In [1]:
```
import pandas pd
```

```
  File "<ipython-input-1-6bfc1f40d5ca>", line 1
    import pandas pd
                  ^
SyntaxError: invalid syntax
```

In [6]:
```
import pandas as pd
import numpy as np
import matplotlib
import seaborn as sb
import matplotlib.pyplot as plt
from matplotlib import style
%matplotlib inline
```

In [7]:
```
data = pd.read_csv(r"C:\Users\IdeaP\Downloads\Weather Data.csv")
```

In [8]:
```
data
```

Out[8]:

	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 [5]:
```
data.head()
```

Out[5]:

	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 [6]:
```
data.shape
```

Out[6]: (8784, 8)

In [7]:
```
data.shape
```

Out[7]: (8784, 8)

In [8]:
```
data.index
```

Out[8]: RangeIndex(start=0, stop=8784, step=1)

In [9]:
```
data.columns
```

Out[9]:
```
Index(['Date/Time', 'Temp_C', 'Dew Point Temp_C', 'Rel Hum_%',
       'Wind Speed_km/h', 'Visibility_km', 'Press_kPa', 'Weather'],
      dtype='object')
```

In [12]:
```
data.dtypes
```

Out[12]:
```
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 [11]:
```
data['Weather'].unique()
```

Out[11]:
```
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 [13]:
```
data['Weather'].nunique()
```

Out[13]:
```
50
```

In [14]:
```
data.nunique()
```

Out[14]:
```
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['Weather'].count
```

Out[15]:
```
<bound method Series.count of 0                       Fog
1                       Fog
2       Freezing Drizzle,Fog
3       Freezing Drizzle,Fog
4                       Fog
                ...
8779                   Snow
8780                   Snow
```

```
8781                    Snow
8782                    Snow
8783                    Snow
Name: Weather, Length: 8784, dtype: object>
```

In [16]:
```
data.count
```

Out[16]:
```
<bound method DataFrame.count of             Date/Time  Temp_C  Dew Point Temp_C  Rel Hum_%  Wind Speed_km/h  \
0        1/1/2012 0:00    -1.8              -3.9         86                4
1        1/1/2012 1:00    -1.8              -3.7         87                4
2        1/1/2012 2:00    -1.8              -3.4         89                7
3        1/1/2012 3:00    -1.5              -3.2         88                6
4        1/1/2012 4:00    -1.5              -3.3         88                7
...                ...     ...               ...        ...              ...
8779  12/31/2012 19:00     0.1              -2.7         81               30
8780  12/31/2012 20:00     0.2              -2.4         83               24
8781  12/31/2012 21:00    -0.5              -1.5         93               28
8782  12/31/2012 22:00    -0.2              -1.8         89               28
8783  12/31/2012 23:00     0.0              -2.1         86               30

      Visibility_km  Press_kPa              Weather
0               8.0     101.24                  Fog
1               8.0     101.24                  Fog
2               4.0     101.26  Freezing Drizzle,Fog
3               4.0     101.27  Freezing Drizzle,Fog
4               4.8     101.23                  Fog
...             ...        ...                  ...
8779            9.7     100.13                 Snow
8780            9.7     100.03                 Snow
8781            4.8      99.95                 Snow
8782            9.7      99.91                 Snow
8783           11.3      99.89                 Snow

[8784 rows x 8 columns]>
```

In [17]:
```
data.count()
```

Out[17]:
```
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 [19]:
```
data['Weather'].value_counts()
```

Out[19]:
```
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
Freezing Fog                           4
```

```
Moderate Snow                               4
Snow Showers,Fog                            4
Freezing Rain,Fog                           4
Rain,Snow,Ice Pellets                       4
Thunderstorms,Rain Showers,Fog              3
Rain,Haze                                   3
Freezing Drizzle,Haze                       3
Thunderstorms,Rain                          3
Drizzle,Snow                                2
Thunderstorms                               2
Freezing Rain,Haze                          2
Moderate Snow,Blowing Snow                  2
Rain Showers,Snow Showers                   2
Rain,Snow Grains                            1
Drizzle,Ice Pellets,Fog                     1
Rain,Ice Pellets                            1
Rain Showers,Fog                            1
Snow Pellets                                1
Rain,Snow,Fog                               1
Freezing Rain,Snow Grains                   1
Thunderstorms,Rain,Fog                      1
Thunderstorms,Heavy Rain Showers            1
Moderate Rain,Fog                           1
Thunderstorms,Moderate Rain Showers,Fog     1
Freezing Rain,Ice Pellets,Fog               1
Name: Weather, dtype: int64
```

In [20]:
```
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 [21]:
```
#unique" wind speed "values
data.head(2)
```

Out[21]:

	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 [22]:
```
data['Wind Speed_km/h'].nunique()
```

Out[22]:
```
34
```

In [23]:
```
data['Wind Speed_km/h'].unique()
```

Out[23]:
```
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 [29]:
```
#no of times weather is clear
data.Weather.value_counts()
```

Out[29]:
```
Mainly Clear                                0.239754
```

```
Mostly Cloudy                            0.235542
Cloudy                                   0.196721
Clear                                    0.150956
Snow                                     0.044399
Rain                                     0.034836
Rain Showers                             0.021403
Fog                                      0.017077
Rain,Fog                                 0.013206
Drizzle,Fog                              0.009107
Snow Showers                             0.006831
Drizzle                                  0.004668
Snow,Fog                                 0.004212
Snow,Blowing Snow                        0.002163
Rain,Snow                                0.002049
Thunderstorms,Rain Showers               0.001821
Haze                                     0.001821
Drizzle,Snow,Fog                         0.001708
Freezing Rain                            0.001594
Freezing Drizzle,Snow                    0.001252
Freezing Drizzle                         0.000797
Snow,Ice Pellets                         0.000683
Freezing Drizzle,Fog                     0.000683
Snow,Haze                                0.000569
Freezing Fog                             0.000455
Moderate Snow                            0.000455
Snow Showers,Fog                         0.000455
Freezing Rain,Fog                        0.000455
Rain,Snow,Ice Pellets                    0.000455
Thunderstorms,Rain Showers,Fog           0.000342
Rain,Haze                                0.000342
Freezing Drizzle,Haze                    0.000342
Thunderstorms,Rain                       0.000342
Drizzle,Snow                             0.000228
Thunderstorms                            0.000228
Freezing Rain,Haze                       0.000228
Moderate Snow,Blowing Snow               0.000228
Rain Showers,Snow Showers                0.000228
Rain,Snow Grains                         0.000114
Drizzle,Ice Pellets,Fog                  0.000114
Rain,Ice Pellets                         0.000114
Rain Showers,Fog                         0.000114
Snow Pellets                             0.000114
Rain,Snow,Fog                            0.000114
Freezing Rain,Snow Grains                0.000114
Thunderstorms,Rain,Fog                   0.000114
Thunderstorms,Heavy Rain Showers         0.000114
Moderate Rain,Fog                        0.000114
Thunderstorms,Moderate Rain Showers,Fog  0.000114
Freezing Rain,Ice Pellets,Fog            0.000114
Name: Weather, dtype: float64
```

In [31]:
```
data[data.Weather == "Clear"]
```

Out[31]:

	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 [34]:
```
data.groupby("Weather").get_group("Clear")
```

Out[34]:

	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 [36]:
```
#wind speed is exactly 4kmph
data[data['Wind Speed_km/h'] ==4]
```

Out[36]:

	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 [38]:
```
#null values in data
data.isnull().sum()
```

Out[38]:
```
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 [39]:
```
data.notnull().sum()
```

Out[39]:
```
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 [40]:
```
#renaming weather to weather condition
data.rename(columns={'Weather' :'Weather Condition'})
```

Out[40]:

	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

In [42]:
```
#mean visibility
#data.head(2)
data.Visibility_km.mean()
```

Out[42]: 27.66444672131151

In [43]:
```
#standard deviation on pressure
data.Press_kPa.std()
```

Out[43]: 0.8440047459486474

In [44]:
```
#variance of relative humidity
data['Rel Hum_%'].var()
```

Out[44]: 286.2485501984998

In [48]:
```
#wind speed is above 24 and visibility is 25
data[(data['Wind Speed_km/h']>24) & (data['Visibility_km']==25)]
```

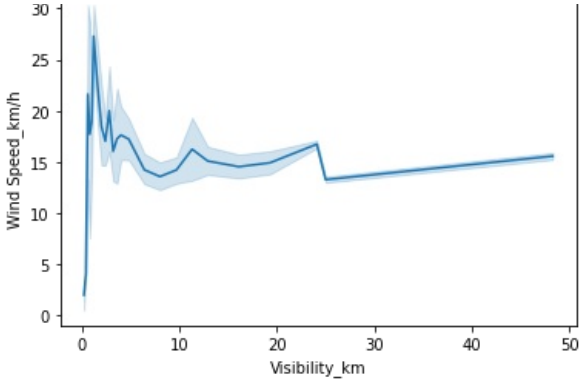
Out[48]:

	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

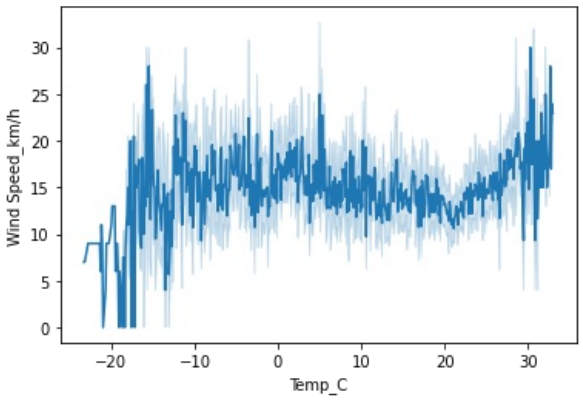
In [9]:
```
sb.lineplot(data=data,x='Visibility_km',y='Wind Speed_km/h')
```

Out[9]: <AxesSubplot:xlabel='Visibility_km', ylabel='Wind Speed_km/h'>



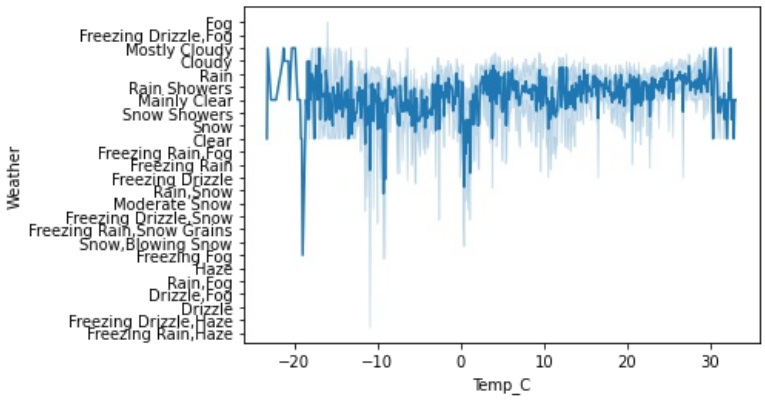
In [10]:
```
sb.lineplot(data=data,x='Temp_C',y='Wind Speed_km/h')
```

Out[10]: <AxesSubplot:xlabel='Temp_C', ylabel='Wind Speed_km/h'>



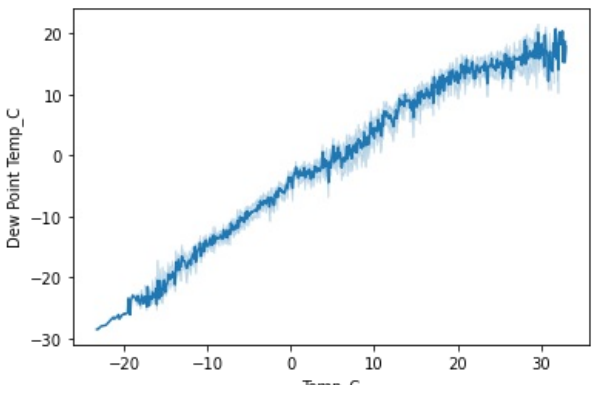
In [11]:
```
sb.lineplot(data=data,x='Temp_C',y='Weather')
```

Out[11]: <AxesSubplot:xlabel='Temp_C', ylabel='Weather'>



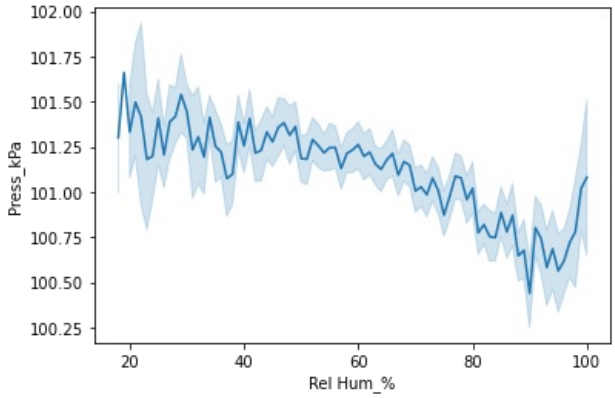
In [13]:
```
sb.lineplot(data=data,x='Temp_C',y='Dew Point Temp_C')
```

Out[13]: <AxesSubplot:xlabel='Temp_C', ylabel='Dew Point Temp_C'>



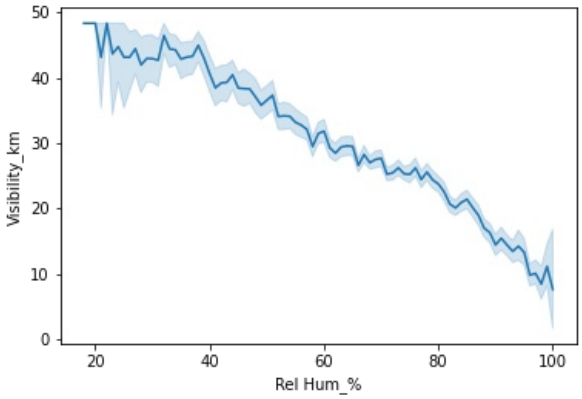
In [14]:
```
sb.lineplot(data=data,x='Rel Hum_%',y='Press_kPa')
```

Out[14]: <AxesSubplot:xlabel='Rel Hum_%', ylabel='Press_kPa'>



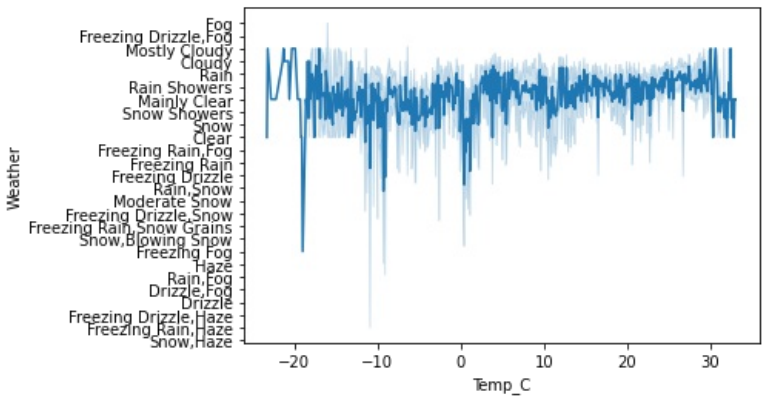
In [15]:
```
sb.lineplot(data=data,x='Rel Hum_%',y='Visibility_km')
```

Out[15]: <AxesSubplot:xlabel='Rel Hum_%', ylabel='Visibility_km'>



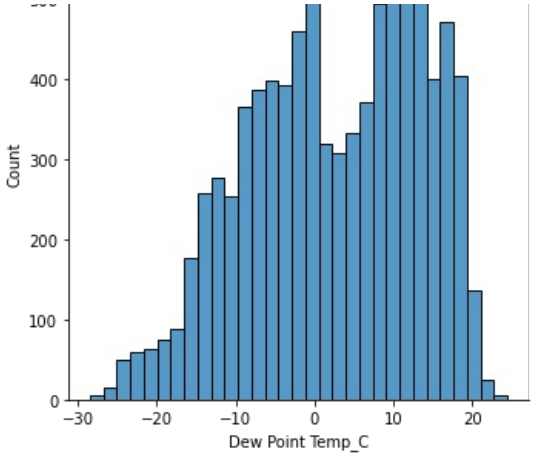
In [16]:
```
sb.lineplot(data=data,x='Temp_C',y='Weather')
```

Out[16]: <AxesSubplot:xlabel='Temp_C', ylabel='Weather'>



In [20]:
```
sb.displot(data['Dew Point Temp_C'])
```

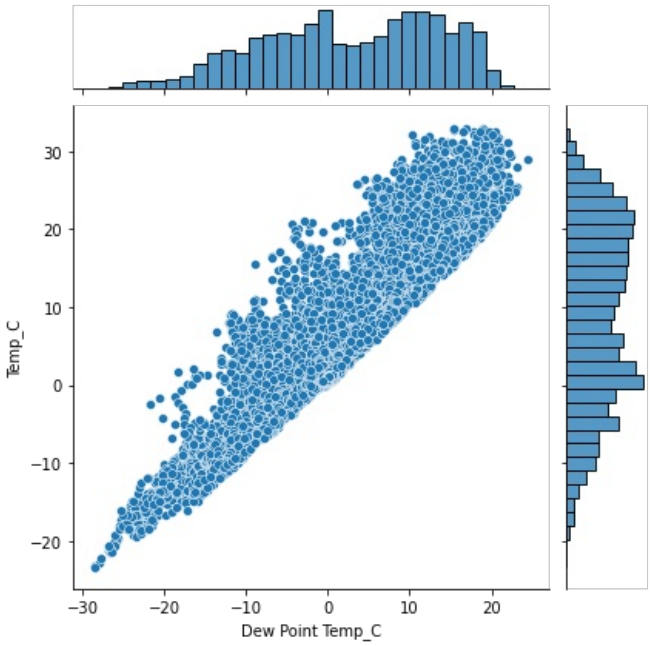
Out[20]: <seaborn.axisgrid.FacetGrid at 0x1e722045970>



In [21]:
```
sb.jointplot(data['Dew Point Temp_C'], data['Temp_C'])
```

```
C:\Users\IdeaP\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other ar
guments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```

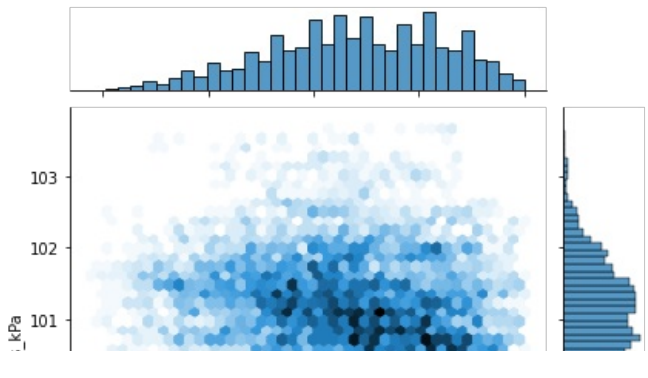
Out[21]: <seaborn.axisgrid.JointGrid at 0x1e7227f2760>

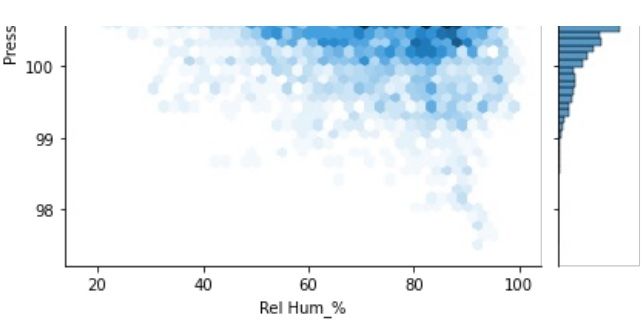


In [22]:
```
sb.jointplot(data['Rel Hum_%'], data['Press_kPa'], kind="hex")
```

```
C:\Users\IdeaP\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other ar
guments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```

Out[22]: <seaborn.axisgrid.JointGrid at 0x1e721e43dc0>





In [ ]: