Weather Data Analysis using Python

Executive Summary

This project conducts a comprehensive analysis of a weather dataset to identify patterns and insights regarding temperature, wind speed, humidity, and other meteorological factors. Through data cleaning, aggregation, and exploration techniques, the project examines various metrics and uncovers notable trends and relationships between weather conditions.

Key Insights

1. **Wind Speed**: The data reveals diverse wind speed values, including rare occurrences of extreme winds.
2. **Clear Weather Occurrences**: Instances where the weather was classified as "clear" are identified, showing patterns and frequency.
3. **Visibility**: Average visibility and cases with poor visibility are explored.
4. **Pressure Analysis**: Analysis of atmospheric pressure provides insights into standard deviation and variance.
5. **Snow Conditions**: Instances with snowfall are extracted and analyzed.

Questions and Answers

Here's a list of questions explored in the notebook, along with brief answers:

1. **Unique Wind Speed Values**: What are all the unique wind speed values in the data?
    - Answer: Listed all unique wind speed values using `nunique()` and `unique()` functions.
2. **Clear Weather Instances**: How many times is the weather classified as "clear"?
    - Answer: Identified using `value_counts()`.
3. **Wind Speed of 4 km/h**: How often is the wind speed exactly 4 km/h?
    - Answer: Filtered for this specific wind speed and calculated occurrences.
4. **Missing Values**: Are there any null values in the dataset?
    - Answer: Used `isnull()` to identify missing data points.
5. **Rename Column**: Rename the 'Weather' column to 'Weather Condition'.
    - Answer: Renamed using the `rename()` method.
6. **Mean Visibility**: What is the mean visibility in kilometers?
    - Answer: Calculated the average visibility.

7. **Pressure Standard Deviation**: What is the standard deviation of pressure?
    - Answer: Calculated with `std()`.
8. **Humidity Variance**: What is the variance of relative humidity?
    - Answer: Computed with `var()`.
9. **Snow Occurrences**: When does snow occur?
    - Answer: Filtered to identify instances with "snow" in weather descriptions.
10. **High Wind and Low Visibility**: When is wind speed above 24 km/h and visibility exactly 25 km?

- Answer: Filtered to meet these conditions.

11. **Mean Values per Weather Condition**: What are the mean values of each column for each weather condition?

- Answer: Computed group means using `groupby()`.

12. **Max and Min by Weather Condition**: What are the maximum and minimum values of each column by weather condition?

- Answer: Calculated max and min values per condition.

13. **Fog Conditions**: What records show a foggy weather condition?

- Answer: Filtered for records with "Fog".

14. **Clear Weather or High Visibility**: When is the weather clear or visibility above 40?

- Answer: Used conditional filtering for visibility and clear conditions.

15. **High Humidity and Clear Weather or Visibility Above 40**: When is humidity high and weather clear, or visibility high?

- Answer: Applied complex conditions to find these instances.