

### **Interpretation of the Weather Feature Histograms**

This image consists of multiple histograms, each representing the distribution of different weather parameters. Below is an interpretation of each variable:

#### **Key Observations by Feature:**

1. **Pressure:**
   * Most values are between **1005 and 1025 hPa**.
   * The distribution is slightly skewed, indicating variations in atmospheric pressure.
2. **Max Temperature:**
   * Peaks around **30-35°C**, suggesting a warm climate.
   * The spread is between **10-35°C**, showing seasonal variations.
3. **Temperature:**
   * Follows a similar pattern to max temperature, with a peak near **30°C**.
4. **Min Temperature:**
   * Skewed distribution with most values between **10-25°C**.
5. **Dewpoint:**
   * Most values are between **10-25°C**, with a sharp rise near 25°C.
   * Indicates moisture presence in the air.
6. **Humidity:**
   * Most values range between **60-90%**, with peaks near **80-90%**.
   * Suggests a generally humid climate.
7. **Cloud Cover:**
   * High concentration at **80-100%**, suggesting frequent cloudiness.
8. **Sunshine Duration:**
   * Skewed towards **0-3 hours**, indicating limited sunshine periods.
9. **Wind Direction:**
   * Spread across various angles, with peaks around **50°, 100°, 200°, and 250°**, indicating dominant wind directions.
10. **Wind Speed:**

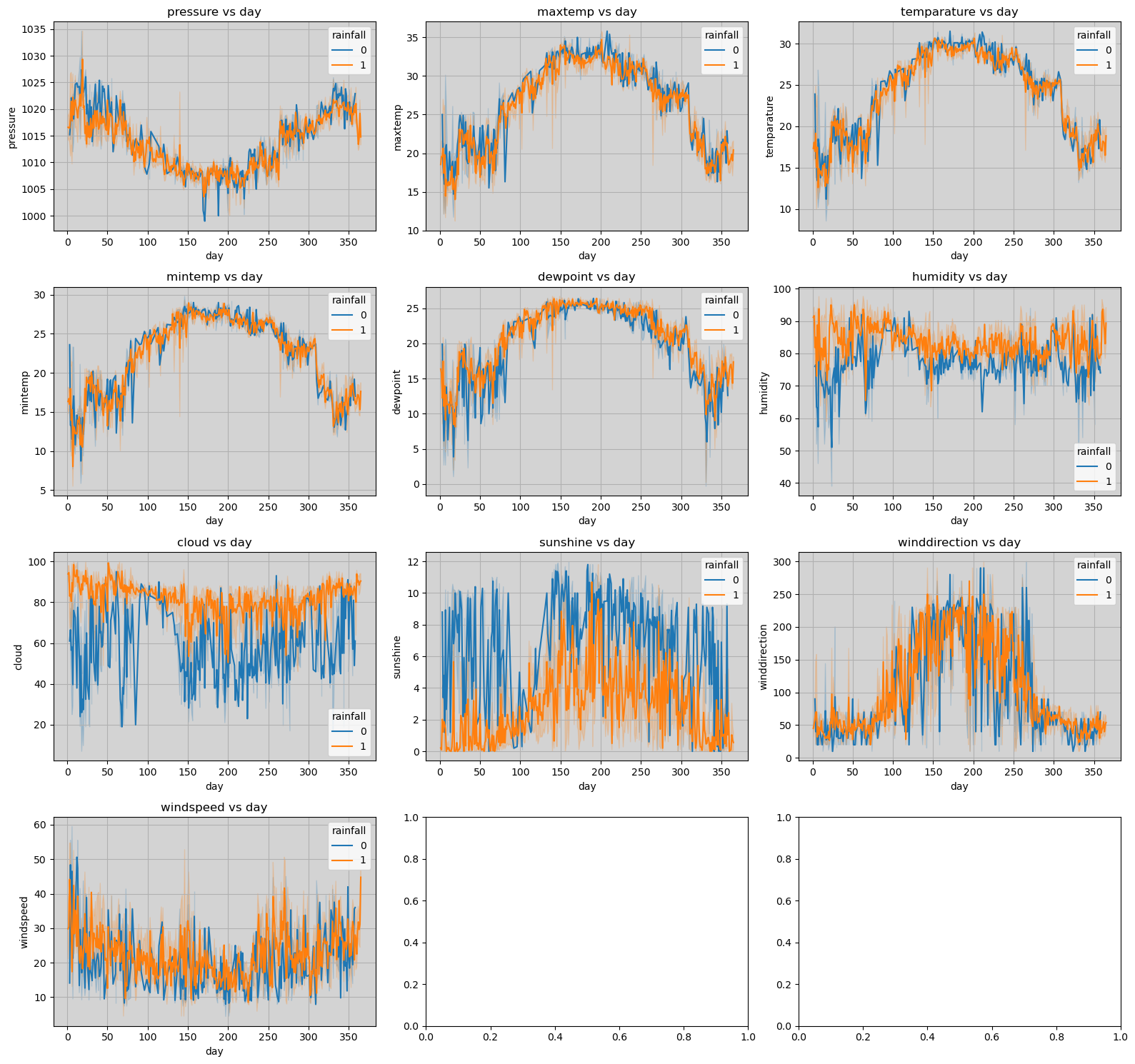
* Ranges between **5-50 km/h**, with peaks around **10-20 km/h**.
* Suggests moderate wind speeds.

1. **Rainfall:**

* Most values are concentrated at **0 and 1**, indicating that rainfall is either absent or occurs in discrete events.

### **Conclusion:**

* The data suggests a humid climate with moderate temperatures and frequent cloud cover.
* Sunshine hours are often low, while rainfall occurs sporadically.
* Wind direction varies, but specific directions dominate.
* Pressure remains stable with slight fluctuations.



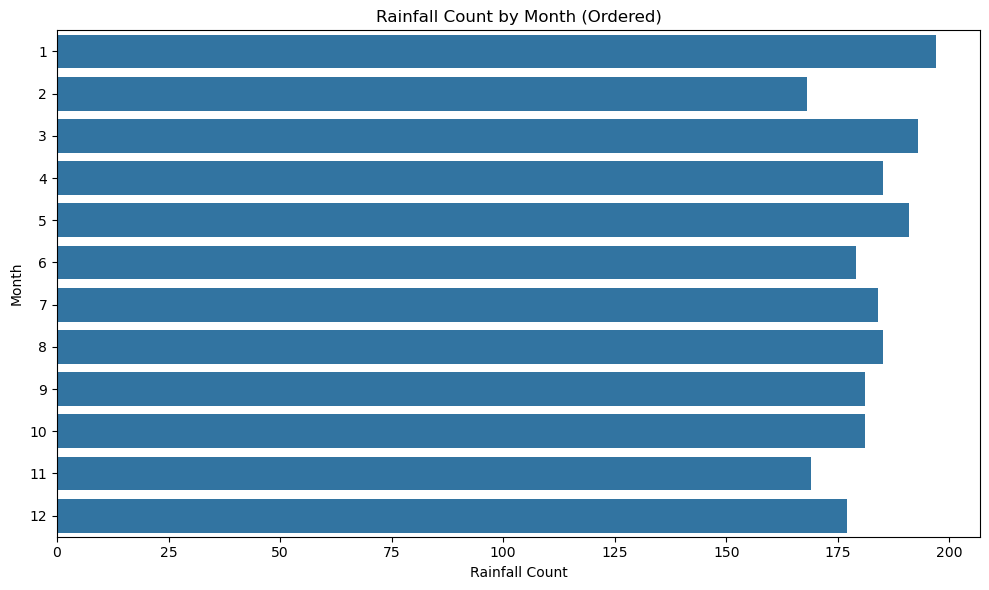
1. Pressure vs. Day:Pressure fluctuates between approximately 1000 and 1035 hPa.No significant trend is observed, but there are periodic drops, especially noticeable around days 100-150 and 250-300, which might correlate with rainfall events.
2. Maximum Temperature vs. Day:Maximum temperature ranges from about 15°C to 35°C, showing a clear seasonal pattern with a peak around days 150-200 (summer) and a dip around days 300-350 (winter).Rainfall (orange) tends to lower the maximum temperature compared to no rainfall (blue).
3. Temperature vs. Day:Similar to maximum temperature, it follows a seasonal trend with a peak around mid-year and a decline toward the year's end.Rainfall slightly reduces the average temperature.
4. Minimum Temperature vs. Day:Ranges from 5°C to 25°C, with a similar seasonal pattern.Rainfall seems to have a less pronounced effect on minimum temperatures compared to maximum temperatures.
5. Dew Point vs. Day:Dew point varies between 5°C and 20°C, following a seasonal trend.Rainfall increases the dew point, indicating higher moisture levels.
6. Humidity vs. Day:Humidity ranges from 40% to 100%, with higher values during rainy periods (orange).A seasonal trend is less clear, but humidity spikes are more frequent with rainfall.
7. Cloud vs. Day:Cloud cover ranges from 20 to 100, with higher values during rainy days, suggesting more cloudiness with rainfall.
8. Sunshine vs. Day:Sunshine hours range from 0 to 12 hours, peaking around mid-year and dropping toward the year's end.Rainfall significantly reduces sunshine hours.
9. Wind Speed vs. Day:Wind speed fluctuates between 10 and 50 units (possibly m/s), with no strong seasonal trend.Rainfall appears to increase wind speed slightly.
10. Wind Direction vs. Day:Wind direction varies widely (0 to 300 degrees), with no clear seasonal pattern.Rainfall does not seem to significantly alter wind direction.

Scatter Plots:

* The three scatter plots (sunshine vs. rainfall, wind speed vs. rainfall, wind direction vs. rainfall) show weak or no correlation between these variables and rainfall, as the data points are scattered without a clear trend.

General Insights:

* The data suggests a strong seasonal influence on temperature-related variables (max temp, min temp, dew point), with a peak in the middle of the year.
* Rainfall consistently affects meteorological conditions, generally increasing humidity, cloud cover, dew point, and wind speed while decreasing temperature and sunshine.
* The lack of strong correlations in the scatter plots indicates that rainfall's impact on sunshine, wind speed, and wind direction might be context-dependent or influenced by other factors not shown here.

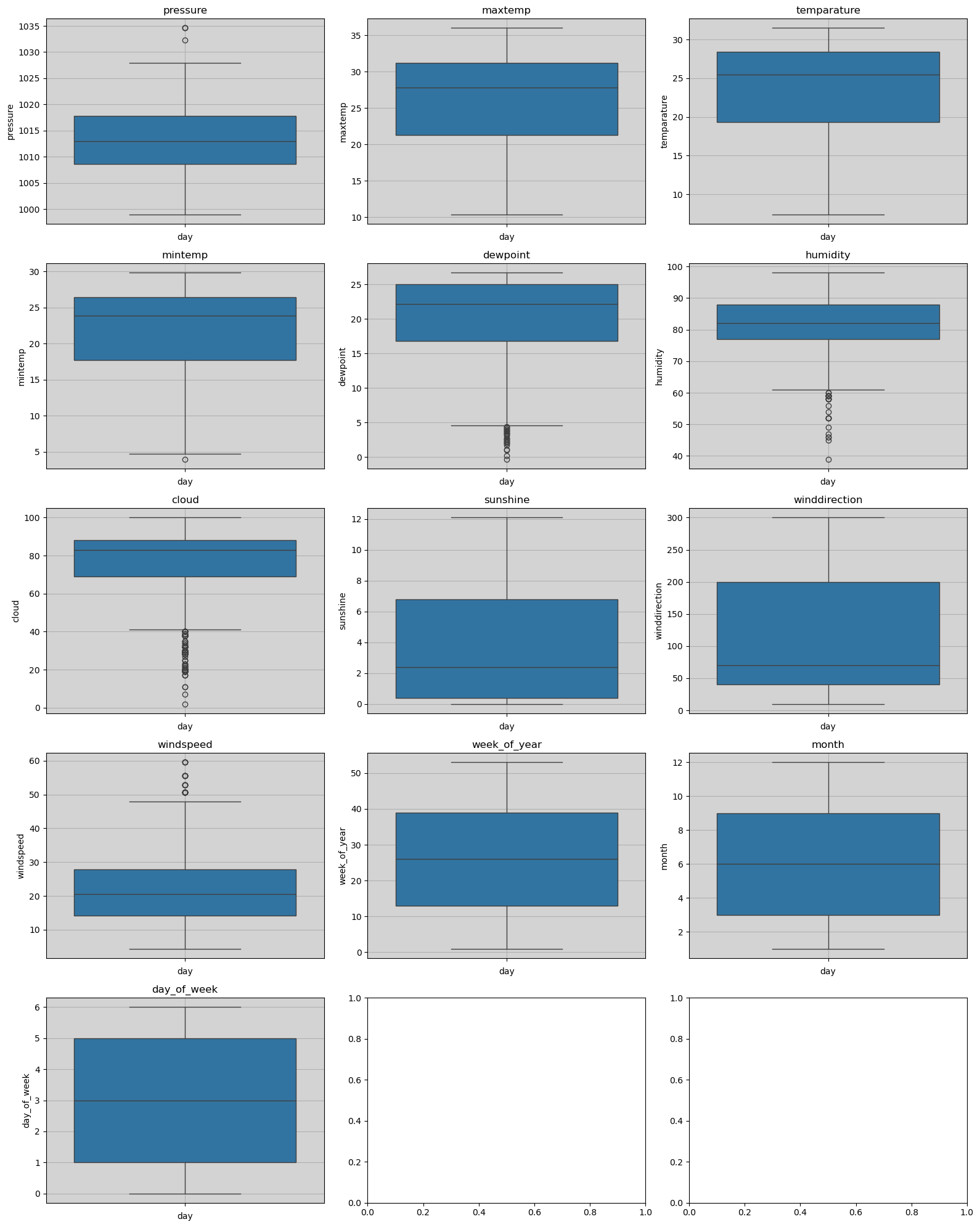


**Data Trends:**

* The rainfall appears to be fairly consistent throughout the year, with no extremely low or high values.
* Months 1 and 3 show slightly higher rainfall counts.
* Month 2 and 11 appear to have the lowest rainfall counts.
* The rainfall count is between 160 and 200 for all months.

**Observations**

* The data includes average values for pressure, temperature, dewpoint, humidity, and rainfall across different months.
* Month 4 has the highest rainfall (0.924324), while Month 7 has the lowest (0.641304).
* Temperature and dewpoint values are highest in Month 7 (29.508152 and 25.266848, respectively), indicating warmer and more humid conditions.
* The pressure values range from 1007.841304 (Month 7) to 1019.794924 (Month 1), suggesting variations in atmospheric pressure.

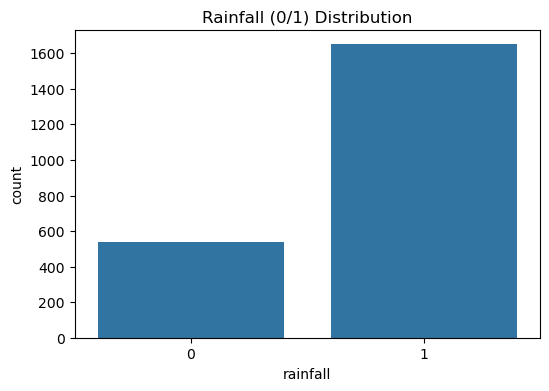


**Visual Description**

* Pressure vs. Day: Pressure values range from approximately 1000 to 1025 hPa, with a median around 1015 hPa. There are a few outliers below 1005 hPa.
* Max Temp vs. Day: Maximum temperature ranges from around 10°C to 35°C, with a median near 25°C.
* Temperature vs. Day: Temperature values range from about 10°C to 30°C, with a median around 22°C.
* Min Temp vs. Day: Minimum temperature ranges from 5°C to 25°C, with a median around 15°C.
* Dewpoint vs. Day: Dewpoint values range from 0°C to 25°C, with a median near 15°C. There are outliers below 5°C.
* Humidity vs. Day: Humidity ranges from 40% to 100%, with a median around 80%. Outliers are present below 60%.
* Cloud vs. Day: Cloud cover ranges from 0 to 100, with a median around 50. Outliers are present below 20.
* Sunshine vs. Day: Sunshine hours range from 0 to 12, with a median around 5.
* Wind Direction vs. Day: Wind direction ranges from 0 to 300 degrees, with a median around 150 degrees.
* Wind Speed vs. Day of Week: Wind speed ranges from 0 to 50, with a median around 30. Outliers are present above 50.
* Week of Year vs. Day: Week of the year ranges from 0 to 50, with a median around 25.
* Month vs. Day: Month ranges from 1 to 12, with a median around 6.
* Day of Week vs. Day: Day of the week ranges from 0 to 6, with a median around 3.
* Rainfall vs. Day (Bottom Two Plots): The last two plots (likely rainfall-related) show values between 0 and 1, suggesting binary or normalized rainfall data (e.g., 0 for no rain, 1 for rain). Both have medians around 0.5, with some spread.

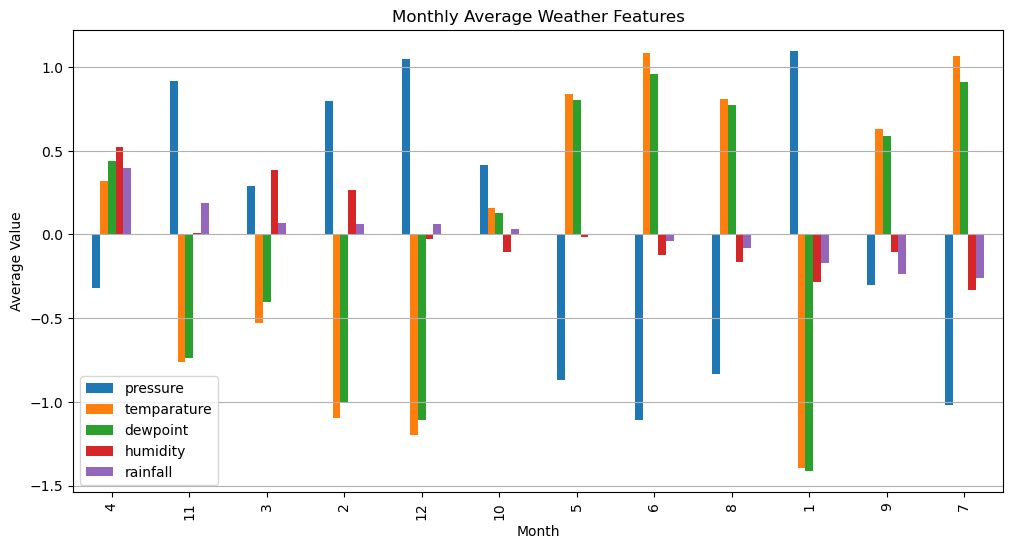
**Observations**

* Pressure: Mostly stable around 1015 hPa, with minor variations across days and a few low-pressure outliers.
* Temperature (Max, Min, Mean): Temperature-related features show significant variation, with maximum temperatures peaking at 35°C and minimum temperatures dropping to 5°C.
* Humidity and Dewpoint: High humidity (median ~80%) and dewpoint (median ~15°C) suggest humid conditions, with some drier outliers.
* Cloud and Sunshine: Cloud cover and sunshine hours show a wide range, indicating variable weather conditions.
* Wind: Wind direction and speed are variable, with wind speeds showing outliers above 50.
* Temporal Features (Month, Week, Day of Week): These features are distributed as expected, with medians reflecting the middle of the year or week.
* Rainfall: The binary-like distribution (0 to 1) suggests this dataset might be used for binary classification (rain vs. no rain), with roughly equal occurrences of both outcomes.



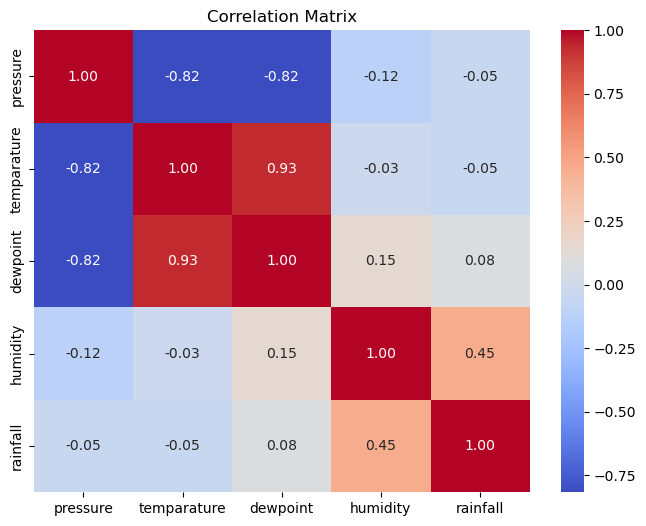
**Observations**

* The dataset contains a total of around 2000 data points (500 + 1500).
* The distribution is imbalanced:
  + No Rain (0): Approximately 500 instances, which is about 25% of the total data.
  + Rain (1): Approximately 1500 instances, which is about 75% of the total data.
* This imbalance suggests that the dataset has significantly more instances of rainy days compared to non-rainy days.
* For a binary classification task this imbalance may affect model performance, potentially leading to a bias towards predicting rain. Techniques like oversampling, undersampling, or using class weights might be necessary to address this issue.



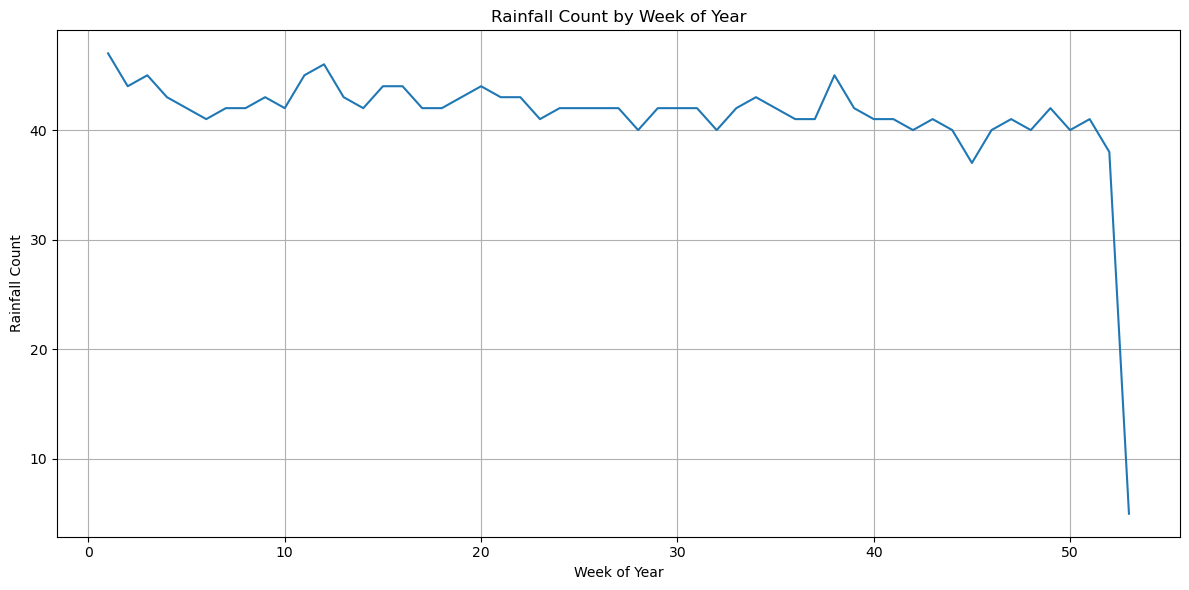
**Observations**

* Pressure: Higher in colder months (January, February, November, December) with values around 0.5, and lower in warmer months (June, July, August) with values around -0.5.
* Temperature and Dewpoint: Both features follow a similar pattern, peaking in July (around 1.0) and dipping in January and December (around -1.0), indicating a strong seasonal trend.
* Humidity: Relatively stable across months, with slight positive values (around 0.2 to 0.5) in most months, peaking in April.
* Rainfall: Peaks in April and November (around 0.5), with lower values in colder months (January, February, December).
* Seasonal Trends: The chart shows clear seasonal patterns, with warmer months (June to August) having higher temperatures and dewpoints, lower pressure, and moderate rainfall, while colder months (December to February) have higher pressure and lower temperatures.



**Observations**

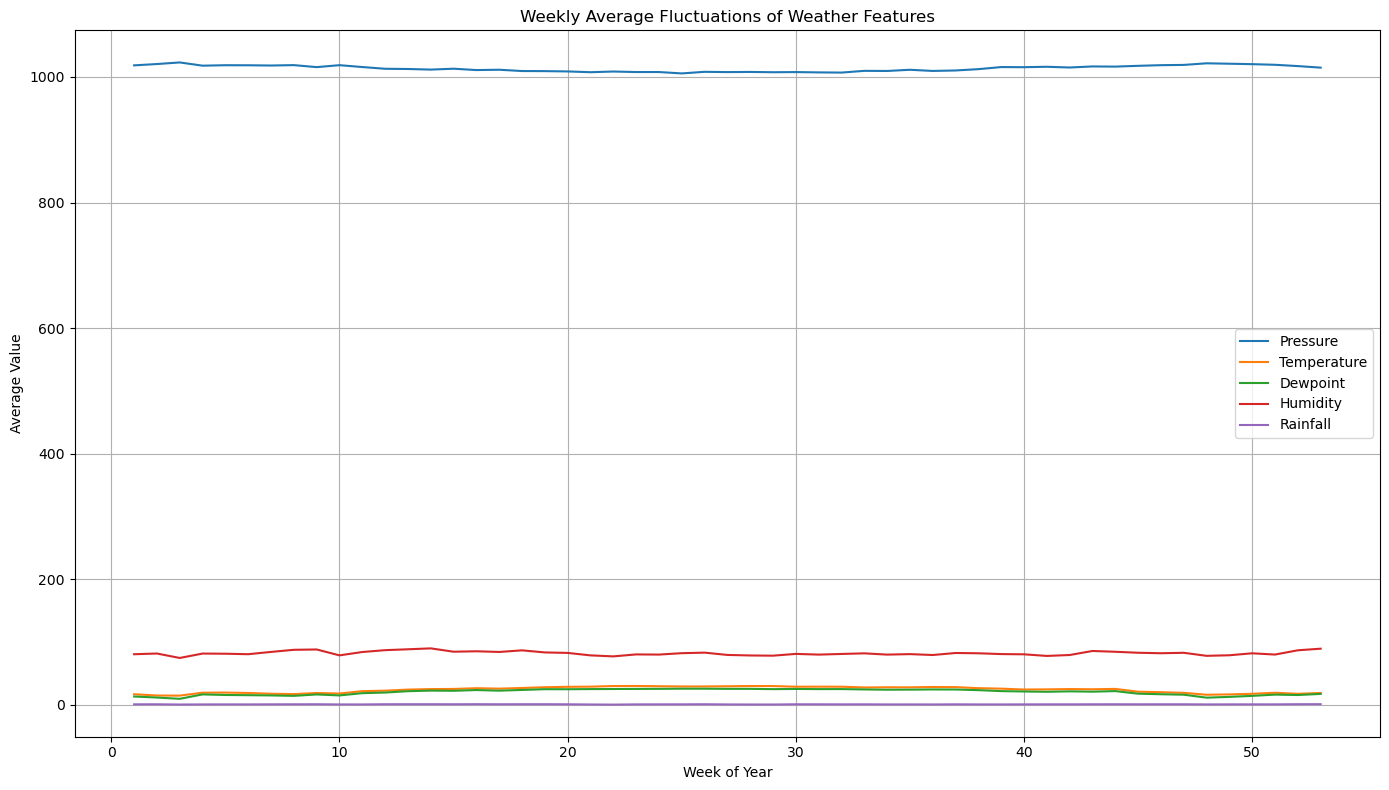
* Strong Correlations:
  + Temperature and Dewpoint: A very strong positive correlation (0.93) indicates that as temperature increases, dewpoint tends to increase as well. This makes sense, as dewpoint is closely related to temperature and moisture in the air.
  + Pressure with Temperature and Dewpoint: Strong negative correlations (-0.82 for both) suggest that higher pressure is associated with lower temperatures and dewpoints, often indicative of cooler, drier conditions.
* Moderate Correlation:
  + Humidity and Rainfall: A moderate positive correlation (0.45) indicates that higher humidity is associated with a higher likelihood of rainfall, which aligns with meteorological expectations.
* Weak Correlations:
  + Pressure and Humidity (-0.12): A weak negative correlation suggests a slight tendency for higher pressure to be associated with lower humidity.
  + Dewpoint and Humidity (0.15): A weak positive correlation indicates a mild relationship between these variables.
  + Rainfall with Temperature (-0.05) and Dewpoint (0.08): Very weak correlations suggest that rainfall is not strongly influenced by temperature or dewpoint in this dataset.
* **Implications:**
  + The strong correlation between temperature and dewpoint suggests potential multicollinearity, which may need to be addressed in a predictive model (e.g., by removing one of the features or using dimensionality reduction techniques like PCA).
  + The moderate correlation between humidity and rainfall suggests that humidity could be a useful predictor for rainfall.
  + The weak correlations between rainfall and other features (pressure, temperature, dewpoint) indicate that these features alone may not be strong predictors of rainfall, and additional features or transformations might be needed.



**Observations**

* Seasonal Pattern: The graph suggests a seasonal trend in rainfall, with higher counts in the earlier and middle parts of the year (weeks 1-40), and a significant decrease towards the end of the year (weeks 45-52).
* Peak Rainfall: The highest rainfall counts occur around weeks 10, 20, and 35-40, indicating possible rainy seasons or periods of higher precipitation.
* Decline: The sharp decline after week 45 suggests a dry season or period with little to no rainfall towards the end of the year.
* Stability: Despite fluctuations, the rainfall count remains relatively stable between 30 and 45 for a significant portion of the year (weeks 5-45).

**Implications:** This pattern could be useful for predicting rainfall, with the end of the year being a period of lower rainfall likelihood. Further analysis could correlate this with temperature, humidity, or other weather features.



**Interpretation of the Weather Features Graph**

This line graph displays the weekly average fluctuations of different weather features over a full year (52 weeks). The key weather parameters included are:

1. Pressure (hPa) - The dominant feature with values around 1000 hPa, showing minor variations.
2. Temperature (°C) - Maintains a consistent trend with slight seasonal variations.
3. Dewpoint (°C) - Generally follows the temperature trend but stays at a lower magnitude.
4. Humidity (%) - Shows moderate fluctuations, typically staying around 80-100%.
5. Rainfall (mm) - Remains at low values with small peaks, indicating periodic rainfall.

Key Observations:

1. Pressure remains relatively stable throughout the year.
2. Temperature and Dewpoint exhibit seasonal changes, likely reflecting summer and winter cycles.
3. Humidity fluctuates slightly but remains relatively high.
4. Rainfall is the lowest among all variables, with periodic increases indicating wet seasons.

This visualization provides insight into how these weather parameters interact over time, useful for climate trend analysis and forecasting.