weather

introduction

- How to import pandas and numpy

```
import pandas as pd
import numpy as np
```

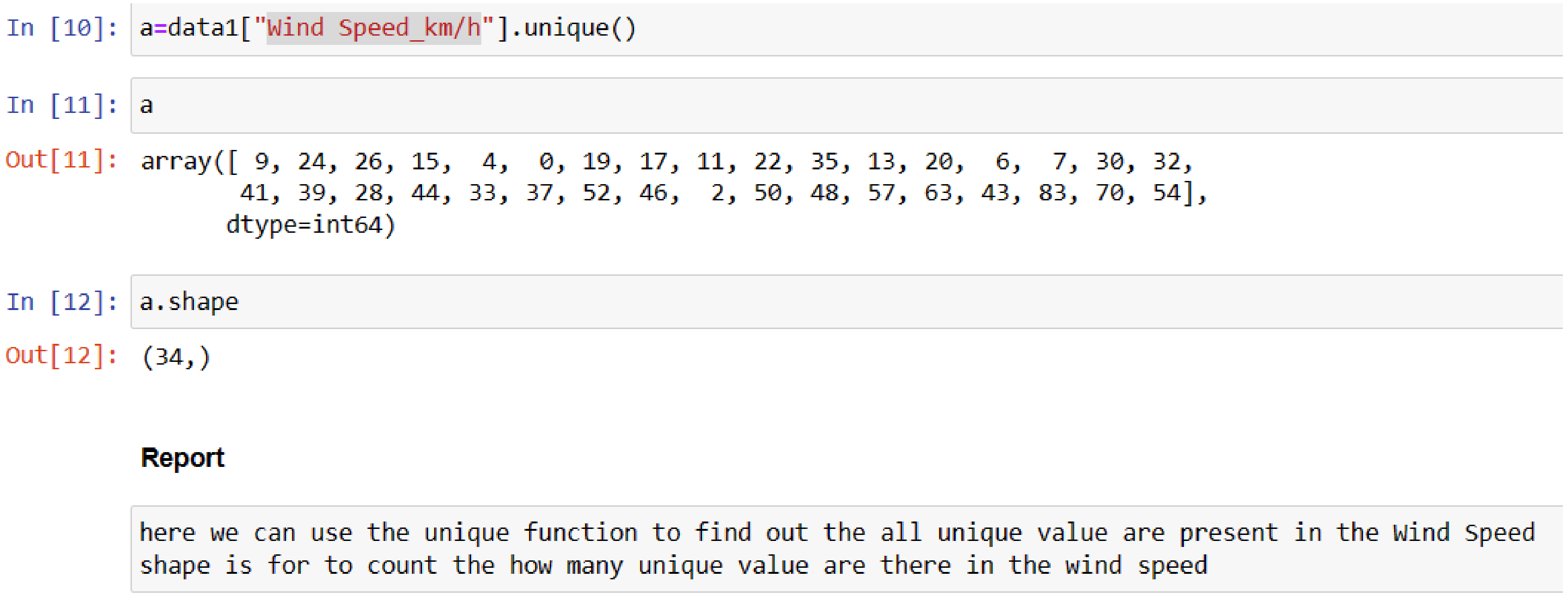
How to import data set into the pandas

```
data=pd.read_csv(r"C:\Users\Intel\Downloads\Weather_Data.csv")
```

To copy the data set we use the copy function for that we are not distribe the original data

```
data1=data.copy()
```

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



```
In [10]: a=data1["Wind Speed_km/h"].unique()
```

```
In [11]: a
```

```
Out[11]: array([ 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],
              dtype=int64)
```

```
In [12]: a.shape
```

```
Out[12]: (34,)
```

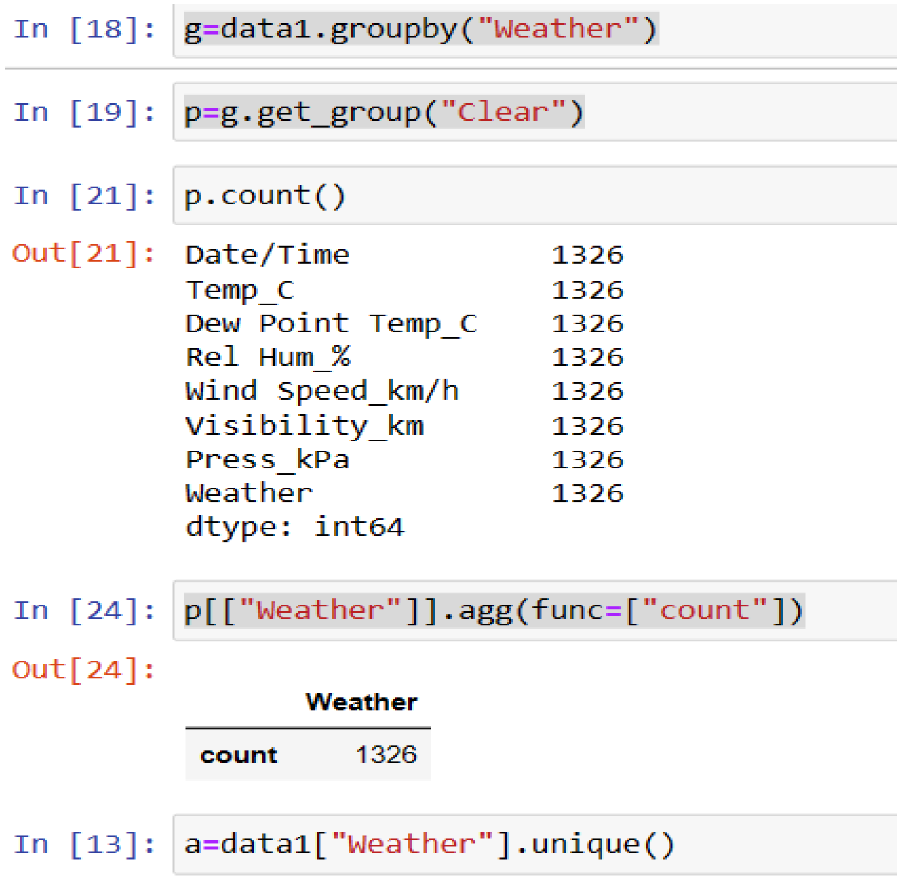
Report

```
here we can use the unique function to find out the all unique value are present in the Wind Speed
shape is for to count the how many unique value are there in the wind speed
```

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

- **Report**

In the data set we have to collect the single column by using groupby function in that column to get the single element we use again group by function than that we have to use count function



```
In [18]: g=data1.groupby("Weather")
```

```
In [19]: p=g.get_group("Clear")
```

```
In [21]: p.count()

Out[21]: Date/Time             1326
         Temp_C                1326
         Dew Point Temp_C      1326
         Rel Hum_%             1326
         Wind Speed_km/h       1326
         Visibility_km         1326
         Press_kPa             1326
         Weather               1326
         dtype: int64
```

```
In [24]: p[["Weather"]].agg(func=["count"])

Out[24]:
```

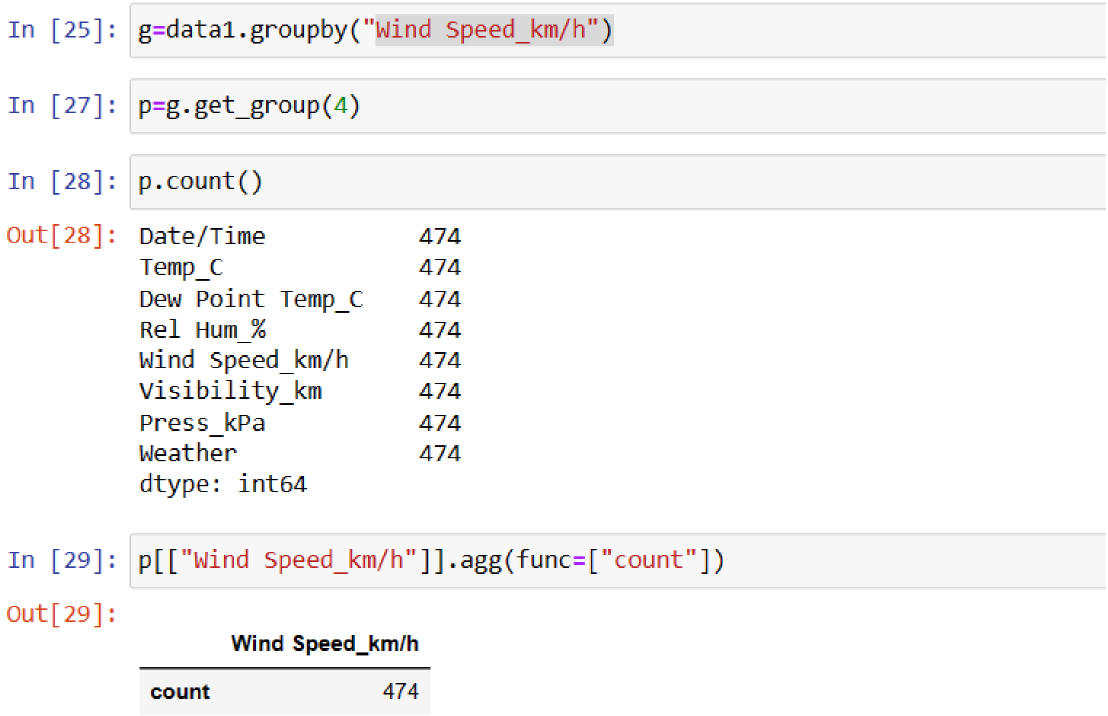
	Weather
count	1326

```
In [13]: a=data1["Weather"].unique()
```

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

- **Report**

In the data set we have to collect the single column by using groupby function in that column to get the single element we use again group by function than that we have to use count function



```
In [25]: g=data1.groupby("Wind Speed_km/h")

In [27]: p=g.get_group(4)

In [28]: p.count()

Out[28]: Date/Time           474
         Temp_C              474
         Dew Point Temp_C    474
         Rel Hum_%           474
         Wind Speed_km/h     474
         Visibility_km       474
         Press_kPa           474
         Weather             474
         dtype: int64

In [29]: p[["Wind Speed_km/h"]].agg(func=["count"])

Out[29]:
```

	Wind Speed_km/h
count	474

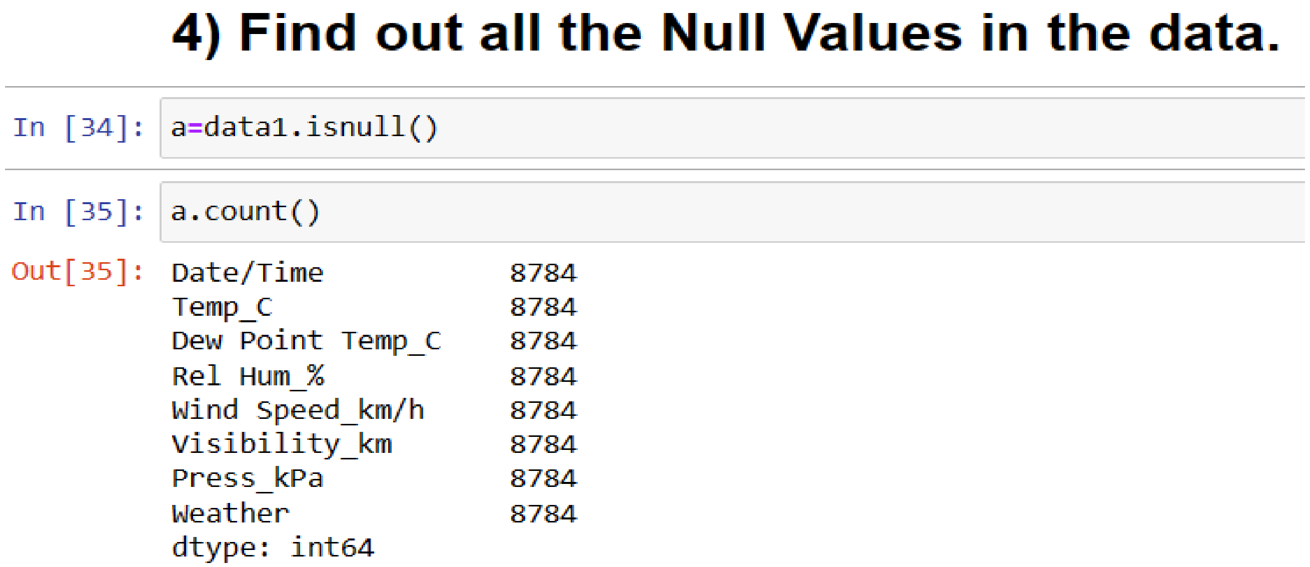
4) Find out all the Null Values in the data.

- Report

We can find all the null

values in the dataset

we use the isnull function



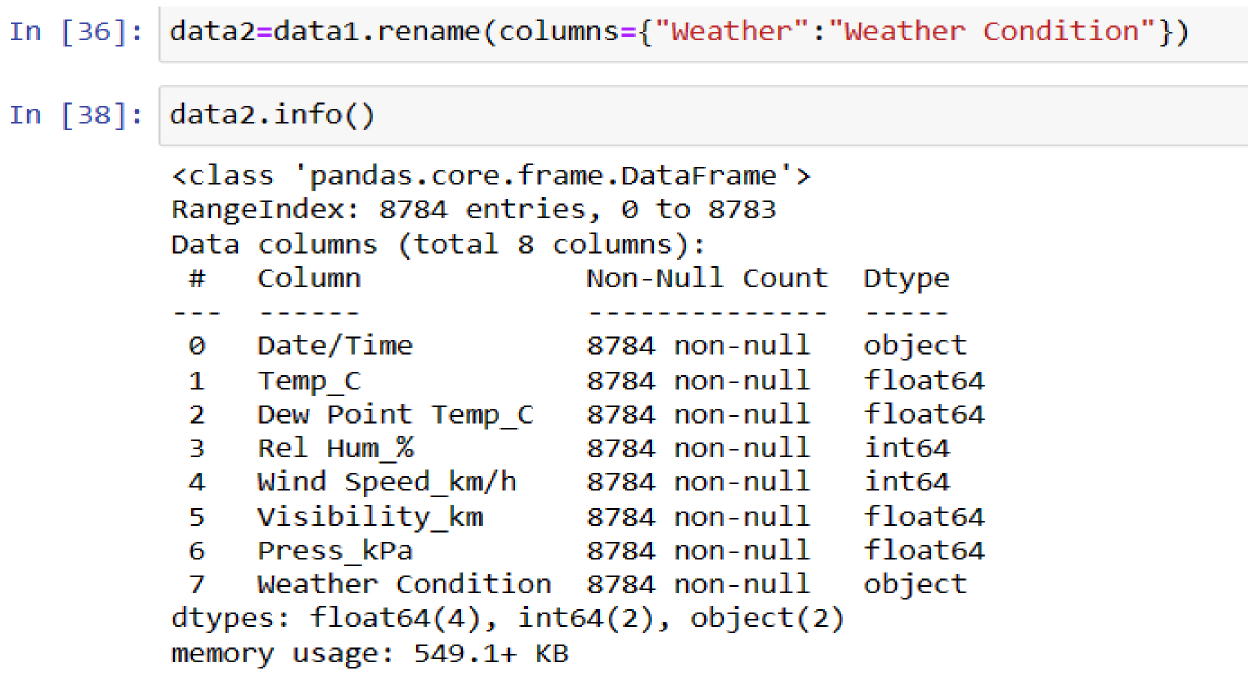
4) Find out all the Null Values in the data.

```
In [34]: a=data1.isnull()

In [35]: a.count()

Out[35]: Date/Time            8784
         Temp_C               8784
         Dew Point Temp_C     8784
         Rel Hum_%            8784
         Wind Speed_km/h      8784
         Visibility_km        8784
         Press_kPa            8784
         Weather              8784
         dtype: int64
```

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

- **Report**

In the dataset to rename the any columns we use the rename function



```
In [36]: data2=data1.rename(columns={"Weather":"Weather Condition"})
```

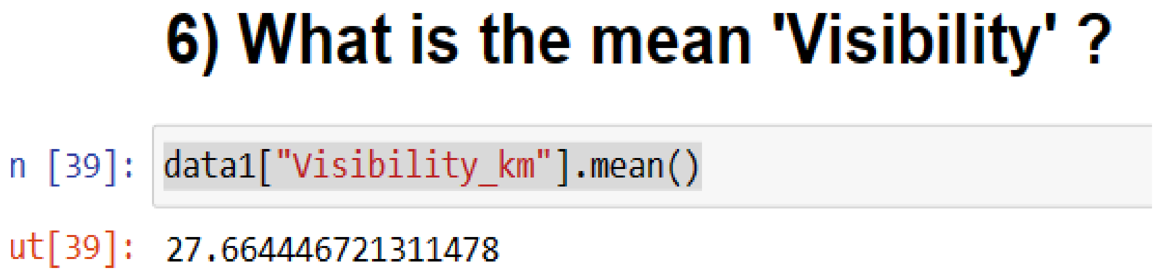
```
In [38]: data2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8784 entries, 0 to 8783
Data columns (total 8 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   Date/Time          8784 non-null   object 
 1   Temp_C             8784 non-null   float64
 2   Dew Point Temp_C   8784 non-null   float64
 3   Rel Hum_%          8784 non-null   int64  
 4   Wind Speed_km/h    8784 non-null   int64  
 5   Visibility_km      8784 non-null   float64
 6   Press_kPa          8784 non-null   float64
 7   Weather Condition  8784 non-null   object 
dtypes: float64(4), int64(2), object(2)
memory usage: 549.1+ KB
```

6) What is the mean 'Visibility' ?

- **Report**

To find the average of an visiblility column in the dataset



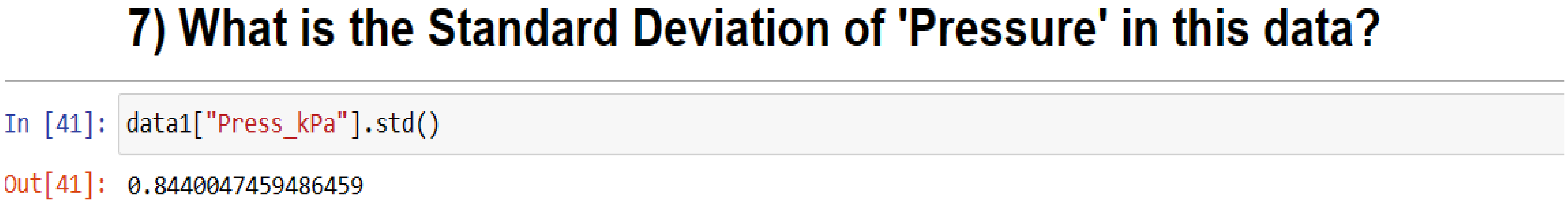
6) What is the mean 'Visibility' ?

```
n [39]: data1["Visibility_km"].mean()

ut[39]: 27.664446721311478
```

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

Report

to find the standard deviation of an pressure column in the datase



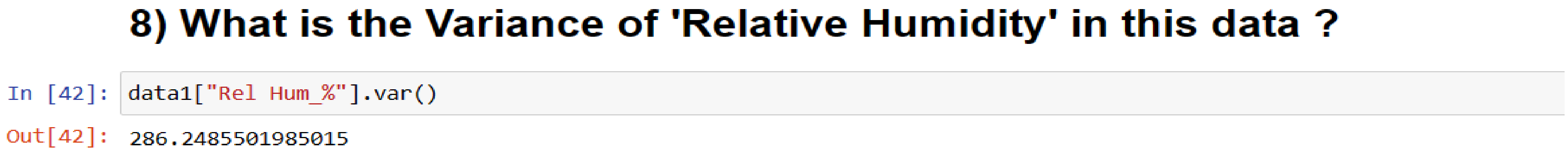
7) What is the Standard Deviation of 'Pressure' in this data?

```
In [41]: data1["Press_kPa"].std()

Out[41]: 0.8440047459486459
```

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

- **Report**

to find the variance of an relative humidity column in the dataset



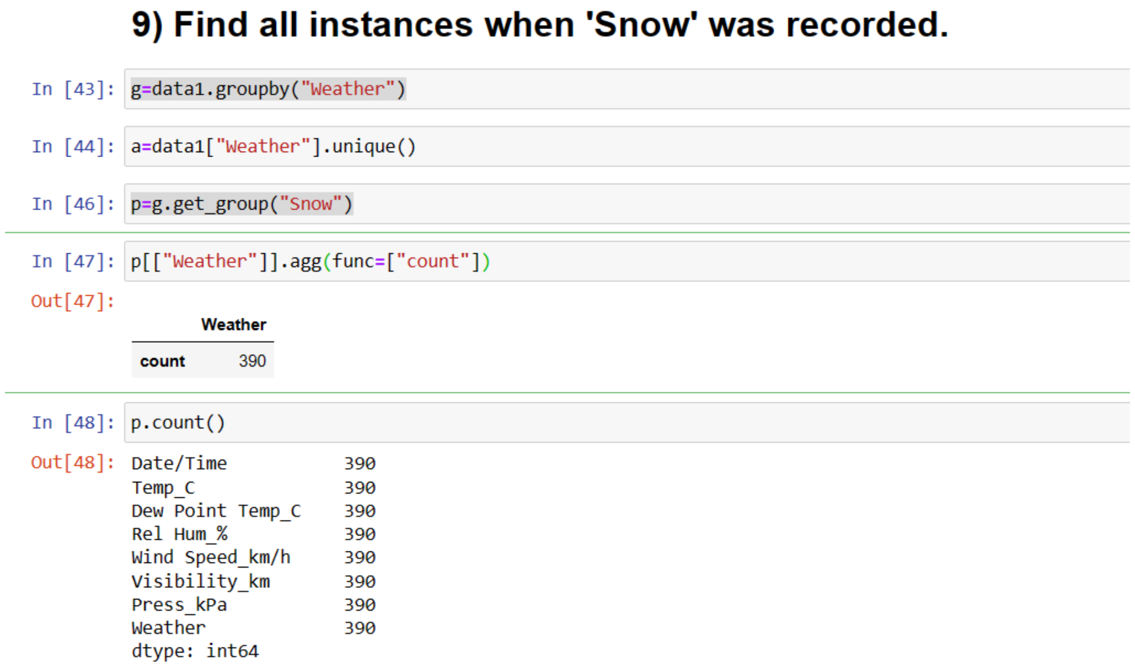
8) What is the Variance of 'Relative Humidity' in this data ?

```
In [42]: data1["Rel Hum_%"].var()

Out[42]: 286.2485501985015
```

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

- **Report**

In the data set we have to collect the single column by using groupby function in that column to get the single element we use again group by function than that we have to use count function



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

```
In [43]: g=data1.groupby("Weather")

In [44]: a=data1["Weather"].unique()

In [46]: p=g.get_group("Snow")

In [47]: p[["Weather"]].agg(func=["count"])
Out[47]:
```

	Weather
count	390

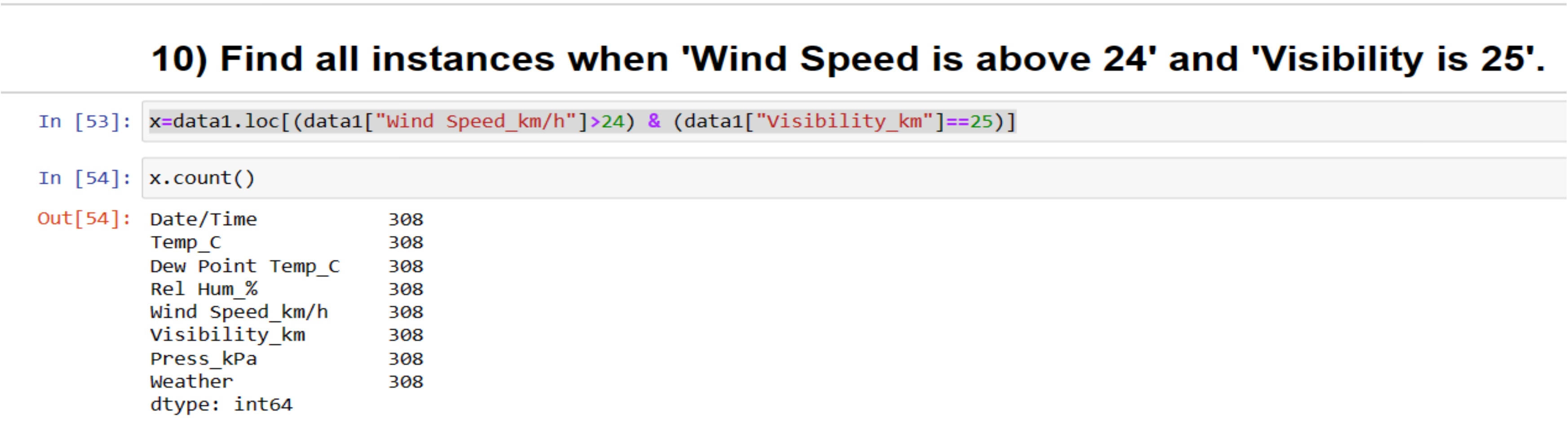
```
In [48]: p.count()

Out[48]: Date/Time           390
         Temp_C              390
         Dew Point Temp_C    390
         Rel Hum_%           390
         Wind Speed_km/h     390
         Visibility_km       390
         Press_kPa           390
         Weather             390
         dtype: int64
```

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

- **Report**

In the dataset to find the all instances in the wind speed column is

Above 24 and visibility column is equal 25



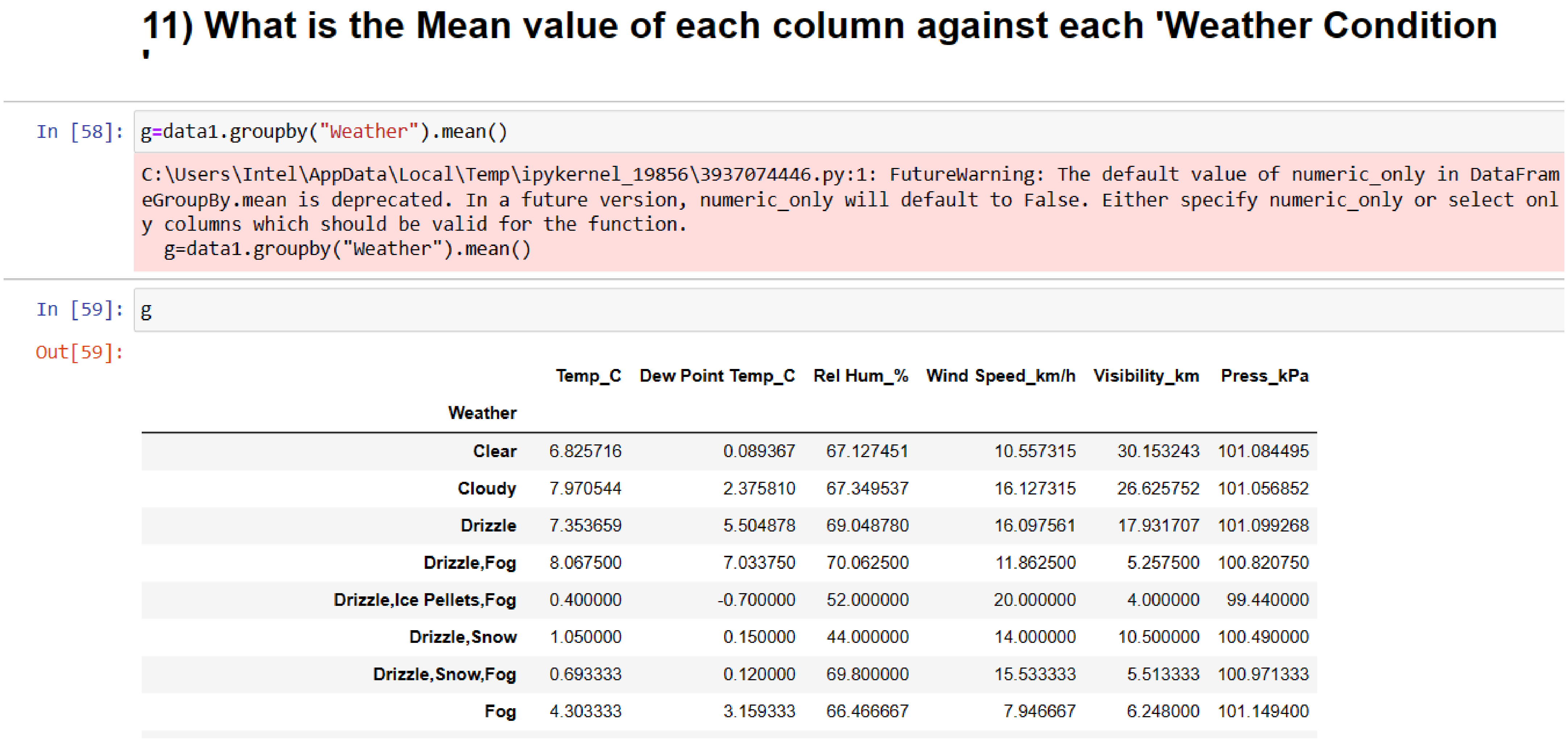
10) Find all instances when 'Wind Speed is above 24' and 'Visibility is 25'.

```
In [53]: x=data1.loc[(data1["Wind Speed_km/h"]>24) & (data1["Visibility_km"]==25)]

In [54]: x.count()

Out[54]: Date/Time           308
         Temp_C              308
         Dew Point Temp_C    308
         Rel Hum_%           308
         Wind Speed_km/h     308
         Visibility_km       308
         Press_kPa           308
         Weather             308
         dtype: int64
```

11) What is the Mean value of each column against each 'Weather Condition '

Report

To find the all mean values of the columns in the dataset except the Weather column



11) What is the Mean value of each column against each 'Weather Condition '

```
In [58]: g=data1.groupby("Weather").mean()
```

```
C:\Users\Intel\AppData\Local\Temp\ipykernel_19856\3937074446.py:1: FutureWarning: The default value of numeric_only in DataFram
eGroupBy.mean is deprecated. In a future version, numeric_only will default to False. Either specify numeric_only or select onl
y columns which should be valid for the function.
  g=data1.groupby("Weather").mean()
```

```
In [59]: g
```

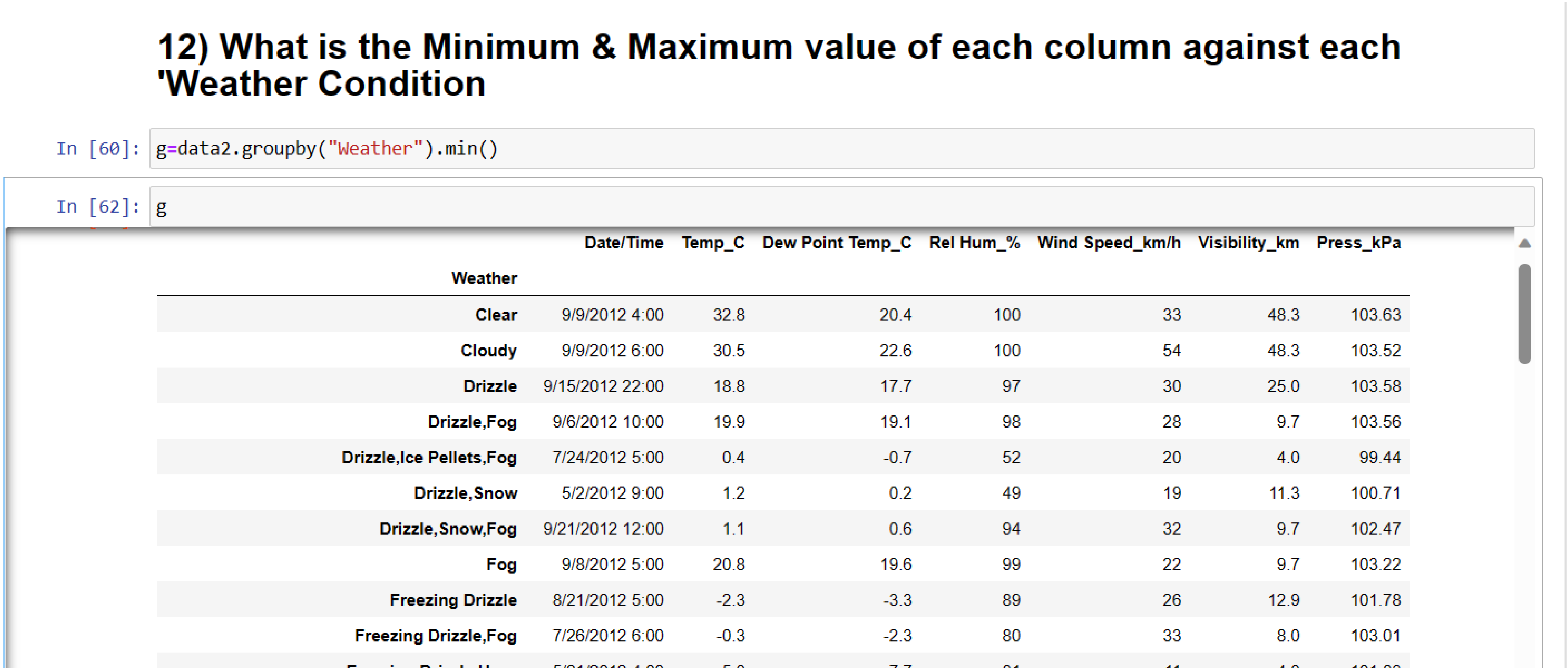
Out[59]:

	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa
Weather						
Clear	6.825716	0.089367	67.127451	10.557315	30.153243	101.084495
Cloudy	7.970544	2.375810	67.349537	16.127315	26.625752	101.056852
Drizzle	7.353659	5.504878	69.048780	16.097561	17.931707	101.099268
Drizzle,Fog	8.067500	7.033750	70.062500	11.862500	5.257500	100.820750
Drizzle,Ice Pellets,Fog	0.400000	-0.700000	52.000000	20.000000	4.000000	99.440000
Drizzle,Snow	1.050000	0.150000	44.000000	14.000000	10.500000	100.490000
Drizzle,Snow,Fog	0.693333	0.120000	69.800000	15.533333	5.513333	100.971333
Fog	4.303333	3.159333	66.466667	7.946667	6.248000	101.149400

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

- **Report**

To find the all maximum and minimum values of the columns in the dataset except the Weather column

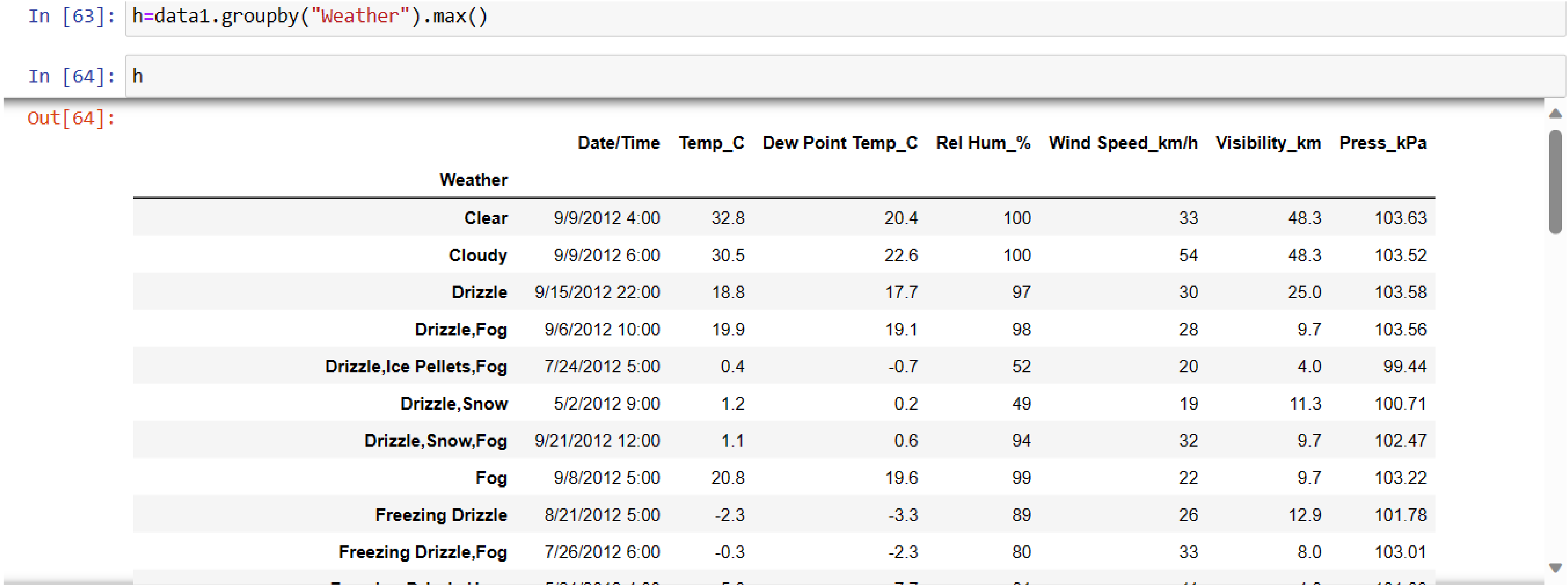


12) What is the Minimum & Maximum value of each column against each 'Weather Condition

```
In [60]: g=data2.groupby("Weather").min()
```

```
In [62]: g
```

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa
Weather							
Clear	9/9/2012 4:00	32.8	20.4	100	33	48.3	103.63
Cloudy	9/9/2012 6:00	30.5	22.6	100	54	48.3	103.52
Drizzle	9/15/2012 22:00	18.8	17.7	97	30	25.0	103.58
Drizzle,Fog	9/6/2012 10:00	19.9	19.1	98	28	9.7	103.56
Drizzle,Ice Pellets,Fog	7/24/2012 5:00	0.4	-0.7	52	20	4.0	99.44
Drizzle,Snow	5/2/2012 9:00	1.2	0.2	49	19	11.3	100.71
Drizzle,Snow,Fog	9/21/2012 12:00	1.1	0.6	94	32	9.7	102.47
Fog	9/8/2012 5:00	20.8	19.6	99	22	9.7	103.22
Freezing Drizzle	8/21/2012 5:00	-2.3	-3.3	89	26	12.9	101.78
Freezing Drizzle,Fog	7/26/2012 6:00	-0.3	-2.3	80	33	8.0	103.01



```
In [63]: h=data1.groupby("Weather").max()
```

```
In [64]: h
```

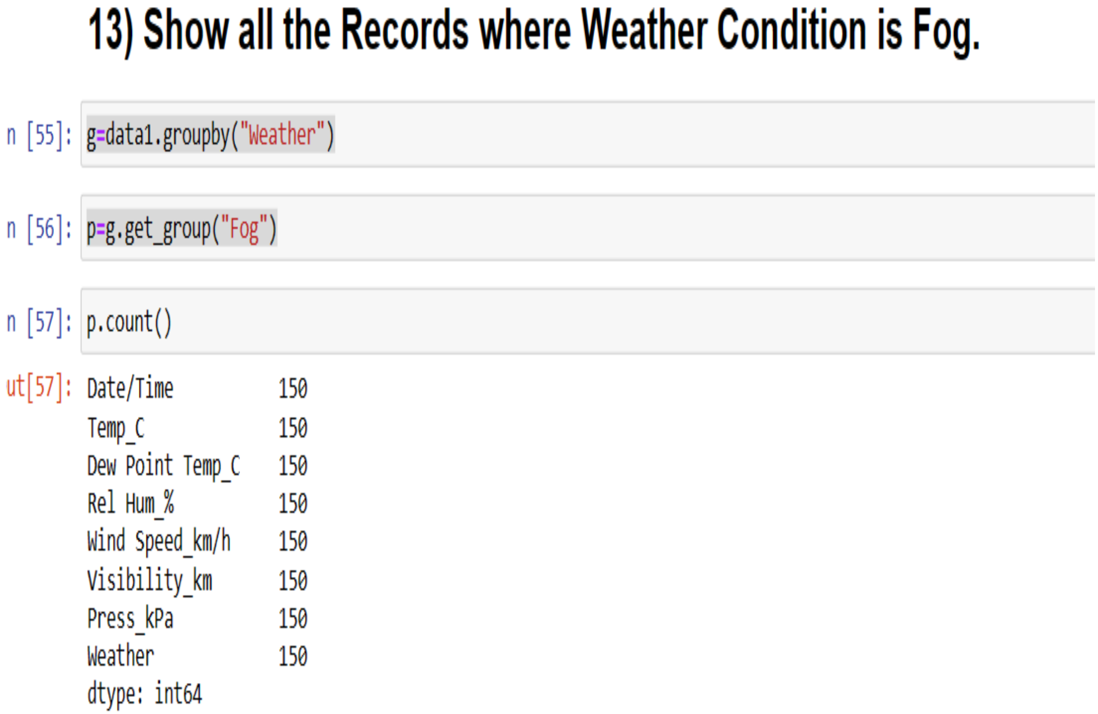
Out[64]:

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa
Weather							
Clear	9/9/2012 4:00	32.8	20.4	100	33	48.3	103.63
Cloudy	9/9/2012 6:00	30.5	22.6	100	54	48.3	103.52
Drizzle	9/15/2012 22:00	18.8	17.7	97	30	25.0	103.58
Drizzle,Fog	9/6/2012 10:00	19.9	19.1	98	28	9.7	103.56
Drizzle,Ice Pellets,Fog	7/24/2012 5:00	0.4	-0.7	52	20	4.0	99.44
Drizzle,Snow	5/2/2012 9:00	1.2	0.2	49	19	11.3	100.71
Drizzle,Snow,Fog	9/21/2012 12:00	1.1	0.6	94	32	9.7	102.47
Fog	9/8/2012 5:00	20.8	19.6	99	22	9.7	103.22
Freezing Drizzle	8/21/2012 5:00	-2.3	-3.3	89	26	12.9	101.78
Freezing Drizzle,Fog	7/26/2012 6:00	-0.3	-2.3	80	33	8.0	103.01

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

- **Report**

In the data set we have to collect the single column by using groupby function in that column to get the single element we use again group by function than that we have to use count function



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

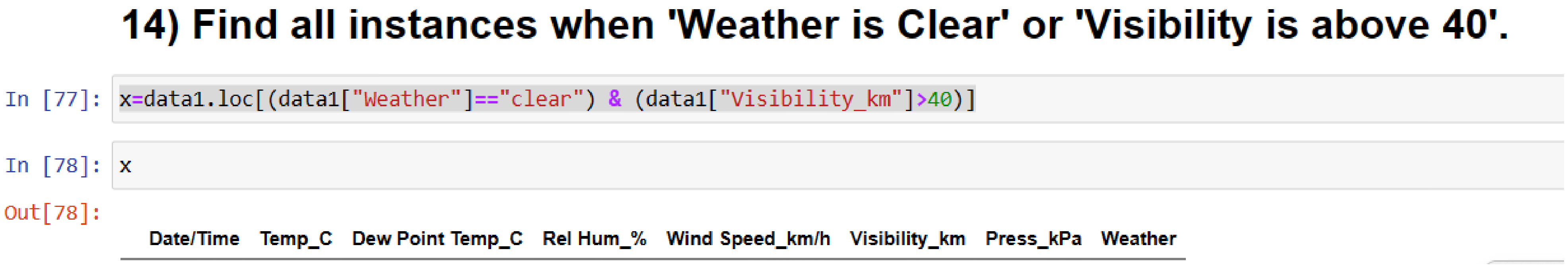
```
n [55]: g=data1.groupby("Weather")

n [56]: p=g.get_group("Fog")

n [57]: p.count()

ut[57]: Date/Time           150
        Temp_C              150
        Dew Point Temp_C    150
        Rel Hum_%           150
        Wind Speed_km/h     150
        Visibility_km       150
        Press_kPa           150
        Weather             150
        dtype: int64
```

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

- **Report**

In the dataset to find the all instances in the weather column is

is equal to clear and visibility column is greather than 25



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

```
In [77]: x=data1.loc[(data1["Weather"]=="clear") & (data1["Visibility_km"]>40)]
```

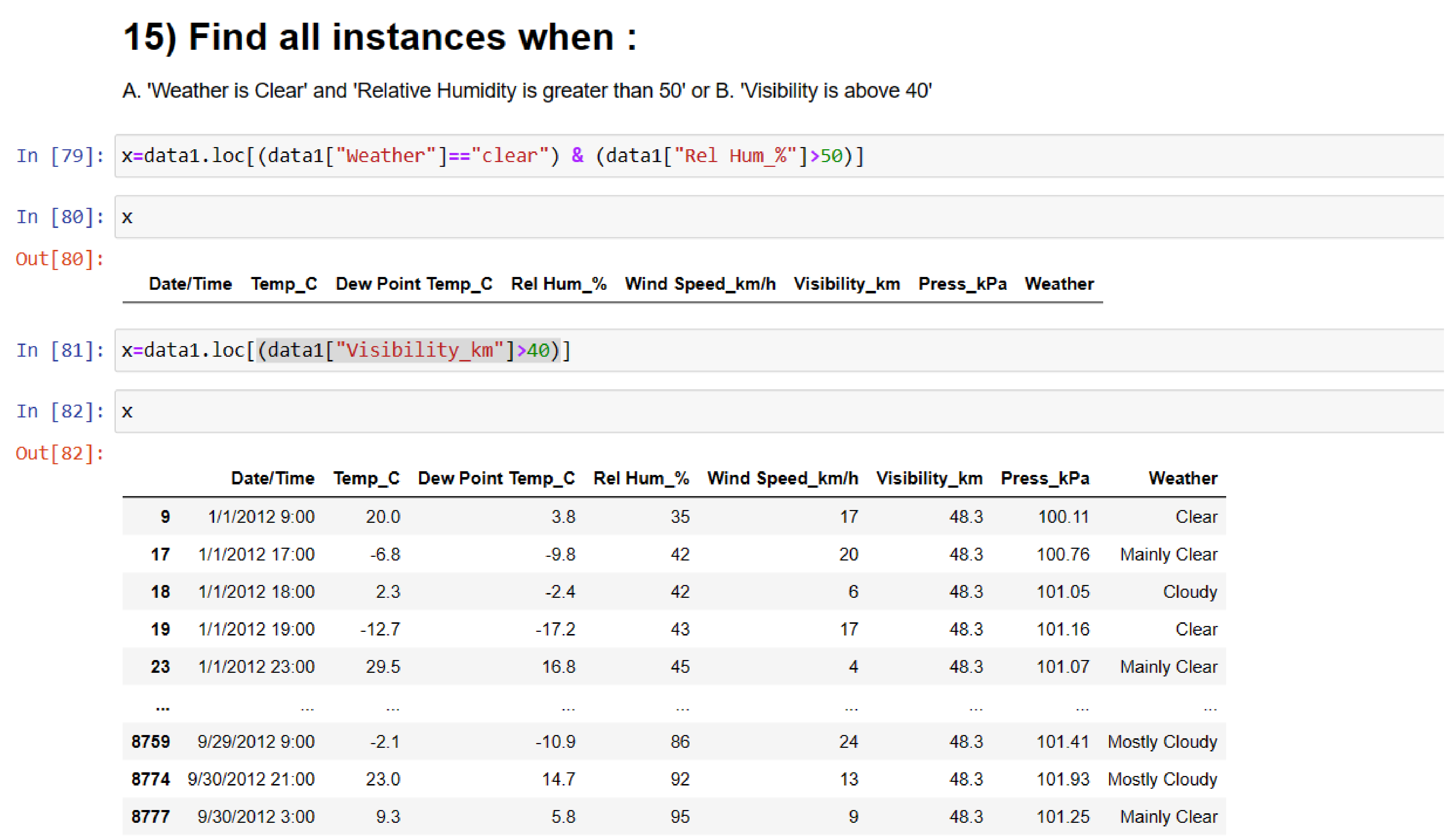
```
In [78]: x
```

```
Out[78]:
   Date/Time   Temp_C   Dew Point Temp_C   Rel Hum_%   Wind Speed_km/h   Visibility_km   Press_kPa   Weather
```

15) Find all instances when :
A. 'Weather is Clear' and 'Relative Humidity is greater than 50' or B. 'Visibility is above 40'

- **Report**

In the dataset to find the all instances in the weather column is equal to clear relative humidity column is above 50 or visibility column is above 25



15) Find all instances when :

A. 'Weather is Clear' and 'Relative Humidity is greater than 50' or B. 'Visibility is above 40'

In [79]:
```
x=data1.loc[(data1["Weather"]=="clear") & (data1["Rel Hum_%"]>50)]
```

In [80]:
```
x
```

Out[80]:

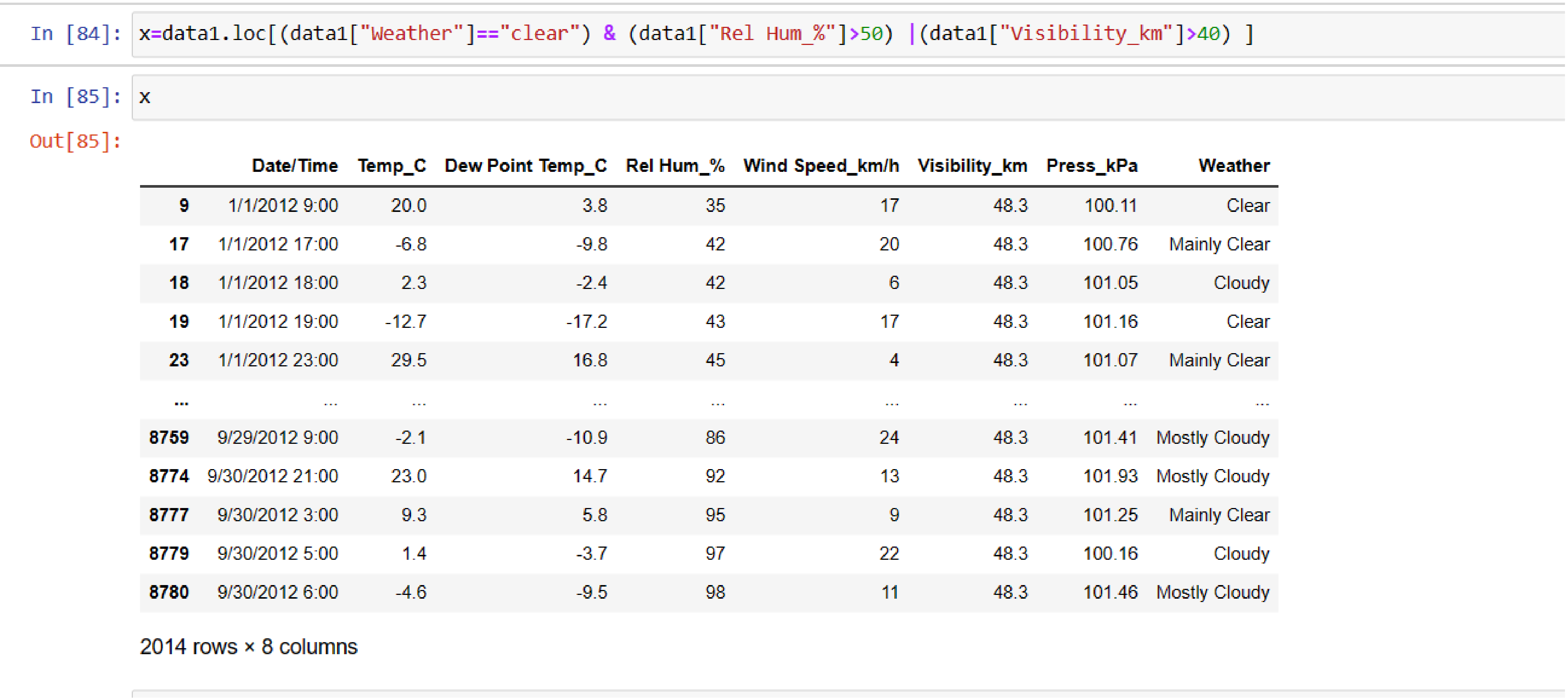
	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather

In [81]:
```
x=data1.loc[(data1["Visibility_km"]>40)]
```

In [82]:
```
x
```

Out[82]:

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
9	1/1/2012 9:00	20.0	3.8	35	17	48.3	100.11	Clear
17	1/1/2012 17:00	-6.8	-9.8	42	20	48.3	100.76	Mainly Clear
18	1/1/2012 18:00	2.3	-2.4	42	6	48.3	101.05	Cloudy
19	1/1/2012 19:00	-12.7	-17.2	43	17	48.3	101.16	Clear
23	1/1/2012 23:00	29.5	16.8	45	4	48.3	101.07	Mainly Clear
...	...	...	...	...	...	...	...	...
8759	9/29/2012 9:00	-2.1	-10.9	86	24	48.3	101.41	Mostly Cloudy
8774	9/30/2012 21:00	23.0	14.7	92	13	48.3	101.93	Mostly Cloudy
8777	9/30/2012 3:00	9.3	5.8	95	9	48.3	101.25	Mainly Clear



```
In [84]: x=data1.loc[(data1["Weather"]=="clear") & (data1["Rel Hum_%"]>50) |(data1["Visibility_km"]>40) ]
```

```
In [85]: x
```

Out[85]:

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
9	1/1/2012 9:00	20.0	3.8	35	17	48.3	100.11	Clear
17	1/1/2012 17:00	-6.8	-9.8	42	20	48.3	100.76	Mainly Clear
18	1/1/2012 18:00	2.3	-2.4	42	6	48.3	101.05	Cloudy
19	1/1/2012 19:00	-12.7	-17.2	43	17	48.3	101.16	Clear
23	1/1/2012 23:00	29.5	16.8	45	4	48.3	101.07	Mainly Clear
...	...	...	...	...	...	...	...	...
8759	9/29/2012 9:00	-2.1	-10.9	86	24	48.3	101.41	Mostly Cloudy
8774	9/30/2012 21:00	23.0	14.7	92	13	48.3	101.93	Mostly Cloudy
8777	9/30/2012 3:00	9.3	5.8	95	9	48.3	101.25	Mainly Clear
8779	9/30/2012 5:00	1.4	-3.7	97	22	48.3	100.16	Cloudy
8780	9/30/2012 6:00	-4.6	-9.5	98	11	48.3	101.46	Mostly Cloudy

2014 rows × 8 columns