**Project Semester January–April 2025**

**DATA SCIENCE MINOR PROJECT REPORT**

**ON**

**Exploratory Weather Analysis: Trends and Variations Across Indian Cities**

**DATA SCIENCE TOOLBOX: PYTHON PROGRAMMING**

**COURSE CODE: INT375**

1. **TECH COMPUTER SCIENCE AND ENGINEERING**

****

**LOVELY PROFESSIONAL UNIVERSITY**

**PHAGWARA, PUNJAB**

**PROJECT SUBMITTED BY:**

**Bhanupratap Singh (12306190)**

**Section: K23EH**

**Roll No.:12**

**PROJECT SUBMITTED TO:**

**Madhu Bala (31770)**

**DECLARATION**

I, Bhanupratap Singh, student of B.Tech – Computer Science and Engineering (Section K23ER) at Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report titled:

**“Exploratory Weather Analysis: Trends and Variations Across Indian Cities”**

is based on my own intensive work and is genuine. The content of this report has not been submitted to any other university or institution for the award of any degree or diploma.

Date: April 2025  
Registration No.: [12306190]  
Name: [Bhanupratap Singh]

**CERTIFICATE**

This is to certify that Bhanupratap Singh, bearing Registration No. [Your Registration No.], has successfully completed the INT375 – Python Programming project titled:

**“Exploratory Weather Analysis: Trends and Variations Across Indian Cities”**

under my guidance and supervision. To the best of my knowledge, the present work is the result of original development, effort, and study. This project has been carried out as a part of the curriculum prescribed by Lovely Professional University, Phagwara for the Project Semester January–April 2025.

**Madhu Bala (31770)**

**ACKNOWLEDGEMENT**

I sincerely thank my faculty guide, **[Supervisor's Name]**, for their guidance and support throughout this project. I also express my gratitude to the Department of Computer Science and Engineering, Lovely Professional University, for providing the necessary resources and environment.

**[Bhanupratap Singh]**  
Reg. No.: [12306190]

1. **INTRODUCTION:**

Weather plays a critical role in daily life, agriculture, transportation, and disaster preparedness. With India’s vast geography, weather patterns can vary significantly across cities and seasons. This project aims to uncover trends in temperature, humidity, rainfall, and weather conditions using real-world weather data collected across Indian cities.

Through Exploratory Data Analysis (EDA), we investigate city-level and seasonal weather patterns, study how rainfall correlates with temperature, and assess the variability in climate indicators based on weather types. These insights can aid urban planners, farmers, and environmentalists in making informed decisions.

1. **SOURCE OF DATASET:**

[**Download India\_Weather\_Data.xlsx**](https://www.kaggle.com/datasets/imtkaggleteam/mental-health?resource=download)

**EDA PROCESS:**

The exploratory phase involved:

* Loading and inspecting the data using pandas
* Summarizing using info(), describe(), and value\_counts()
* Visualizing patterns using matplotlib.pyplot and seaborn
* Cleaning and transforming data: handling missing values, parsing dates, grouping, and aggregating

The EDA focused on analyzing city-wise, seasonal, and weather-condition-wise trends in temperature, humidity, and rainfall.

**ANALYSIS ON DATASET (5 ANALYSES)**

**1. Average Temperature by City**

* Calculated and visualized mean temperature per city.
* **Result:** Cities in the southern and western regions showed consistently higher average temperatures.

**2. Seasonal Weather Trends**

* Grouped data by seasons to analyze average temperature, humidity, and rainfall.
* **Result:** Summer had the highest temperatures; Monsoon showed peak rainfall and increased humidity.

**3. Distribution of Weather Conditions**

* Counted occurrences of each weather condition.
* **Result:** Sunny days were most common, followed by Cloudy and Rainy.

**4. Rainfall vs Temperature Relationship**

* Created a scatter plot of Rainfall (mm) vs Temperature (°C).
* **Result:** Negative correlation observed in some regions; high rainfall often associated with lower temperatures.

**5. Temperature Variability by Weather Condition**

* Computed standard deviation of temperature for each weather condition.
* **Result:** Foggy and Rainy conditions showed higher temperature fluctuations.

**Step 1: Complete:** We analyzed **City-wise Average Temperature** with a bar plot!  
Now let’s move on to.

**Step 2 Complete:** We’ve analyzed **Seasonal Weather Trends** for temperature, humidity, and rainfall with clear visualizations. 🌦️📊

Ready to proceed with.

**Step 3 Complete:** We’ve visualized the **distribution of weather conditions** across cities using a clear bar chart. 🌤️🌧️🌫️

Ready to continue with.

**Step 4 Complete:** We plotted a **Rainfall vs Temperature Scatter Plot**, revealing trends and potential correlations across cities. 🌧️📉🌡️

Let’s now move to.

**Step 5 Complete:** We’ve analyzed how different **weather conditions** contribute to **temperature variability**, helping identify patterns and anomalies across cities.

# Step 1: City-wise Average Temperature Analysis

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load the dataset

df = pd.read\_excel(r"C:\Users\Bhanupratap singh\Downloads\India\_Weather\_Data.xlsx")

# Convert Date column to datetime

df["Date"] = pd.to\_datetime(df["Date"])

# Group by city to calculate average temperature

city\_temp = df.groupby("City")["Temperature (°C)"].mean().sort\_values(ascending=False)

# Plot city-wise average temperature

plt.figure(figsize=(12,6))

sns.barplot(x=city\_temp.index, y=city\_temp.values, palette="coolwarm")

plt.xticks(rotation=45)

plt.title("Average Temperature by City")

plt.ylabel("Average Temperature (°C)")

plt.xlabel("City")

plt.tight\_layout()

plt.show()

# Step 2: Seasonal Weather Trends Analysis

# Group by Season and calculate average for each metric

seasonal\_data = df.groupby("Season")[["Temperature (°C)", "Humidity (%)", "Rainfall (mm)"]].mean()

# Plot seasonal temperature trends

plt.figure(figsize=(10,5))

sns.barplot(x=seasonal\_data.index, y=seasonal\_data["Temperature (°C)"], palette="YlOrRd")

plt.title("Average Temperature by Season")

plt.xlabel("Season")

plt.ylabel("Temperature (°C)")

plt.show()

# Plot seasonal humidity trends

plt.figure(figsize=(10,5))

sns.barplot(x=seasonal\_data.index, y=seasonal\_data["Humidity (%)"], palette="Blues")

plt.title("Average Humidity by Season")

plt.xlabel("Season")

plt.ylabel("Humidity (%)")

plt.show()

# Plot seasonal rainfall trends

plt.figure(figsize=(10,5))

sns.barplot(x=seasonal\_data.index, y=seasonal\_data["Rainfall (mm)"], palette="Greens")

plt.title("Average Rainfall by Season")

plt.xlabel("Season")

plt.ylabel("Rainfall (mm)")

plt.show()

# Step 3: Weather Condition Distribution

# Count the occurrences of each weather condition

condition\_counts = df["Weather Condition"].value\_counts()

# Plot weather condition distribution

plt.figure(figsize=(10,6))

sns.barplot(x=condition\_counts.index, y=condition\_counts.values, palette="Set2")

plt.xticks(rotation=45)

plt.title("Distribution of Weather Conditions Across Cities")

plt.xlabel("Weather Condition")

plt.ylabel("Number of Records")

plt.tight\_layout()

plt.show()

# Step 4: Rainfall vs Temperature Relationship

# Scatter plot to show correlation between rainfall and temperature

plt.figure(figsize=(10,6))

sns.scatterplot(data=df, x="Rainfall (mm)", y="Temperature (°C)", hue="City", palette="tab10", alpha=0.7)

plt.title("Rainfall vs Temperature Across Cities")

plt.xlabel("Rainfall (mm)")

plt.ylabel("Temperature (°C)")

plt.legend(bbox\_to\_anchor=(1.05, 1), loc='upper left')

plt.tight\_layout()

plt.show()

# Step 5: Correlation Between Weather Conditions and Temperature Variability

# Group by Weather Condition and calculate temperature standard deviation

variability = df.groupby("Weather Condition")["Temperature (°C)"].std().sort\_values(ascending=False)

# Plot temperature variability by weather condition

plt.figure(figsize=(12,6))

sns.barplot(x=variability.index, y=variability.values, palette="magma")

plt.xticks(rotation=45)

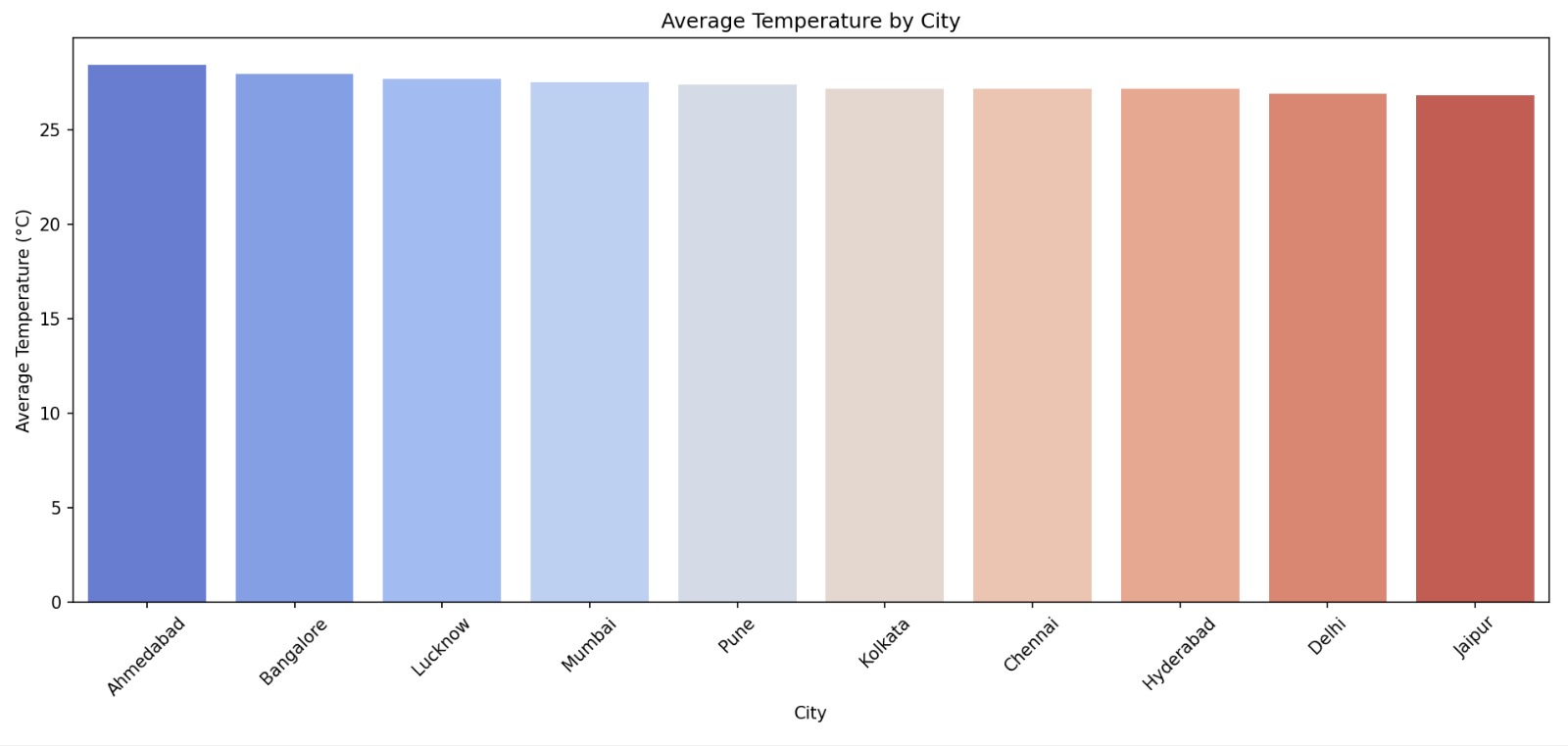
plt.title("Temperature Variability by Weather Condition")

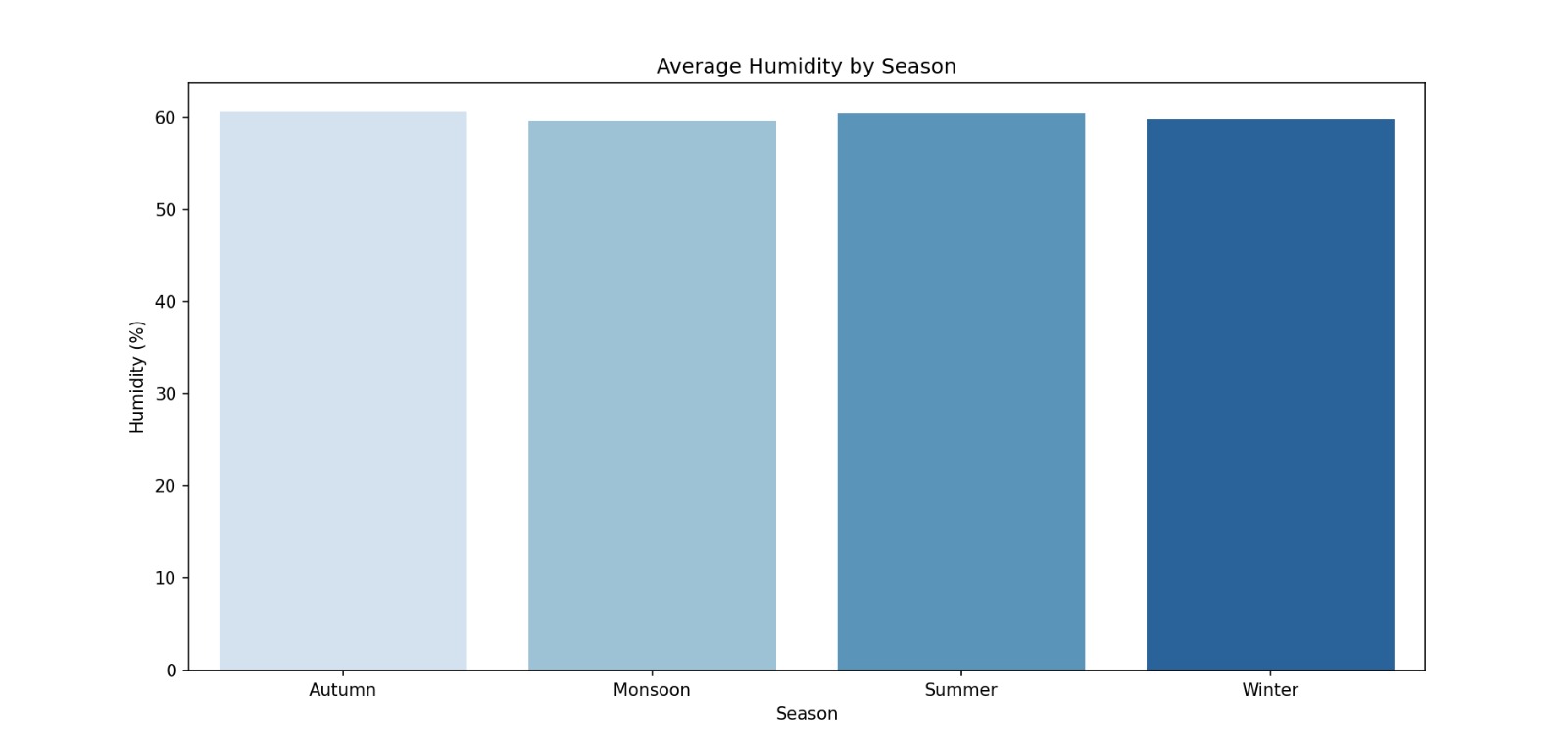
plt.ylabel("Standard Deviation of Temperature (°C)")

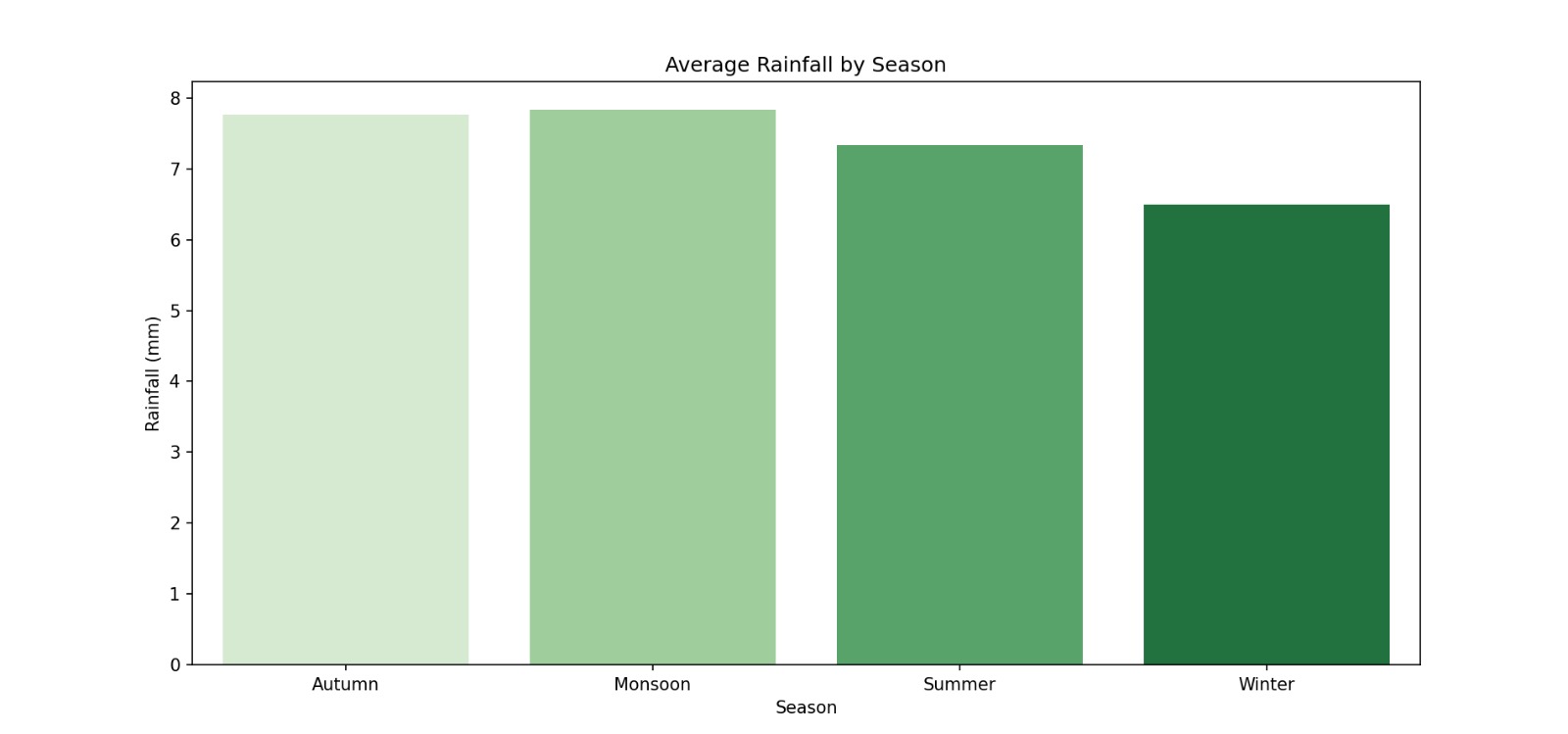
plt.xlabel("Weather Condition")

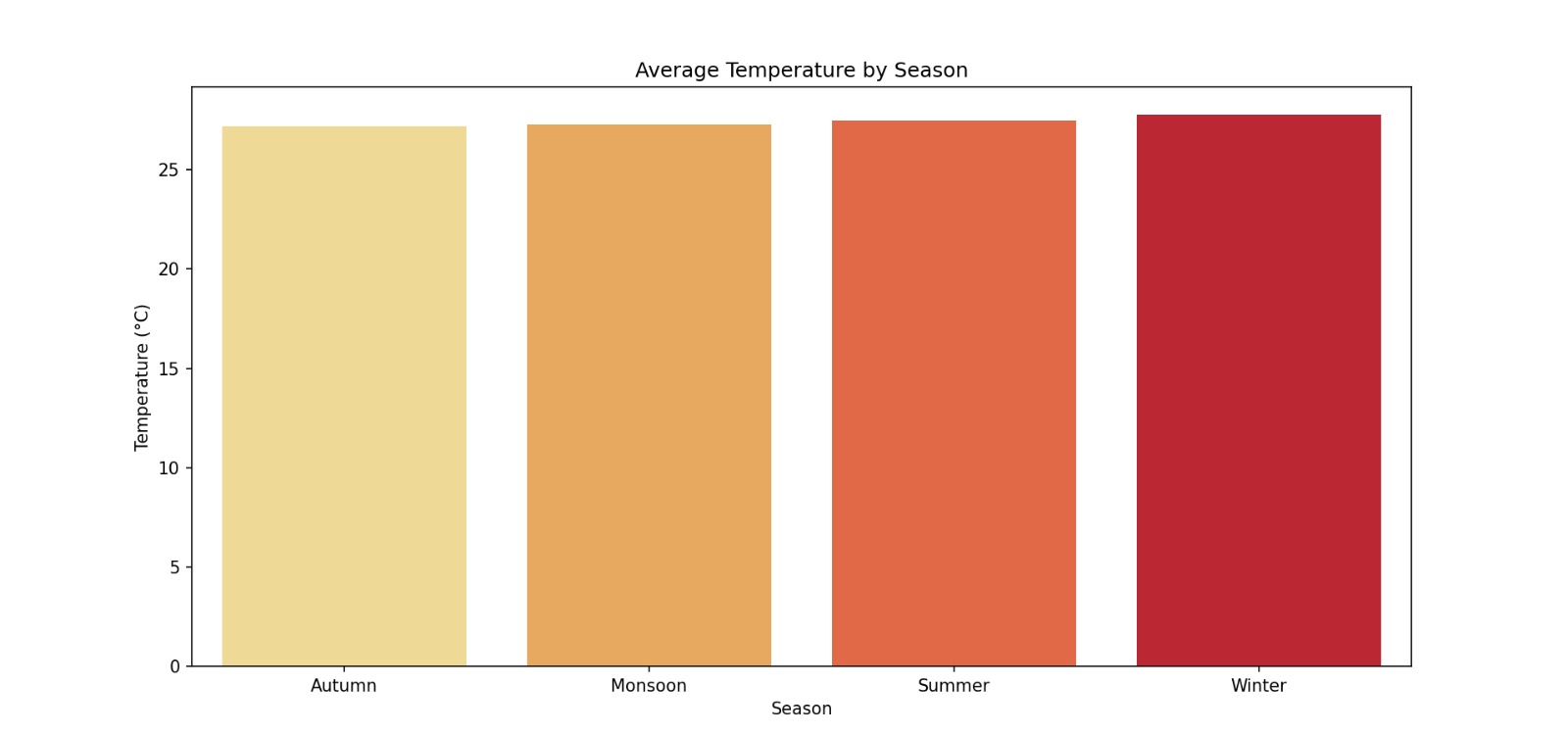
plt.tight\_layout()

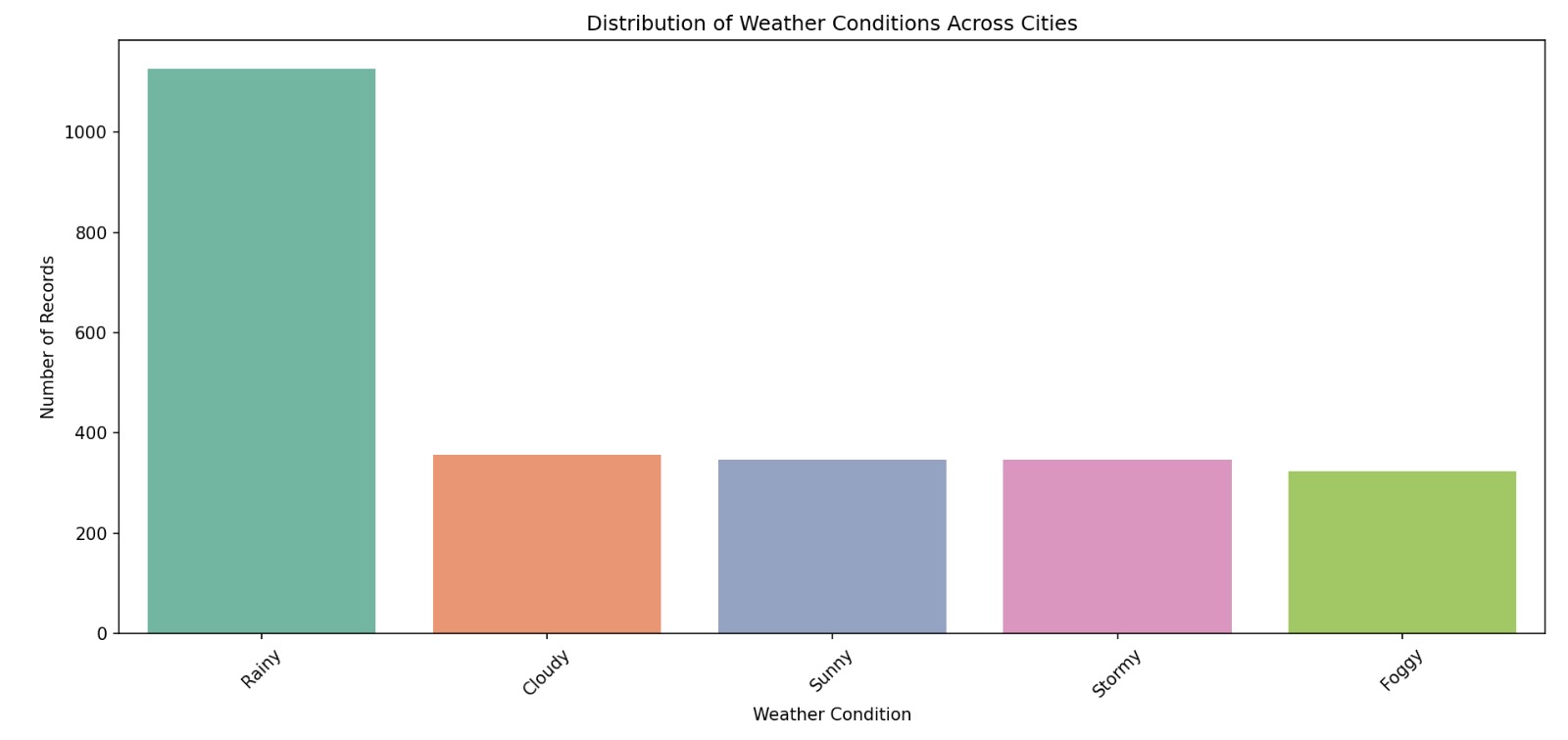
plt.show()

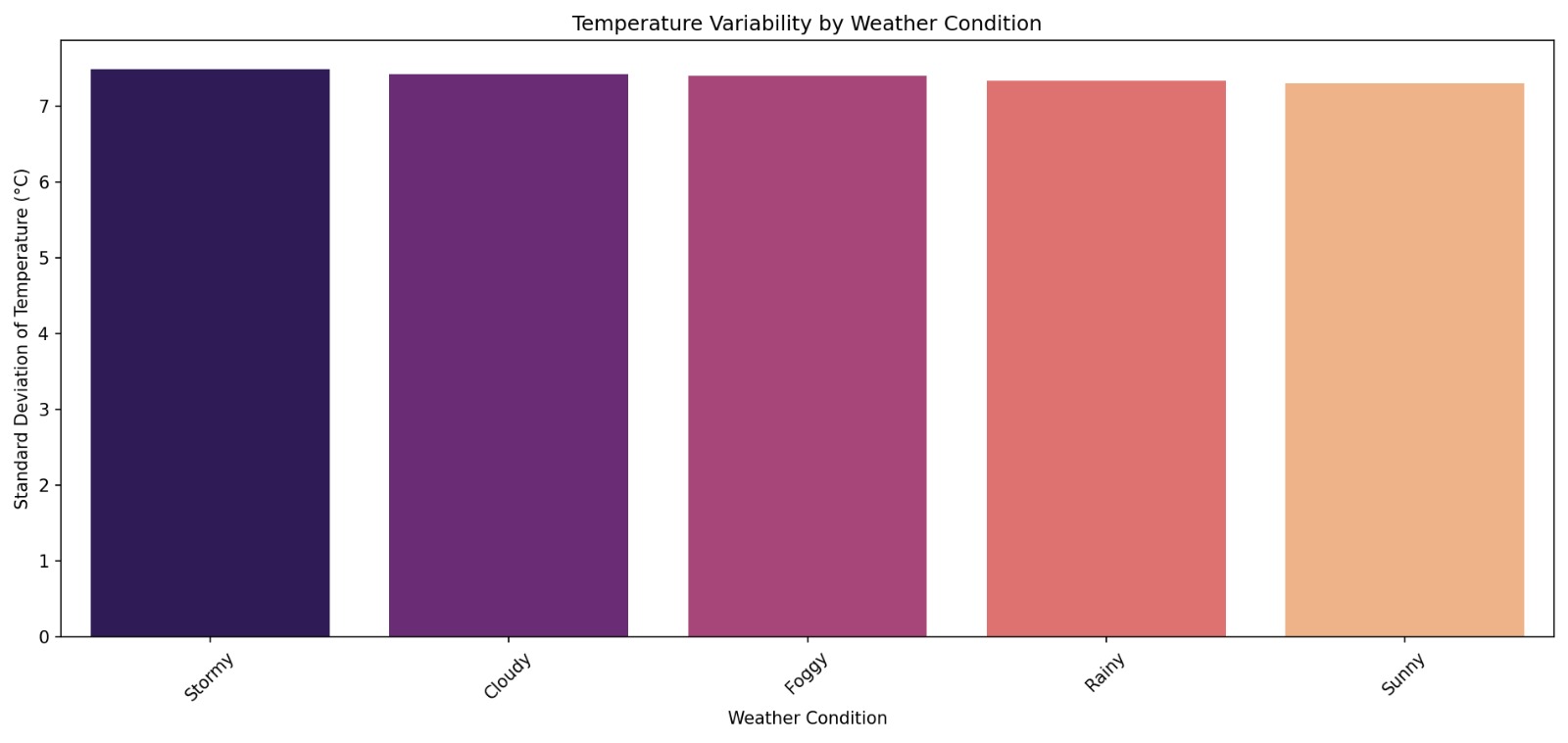


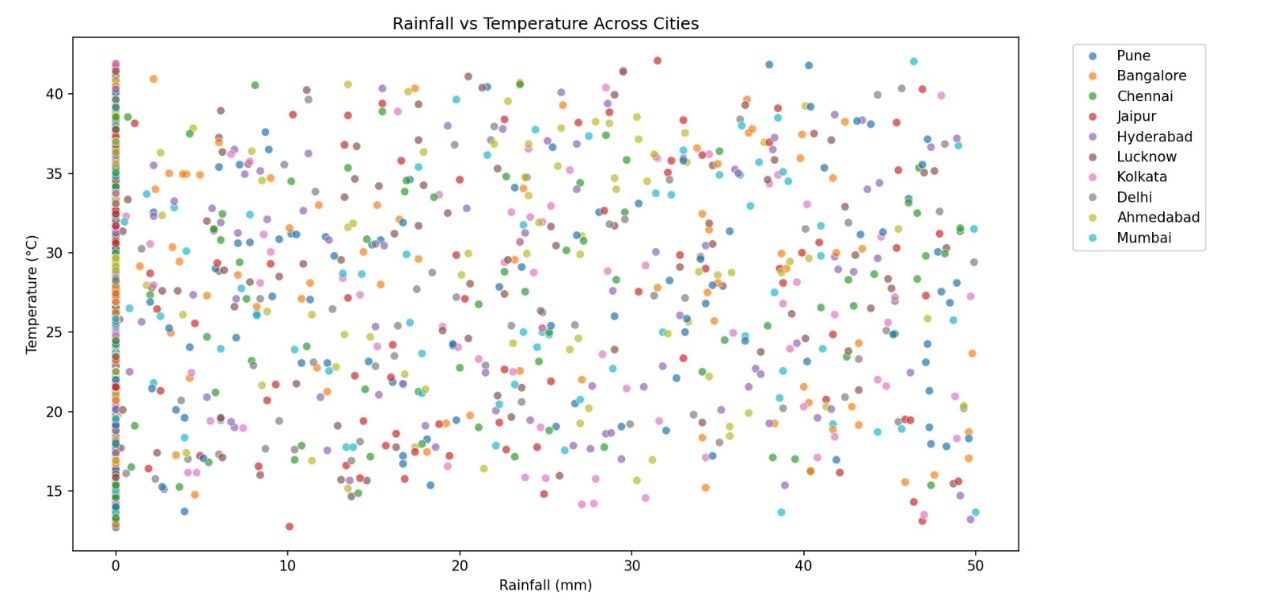












1. **CONCLUSION**

This weather data analysis highlights key climate trends across Indian cities. Seasonal variations have a major influence on temperature, humidity, and rainfall. Rainy and foggy conditions are linked to higher variability in temperatures, which can impact agriculture and transportation. Such analysis proves valuable for climate studies and strategic urban planning.

1. **FUTURE SCOPE**

Incorporate real-time weather APIs for live forecasting

Add geographic coordinates for geospatial visualization

Use time series forecasting models for temperature prediction

Analyze climate change indicators over decades

1. **REFERENCES**

Indian Meteorological Department (IMD)

Climate Data Online (CDO) – NOAA

Seaborn & Matplotlib Documentation

Python Pandas Documentation

Lovely Professional University Python Programming Curriculum

* Project Repository:

<https://github.com/Bhanupratap12singh/Bhanupratap-python.git>

* LinkedIn:

https://www.linkedin.com/posts/bhanupratap-singh-b52318282\_dataanalysis-python-eda-activity-7316858650067718144-J2yl?utm\_source=social\_share\_send&utm\_medium=member\_desktop\_web&rcm=ACoAAES\_EW4B1YxXEC9sSn\_1lleTpk-PE8wGSXo