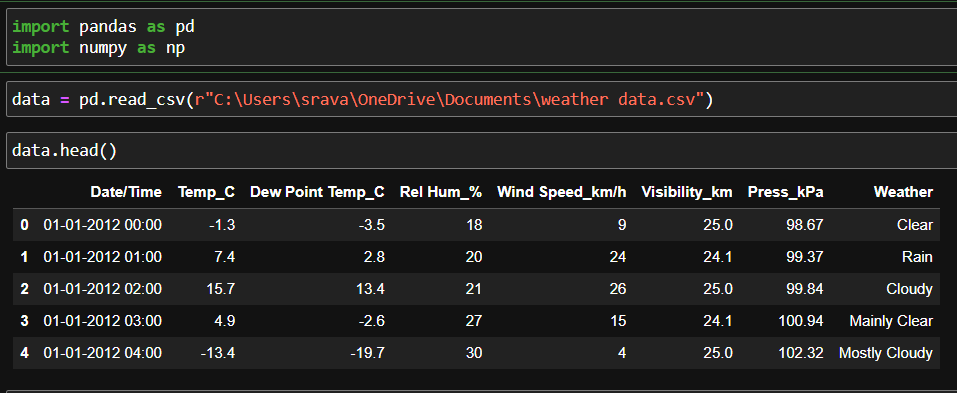
**Report on weather**

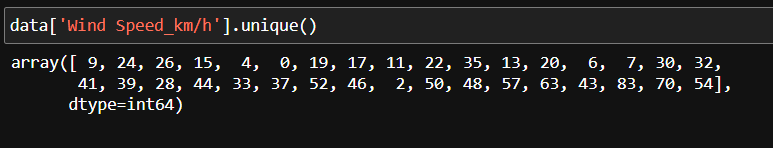
The data that we have collected should be understandable by the python to do analysis so that we have to load the data in the python IDE using the syntax,  
data=pd.read\_csv(r”copy path”)

# ‘data’ is the variable name that we are given to the data we collected.



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

* To know the unique wind speed values in the data we have to use the code : data[“Wind speed\_km/h”].unique()
* After running the above code it gives all the unique wind speed values in the wind speed column.

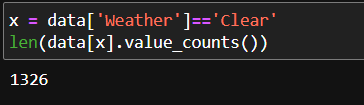


* The unique function gives the unique values in the data. So, the unique wind speed values are: 9, 24, 26, 15, 4, 0, 19, 17, 11, 22, 35, 13, 20, 6, 7, 30, 32,41, 39, 28, 44, 33, 37, 52, 46, 2, 50, 48, 57, 63, 43, 83, 70, 54.

Q. 2) Find the number of times when the

'Weather is exactly Clear'.

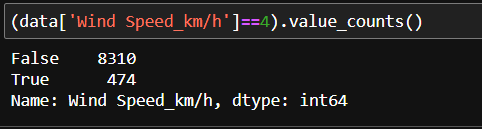
* To find the number of times when the data is clear , we have to use the function called ‘loc’.
* It gives us the locations where the data is clear upon that we have to give length function, so that we’ll get the number of times the weather is exactly clear.
* It gives us the locations where the data is clear upon that we have to give length function, so that we’ll get the number of times the weather is exactly clear.



* The number of times the weather is clear is 1326.

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

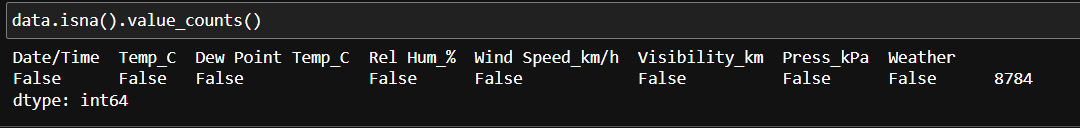
* To know the number of times that wind speed was exactly 4km/h is we have to access the wind speed column by applying assignment operator.
* (data[“Wind Speed\_km/h”]==4).value\_counts()
* The above line will give the True and False counts where the True are the number of times the wind speed is exactly 4.



* Here we got the total number of times that wind speed is exactly 4 is 474.

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

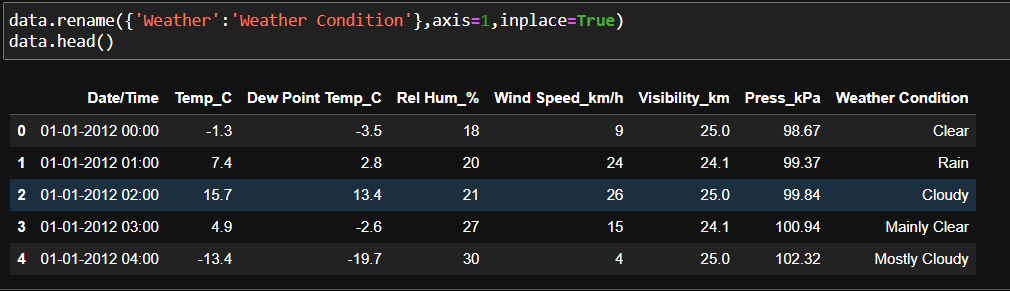
* To find the null values in the data we have to check the data using the code : data.isna()
* This will provides us a Boolean masking type table where in the table gives false if we do not have any null value and gives True if it consist of null values.



* By looking at the output we understand that the data do not have any null values because the value counts given false . so that we can understand that there are no null values in the data.

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

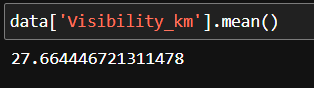
* To change the column name in a dataframe we have to use: datac.rename({"Weather":"WeatherCondition"},axis=1,inplace=True)
* By the above code we can change the name of any column in the data



* We can see that the Weather has changed to weather condition .

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

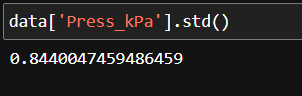
* To find the mean for the column “visibility” we simply use the code : data[‘Visibility\_km’].mean()
* It will gives the average (or) mean value of the required column as as central information of that column.



* For the data the variable (or) column name “Visibility\_km” the mean is 27.6644

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

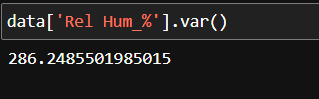
* We have to know the Standard deviation of ‘Pressure’ in the data.
* So, Now to find the standard deviation of the data we have a function called ‘.std()’.
* We have to use that function on the column.
* data[‘Press\_kPa’].std()
* It will provide the information of standard deviation for the required column as



* From the data the standard deviation for the column name “Press\_kPa” is 0.84400

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

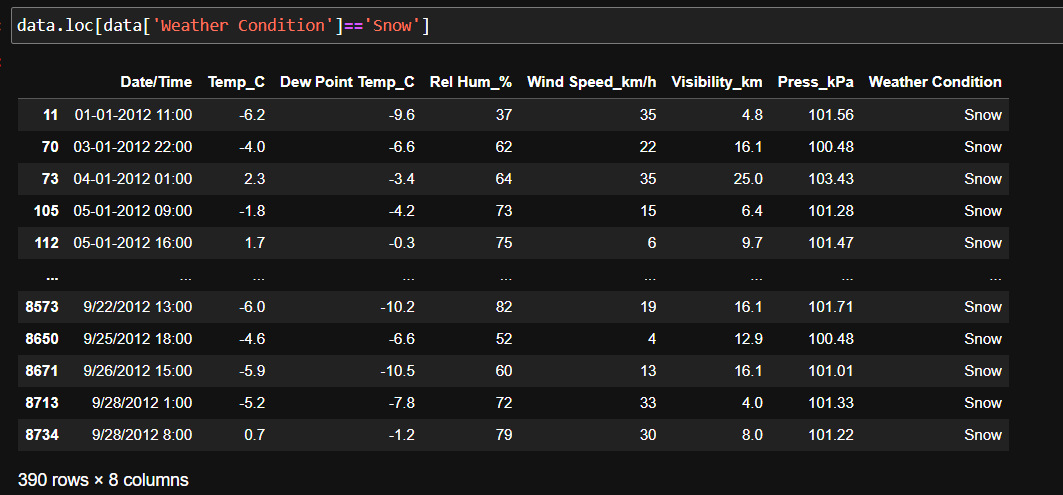
* To find the variance for the data in the column “Relative Humidity” is simply by using the function ‘.var()’
* It will gives the variance for the column name “'Relative Humidity” for the data



* The variance for the column name “Relative Humidity” is 286.24855

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

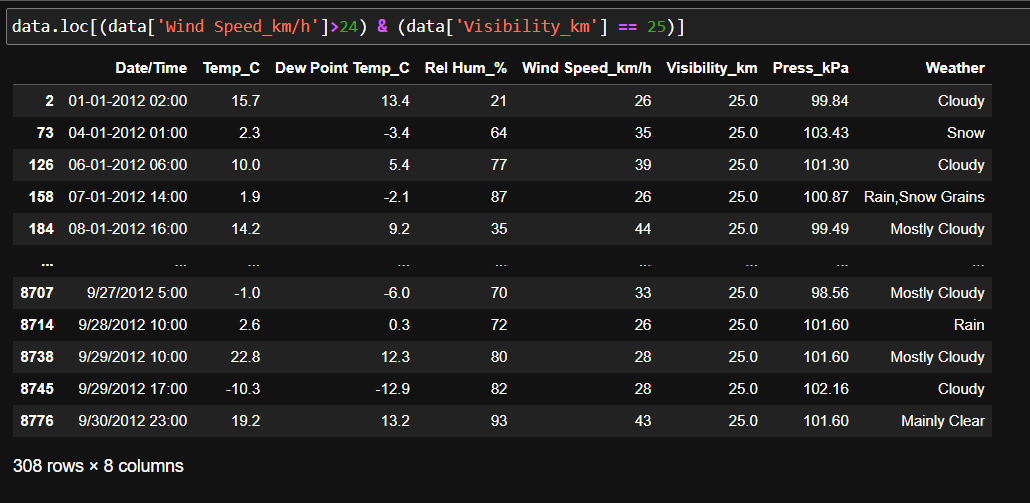
* To know the data whose weather condition is “snow” is by using the code  
  data.loc[data["Weather Condition"]=="Snow"]
* Here the data gives the dataframe table where the data consist of weather condition as snow



* The above output shows the information about the dataframe of the data whose weather condition is snow.
* There are 390 data points whose weather condition is snow.

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

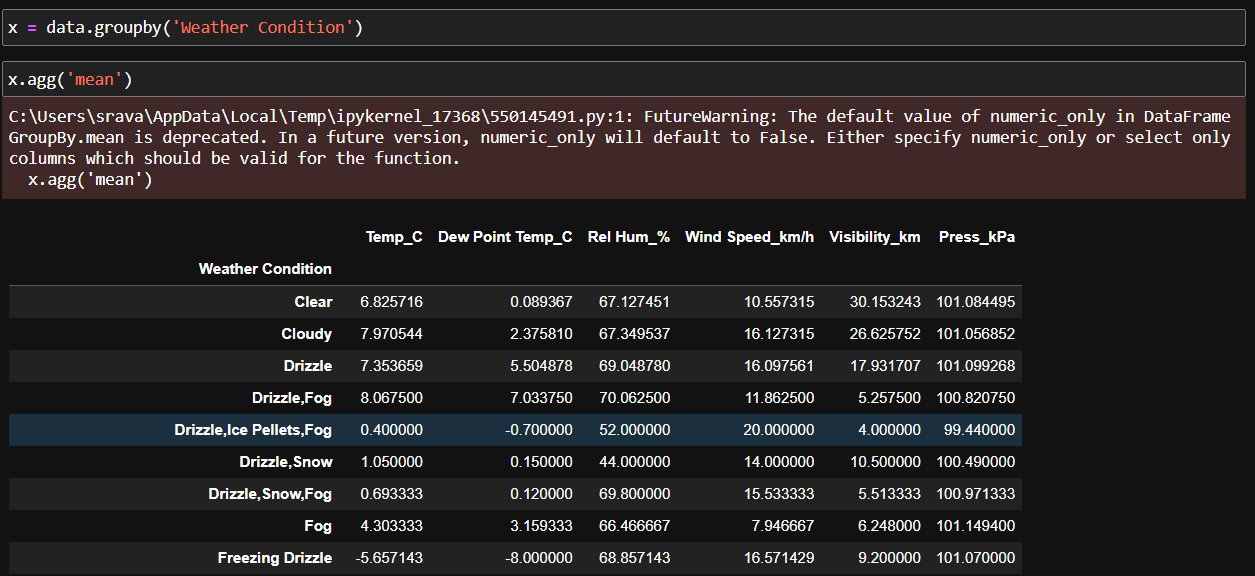
* We have to know the data whose wind speed is greater than 24 and visibility is equal to 25 we use the code as   
  data.loc[(data["Wind Speed\_km/h"]>24) & (data["Visibility\_km"]==25)]
* After performing above code it gives the data of column name “wind speed” greater than 24 and “visibility” is equal to 25.

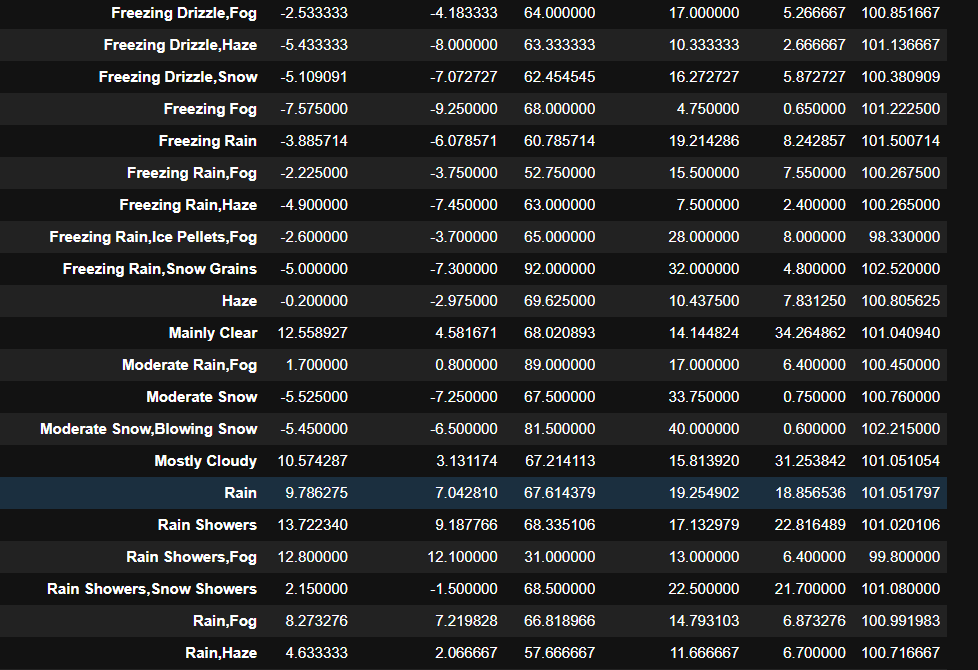


* There are 308 data points which satisfies both the conditions.

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

* The mean for each column on the specified weather condition is, to find that we have to group all the weather conditions by using the function ‘groupby()’.
* The code is:
* x = data.groupby(“Weather Condition”)
* x.agg(“mean”)
* Where the groupby will specify the data into different categories in the particular column and also we can know different values like ‘mean’,’max’,’min’,etc., by using aggregate function with it.



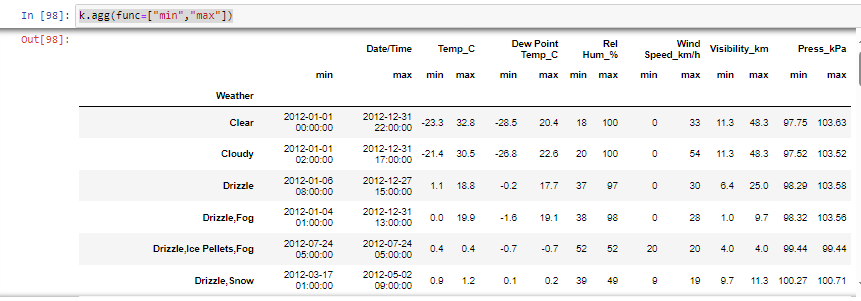




* The above data is the mean of all the data with the specified weather condition.

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

* To know the minimum and maximum data for all the columns on the specified column we use the code as x.agg(func=["min","max"])
* Where the x means the data which is already grouped by the condition for the particular required column name weather condition
* Also we use the aggregate function for the data wher we can use any type of function inside it



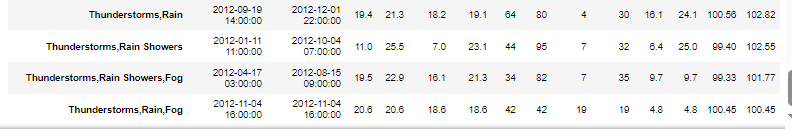








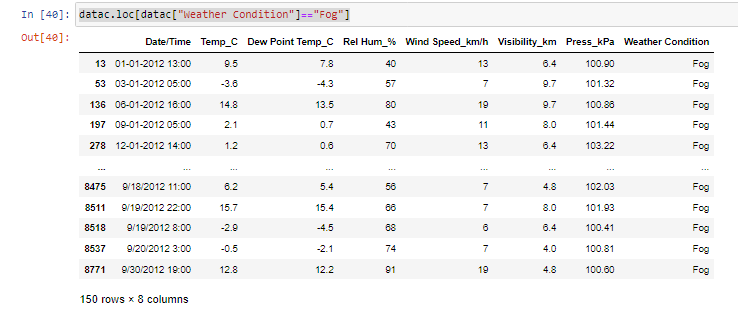




* Here above is the data for the different weather conditions minimum and maximum values of the other parameters (or) columns in the data likely the data is in the above is given the maximum and minimum values of all the data on the specified weather condition

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

* To know the data records whose weather condition is Fog we use the code as datac.loc[datac["Weather Condition"]=="Fog"]
* It gives the entire data for the weather condition is equal to fog



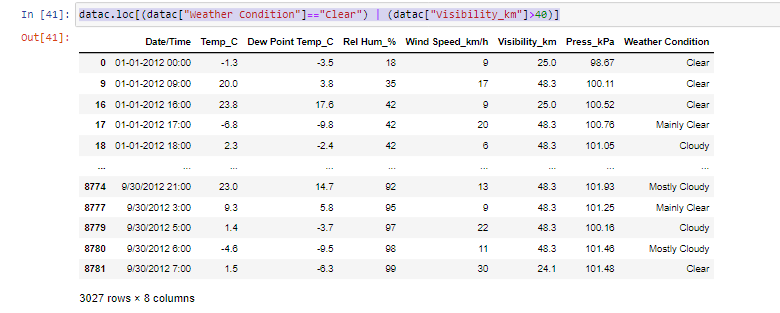
* Here is the data whose weather condition is fog
* We have 150 different data points in the data whose weather condition is fog

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

* To get the data for the conditions either weather condition is clear or visibility is 40
* We use the code as

datac.loc[(datac["Weather Condition"]=="Clear") | (datac["Visibility\_km"]>40)]

* We get the data which satisfies the condition that weather condition is clear or visibility is greater than 40



* In the condition we should use the bitwise or (|)operator to satisfy the condition
* And the above data is the which satisfied the conditions
* We have 3027 number of data points in the data which satisfying the condition for weather condition is clear or visibility is greater than 40

Q. 15) Find all instances when :

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

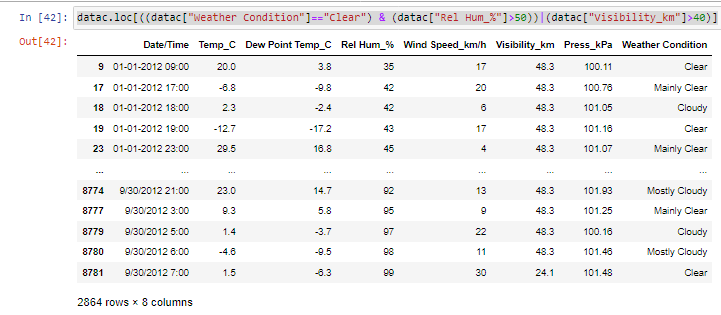
or

B. 'Visibility is above 40'

* To get the data for the conditions either weather condition is clear and Relative humidity greater than 50 or visibility is greater than 40
* We use the code as

datac.loc[((datac["Weather Condition"]=="Clear") & (datac["Rel Hum\_%"]>50))|(datac["Visibility\_km"]>40)]

* This condition satisfies the requirement for the weather condition is clear and relative humidity is greater than 50 or else the visibility is greater than 40



* In the condition we should use the bitwise and (&) and bitwise or (|)operators to satisfy the conditions
* And the above data is the which satisfied the conditions
* We have the total count of 2864 datapoints from the data which satisfies the condition for the weather condition should be clear and relative humidity greater than 50 orelse the visibility is greater than 40