29.82	1.737	
2	1301.74	25.75	1.719	
3	1154.36	13.91	3.890	
4	1627.48	11.81	1.080	

	Extreme_Weather_Events	Irrigation_Access_%	Pesticide_Use_KG_per_HA	\
0	8	14.54	10.08	
1	8	11.05	33.06	
2	5	84.42	27.41	
3	5	94.06	14.38	
4	9	95.75	44.35	

	Fertilizer_Use_KG_per_HA	Soil_Health_Index	Adaptation_Strategies	\
0	14.78	83.25	Water Management	
1	23.25	54.02	Crop Rotation	
2	65.53	67.78	Water Management	

3	87.58	91.39	No Adaptation
4	88.08	49.61	Crop Rotation

	Economic_Impact_Million_USD
0	808.13
1	616.22
2	796.96
3	790.32
4	401.72

Dataset information:

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 10000 entries, 0 to 9999

Data columns (total 15 columns):

#	Column	Non-Null Count	Dtype
0	Year	10000 non-null	int64
1	Country	10000 non-null	object
2	Region	10000 non-null	object
3	Crop_Type	10000 non-null	object
4	Average_Temperature_C	10000 non-null	float64
5	Total_Precipitation_mm	10000 non-null	float64
6	CO2_Emissions_MT	10000 non-null	float64
7	Crop_Yield_MT_per_HA	10000 non-null	float64
8	Extreme_Weather_Events	10000 non-null	int64
9	Irrigation_Access_%	10000 non-null	float64
10	Pesticide_Use_KG_per_HA	10000 non-null	float64
11	Fertilizer_Use_KG_per_HA	10000 non-null	float64
12	Soil_Health_Index	10000 non-null	float64
13	Adaptation_Strategies	10000 non-null	object
14	Economic_Impact_Million_USD	10000 non-null	float64

dtypes: float64(9), int64(2), object(4)

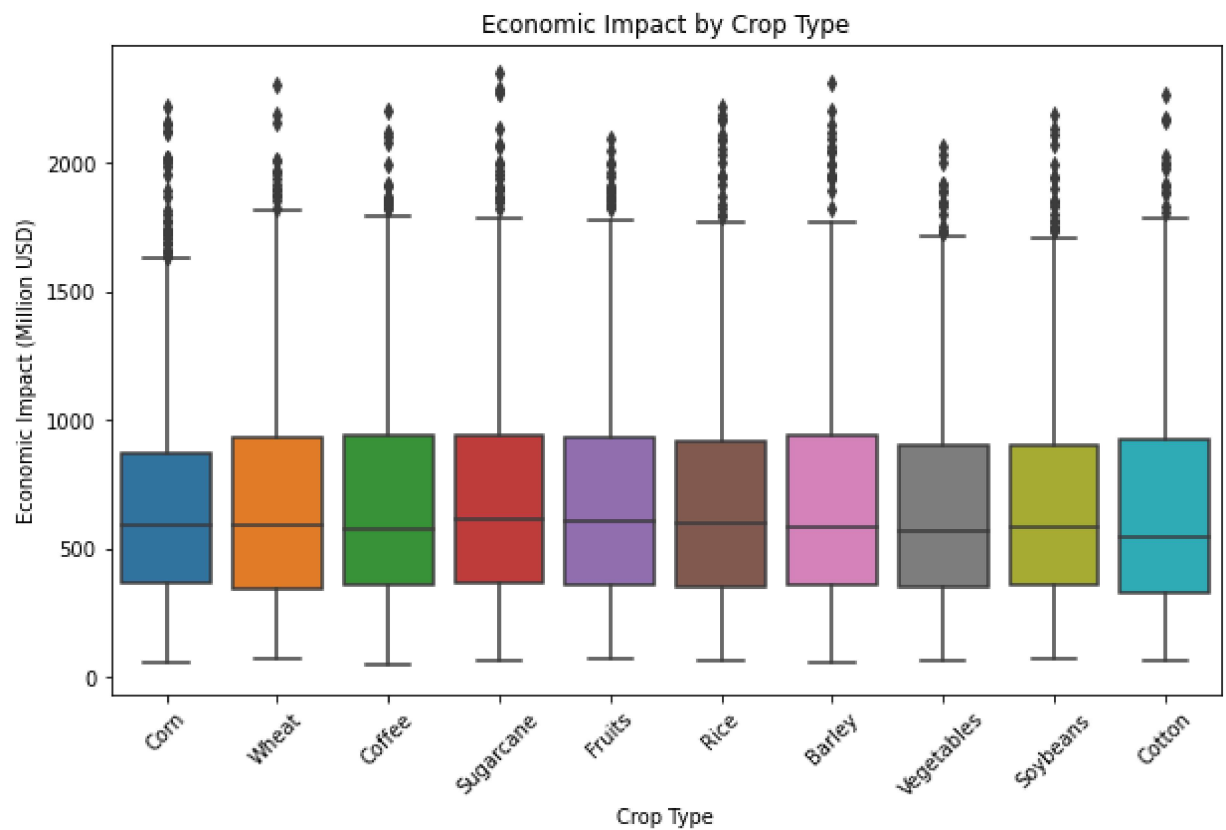
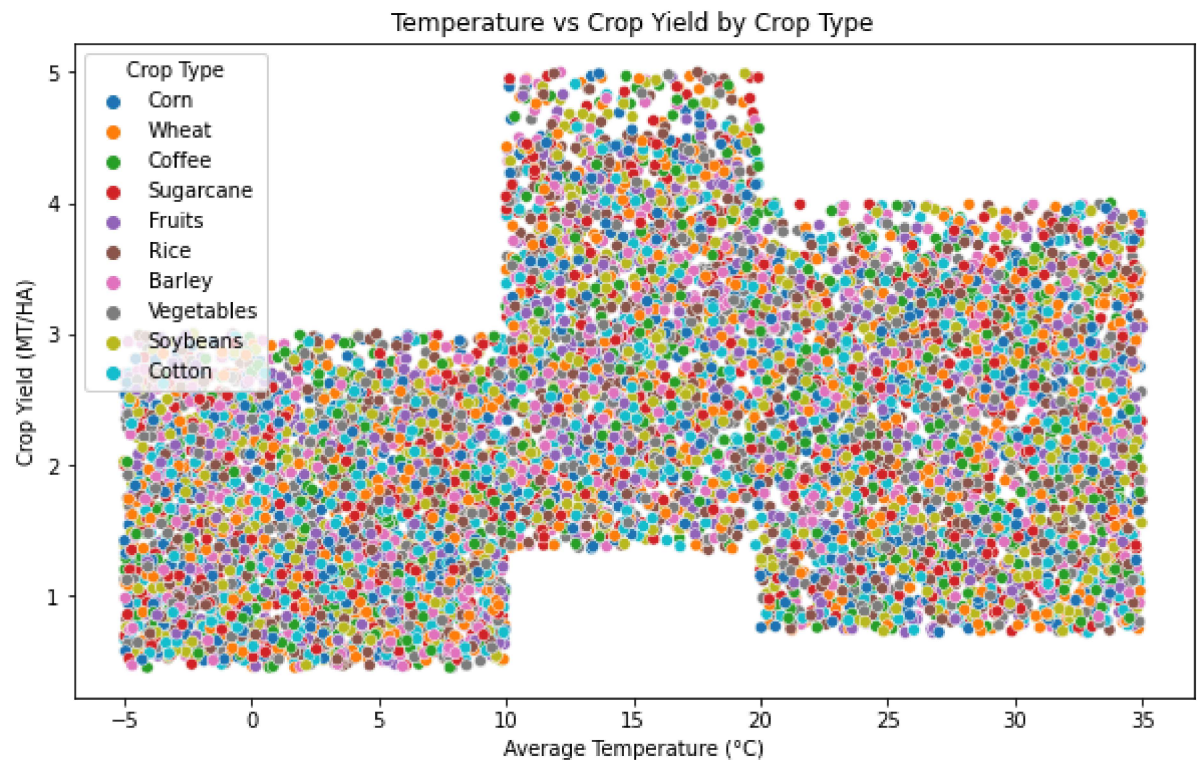
memory usage: 1.1+ MB

None

Missing values in each column:

Year	0
Country	0
Region	0
Crop_Type	0
Average_Temperature_C	0
Total_Precipitation_mm	0
CO2_Emissions_MT	0
Crop_Yield_MT_per_HA	0
Extreme_Weather_Events	0
Irrigation_Access_%	0
Pesticide_Use_KG_per_HA	0
Fertilizer_Use_KG_per_HA	0
Soil_Health_Index	0
Adaptation_Strategies	0
Economic_Impact_Million_USD	0

dtype: int64



In []:

In []:

In []:

In []:

In []:

In []: