Report

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

```
In [5]: data_copy["Wind Speed_km/h"].unique()
Out[5]: array([ 4,  7,  6,  9, 15, 13, 20, 22, 19, 24, 30, 35, 39, 32, 33, 26, 44,
                43, 48, 37, 28, 17, 11,  0, 83, 70, 57, 46, 41, 52, 50, 63, 54,  2],
               dtype=int64)
```

These are the unique wind speeds in the given dataset.

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

```
In [6]: weather=data_copy.groupby("Weather")
In [7]: weather.ngroups
Out[7]: 50
In [8]: len(weather.get_group("Clear"))
Out[8]: 1326
```

1326 times 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             0
         Dew Point Temp_C   0
         Rel Hum_%          0
         Wind Speed_km/h    0
         Visibility_km      0
         Press_kPa          0
         Weather            0
         dtype: int64
```

there are 0 null values in data set

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

```
In [11]: data_copy.rename(columns={"Weather":"Weather_Condition"})
```

Out[11]:

	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

8784 rows × 8 columns

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.

```
In [15]: d=data_copy.groupby("Weather")
```

```
In [16]: d.get_group("Snow")
```

Out[16]:

	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

390 rows × 8 columns

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.

```
In [17]: data_copy.loc[(data_copy["Wind Speed_km/h"]>24) & (data_copy["Visibility_km"]==25)]
```

Out[17]:

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
23	1/1/2012 23:00	5.3	2.0	79	30	25.0	99.31	Cloudy
24	1/2/2012 0:00	5.2	1.5	77	35	25.0	99.26	Rain Showers
25	1/2/2012 1:00	4.6	0.0	72	39	25.0	99.26	Cloudy
26	1/2/2012 2:00	3.9	-0.9	71	32	25.0	99.26	Mostly Cloudy
27	1/2/2012 3:00	3.7	-1.5	69	33	25.0	99.30	Mostly Cloudy
...	...	...	...	...	...	...	...	...
8705	12/28/2012 17:00	-8.6	-12.0	76	26	25.0	101.34	Mainly Clear
8753	12/30/2012 17:00	-12.1	-15.8	74	28	25.0	101.26	Mainly Clear
8755	12/30/2012 19:00	-13.4	-16.5	77	26	25.0	101.47	Mainly Clear
8759	12/30/2012 23:00	-12.1	-15.1	78	28	25.0	101.52	Mostly Cloudy
8760	12/31/2012 0:00	-11.1	-14.4	77	26	25.0	101.51	Cloudy

308 rows × 8 columns

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]:

Weather	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa
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	11.862500	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?

```
In [19]: data_copy.groupby("Weather").min()
```

Out[19]:

Weather	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa
Clear	1/11/2012 1:00	-23.3	-28.5	20	0	11.3	99.52
Cloudy	1/1/2012 17:00	-21.4	-26.8	18	0	11.3	98.39
Drizzle	1/23/2012 21:00	1.1	-0.2	74	0	6.4	97.84
Drizzle,Fog	1/23/2012 20:00	0.0	-1.6	85	0	1.0	98.65
Drizzle,Ice Pellets,Fog	12/17/2012 9:00	0.4	-0.7	92	20	4.0	100.79
Drizzle,Snow	12/17/2012 15:00	0.9	0.1	92	9	9.7	100.63
Drizzle,Snow,Fog	12/18/2012 21:00	0.3	-0.1	92	7	2.4	97.79
Fog	1/1/2012 0:00	-16.0	-17.2	80	0	0.2	98.31
Freezing Drizzle	1/13/2012 10:00	-9.0	-12.2	78	6	4.8	98.44
Freezing Drizzle,Fog	1/1/2012 2:00	-6.4	-9.0	82	6	3.6	98.74
Freezing Drizzle,Haze	2/1/2012 11:00	-5.8	-8.3	81	9	2.0	100.28

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

```
In [20]: data_copy.groupby("Weather").max()
```

Out[20]:

Weather	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa
Clear	9/9/2012 5:00	32.8	20.4	99	33	48.3	103.63
Cloudy	9/9/2012 23:00	30.5	22.6	99	54	48.3	103.65
Drizzle	9/30/2012 3:00	18.8	17.7	96	30	25.0	101.56
Drizzle,Fog	9/30/2012 2:00	19.9	19.1	100	28	9.7	102.07
Drizzle,Ice Pellets,Fog	12/17/2012 9:00	0.4	-0.7	92	20	4.0	100.79
Drizzle,Snow	12/19/2012 18:00	1.2	0.2	95	19	11.3	101.15
Drizzle,Snow,Fog	12/22/2012 3:00	1.1	0.6	98	32	9.7	100.15
Fog	9/22/2012 0:00	20.8	19.6	100	22	9.7	103.04
Freezing Drizzle	2/1/2012 5:00	-2.3	-3.3	93	26	12.9	101.02
Freezing Drizzle,Fog	12/10/2012 5:00	-0.3	-2.3	94	33	8.0	101.27
Freezing Drizzle,Haze	2/1/2012 13:00	-5.0	-7.7	83	11	4.0	100.36

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

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

```
In [21]: d=data_copy.groupby("Weather")
```

```
In [22]: d.get_group("Fog")
```

Out[22]:

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
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
4	1/1/2012 4:00	-1.5	-3.3	88	7	4.8	101.23	Fog
5	1/1/2012 5:00	-1.4	-3.3	87	9	6.4	101.27	Fog
6	1/1/2012 6:00	-1.5	-3.1	89	7	6.4	101.29	Fog
...	...	...	...	...	...	...	...	...
8716	12/29/2012 4:00	-16.0	-17.2	90	6	9.7	101.25	Fog
8717	12/29/2012 5:00	-14.8	-15.9	91	4	6.4	101.25	Fog
8718	12/29/2012 6:00	-13.8	-15.3	88	4	9.7	101.25	Fog
8719	12/29/2012 7:00	-14.8	-16.4	88	7	8.0	101.22	Fog
8722	12/29/2012 10:00	-12.0	-13.3	90	7	6.4	101.15	Fog

150 rows × 8 columns

These are the records of weather condition where weather is ‘fog’.

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

```
In [23]: data_copy.loc[(data_copy["Weather"]=="Clear") | (data_copy["Visibility_km"]>40)]
```

Out[23]:

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
67	1/3/2012 19:00	-16.9	-24.8	50	24	25.0	101.74	Clear
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
...	...	...	...	...	...	...	...	...
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	100.96	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
8756	12/30/2012 20:00	-13.8	-16.5	80	24	25.0	101.52	Clear

3027 rows × 8 columns

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.

```
In [24]: data_copy.loc[(data_copy["Weather"]=="Clear") & (data_copy["Rel Hum_%"]>50)]
```

Out[24]:

	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
114	1/5/2012 18:00	-7.1	-14.4	56	11	25.0	100.71	Clear
115	1/5/2012 19:00	-9.2	-15.4	61	7	25.0	100.80	Clear
116	1/5/2012 20:00	-9.8	-15.7	62	9	25.0	100.83	Clear
117	1/5/2012 21:00	-9.0	-14.8	63	13	25.0	100.83	Clear
241	1/11/2012 1:00	-10.7	-17.8	56	17	25.0	101.49	Clear
...	...	...	...	...	...	...	...	...
8646	12/26/2012 6:00	-13.4	-14.8	89	4	25.0	102.47	Clear
8698	12/28/2012 10:00	-6.1	-8.6	82	19	24.1	101.27	Clear
8713	12/29/2012 1:00	-11.9	-13.6	87	11	25.0	101.31	Clear
8714	12/29/2012 2:00	-11.8	-13.1	90	13	25.0	101.33	Clear
8756	12/30/2012 20:00	-13.8	-16.5	80	24	25.0	101.52	Clear

1070 rows × 8 columns

There are 1070 records when "weather is clear and Relative humidity is greater than 50.

b. 'Visibility is above 40'.

```
In [25]: data_copy.loc[data_copy["Visibility_km"]>40]
```

Out[25]:

	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	100.96	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

2014 rows × 8 columns

There are 2014 records when visibility is > 40.