**REPORT ON**

**WEATHER DATA SET**

**Import Necessary Libraries:**

**S**tart by importing the necessary Python libraries such as Pandas, Numpy and other libraries.

import pandas as pd

import numpy as np

and etc….,

**Load Given Data Set**

Load the weather dataset into a DataFrame and inspect it to understand its structure, available columns, and data types.

df = pd.read\_csv(“File Location of dataset”)

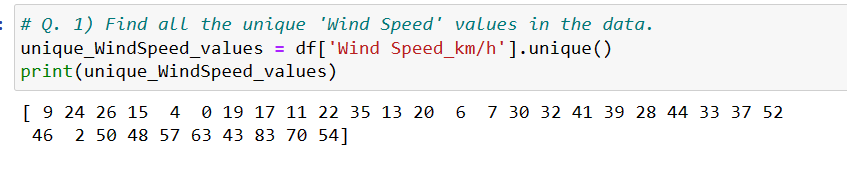
print(df.head()

**NOTE: This Analysis is done in the JupiterNoteBook ,there may be change in code when you use different platfroms.**

**Iterating on the Data Frame as per the Problem Statement**

**Q. 1) Find all the unique 'Wind Speed' values in the data.**

**Python Code:**

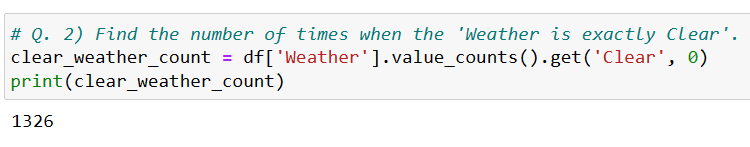


**Code Explanation:**

This code will extract the 'Wind Speed' column from weather DataFrame and then use the unique() method to find all the unique values in that column.

**Q. 2) Find the number of times when the 'Weather is exactly Clear'.**

**Python Code:**

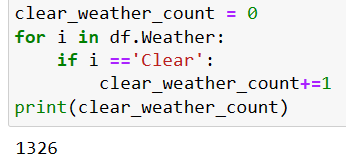


**Code Explanation:**

In this code, we use the value\_counts() method on the 'Weather' column to get a count of unique values in that column. Then, we use the get() method to retrieve the count for the 'Clear' value. If 'Clear' is not present in the 'Weather' column, it returns 0. The result is printed to the console.

**Or**

**Python Code Using For Loop:**

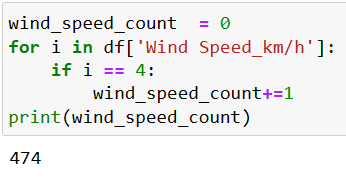


**Code Explanation:**

In this code, we initialize a counter variable clear\_weather\_count to 0. Then, we iterate through each row in the DataFrame using a for loop. For each row, we check if the value in the 'Weather' column is equal to 'Clear'. If it is, we increment the clear\_weather\_count by 1. Finally, we print the count after the loop is done.

**Q. 3) Find the number of times when the 'Wind Speed was exactly 4 km/h'.**

**Python Code Using For Loop:**

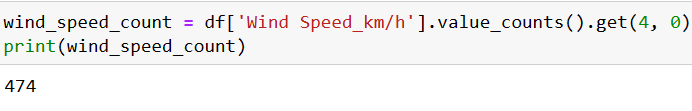


**Code Explanation:**

In this code, we initialize a counter variable wind\_speed\_count to 0. Then, we iterate through each row in the DataFrame using a for loop. For each row, we check if the value in the 'Wind Speed' column is equal to '4 km/h'. If it is, we increment the wind\_speed\_count by 1. Finally, we print the count after the loop is done.

**Or**

**Python Code:**

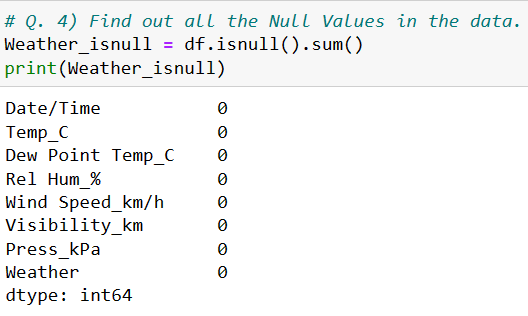


**Code Explanation:**

In this code, we first use the value\_counts() method to count the occurrences of each unique value in the 'Wind Speed' column. Then, we use the get() method to retrieve the count for '4 km/h' specifically. If '4 km/h' is not found in the series (i.e., it doesn't exist in the dataframe), we default to 0. Finally, we print the count

**Q. 4) Find out all the Null Values in the data.**

**Python Code:**

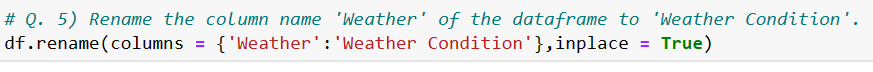


**Code Explanation:**

In this code, weather.isnull() creates a DataFrame of the same shape as your original 'weather' DataFrame, where each cell is True if it's a null value and False otherwise. Then, sum() is used to count the number of True values (which are nulls) along each column. The result is a Series where the column names are the columns in your DataFrame, and the values represent the count of null values in each column.

**Q. 5) Rename the column name 'Weather' of the dataframe to 'Weather Condition'**

**Python Code :**

****

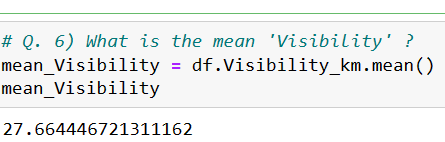
**Code Explanation:**

df.rename(columns={'Weather': 'Weather Condition'}, inplace=True): This renames the 'Weather' column to 'Weather Condition' in the DataFrame df.

The inplace=True argument allows you to modify the DataFrame in place without the need to assign it back to a variable.

**Q. 6) What is the mean 'Visibility' ?**

**Python Code :**

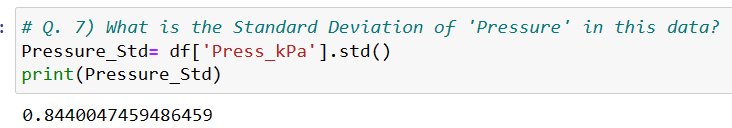
****

**Code Explanation:**

mean\_visibility will now contain the mean visibility value for weather data.

**Q. 7) What is the Standard Deviation of 'Pressure' in this data?**

**Python Code :**

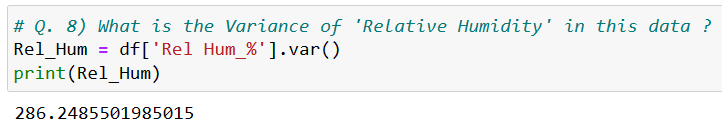
****

**Code Explanation:**

This code will compute the standard deviation of the 'Pressure' column in your weather dataframe and print the result.

**Q. 8) What is the Variance of 'Relative Humidity' in this data ?**

**Python Code :**

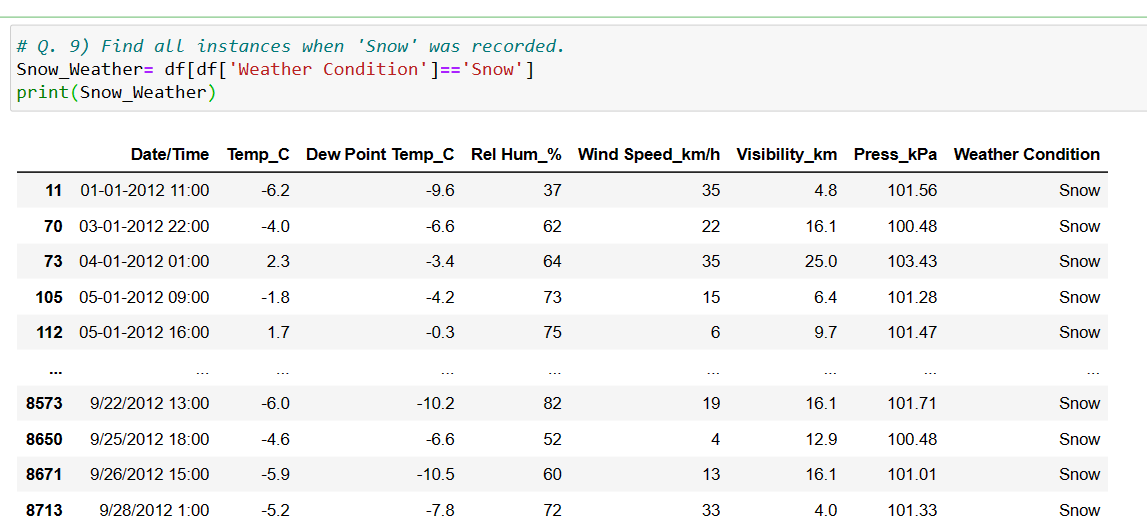
****

**Code Explanation:**

This code will compute and print the variance of the 'Relative Humidity' column in your weather dataframe.

**Q. 9) Find all instances when 'Snow' was recorded.**

**Python Code :**

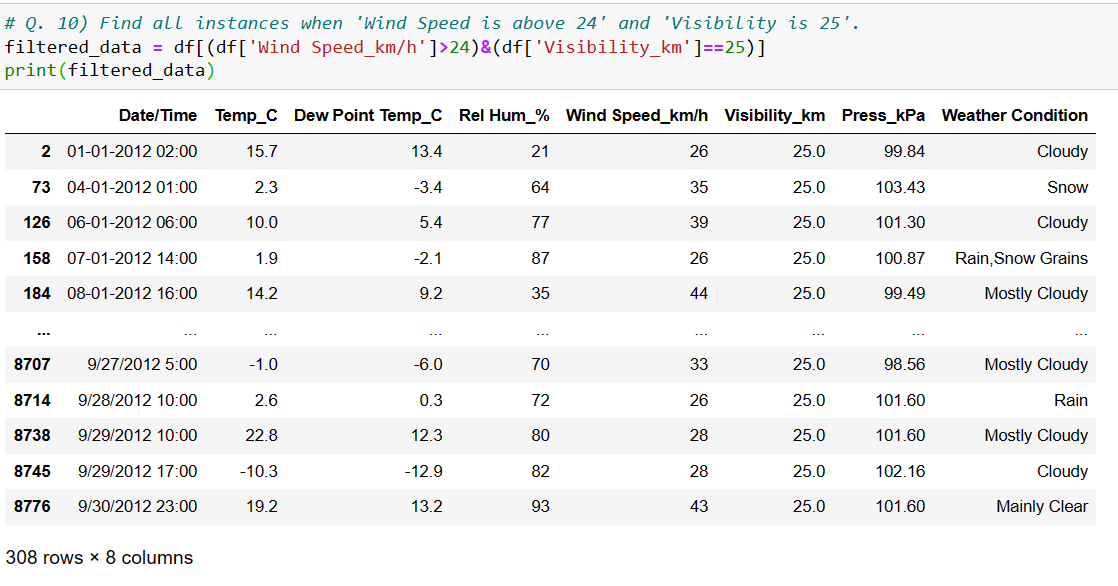
****

**Code Explanation:**

This code will create a new dataframe called snow\_Weather that contains only the rows where 'Snow' is recorded in the 'Weather Condition' column. You can then use this dataframe for further analysis or display the specific instances when 'Snow' occurred in your weather data.

**Q. 10) Find all instances when 'Wind Speed is above 24' and 'Visibility is 25'.**

**Python Code :**

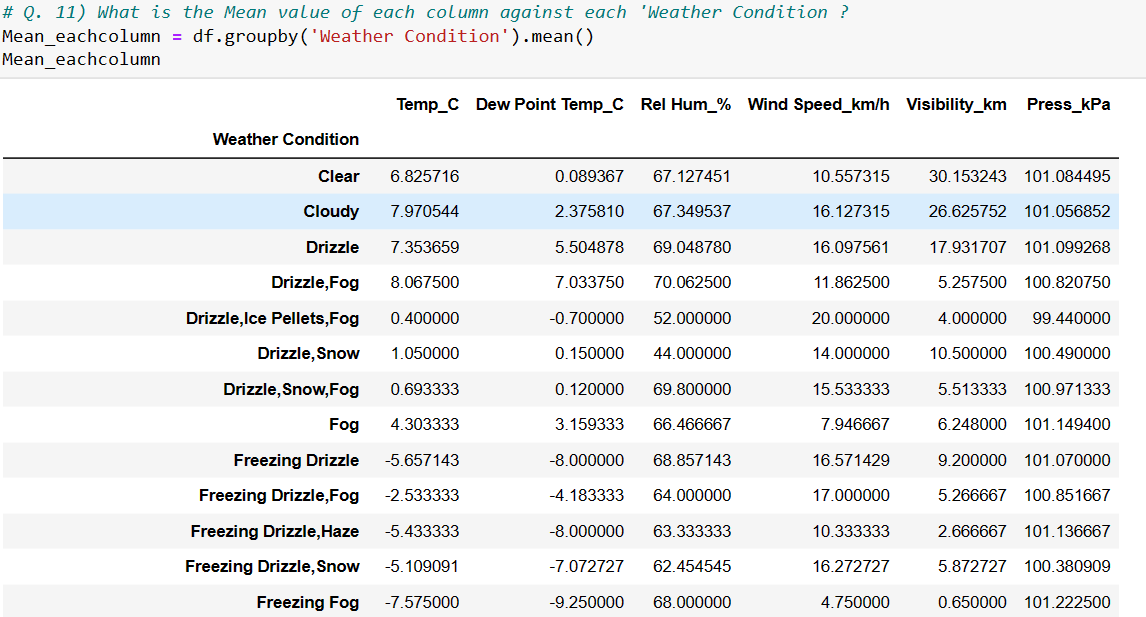
****

**Code Explanation:**

This code will create a new dataframe called filtered\_data that contains only the rows where 'Wind Speed' is above 24 and 'Visibility' is 25.

**Q. 11) What is the Mean value of each column against each 'Weather Condition ?**

**Python Code :**

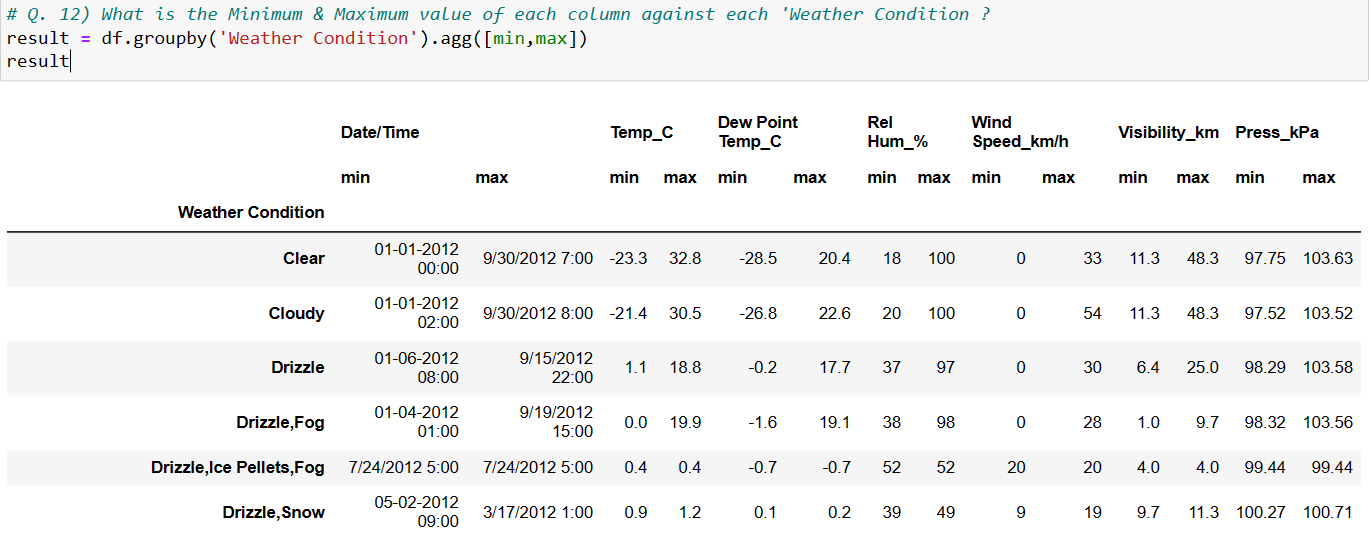
****

**Code Explanation:**

This code will group your DataFrame by the 'Weather Condition' column and then calculate the mean for each numerical column in the grouped data for each unique 'Weather Condition.' The resulting DataFrame, means\_by\_weather\_condition, will contain the mean values for each column against each 'Weather Condition.

**Q. 12) What is the Minimum & Maximum value of each column against each 'Weather Condition ?**

**Python Code :**

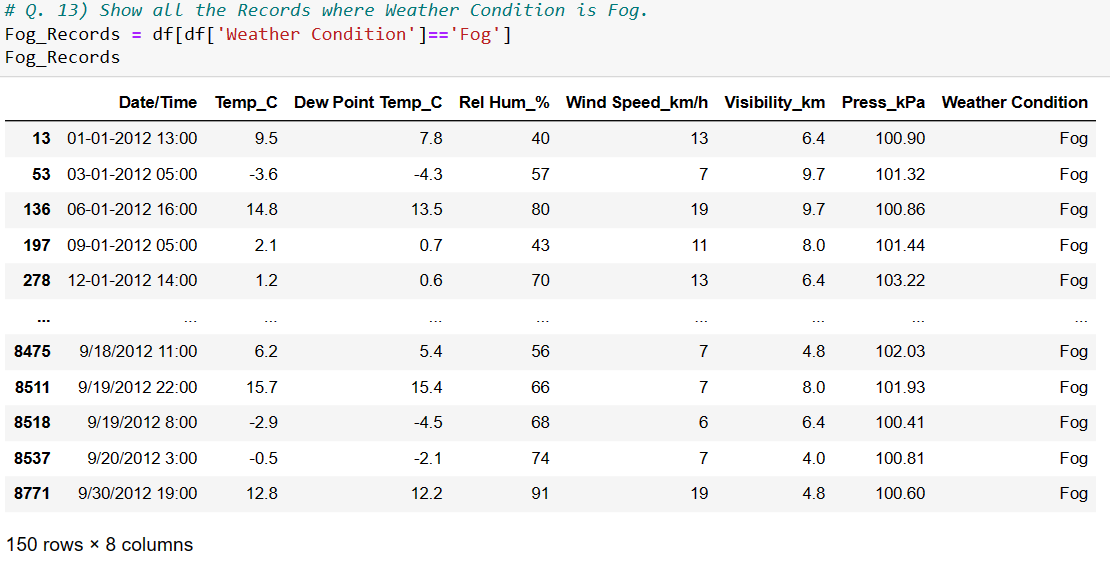
****

**Code Explanation:**

This code will group the DataFrame by 'Weather Condition' and then calculate the minimum and maximum values for each group.The resulting DataFrame result will display the minimum and maximum values for each column for each 'Weather Condition'.

**Q. 13) Show all the Records where Weather Condition is Fog.**

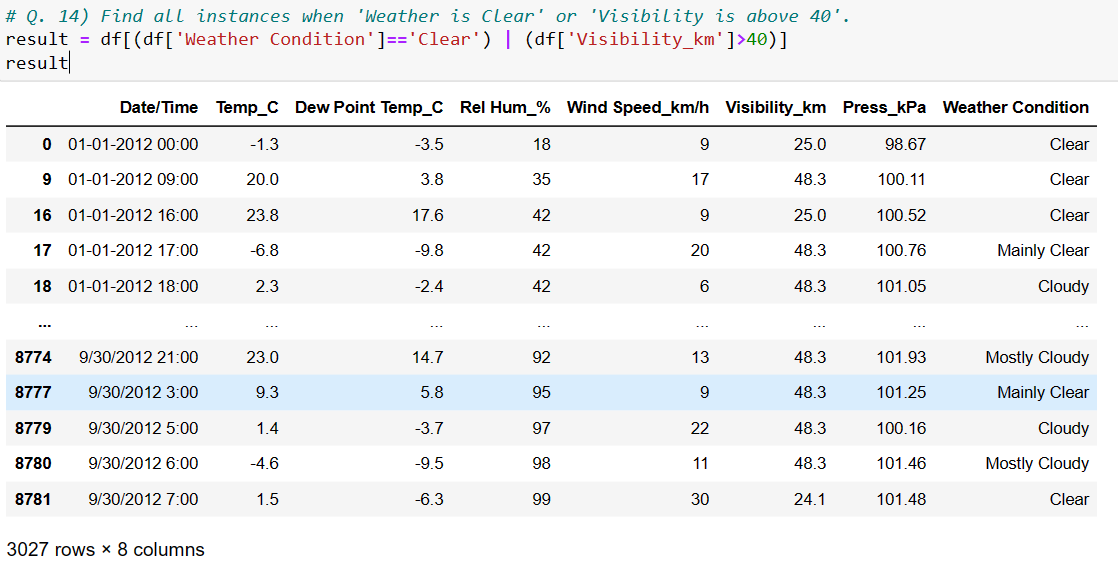
**Python Code :**

****

**Code Explanation:** This code will display all records where the weather condition is "Fog."

**Q. 14) Find all instances when 'Weather is Clear' or 'Visibility is above 40'.**

**Python Code :**

****

**Code Explanation:**

This code will create a new dataframe result data frame that contains only the rows where 'Weather' is 'Clear' or 'Visibility' is above 40.

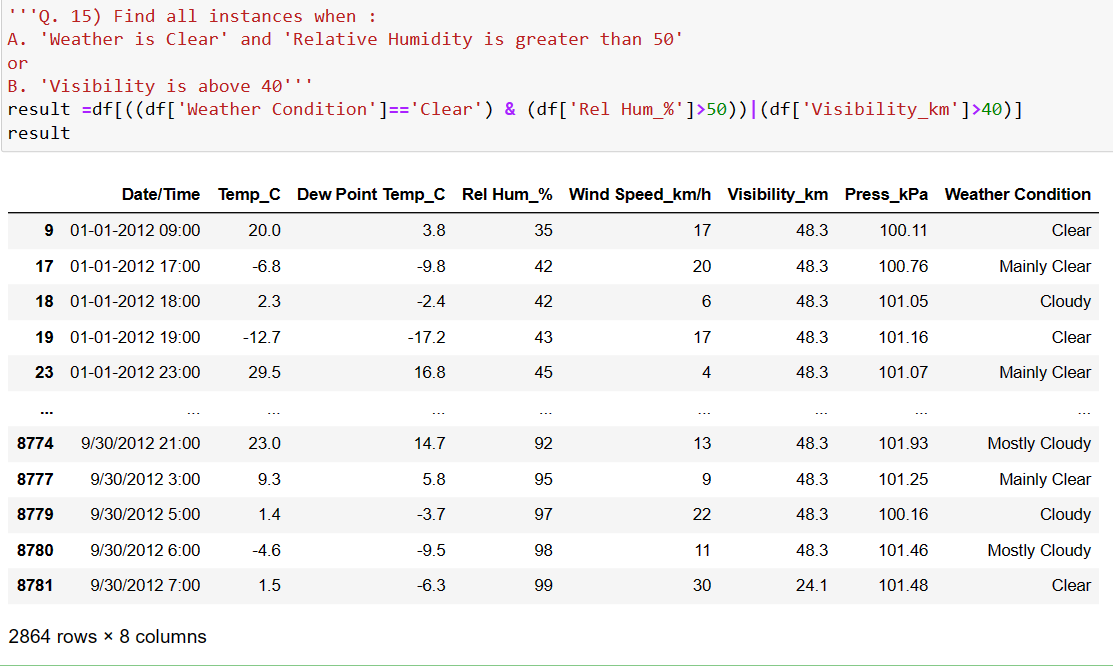
**Q. 15) Find all instances when :**

**A. 'Weather is Clear' and 'Relative Humidity is greater than 50'**

**or**

**B. 'Visibility is above 40'.**

**Python Code :**

****

**Code Explanation:**

The **result** will contain the rows that satisfy either condition A or condition B or both.