

Crop Yield Prediction Using Weather Data



Project Overview

This project predicts **crop yield** based on weather parameters such as **temperature, humidity, and rainfall**. It uses Python for data analysis, visualization, and can be extended with machine learning models.

This project is suitable for:

- Academic mini / major projects
- Resume & portfolio
- GitHub showcase



Problem Statement

Agriculture heavily depends on weather conditions. Predicting crop yield using historical weather data helps farmers and policymakers make better decisions.



Technologies Used

- Python
- Pandas
- NumPy
- Matplotlib
- Seaborn
- Jupyter Notebook



Dataset Description

weather.csv contains:

- Temperature (°C)
- Humidity (%)
- Rainfall_mm (mm)
- Crop_Yield (tons/hectare)
- Date
- City



Project Workflow

1. Load weather dataset
2. Handle missing values
3. Perform statistical analysis
4. Visualize weather vs crop yield
5. (Optional) Train ML models



Python Source Code (`crop_yield_prediction.py`)

```
import numpy as np
import pandas as pd
import statistics
import matplotlib.pyplot as plt
import seaborn as sns

# Load dataset
df = pd.read_csv("data/weather.csv")

# Handle missing values
```

```
df.fillna(df.mean(numeric_only=True), inplace=True)

# Feature selection
X = df[['Temperature', 'Humidity', 'Rainfall_mm']]
y = df['Crop_Yield']

# Statistical analysis
print("Mean Temperature:", statistics.mean(df['Temperature']))
print("Median Temperature:", statistics.median(df['Temperature']))

# Visualization
plt.figure(figsize=(10, 5))
sns.scatterplot(x='Temperature', y='Crop_Yield', data=df)
plt.title("Temperature vs Crop Yield")
plt.show()

plt.figure(figsize=(10, 5))
sns.scatterplot(x='Rainfall_mm', y='Crop_Yield', data=df)
plt.title("Rainfall vs Crop Yield")
plt.show()
```



Future Enhancements

- Add Machine Learning models (Linear Regression, Random Forest)
- Accuracy evaluation
- Web interface using Flask or Streamlit



Output

Date	City	Temperature_C	Humidity_%	Wind_Speed_kmph	Rainfall_mm	Weather_Condition
0	2024-01-01	Kolkata	20.6	91	34.0	47.31
1	2024-01-02	Chennai	12.0	55	9.4	39.06
2	2024-01-03	Mumbai	20.0	57	17.6	5.67
3	2024-01-04	Chennai	20.4	44	28.6	46.55
4	2024-01-05	Chennai	33.3	37	7.3	48.71
...
115	2024-04-25	Delhi	38.7	27	5.2	36.31
116	2024-04-26	Bhopal	38.4	72	5.5	37.85
117	2024-04-27	Kolkata	35.0	79	24.9	18.89
118	2024-04-28	Bhopal	30.5	24	23.0	12.05
119	2024-04-29	Kolkata	12.7	87	10.1	10.25

(120, 7)

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

RangeIndex: 120 entries, 0 to 119

Data columns (total 7 columns):

#	Column	Non-Null Count	Dtype
0	Date	120 non-null	object
1	City	120 non-null	object
2	Temperature_C	120 non-null	float64
3	Humidity_%	120 non-null	int64
4	Wind_Speed_kmph	120 non-null	float64
5	Rainfall_mm	120 non-null	float64

6 Weather_Condition 120 non-null object

dtypes: float64(3), int64(1), object(3)

memory usage: 6.7+ KB

Temperature_C	Humidity_%	Wind_Speed_kmph	Rainfall_mm	
count	120.000000	120.000000	120.000000	120.000000
mean	25.673333	58.633333	20.325000	25.534667
std	9.238854	20.148997	10.859363	14.676269
min	10.200000	21.000000	2.600000	0.680000
25%	17.700000	42.750000	11.100000	12.672500
50%	26.000000	57.500000	20.750000	24.050000
75%	33.525000	77.000000	28.225000	37.587500
max	41.500000	99.000000	39.200000	49.800000

7

Index(['Date', 'City', 'Temperature_C', 'Humidity_%', 'Wind_Speed_kmph',

'Rainfall_mm', 'Weather_Condition'],

dtype='object')

Index(['Date', 'City', 'Temperature_C', 'Humidity_%', 'Wind_Speed_kmph',

'Rainfall_mm', 'Weather_Condition'],

dtype='object')

