In [19]:
```python
# import data set

import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

In [6]:
```python
data = pd.read_csv(r"C:\Users\acer\Desktop\Weather data\1. Weather Data.csv")
```

In [14]:
```python
data
```

Out[14]:

	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
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

In [11]: `data.head()`

Out[11]:

	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
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

In [10]: `data.shape`

Out[10]: `(8784, 8)`

In [11]: `data.index`

Out[11]: `RangeIndex(start=0, stop=8784, step=1)`

In [12]: `data.columns`

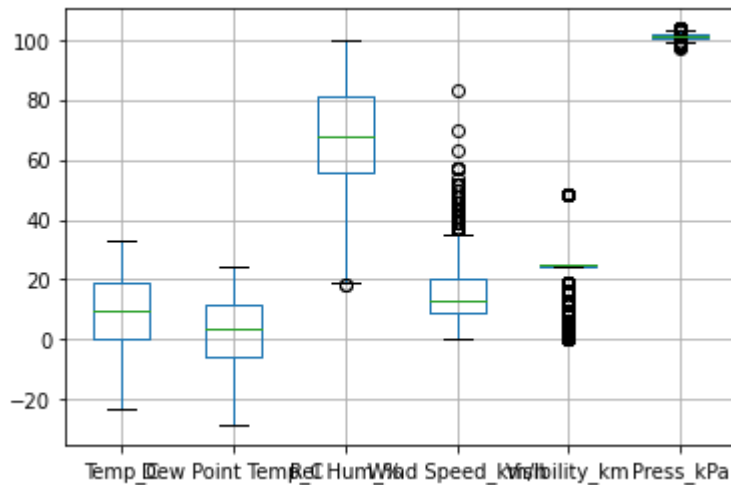Out[12]:
```
Index(['Date/Time', 'Temp_C', 'Dew Point Temp_C', 'Rel Hum_%',
       'Wind Speed_km/h', 'Visibility_km', 'Press_kPa', 'Weather'],
      dtype='object')
```

In [13]: `data.dtypes`

Out[13]:
```
Date/Time           object
Temp_C             float64
Dew Point Temp_C   float64
Rel Hum_%            int64
Wind Speed_km/h      int64
Visibility_km      float64
Press_kPa          float64
Weather             object
dtype: object
```

In [12]:
```python
data.boxplot()
```

Out[12]: `<AxesSubplot:>`

In [14]:
```python
data['Weather'].unique()
```

Out[14]:
```
array(['Fog', 'Freezing Drizzle,Fog', 'Mostly Cloudy', 'Cloudy', 'Rain',
       'Rain Showers', 'Mainly Clear', 'Snow Showers', 'Snow', 'Clear',
       'Freezing Rain,Fog', 'Freezing Rain', 'Freezing Drizzle',
       'Rain,Snow', 'Moderate Snow', 'Freezing Drizzle,Snow',
       'Freezing Rain,Snow Grains', 'Snow,Blowing Snow', 'Freezing Fog',
       'Haze', 'Rain,Fog', 'Drizzle,Fog', 'Drizzle',
       'Freezing Drizzle,Haze', 'Freezing Rain,Haze', 'Snow,Haze',
       'Snow,Fog', 'Snow,Ice Pellets', 'Rain,Haze', 'Thunderstorms,Rain',
       'Thunderstorms,Rain Showers', 'Thunderstorms,Heavy Rain Showers',
       'Thunderstorms,Rain Showers,Fog', 'Thunderstorms',
       'Thunderstorms,Rain,Fog',
       'Thunderstorms,Moderate Rain Showers,Fog', 'Rain Showers,Fog',
       'Rain Showers,Snow Showers', 'Snow Pellets', 'Rain,Snow,Fog',
       'Moderate Rain,Fog', 'Freezing Rain,Ice Pellets,Fog',
       'Drizzle,Ice Pellets,Fog', 'Drizzle,Snow', 'Rain,Ice Pellets',
       'Drizzle,Snow,Fog', 'Rain,Snow Grains', 'Rain,Snow,Ice Pellets',
       'Snow Showers,Fog', 'Moderate Snow,Blowing Snow'], dtype=object)
```

In [22]: `data.nunique()`

Out[22]:
```
Date/Time          8784
Temp_C              533
Dew Point Temp_C    489
Rel Hum_%            83
Wind Speed_km/h      34
Visibility_km        24
Press_kPa           518
Weather              50
dtype: int64
```

In [15]: `data.count()`

Out[15]:
```
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
```

In [16]: `data.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           8784 non-null   object
dtypes: float64(4), int64(2), object(2)
memory usage: 549.1+ KB
```

In [17]:
```python
# all the unique 'Wind speed' values recorded in dataset
data.head(2)
```

Out[17]:

	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

In [18]: `data["Wind Speed_km/h"].nunique()`

Out[18]: 34

In [19]: `data["Wind Speed_km/h"].unique()`

Out[19]:
```
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)
```

In [20]:
```python
# no. of times when weather was exactly 'clear'

# 1. value_counts()
data.head()
data["Weather"].value_counts()
```

Out[20]:

```
Mainly Clear                                    2106
Mostly Cloudy                                   2069
Cloudy                                          1728
Clear                                           1326
Snow                                             390
Rain                                             306
Rain Showers                                     188
Fog                                              150
Rain,Fog                                         116
Drizzle,Fog                                       80
Snow Showers                                      60
Drizzle                                           41
Snow,Fog                                          37
Snow,Blowing Snow                                 19
Rain,Snow                                         18
Thunderstorms,Rain Showers                        16
Haze                                              16
Drizzle,Snow,Fog                                  15
Freezing Rain                                     14
Freezing Drizzle,Snow                             11
Freezing Drizzle                                   7
Snow,Ice Pellets                                   6
Freezing Drizzle,Fog                               6
Snow,Haze                                          5
Rain,Snow,Ice Pellets                              4
Snow Showers,Fog                                   4
Freezing Fog                                       4
Moderate Snow                                      4
Freezing Rain,Fog                                  4
Thunderstorms,Rain                                 3
Thunderstorms,Rain Showers,Fog                     3
Rain,Haze                                          3
Freezing Drizzle,Haze                              3
Moderate Snow,Blowing Snow                         2
Thunderstorms                                      2
Rain Showers,Snow Showers                          2
Freezing Rain,Haze                                 2
Drizzle,Snow                                       2
Snow Pellets                                       1
Drizzle,Ice Pellets,Fog                            1
Freezing Rain,Ice Pellets,Fog                      1
Rain,Snow,Fog                                      1
Rain,Ice Pellets                                   1
Freezing Rain,Snow Grains                          1
Moderate Rain,Fog                                  1
Rain,Snow Grains                                   1
Rain Showers,Fog                                   1
Thunderstorms,Moderate Rain Showers,Fog            1
Thunderstorms,Heavy Rain Showers                   1
Thunderstorms,Rain,Fog                             1
Name: Weather, dtype: int64
```

In [21]: `# 2.using filtering`
`data[data["Weather"] =="Clear"]`

Out[21]:

	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
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
...	...	...	...	...	...	...	...	...
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

1326 rows × 8 columns

In [22]: `# 3.groupby()`
`data.groupby("Weather").get_group("Clear")`

Out[22]:

	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
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
...	...	...	...	...	...	...	...	...
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

1326 rows × 8 columns

In [29]:
```python
# the no. of times when the 'wind speed' was 4 km/h
data.head(2)
data[data["Wind Speed_km/h"] == 4]
```

Out[29]:

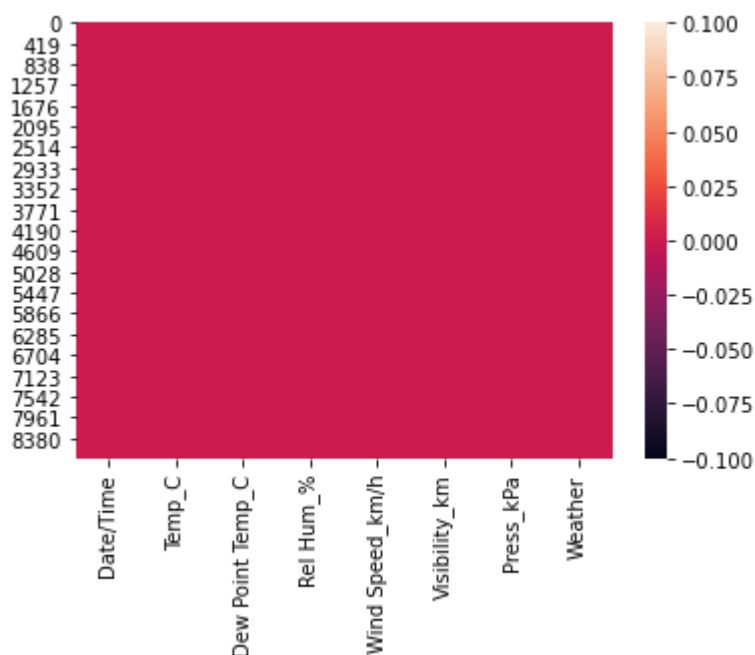|	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather
0	1/1/2012 0:00	-1.8	-3.9	86	4	8.0	101.24	Fog
1	1/1/2012 1:00	-1.8	-3.7	87	4	8.0	101.24	Fog
96	1/5/2012 0:00	-8.8	-11.7	79	4	9.7	100.32	Snow
101	1/5/2012 5:00	-7.0	-9.5	82	4	4.0	100.19	Snow
146	1/7/2012 2:00	-8.1	-11.1	79	4	19.3	100.15	Cloudy
...	...	...	...	...	...	...	...	...
8768	12/31/2012 8:00	-8.6	-10.3	87	4	3.2	101.14	Snow Showers
8769	12/31/2012 9:00	-8.1	-9.6	89	4	2.4	101.09	Snow
8770	12/31/2012 10:00	-7.4	-8.9	89	4	6.4	101.05	Snow,Fog
8772	12/31/2012 12:00	-5.8	-7.5	88	4	12.9	100.78	Snow
8773	12/31/2012 13:00	-4.6	-6.6	86	4	12.9	100.63	Snow

474 rows × 8 columns

In [30]:
```python
# isnull()
data.isnull().sum()
```

Out[30]:
```
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
```

In [20]:
```python
sns.heatmap(data.isnull())
```

Out[20]: <AxesSubplot:>

In [42]:
```python
#rename the column name
data.rename(columns={"Weather":"Weather Condition"}, inplace=True)
```

In [43]:
```python
data.head(2)
```

Out[43]:

	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

In [48]:
```python
#mean value of the 'visibility column'
data["Visibility_km"].mean()
```

Out[48]: 27.66444672131151

In [50]:
```python
data["Press_kPa"].std()
```

Out[50]: 0.8440047459486474

In [51]: `data["Rel Hum_%"].var()`

Out[51]: 286.2485501984998

In [56]: `data.groupby("Weather Condition").min().head()`

Out[56]:

Weather Condition	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

In [55]: `data.groupby("Weather Condition").max().head()`

Out[55]:

Weather Condition	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

In [59]: `data[data["Weather Condition"] == "Fog"]`

Out[59]:

	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
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

In [64]:
```python
data[(data["Weather Condition"] == "Clear") | (data["Visibility_km"] > 40)]
```

Out[64]:

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

In [66]: