

Working on Real Time Project with Python

Weather Dataset

A big dataset has daily weather info from lots of major cities. At first, it only had capitals, but now it covers more cities and even has hourly data. There are about 1250 cities included. Some places have weather records going way back to January 2, 1833. This helps users see how weather patterns have changed over a long time.

```
In [1]: import pandas as pd  
import warnings  
warnings.filterwarnings("ignore")
```

```
In [2]: df = pd.read_csv(r'C:\Users\LENOVO\Desktop\Self Prep\Python DA Project-1\weatherHistor
```

```
In [3]: df
```

Out[3]:

		Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0		2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.0
1		2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.0
2		2006-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204.0	14.0
3		2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269.0	15.0
4		2006-04-01 04:00:00.000 +0200	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259.0	15.0
...
96448		2016-09-09 19:00:00.000 +0200	Partly Cloudy	rain	26.016667	26.016667	0.43	10.9963	31.0	16.0
96449		2016-09-09 20:00:00.000 +0200	Partly Cloudy	rain	24.583333	24.583333	0.48	10.0947	20.0	15.0
96450		2016-09-09 21:00:00.000 +0200	Partly Cloudy	rain	22.038889	22.038889	0.56	8.9838	30.0	16.0
96451		2016-09-09 22:00:00.000 +0200	Partly Cloudy	rain	21.522222	21.522222	0.60	10.5294	20.0	16.0
96452		2016-09-09 23:00:00.000 +0200	Partly Cloudy	rain	20.438889	20.438889	0.61	5.8765	39.0	15.0

96453 rows × 12 columns

Analyzing DataFrames

- `.head()` - Shows first N rows in the data. By default N=5

In [4]: `df.head()`

Out[4]:

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0	2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.8263
1	2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.8263
2	2006-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204.0	14.9569
3	2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269.0	15.8263
4	2006-04-01 04:00:00.000 +0200	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259.0	15.8263



- `.shape` - Shows total no of rows & columns in Dataframe

In [5]: `df.shape`

Out[5]: `(96453, 12)`

- `.index` - provide index of dataframe

In [6]: `df.index`

Out[6]: `RangeIndex(start=0, stop=96453, step=1)`

- `.columns` - shows name of each column

```
In [7]: df.columns
```

```
Out[7]: 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')
```

- `.dtypes` - Shows data-type of each column

```
In [8]: df.dtypes
```

```
Out[8]: 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
```

- `.unique()` - shows unique values in column. Apply only at single column

```
In [9]: df['Summary'].unique()
```

```
Out[9]: 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)
```

- `.nunique()` - Shows total count of unique values. Apply on single column or whole DataFrame

```
In [10]: df.nunique()
```

```
Out[10]: Formatted Date      96429  
Summary          27  
Precip Type      2  
Temperature (C)   7574  
Apparent Temperature (C) 8984  
Humidity         90  
Wind Speed (km/h) 2484  
Wind Bearing (degrees) 360  
Visibility (km)    949  
Loud Cover        1  
Pressure (millibars) 4979  
Daily Summary     214  
dtype: int64
```

- `.count()` - shows total no of non-null values in each column. Can be applied on single column as well as whole dataframe

```
In [11]: df.count()
```

```
Out[11]: Formatted Date      96453  
Summary          96453  
Precip Type      95936  
Temperature (C)   96453  
Apparent Temperature (C) 96453  
Humidity         96453  
Wind Speed (km/h) 96453  
Wind Bearing (degrees) 96453  
Visibility (km)    96453  
Loud Cover        96453  
Pressure (millibars) 96453  
Daily Summary     96453  
dtype: int64
```

- `.value_counts` - shows the unique values with their count. Applied on single column only

```
In [12]: df.value_counts
```

```

Out[12]: <bound method DataFrame.value_counts of
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.

[96453 rows x 12 columns]>

```

- .info - Provides basic information about the dataframe

In [13]: df.info

```

Out[13]: <bound method DataFrame.info of
          recip 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.

[96453 rows x 12 columns]>

```

Finding the unique wind speed in the dataset

```
In [14]: df.head()
```

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0	2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.8263
1	2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.8263
2	2006-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204.0	14.9569
3	2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269.0	15.8263
4	2006-04-01 04:00:00.000 +0200	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259.0	15.8263

- .nunique() - Shows total no of unique values. Can be applied on single column or whole Dataframe

```
In [15]: df.nunique()
```

```
Out[15]: Formatted Date      96429
Summary          27
Precip Type       2
Temperature (C)   7574
Apparent Temperature (C)  8984
Humidity          90
Wind Speed (km/h) 2484
Wind Bearing (degrees) 360
Visibility (km)    949
Loud Cover         1
Pressure (millibars) 4979
Daily Summary      214
dtype: int64
```

```
In [16]: df['Wind Speed (km/h)'].nunique()
```

```
Out[16]: 2484
```

- .unique() - Shows unique values in columns. Work on single column only

```
In [17]: df['Wind Speed (km/h)'].unique()
```

```
Out[17]: array([14.1197, 14.2646, 3.9284, ..., 37.0622, 35.5971, 30.751 ])
```

Analyzing where the weather is rainy

```
In [18]: df.head(2)
```

```
Out[18]:
```

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0	2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.8263
1	2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.8263



- Value_count() - shows no of common counts (yes / no / neutral)

```
In [19]: df['Precip Type'].value_counts()
```

```
Out[19]: rain    85224  
snow    10712  
Name: Precip Type, dtype: int64
```

- Filtering - Select specific rows based on given conditions

```
In [20]: df[df['Precip Type'] == 'rain']
```

Out[20]:

		Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0		2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.0
1		2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.0
2		2006-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204.0	14.0
3		2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269.0	15.0
4		2006-04-01 04:00:00.000 +0200	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259.0	15.0
...
96448		2016-09-09 19:00:00.000 +0200	Partly Cloudy	rain	26.016667	26.016667	0.43	10.9963	31.0	16.0
96449		2016-09-09 20:00:00.000 +0200	Partly Cloudy	rain	24.583333	24.583333	0.48	10.0947	20.0	15.0
96450		2016-09-09 21:00:00.000 +0200	Partly Cloudy	rain	22.038889	22.038889	0.56	8.9838	30.0	16.0
96451		2016-09-09 22:00:00.000 +0200	Partly Cloudy	rain	21.522222	21.522222	0.60	10.5294	20.0	16.0
96452		2016-09-09 23:00:00.000 +0200	Partly Cloudy	rain	20.438889	20.438889	0.61	5.8765	39.0	15.0

85224 rows × 12 columns

- Groupby () - Split DataFrame based on specified column

```
In [21]: df.groupby('Summary').get_group('Partly Cloudy')
```

Out[21]:

		Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0		2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.0
1		2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.0
3		2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269.0	15.0
5		2006-04-01 05:00:00.000 +0200	Partly Cloudy	rain	9.222222	7.111111	0.85	13.9587	258.0	14.0
6		2006-04-01 06:00:00.000 +0200	Partly Cloudy	rain	7.733333	5.522222	0.95	12.3648	259.0	9.0
...
96448		2016-09-09 19:00:00.000 +0200	Partly Cloudy	rain	26.016667	26.016667	0.43	10.9963	31.0	16.0
96449		2016-09-09 20:00:00.000 +0200	Partly Cloudy	rain	24.583333	24.583333	0.48	10.0947	20.0	15.0
96450		2016-09-09 21:00:00.000 +0200	Partly Cloudy	rain	22.038889	22.038889	0.56	8.9838	30.0	16.0
96451		2016-09-09 22:00:00.000 +0200	Partly Cloudy	rain	21.522222	21.522222	0.60	10.5294	20.0	16.0
96452		2016-09-09 23:00:00.000 +0200	Partly Cloudy	rain	20.438889	20.438889	0.61	5.8765	39.0	15.0

31733 rows × 12 columns

Finding the Wind Speed when it is exactly 5.8765

```
In [22]: df[df['Wind Speed (km/h)'] == 5.8765]
```

Out[22]:

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visib (
1801	2006-12-22 01:00:00.000 +0100	Mostly Cloudy	rain	0.683333	-1.172222	0.92	5.8765	311.0	3.7
1949	2006-12-28 05:00:00.000 +0100	Foggy	snow	-7.622222	-10.705556	0.92	5.8765	188.0	0.1
3334	2006-01-26 22:00:00.000 +0100	Partly Cloudy	snow	-8.944444	-12.216667	0.78	5.8765	81.0	14.7
4121	2006-07-28 17:00:00.000 +0200	Partly Cloudy	rain	32.877778	32.155556	0.32	5.8765	338.0	10.3
4419	2006-06-11 03:00:00.000 +0200	Foggy	rain	6.861111	5.916667	1.00	5.8765	291.0	2.0
5308	2006-03-18 04:00:00.000 +0100	Mostly Cloudy	rain	1.205556	-0.577778	0.92	5.8765	335.0	7.7
6215	2006-05-24 00:00:00.000 +0200	Partly Cloudy	rain	19.972222	19.972222	0.73	5.8765	36.0	15.8
7906	2006-10-04 10:00:00.000 +0200	Partly Cloudy	rain	21.088889	21.088889	0.59	5.8765	324.0	9.9
8981	2007-04-17 05:00:00.000 +0200	Partly Cloudy	rain	1.750000	0.055556	0.85	5.8765	43.0	14.9
9173	2007-04-24 05:00:00.000 +0200	Partly Cloudy	rain	4.838889	3.600000	0.77	5.8765	172.0	15.3
12891	2007-07-28 03:00:00.000 +0200	Partly Cloudy	rain	18.361111	18.361111	0.55	5.8765	339.0	15.8

		Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visib (
13673		2007-06-29 17:00:00.000 +0200	Mostly Cloudy	rain	25.116667	25.116667	0.30	5.8765	247.0	10.3
16223		2007-10-15 00:00:00.000 +0200	Clear	snow	-0.044444	-2.005556	0.82	5.8765	239.0	9.9
17546		2008-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	5.827778	4.727778	0.78	5.8765	15.0	14.1
19799		2008-02-10 23:00:00.000 +0100	Partly Cloudy	snow	-1.494444	-3.672222	0.85	5.8765	146.0	7.4
19846		2008-02-12 22:00:00.000 +0100	Clear	snow	-4.261111	-6.850000	0.82	5.8765	321.0	9.3
21623		2008-07-25 23:00:00.000 +0200	Partly Cloudy	rain	18.655556	18.655556	0.88	5.8765	1.0	14.9
21944		2008-06-01 08:00:00.000 +0200	Mostly Cloudy	rain	21.844444	21.844444	0.66	5.8765	265.0	9.9
26294		2008-09-08 14:00:00.000 +0200	Mostly Cloudy	rain	19.205556	19.205556	0.88	5.8765	317.0	11.4
30935		2009-06-18 23:00:00.000 +0200	Clear	rain	18.022222	18.022222	0.64	5.8765	206.0	15.1
35158		2010-04-11 22:00:00.000 +0200	Overcast	rain	8.222222	7.472222	0.81	5.8765	155.0	11.9
35417		2010-04-21 17:00:00.000 +0200	Mostly Cloudy	rain	18.800000	18.800000	0.42	5.8765	330.0	10.3
36480		2010-08-07 00:00:00.000 +0200	Partly Cloudy	rain	16.672222	16.672222	1.00	5.8765	238.0	15.6
38407		2010-01-26 07:00:00.000 +0100	Overcast	snow	-7.950000	-11.083333	0.73	5.8765	91.0	9.7

		Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visib (
39675		2010-06-18 03:00:00.000 +0200	Mostly Cloudy	rain	17.244444	17.244444	0.93	5.8765	0.0	8.1
41331		2010-05-25 04:00:00.000 +0200	Mostly Cloudy	rain	15.950000	15.950000	0.88	5.8765	21.0	14.9
41641		2010-05-09 02:00:00.000 +0200	Partly Cloudy	rain	11.300000	11.300000	0.96	5.8765	140.0	8.4
44015		2011-04-15 23:00:00.000 +0200	Overcast	rain	8.900000	8.250000	0.79	5.8765	1.0	11.6
44018		2011-04-16 02:00:00.000 +0200	Partly Cloudy	rain	4.655556	3.383333	0.89	5.8765	16.0	14.1
46680		2011-02-08 00:00:00.000 +0100	Foggy	snow	-1.622222	-3.822222	1.00	5.8765	209.0	1.0
47978		2011-07-29 02:00:00.000 +0200	Mostly Cloudy	rain	17.277778	17.277778	0.99	5.8765	276.0	7.9
50787		2011-11-23 04:00:00.000 +0100	Overcast	snow	-2.116667	-4.388889	0.89	5.8765	35.0	4.9
50928		2011-11-29 01:00:00.000 +0100	Foggy	snow	-2.500000	-4.827778	0.96	5.8765	37.0	0.1
52442		2011-09-30 02:00:00.000 +0200	Clear	rain	10.861111	10.861111	0.80	5.8765	337.0	14.9
57817		2012-03-13 01:00:00.000 +0100	Mostly Cloudy	rain	3.227778	1.750000	0.57	5.8765	322.0	15.5
59397		2012-11-16 22:00:00.000 +0100	Foggy	rain	2.722222	1.166667	1.00	5.8765	12.0	0.1

		Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visib (
62597	2013-08-28 05:00:00.000 +0200	Overcast	rain	17.194444	17.194444	0.96	5.8765	21.0	8.2	
63772	2013-02-16 04:00:00.000 +0100	Foggy	rain	2.788889	1.244444	1.00	5.8765	303.0	2.7	
64287	2013-01-01 15:00:00.000 +0100	Foggy	snow	-4.772222	-7.433333	0.96	5.8765	161.0	0.1	
64540	2013-01-02 04:00:00.000 +0100	Foggy	snow	-4.772222	-7.433333	0.96	5.8765	161.0	0.1	
66843	2013-03-23 03:00:00.000 +0100	Partly Cloudy	snow	-0.644444	-2.700000	0.81	5.8765	320.0	15.7	
67920	2013-05-08 01:00:00.000 +0200	Overcast	rain	16.650000	16.650000	0.96	5.8765	43.0	15.7	
68448	2013-11-28 01:00:00.000 +0100	Mostly Cloudy	snow	-0.255556	-2.244444	0.92	5.8765	192.0	11.0	
71495	2014-08-04 23:00:00.000 +0200	Mostly Cloudy	rain	19.988889	19.988889	0.94	5.8765	322.0	15.0	
72670	2014-02-20 22:00:00.000 +0100	Overcast	rain	6.772222	5.811111	0.94	5.8765	183.0	10.7	
78811	2014-09-05 22:00:00.000 +0200	Overcast	rain	17.294444	17.294444	0.92	5.8765	39.0	10.2	
82703	2015-07-16 02:00:00.000 +0200	Partly Cloudy	rain	16.288889	16.288889	0.88	5.8765	320.0	15.1	
84931	2015-05-16 23:00:00.000	Partly Cloudy	rain	13.466667	13.466667	0.83	5.8765	252.0	14.9	

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visib (
+0200									
86473	2015-10-02 05:00:00.000 +0200	Clear	rain	4.216667	2.883333	0.92	5.8765	334.0	9.9
+0200									
86662	2015-10-27 01:00:00.000 +0100	Mostly Cloudy	rain	7.516667	6.666667	0.86	5.8765	41.0	7.7
+0200									
87194	2015-09-19 05:00:00.000 +0200	Partly Cloudy	rain	18.144444	18.144444	0.85	5.8765	351.0	15.5
+0200									
93217	2016-03-26 04:00:00.000 +0100	Overcast	rain	3.961111	2.583333	0.83	5.8765	179.0	15.1
+0200									
96452	2016-09-09 23:00:00.000 +0200	Partly Cloudy	rain	20.438889	20.438889	0.61	5.8765	39.0	15.5
+0200									

Finding Null Values in the Data

In [23]: `df.isna()`

Out[23]:

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (k)
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...
96448	False	False	False	False	False	False	False	False	False
96449	False	False	False	False	False	False	False	False	False
96450	False	False	False	False	False	False	False	False	False
96451	False	False	False	False	False	False	False	False	False
96452	False	False	False	False	False	False	False	False	False

96453 rows × 12 columns

In [24]: `df.isnull()`

Out[24]:

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (k)
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...
96448	False	False	False	False	False	False	False	False	False
96449	False	False	False	False	False	False	False	False	False
96450	False	False	False	False	False	False	False	False	False
96451	False	False	False	False	False	False	False	False	False
96452	False	False	False	False	False	False	False	False	False

96453 rows × 12 columns

In [25]: `df.notnull().sum()`

```
Out[25]: Formatted Date      96453
          Summary           96453
          Precip Type       95936
          Temperature (C)   96453
          Apparent Temperature (C) 96453
          Humidity          96453
          Wind Speed (km/h)  96453
          Wind Bearing (degrees) 96453
          Visibility (km)    96453
          Loud Cover         96453
          Pressure (millibars) 96453
          Daily Summary      96453
          dtype: int64
```

Renaming column name "Summary" to "Weather" in dataframe

- Temporary Rename

```
In [26]: df.rename(columns = {'Summary' : 'Weather'})
```

Out[26]:

		Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0		2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.8
1		2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.8
2		2006-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204.0	14.9
3		2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269.0	15.8
4		2006-04-01 04:00:00.000 +0200	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259.0	15.8
...
96448		2016-09-09 19:00:00.000 +0200	Partly Cloudy	rain	26.016667	26.016667	0.43	10.9963	31.0	16.1
96449		2016-09-09 20:00:00.000 +0200	Partly Cloudy	rain	24.583333	24.583333	0.48	10.0947	20.0	15.5
96450		2016-09-09 21:00:00.000 +0200	Partly Cloudy	rain	22.038889	22.038889	0.56	8.9838	30.0	16.1
96451		2016-09-09 22:00:00.000 +0200	Partly Cloudy	rain	21.522222	21.522222	0.60	10.5294	20.0	16.1
96452		2016-09-09 23:00:00.000 +0200	Partly Cloudy	rain	20.438889	20.438889	0.61	5.8765	39.0	15.5

96453 rows × 12 columns

- Permanently Rename

```
In [27]: df.rename(columns = {'Summary' : 'Weather'}, inplace = True)
```

```
In [28]: df.head(1)
```

Out[28]:

	Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0	2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.8263



Analyzing Mean Visibility

```
In [29]: df['Visibility (km)'].mean()
```

```
Out[29]: 10.347324929240148
```

Analyzing Standard Deviation of 'Pressure' in Data

```
In [30]: df['Pressure (millibars)'].std()
```

```
Out[30]: 116.96990568258147
```

Finding Variance of "Humidity"

```
In [31]: df['Humidity'].var()
```

```
Out[31]: 0.03820959171844407
```

Analyzing all the instances when 'Dry' was Recorded

```
In [32]: df['Weather'].value_counts()
```

```
Out[32]:
```

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: Weather, dtype: int64

```
In [33]: df[df['Weather'] == 'Dry']
```

Out[33]:

		Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
12706		2007-07-20 10:00:00.000 +0200	Dry	rain	34.055556	33.005556	0.28	13.8138	179.0	9.9
12707		2007-07-20 11:00:00.000 +0200	Dry	rain	35.944444	34.927778	0.25	13.1054	168.0	10.3
12714		2007-07-20 18:00:00.000 +0200	Dry	rain	37.894444	35.761111	0.16	9.6922	179.0	9.9
12715		2007-07-20 19:00:00.000 +0200	Dry	rain	36.272222	34.772222	0.22	8.1144	170.0	9.9
12754		2007-07-22 10:00:00.000 +0200	Dry	rain	34.955556	34.166667	0.28	14.4417	248.0	9.9
12755		2007-07-22 11:00:00.000 +0200	Dry	rain	36.633333	35.077778	0.21	20.0606	223.0	10.3
12756		2007-07-22 12:00:00.000 +0200	Dry	rain	37.711111	35.605556	0.16	25.1643	220.0	9.9
12757		2007-07-22 13:00:00.000 +0200	Dry	rain	38.750000	36.622222	0.15	21.8799	230.0	9.9
12802		2007-07-24 10:00:00.000 +0200	Dry	rain	33.750000	32.538889	0.27	17.0821	251.0	9.9

		Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
		2007-07-24								
12803	11:00:00.000 +0200		Dry	rain	36.577778	35.738889	0.25	15.6331	231.0	10.3
		2007-07-24								
12804	12:00:00.000 +0200		Dry	rain	37.566667	35.677778	0.18	21.8477	222.0	9.9
		2007-07-26								
12850	10:00:00.000 +0200		Dry	rain	25.044444	25.044444	0.28	10.8031	352.0	9.9
		2007-07-27								
12875	11:00:00.000 +0200		Dry	rain	28.838889	27.444444	0.24	7.6314	28.0	10.3
		2007-07-27								
12876	12:00:00.000 +0200		Dry	rain	29.977778	28.261111	0.21	8.9033	358.0	9.9
		2008-09-06								
26250	18:00:00.000 +0200		Dry	rain	33.838889	32.127778	0.23	12.8317	211.0	9.9
		2008-09-07								
26268	12:00:00.000 +0200		Dry	rain	33.961111	32.888889	0.28	14.5061	220.0	9.9
		2008-09-07								
26271	15:00:00.000 +0200		Dry	rain	37.150000	35.444444	0.19	20.2699	249.0	9.9
		2009-04-08								
27013	13:00:00.000 +0200		Dry	rain	23.888889	23.888889	0.22	16.1000	180.0	9.9
		2009-08-21								
27371	11:00:00.000 +0200		Dry	rain	28.750000	27.577778	0.28	8.1144	74.0	10.3
		2012-08-20								
53626	10:00:00.000 +0200		Dry	rain	30.000000	28.677778	0.29	4.4919	355.0	9.9
		2012-08-25								
53748	12:00:00.000 +0200		Dry	rain	34.788889	33.555556	0.25	12.2843	261.0	9.9

	Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
57923	2012-03-17 11:00:00.000 +0100	Dry	rain	20.000000	20.000000	0.26	14.4900	200.0	9.9
57924	2012-03-17 12:00:00.000 +0100	Dry	rain	21.111111	21.111111	0.21	12.8800	230.0	9.9
57925	2012-03-17 13:00:00.000 +0100	Dry	rain	21.822222	21.822222	0.23	16.7601	205.0	10.3
57926	2012-03-17 14:00:00.000 +0100	Dry	rain	22.222222	22.222222	0.21	20.9300	200.0	9.9
57927	2012-03-17 15:00:00.000 +0100	Dry	rain	22.777778	22.777778	0.18	17.7100	200.0	9.9
57928	2012-03-17 16:00:00.000 +0100	Dry	rain	21.961111	21.961111	0.19	15.9229	209.0	10.3
57929	2012-03-17 17:00:00.000 +0100	Dry	rain	20.000000	20.000000	0.24	12.8800	190.0	9.9
57948	2012-03-18 12:00:00.000 +0100	Dry	rain	22.222222	22.222222	0.25	27.3700	220.0	9.9
58047	2012-03-21 15:00:00.000 +0100	Dry	rain	20.000000	20.000000	0.21	14.4900	320.0	9.9
58094	2012-03-23 14:00:00.000 +0100	Dry	rain	22.222222	22.222222	0.19	9.6600	300.0	9.9
58140	2012-03-25 13:00:00.000 +0200	Dry	rain	21.111111	21.111111	0.29	12.8800	310.0	9.9

	Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
58141	2012-03-25 14:00:00.000 +0200	Dry	rain	22.050000	22.050000	0.25	16.7118	288.0	10.3
70098	2013-09-07 18:00:00.000 +0200	Dry	rain	24.994444	24.994444	0.24	10.8192	349.0	16.1

- str.contains - filters rows in the DataFrame where the values in a particular column contain the specified substring.

```
In [34]: df[df['Weather'].str.contains('Dry')]
```

Out[34]:

	Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
9157	2007-04-23 13:00:00.000 +0200	Dry and Partly Cloudy	rain	20.072222	20.072222	0.17	2.8014	310.0	9.9
9160	2007-04-23 16:00:00.000 +0200	Dry and Partly Cloudy	rain	21.205556	21.205556	0.15	5.2969	128.0	9.9
9161	2007-04-23 17:00:00.000 +0200	Dry and Partly Cloudy	rain	21.244444	21.244444	0.15	2.5921	198.0	10.3
9162	2007-04-23 18:00:00.000 +0200	Dry and Partly Cloudy	rain	20.161111	20.161111	0.18	3.6708	172.0	9.9
12706	2007-07-20 10:00:00.000 +0200	Dry	rain	34.055556	33.005556	0.28	13.8138	179.0	9.9
75586	2014-03-22 13:00:00.000 +0100	Dry and Partly Cloudy	rain	22.344444	22.344444	0.27	29.7528	193.0	10.5
75587	2014-03-22 14:00:00.000 +0100	Dry and Partly Cloudy	rain	22.150000	22.150000	0.27	26.6616	220.0	16.1
75588	2014-03-22 15:00:00.000 +0100	Dry and Partly Cloudy	rain	22.172222	22.172222	0.27	24.9389	229.0	16.1
79664	2015-08-10 11:00:00.000 +0200	Dry and Partly Cloudy	rain	31.955556	30.266667	0.25	9.4990	172.0	15.5
79665	2015-08-10 12:00:00.000	Dry and Partly	rain	32.733333	30.794444	0.22	9.1448	157.0	16.1

Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (k)
+0200	Cloudy							

1 row × 9 columns

Finding mean value of each column against each 'Weather'

In [35]: `df.groupby('Weather').mean()`

Out[35]:

	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover	Press (millib)
Weather								
Breezy	7.922016	3.387654	0.637778	32.143948	233.018519	9.577115	0.0	563.917
Breezy and Dry	21.111111	21.111111	0.260000	33.810000	240.000000	9.982000	0.0	1021.600
Breezy and Foggy	-0.510317	-7.403492	0.938571	33.477880	160.628571	1.621960	0.0	1008.934
Breezy and Mostly Cloudy	11.093411	8.680588	0.637054	33.386345	227.639535	11.478302	0.0	1000.622
Breezy and Overcast	7.241614	3.492235	0.763144	33.037566	213.526515	11.067012	0.0	1002.114
Breezy and Partly Cloudy	12.492761	9.989349	0.545803	33.532796	259.282383	11.326058	0.0	996.398
Clear	11.925109	11.040338	0.729708	8.141352	179.180257	11.441788	0.0	951.763
Dangerously Windy and Partly Cloudy	8.944444	3.483333	0.490000	63.852600	307.000000	11.447100	0.0	1009.050
Drizzle	10.847578	10.011681	0.867949	10.356428	177.307692	8.069815	0.0	1014.931
Dry	29.083660	28.273529	0.230294	14.713979	230.294118	10.250965	0.0	1016.391
Dry and Mostly Cloudy	26.838492	25.929365	0.242143	13.667750	187.785714	10.115400	0.0	1014.872
Dry and Partly Cloudy	26.605749	25.982235	0.240814	12.304519	224.465116	10.987501	0.0	1017.242
Foggy	1.464035	-0.210419	0.950765	7.171649	168.668439	1.551411	0.0	1007.475
Humid and Mostly Cloudy	20.886389	20.886389	0.874250	10.058877	153.425000	9.732852	0.0	1012.887
Humid and Overcast	21.515079	21.515079	0.881429	9.740500	138.857143	9.052800	0.0	1014.550
Humid and Partly Cloudy	21.568301	21.568301	0.848824	9.938435	201.647059	10.633576	0.0	1011.974
Light Rain	10.021517	8.560317	0.888095	14.673489	180.761905	6.657989	0.0	1011.054
Mostly Cloudy	12.629334	11.624994	0.725069	11.418404	192.049299	11.117234	0.0	1010.840
Overcast	7.516502	5.789636	0.837232	12.027738	183.532747	9.275112	0.0	1005.632
Partly	16.024782	15.394033	0.648571	10.115130	190.161094	11.811517	0.0	1013.079

	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover	Press (millib)
Weather								
Cloudy								
Rain	10.096111	9.607222	0.947000	5.797610	211.800000	1.779050	0.0	1017.318
Windy	6.804861	2.009028	0.572500	42.165900	319.750000	10.712537	0.0	127.376
Windy and Dry	27.222222	26.344444	0.240000	40.250000	150.000000	9.982000	0.0	1020.200
Windy and Foggy	11.876389	9.769444	0.900000	44.178400	155.000000	1.903825	0.0	1011.975
Windy and Mostly Cloudy	11.834603	9.754286	0.600000	43.117640	261.428571	11.159600	0.0	979.825
Windy and Overcast	7.932963	3.696543	0.708667	43.378409	244.311111	9.806331	0.0	1006.446
Windy and Partly Cloudy	9.968076	6.551244	0.528806	44.610937	295.119403	11.484106	0.0	953.291

Analyzing Min & Max value of each column against each 'Weather'

In [36]: `df.groupby('Weather').min()`

Out[36]:

		Formatted Date	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover
Weather									
Breezy		2006-09-17 12:00:00.000 +0200	-2.777778	-10.411111	0.18	28.9800	0.0	0.0000	0.0
Breezy and Dry		2012-03-18 11:00:00.000 +0100	21.111111	21.111111	0.26	33.8100	240.0	9.9820	0.0
Breezy and Foggy		2006-03-13 07:00:00.000 +0100	-3.827778	-11.905556	0.83	29.1571	0.0	0.3542	0.0
Breezy and Mostly Cloudy		2006-01-22 17:00:00.000 +0100	-6.966667	-15.972222	0.15	28.9800	0.0	4.7817	0.0
Breezy and Overcast		2006-01-19 09:00:00.000 +0100	-11.533333	-21.716667	0.15	28.9800	0.0	3.3488	0.0
Breezy and Partly Cloudy		2006-03-25 08:00:00.000 +0100	-5.050000	-13.416667	0.17	28.9800	10.0	5.5384	0.0
Clear		2006-01-08 01:00:00.000 +0100	-20.000000	-27.716667	0.00	0.0000	0.0	0.0000	0.0
Dangerously Windy and Partly Cloudy		2007-01-29 13:00:00.000 +0100	8.944444	3.483333	0.49	63.8526	307.0	11.4471	0.0
Drizzle		2012-04-12 10:00:00.000 +0200	7.805556	4.938889	0.58	1.7388	5.0	0.0000	0.0
Dry		2007-07-20 10:00:00.000 +0200	20.000000	20.000000	0.15	4.4919	28.0	9.9820	0.0

		Formatted Date	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover
Weather									
Dry and Mostly Cloudy	2012-03-18 16:00:00.000 +0100	20.072222	20.072222	0.17	4.7334	30.0	9.9820	0.0	
Dry and Partly Cloudy	2007-04-23 13:00:00.000 +0200	20.000000	20.000000	0.15	0.0000	0.0	9.9820	0.0	
Foggy	2006-01-05 01:00:00.000 +0100	-21.822222	-26.227778	0.00	0.0000	0.0	0.0000	0.0	
Humid and Mostly Cloudy	2006-06-29 16:00:00.000 +0200	20.000000	20.000000	0.80	0.0000	0.0	4.5724	0.0	
Humid and Overcast	2010-08-06 13:00:00.000 +0200	20.105556	20.105556	0.85	1.7227	20.0	6.1663	0.0	
Humid and Partly Cloudy	2006-06-30 18:00:00.000 +0200	20.055556	20.055556	0.80	2.5277	1.0	6.1180	0.0	
Light Rain	2012-04-12 09:00:00.000 +0200	6.811111	3.516667	0.62	1.8998	8.0	0.0000	0.0	
Mostly Cloudy	2006-01-01 01:00:00.000 +0100	-16.894444	-23.872222	0.00	0.0000	0.0	0.0000	0.0	
Overcast	2006-01-01 03:00:00.000 +0100	-16.111111	-23.027778	0.00	0.0000	0.0	0.0000	0.0	
Partly Cloudy	2006-01-01 00:00:00.000 +0100	-20.555556	-24.233333	0.00	0.0000	0.0	0.0000	0.0	
Rain	2016-10-21 23:00:00.000 +0200	8.205556	6.105556	0.93	0.9821	1.0	0.0000	0.0	

	Formatted Date	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover
Weather								
Windy	2008-01-28 02:00:00.000 +0100	3.938889	-2.255556	0.41	40.2983	310.0	9.9820	0.0
Windy and Dry	2012-04-29 12:00:00.000 +0200	27.222222	26.344444	0.24	40.2500	150.0	9.9820	0.0
Windy and Foggy	2006-03-13 06:00:00.000 +0100	-2.150000	-10.577778	0.76	40.3144	20.0	0.7567	0.0
Windy and Mostly Cloudy	2007-01-19 05:00:00.000 +0100	3.888889	-2.522222	0.29	40.2500	40.0	9.9820	0.0
Windy and Overcast	2006-03-12 23:00:00.000 +0100	-2.266667	-10.644444	0.19	40.2500	20.0	3.2844	0.0
Windy and Partly Cloudy	2006-03-29 17:00:00.000 +0200	1.111111	-6.272222	0.28	40.2500	20.0	6.9874	0.0

In [37]: `df.groupby('Weather').max()`

Out[37]:

		Formatted Date	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover
Weather									
Breezy	2016-08-05 10:00:00.000 +0200	37.588889	35.700000	0.82	38.2697	350.0	16.1000	0.0	
Breezy and Dry	2012-03-18 11:00:00.000 +0100	21.111111	21.111111	0.26	33.8100	240.0	9.9820	0.0	
Breezy and Foggy	2015-02-08 17:00:00.000 +0100	17.616667	17.616667	1.00	38.0765	351.0	2.5277	0.0	
Breezy and Mostly Cloudy	2016-09-17 15:00:00.000 +0200	27.522222	27.766667	0.93	40.1373	350.0	16.1000	0.0	
Breezy and Overcast	2016-10-04 23:00:00.000 +0200	22.433333	22.433333	1.00	40.1212	359.0	16.1000	0.0	
Breezy and Partly Cloudy	2016-08-10 19:00:00.000 +0200	38.438889	36.922222	0.93	40.1373	350.0	16.1000	0.0	
Clear	2016-11-02 23:00:00.000 +0100	39.905556	38.661111	1.00	29.7850	359.0	16.1000	0.0	
Dangerously Windy and Partly Cloudy	2007-01-29 13:00:00.000 +0100	8.944444	3.483333	0.49	63.8526	307.0	11.4471	0.0	
Drizzle	2016-10-26 14:00:00.000 +0200	17.844444	17.844444	0.94	24.7940	355.0	15.7297	0.0	
Dry	2013-09-07 18:00:00.000 +0200	38.750000	36.622222	0.29	27.3700	358.0	16.1000	0.0	
Dry and Mostly Cloudy	2012-10-05 16:00:00.000 +0200	33.905556	31.611111	0.28	22.1375	306.0	10.3684	0.0	
Dry and Partly Cloudy	2015-08-10 12:00:00.000 +0200	37.800000	35.672222	0.29	29.7528	358.0	16.1000	0.0	
Foggy	2016-10-24 08:00:00.000 +0200	23.661111	23.661111	1.00	31.8780	359.0	3.2039	0.0	

		Formatted Date	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover
Weather									
Humid and Mostly Cloudy	2016-08-17 15:00:00.000 +0200	23.111111	23.111111	0.99	26.9514	342.0	15.8263	0.0	
Humid and Overcast	2016-07-26 13:00:00.000 +0200	24.055556	24.055556	0.92	19.4327	227.0	11.2056	0.0	
Humid and Partly Cloudy	2013-08-26 19:00:00.000 +0200	23.566667	23.566667	0.91	19.7547	339.0	11.2700	0.0	
Light Rain	2016-10-26 12:00:00.000 +0200	16.011111	16.011111	0.96	27.2734	345.0	16.1000	0.0	
Mostly Cloudy	2016-12-31 23:00:00.000 +0100	37.155556	38.116667	1.00	32.8279	359.0	16.1000	0.0	
Overcast	2016-10-29 14:00:00.000 +0200	33.644444	32.250000	1.00	32.7957	359.0	16.1000	0.0	
Partly Cloudy	2016-12-07 03:00:00.000 +0100	39.588889	39.344444	1.00	32.7957	359.0	16.1000	0.0	
Rain	2016-10-26 09:00:00.000 +0200	11.500000	11.500000	0.96	12.2843	355.0	5.5062	0.0	
Windy	2015-01-04 14:00:00.000 +0100	12.172222	12.172222	0.70	45.0800	340.0	15.8263	0.0	
Windy and Dry	2012-04-29 12:00:00.000 +0200	27.222222	26.344444	0.24	40.2500	150.0	9.9820	0.0	
Windy and Foggy	2015-05-14 01:00:00.000 +0200	19.633333	19.633333	0.96	46.6095	279.0	3.1395	0.0	
Windy and Mostly Cloudy	2015-05-06 21:00:00.000 +0200	19.905556	19.905556	0.89	53.1944	320.0	16.1000	0.0	

	Formatted Date	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover
Weather								
Windy and Overcast	2015-08-16 21:00:00.000 +0200	21.305556	21.305556	0.93	54.8849	340.0	16.1000	0.0
Windy and Partly Cloudy	2015-05-10 16:00:00.000 +0200	22.705556	22.705556	0.82	55.9314	330.0	16.1000	0.0

Finding all the records when Weather Condition is Fog

In [38]: `df[df['Weather'] == 'Foggy']`

Out[38]:

	Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
66	2006-04-11 18:00:00.000 +0200	Foggy	rain	10.911111	10.911111	0.86	22.3951	311.0	2.6
67	2006-04-11 19:00:00.000 +0200	Foggy	rain	8.800000	5.294444	0.99	26.5006	339.0	2.6
72	2006-04-12 00:00:00.000 +0200	Foggy	rain	8.200000	5.072222	0.96	20.4470	341.0	3.1
73	2006-04-12 01:00:00.000 +0200	Foggy	rain	8.177778	4.372222	0.93	27.8691	19.0	3.2
126	2006-04-14 06:00:00.000 +0200	Foggy	rain	5.211111	5.211111	0.92	4.7656	178.0	1.2
...
95858	2016-09-14 05:00:00.000 +0200	Foggy	rain	13.144444	13.144444	0.99	3.2039	2.0	2.2
95859	2016-09-14 06:00:00.000 +0200	Foggy	rain	13.822222	13.822222	0.94	4.8300	10.0	1.9
95860	2016-09-14 07:00:00.000 +0200	Foggy	rain	13.872222	13.872222	0.94	3.2039	321.0	1.9
95907	2016-09-16 06:00:00.000 +0200	Foggy	rain	14.905556	14.905556	1.00	7.5992	169.0	1.6
96228	2016-09-28 15:00:00.000 +0200	Foggy	rain	22.800000	22.800000	0.32	10.9158	231.0	0.7

7148 rows × 12 columns



Analyzing the instance when Weather is Foggy or Visibility is Above 0.7245

```
In [39]: df[(df['Weather'] == 'Foggy') | df['Visibility (km)'] > 0.7245]
```

Out[39]:

		Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0		2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.8
1		2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.8
2		2006-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204.0	14.9
3		2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269.0	15.8
4		2006-04-01 04:00:00.000 +0200	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259.0	15.8
...
96448		2016-09-09 19:00:00.000 +0200	Partly Cloudy	rain	26.016667	26.016667	0.43	10.9963	31.0	16.1
96449		2016-09-09 20:00:00.000 +0200	Partly Cloudy	rain	24.583333	24.583333	0.48	10.0947	20.0	15.5
96450		2016-09-09 21:00:00.000 +0200	Partly Cloudy	rain	22.038889	22.038889	0.56	8.9838	30.0	16.1
96451		2016-09-09 22:00:00.000 +0200	Partly Cloudy	rain	21.522222	21.522222	0.60	10.5294	20.0	16.1
96452		2016-09-09 23:00:00.000 +0200	Partly Cloudy	rain	20.438889	20.438889	0.61	5.8765	39.0	15.5

96089 rows × 12 columns

Finding all instances when -

A. Whether is Dry & HUmidity is above 0.26

In [40]: `df.head(1)`

Out[40]:

	Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0	2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.8263

In [41]: `df[(df['Weather'] == 'Dry') | df['Humidity'] > 0.26]`

Out[41]:

		Formatted Date	Weather	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)
0		2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251.0	15.8
1		2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259.0	15.8
2		2006-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204.0	14.9
3		2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269.0	15.8
4		2006-04-01 04:00:00.000 +0200	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259.0	15.8
...
96448		2016-09-09 19:00:00.000 +0200	Partly Cloudy	rain	26.016667	26.016667	0.43	10.9963	31.0	16.1
96449		2016-09-09 20:00:00.000 +0200	Partly Cloudy	rain	24.583333	24.583333	0.48	10.0947	20.0	15.5
96450		2016-09-09 21:00:00.000 +0200	Partly Cloudy	rain	22.038889	22.038889	0.56	8.9838	30.0	16.1
96451		2016-09-09 22:00:00.000 +0200	Partly Cloudy	rain	21.522222	21.522222	0.60	10.5294	20.0	16.1
96452		2016-09-09 23:00:00.000 +0200	Partly Cloudy	rain	20.438889	20.438889	0.61	5.8765	39.0	15.5

96431 rows × 12 columns