Date	City	Temperature_C	Wind_Speed_kmph	Rainfall_mm	Weather_Condition	
2024						
0 -01-01	Kolkata	20.6		34.0	47.31	Sunny
2024						
1 -01-02	Chennai	12.0		9.4	39.06	Rainy
2024						
2 -01-03	Mumbai	20.0		17.6	5.67	Foggy
2024						
3 -01-04	Chennai	20.4		28.6	46.55	Foggy
2024						
4 -01-05	Chennai	33.3		7.3	48.71	Stormy
...
2024						
115 -04-25	Delhi	38.7		5.2	36.31	Hazy
2024						
116 -04-26	Bhopal	38.4		5.5	37.85	Rainy
2024						
117 -04-27	Kolkata	35.0		24.9	18.89	Foggy

Date	City	Temperature_C	Wind_Speed_kmph	Rainfall_mm	Weather_Condition
2024					
118	-04-28	Bhopal	30.5	23.0	12.05
2024					
119	-04-29	Kolkata	12.7	10.1	10.25

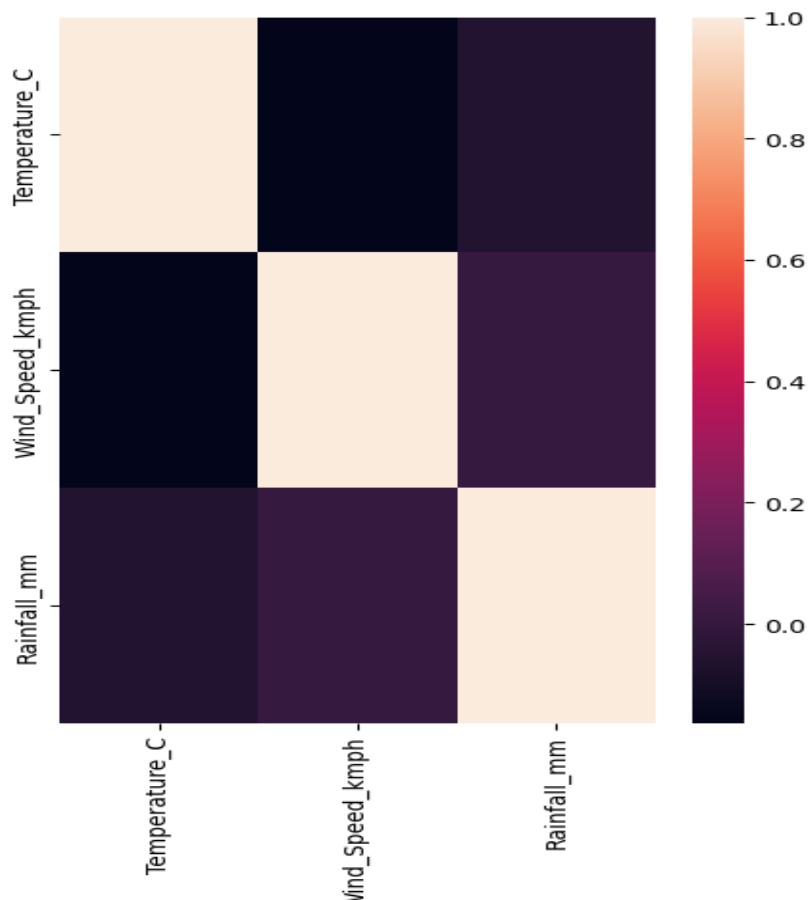
Temperature_C Wind_Speed_kmph Rainfall_mm

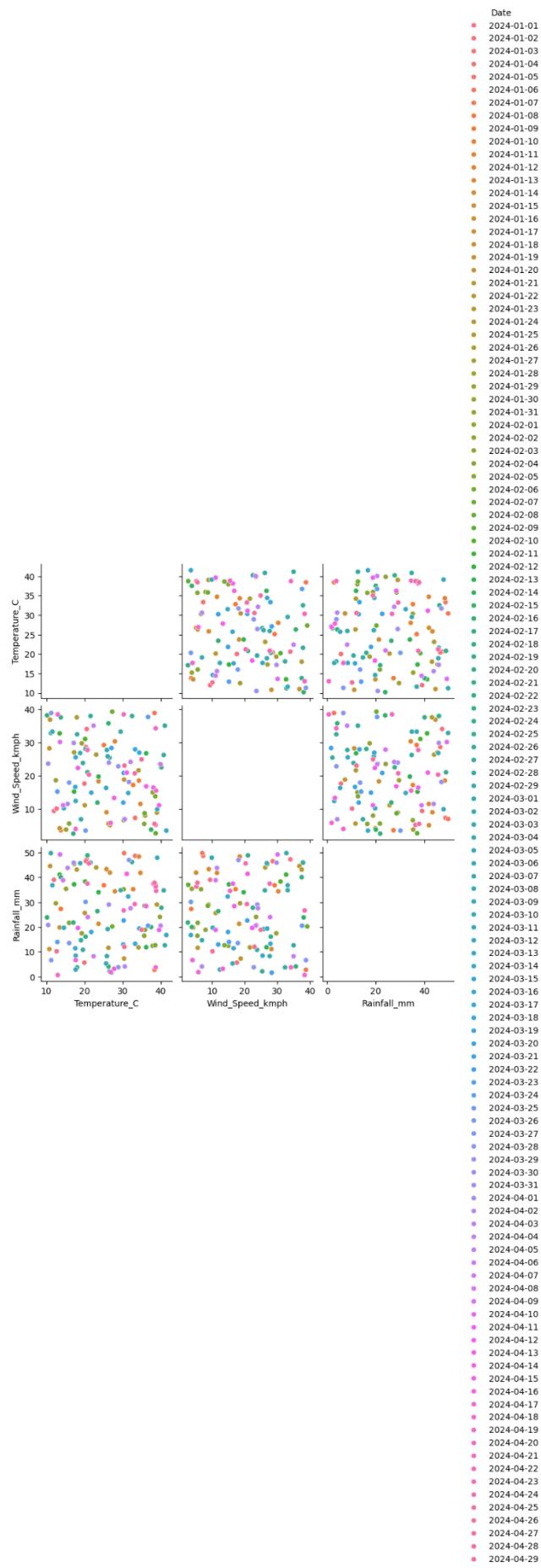
Temperature_C 1.000000 -0.162492 -0.057216

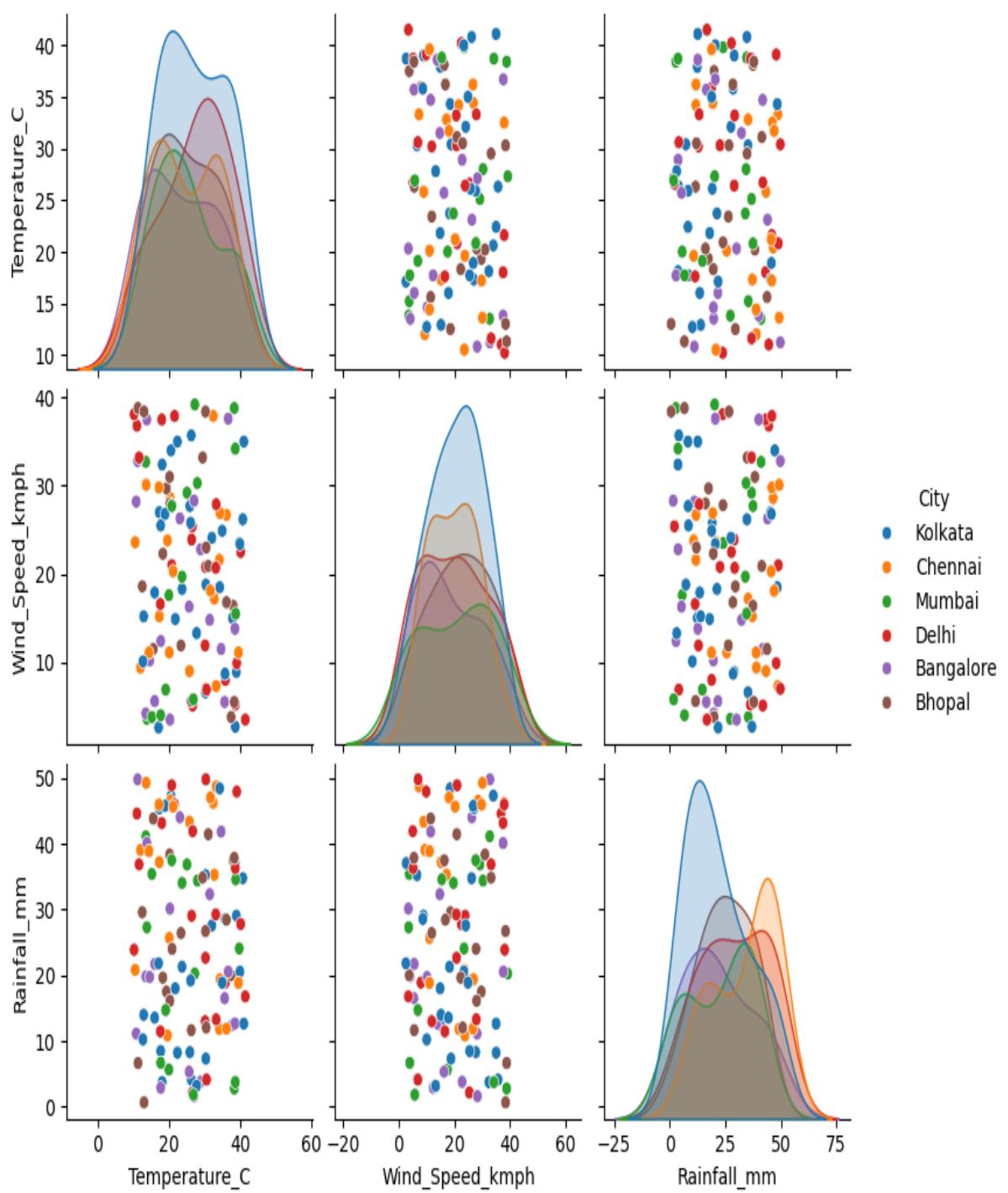
Wind_Speed_kmph -0.162492 1.000000 0.001789

