Weather Data Analysis

---

1 The Weather Dataset

Here, The Weather Dataset is a time-series data set with per-hour information about the weather conditions at a particular location. It records Temperature, Dew Point Temperature, Relative Humidity, Wind Speed, Visibility, Pressure, and Conditions.

This data is available as a CSV file.

```
## # A tibble: 8,784 x 8
##    `Date/Time`   Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>          <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/1/2012 0:00   -1.8               -3.9          86                  4
##  2 1/1/2012 1:00   -1.8               -3.7          87                  4
##  3 1/1/2012 2:00   -1.8               -3.4          89                  7
##  4 1/1/2012 3:00   -1.5               -3.2          88                  6
##  5 1/1/2012 4:00   -1.5               -3.3          88                  7
##  6 1/1/2012 5:00   -1.4               -3.3          87                  9
##  7 1/1/2012 6:00   -1.5               -3.1          89                  7
##  8 1/1/2012 7:00   -1.4               -3.6          85                  7
##  9 1/1/2012 8:00   -1.4               -3.6          85                  9
## 10 1/1/2012 9:00   -1.3               -3.1          88                 15
## # i 8,774 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

1.1 .head()

It shows the first N rows in the data (by default, N=5).

```
## # A tibble: 6 x 8
##   `Date/Time`   Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##   <chr>          <dbl>              <dbl>       <dbl>              <dbl>
## 1 1/1/2012 0:00   -1.8               -3.9          86                  4
## 2 1/1/2012 1:00   -1.8               -3.7          87                  4
## 3 1/1/2012 2:00   -1.8               -3.4          89                  7
## 4 1/1/2012 3:00   -1.5               -3.2          88                  6
## 5 1/1/2012 4:00   -1.5               -3.3          88                  7
## 6 1/1/2012 5:00   -1.4               -3.3          87                  9
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

1.2 .shape

It shows the total no. of rows and no. of columns of the dataframe.

```
## (8784, 8)
```

1.3 .index

This attribute provides the index of the dataframe.

```
## RangeIndex(start=0, stop=8784, step=1)
```

1.4 .columns

It shows the name of each column.

```
## [1] "Date/Time"        "Temp_C"           "Dew Point Temp_C" "Rel Hum_%"
## [5] "Wind Speed_km/h"  "Visibility_km"    "Press_kPa"        "Weather"
```

1.5 .dtypes

It shows the data-type of each column.

```
## spc_tbl_ [8,784 x 8] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ Date/Time       : chr [1:8784] "1/1/2012 0:00" "1/1/2012 1:00" "1/1/2012 2:00" "1/1/2012 3:00" ..
##  $ Temp_C          : num [1:8784] -1.8 -1.8 -1.8 -1.5 -1.5 -1.4 -1.5 -1.4 -1.4 -1.3 ...
##  $ Dew Point Temp_C: num [1:8784] -3.9 -3.7 -3.4 -3.2 -3.3 -3.3 -3.1 -3.6 -3.6 -3.1 ...
##  $ Rel Hum_%       : num [1:8784] 86 87 89 88 88 87 89 85 85 88 ...
##  $ Wind Speed_km/h : num [1:8784] 4 4 7 6 7 9 7 7 9 15 ...
##  $ Visibility_km   : num [1:8784] 8 8 4 4 4.8 6.4 6.4 8 8 4 ...
##  $ Press_kPa       : num [1:8784] 101 101 101 101 101 ...
##  $ Weather         : chr [1:8784] "Fog" "Fog" "Freezing Drizzle,Fog" "Freezing Drizzle,Fog" ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   `Date/Time` = col_character(),
##   ..   Temp_C = col_double(),
##   ..   `Dew Point Temp_C` = col_double(),
##   ..   `Rel Hum_%` = col_double(),
##   ..   `Wind Speed_km/h` = col_double(),
##   ..   Visibility_km = col_double(),
##   ..   Press_kPa = col_double(),
##   ..   Weather = col_character()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

with python code the output is

```
## Date/Time            object
## Temp_C              float64
## Dew Point Temp_C    float64
```

```
## Rel Hum_%           float64
## Wind Speed_km/h     float64
## Visibility_km       float64
## Press_kPa           float64
## Weather              object
## dtype: object
```

1.6 .unique()

In a column, it shows all the unique values. It can be applied on a single column only, not on the whole dataframe.

```
##  [1] "Fog"
##  [2] "Freezing Drizzle,Fog"
##  [3] "Mostly Cloudy"
##  [4] "Cloudy"
##  [5] "Rain"
##  [6] "Rain Showers"
##  [7] "Mainly Clear"
##  [8] "Snow Showers"
##  [9] "Snow"
## [10] "Clear"
## [11] "Freezing Rain,Fog"
## [12] "Freezing Rain"
## [13] "Freezing Drizzle"
## [14] "Rain,Snow"
## [15] "Moderate Snow"
## [16] "Freezing Drizzle,Snow"
## [17] "Freezing Rain,Snow Grains"
## [18] "Snow,Blowing Snow"
## [19] "Freezing Fog"
## [20] "Haze"
## [21] "Rain,Fog"
## [22] "Drizzle,Fog"
## [23] "Drizzle"
## [24] "Freezing Drizzle,Haze"
## [25] "Freezing Rain,Haze"
## [26] "Snow,Haze"
## [27] "Snow,Fog"
## [28] "Snow,Ice Pellets"
## [29] "Rain,Haze"
## [30] "Thunderstorms,Rain"
## [31] "Thunderstorms,Rain Showers"
## [32] "Thunderstorms,Heavy Rain Showers"
## [33] "Thunderstorms,Rain Showers,Fog"
## [34] "Thunderstorms"
## [35] "Thunderstorms,Rain,Fog"
## [36] "Thunderstorms,Moderate Rain Showers,Fog"
## [37] "Rain Showers,Fog"
## [38] "Rain Showers,Snow Showers"
## [39] "Snow Pellets"
## [40] "Rain,Snow,Fog"
## [41] "Moderate Rain,Fog"
```

```
## [42] "Freezing Rain,Ice Pellets,Fog"
## [43] "Drizzle,Ice Pellets,Fog"
## [44] "Drizzle,Snow"
## [45] "Rain,Ice Pellets"
## [46] "Drizzle,Snow,Fog"
## [47] "Rain,Snow Grains"
## [48] "Rain,Snow,Ice Pellets"
## [49] "Snow Showers,Fog"
## [50] "Moderate Snow,Blowing Snow"
```

1.7 .is.na

SHow the total number of non-null Values in each column. It can be applied in both the Dataframe and a single column

```
## [1] 0
```

In a column, it shows all the unique values with their count. It can be applied on single column only.

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

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

1.8 .count

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1  8784
```

1.9 .info()

Provides basic information about the dataframe.

```
##                  vars    n    mean      sd  median trimmed     mad    min
## Date/Time*          1 8784 4392.50 2535.87 4392.50 4392.50 3255.79   1.00
## Temp_C              2 8784    8.80   11.69    9.30    9.11   13.94 -23.30
## Dew Point Temp_C    3 8784    2.56   10.88    3.30    3.03   12.97 -28.50
## Rel Hum_%           4 8784   67.43   16.92   68.00   68.15   19.27  18.00
## Wind Speed_km/h     5 8784   14.95    8.69   13.00   14.27    8.90   0.00
## Visibility_km       6 8784   27.66   12.62   25.00   27.76    1.33   0.20
## Press_kPa           7 8784  101.05    0.84  101.07  101.07    0.76  97.52
## Weather*            8 8784   15.58   11.49   20.00   15.14    7.41   1.00
##                     max   range  skew kurtosis    se
## Date/Time*       8784.00 8783.00  0.00    -1.20 27.06
## Temp_C             33.00   56.30 -0.18    -0.92  0.12
## Dew Point Temp_C   24.40   52.90 -0.32    -0.82  0.12
## Rel Hum_%         100.00   82.00 -0.32    -0.55  0.18
## Wind Speed_km/h    83.00   83.00  0.87     1.54  0.09
## Visibility_km      48.30   48.10  0.41    -0.35  0.13
## Press_kPa         103.65    6.13 -0.23     0.71  0.01
## Weather*           50.00   49.00 -0.03    -1.14  0.12
```

Lets Dive into Answering Some Useful Analysis Questions

---

2 Weather Analysis

2.1 Q) 1. Find all the unique 'Wind Speed' values in the data.

```
## # A tibble: 34 x 2
##    `Wind Speed_km/h`     n
##                <dbl> <int>
##  1                 0   309
##  2                 2     2
##  3                 4   474
##  4                 6   609
##  5                 7   677
##  6                 9   830
##  7                11   791
##  8                13   735
##  9                15   719
## 10                17   666
## # i 24 more rows
```

Using Unique

```
##  [1]  4  7  6  9 15 13 20 22 19 24 30 35 39 32 33 26 44 43 48 37 28 17 11  0 83
## [26] 70 57 46 41 52 50 63 54  2
```

2.2 Q) 2. Find the number of times when the 'Weather is exactly Clear'.

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1  1326
```

```
## # A tibble: 1,326 x 8
##    `Date/Time`    Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>           <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/3/2012 19:00  -16.9              -24.8          50                 24
##  2 1/5/2012 18:00   -7.1              -14.4          56                 11
##  3 1/5/2012 19:00   -9.2              -15.4          61                  7
##  4 1/5/2012 20:00   -9.8              -15.7          62                  9
##  5 1/5/2012 21:00   -9                -14.8          63                 13
##  6 1/11/2012 1:00  -10.7              -17.8          56                 17
##  7 1/11/2012 2:00  -12                -18.9          56                 19
##  8 1/11/2012 3:00  -12.7              -19.4          57                 19
##  9 1/11/2012 4:00  -13.4              -20.1          57                 17
## 10 1/15/2012 8:00  -23.3              -28.5          62                  7
## # i 1,316 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

2.3 Q) 3. Find the number of times when the 'Wind Speed was exactly 4 km/h'.

```
## # A tibble: 1 x 1
##       n
```

```
##   <int>
## 1   474
```

```
## # A tibble: 474 x 8
##    `Date/Time`    Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>           <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/1/2012 0:00    -1.8               -3.9          86                  4
##  2 1/1/2012 1:00    -1.8               -3.7          87                  4
##  3 1/5/2012 0:00    -8.8              -11.7          79                  4
##  4 1/5/2012 5:00    -7                 -9.5          82                  4
##  5 1/7/2012 2:00    -8.1              -11.1          79                  4
##  6 1/7/2012 3:00    -7.8              -10.8          79                  4
##  7 1/7/2012 5:00    -6.9               -9.7          80                  4
##  8 1/7/2012 20:00   -1.8               -3.7          87                  4
##  9 1/7/2012 22:00   -1.5               -3            89                  4
## 10 1/9/2012 2:00    -9                -14.1          66                  4
## # i 464 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

2.4 Q. 4) Find out all the Null Values in the data.

```
## [1] FALSE
```

This means that there are no null Values in the dataset ## Q. 5) Rename the column name 'Weather' of the dataframe to 'Weather Condition'.

```
## [1] "Date/Time"        "Temp_C"           "Dew Point Temp_C" "Rel Hum_%"
## [5] "Wind Speed_km/h"  "Visibility_km"    "Press_kPa"        "Weather_Dataset"
```

2.5 Q.6) What is the mean 'Visibility' ?

```
##                   vars    n    mean      sd  median trimmed     mad    min
## Date/Time*           1 8784 4392.50 2535.87 4392.50 4392.50 3255.79   1.00
## Temp_C               2 8784    8.80   11.69    9.30    9.11   13.94 -23.30
## Dew Point Temp_C     3 8784    2.56   10.88    3.30    3.03   12.97 -28.50
## Rel Hum_%            4 8784   67.43   16.92   68.00   68.15   19.27  18.00
## Wind Speed_km/h      5 8784   14.95    8.69   13.00   14.27    8.90   0.00
## Visibility_km        6 8784   27.66   12.62   25.00   27.76    1.33   0.20
## Press_kPa            7 8784  101.05    0.84  101.07  101.07    0.76  97.52
## Weather*             8 8784   15.58   11.49   20.00   15.14    7.41   1.00
##                     max   range  skew kurtosis    se
## Date/Time*       8784.00 8783.00  0.00    -1.20 27.06
## Temp_C             33.00   56.30 -0.18    -0.92  0.12
## Dew Point Temp_C   24.40   52.90 -0.32    -0.82  0.12
## Rel Hum_%         100.00   82.00 -0.32    -0.55  0.18
## Wind Speed_km/h    83.00   83.00  0.87     1.54  0.09
## Visibility_km      48.30   48.10  0.41    -0.35  0.13
## Press_kPa         103.65    6.13 -0.23     0.71  0.01
## Weather*           50.00   49.00 -0.03    -1.14  0.12
```

Singularly

```
## [1] 27.66445
```

2.6 Q. 7) What is the Standard Deviation of 'Pressure' in this data?

```
## [1] 0.8440047
```

2.7 Q. 8) Whats is the Variance of 'Relative Humidity' in this data ?

```
## [1] 286.2486
```

2.8 Q. 9) Find all instances when 'Snow' was recorded.

looking for just the Instance "Snow"

```
## # A tibble: 390 x 8
##    `Date/Time`    Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>           <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/3/2012 7:00   -14                -19.5          63                 19
##  2 1/4/2012 12:00  -13.7              -21.7          51                 11
##  3 1/4/2012 14:00  -11.3              -19            53                  7
##  4 1/4/2012 15:00  -10.2              -16.3          61                 11
##  5 1/4/2012 16:00   -9.4              -15.5          61                 13
##  6 1/4/2012 17:00   -8.9              -13.2          71                  9
##  7 1/4/2012 18:00   -8.9              -12.6          75                 11
##  8 1/4/2012 19:00   -8.4              -12.7          71                  9
##  9 1/4/2012 20:00   -7.8              -12.1          71                  9
## 10 1/4/2012 21:00   -7.6              -11.6          73                  7
## # i 380 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

But when we want to get all the columns that has "Snow" in it we use `grepl` in R and `str.contains` in python

```
## # A tibble: 583 x 8
##    `Date/Time`    Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>           <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/2/2012 17:00   -2.1               -9.5          57                 22
##  2 1/2/2012 20:00   -5.6              -13.4          54                 24
##  3 1/2/2012 21:00   -5.8              -12.8          58                 26
##  4 1/2/2012 23:00   -7.4              -14.1          59                 17
##  5 1/3/2012 0:00    -9                -16            57                 28
##  6 1/3/2012 2:00   -10.5              -15.8          65                 22
##  7 1/3/2012 3:00   -11.3              -18.7          54                 33
##  8 1/3/2012 5:00   -12.9              -19.1          60                 22
##  9 1/3/2012 6:00   -13.3              -19.3          61                 19
## 10 1/3/2012 7:00   -14                -19.5          63                 19
## # i 573 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

2.9 Q. 10) Find all instances when 'Wind Speed is above 24' and 'Visibility is 25'.

```
## # A tibble: 3,324 x 8
```

```
##    `Date/Time`    Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>           <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/1/2012 20:00    3.2                1.3          87                 19
##  2 1/1/2012 21:00    4                  1.7          85                 20
##  3 1/1/2012 23:00    5.3                2            79                 30
##  4 1/2/2012 0:00     5.2                1.5          77                 35
##  5 1/2/2012 1:00     4.6                0            72                 39
##  6 1/2/2012 2:00     3.9               -0.9          71                 32
##  7 1/2/2012 3:00     3.7               -1.5          69                 33
##  8 1/2/2012 4:00     2.9               -2.3          69                 32
##  9 1/2/2012 5:00     2.6               -2.3          70                 32
## 10 1/2/2012 6:00     2.3               -2.6          70                 26
## # i 3,314 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

2.10 Q. 11) What is the Mean value of each column against each 'Weather Conditon' ?

```
## # A tibble: 50 x 7
##    Weather Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h` Visibility_km
##    <chr>    <dbl>              <dbl>       <dbl>             <dbl>         <dbl>
##  1 Clear    6.83              0.0894        64.5              10.6          30.2
##  2 Cloudy   7.97              2.38          69.6              16.1          26.6
##  3 Drizzle  7.35              5.50          88.2              16.1          17.9
##  4 Drizzl~  8.07              7.03          93.3              11.9           5.26
##  5 Drizzl~  0.4              -0.7           92                20             4
##  6 Drizzl~  1.05              0.15          93.5              14            10.5
##  7 Drizzl~  0.693             0.12          95.9              15.5           5.51
##  8 Fog      4.30              3.16          92.3               7.95          6.25
##  9 Freezi~ -5.66             -8             83.6              16.6           9.2
## 10 Freezi~ -2.53             -4.18          88.5              17             5.27
## # i 40 more rows
## # i 1 more variable: Press_kPa <dbl>
```

Though the code is longer using R, Python makes It Easy for us

```
r.WeatherDataset.drop('Date/Time', axis=1).groupby('Weather').mean()
```

```
##                                      Temp_C  ...   Press_kPa
## Weather                                      ...
## Clear                              6.825716  ...  101.587443
## Cloudy                             7.970544  ...  100.911441
## Drizzle                            7.353659  ...  100.435366
## Drizzle,Fog                        8.067500  ...  100.786625
## Drizzle,Ice Pellets,Fog            0.400000  ...  100.790000
## Drizzle,Snow                       1.050000  ...  100.890000
## Drizzle,Snow,Fog                   0.693333  ...   99.281333
## Fog                                4.303333  ...  101.184067
## Freezing Drizzle                  -5.657143  ...  100.202857
## Freezing Drizzle,Fog              -2.533333  ...  100.441667
## Freezing Drizzle,Haze             -5.433333  ...  100.316667
## Freezing Drizzle,Snow             -5.109091  ...  100.520909
```

```
## Freezing Fog                                -7.575000  ...  102.320000
## Freezing Rain                               -3.885714  ...   99.647143
## Freezing Rain,Fog                           -2.225000  ...   99.945000
## Freezing Rain,Haze                          -4.900000  ...  100.375000
## Freezing Rain,Ice Pellets,Fog               -2.600000  ...  100.950000
## Freezing Rain,Snow Grains                   -5.000000  ...   98.560000
## Haze                                        -0.200000  ...  101.482500
## Mainly Clear                                12.558927  ...  101.248832
## Moderate Rain,Fog                            1.700000  ...   99.980000
## Moderate Snow                               -5.525000  ...  100.275000
## Moderate Snow,Blowing Snow                  -5.450000  ...  100.570000
## Mostly Cloudy                               10.574287  ...  101.025288
## Rain                                         9.786275  ...  100.233333
## Rain Showers                                13.722340  ...  100.404043
## Rain Showers,Fog                            12.800000  ...   99.830000
## Rain Showers,Snow Showers                    2.150000  ...  101.100000
## Rain,Fog                                     8.273276  ...  100.500862
## Rain,Haze                                    4.633333  ...  100.540000
## Rain,Ice Pellets                             0.600000  ...  100.120000
## Rain,Snow                                    1.055556  ...   99.951111
## Rain,Snow Grains                             1.900000  ...  100.600000
## Rain,Snow,Fog                                0.800000  ...  100.730000
## Rain,Snow,Ice Pellets                        1.100000  ...  100.105000
## Snow                                        -4.524103  ...  100.536103
## Snow Pellets                                 0.700000  ...   99.700000
## Snow Showers                                -3.506667  ...  100.963500
## Snow Showers,Fog                           -10.675000  ...  101.292500
## Snow,Blowing Snow                           -5.410526  ...   99.704737
## Snow,Fog                                    -5.075676  ...  100.688649
## Snow,Haze                                   -4.020000  ...  100.782000
## Snow,Ice Pellets                            -1.883333  ...  100.548333
## Thunderstorms                               24.150000  ...  100.230000
## Thunderstorms,Heavy Rain Showers            10.900000  ...  100.260000
## Thunderstorms,Moderate Rain Showers,Fog     19.600000  ...  100.010000
## Thunderstorms,Rain                          20.433333  ...  100.420000
## Thunderstorms,Rain Showers                  20.037500  ...  100.233750
## Thunderstorms,Rain Showers,Fog              21.600000  ...  100.063333
## Thunderstorms,Rain,Fog                      20.600000  ...  100.080000
##
## [50 rows x 6 columns]
```

The Date/Time column is recorded as a calculated column, so we hav to drop it ## Q. 12) What is the Minimum & Maximum value of each column against each 'Weather Conditon' ?

```
## # A tibble: 50 x 7
##    Weather Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h` Visibility_km
##    <chr>    <dbl>              <dbl>       <dbl>             <dbl>         <dbl>
##  1 Clear    -23.3              -28.5          20                 0          11.3
##  2 Cloudy   -21.4              -26.8          18                 0          11.3
##  3 Drizzle    1.1               -0.2          74                 0           6.4
##  4 Drizzl~    0                 -1.6          85                 0           1
##  5 Drizzl~    0.4               -0.7          92                20           4
##  6 Drizzl~    0.9                0.1          92                 9           9.7
##  7 Drizzl~    0.3               -0.1          92                 7           2.4
```

```
##  8 Fog      -16               -17.2          80                 0           0.2
##  9 Freezi~   -9               -12.2          78                 6           4.8
## 10 Freezi~   -6.4              -9            82                 6           3.6
## # i 40 more rows
## # i 1 more variable: Press_kPa <dbl>
```

The Above is for the Minimun, The Maximum numbers include

```
## # A tibble: 50 x 7
##    Weather Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h` Visibility_km
##    <chr>    <dbl>              <dbl>       <dbl>             <dbl>         <dbl>
##  1 Clear     32.8               20.4          99                33          48.3
##  2 Cloudy    30.5               22.6          99                54          48.3
##  3 Drizzle   18.8               17.7          96                30          25
##  4 Drizzl~   19.9               19.1         100                28           9.7
##  5 Drizzl~    0.4               -0.7          92                20           4
##  6 Drizzl~    1.2                0.2          95                19          11.3
##  7 Drizzl~    1.1                0.6          98                32           9.7
##  8 Fog       20.8               19.6         100                22           9.7
##  9 Freezi~   -2.3               -3.3          93                26          12.9
## 10 Freezi~   -0.3               -2.3          94                33           8
## # i 40 more rows
## # i 1 more variable: Press_kPa <dbl>
```

Lets try python code

```
r.WeatherDataset.groupby('Weather').min()
```

```
##                                        Date/Time  ...  Press_kPa
## Weather                                           ...
## Clear                            1/11/2012 1:00  ...      99.52
## Cloudy                           1/1/2012 17:00  ...      98.39
## Drizzle                         1/23/2012 21:00  ...      97.84
## Drizzle,Fog                     1/23/2012 20:00  ...      98.65
## Drizzle,Ice Pellets,Fog         12/17/2012 9:00  ...     100.79
## Drizzle,Snow                   12/17/2012 15:00  ...     100.63
## Drizzle,Snow,Fog               12/18/2012 21:00  ...      97.79
## Fog                               1/1/2012 0:00  ...      98.31
## Freezing Drizzle                1/13/2012 10:00  ...      98.44
## Freezing Drizzle,Fog              1/1/2012 2:00  ...      98.74
## Freezing Drizzle,Haze             2/1/2012 11:00  ...     100.28
## Freezing Drizzle,Snow             1/13/2012 3:00  ...      99.19
## Freezing Fog                      1/22/2012 6:00  ...     101.97
## Freezing Rain                   1/13/2012 11:00  ...      98.22
## Freezing Rain,Fog               1/17/2012 23:00  ...      98.32
## Freezing Rain,Haze                2/1/2012 14:00  ...     100.34
## Freezing Rain,Ice Pellets,Fog    12/17/2012 3:00  ...     100.95
## Freezing Rain,Snow Grains         1/13/2012 9:00  ...      98.56
## Haze                            1/22/2012 12:00  ...     100.35
## Mainly Clear                    1/10/2012 11:00  ...      98.67
## Moderate Rain,Fog                12/10/2012 8:00  ...      99.98
## Moderate Snow                    1/12/2012 15:00  ...      99.88
## Moderate Snow,Blowing Snow      12/27/2012 10:00  ...     100.50
```

```
## Mostly Cloudy                                1/1/2012 16:00  ...      98.36
## Rain                                         1/1/2012 18:00  ...      97.52
## Rain Showers                                 1/1/2012 22:00  ...      98.51
## Rain Showers,Fog                            10/20/2012 3:00  ...      99.83
## Rain Showers,Snow Showers                    11/4/2012 8:00  ...     101.09
## Rain,Fog                                    1/23/2012 18:00  ...      98.61
## Rain,Haze                                    3/13/2012 7:00  ...     100.50
## Rain,Ice Pellets                            12/18/2012 5:00  ...     100.12
## Rain,Snow                                    1/10/2012 5:00  ...      98.18
## Rain,Snow Grains                            12/21/2012 0:00  ...     100.60
## Rain,Snow,Fog                               12/8/2012 21:00  ...     100.73
## Rain,Snow,Ice Pellets                       12/21/2012 1:00  ...      99.85
## Snow                                         1/10/2012 1:00  ...      97.75
## Snow Pellets                                11/24/2012 15:00  ...      99.70
## Snow Showers                                 1/12/2012 7:00  ...      99.49
## Snow Showers,Fog                            12/26/2012 9:00  ...     100.63
## Snow,Blowing Snow                           1/13/2012 21:00  ...      98.11
## Snow,Fog                                    12/16/2012 15:00  ...      99.38
## Snow,Haze                                    2/1/2012 17:00  ...     100.61
## Snow,Ice Pellets                            12/10/2012 3:00  ...      99.40
## Thunderstorms                                7/16/2012 1:00  ...      99.84
## Thunderstorms,Heavy Rain Showers             5/29/2012 6:00  ...     100.26
## Thunderstorms,Moderate Rain Showers,Fog      7/17/2012 6:00  ...     100.01
## Thunderstorms,Rain                          5/25/2012 20:00  ...     100.19
## Thunderstorms,Rain Showers                  5/29/2012 16:00  ...      99.65
## Thunderstorms,Rain Showers,Fog               6/29/2012 3:00  ...      99.71
## Thunderstorms,Rain,Fog                       7/17/2012 5:00  ...     100.08
##
## [50 rows x 7 columns]
```

2.11 Q. 13) Show all the Records where Weather Condition is Fog.

```
WeatherDataset |>
  filter(Weather == 'Fog')
```

```
## # A tibble: 150 x 8
##    `Date/Time`    Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>           <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/1/2012 0:00    -1.8               -3.9          86                  4
##  2 1/1/2012 1:00    -1.8               -3.7          87                  4
##  3 1/1/2012 4:00    -1.5               -3.3          88                  7
##  4 1/1/2012 5:00    -1.4               -3.3          87                  9
##  5 1/1/2012 6:00    -1.5               -3.1          89                  7
##  6 1/1/2012 7:00    -1.4               -3.6          85                  7
##  7 1/1/2012 8:00    -1.4               -3.6          85                  9
##  8 1/1/2012 9:00    -1.3               -3.1          88                 15
##  9 1/1/2012 10:00   -1                 -2.3          91                  9
## 10 1/1/2012 11:00   -0.5               -2.1          89                  7
## # i 140 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

2.12 Q. 14) Find all instances when 'Weather is Clear' or 'Visibility is above 40'.

```
WeatherDataset |>
  filter((Visibility_km > 40) & grepl('Clear', Weather))
```

```
## # A tibble: 1,184 x 8
##    `Date/Time`     Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>            <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/5/2012 10:00    -6                -10            73                 17
##  2 1/5/2012 11:00    -5.6              -10.2          70                 22
##  3 1/5/2012 12:00    -4.7               -9.6          69                 20
##  4 1/5/2012 13:00    -4.4               -9.7          66                 26
##  5 1/5/2012 14:00    -5.1              -10.7          65                 22
##  6 1/5/2012 15:00    -4.3              -12            55                 26
##  7 1/14/2012 13:00  -17.1              -24.1          55                 17
##  8 1/15/2012 9:00   -22.2              -27.8          60                  9
##  9 1/15/2012 10:00  -20.6              -26.8          58                  9
## 10 1/15/2012 11:00  -19.3              -26.1          55                  9
## # i 1,174 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

2.13 Q. 15) Find all instances when :

2.13.1 A. 'Weather is Clear' and 'Relative Humidity is greater than 50'

2.13.2 or

2.13.3 B. 'Visibility is above 40'

```
## # A tibble: 4,034 x 8
##    `Date/Time`    Temp_C `Dew Point Temp_C` `Rel Hum_%` `Wind Speed_km/h`
##    <chr>           <dbl>              <dbl>       <dbl>              <dbl>
##  1 1/2/2012 12:00    1.7               -6.2          56                 48
##  2 1/3/2012 12:00  -14.9              -22.6          52                 20
##  3 1/3/2012 13:00  -15.1              -22.4          54                 22
##  4 1/3/2012 15:00  -14.8              -22.2          53                 19
##  5 1/3/2012 16:00  -15.3              -22.9          52                 22
##  6 1/3/2012 17:00  -15.8              -23.2          53                 22
##  7 1/3/2012 18:00  -16.3              -23.8          52                 24
##  8 1/4/2012 1:00   -17.9              -24.1          58                 11
##  9 1/4/2012 2:00   -18.1              -23.8          61                 15
## 10 1/4/2012 3:00   -18.5              -24.6          59                 13
## # i 4,024 more rows
## # i 3 more variables: Visibility_km <dbl>, Press_kPa <dbl>, Weather <chr>
```

We have come to the End of the Descriptive Analysis.

By Precious Ikebude.
