paper05_data_aggregation

October 10, 2021

1 Data Aggregation Across Data Sources

We have 3 different sources of data:

1. Our sensor data: that has the Indoor Air Quality and Indoor Environmental Data.

2. SINAICA: Outdoor Air Quality Monitoring Data from the Government.

3. OpenWeatherData: Outdoor Environmental Data.

We need it to be available that data to the models we plan to train. In the following sections this process is detailed.

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/importlib/_bootstrap.py:219:
RuntimeWarning: numpy.ufunc size changed, may indicate binary incompatibility.
Expected 192 from C header, got 216 from PyObject
/home/jaa6766/.conda/envs/cuda/lib/python3.7/importlib/_bootstrap.py:219:
RuntimeWarning: numpy.ufunc size changed, may indicate binary incompatibility.
Expected 192 from C header, got 216 from PyObject
/home/jaa6766/.conda/envs/cuda/lib/python3.7/importlib/_bootstrap.py:219:
RuntimeWarning: numpy.ufunc size changed, may indicate binary incompatibility.
Expected 192 from C header, got 216 from PyObject
/home/jaa6766/.conda/envs/cuda/lib/python3.7/importlib/_bootstrap.py:219:
RuntimeWarning: numpy.ufunc size changed, may indicate binary incompatibility.
Expected 192 from C header, got 216 from PyObject
```

Default Values:

- pandas max_columns=20
- pandas max_rows=60

New Values:

- pandas max_columns=35
- pandas max_rows=60

1.1 Indoor Data

datetime	temperature	pressure	humidity	gasResistance
2021-02-12 06:04:09.089621067	21.54	777.41	43.93	151328
2021-02-12 06:04:12.087778807	21.56	777.41	43.89	152702
2021-02-12 06:04:15.072475433	21.53	777.41	43.97	151328
2021-02-12 06:04:18.070170164	21.51	777.41	44.03	151464
2021-02-12 06:04:21.061994791	21.51	777.41	44.05	152425
...	...	...	...	...
2021-09-18 01:20:38.889113188	25.84	782.96	56.64	928867
2021-09-18 01:20:41.882042885	25.83	782.94	56.66	923130
2021-09-18 01:20:44.877856970	25.83	782.94	56.63	925034
2021-09-18 01:20:47.872255564	25.83	782.94	56.62	923130
2021-09-18 01:20:50.866486311	25.83	782.96	56.63	925034

datetime	IAQ	iaqAccuracy	year	month	day	hour
2021-02-12 06:04:09.089621067	37.5	1	2021	2	12	6

```
2021-02-12 06:04:12.087778807    35.6            1    2021      2    12      6
2021-02-12 06:04:15.072475433    37.5            1    2021      2    12      6
2021-02-12 06:04:18.070170164    38.5            1    2021      2    12      6
2021-02-12 06:04:21.061994791    36.9            1    2021      2    12      6
...                              ...          ...  ...        ...  ...    ...
2021-09-18 01:20:38.889113188   130.8            1    2021      9    18      1
2021-09-18 01:20:41.882042885   131.5            1    2021      9    18      1
2021-09-18 01:20:44.877856970   131.3            1    2021      9    18      1
2021-09-18 01:20:47.872255564   131.9            1    2021      9    18      1
2021-09-18 01:20:50.866486311   131.6            1    2021      9    18      1


                                 minute   second
datetime
2021-02-12 06:04:09.089621067         4        9
2021-02-12 06:04:12.087778807         4       12
2021-02-12 06:04:15.072475433         4       15
2021-02-12 06:04:18.070170164         4       18
2021-02-12 06:04:21.061994791         4       21
...                                 ...      ...
2021-09-18 01:20:38.889113188        20       38
2021-09-18 01:20:41.882042885        20       41
2021-09-18 01:20:44.877856970        20       44
2021-09-18 01:20:47.872255564        20       47
2021-09-18 01:20:50.866486311        20       50

[6285103 rows x 12 columns]
```

1.2 Outdoor Air Quality Data

```
                          NO         CO        NO2        NOx        O3  \
Fecha
2021-01-01 00:00:00  0.006000   1.000000   0.032000   0.036000   0.006000
2021-01-01 01:00:00  0.021000        NaN        NaN        NaN        NaN
2021-01-01 02:00:00  0.013000   1.100000   0.032000   0.039000   0.004000
2021-01-01 03:00:00  0.031000   1.200000   0.033000   0.043000   0.001000
2021-01-01 04:00:00  0.005000   1.200000   0.031000   0.039000   0.002000
...                       ...        ...        ...        ...        ...
2021-10-04 00:00:00  0.008292   0.545833   0.019083   0.026875   0.015833
2021-10-05 00:00:00  0.010000   0.563158   0.019722   0.030500   0.012278
2021-10-06 00:00:00  0.007571   0.672222   0.026111   0.035611   0.011000
2021-10-07 00:00:00  0.011565   0.713636   0.028636   0.040318   0.017909
2021-10-08 00:00:00  0.023778   0.758824   0.029412   0.050588   0.017941


                          PM10       PM2.5        SO2  year  month  day  hour  \
Fecha
2021-01-01 00:00:00  31.000000   19.000000   0.003000  2021      1    1     0
2021-01-01 01:00:00        NaN         NaN        NaN  2021      1    1     1
2021-01-01 02:00:00  37.000000   24.000000   0.003000  2021      1    1     2
2021-01-01 03:00:00  49.000000   39.000000   0.003000  2021      1    1     3
2021-01-01 04:00:00  80.000000   65.000000   0.003000  2021      1    1     4
...                        ...         ...        ...   ...    ...  ...   ...
2021-10-04 00:00:00  11.826087    7.913043   0.000750  2021     10    4     0
2021-10-05 00:00:00  11.090909    6.772727   0.000556  2021     10    5     0
2021-10-06 00:00:00  18.722222   11.833333   0.000111  2021     10    6     0
2021-10-07 00:00:00  26.772727   17.000000   0.001045  2021     10    7     0
2021-10-08 00:00:00  29.000000   17.705882   0.008176  2021     10    8     0


                        minute
Fecha
```

```
2021-01-01 00:00:00       0
2021-01-01 01:00:00       0
2021-01-01 02:00:00       0
2021-01-01 03:00:00       0
2021-01-01 04:00:00       0
...                      ...
2021-10-04 00:00:00       0
2021-10-05 00:00:00       0
2021-10-06 00:00:00       0
2021-10-07 00:00:00       0
2021-10-08 00:00:00       0

[2352 rows x 13 columns]
```

1.3 Outdoor Weather Data

```
                     temperature  feels_like  pressure  humidity  wind_speed  \
dt
2021-02-12 07:00:00        13.87       12.46      1020        44        0.00
2021-02-12 08:00:00        12.81       11.37      1020        47        0.00
2021-02-12 09:00:00        10.83        9.35      1019        53        1.54
2021-02-12 10:00:00         6.40        3.51      1019        61        4.12
2021-02-12 11:00:00         6.23        6.23      1019        57        0.00
...                          ...         ...       ...       ...         ...
2021-09-27 19:00:00        21.51       20.89      1006        45        0.89
2021-09-27 20:00:00        23.18       22.81      1005        48        0.45
2021-09-27 21:00:00        22.21       21.69      1025        46        6.17
2021-09-27 22:00:00        21.03       20.68      1004        57        0.45
2021-09-27 23:00:00        20.17       19.81      1004        60        5.66


                     wind_deg weather_main
dt
2021-02-12 07:00:00         0        Clear
2021-02-12 08:00:00         0        Clear
2021-02-12 09:00:00        60        Clear
2021-02-12 10:00:00        40        Clear
2021-02-12 11:00:00         0        Clear
...                       ...          ...
2021-09-27 19:00:00       139        Clear
2021-09-27 20:00:00       224         Rain
2021-09-27 21:00:00       220         Rain
2021-09-27 22:00:00       242         Rain
2021-09-27 23:00:00       140       Clouds

[5601 rows x 7 columns]
```

1.4 Merging the 3 Datasets: Indoor Data, Outdoor Air Quality Data, Outdoor Weather Data.

1.4.1 Merging Air Quality and Weather Data

```
                          NO        CO       NO2       NOx        O3  \
2021-02-12 07:00:00  0.244000  2.500000  0.035000  0.205000  0.002000
2021-02-12 08:00:00  0.146000  1.600000  0.030000  0.089000  0.004000
2021-02-12 09:00:00  0.099000  1.500000  0.039000  0.072000  0.012000
2021-02-12 10:00:00  0.024000  1.200000  0.030000  0.047000  0.025000
2021-02-12 11:00:00  0.009000  0.900000  0.016000  0.026000  0.033000
...                       ...       ...       ...       ...       ...
2021-09-14 00:00:00  0.017000  0.716667  0.015125  0.024625  0.013333
```

3

```
2021-09-15 00:00:00  0.027458  0.954167  0.028167  0.048625  0.019375
2021-09-16 00:00:00  0.006875  0.883333  0.028000  0.034458  0.022792
2021-09-17 00:00:00  0.010250  0.947619  0.032700  0.042750  0.026650
2021-09-18 00:00:00  0.009174  0.765217  0.027826  0.039043  0.015304


                          PM10       PM2.5        SO2  temperature  feels_like  \
2021-02-12 07:00:00  57.000000   25.000000   0.005000        13.87       12.46
2021-02-12 08:00:00  67.000000   33.000000   0.003000        12.81       11.37
2021-02-12 09:00:00  50.000000   28.000000   0.002000        10.83        9.35
2021-02-12 10:00:00  40.000000   21.000000   0.002000         6.40        3.51
2021-02-12 11:00:00  33.000000   19.000000   0.001000         6.23        6.23
...                        ...         ...        ...          ...         ...
2021-09-14 00:00:00   7.047619    4.500000   0.000125        15.85       14.95
2021-09-15 00:00:00  24.416667   17.333333   0.000875        17.95       17.08
2021-09-16 00:00:00  46.833333   40.041667   0.001542        18.45       17.63
2021-09-17 00:00:00  29.666667   24.875000   0.003350        18.34       17.69
2021-09-18 00:00:00  20.304348   16.391304   0.001304        17.61       17.85


                     pressure  humidity  wind_speed  wind_deg weather_main
2021-02-12 07:00:00    1020.0      44.0        0.00       0.0        Clear
2021-02-12 08:00:00    1020.0      47.0        0.00       0.0        Clear
2021-02-12 09:00:00    1019.0      53.0        1.54      60.0        Clear
2021-02-12 10:00:00    1019.0      61.0        4.12      40.0        Clear
2021-02-12 11:00:00    1019.0      57.0        0.00       0.0        Clear
...                       ...       ...         ...       ...          ...
2021-09-14 00:00:00    1025.0      56.0        4.12     170.0         Rain
2021-09-15 00:00:00    1023.0      49.0        7.72     130.0       Clouds
2021-09-16 00:00:00    1022.0      49.0        6.17     130.0        Smoke
2021-09-17 00:00:00    1024.0      56.0        4.63     300.0         Rain
2021-09-18 00:00:00    1015.0      93.0        1.37     199.0         Rain


[1341 rows x 15 columns]

                  NO        CO       NO2       NOx        O3       PM10  \
2021-09-16  0.006875  0.883333  0.028000  0.034458  0.022792  46.833333
2021-09-17  0.010250  0.947619  0.032700  0.042750  0.026650  29.666667
2021-09-18  0.009174  0.765217  0.027826  0.039043  0.015304  20.304348


                PM2.5       SO2  temperature  feels_like  pressure  humidity  \
2021-09-16  40.041667  0.001542        18.45       17.63    1022.0      49.0
2021-09-17  24.875000  0.003350        18.34       17.69    1024.0      56.0
2021-09-18  16.391304  0.001304        17.61       17.85    1015.0      93.0


            wind_speed  wind_deg weather_main
2021-09-16        6.17     130.0        Smoke
2021-09-17        4.63     300.0         Rain
2021-09-18        1.37     199.0         Rain
```

1.4.2 Merging Indoor and Outdoor (Air Quality and Weather) Data

```
                                 temperature  pressure  humidity  gasResistance  \
datetime
2021-09-16 23:59:42.445474625          25.86    780.44     57.00         877365
2021-09-16 23:59:45.439809322          25.85    780.40     57.04         874512
2021-09-16 23:59:48.434415102          25.85    780.42     57.07         877365
2021-09-16 23:59:51.428925753          25.84    780.42     57.07         872810
2021-09-16 23:59:54.423571348          25.84    780.42     57.07         872244
2021-09-16 23:59:57.418201685          25.84    780.44     57.06         868302
2021-09-17 00:00:00.412962675          25.84    780.42     57.02         869987
```

```
2021-09-17 00:00:03.407538652          25.84    780.44    57.02          877365
2021-09-17 00:00:06.402314186          25.83    780.42    57.05          868302
2021-09-17 00:00:09.396840096          25.84    780.42    57.00          875651


                                  IAQ  iaqAccuracy  year  month  day  hour  \
datetime
2021-09-16 23:59:42.445474625   246.4            1  2021      9   16    23
2021-09-16 23:59:45.439809322   246.7            1  2021      9   16    23
2021-09-16 23:59:48.434415102   245.9            1  2021      9   16    23
2021-09-16 23:59:51.428925753   246.9            1  2021      9   16    23
2021-09-16 23:59:54.423571348   247.7            1  2021      9   16    23
2021-09-16 23:59:57.418201685   249.6            1  2021      9   16    23
2021-09-17 00:00:00.412962675   250.0            1  2021      9   17     0
2021-09-17 00:00:03.407538652   248.1            1  2021      9   17     0
2021-09-17 00:00:06.402314186   249.8            1  2021      9   17     0
2021-09-17 00:00:09.396840096   248.6            1  2021      9   17     0