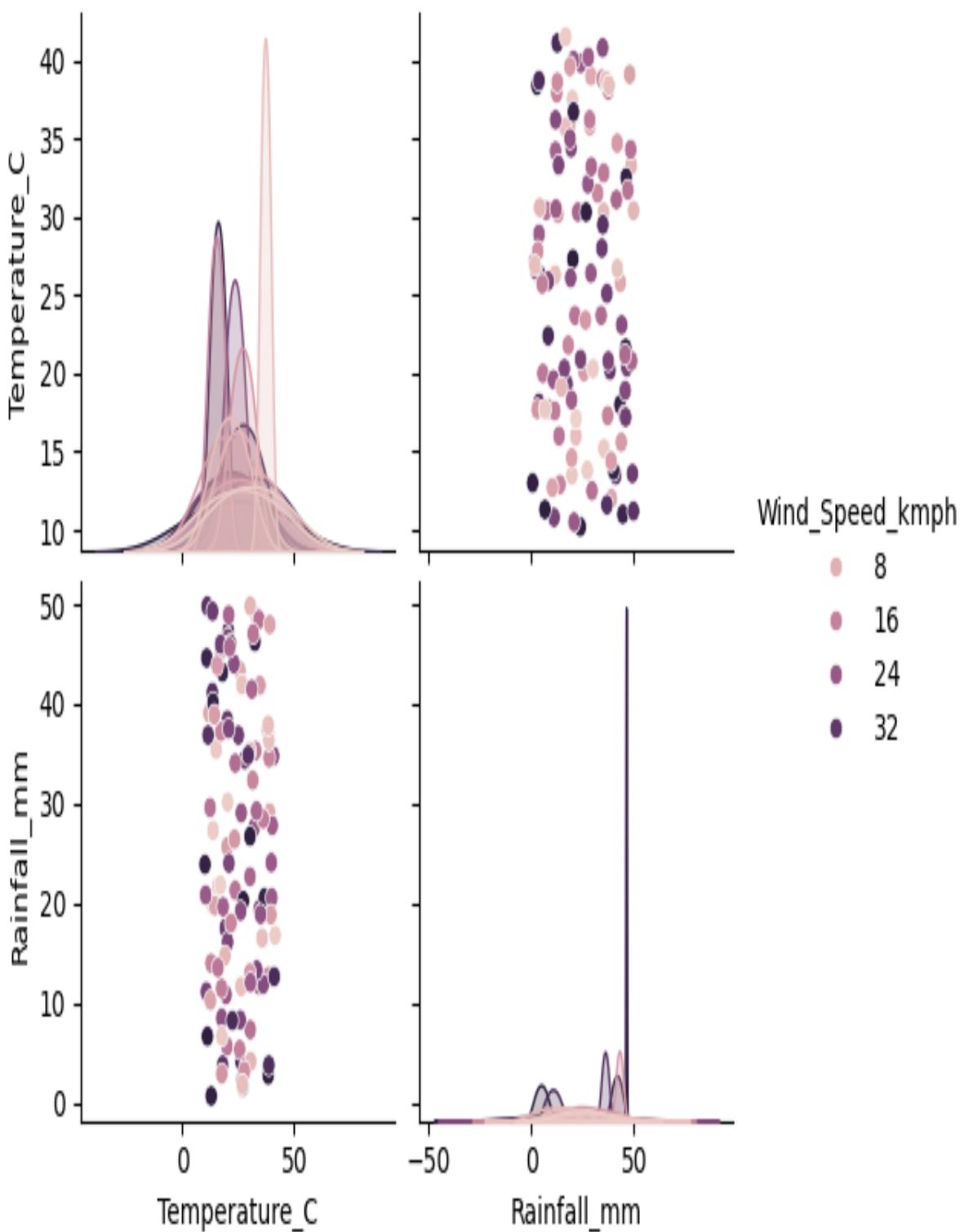
Rainfall_mm -0.057216 0.001789 1.000000

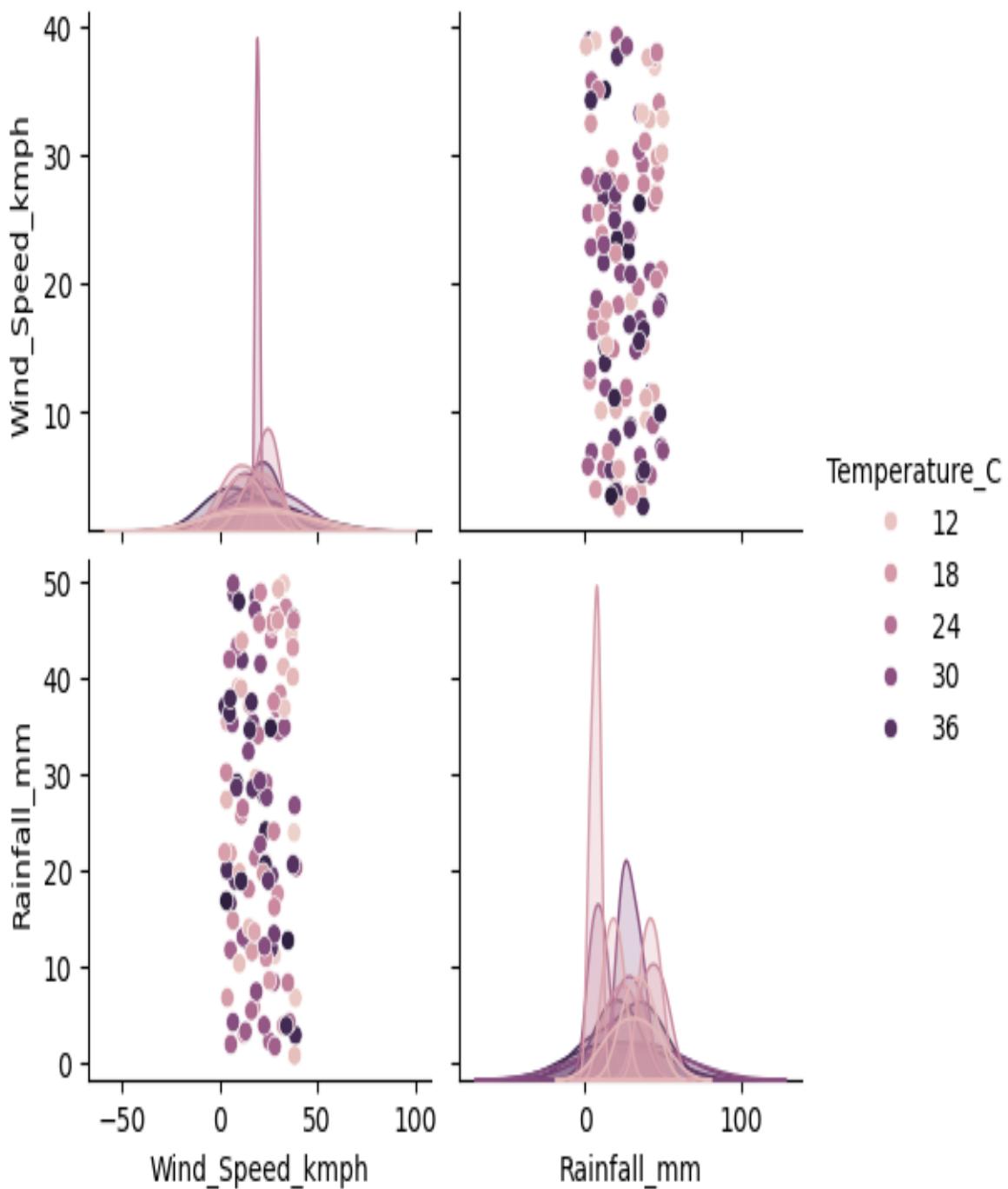
<Axes: >

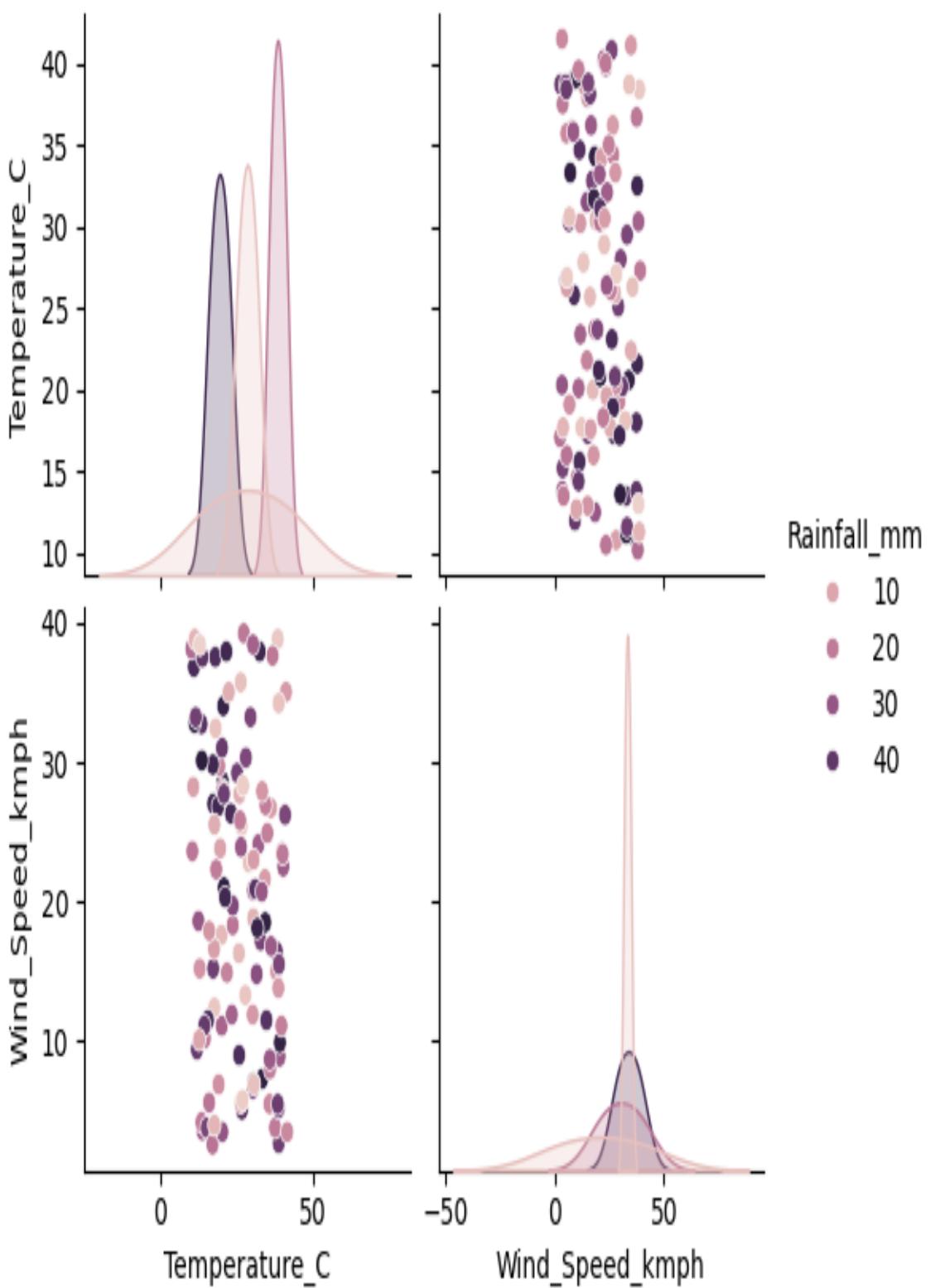


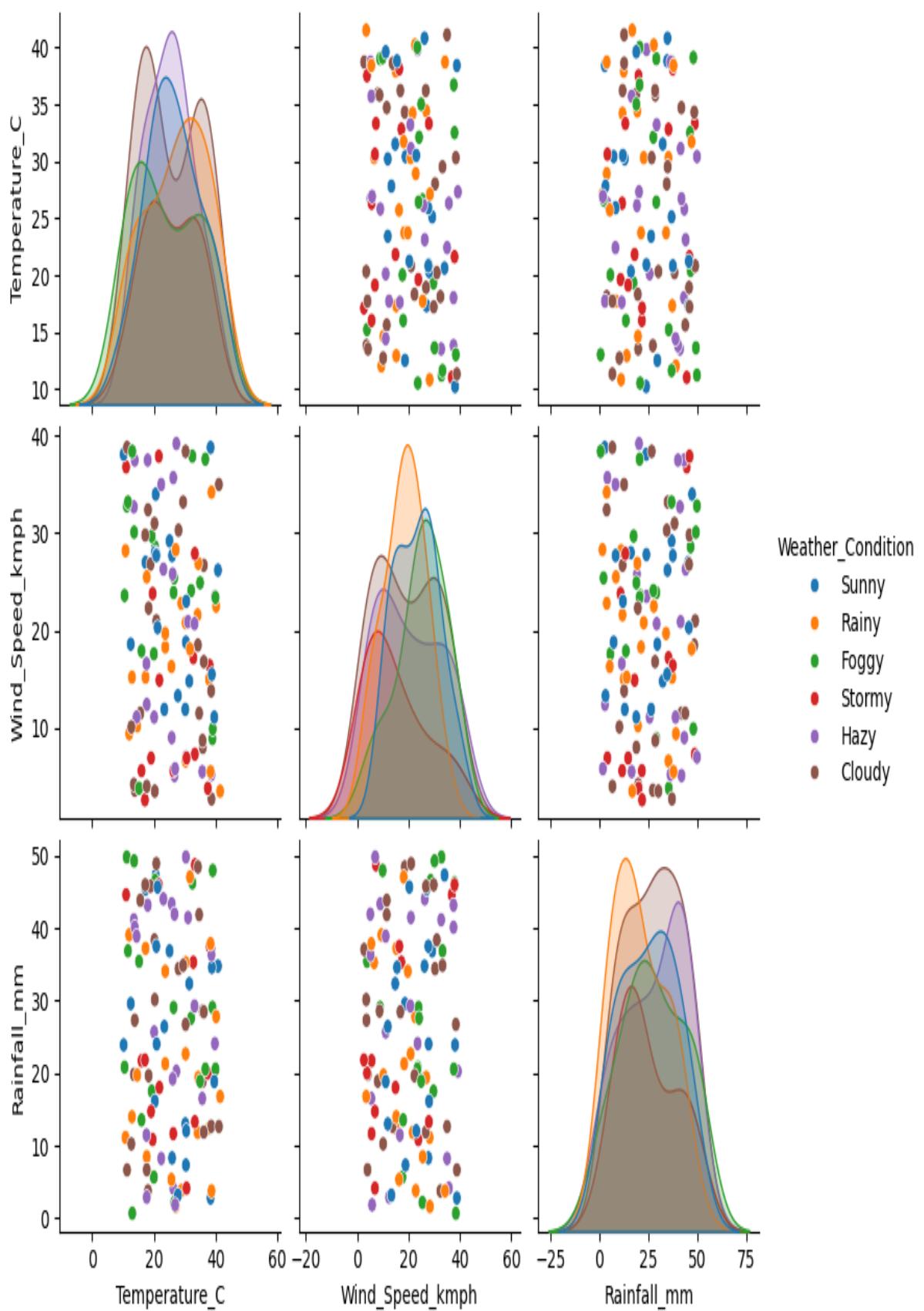












New Shape: (120, 6)

