Transformation of Weather Data

February 9, 2024

0.0.1 About Dataset:

0.0.2 Context:

The data contains the weather report and the daily summary of weather

0.0.3 Contents:

The data contains: Date, temperature, Humidity,Windspeed, Visibility, Loud Cover, Daily Summary

0.0.4 Problem Statement :

Transform the given Data.

```python
[1]: #Import library:
     import pandas as pd
```

```python
[2]: #load the data
     df=pd.read_csv("C:/Users/amitm/Desktop/New folder/Task Impetus/Class/Python/
       ↪Case Study/Weather/Weather.csv")
     print(type(df))
```

```
<class 'pandas.core.frame.DataFrame'>
```

Data Description:

```python
[3]: #df.head() show the first 5 rows
     df.head()
```

```
[3]:                  Formatted Date        Summary Precip Type  Temperature (C)  \
     0  2006-04-01 00:00:00.000 +0200  Partly Cloudy        rain         9.472222
     1  2006-04-01 01:00:00.000 +0200  Partly Cloudy        rain         9.355556
     2  2006-04-01 02:00:00.000 +0200  Mostly Cloudy        rain         9.377778
     3  2006-04-01 03:00:00.000 +0200  Partly Cloudy        rain         8.288889
     4  2006-04-01 04:00:00.000 +0200  Mostly Cloudy        rain         8.755556

        Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
     0                  7.388889      0.89            14.1197
     1                  7.227778      0.86            14.2646
     2                  9.377778      0.89             3.9284
     3                  5.944444      0.83            14.1036
```

```
4                              6.977778        0.83                11.0446
```

```
     Wind Bearing (degrees)  Visibility (km)  Loud Cover  Pressure (millibars)  \
0                     251.0          15.8263         0.0               1015.13
1                     259.0          15.8263         0.0               1015.63
2                     204.0          14.9569         0.0               1015.94
3                     269.0          15.8263         0.0               1016.41
4                     259.0          15.8263         0.0               1016.51
```

```
                         Daily Summary
0  Partly cloudy throughout the day.
1  Partly cloudy throughout the day.
2  Partly cloudy throughout the day.
3  Partly cloudy throughout the day.
4  Partly cloudy throughout the day.
```

[4]:
```python
#df.head(n) show the first n rows
df.head(10)
```

[4]:
```
                 Formatted Date        Summary Precip Type  Temperature (C)  \
0  2006-04-01 00:00:00.000 +0200  Partly Cloudy        rain         9.472222
1  2006-04-01 01:00:00.000 +0200  Partly Cloudy        rain         9.355556
2  2006-04-01 02:00:00.000 +0200  Mostly Cloudy        rain         9.377778
3  2006-04-01 03:00:00.000 +0200  Partly Cloudy        rain         8.288889
4  2006-04-01 04:00:00.000 +0200  Mostly Cloudy        rain         8.755556
5  2006-04-01 05:00:00.000 +0200  Partly Cloudy        rain         9.222222
6  2006-04-01 06:00:00.000 +0200  Partly Cloudy        rain         7.733333
7  2006-04-01 07:00:00.000 +0200  Partly Cloudy        rain         8.772222
8  2006-04-01 08:00:00.000 +0200  Partly Cloudy        rain        10.822222
9  2006-04-01 09:00:00.000 +0200  Partly Cloudy        rain        13.772222
```

```
   Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
0                  7.388889      0.89            14.1197
1                  7.227778      0.86            14.2646
2                  9.377778      0.89             3.9284
3                  5.944444      0.83            14.1036
4                  6.977778      0.83            11.0446
5                  7.111111      0.85            13.9587
6                  5.522222      0.95            12.3648
7                  6.527778      0.89            14.1519
8                 10.822222      0.82            11.3183
9                 13.772222      0.72            12.5258
```

```
   Wind Bearing (degrees)  Visibility (km)  Loud Cover  Pressure (millibars)  \
0                   251.0          15.8263         0.0               1015.13
1                   259.0          15.8263         0.0               1015.63
2                   204.0          14.9569         0.0               1015.94
```

3	269.0	15.8263	0.0		1016.41
4	259.0	15.8263	0.0		1016.51
5	258.0	14.9569	0.0		1016.66
6	259.0	9.9820	0.0		1016.72
7	260.0	9.9820	0.0		1016.84
8	259.0	9.9820	0.0		1017.37
9	279.0	9.9820	0.0		1017.22

```
                     Daily Summary
0  Partly cloudy throughout the day.
1  Partly cloudy throughout the day.
2  Partly cloudy throughout the day.
3  Partly cloudy throughout the day.
4  Partly cloudy throughout the day.
5  Partly cloudy throughout the day.
6  Partly cloudy throughout the day.
7  Partly cloudy throughout the day.
8  Partly cloudy throughout the day.
9  Partly cloudy throughout the day.
```

[5]:
```
#tail() show th last 5 records
df.tail()
```

[5]:
```
                    Formatted Date          Summary Precip Type  \
96448  2016-09-09 19:00:00.000 +0200  Partly Cloudy        rain
96449  2016-09-09 20:00:00.000 +0200  Partly Cloudy        rain
96450  2016-09-09 21:00:00.000 +0200  Partly Cloudy        rain
96451  2016-09-09 22:00:00.000 +0200  Partly Cloudy        rain
96452  2016-09-09 23:00:00.000 +0200  Partly Cloudy        rain


       Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
96448        26.016667                 26.016667      0.43            10.9963
96449        24.583333                 24.583333      0.48            10.0947
96450        22.038889                 22.038889      0.56             8.9838
96451        21.522222                 21.522222      0.60            10.5294
96452        20.438889                 20.438889      0.61             5.8765


       Wind Bearing (degrees)  Visibility (km)  Loud Cover  \
96448                    31.0          16.1000         0.0
96449                    20.0          15.5526         0.0
96450                    30.0          16.1000         0.0
96451                    20.0          16.1000         0.0
96452                    39.0          15.5204         0.0


       Pressure (millibars)                        Daily Summary
96448               1014.36  Partly cloudy starting in the morning.
96449               1015.16  Partly cloudy starting in the morning.
```

```
96450                        1015.66  Partly cloudy starting in the morning.
96451                        1015.95  Partly cloudy starting in the morning.
96452                        1016.16  Partly cloudy starting in the morning.
```

[6]: *#df.tail(n) show the last n no rows*
     df.tail(5)

[6]:
```
                   Formatted Date        Summary Precip Type  \
96448  2016-09-09 19:00:00.000 +0200  Partly Cloudy        rain
96449  2016-09-09 20:00:00.000 +0200  Partly Cloudy        rain
96450  2016-09-09 21:00:00.000 +0200  Partly Cloudy        rain
96451  2016-09-09 22:00:00.000 +0200  Partly Cloudy        rain
96452  2016-09-09 23:00:00.000 +0200  Partly Cloudy        rain

       Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
96448        26.016667                 26.016667      0.43            10.9963
96449        24.583333                 24.583333      0.48            10.0947
96450        22.038889                 22.038889      0.56             8.9838
96451        21.522222                 21.522222      0.60            10.5294
96452        20.438889                 20.438889      0.61             5.8765

       Wind Bearing (degrees)  Visibility (km)  Loud Cover  \
96448                    31.0          16.1000         0.0
96449                    20.0          15.5526         0.0
96450                    30.0          16.1000         0.0
96451                    20.0          16.1000         0.0
96452                    39.0          15.5204         0.0

       Pressure (millibars)                         Daily Summary
96448               1014.36  Partly cloudy starting in the morning.
96449               1015.16  Partly cloudy starting in the morning.
96450               1015.66  Partly cloudy starting in the morning.
96451               1015.95  Partly cloudy starting in the morning.
96452               1016.16  Partly cloudy starting in the morning.
```

[7]: *# No of rows and cols:*
     print("No of rows and cols :",df.shape)

```
No of rows and cols : (96453, 12)
```

[8]: *# No of rows:*
     print("No of rows :",df.shape[0])
     print("No of rows :",len(df))

```
No of rows  : 96453
No of rows  : 96453
```

```
[9]: # No of cols
     print("No of cols  :",df.shape[1])
     print("No of cols  :",len(df.columns))
```

```
No of cols  : 12
No of cols  : 12
```

```
[10]: # List of column name
      df.columns
```

```
[10]: Index(['Formatted Date', 'Summary', 'Precip Type', 'Temperature (C)',
             'Apparent Temperature (C)', 'Humidity', 'Wind Speed (km/h)',
             'Wind Bearing (degrees)', 'Visibility (km)', 'Loud Cover',
             'Pressure (millibars)', 'Daily Summary'],
            dtype='object')
```

```
[11]: # Show data type of each columns
      df.dtypes
```

```
[11]: Formatted Date              object
      Summary                     object
      Precip Type                 object
      Temperature (C)             float64
      Apparent Temperature (C)    float64
      Humidity                    float64
      Wind Speed (km/h)           float64
      Wind Bearing (degrees)      float64
      Visibility (km)             float64
      Loud Cover                  float64
      Pressure (millibars)        float64
      Daily Summary               object
      dtype: object
```

```
[12]: # Get infformation of the dataframe
      df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 96453 entries, 0 to 96452
Data columns (total 12 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   Formatted Date            96453 non-null  object
 1   Summary                   96453 non-null  object
 2   Precip Type               95936 non-null  object
 3   Temperature (C)           96453 non-null  float64
 4   Apparent Temperature (C)  96453 non-null  float64
 5   Humidity                  96453 non-null  float64
 6   Wind Speed (km/h)         96453 non-null  float64
 7   Wind Bearing (degrees)    96453 non-null  float64
```

```
8    Visibility (km)       96453 non-null  float64
9    Loud Cover            96453 non-null  float64
10   Pressure (millibars)  96453 non-null  float64
11   Daily Summary         96453 non-null  object
dtypes: float64(8), object(4)
memory usage: 8.8+ MB
```

[13]: ```
# Statiscal information of numeric columns
df.describe()
```

[13]:
|        | Temperature (C) | Apparent Temperature (C) | Humidity   \ |
|--------|-----------------|--------------------------|--------------|
| count  | 96453.000000    | 96453.000000             | 96453.000000 |
| mean   | 11.932678       | 10.855029                | 0.734899     |
| std    | 9.551546        | 10.696847                | 0.195473     |
| min    | -21.822222      | -27.716667               | 0.000000     |
| 25%    | 4.688889        | 2.311111                 | 0.600000     |
| 50%    | 12.000000       | 12.000000                | 0.780000     |
| 75%    | 18.838889       | 18.838889                | 0.890000     |
| max    | 39.905556       | 39.344444                | 1.000000     |

|        | Wind Speed (km/h) | Wind Bearing (degrees) | Visibility (km) | Loud Cover  \ |
|--------|-------------------|------------------------|-----------------|---------------|
| count  | 96453.000000      | 96453.000000           | 96453.000000    | 96453.0       |
| mean   | 10.810640         | 187.509232             | 10.347325       | 0.0           |
| std    | 6.913571          | 107.383428             | 4.192123        | 0.0           |
| min    | 0.000000          | 0.000000               | 0.000000        | 0.0           |
| 25%    | 5.828200          | 116.000000             | 8.339800        | 0.0           |
| 50%    | 9.965900          | 180.000000             | 10.046400       | 0.0           |
| 75%    | 14.135800         | 290.000000             | 14.812000       | 0.0           |
| max    | 63.852600         | 359.000000             | 16.100000       | 0.0           |

|        | Pressure (millibars) |
|--------|----------------------|
| count  | 96453.000000         |
| mean   | 1003.235956          |
| std    | 116.969906           |
| min    | 0.000000             |
| 25%    | 1011.900000          |
| 50%    | 1016.450000          |
| 75%    | 1021.090000          |
| max    | 1046.380000          |

[14]: ```
# Statical information of the data type object
df.describe(include='object')
```

[14]:
	Formatted Date	Summary	Precip Type \
count	96453	96453	95936
unique	96429	27	2
top	2010-08-02 00:00:00.000 +0200	Partly Cloudy	rain
freq	2	31733	85224

```
                       Daily Summary
        count                            96453
        unique                             214
        top     Mostly cloudy throughout the day.
        freq                             20085
```

```
[15]: # Get unique columns
      df['Precip Type'].unique()
```

```
[15]: array(['rain', 'snow', nan], dtype=object)
```

```
[16]: df['Summary'].unique()
```

```
[16]: array(['Partly Cloudy', 'Mostly Cloudy', 'Overcast', 'Foggy',
             'Breezy and Mostly Cloudy', 'Clear', 'Breezy and Partly Cloudy',
             'Breezy and Overcast', 'Humid and Mostly Cloudy',
             'Humid and Partly Cloudy', 'Windy and Foggy', 'Windy and Overcast',
             'Breezy and Foggy', 'Windy and Partly Cloudy', 'Breezy',
             'Dry and Partly Cloudy', 'Windy and Mostly Cloudy',
             'Dangerously Windy and Partly Cloudy', 'Dry', 'Windy',
             'Humid and Overcast', 'Light Rain', 'Drizzle', 'Windy and Dry',
             'Dry and Mostly Cloudy', 'Breezy and Dry', 'Rain'], dtype=object)
```

```
[17]: df['Daily Summary'].unique()
```

```
[17]: array(['Partly cloudy throughout the day.',
             'Mostly cloudy throughout the day.', 'Foggy in the evening.',
             'Foggy overnight and breezy in the morning.',
             'Overcast throughout the day.', 'Partly cloudy until night.',
             'Mostly cloudy until night.',
             'Foggy starting overnight continuing until morning.',
             'Foggy in the morning.', 'Partly cloudy until evening.',
             'Partly cloudy starting in the morning.',
             'Mostly cloudy starting overnight continuing until night.',
             'Mostly cloudy until evening.',
             'Partly cloudy starting in the morning continuing until evening.',
             'Partly cloudy starting in the afternoon.',
             'Partly cloudy starting overnight.',
             'Partly cloudy until morning.',
             'Partly cloudy starting overnight continuing until night.',
             'Partly cloudy starting in the afternoon continuing until night.',
             'Mostly cloudy starting overnight.',
             'Partly cloudy until afternoon.',
             'Mostly cloudy until night and breezy in the afternoon.',
             'Foggy starting in the evening.', 'Foggy throughout the day.',
             'Foggy starting in the evening continuing until night.',
```

```
    'Mostly cloudy until morning.',
    'Foggy starting in the morning continuing until evening.',
    'Foggy starting overnight continuing until afternoon.',
    'Partly cloudy starting in the morning continuing until afternoon.',
    'Foggy starting overnight.', 'Foggy until morning.',
    'Foggy starting overnight continuing until evening.',
    'Foggy starting in the afternoon.',
    'Partly cloudy starting overnight continuing until afternoon.',
    'Partly cloudy starting in the morning continuing until night.',
    'Overcast until night.',
    'Mostly cloudy starting overnight continuing until evening.',
    'Foggy overnight.', 'Partly cloudy in the morning.',
    'Mostly cloudy starting in the morning.',
    'Foggy starting in the afternoon continuing until evening.',
    'Mostly cloudy until afternoon.',
    'Foggy starting overnight continuing until night.',
    'Mostly cloudy throughout the day and breezy in the evening.',
    'Foggy starting in the morning continuing until afternoon.',
    'Partly cloudy in the afternoon.', 'Clear throughout the day.',
    'Partly cloudy starting in the afternoon continuing until evening.',
    'Partly cloudy overnight.', 'Overcast until evening.',
    'Foggy in the morning and breezy starting in the afternoon continuing
until night.',
    'Breezy starting overnight continuing until afternoon and foggy starting
in the morning continuing until evening.',
    'Partly cloudy starting overnight continuing until morning.',
    'Mostly cloudy throughout the day and breezy in the afternoon.',
    'Mostly cloudy starting overnight and breezy in the afternoon.',
    'Partly cloudy throughout the day and breezy starting in the morning
continuing until night.',
    'Mostly cloudy throughout the day and breezy in the morning.',
    'Partly cloudy starting in the evening continuing until night.',
    'Mostly cloudy until night and breezy starting in the morning continuing
until afternoon.',
    'Partly cloudy starting in the morning continuing until evening and
breezy starting in the morning continuing until afternoon.',
    'Partly cloudy throughout the day and breezy starting in the morning
continuing until afternoon.',
    'Partly cloudy throughout the day and breezy starting in the morning
continuing until evening.',
    'Foggy until afternoon.',
    'Overcast until night and breezy overnight.',
    'Breezy until morning and mostly cloudy throughout the day.',
    'Mostly cloudy starting in the morning continuing until night.',
    'Breezy starting overnight continuing until morning and partly cloudy
starting overnight continuing until evening.',
    'Partly cloudy in the evening.',
```

```
    'Mostly cloudy starting overnight continuing until afternoon.',
    'Mostly cloudy starting in the morning continuing until afternoon.',
    'Mostly cloudy starting in the afternoon.',
    'Mostly cloudy starting in the morning continuing until evening.',
    'Partly cloudy starting overnight continuing until afternoon and breezy
in the afternoon.',
    'Partly cloudy starting overnight and breezy in the afternoon.',
    'Mostly cloudy starting in the morning and breezy in the evening.',
    'Foggy starting in the afternoon continuing until night.',
    'Foggy until night.',
    'Foggy starting in the morning continuing until night.',
    'Foggy until evening.', 'Foggy starting in the morning.',
    'Partly cloudy starting overnight continuing until evening.',
    'Partly cloudy starting overnight continuing until evening and breezy
starting in the morning continuing until evening.',
    'Breezy starting overnight continuing until morning and foggy in the
evening.',
    'Mostly cloudy throughout the day and breezy starting in the morning
continuing until evening.',
    'Partly cloudy until evening and breezy starting in the morning
continuing until afternoon.',
    'Mostly cloudy starting in the afternoon continuing until night.',
    'Breezy starting overnight continuing until afternoon and mostly cloudy
starting overnight continuing until evening.',
    'Mostly cloudy throughout the day and windy starting in the morning
continuing until evening.',
    'Breezy and partly cloudy in the afternoon.',
    'Mostly cloudy starting overnight and breezy starting in the morning
continuing until afternoon.',
    'Partly cloudy until night and breezy starting in the morning continuing
until afternoon.',
    'Breezy and mostly cloudy overnight.',
    'Mostly cloudy throughout the day and breezy overnight.',
    'Mostly cloudy throughout the day and breezy starting in the morning
continuing until afternoon.',
    'Partly cloudy throughout the day and breezy in the morning.',
    'Partly cloudy starting in the morning continuing until evening and
breezy starting in the afternoon continuing until evening.',
    'Partly cloudy throughout the day and breezy starting in the afternoon
continuing until evening.',
    'Mostly cloudy starting overnight and breezy in the morning.',
    'Partly cloudy starting in the afternoon and breezy in the afternoon.',
    'Partly cloudy starting in the morning and breezy in the evening.',
    'Partly cloudy until evening and breezy in the morning.',
    'Partly cloudy starting overnight continuing until evening and breezy
starting in the morning continuing until afternoon.',
    'Partly cloudy starting overnight continuing until evening and breezy in
```

the evening.',
        'Mostly cloudy throughout the day and breezy starting in the evening.',
        'Mostly cloudy throughout the day and windy starting in the morning
continuing until night.',
        'Breezy starting overnight continuing until morning and partly cloudy
starting in the morning.',
        'Mostly cloudy starting in the morning and breezy overnight.',
        'Overcast throughout the day and breezy starting overnight continuing
until morning.',
        'Partly cloudy throughout the day and breezy in the evening.',
        'Mostly cloudy until evening and breezy starting in the morning
continuing until afternoon.',
        'Mostly cloudy until night and breezy in the evening.',
        'Partly cloudy starting in the evening.',
        'Overcast starting in the morning.',
        'Mostly cloudy starting overnight continuing until evening and breezy
starting overnight continuing until morning.',
        'Partly cloudy starting overnight continuing until morning and breezy
starting in the morning continuing until afternoon.',
        'Partly cloudy until evening and breezy starting in the morning
continuing until evening.',
        'Breezy starting in the morning continuing until afternoon and partly
cloudy starting in the morning.',
        'Partly cloudy starting in the morning and breezy starting in the
afternoon continuing until evening.',
        'Mostly cloudy starting overnight continuing until morning.',
        'Mostly cloudy throughout the day and breezy starting overnight
continuing until afternoon.',
        'Breezy starting overnight continuing until morning and foggy
overnight.',
        'Mostly cloudy throughout the day and breezy starting overnight
continuing until morning.',
        'Overcast throughout the day and breezy in the morning.',
        'Overcast throughout the day and breezy in the evening.',
        'Mostly cloudy starting in the morning continuing until night and breezy
in the afternoon.',
        'Mostly cloudy until night and breezy starting in the evening continuing
until night.',
        'Partly cloudy until night and breezy in the morning.',
        'Partly cloudy until evening and breezy overnight.',
        'Partly cloudy starting overnight continuing until night and windy
starting in the morning continuing until afternoon.',
        'Breezy starting in the morning continuing until afternoon and mostly
cloudy starting in the morning.',
        'Foggy starting overnight continuing until morning and breezy starting in
the evening.',
        'Mostly cloudy until night and breezy starting in the afternoon.',

10

'Foggy in the afternoon.',
        'Mostly cloudy until night and breezy starting in the afternoon
continuing until night.',
        'Foggy starting overnight continuing until morning and breezy starting in
the evening continuing until night.',
        'Breezy until afternoon and mostly cloudy throughout the day.',
        'Mostly cloudy throughout the day and breezy starting in the morning
continuing until night.',
        'Partly cloudy starting overnight continuing until evening and breezy in
the morning.',
        'Mostly cloudy starting in the morning and breezy in the afternoon.',
        'Mostly cloudy starting overnight continuing until night and breezy
starting in the morning continuing until evening.',
        'Foggy starting overnight continuing until morning and breezy starting in
the morning continuing until afternoon.',
        'Mostly cloudy until evening and windy starting in the morning continuing
until afternoon.',
        'Foggy starting overnight continuing until afternoon and breezy in the
morning.',
        'Foggy starting in the morning continuing until afternoon and breezy
starting in the evening.',
        'Partly cloudy starting overnight and breezy starting in the morning
continuing until afternoon.',
        'Foggy starting overnight continuing until morning and breezy in the
afternoon.',
        'Mostly cloudy starting overnight and breezy starting in the afternoon
continuing until evening.',
        'Overcast throughout the day and breezy starting overnight continuing
until afternoon.',
        'Partly cloudy starting in the morning continuing until evening and
breezy in the afternoon.',
        'Partly cloudy starting in the morning continuing until night and breezy
starting in the afternoon continuing until evening.',
        'Mostly cloudy until night and breezy starting in the evening.',
        'Breezy in the morning and mostly cloudy starting in the morning.',
        'Mostly cloudy until night and breezy starting in the morning continuing
until evening.',
        'Partly cloudy starting overnight continuing until evening and windy
starting in the morning continuing until evening.',
        'Breezy in the morning and partly cloudy starting in the evening
continuing until night.',
        'Partly cloudy overnight and breezy starting in the morning continuing
until afternoon.',
        'Light rain in the morning.', 'Light rain until morning.',
        'Light rain in the morning and afternoon.',
        'Partly cloudy starting in the morning continuing until night and breezy
starting in the morning continuing until afternoon.',

'Breezy starting in the afternoon continuing until night and mostly
cloudy starting in the evening.',
        'Mostly cloudy throughout the day and breezy starting in the evening
continuing until night.',
        'Foggy starting in the afternoon and breezy starting in the afternoon
continuing until evening.',
        'Breezy and foggy until morning.',
        'Mostly cloudy until night and breezy starting overnight continuing until
morning.',
        'Partly cloudy starting overnight continuing until night and breezy in
the morning.',
        'Partly cloudy starting overnight continuing until night and breezy in
the afternoon.',
        'Mostly cloudy starting in the morning and breezy starting in the
afternoon continuing until evening.',
        'Partly cloudy starting overnight and breezy starting in the evening.',
        'Breezy overnight and overcast throughout the day.',
        'Partly cloudy until night and breezy in the afternoon.',
        'Mostly cloudy starting overnight and breezy starting in the evening.',
        'Breezy overnight and partly cloudy until evening.',
        'Mostly cloudy starting in the evening.',
        'Mostly cloudy throughout the day and breezy starting in the afternoon.',
        'Mostly cloudy throughout the day and breezy starting in the afternoon
continuing until evening.',
        'Mostly cloudy until night and windy starting in the morning continuing
until afternoon.',
        'Breezy and foggy starting in the evening.',
        'Breezy overnight and partly cloudy throughout the day.',
        'Overcast throughout the day and breezy starting in the evening.',
        'Breezy until evening and foggy in the morning.',
        'Breezy overnight and mostly cloudy throughout the day.',
        'Partly cloudy until evening and breezy in the afternoon.',
        'Partly cloudy starting in the morning and breezy starting in the morning
continuing until afternoon.',
        'Mostly cloudy until evening and breezy in the evening.',
        'Windy in the afternoon.', 'Overcast until morning.',
        'Mostly cloudy overnight.',
        'Foggy starting in the morning continuing until evening and breezy in the
evening.',
        'Breezy starting overnight continuing until morning.',
        'Breezy starting in the afternoon continuing until evening and foggy
starting in the evening.',
        'Mostly cloudy until night and breezy overnight.',
        'Mostly cloudy starting in the morning and windy in the evening.',
        'Partly cloudy throughout the day and windy starting in the morning
continuing until afternoon.',
        'Breezy until afternoon and overcast throughout the day.',

```
        'Breezy in the morning and foggy in the evening.',
        'Breezy starting in the afternoon continuing until evening and foggy in
the evening.',
        'Breezy starting in the morning continuing until night.',
        'Breezy in the morning and mostly cloudy starting in the evening.',
        'Mostly cloudy until evening and breezy in the afternoon.',
        'Mostly cloudy until night and breezy starting in the afternoon
continuing until evening.',
        'Mostly cloudy until evening and breezy starting overnight continuing
until morning.',
        'Overcast throughout the day and breezy in the afternoon.',
        'Overcast throughout the day and breezy starting in the morning
continuing until evening.',
        'Overcast throughout the day and breezy overnight.',
        'Overcast starting in the afternoon.',
        'Partly cloudy throughout the day and breezy in the afternoon.',
        'Light rain starting overnight.',
        'Drizzle starting in the evening.', 'Drizzle until morning.',
        'Rain throughout the day.', 'Rain until morning.',
        'Light rain overnight.', 'Rain until afternoon.'], dtype=object)
```

[18]:
```python
# Show values of the Columns Specified
df[['Formatted Date', 'Summary', 'Precip Type',         'Temperature (C)',
 ↪'Humidity', 'Wind Speed (km/h)']]
```

[18]:
```
                        Formatted Date         Summary Precip Type  \
0      2006-04-01 00:00:00.000 +0200  Partly Cloudy        rain
1      2006-04-01 01:00:00.000 +0200  Partly Cloudy        rain
2      2006-04-01 02:00:00.000 +0200  Mostly Cloudy        rain
3      2006-04-01 03:00:00.000 +0200  Partly Cloudy        rain
4      2006-04-01 04:00:00.000 +0200  Mostly Cloudy        rain
...                              ...            ...         ...
96448  2016-09-09 19:00:00.000 +0200  Partly Cloudy        rain
96449  2016-09-09 20:00:00.000 +0200  Partly Cloudy        rain
96450  2016-09-09 21:00:00.000 +0200  Partly Cloudy        rain
96451  2016-09-09 22:00:00.000 +0200  Partly Cloudy        rain
96452  2016-09-09 23:00:00.000 +0200  Partly Cloudy        rain

       Temperature (C)  Humidity  Wind Speed (km/h)
0             9.472222      0.89            14.1197
1             9.355556      0.86            14.2646
2             9.377778      0.89             3.9284
3             8.288889      0.83            14.1036
4             8.755556      0.83            11.0446
...                ...       ...                ...
96448        26.016667      0.43            10.9963
96449        24.583333      0.48            10.0947
```

```
96450        22.038889      0.56              8.9838
96451        21.522222      0.60             10.5294
96452        20.438889      0.61              5.8765

[96453 rows x 6 columns]
```

Slicing:

```
[19]: #first value of the first row
      df.loc[0]
```

```
[19]: Formatted Date                 2006-04-01 00:00:00.000 +0200
      Summary                                      Partly Cloudy
      Precip Type                                           rain
      Temperature (C)                                  9.472222
      Apparent Temperature (C)                         7.388889
      Humidity                                             0.89
      Wind Speed (km/h)                                 14.1197
      Wind Bearing (degrees)                              251.0
      Visibility (km)                                   15.8263
      Loud Cover                                            0.0
      Pressure (millibars)                              1015.13
      Daily Summary                 Partly cloudy throughout the day.
      Name: 0, dtype: object
```

```
[20]: #show the values rom rows 10 to 20
      df.loc[10:20]
```

```
[20]:                  Formatted Date         Summary Precip Type  Temperature (C)  \
      10  2006-04-01 10:00:00.000 +0200  Partly Cloudy        rain        16.016667
      11  2006-04-01 11:00:00.000 +0200  Partly Cloudy        rain        17.144444
      12  2006-04-01 12:00:00.000 +0200  Partly Cloudy        rain        17.800000
      13  2006-04-01 13:00:00.000 +0200  Partly Cloudy        rain        17.333333
      14  2006-04-01 14:00:00.000 +0200  Partly Cloudy        rain        18.877778
      15  2006-04-01 15:00:00.000 +0200  Partly Cloudy        rain        18.911111
      16  2006-04-01 16:00:00.000 +0200  Partly Cloudy        rain        15.388889
      17  2006-04-01 17:00:00.000 +0200  Mostly Cloudy        rain        15.550000
      18  2006-04-01 18:00:00.000 +0200  Mostly Cloudy        rain        14.255556
      19  2006-04-01 19:00:00.000 +0200  Mostly Cloudy        rain        13.144444
      20  2006-04-01 20:00:00.000 +0200  Mostly Cloudy        rain        11.550000

          Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
      10                 16.016667      0.67            17.5651
      11                 17.144444      0.54            19.7869
      12                 17.800000      0.55            21.9443
      13                 17.333333      0.51            20.6885
      14                 18.877778      0.47            15.3755
      15                 18.911111      0.46            10.4006
```

```
16                15.388889      0.60          14.4095
17                15.550000      0.63          11.1573
18                14.255556      0.69           8.5169
19                13.144444      0.70           7.6314
20                11.550000      0.77           7.3899
```

```
    Wind Bearing (degrees)  Visibility (km)  Loud Cover  Pressure (millibars)  \
10                   290.0          11.2056         0.0                1017.42
11                   316.0          11.4471         0.0                1017.74
12                   281.0          11.2700         0.0                1017.59
13                   289.0          11.2700         0.0                1017.48
14                   262.0          11.4471         0.0                1017.17
15                   288.0          11.2700         0.0                1016.47
16                   251.0          11.2700         0.0                1016.15
17                   230.0          11.4471         0.0                1016.17
18                   163.0          11.2056         0.0                1015.82
19                   139.0          11.2056         0.0                1015.83
20                   147.0          11.0285         0.0                1015.85
```

```
                     Daily Summary
10  Partly cloudy throughout the day.
11  Partly cloudy throughout the day.
12  Partly cloudy throughout the day.
13  Partly cloudy throughout the day.
14  Partly cloudy throughout the day.
15  Partly cloudy throughout the day.
16  Partly cloudy throughout the day.
17  Partly cloudy throughout the day.
18  Partly cloudy throughout the day.
19  Partly cloudy throughout the day.
20  Partly cloudy throughout the day.
```

```
[21]: #show all rows and the first 4 columns o to 3
      df.iloc[:,0:4]
```

```
[21]:                   Formatted Date        Summary Precip Type  \
      0      2006-04-01 00:00:00.000 +0200  Partly Cloudy       rain
      1      2006-04-01 01:00:00.000 +0200  Partly Cloudy       rain
      2      2006-04-01 02:00:00.000 +0200  Mostly Cloudy       rain
      3      2006-04-01 03:00:00.000 +0200  Partly Cloudy       rain
      4      2006-04-01 04:00:00.000 +0200  Mostly Cloudy       rain
      ...                              ...            ...        ...
      96448  2016-09-09 19:00:00.000 +0200  Partly Cloudy       rain
      96449  2016-09-09 20:00:00.000 +0200  Partly Cloudy       rain
      96450  2016-09-09 21:00:00.000 +0200  Partly Cloudy       rain
      96451  2016-09-09 22:00:00.000 +0200  Partly Cloudy       rain
      96452  2016-09-09 23:00:00.000 +0200  Partly Cloudy       rain
```

```
        Temperature (C)
0              9.472222
1              9.355556
2              9.377778
3              8.288889
4              8.755556
...                 ...
96448         26.016667
96449         24.583333
96450         22.038889
96451         21.522222
96452         20.438889

[96453 rows x 4 columns]
```

[22]: `#show the values from rows 10 to 19 and all the columns`
`df.iloc[10:20,:]`

[22]:
```
                  Formatted Date       Summary Precip Type  Temperature (C)  \
10  2006-04-01 10:00:00.000 +0200  Partly Cloudy       rain        16.016667
11  2006-04-01 11:00:00.000 +0200  Partly Cloudy       rain        17.144444
12  2006-04-01 12:00:00.000 +0200  Partly Cloudy       rain        17.800000
13  2006-04-01 13:00:00.000 +0200  Partly Cloudy       rain        17.333333
14  2006-04-01 14:00:00.000 +0200  Partly Cloudy       rain        18.877778
15  2006-04-01 15:00:00.000 +0200  Partly Cloudy       rain        18.911111
16  2006-04-01 16:00:00.000 +0200  Partly Cloudy       rain        15.388889
17  2006-04-01 17:00:00.000 +0200  Mostly Cloudy       rain        15.550000
18  2006-04-01 18:00:00.000 +0200  Mostly Cloudy       rain        14.255556
19  2006-04-01 19:00:00.000 +0200  Mostly Cloudy       rain        13.144444


    Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
10                 16.016667      0.67            17.5651
11                 17.144444      0.54            19.7869
12                 17.800000      0.55            21.9443
13                 17.333333      0.51            20.6885
14                 18.877778      0.47            15.3755
15                 18.911111      0.46            10.4006
16                 15.388889      0.60            14.4095
17                 15.550000      0.63            11.1573
18                 14.255556      0.69             8.5169
19                 13.144444      0.70             7.6314


    Wind Bearing (degrees)  Visibility (km)  Loud Cover  Pressure (millibars)  \
10                   290.0          11.2056         0.0               1017.42
11                   316.0          11.4471         0.0               1017.74
12                   281.0          11.2700         0.0               1017.59
```

```
13                  289.0           11.2700         0.0                 1017.48
14                  262.0           11.4471         0.0                 1017.17
15                  288.0           11.2700         0.0                 1016.47
16                  251.0           11.2700         0.0                 1016.15
17                  230.0           11.4471         0.0                 1016.17
18                  163.0           11.2056         0.0                 1015.82
19                  139.0           11.2056         0.0                 1015.83


                           Daily Summary
10   Partly cloudy throughout the day.
11   Partly cloudy throughout the day.
12   Partly cloudy throughout the day.
13   Partly cloudy throughout the day.
14   Partly cloudy throughout the day.
15   Partly cloudy throughout the day.
16   Partly cloudy throughout the day.
17   Partly cloudy throughout the day.
18   Partly cloudy throughout the day.
19   Partly cloudy throughout the day.
```

[23]: `#show all the rows and all the columns except th last col`
`df.iloc[:,:-1]`

[23]:
```
                         Formatted Date           Summary Precip Type  \
0        2006-04-01 00:00:00.000 +0200  Partly Cloudy        rain
1        2006-04-01 01:00:00.000 +0200  Partly Cloudy        rain
2        2006-04-01 02:00:00.000 +0200  Mostly Cloudy        rain
3        2006-04-01 03:00:00.000 +0200  Partly Cloudy        rain
4        2006-04-01 04:00:00.000 +0200  Mostly Cloudy        rain
...                                ...            ...         ...
96448    2016-09-09 19:00:00.000 +0200  Partly Cloudy        rain
96449    2016-09-09 20:00:00.000 +0200  Partly Cloudy        rain
96450    2016-09-09 21:00:00.000 +0200  Partly Cloudy        rain
96451    2016-09-09 22:00:00.000 +0200  Partly Cloudy        rain
96452    2016-09-09 23:00:00.000 +0200  Partly Cloudy        rain


        Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
0              9.472222                  7.388889      0.89            14.1197
1              9.355556                  7.227778      0.86            14.2646
2              9.377778                  9.377778      0.89             3.9284
3              8.288889                  5.944444      0.83            14.1036
4              8.755556                  6.977778      0.83            11.0446
...                 ...                       ...       ...                ...
96448         26.016667                 26.016667      0.43            10.9963
96449         24.583333                 24.583333      0.48            10.0947
96450         22.038889                 22.038889      0.56             8.9838
96451         21.522222                 21.522222      0.60            10.5294
```

96452	20.438889	20.438889	0.61	5.8765

	Wind Bearing (degrees)	Visibility (km)	Loud Cover \
0	251.0	15.8263	0.0
1	259.0	15.8263	0.0
2	204.0	14.9569	0.0
3	269.0	15.8263	0.0
4	259.0	15.8263	0.0
...	...	...	...
96448	31.0	16.1000	0.0
96449	20.0	15.5526	0.0
96450	30.0	16.1000	0.0
96451	20.0	16.1000	0.0
96452	39.0	15.5204	0.0

	Pressure (millibars)
0	1015.13
1	1015.63
2	1015.94
3	1016.41
4	1016.51
...	...
96448	1014.36
96449	1015.16
96450	1015.66
96451	1015.95
96452	1016.16

[96453 rows x 11 columns]

USing Conditions:

```
[24]: #show me the records where humidity=1
      df[df['Humidity']==1]
```

```
[24]:
```

	Formatted Date	Summary	Precip Type \
319	2006-04-21 07:00:00.000 +0200	Foggy	rain
342	2006-04-22 06:00:00.000 +0200	Foggy	rain
390	2006-04-24 06:00:00.000 +0200	Foggy	rain
535	2006-04-03 07:00:00.000 +0200	Clear	rain
536	2006-04-03 08:00:00.000 +0200	Clear	rain
...	...	...	...
95907	2016-09-16 06:00:00.000 +0200	Foggy	rain
96123	2016-09-24 06:00:00.000 +0200	Partly Cloudy	rain
96148	2016-09-25 07:00:00.000 +0200	Mostly Cloudy	rain
96363	2016-09-06 06:00:00.000 +0200	Mostly Cloudy	rain
96364	2016-09-06 07:00:00.000 +0200	Mostly Cloudy	rain

	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h) \
319	11.972222	11.972222	1.0	9.6600
342	7.155556	7.155556	1.0	1.8515
390	7.688889	7.688889	1.0	0.0000
535	5.933333	5.933333	1.0	0.1449
536	10.405556	10.405556	1.0	0.7245
…	…	…	…	…
95907	14.905556	14.905556	1.0	7.5992
96123	7.594444	7.594444	1.0	4.8300
96148	8.838889	8.838889	1.0	4.5402
96363	12.338889	12.338889	1.0	3.5581
96364	12.872222	12.872222	1.0	4.5241

	Wind Bearing (degrees)	Visibility (km)	Loud Cover \
319	10.0	1.2558	0.0
342	319.0	0.2898	0.0
390	0.0	0.7406	0.0
535	230.0	5.8443	0.0
536	18.0	5.3935	0.0
…	…	…	…
95907	169.0	1.6100	0.0
96123	160.0	6.2951	0.0
96148	349.0	6.1985	0.0
96363	301.0	8.0500	0.0
96364	339.0	9.9820	0.0

	Pressure (millibars)	Daily Summary
319	1016.45	Foggy in the morning.
342	1017.15	Foggy starting overnight continuing until morn…
390	1018.13	Partly cloudy until evening.
535	1014.55	Partly cloudy until night.
536	1014.91	Partly cloudy until night.
…	…	…
95907	1016.51	Partly cloudy starting in the afternoon.
96123	1026.93	Partly cloudy starting in the morning.
96148	1022.88	Mostly cloudy until night.
96363	1014.45	Mostly cloudy starting overnight.
96364	1014.84	Mostly cloudy starting overnight.

[2890 rows x 12 columns]

```
[25]: # show all the columns and rows where humidity=1
      df[df['Humidity']==1]
```

```
[25]:                    Formatted Date        Summary Precip Type \
      319   2006-04-21 07:00:00.000 +0200        Foggy       rain
      342   2006-04-22 06:00:00.000 +0200        Foggy       rain
```

```
390     2006-04-24 06:00:00.000 +0200              Foggy        rain
535     2006-04-03 07:00:00.000 +0200              Clear        rain
536     2006-04-03 08:00:00.000 +0200              Clear        rain
…                               …        …        …
95907   2016-09-16 06:00:00.000 +0200              Foggy        rain
96123   2016-09-24 06:00:00.000 +0200      Partly Cloudy        rain
96148   2016-09-25 07:00:00.000 +0200      Mostly Cloudy        rain
96363   2016-09-06 06:00:00.000 +0200      Mostly Cloudy        rain
96364   2016-09-06 07:00:00.000 +0200      Mostly Cloudy        rain


        Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
319           11.972222                 11.972222       1.0             9.6600
342            7.155556                  7.155556       1.0             1.8515
390            7.688889                  7.688889       1.0             0.0000
535            5.933333                  5.933333       1.0             0.1449
536           10.405556                 10.405556       1.0             0.7245
…                   …                         …         …                  …
95907         14.905556                 14.905556       1.0             7.5992
96123          7.594444                  7.594444       1.0             4.8300
96148          8.838889                  8.838889       1.0             4.5402
96363         12.338889                 12.338889       1.0             3.5581
96364         12.872222                 12.872222       1.0             4.5241


        Wind Bearing (degrees)  Visibility (km)  Loud Cover  \
319                       10.0           1.2558         0.0
342                      319.0           0.2898         0.0
390                        0.0           0.7406         0.0
535                      230.0           5.8443         0.0
536                       18.0           5.3935         0.0
…                          …               …            …
95907                    169.0           1.6100         0.0
96123                    160.0           6.2951         0.0
96148                    349.0           6.1985         0.0
96363                    301.0           8.0500         0.0
96364                    339.0           9.9820         0.0


        Pressure (millibars)                            Daily Summary
319                  1016.45                     Foggy in the morning.
342                  1017.15   Foggy starting overnight continuing until morn…
390                  1018.13                Partly cloudy until evening.
535                  1014.55                  Partly cloudy until night.
536                  1014.91                  Partly cloudy until night.
…                        …                                          …
95907                1016.51   Partly cloudy starting in the afternoon.
96123                1026.93   Partly cloudy starting in the morning.
96148                1022.88                Mostly cloudy until night.
96363                1014.45          Mostly cloudy starting overnight.
```

```
96364                1014.84                Mostly cloudy starting overnight.

[2890 rows x 12 columns]
```

[26]:
```
# Show the no of values contains each column
df[df['Humidity']==1].count()
```

[26]:
```
Formatted Date              2890
Summary                     2890
Precip Type                 2888
Temperature (C)             2890
Apparent Temperature (C)    2890
Humidity                    2890
Wind Speed (km/h)           2890
Wind Bearing (degrees)      2890
Visibility (km)             2890
Loud Cover                  2890
Pressure (millibars)        2890
Daily Summary               2890
dtype: int64
```

[27]:
```
# show no of rows satisfies the condition and only the value of a particular
 ↪column
df[df['Humidity']==1]['Precip Type'].count()
```

[27]: 2888

[28]:
```
df[df['Humidity']==1].shape[0]
```

[28]: 2890

[29]:
```
#show the records where humidity >0.95 Wind Speed (km/h)>=8.9
df[(df['Humidity']>=0.95) & (df['Wind Speed (km/h)']>=8.9)]
```

[29]:
```
                     Formatted Date        Summary Precip Type  \
6      2006-04-01 06:00:00.000 +0200  Partly Cloudy        rain
53     2006-04-11 05:00:00.000 +0200       Overcast        rain
67     2006-04-11 19:00:00.000 +0200          Foggy        rain
69     2006-04-11 21:00:00.000 +0200       Overcast        rain
70     2006-04-11 22:00:00.000 +0200       Overcast        rain
…                                …              …           …
95991  2016-09-19 18:00:00.000 +0200       Overcast        rain
95992  2016-09-19 19:00:00.000 +0200       Overcast        rain
95995  2016-09-19 22:00:00.000 +0200       Overcast        rain
96024  2016-09-20 03:00:00.000 +0200       Overcast        rain
96386  2016-09-07 05:00:00.000 +0200  Mostly Cloudy        rain

       Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
```

```
6              7.733333             5.522222       0.95         12.3648
53            10.694444            10.694444       0.95         10.4006
67             8.800000             5.294444       0.99         26.5006
69             8.200000             4.605556       0.96         25.4219
70             7.688889             3.722222       0.96         28.1267
...                  ...                  ...        ...             ...
95991         17.688889            17.688889       0.97         11.1573
95992         17.688889            17.688889       0.97         12.1716
95995         16.550000            16.550000       0.97         13.7977
96024         14.422222            14.422222       0.96         16.8406
96386         16.705556            16.705556       0.97         13.3147

        Wind Bearing (degrees)  Visibility (km)  Loud Cover  \
6                        259.0           9.9820         0.0
53                       161.0           6.6976         0.0
67                       339.0           2.6565         0.0
69                       330.0           4.4919         0.0
70                       339.0           4.4436         0.0
...                        ...              ...         ...
95991                    342.0           8.0500         0.0
95992                    312.0           6.2951         0.0
95995                      1.0           9.8049         0.0
96024                    350.0           4.1699         0.0
96386                     21.0          14.1197         0.0

        Pressure (millibars)                       Daily Summary
6                    1016.72  Partly cloudy throughout the day.
53                   1006.59              Foggy in the evening.
67                   1004.99              Foggy in the evening.
69                   1004.96              Foggy in the evening.
70                   1005.14              Foggy in the evening.
...                      ...                                ...
95991                1004.80       Overcast throughout the day.
95992                1004.99       Overcast throughout the day.
95995                1005.54       Overcast throughout the day.
96024                1006.53  Mostly cloudy throughout the day.
96386                1021.47  Partly cloudy throughout the day.

[3726 rows x 12 columns]
```

[30]: *#show the records where humidity >0.95 Wind Speed (km/h)>=8.9 values of th*
      ↪columns Formatted Date   Summary   Precip Type
      ↪Humidity   Wind Speed (km/h)
```
df[['Formatted Date',        'Summary',        'Precip Type',
 ↪'Humidity'      ,'Wind Speed (km/h)']][(df['Humidity']>=0.95) & (df['Wind
 ↪Speed (km/h)']>=8.9)]
```

```
[30]:                      Formatted Date          Summary Precip Type  Humidity  \
       6      2006-04-01 06:00:00.000 +0200  Partly Cloudy        rain      0.95
       53     2006-04-11 05:00:00.000 +0200       Overcast        rain      0.95
       67     2006-04-11 19:00:00.000 +0200          Foggy        rain      0.99
       69     2006-04-11 21:00:00.000 +0200       Overcast        rain      0.96
       70     2006-04-11 22:00:00.000 +0200       Overcast        rain      0.96
       ...                              ...            ...         ...       ...
       95991  2016-09-19 18:00:00.000 +0200       Overcast        rain      0.97
       95992  2016-09-19 19:00:00.000 +0200       Overcast        rain      0.97
       95995  2016-09-19 22:00:00.000 +0200       Overcast        rain      0.97
       96024  2016-09-20 03:00:00.000 +0200       Overcast        rain      0.96
       96386  2016-09-07 05:00:00.000 +0200  Mostly Cloudy        rain      0.97

              Wind Speed (km/h)
       6                12.3648
       53               10.4006
       67               26.5006
       69               25.4219
       70               28.1267
       ...                  ...
       95991            11.1573
       95992            12.1716
       95995            13.7977
       96024            16.8406
       96386            13.3147

       [3726 rows x 5 columns]

[31]:  df.describe(include='object')

[31]:                       Formatted Date        Summary Precip Type  \
       count                         96453          96453        95936
       unique                        96429             27            2
       top     2010-08-02 00:00:00.000 +0200  Partly Cloudy         rain
       freq                              2          31733        85224

                             Daily Summary
       count                         96453
       unique                          214
       top     Mostly cloudy throughout the day.
       freq                          20085

[32]:  #show the no of values we have for each unique values of Summmary
       df['Summary'].value_counts()

[32]:  Summary
       Partly Cloudy                            31733
```

```
Mostly Cloudy                               28094
Overcast                                    16597
Clear                                       10890
Foggy                                        7148
Breezy and Overcast                           528
Breezy and Mostly Cloudy                      516
Breezy and Partly Cloudy                      386
Dry and Partly Cloudy                          86
Windy and Partly Cloudy                        67
Light Rain                                     63
Breezy                                         54
Windy and Overcast                             45
Humid and Mostly Cloudy                        40
Drizzle                                        39
Breezy and Foggy                               35
Windy and Mostly Cloudy                        35
Dry                                            34
Humid and Partly Cloudy                        17
Dry and Mostly Cloudy                          14
Rain                                           10
Windy                                           8
Humid and Overcast                              7
Windy and Foggy                                 4
Windy and Dry                                   1
Dangerously Windy and Partly Cloudy             1
Breezy and Dry                                  1
Name: count, dtype: int64
```

[33]: `df['Precip Type'].unique()`

[33]: `array(['rain', 'snow', nan], dtype=object)`

[34]:
```python
#show me the reocrds for the summary=Partly cloudy ,Mostly Cloudy
df[['Formatted Date',        'Summary',        'Precip Type',␣
 ↪'Humidity'         ,'Wind Speed (km/h)']][(df['Summary']=='Windy and Foggy')
                    | (df['Summary']=='Windy and␣
 ↪Dry')|(df['Summary']=='Windy')]
```

[34]:
```
              Formatted Date          Summary Precip Type  Humidity  \
4338   2006-07-08 18:00:00.000 +0200  Windy and Foggy        rain      0.76
5190   2006-03-13 06:00:00.000 +0100  Windy and Foggy        snow      0.95
20930  2008-01-28 02:00:00.000 +0100            Windy        rain      0.45
31640  2009-03-18 08:00:00.000 +0100            Windy        rain      0.60
31644  2009-03-18 12:00:00.000 +0100            Windy        rain      0.41
43944  2011-04-13 00:00:00.000 +0200            Windy        rain      0.57
53124  2012-04-29 12:00:00.000 +0200  Windy and Dry          rain      0.24
67437  2013-05-17 22:00:00.000 +0200  Windy and Foggy        rain      0.96
```

```
71942  2014-12-21 14:00:00.000 +0100               Windy      rain      0.55
82400  2015-01-04 11:00:00.000 +0100               Windy      rain      0.70
82401  2015-01-04 12:00:00.000 +0100               Windy      rain      0.65
82403  2015-01-04 14:00:00.000 +0100               Windy      rain      0.65
84861  2015-05-14 01:00:00.000 +0200    Windy and Foggy      rain      0.93


       Wind Speed (km/h)
4338             45.9333
5190             43.8564
20930            42.5362
31640            45.0800
31644            45.0800
43944            41.1999
53124            40.2500
67437            40.3144
71942            40.2983
82400            42.0854
82401            40.5398
82403            40.5076
84861            46.6095
```

[35]:
```python
#isin([<list of values>])
df[['Formatted Date',        'Summary',        'Precip Type',
 ↪'Humidity'      ,'Wind Speed (km/h)']][df['Summary'].isin(['Windy and
 ↪Foggy','Windy and Dry','Windy'])]
```

[35]:
```
                       Formatted Date          Summary Precip Type   Humidity  \
4338   2006-07-08 18:00:00.000 +0200   Windy and Foggy      rain      0.76
5190   2006-03-13 06:00:00.000 +0100   Windy and Foggy      snow      0.95
20930  2008-01-28 02:00:00.000 +0100             Windy      rain      0.45
31640  2009-03-18 08:00:00.000 +0100             Windy      rain      0.60
31644  2009-03-18 12:00:00.000 +0100             Windy      rain      0.41
43944  2011-04-13 00:00:00.000 +0200             Windy      rain      0.57
53124  2012-04-29 12:00:00.000 +0200     Windy and Dry      rain      0.24
67437  2013-05-17 22:00:00.000 +0200   Windy and Foggy      rain      0.96
71942  2014-12-21 14:00:00.000 +0100             Windy      rain      0.55
82400  2015-01-04 11:00:00.000 +0100             Windy      rain      0.70
82401  2015-01-04 12:00:00.000 +0100             Windy      rain      0.65
82403  2015-01-04 14:00:00.000 +0100             Windy      rain      0.65
84861  2015-05-14 01:00:00.000 +0200   Windy and Foggy      rain      0.93


       Wind Speed (km/h)
4338             45.9333
5190             43.8564
20930            42.5362
31640            45.0800
31644            45.0800
```

```
43944                41.1999
53124                40.2500
67437                40.3144
71942                40.2983
82400                42.0854
82401                40.5398
82403                40.5076
84861                46.6095
```

[36]: `# Show all the records where precip type not equal to rain`
`df[df['Precip Type']!='rain']`

[36]:
```
                    Formatted Date        Summary Precip Type  \
1562   2006-12-13 02:00:00.000 +0100        Foggy        snow
1563   2006-12-13 03:00:00.000 +0100        Foggy        snow
1564   2006-12-13 04:00:00.000 +0100        Foggy        snow
1565   2006-12-13 05:00:00.000 +0100        Foggy        snow
1566   2006-12-13 06:00:00.000 +0100        Foggy        snow
...                              ...          ...         ...
95584  2016-10-31 19:00:00.000 +0100  Mostly Cloudy        NaN
95585  2016-10-31 20:00:00.000 +0100  Mostly Cloudy        NaN
95586  2016-10-31 21:00:00.000 +0100  Mostly Cloudy        NaN
95587  2016-10-31 22:00:00.000 +0100  Partly Cloudy        NaN
95588  2016-10-31 23:00:00.000 +0100  Mostly Cloudy        NaN

       Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
1562         -0.483333                 -4.150000      1.00            11.0929
1563         -0.483333                 -4.061111      0.96            10.7387
1564         -0.922222                 -3.477778      1.00             7.0679
1565         -1.038889                 -4.400000      1.00             9.4990
1566         -1.088889                 -4.438889      1.00             9.4346
...                ...                       ...       ...                ...
95584         8.322222                  7.044444      0.85             8.0339
95585         7.627778                  6.183333      0.87             8.2271
95586         7.111111                  5.511111      0.89             8.5008
95587         6.672222                  4.961111      0.90             8.6457
95588         6.322222                  4.588889      0.91             8.4686

       Wind Bearing (degrees)  Visibility (km)  Loud Cover  \
1562                    219.0            0.483         0.0
1563                    200.0            0.322         0.0
1564                    206.0            0.161         0.0
1565                    199.0            0.161         0.0
1566                    219.0            0.322         0.0
...                      ...              ...          ...
95584                  290.0            0.000         0.0
95585                  293.0            0.000         0.0
```

```
95586                     297.0            0.000            0.0
95587                     299.0            0.000            0.0
95588                     299.0            0.000            0.0

        Pressure (millibars)                             Daily Summary
1562                 1031.56                 Foggy throughout the day.
1563                 1031.47                 Foggy throughout the day.
1564                 1031.23                 Foggy throughout the day.
1565                 1031.41                 Foggy throughout the day.
1566                 1031.98                 Foggy throughout the day.
...                      ...                                       ...
95584                1021.73  Mostly cloudy starting in the afternoon.
95585                1021.76  Mostly cloudy starting in the afternoon.
95586                1021.81  Mostly cloudy starting in the afternoon.
95587                1021.83  Mostly cloudy starting in the afternoon.
95588                1021.80  Mostly cloudy starting in the afternoon.

[11229 rows x 12 columns]
```

Using Mathematical Finctions

```python
[37]:  #show the max value of each column
       df.max(numeric_only=True)
```

```
[37]:  Temperature (C)                39.905556
       Apparent Temperature (C)       39.344444
       Humidity                        1.000000
       Wind Speed (km/h)              63.852600
       Wind Bearing (degrees)        359.000000
       Visibility (km)                16.100000
       Loud Cover                      0.000000
       Pressure (millibars)         1046.380000
       dtype: float64
```

```python
[38]:  #max of of particular column
       df['Humidity'].min()
```

```
[38]:  0.0
```

```python
[39]:  df['Humidity'].max()
```

```
[39]:  1.0
```

```python
[40]:  df['Humidity'].mean()
```

```
[40]:  0.7348989663359355
```

Data Cleaning and Manipulation

```
[41]: #Create a new df with values>0
      df_clean=df[(df['Humidity']>0) & (df['Wind Speed (km/h)']>0)& (df['Temperature␣
       ↪(C)']>0)& (df['Apparent Temperature (C)']>0) & (df['Wind Bearing␣
       ↪(degrees)']>0)]
```

```
[42]: df_clean.describe()
```

```
[42]:        Temperature (C)  Apparent Temperature (C)      Humidity  \
      count     76534.000000              76534.000000  76534.000000
      mean         14.985860                 14.406166      0.706184
      std           7.796182                  8.448812      0.200611
      min           0.005556                  0.005556      0.120000
      25%           8.611111                  6.933333      0.550000
      50%          14.422222                 14.422222      0.740000
      75%          20.458333                 20.458333      0.870000
      max          39.905556                 39.344444      1.000000

             Wind Speed (km/h)  Wind Bearing (degrees)  Visibility (km)  Loud Cover  \
      count       76534.000000            76534.000000     76534.000000     76534.0
      mean           10.776869              193.989913        10.988979         0.0
      std             6.784777              104.542308         3.775415         0.0
      min             0.032200                1.000000         0.000000         0.0
      25%             5.779900              128.000000         9.982000         0.0
      50%             9.837100              187.000000        10.352300         0.0
      75%            14.087500              292.000000        14.956900         0.0
      max            63.852600              359.000000        16.100000         0.0

             Pressure (millibars)
      count          76534.000000
      mean            1003.154809
      std              112.614199
      min                0.000000
      25%             1011.510000
      50%             1015.710000
      75%             1019.800000
      max             1042.380000
```

```
[43]: #drop the column loud Cover
      df_clean.drop(columns=['Loud Cover'],axis=1) #axis=1 when we want ot drop␣
       ↪column by default axis=0 i.e rows
```

```
[43]:                       Formatted Date          Summary Precip Type  \
      0      2006-04-01 00:00:00.000 +0200  Partly Cloudy        rain
      1      2006-04-01 01:00:00.000 +0200  Partly Cloudy        rain
      2      2006-04-01 02:00:00.000 +0200  Mostly Cloudy        rain
      3      2006-04-01 03:00:00.000 +0200  Partly Cloudy        rain
      4      2006-04-01 04:00:00.000 +0200  Mostly Cloudy        rain
```

```
…                        …           …          …
96448  2016-09-09 19:00:00.000 +0200  Partly Cloudy        rain
96449  2016-09-09 20:00:00.000 +0200  Partly Cloudy        rain
96450  2016-09-09 21:00:00.000 +0200  Partly Cloudy        rain
96451  2016-09-09 22:00:00.000 +0200  Partly Cloudy        rain
96452  2016-09-09 23:00:00.000 +0200  Partly Cloudy        rain


       Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
0            9.472222                  7.388889      0.89            14.1197
1            9.355556                  7.227778      0.86            14.2646
2            9.377778                  9.377778      0.89             3.9284
3            8.288889                  5.944444      0.83            14.1036
4            8.755556                  6.977778      0.83            11.0446
…                 …                        …          …                  …
96448       26.016667                 26.016667      0.43            10.9963
96449       24.583333                 24.583333      0.48            10.0947
96450       22.038889                 22.038889      0.56             8.9838
96451       21.522222                 21.522222      0.60            10.5294
96452       20.438889                 20.438889      0.61             5.8765


       Wind Bearing (degrees)  Visibility (km)  Pressure (millibars)  \
0                       251.0          15.8263              1015.13
1                       259.0          15.8263              1015.63
2                       204.0          14.9569              1015.94
3                       269.0          15.8263              1016.41
4                       259.0          15.8263              1016.51
…                          …               …                    …
96448                    31.0          16.1000              1014.36
96449                    20.0          15.5526              1015.16
96450                    30.0          16.1000              1015.66
96451                    20.0          16.1000              1015.95
96452                    39.0          15.5204              1016.16


                             Daily Summary
0              Partly cloudy throughout the day.
1              Partly cloudy throughout the day.
2              Partly cloudy throughout the day.
3              Partly cloudy throughout the day.
4              Partly cloudy throughout the day.
…                                 …
96448  Partly cloudy starting in the morning.
96449  Partly cloudy starting in the morning.
96450  Partly cloudy starting in the morning.
96451  Partly cloudy starting in the morning.
96452  Partly cloudy starting in the morning.

[76534 rows x 11 columns]
```

```
[44]: df_clean.head()
```

```
[44]:              Formatted Date          Summary Precip Type  Temperature (C)  \
      0  2006-04-01 00:00:00.000 +0200  Partly Cloudy       rain         9.472222
      1  2006-04-01 01:00:00.000 +0200  Partly Cloudy       rain         9.355556
      2  2006-04-01 02:00:00.000 +0200  Mostly Cloudy       rain         9.377778
      3  2006-04-01 03:00:00.000 +0200  Partly Cloudy       rain         8.288889
      4  2006-04-01 04:00:00.000 +0200  Mostly Cloudy       rain         8.755556


         Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
      0                  7.388889      0.89            14.1197
      1                  7.227778      0.86            14.2646
      2                  9.377778      0.89             3.9284
      3                  5.944444      0.83            14.1036
      4                  6.977778      0.83            11.0446


         Wind Bearing (degrees)  Visibility (km)  Loud Cover  Pressure (millibars)  \
      0                   251.0          15.8263         0.0               1015.13
      1                   259.0          15.8263         0.0               1015.63
      2                   204.0          14.9569         0.0               1015.94
      3                   269.0          15.8263         0.0               1016.41
      4                   259.0          15.8263         0.0               1016.51


                            Daily Summary
      0  Partly cloudy throughout the day.
      1  Partly cloudy throughout the day.
      2  Partly cloudy throughout the day.
      3  Partly cloudy throughout the day.
      4  Partly cloudy throughout the day.
```

```
[45]: #drop the column loud Cover and update the dataFrame
      df_clean.drop(columns=['Loud Cover']) #axis=1 when we want ot drop column by␣
       ↪default axis=0 i.e rows
```

```
[45]:                     Formatted Date          Summary Precip Type  \
      0         2006-04-01 00:00:00.000 +0200  Partly Cloudy       rain
      1         2006-04-01 01:00:00.000 +0200  Partly Cloudy       rain
      2         2006-04-01 02:00:00.000 +0200  Mostly Cloudy       rain
      3         2006-04-01 03:00:00.000 +0200  Partly Cloudy       rain
      4         2006-04-01 04:00:00.000 +0200  Mostly Cloudy       rain
      ...                                 ...            ...        ...
      96448     2016-09-09 19:00:00.000 +0200  Partly Cloudy       rain
      96449     2016-09-09 20:00:00.000 +0200  Partly Cloudy       rain
      96450     2016-09-09 21:00:00.000 +0200  Partly Cloudy       rain
      96451     2016-09-09 22:00:00.000 +0200  Partly Cloudy       rain
      96452     2016-09-09 23:00:00.000 +0200  Partly Cloudy       rain
```

```
       Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
0             9.472222                  7.388889      0.89            14.1197
1             9.355556                  7.227778      0.86            14.2646
2             9.377778                  9.377778      0.89             3.9284
3             8.288889                  5.944444      0.83            14.1036
4             8.755556                  6.977778      0.83            11.0446
...                ...                       ...       ...                ...
96448        26.016667                 26.016667      0.43            10.9963
96449        24.583333                 24.583333      0.48            10.0947
96450        22.038889                 22.038889      0.56             8.9838
96451        21.522222                 21.522222      0.60            10.5294
96452        20.438889                 20.438889      0.61             5.8765

       Wind Bearing (degrees)  Visibility (km)  Pressure (millibars)  \
0                       251.0          15.8263               1015.13
1                       259.0          15.8263               1015.63
2                       204.0          14.9569               1015.94
3                       269.0          15.8263               1016.41
4                       259.0          15.8263               1016.51
...                       ...              ...                   ...
96448                    31.0          16.1000               1014.36
96449                    20.0          15.5526               1015.16
96450                    30.0          16.1000               1015.66
96451                    20.0          16.1000               1015.95
96452                    39.0          15.5204               1016.16

                                      Daily Summary
0              Partly cloudy throughout the day.
1              Partly cloudy throughout the day.
2              Partly cloudy throughout the day.
3              Partly cloudy throughout the day.
4              Partly cloudy throughout the day.
...                                          ...
96448    Partly cloudy starting in the morning.
96449    Partly cloudy starting in the morning.
96450    Partly cloudy starting in the morning.
96451    Partly cloudy starting in the morning.
96452    Partly cloudy starting in the morning.

[76534 rows x 11 columns]
```

[46]: `df.drop(columns=['Loud Cover'],axis=1,inplace=True)`

[47]: `df`

[47]:
```
                    Formatted Date       Summary Precip Type  \
0     2006-04-01 00:00:00.000 +0200  Partly Cloudy        rain
```

```
1      2006-04-01 01:00:00.000 +0200  Partly Cloudy       rain
2      2006-04-01 02:00:00.000 +0200  Mostly Cloudy       rain
3      2006-04-01 03:00:00.000 +0200  Partly Cloudy       rain
4      2006-04-01 04:00:00.000 +0200  Mostly Cloudy       rain
…                               …            …           …
96448  2016-09-09 19:00:00.000 +0200  Partly Cloudy       rain
96449  2016-09-09 20:00:00.000 +0200  Partly Cloudy       rain
96450  2016-09-09 21:00:00.000 +0200  Partly Cloudy       rain
96451  2016-09-09 22:00:00.000 +0200  Partly Cloudy       rain
96452  2016-09-09 23:00:00.000 +0200  Partly Cloudy       rain


       Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
0           9.472222                  7.388889      0.89             14.1197
1           9.355556                  7.227778      0.86             14.2646
2           9.377778                  9.377778      0.89              3.9284
3           8.288889                  5.944444      0.83             14.1036
4           8.755556                  6.977778      0.83             11.0446
…              …                         …           …                 …
96448      26.016667                 26.016667      0.43             10.9963
96449      24.583333                 24.583333      0.48             10.0947
96450      22.038889                 22.038889      0.56              8.9838
96451      21.522222                 21.522222      0.60             10.5294
96452      20.438889                 20.438889      0.61              5.8765


       Wind Bearing (degrees)  Visibility (km)  Pressure (millibars)  \
0                       251.0          15.8263               1015.13
1                       259.0          15.8263               1015.63
2                       204.0          14.9569               1015.94
3                       269.0          15.8263               1016.41
4                       259.0          15.8263               1016.51
…                         …              …                      …
96448                    31.0          16.1000               1014.36
96449                    20.0          15.5526               1015.16
96450                    30.0          16.1000               1015.66
96451                    20.0          16.1000               1015.95
96452                    39.0          15.5204               1016.16


                              Daily Summary
0           Partly cloudy throughout the day.
1           Partly cloudy throughout the day.
2           Partly cloudy throughout the day.
3           Partly cloudy throughout the day.
4           Partly cloudy throughout the day.
…                            …
96448  Partly cloudy starting in the morning.
96449  Partly cloudy starting in the morning.
96450  Partly cloudy starting in the morning.
```

```
96451  Partly cloudy starting in the morning.
96452  Partly cloudy starting in the morning.

[96453 rows x 11 columns]
```

[48]: `df_clean`

[48]:
```
                      Formatted Date          Summary Precip Type  \
0      2006-04-01 00:00:00.000 +0200  Partly Cloudy         rain
1      2006-04-01 01:00:00.000 +0200  Partly Cloudy         rain
2      2006-04-01 02:00:00.000 +0200  Mostly Cloudy         rain
3      2006-04-01 03:00:00.000 +0200  Partly Cloudy         rain
4      2006-04-01 04:00:00.000 +0200  Mostly Cloudy         rain
...                              ...            ...          ...
96448  2016-09-09 19:00:00.000 +0200  Partly Cloudy         rain
96449  2016-09-09 20:00:00.000 +0200  Partly Cloudy         rain
96450  2016-09-09 21:00:00.000 +0200  Partly Cloudy         rain
96451  2016-09-09 22:00:00.000 +0200  Partly Cloudy         rain
96452  2016-09-09 23:00:00.000 +0200  Partly Cloudy         rain

       Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
0             9.472222                  7.388889      0.89            14.1197
1             9.355556                  7.227778      0.86            14.2646
2             9.377778                  9.377778      0.89             3.9284
3             8.288889                  5.944444      0.83            14.1036
4             8.755556                  6.977778      0.83            11.0446
...                ...                       ...       ...                ...
96448        26.016667                 26.016667      0.43            10.9963
96449        24.583333                 24.583333      0.48            10.0947
96450        22.038889                 22.038889      0.56             8.9838
96451        21.522222                 21.522222      0.60            10.5294
96452        20.438889                 20.438889      0.61             5.8765

       Wind Bearing (degrees)  Visibility (km)  Loud Cover  \
0                       251.0          15.8263         0.0
1                       259.0          15.8263         0.0
2                       204.0          14.9569         0.0
3                       269.0          15.8263         0.0
4                       259.0          15.8263         0.0
...                       ...              ...         ...
96448                    31.0          16.1000         0.0
96449                    20.0          15.5526         0.0
96450                    30.0          16.1000         0.0
96451                    20.0          16.1000         0.0
96452                    39.0          15.5204         0.0

       Pressure (millibars)                       Daily Summary
```

```
0                    1015.13        Partly cloudy throughout the day.
1                    1015.63        Partly cloudy throughout the day.
2                    1015.94        Partly cloudy throughout the day.
3                    1016.41        Partly cloudy throughout the day.
4                    1016.51        Partly cloudy throughout the day.
...                     ...                               ...
96448                1014.36  Partly cloudy starting in the morning.
96449                1015.16  Partly cloudy starting in the morning.
96450                1015.66  Partly cloudy starting in the morning.
96451                1015.95  Partly cloudy starting in the morning.
96452                1016.16  Partly cloudy starting in the morning.

[76534 rows x 12 columns]
```

[49]: ```python
#check null values
df.isna() #returns True (Nan) False(Not NaN)
```

[49]:
```
       Formatted Date Summary Precip Type Temperature (C) \
0   False   False   False   False
1   False   False   False   False
2   False   False   False   False
3   False   False   False   False
4   False   False   False   False
...   .
96448   False   False   False   False
96449   False   False   False   False
96450   False   False   False   False
96451   False   False   False   False
96452   False   False   False   False

       Apparent Temperature (C) Humidity Wind Speed (km/h) \
0   False   False   False
1   False   False   False
2   False   False   False
3   False   False   False
4   False   False   False
...   .
96448   False   False   False
96449   False   False   False
96450   False   False   False
96451   False   False   False
96452   False   False   False

       Wind Bearing (degrees) Visibility (km) Pressure (millibars) \
0   False   False   False
1   False   False   False
2   False   False   False
```

```
3                              False        False                False
4                              False        False                False
…                              …            …             …
96448                          False        False                False
96449                          False        False                False
96450                          False        False                False
96451                          False        False                False
96452                          False        False                False


       Daily Summary
0               False
1               False
2               False
3               False
4               False
…               …
96448           False
96449           False
96450           False
96451           False
96452           False

[96453 rows x 11 columns]
```

[50]: `df.isna().sum()`

```
[50]: Formatted Date                 0
      Summary                        0
      Precip Type                  517
      Temperature (C)                0
      Apparent Temperature (C)       0
      Humidity                       0
      Wind Speed (km/h)              0
      Wind Bearing (degrees)         0
      Visibility (km)                0
      Pressure (millibars)           0
      Daily Summary                  0
      dtype: int64
```

[51]: ```python
#show me rows of all the nan values of the columns Precip Type
df[df['Precip Type'].isna()==True]
```

```
[51]:                   Formatted Date       Summary Precip Type  \
      52672  2012-04-11 16:00:00.000 +0200  Mostly Cloudy         NaN
      52674  2012-04-11 18:00:00.000 +0200  Mostly Cloudy         NaN
      52675  2012-04-11 19:00:00.000 +0200  Mostly Cloudy         NaN
      52677  2012-04-11 21:00:00.000 +0200  Mostly Cloudy         NaN
```

```
52678   2012-04-11 22:00:00.000 +0200   Mostly Cloudy           NaN
...                             ...         ...         ...
95584   2016-10-31 19:00:00.000 +0100   Mostly Cloudy           NaN
95585   2016-10-31 20:00:00.000 +0100   Mostly Cloudy           NaN
95586   2016-10-31 21:00:00.000 +0100   Mostly Cloudy           NaN
95587   2016-10-31 22:00:00.000 +0100   Partly Cloudy           NaN
95588   2016-10-31 23:00:00.000 +0100   Mostly Cloudy           NaN


        Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
52672         19.016667                 19.016667      0.26            14.8764
52674         17.850000                 17.850000      0.28            13.7977
52675         16.322222                 16.322222      0.32            10.8192
52677         12.566667                 12.566667      0.43             9.0160
52678         12.927778                 12.927778      0.47            17.6295
...                 ...                       ...       ...                ...
95584          8.322222                  7.044444      0.85             8.0339
95585          7.627778                  6.183333      0.87             8.2271
95586          7.111111                  5.511111      0.89             8.5008
95587          6.672222                  4.961111      0.90             8.6457
95588          6.322222                  4.588889      0.91             8.4686


        Wind Bearing (degrees)  Visibility (km)  Pressure (millibars)  \
52672                    163.0            9.982               1002.40
52674                    169.0            9.982               1001.79
52675                    151.0            9.982               1001.60
52677                    159.0            9.982               1001.92
52678                    197.0           16.100               1002.20
...                        ...              ...                   ...
95584                    290.0            0.000               1021.73
95585                    293.0            0.000               1021.76
95586                    297.0            0.000               1021.81
95587                    299.0            0.000               1021.83
95588                    299.0            0.000               1021.80


                                    Daily Summary
52672                 Mostly cloudy until night.
52674                 Mostly cloudy until night.
52675                 Mostly cloudy until night.
52677                 Mostly cloudy until night.
52678                 Mostly cloudy until night.
...                                           ...
95584  Mostly cloudy starting in the afternoon.
95585  Mostly cloudy starting in the afternoon.
95586  Mostly cloudy starting in the afternoon.
95587  Mostly cloudy starting in the afternoon.
95588  Mostly cloudy starting in the afternoon.
```

```
[517 rows x 11 columns]
```

```
[52]: df_clean1=df[df['Precip Type'].isna()==False]
```

```
[53]: #drop all the null values
      df.dropna()
```

```
[53]:                        Formatted Date          Summary Precip Type  \
      0      2006-04-01 00:00:00.000 +0200    Partly Cloudy        rain
      1      2006-04-01 01:00:00.000 +0200    Partly Cloudy        rain
      2      2006-04-01 02:00:00.000 +0200    Mostly Cloudy        rain
      3      2006-04-01 03:00:00.000 +0200    Partly Cloudy        rain
      4      2006-04-01 04:00:00.000 +0200    Mostly Cloudy        rain
      ...                              ...              ...         ...
      96448  2016-09-09 19:00:00.000 +0200    Partly Cloudy        rain
      96449  2016-09-09 20:00:00.000 +0200    Partly Cloudy        rain
      96450  2016-09-09 21:00:00.000 +0200    Partly Cloudy        rain
      96451  2016-09-09 22:00:00.000 +0200    Partly Cloudy        rain
      96452  2016-09-09 23:00:00.000 +0200    Partly Cloudy        rain

             Temperature (C)  Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
      0             9.472222                  7.388889      0.89            14.1197
      1             9.355556                  7.227778      0.86            14.2646
      2             9.377778                  9.377778      0.89             3.9284
      3             8.288889                  5.944444      0.83            14.1036
      4             8.755556                  6.977778      0.83            11.0446
      ...                ...                       ...       ...                ...
      96448        26.016667                 26.016667      0.43            10.9963
      96449        24.583333                 24.583333      0.48            10.0947
      96450        22.038889                 22.038889      0.56             8.9838
      96451        21.522222                 21.522222      0.60            10.5294
      96452        20.438889                 20.438889      0.61             5.8765

             Wind Bearing (degrees)  Visibility (km)  Pressure (millibars)  \
      0                       251.0          15.8263               1015.13
      1                       259.0          15.8263               1015.63
      2                       204.0          14.9569               1015.94
      3                       269.0          15.8263               1016.41
      4                       259.0          15.8263               1016.51
      ...                       ...              ...                   ...
      96448                    31.0          16.1000               1014.36
      96449                    20.0          15.5526               1015.16
      96450                    30.0          16.1000               1015.66
      96451                    20.0          16.1000               1015.95
      96452                    39.0          15.5204               1016.16

                                Daily Summary
```

```
0                 Partly cloudy throughout the day.
1                 Partly cloudy throughout the day.
2                 Partly cloudy throughout the day.
3                 Partly cloudy throughout the day.
4                 Partly cloudy throughout the day.
...                                             ...
96448   Partly cloudy starting in the morning.
96449   Partly cloudy starting in the morning.
96450   Partly cloudy starting in the morning.
96451   Partly cloudy starting in the morning.
96452   Partly cloudy starting in the morning.

[95936 rows x 11 columns]
```

[54]:
```
#duplicate rows
df.duplicated().sum()
```

[54]: 24

[55]:
```
df.drop_duplicates(keep='first',inplace=True)
```

[56]:
```
df_clean1.isna().sum()
```

[56]:
```
Formatted Date              0
Summary                     0
Precip Type                 0
Temperature (C)             0
Apparent Temperature (C)    0
Humidity                    0
Wind Speed (km/h)           0
Wind Bearing (degrees)      0
Visibility (km)             0
Pressure (millibars)        0
Daily Summary               0
dtype: int64
```

Conclusion:

The data above is completely transformed and ready for further analysis