## Report

1. Find all the unique "wind speed" values in the data.

These are the unique wind speeds in the given dataset.

2. Find the number of times when the "Weather is exactly clear".

3. Find the number of times when the "Wind speed was exactly 4 km/h".

```
In [9]: len(data_copy.loc[data_copy["Wind Speed_km/h"]==4])
Out[9]: 474

474 times wind speed is exactly 4km/h
```

4. Find out all the null values in the data.

```
In [10]: data_copy.isnull().sum()
Out[10]: Date/Time
                              0
         Temp_C
         Dew Point Temp_C
                              0
         Rel Hum_%
                              0
         Wind Speed_km/h
                              0
         Visibility_km
                              0
         Press_kPa
                              0
         Weather
         dtype: int64
         there are 0 null values in data set
```

5. Rename the column name "weather" of the data frame to "weather condition".

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather_Condition
0	1/1/2012 0:00	-1.8	-3.9	86	4	8.0	101.24	Fog
1	1/1/2012 1:00	-1.8	-3.7	87	4	8.0	101.24	Fog
2	1/1/2012 2:00	-1.8	-3.4	89	7	4.0	101.26	Freezing Drizzle,Fog
3	1/1/2012 3:00	-1.5	-3.2	88	6	4.0	101.27	Freezing Drizzle,Fog
4	1/1/2012 4:00	-1.5	-3.3	88	7	4.8	101.23	Fog
	***		***					***
8779	12/31/2012 19:00	0.1	-2.7	81	30	9.7	100.13	Snow
8780	12/31/2012 20:00	0.2	-2.4	83	24	9.7	100.03	Snow
8781	12/31/2012 21:00	-0.5	-1.5	93	28	4.8	99.95	Snow
8782	12/31/2012 22:00	-0.2	-1.8	89	28	9.7	99.91	Snow
8783	12/31/2012 23:00	0.0	-2.1	86	30	11.3	99.89	Snow

These are the column names of a data frame after renaming the column name "weather" to "weather condition".

6. What is the mean "visibility".

```
In [12]: data_copy["Visibility_km"].mean()
Out[12]: 27.66444672131151
```

The mean for "visibility" is 27.66.

7. What is the standard deviation of pressure in this data.

```
In [13]: data_copy["Press_kPa"].std()
Out[13]: 0.8440047459486474
```

The standard deviation of pressure is 0.8744.

8. What is the variance of Relative humidity in this data?

```
In [14]: data_copy["Rel Hum_%"].var()
Out[14]: 286.2485501984998
```

The variance of Relative humidity is 286.248.

9. Find all instances when snow was recorded.

.get_group("Snow")										
	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather		
55	1/3/2012 7:00	-14.0	-19.5	63	19	25.0	100.95	Snow		
84	1/4/2012 12:00	-13.7	-21.7	51	11	24.1	101.25	Snow		
86	1/4/2012 14:00	-11.3	-19.0	53	7	19.3	100.97	Snow		
87	1/4/2012 15:00	-10.2	-16.3	61	11	9.7	100.89	Snow		
88	1/4/2012 16:00	-9.4	-15.5	61	13	19.3	100.79	Snow		
8779	12/31/2012 19:00	0.1	-2.7	81	30	9.7	100.13	Snow		
8780	12/31/2012 20:00	0.2	-2.4	83	24	9.7	100.03	Snow		
8781	12/31/2012 21:00	-0.5	-1.5	93	28	4.8	99.95	Snow		
8782	12/31/2012 22:00	-0.2	-1.8	89	28	9.7	99.91	Snow		
8783	12/31/2012 23:00	0.0	-2.1	86	30	11.3	99.89	Snow		

There are 390 data points where the weather was recorded as snow.

10. Find all the instances when wind speed is above 24 and visibility is 25.

17]:		Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
2	3	1/1/2012 23:00	5.3	2.0	79	30	25.0	99.31	Cloudy
2	4	1/2/2012 0:00	5.2	1.5	77	35	25.0	99.26	Rain Showers
2	5	1/2/2012 1:00	4.6	0.0	72	39	25.0	99.26	Cloudy
2	6	1/2/2012 2:00	3.9	-0.9	71	32	25.0	99.26	Mostly Cloudy
2	7	1/2/2012 3:00	3.7	-1.5	69	33	25.0	99.30	Mostly Cloudy
		***							
870	5 1	12/28/2012 17:00	-8.6	-12.0	76	26	25.0	101.34	Mainly Clear
875	3 1	12/30/2012 17:00	-12.1	-15.8	74	28	25.0	101.26	Mainly Clear
875	5 1	12/30/2012 19:00	-13.4	-16.5	77	26	25.0	101.47	Mainly Clear
875	9 1	12/30/2012 23:00	-12.1	-15.1	78	28	25.0	101.52	Mostly Cloudy
876	0	12/31/2012 0:00	-11.1	-14.4	77	26	25.0	101.51	Cloudy

There are 308 records when 'wind speed \_km/h' >24 and 'visibility \_km' ==25.

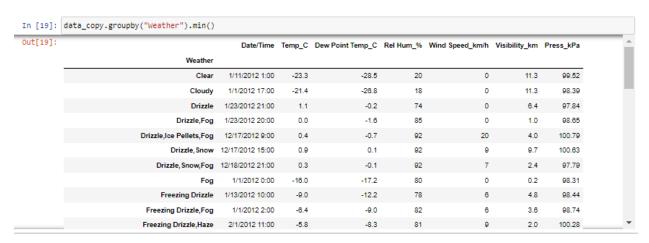
11. What is the mean value of each column against the weather condition?

```
In [6]: data_copy.groupby("Weather").mean()
```

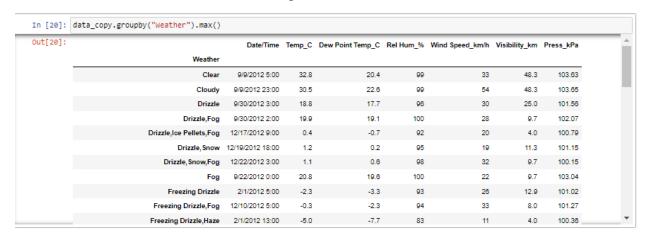
This returns all the mean value of each column against the weather condition.

In [6]:	data_copy.groupby(	"Weather	").mean().head()	)			
Out[6]:		Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa
	Weather						
	Clear	6.825716	0.089367	64.497738	10.557315	30.153243	101.587443
	Cloudy	7.970544	2.375810	69.592593	16.127315	26.625752	100.911441
	Drizzle	7.353659	5.504878	88.243902	16.097561	17.931707	100.435366
	Drizzle,Fog	8.067500	7.033750	93.275000	11.882500	5.257500	100.786625
	Drizzle,Ice Pellets,Fog	0.400000	-0.700000	92.000000	20.000000	4.000000	100.790000

12. What is the min and max value of each column against the weather condition?

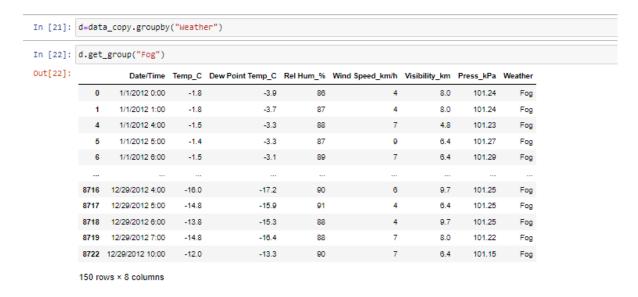


These are the min values of each column against the weather condition.



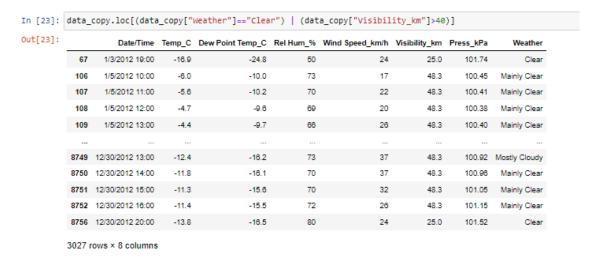
These are the max values of each column against the weather condition.

13. Show all the records where weather condition is fog.



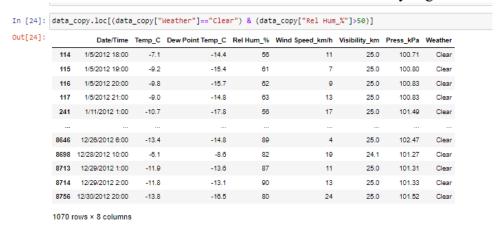
These are the records of weather condition where weather is 'fog'.

14. Find all instances when "weather is clear" or "visibility is above 40".



There are 3027 records when 'weather' is clear or 'visibility' > 40.

15. a. Find all instances when "weather is clear and Relative humidity is greater than 50.



There are 1070 records when "weather is clear and Relative humidity is greater than 50. b. 'Visibility is above 40'.

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
106	1/5/2012 10:00	-6.0	-10.0	73	17	48.3	100.45	Mainly Clear
107	1/5/2012 11:00	-5.6	-10.2	70	22	48.3	100.41	Mainly Clear
108	1/5/2012 12:00	-4.7	-9.6	69	20	48.3	100.38	Mainly Clear
109	1/5/2012 13:00	-4.4	-9.7	66	26	48.3	100.40	Mainly Clear
110	1/5/2012 14:00	-5.1	-10.7	65	22	48.3	100.46	Mainly Clear
				***				
8748	12/30/2012 12:00	-12.2	-15.7	75	26	48.3	100.91	Mostly Cloudy
8749	12/30/2012 13:00	-12.4	-16.2	73	37	48.3	100.92	Mostly Cloudy
8750	12/30/2012 14:00	-11.8	-16.1	70	37	48.3	100.98	Mainly Clear
8751	12/30/2012 15:00	-11.3	-15.6	70	32	48.3	101.05	Mainly Clear
8752	12/30/2012 16:00	-11.4	-15.5	72	26	48.3	101.15	Mainly Clear

There are 2014 records when visibility is > 40.