                                minute  second       NO        CO     NO2  \
datetime
2021-09-16 23:59:42.445474625       59      42      NaN       NaN     NaN
2021-09-16 23:59:45.439809322       59      45      NaN       NaN     NaN
2021-09-16 23:59:48.434415102       59      48      NaN       NaN     NaN
2021-09-16 23:59:51.428925753       59      51      NaN       NaN     NaN
2021-09-16 23:59:54.423571348       59      54      NaN       NaN     NaN
2021-09-16 23:59:57.418201685       59      57      NaN       NaN     NaN
2021-09-17 00:00:00.412962675        0       0  0.01025  0.947619  0.0327
2021-09-17 00:00:03.407538652        0       3      NaN       NaN     NaN
2021-09-17 00:00:06.402314186        0       6      NaN       NaN     NaN
2021-09-17 00:00:09.396840096        0       9      NaN       NaN     NaN


                                   NOx       O3       PM10   PM2.5      SO2  \
datetime
2021-09-16 23:59:42.445474625      NaN      NaN        NaN     NaN      NaN
2021-09-16 23:59:45.439809322      NaN      NaN        NaN     NaN      NaN
2021-09-16 23:59:48.434415102      NaN      NaN        NaN     NaN      NaN
2021-09-16 23:59:51.428925753      NaN      NaN        NaN     NaN      NaN
2021-09-16 23:59:54.423571348      NaN      NaN        NaN     NaN      NaN
2021-09-16 23:59:57.418201685      NaN      NaN        NaN     NaN      NaN
2021-09-17 00:00:00.412962675  0.04275  0.02665  29.666667  24.875  0.00335
2021-09-17 00:00:03.407538652      NaN      NaN        NaN     NaN      NaN
2021-09-17 00:00:06.402314186      NaN      NaN        NaN     NaN      NaN
2021-09-17 00:00:09.396840096      NaN      NaN        NaN     NaN      NaN


                                temperature_outdoor  feels_like  \
datetime
2021-09-16 23:59:42.445474625                   NaN         NaN
2021-09-16 23:59:45.439809322                   NaN         NaN
2021-09-16 23:59:48.434415102                   NaN         NaN
2021-09-16 23:59:51.428925753                   NaN         NaN
2021-09-16 23:59:54.423571348                   NaN         NaN
2021-09-16 23:59:57.418201685                   NaN         NaN
2021-09-17 00:00:00.412962675                 18.34       17.69
2021-09-17 00:00:03.407538652                   NaN         NaN
2021-09-17 00:00:06.402314186                   NaN         NaN
2021-09-17 00:00:09.396840096                   NaN         NaN


                                pressure_outdoor  humidity_outdoor  wind_speed  \
datetime
2021-09-16 23:59:42.445474625               NaN               NaN         NaN
```

```
2021-09-16 23:59:45.439809322                    NaN              NaN         NaN
2021-09-16 23:59:48.434415102                    NaN              NaN         NaN
2021-09-16 23:59:51.428925753                    NaN              NaN         NaN
2021-09-16 23:59:54.423571348                    NaN              NaN         NaN
2021-09-16 23:59:57.418201685                    NaN              NaN         NaN
2021-09-17 00:00:00.412962675                 1024.0             56.0        4.63
2021-09-17 00:00:03.407538652                    NaN              NaN         NaN
2021-09-17 00:00:06.402314186                    NaN              NaN         NaN
2021-09-17 00:00:09.396840096                    NaN              NaN         NaN


                               wind_deg weather_main
datetime
2021-09-16 23:59:42.445474625      NaN          NaN
2021-09-16 23:59:45.439809322      NaN          NaN
2021-09-16 23:59:48.434415102      NaN          NaN
2021-09-16 23:59:51.428925753      NaN          NaN
2021-09-16 23:59:54.423571348      NaN          NaN
2021-09-16 23:59:57.418201685      NaN          NaN
2021-09-17 00:00:00.412962675    300.0         Rain
2021-09-17 00:00:03.407538652      NaN          NaN
2021-09-17 00:00:06.402314186      NaN          NaN
2021-09-17 00:00:09.396840096      NaN          NaN
```

Dataset with Indoor and Outdoor Data: * 6285103 Rows * 27 Columns.

```
                               temperature  pressure  humidity  gasResistance  \
datetime
2021-02-12 07:00:02.979657173        21.51    777.30     43.74         144561
2021-02-12 08:00:01.982832432        21.04    776.92     42.35         153539
2021-02-12 09:00:00.729691744        20.41    776.33     42.56         153820
2021-02-12 10:00:02.449775934        20.27    776.20     42.21         144066
2021-02-12 11:00:01.044736862        19.91    776.25     42.26         142117
...                                    ...       ...       ...            ...
2021-09-14 00:00:02.855808020        24.31    780.42     55.46         873944
2021-09-15 00:00:01.543255568        25.52    779.38     54.92         938590
2021-09-16 00:00:01.135869265        27.09    778.30     48.41        1222727
2021-09-17 00:00:00.412962675        25.84    780.42     57.02         869987
2021-09-18 00:00:02.824432135        26.37    782.04     54.60         950530


                                 IAQ  iaqAccuracy  year  month  day  hour  \
datetime
2021-02-12 07:00:02.979657173   95.2            1  2021      2   12     7
2021-02-12 08:00:01.982832432   78.3            1  2021      2   12     8
2021-02-12 09:00:00.729691744   99.0            1  2021      2   12     9
2021-02-12 10:00:02.449775934  178.7            1  2021      2   12    10
2021-02-12 11:00:01.044736862  212.6            1  2021      2   12    11
...                              ...          ...   ...    ...  ...   ...
2021-09-14 00:00:02.855808020  230.3            3  2021      9   14     0
2021-09-15 00:00:01.543255568  125.5            1  2021      9   15     0
2021-09-16 00:00:01.135869265  148.3            3  2021      9   16     0
2021-09-17 00:00:00.412962675  250.0            1  2021      9   17     0
2021-09-18 00:00:02.824432135  111.9            1  2021      9   18     0


                               minute  second        NO        CO       NO2  \
datetime
2021-02-12 07:00:02.979657173       0       2  0.244000  2.500000  0.035000
2021-02-12 08:00:01.982832432       0       1  0.146000  1.600000  0.030000
2021-02-12 09:00:00.729691744       0       0  0.099000  1.500000  0.039000
2021-02-12 10:00:02.449775934       0       2  0.024000  1.200000  0.030000
2021-02-12 11:00:01.044736862       0       1  0.009000  0.900000  0.016000
```

```
...                               ...  ...       ...       ...       ...
2021-09-14 00:00:02.855808020       0    2  0.017000  0.716667  0.015125
2021-09-15 00:00:01.543255568       0    1  0.027458  0.954167  0.028167
2021-09-16 00:00:01.135869265       0    1  0.006875  0.883333  0.028000
2021-09-17 00:00:00.412962675       0    0  0.010250  0.947619  0.032700
2021-09-18 00:00:02.824432135       0    2  0.009174  0.765217  0.027826

                                 NOx        O3       PM10      PM2.5  \
datetime
2021-02-12 07:00:02.979657173  0.205000  0.002000  57.000000  25.000000
2021-02-12 08:00:01.982832432  0.089000  0.004000  67.000000  33.000000
2021-02-12 09:00:00.729691744  0.072000  0.012000  50.000000  28.000000
2021-02-12 10:00:02.449775934  0.047000  0.025000  40.000000  21.000000
2021-02-12 11:00:01.044736862  0.026000  0.033000  33.000000  19.000000
...                                 ...       ...       ...       ...
2021-09-14 00:00:02.855808020  0.024625  0.013333   7.047619   4.500000
2021-09-15 00:00:01.543255568  0.048625  0.019375  24.416667  17.333333
2021-09-16 00:00:01.135869265  0.034458  0.022792  46.833333  40.041667
2021-09-17 00:00:00.412962675  0.042750  0.026650  29.666667  24.875000
2021-09-18 00:00:02.824432135  0.039043  0.015304  20.304348  16.391304

                                 SO2   temperature_outdoor   feels_like  \
datetime
2021-02-12 07:00:02.979657173  0.005000                13.87        12.46
2021-02-12 08:00:01.982832432  0.003000                12.81        11.37
2021-02-12 09:00:00.729691744  0.002000                10.83         9.35
2021-02-12 10:00:02.449775934  0.002000                 6.40         3.51
2021-02-12 11:00:01.044736862  0.001000                 6.23         6.23
...                                 ...                  ...          ...
2021-09-14 00:00:02.855808020  0.000125                15.85        14.95
2021-09-15 00:00:01.543255568  0.000875                17.95        17.08
2021-09-16 00:00:01.135869265  0.001542                18.45        17.63
2021-09-17 00:00:00.412962675  0.003350                18.34        17.69
2021-09-18 00:00:02.824432135  0.001304                17.61        17.85

                                 pressure_outdoor   humidity_outdoor   wind_speed  \
datetime
2021-02-12 07:00:02.979657173              1020.0               44.0         0.00
2021-02-12 08:00:01.982832432              1020.0               47.0         0.00
2021-02-12 09:00:00.729691744              1019.0               53.0         1.54
2021-02-12 10:00:02.449775934              1019.0               61.0         4.12
2021-02-12 11:00:01.044736862              1019.0               57.0         0.00
...                                            ...                ...          ...
2021-09-14 00:00:02.855808020              1025.0               56.0         4.12
2021-09-15 00:00:01.543255568              1023.0               49.0         7.72
2021-09-16 00:00:01.135869265              1022.0               49.0         6.17
2021-09-17 00:00:00.412962675              1024.0               56.0         4.63
2021-09-18 00:00:02.824432135              1015.0               93.0         1.37

                                 wind_deg weather_main
datetime
2021-02-12 07:00:02.979657173       0.0          Clear
2021-02-12 08:00:01.982832432       0.0          Clear
2021-02-12 09:00:00.729691744      60.0          Clear
2021-02-12 10:00:02.449775934      40.0          Clear
2021-02-12 11:00:01.044736862       0.0          Clear
...                                 ...            ...
2021-09-14 00:00:02.855808020     170.0           Rain
2021-09-15 00:00:01.543255568     130.0         Clouds
```

```
2021-09-16 00:00:01.135869265        130.0          Smoke
2021-09-17 00:00:00.412962675        300.0           Rain
2021-09-18 00:00:02.824432135        199.0           Rain


[1186 rows x 27 columns]
```

1.5 Imputations

We found that the resulting dataframe after merging 2 datasets (Outdoor Data that is sampled every 1 hour and Indoor Data that is sampled every 3 seconds) contains repeated records on the columns of hourly data: SINAICA Gov't Air Quality Monitoring and OpenWeatherData.

We think that the repeated data can be an issue, as the data moves very abruptly from a record call it at 10:57 and 11:00. This is relevant as the real world is not represented by the data correctly. Temperature, pressure and general natural features move slowly from one value to other. But we don't have that data, and it's not easily obtainable.

Therefore, we propose an approach similar to the imputations using the interpolation incorporating noise, that could avert the overfitting issue on our machine learning and deep learning training.

```
CPU times: user 999 ms, sys: 176 ms, total: 1.18 s
Wall time: 1.17 s

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel_launcher.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame


See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

CPU times: user 1.9 s, sys: 1.09 s, total: 2.99 s
Wall time: 2.99 s

                                  temperature   pressure   humidity   gasResistance  \
datetime
2021-02-12 06:04:09.089621067        21.54      777.41      43.93          151328
2021-02-12 06:04:12.087778807        21.56      777.41      43.89          152702
2021-02-12 06:04:15.072475433        21.53      777.41      43.97          151328
2021-02-12 06:04:18.070170164        21.51      777.41      44.03          151464
2021-02-12 06:04:21.061994791        21.51      777.41      44.05          152425
...                                  ...         ...         ...            ...
2021-09-18 01:20:38.889113188        25.84      782.96      56.64          928867
2021-09-18 01:20:41.882042885        25.83      782.94      56.66          923130
2021-09-18 01:20:44.877856970        25.83      782.94      56.63          925034
2021-09-18 01:20:47.872255564        25.83      782.94      56.62          923130
2021-09-18 01:20:50.866486311        25.83      782.96      56.63          925034

                                   IAQ   iaqAccuracy   year   month   day   hour  \
datetime
2021-02-12 06:04:09.089621067      37.5            1   2021       2    12      6
2021-02-12 06:04:12.087778807      35.6            1   2021       2    12      6
2021-02-12 06:04:15.072475433      37.5            1   2021       2    12      6
2021-02-12 06:04:18.070170164      38.5            1   2021       2    12      6
2021-02-12 06:04:21.061994791      36.9            1   2021       2    12      6
...                                ...           ...    ...     ...   ...    ...
2021-09-18 01:20:38.889113188     130.8            1   2021       9    18      1
2021-09-18 01:20:41.882042885     131.5            1   2021       9    18      1
2021-09-18 01:20:44.877856970     131.3            1   2021       9    18      1
2021-09-18 01:20:47.872255564     131.9            1   2021       9    18      1
2021-09-18 01:20:50.866486311     131.6            1   2021       9    18      1

                                  minute   second        NO        CO       NO2  \
datetime
```

```
2021-02-12 06:04:09.089621067          4        9  0.205000  2.200000  0.031000
2021-02-12 06:04:12.087778807          4       12       NaN       NaN       NaN
2021-02-12 06:04:15.072475433          4       15       NaN       NaN       NaN
2021-02-12 06:04:18.070170164          4       18       NaN       NaN       NaN
2021-02-12 06:04:21.061994791          4       21       NaN       NaN       NaN
...                                  ...      ...       ...       ...       ...
2021-09-18 01:20:38.889113188         20       38       NaN       NaN       NaN
2021-09-18 01:20:41.882042885         20       41       NaN       NaN       NaN
2021-09-18 01:20:44.877856970         20       44       NaN       NaN       NaN
2021-09-18 01:20:47.872255564         20       47       NaN       NaN       NaN
2021-09-18 01:20:50.866486311         20       50  0.009174  0.765217  0.027826

                                    NOx        O3      PM10     PM2.5  \
datetime
2021-02-12 06:04:09.089621067  0.207000  0.002000  45.000000  22.000000
2021-02-12 06:04:12.087778807       NaN       NaN       NaN       NaN
2021-02-12 06:04:15.072475433       NaN       NaN       NaN       NaN
2021-02-12 06:04:18.070170164       NaN       NaN       NaN       NaN
2021-02-12 06:04:21.061994791       NaN       NaN       NaN       NaN
...                                 ...       ...       ...       ...
2021-09-18 01:20:38.889113188       NaN       NaN       NaN       NaN
2021-09-18 01:20:41.882042885       NaN       NaN       NaN       NaN
2021-09-18 01:20:44.877856970       NaN       NaN       NaN       NaN
2021-09-18 01:20:47.872255564       NaN       NaN       NaN       NaN
2021-09-18 01:20:50.866486311  0.039043  0.015304  20.304348  16.391304

                                    SO2  temperature_outdoor  feels_like  \
datetime
2021-02-12 06:04:09.089621067  0.004000                14.93        13.6
2021-02-12 06:04:12.087778807       NaN                  NaN         NaN
2021-02-12 06:04:15.072475433       NaN                  NaN         NaN
2021-02-12 06:04:18.070170164       NaN                  NaN         NaN
2021-02-12 06:04:21.061994791       NaN                  NaN         NaN
...                                 ...                  ...         ...
2021-09-18 01:20:38.889113188       NaN                  NaN         NaN
2021-09-18 01:20:41.882042885       NaN                  NaN         NaN
2021-09-18 01:20:44.877856970       NaN                  NaN         NaN
2021-09-18 01:20:47.872255564       NaN                  NaN         NaN
2021-09-18 01:20:50.866486311  0.001304                14.29        14.1

                               pressure_outdoor  humidity_outdoor  wind_speed  \
datetime
2021-02-12 06:04:09.089621067            1021.0              43.0        2.57
2021-02-12 06:04:12.087778807               NaN               NaN         NaN
2021-02-12 06:04:15.072475433               NaN               NaN         NaN
2021-02-12 06:04:18.070170164               NaN               NaN         NaN
2021-02-12 06:04:21.061994791               NaN               NaN         NaN
...                                         ...               ...         ...
2021-09-18 01:20:38.889113188               NaN               NaN         NaN
2021-09-18 01:20:41.882042885               NaN               NaN         NaN
2021-09-18 01:20:44.877856970               NaN               NaN         NaN
2021-09-18 01:20:47.872255564               NaN               NaN         NaN
2021-09-18 01:20:50.866486311            1018.0              89.0        2.11

                               wind_deg weather_main
datetime
2021-02-12 06:04:09.089621067    110.0         Clear
2021-02-12 06:04:12.087778807      NaN           NaN
2021-02-12 06:04:15.072475433      NaN           NaN
```

```
2021-02-12 06:04:18.070170164         NaN            NaN
2021-02-12 06:04:21.061994791         NaN            NaN

...                                    ...            ...
2021-09-18 01:20:38.889113188         NaN            NaN
2021-09-18 01:20:41.882042885         NaN            NaN
2021-09-18 01:20:44.877856970         NaN            NaN
2021-09-18 01:20:47.872255564         NaN            NaN
2021-09-18 01:20:50.866486311       350.0           Rain


[6285103 rows x 27 columns]
```

Here we can see the first and last data points to create the interpolation for the first and last values:

```
CPU times: user 340 ms, sys: 293 ms, total: 633 ms
Wall time: 632 ms

                                 temperature   pressure   humidity   gasResistance  \
datetime
2021-02-12 06:04:09.089621067          21.54     777.41      43.93          151328
2021-02-12 06:04:12.087778807          21.56     777.41      43.89          152702
2021-02-12 06:04:15.072475433          21.53     777.41      43.97          151328
2021-02-12 06:04:18.070170164          21.51     777.41      44.03          151464
2021-02-12 06:04:21.061994791          21.51     777.41      44.05          152425

...                                      ...        ...        ...             ...
2021-09-18 01:20:38.889113188          25.84     782.96      56.64          928867
2021-09-18 01:20:41.882042885          25.83     782.94      56.66          923130
2021-09-18 01:20:44.877856970          25.83     782.94      56.63          925034
2021-09-18 01:20:47.872255564          25.83     782.94      56.62          923130
2021-09-18 01:20:50.866486311          25.83     782.96      56.63          925034


                                 IAQ   iaqAccuracy   year   month   day   hour  \
datetime
2021-02-12 06:04:09.089621067   37.5             1   2021       2    12      6
2021-02-12 06:04:12.087778807   35.6             1   2021       2    12      6
2021-02-12 06:04:15.072475433   37.5             1   2021       2    12      6
2021-02-12 06:04:18.070170164   38.5             1   2021       2    12      6
2021-02-12 06:04:21.061994791   36.9             1   2021       2    12      6
...                              ...           ...    ...     ...   ...
2021-09-18 01:20:38.889113188  130.8             1   2021       9    18      1
2021-09-18 01:20:41.882042885  131.5             1   2021       9    18      1
2021-09-18 01:20:44.877856970  131.3             1   2021       9    18      1
2021-09-18 01:20:47.872255564  131.9             1   2021       9    18      1
2021-09-18 01:20:50.866486311  131.6             1   2021       9    18      1


                                 minute   second         NO         CO        NO2  \
datetime
2021-02-12 06:04:09.089621067         4        9   0.205000   2.200000   0.031000
2021-02-12 06:04:12.087778807         4       12        NaN        NaN        NaN
2021-02-12 06:04:15.072475433         4       15        NaN        NaN        NaN
2021-02-12 06:04:18.070170164         4       18        NaN        NaN        NaN
2021-02-12 06:04:21.061994791         4       21        NaN        NaN        NaN

...                                  ...      ...        ...        ...        ...
2021-09-18 01:20:38.889113188        20       38        NaN        NaN        NaN
2021-09-18 01:20:41.882042885        20       41        NaN        NaN        NaN
2021-09-18 01:20:44.877856970        20       44        NaN        NaN        NaN
2021-09-18 01:20:47.872255564        20       47        NaN        NaN        NaN
2021-09-18 01:20:50.866486311        20       50   0.009174   0.765217   0.027826


                                      NOx         O3       PM10      PM2.5  \
datetime
```

```
2021-02-12 06:04:09.089621067  0.207000  0.002000  45.000000  22.000000
2021-02-12 06:04:12.087778807       NaN       NaN        NaN        NaN
2021-02-12 06:04:15.072475433       NaN       NaN        NaN        NaN
2021-02-12 06:04:18.070170164       NaN       NaN        NaN        NaN
2021-02-12 06:04:21.061994791       NaN       NaN        NaN        NaN
...                                  ...       ...        ...        ...
2021-09-18 01:20:38.889113188       NaN       NaN        NaN        NaN
2021-09-18 01:20:41.882042885       NaN       NaN        NaN        NaN
2021-09-18 01:20:44.877856970       NaN       NaN        NaN        NaN
2021-09-18 01:20:47.872255564       NaN       NaN        NaN        NaN
2021-09-18 01:20:50.866486311  0.039043  0.015304  20.304348  16.391304

                                     SO2  temperature_outdoor  feels_like  \
datetime
2021-02-12 06:04:09.089621067  0.004000                14.93        13.6
2021-02-12 06:04:12.087778807       NaN                  NaN         NaN
2021-02-12 06:04:15.072475433       NaN                  NaN         NaN
2021-02-12 06:04:18.070170164       NaN                  NaN         NaN
2021-02-12 06:04:21.061994791       NaN                  NaN         NaN
...                                  ...                  ...         ...
2021-09-18 01:20:38.889113188       NaN                  NaN         NaN
2021-09-18 01:20:41.882042885       NaN                  NaN         NaN
2021-09-18 01:20:44.877856970       NaN                  NaN         NaN
2021-09-18 01:20:47.872255564       NaN                  NaN         NaN
2021-09-18 01:20:50.866486311  0.001304                14.29        14.1

                               pressure_outdoor  humidity_outdoor  wind_speed  \
datetime
2021-02-12 06:04:09.089621067            1021.0              43.0        2.57
2021-02-12 06:04:12.087778807               NaN               NaN         NaN
2021-02-12 06:04:15.072475433               NaN               NaN         NaN
2021-02-12 06:04:18.070170164               NaN               NaN         NaN
2021-02-12 06:04:21.061994791               NaN               NaN         NaN
...                                         ...               ...         ...
2021-09-18 01:20:38.889113188               NaN               NaN         NaN
2021-09-18 01:20:41.882042885               NaN               NaN         NaN
2021-09-18 01:20:44.877856970               NaN               NaN         NaN
2021-09-18 01:20:47.872255564               NaN               NaN         NaN
2021-09-18 01:20:50.866486311            1018.0              89.0        2.11

                               wind_deg weather_main
datetime
2021-02-12 06:04:09.089621067     110.0        Clear
2021-02-12 06:04:12.087778807       NaN          NaN
2021-02-12 06:04:15.072475433       NaN          NaN
2021-02-12 06:04:18.070170164       NaN          NaN
2021-02-12 06:04:21.061994791       NaN          NaN
...                                 ...          ...
2021-09-18 01:20:38.889113188       NaN          NaN
2021-09-18 01:20:41.882042885       NaN          NaN
2021-09-18 01:20:44.877856970       NaN          NaN
2021-09-18 01:20:47.872255564       NaN          NaN
2021-09-18 01:20:50.866486311     350.0         Rain

[6285103 rows x 27 columns]

                               temperature  pressure  humidity  gasResistance  \
datetime
2021-02-12 06:04:09.089621067        21.54    777.41     43.93         151328
2021-02-12 06:04:12.087778807        21.56    777.41     43.89         152702
```

```
2021-02-12 06:04:15.072475433          21.53    777.41    43.97           151328
2021-02-12 06:04:18.070170164          21.51    777.41    44.03           151464
2021-02-12 06:04:21.061994791          21.51    777.41    44.05           152425
…                                      …        …         …               …
2021-02-12 09:04:38.933268309          20.65    776.29    42.03           152564
2021-02-12 09:04:41.924828053          20.66    776.29    42.01           152149
2021-02-12 09:04:44.916538477          20.62    776.29    42.08           151737
2021-02-12 09:04:47.907913446          20.63    776.29    42.05           151464
2021-02-12 09:04:50.899550915          20.59    776.25    42.12           151601


                                IAQ  iaqAccuracy  year  month  day  hour  \
datetime
2021-02-12 06:04:09.089621067  37.5            1  2021      2   12     6
2021-02-12 06:04:12.087778807  35.6            1  2021      2   12     6
2021-02-12 06:04:15.072475433  37.5            1  2021      2   12     6
2021-02-12 06:04:18.070170164  38.5            1  2021      2   12     6
2021-02-12 06:04:21.061994791  36.9            1  2021      2   12     6
…                               …           …    …   …   …    …
2021-02-12 09:04:38.933268309  107.7           1  2021      2   12     9
2021-02-12 09:04:41.924828053  107.9           1  2021      2   12     9
2021-02-12 09:04:44.916538477  109.2           1  2021      2   12     9
2021-02-12 09:04:47.907913446  110.7           1  2021      2   12     9
2021-02-12 09:04:50.899550915  111.5           1  2021      2   12     9


                               minute  second        NO        CO       NO2  \
datetime
2021-02-12 06:04:09.089621067       4       9  0.205000  2.200000  0.031000
2021-02-12 06:04:12.087778807       4      12  0.205036  2.200274  0.031004
2021-02-12 06:04:15.072475433       4      15  0.205071  2.200547  0.031007
2021-02-12 06:04:18.070170164       4      18  0.205107  2.200821  0.031011
2021-02-12 06:04:21.061994791       4      21  0.205142  2.201095  0.031015
…                                   …       …         …         …         …
2021-02-12 09:04:38.933268309       4      38  0.093207  1.476827  0.038305
2021-02-12 09:04:41.924828053       4      41  0.093145  1.476578  0.038297
2021-02-12 09:04:44.916538477       4      44  0.093082  1.476329  0.038290
2021-02-12 09:04:47.907913446       4      47  0.093020  1.476080  0.038282
2021-02-12 09:04:50.899550915       4      50  0.092958  1.475831  0.038275


                                    NOx        O3      PM10     PM2.5  \
datetime
2021-02-12 06:04:09.089621067  0.207000  0.002000  45.000000  22.000000
2021-02-12 06:04:12.087778807  0.206998  0.002000  45.010949  22.002737
2021-02-12 06:04:15.072475433  0.206996  0.002000  45.021898  22.005474
2021-02-12 06:04:18.070170164  0.206995  0.002000  45.032847  22.008212
2021-02-12 06:04:21.061994791  0.206993  0.002000  45.043796  22.010949
…                                   …         …         …          …
2021-02-12 09:04:38.933268309  0.070069  0.013004  49.227575  27.459302
2021-02-12 09:04:41.924828053  0.070048  0.013015  49.219269  27.453488
2021-02-12 09:04:44.916538477  0.070027  0.013026  49.210963  27.447674
2021-02-12 09:04:47.907913446  0.070007  0.013037  49.202658  27.441860
2021-02-12 09:04:50.899550915  0.069986  0.013047  49.194352  27.436047


                                    SO2  temperature_outdoor  feels_like  \
datetime
2021-02-12 06:04:09.089621067  0.004000             14.930000   13.600000
2021-02-12 06:04:12.087778807  0.004001             14.929033   13.598960
2021-02-12 06:04:15.072475433  0.004002             14.928066   13.597920
2021-02-12 06:04:18.070170164  0.004003             14.927099   13.596880
2021-02-12 06:04:21.061994791  0.004004             14.926131   13.595839
…                                   …                     …            …
```

```
2021-02-12 09:04:38.933268309  0.002000        10.487816   8.898904
2021-02-12 09:04:41.924828053  0.002000        10.484136   8.894053
2021-02-12 09:04:44.916538477  0.002000        10.480457   8.889203
2021-02-12 09:04:47.907913446  0.002000        10.476777   8.884352
2021-02-12 09:04:50.899550915  0.002000        10.473098   8.879502
```

	pressure_outdoor	humidity_outdoor	wind_speed \
datetime			
2021-02-12 06:04:09.089621067	1021.000000	43.000000	2.570000
2021-02-12 06:04:12.087778807	1020.999088	43.000912	2.567655
2021-02-12 06:04:15.072475433	1020.998175	43.001825	2.565310
2021-02-12 06:04:18.070170164	1020.997263	43.002737	2.562965
2021-02-12 06:04:21.061994791	1020.996350	43.003650	2.560620
...	...	...	...
2021-02-12 09:04:38.933268309	1019.000000	53.617940	1.739286
2021-02-12 09:04:41.924828053	1019.000000	53.624585	1.741429
2021-02-12 09:04:44.916538477	1019.000000	53.631229	1.743571
2021-02-12 09:04:47.907913446	1019.000000	53.637874	1.745714
2021-02-12 09:04:50.899550915	1019.000000	53.644518	1.747857

	wind_deg	weather_main
datetime		
2021-02-12 06:04:09.089621067	110.000000	Clear
2021-02-12 06:04:12.087778807	109.899635	Clear
2021-02-12 06:04:15.072475433	109.799270	Clear
2021-02-12 06:04:18.070170164	109.698905	Clear
2021-02-12 06:04:21.061994791	109.598540	Clear
...	...	...
2021-02-12 09:04:38.933268309	58.455150	Clear
2021-02-12 09:04:41.924828053	58.438538	Clear
2021-02-12 09:04:44.916538477	58.421927	Clear
2021-02-12 09:04:47.907913446	58.405316	Clear
2021-02-12 09:04:50.899550915	58.388704	Clear

```
[3600 rows x 27 columns]
```

	temperature	pressure	humidity	gasResistance \
datetime				
2021-09-17 22:21:13.529673576	27.02	779.99	49.73	1049961
2021-09-17 22:21:16.524485349	27.02	780.01	49.74	1050781
2021-09-17 22:21:19.519613981	27.02	780.01	49.70	1044255
2021-09-17 22:21:22.514595508	27.02	779.99	49.69	1042636
2021-09-17 22:21:25.509442091	27.01	780.01	49.68	1047508
...	...	...	...	...
2021-09-18 01:20:38.889113188	25.84	782.96	56.64	928867
2021-09-18 01:20:41.882042885	25.83	782.94	56.66	923130
2021-09-18 01:20:44.877856970	25.83	782.94	56.63	925034
2021-09-18 01:20:47.872255564	25.83	782.94	56.62	923130
2021-09-18 01:20:50.866486311	25.83	782.96	56.63	925034

	IAQ	iaqAccuracy	year	month	day	hour \
datetime						
2021-09-17 22:21:13.529673576	32.3	1	2021	9	17	22
2021-09-17 22:21:16.524485349	31.8	1	2021	9	17	22
2021-09-17 22:21:19.519613981	33.4	1	2021	9	17	22
2021-09-17 22:21:22.514595508	35.0	1	2021	9	17	22
2021-09-17 22:21:25.509442091	34.7	1	2021	9	17	22
...	...	...	...	...	...	...
2021-09-18 01:20:38.889113188	130.8	1	2021	9	18	1
2021-09-18 01:20:41.882042885	131.5	1	2021	9	18	1

datetime							
2021-09-18 01:20:44.877856970	131.3	1	2021	9	18	1
2021-09-18 01:20:47.872255564	131.9	1	2021	9	18	1
2021-09-18 01:20:50.866486311	131.6	1	2021	9	18	1

	minute	second	NO	CO	NO2 \
datetime					
2021-09-17 22:21:13.529673576	21	13	0.009248	0.777732	0.028160
2021-09-17 22:21:16.524485349	21	16	0.009248	0.777726	0.028160
2021-09-17 22:21:19.519613981	21	19	0.009248	0.777719	0.028160
2021-09-17 22:21:22.514595508	21	22	0.009248	0.777713	0.028160
2021-09-17 22:21:25.509442091	21	25	0.009248	0.777707	0.028160
...	...	...	...	...	...
2021-09-18 01:20:38.889113188	20	38	0.009174	0.765217	0.027826
2021-09-18 01:20:41.882042885	20	41	0.009174	0.765217	0.027826
2021-09-18 01:20:44.877856970	20	44	0.009174	0.765217	0.027826
2021-09-18 01:20:47.872255564	20	47	0.009174	0.765217	0.027826
2021-09-18 01:20:50.866486311	20	50	0.009174	0.765217	0.027826

	NOx	O3	PM10	PM2.5 \
datetime				
2021-09-17 22:21:13.529673576	0.039298	0.016083	20.946691	16.973366
2021-09-17 22:21:16.524485349	0.039298	0.016082	20.946367	16.973072
2021-09-17 22:21:19.519613981	0.039298	0.016082	20.946043	16.972778
2021-09-17 22:21:22.514595508	0.039297	0.016082	20.945718	16.972484
2021-09-17 22:21:25.509442091	0.039297	0.016081	20.945394	16.972190
...	...	...	...	...
2021-09-18 01:20:38.889113188	0.039043	0.015304	20.304348	16.391304
2021-09-18 01:20:41.882042885	0.039043	0.015304	20.304348	16.391304
2021-09-18 01:20:44.877856970	0.039043	0.015304	20.304348	16.391304
2021-09-18 01:20:47.872255564	0.039043	0.015304	20.304348	16.391304
2021-09-18 01:20:50.866486311	0.039043	0.015304	20.304348	16.391304

	SO2	temperature_outdoor	feels_like \
datetime			
2021-09-17 22:21:13.529673576	0.001445	17.660085	17.839022
2021-09-17 22:21:16.524485349	0.001445	17.660060	17.839028
2021-09-17 22:21:19.519613981	0.001445	17.660034	17.839034
2021-09-17 22:21:22.514595508	0.001444	17.660009	17.839039
2021-09-17 22:21:25.509442091	0.001444	17.659984	17.839045
...	...	...	...
2021-09-18 01:20:38.889113188	0.001304	14.298203	14.109265
2021-09-18 01:20:41.882042885	0.001304	14.296152	14.106949
2021-09-18 01:20:44.877856970	0.001304	14.294101	14.104632
2021-09-18 01:20:47.872255564	0.001304	14.292051	14.102316
2021-09-18 01:20:50.866486311	0.001304	14.290000	14.100000

	pressure_outdoor	humidity_outdoor	wind_speed \
datetime			
2021-09-17 22:21:13.529673576	1015.617485	90.461451	1.593667
2021-09-17 22:21:16.524485349	1015.617173	90.462733	1.593554
2021-09-17 22:21:19.519613981	1015.616861	90.464015	1.593441
2021-09-17 22:21:22.514595508	1015.616549	90.465297	1.593328
2021-09-17 22:21:25.509442091	1015.616238	90.466579	1.593215
...	...	...	...
2021-09-18 01:20:38.889113188	1017.992588	89.009883	2.108172
2021-09-18 01:20:41.882042885	1017.994441	89.007412	2.108629
2021-09-18 01:20:44.877856970	1017.996294	89.004941	2.109086
2021-09-18 01:20:47.872255564	1017.998147	89.002471	2.109543
2021-09-18 01:20:50.866486311	1018.000000	89.000000	2.110000

```
                            wind_deg weather_main
datetime
2021-09-17 22:21:13.529673576   205.929554          Rain
2021-09-17 22:21:16.524485349   205.926054          Rain
2021-09-17 22:21:19.519613981   205.922554          Rain
2021-09-17 22:21:22.514595508   205.919055          Rain
2021-09-17 22:21:25.509442091   205.915555          Rain

...                                  ...            ...
2021-09-18 01:20:38.889113188   349.626930          Rain
2021-09-18 01:20:41.882042885   349.720198          Rain
2021-09-18 01:20:44.877856970   349.813465          Rain
2021-09-18 01:20:47.872255564   349.906733          Rain
2021-09-18 01:20:50.866486311   350.000000          Rain

[3600 rows x 27 columns]

CPU times: user 20.7 s, sys: 11.3 s, total: 32.1 s
Wall time: 32 s
```

1.6 Resampling

To reduce training time we propose to have a resampling of the data.

In the following subsections we create those resampled-data dataframes.

1.6.1 1 Minute Resampling

```
CPU times: user 6.83 s, sys: 605 ms, total: 7.44 s
Wall time: 7.43 s

                        temperature     pressure    humidity   gasResistance  \
datetime
2021-02-12 06:04:00     21.530000   777.410000   43.974000   151849.400000
2021-02-12 06:05:00     21.526250   777.408750   43.840000   152790.000000
2021-02-12 06:06:00     21.693000   777.409000   43.426000   152220.550000
2021-02-12 06:07:00     21.759000   777.410500   43.245500   151978.450000
2021-02-12 06:08:00     21.750500   777.390500   43.056000   150300.400000

...                          ...          ...         ...          ...
2021-09-18 01:16:00     25.871000   782.805000   56.607000   921467.050000
2021-09-18 01:17:00     25.861000   782.832000   56.587500   921211.600000
2021-09-18 01:18:00     25.850000   782.866000   56.597500   922348.700000
2021-09-18 01:19:00     25.836190   782.900000   56.683810   921997.095238
2021-09-18 01:20:00     25.836471   782.932941   56.662941   922768.882353

                            IAQ  iaqAccuracy     year   month    day   hour  \
datetime
2021-02-12 06:04:00     37.200000          1.0   2021.0    2.0   12.0    6.0
2021-02-12 06:05:00     32.162500          1.0   2021.0    2.0   12.0    6.0
2021-02-12 06:06:00     34.325000          1.0   2021.0    2.0   12.0    6.0
2021-02-12 06:07:00     36.190000          1.0   2021.0    2.0   12.0    6.0
2021-02-12 06:08:00     46.600000          1.0   2021.0    2.0   12.0    6.0

...                          ...          ...      ...     ...    ...    ...
2021-09-18 01:16:00    133.940000          1.0   2021.0    9.0   18.0    1.0
2021-09-18 01:17:00    134.815000          1.0   2021.0    9.0   18.0    1.0
2021-09-18 01:18:00    133.850000          1.0   2021.0    9.0   18.0    1.0
2021-09-18 01:19:00    134.190476          1.0   2021.0    9.0   18.0    1.0
2021-09-18 01:20:00    133.876471          1.0   2021.0    9.0   18.0    1.0

                        minute      second        NO        CO       NO2  \
datetime
2021-02-12 06:04:00        4.0   15.000000  0.205071  2.200547  0.031007
```

15

```
2021-02-12 06:05:00      5.0   45.875000   0.205302   2.202327   0.031031
2021-02-12 06:06:00      6.0   30.500000   0.205801   2.206159   0.031082
2021-02-12 06:07:00      7.0   30.500000   0.206512   2.211633   0.031155
2021-02-12 06:08:00      8.0   30.500000   0.207224   2.217108   0.031228
...                      ...   ...         ...        ...        ...
2021-09-18 01:16:00     16.0   28.500000   0.009174   0.765217   0.027826
2021-09-18 01:17:00     17.0   28.500000   0.009174   0.765217   0.027826
2021-09-18 01:18:00     18.0   28.500000   0.009174   0.765217   0.027826
2021-09-18 01:19:00     19.0   29.619048   0.009174   0.765217   0.027826
2021-09-18 01:20:00     20.0   26.000000   0.009174   0.765217   0.027826

                         NOx          O3       PM10       PM2.5        SO2  \
datetime
2021-02-12 06:04:00  0.206996   0.002000  45.021898  22.005474   0.004002
2021-02-12 06:05:00  0.206984   0.002000  45.093066  22.023266   0.004008
2021-02-12 06:06:00  0.206959   0.002000  45.246350  22.061588   0.004021
2021-02-12 06:07:00  0.206922   0.002000  45.465328  22.116332   0.004039
2021-02-12 06:08:00  0.206886   0.002000  45.684307  22.171077   0.004057
...                  ...        ...        ...        ...         ...
2021-09-18 01:16:00  0.039043   0.015304  20.304348  16.391304   0.001304
2021-09-18 01:17:00  0.039043   0.015304  20.304348  16.391304   0.001304
2021-09-18 01:18:00  0.039043   0.015304  20.304348  16.391304   0.001304
2021-09-18 01:19:00  0.039043   0.015304  20.304348  16.391304   0.001304
2021-09-18 01:20:00  0.039043   0.015304  20.304348  16.391304   0.001304

                     temperature_outdoor   feels_like   pressure_outdoor  \
datetime
2021-02-12 06:04:00            14.928066    13.597920        1020.998175
2021-02-12 06:05:00            14.921779    13.591159        1020.992245
2021-02-12 06:06:00            14.908239    13.576597        1020.979471
2021-02-12 06:07:00            14.888896    13.555794        1020.961223
2021-02-12 06:08:00            14.869553    13.534991        1020.942974
...                           ...          ...               ...
2021-09-18 01:16:00           14.469432    14.302671        1017.837863
2021-09-18 01:17:00           14.428419    14.256347        1017.874923
2021-09-18 01:18:00           14.387406    14.210022        1017.911983
2021-09-18 01:19:00           14.345368    14.162539        1017.949969
2021-09-18 01:20:00           14.306405    14.118530        1017.985176

                     humidity_outdoor   wind_speed    wind_deg
datetime
2021-02-12 06:04:00         43.001825     2.565310  109.799270
2021-02-12 06:05:00         43.007755     2.550068  109.146898
2021-02-12 06:06:00         43.020529     2.517240  107.741788
2021-02-12 06:07:00         43.038777     2.470342  105.734489
2021-02-12 06:08:00         43.057026     2.423444  103.727190
...                        ...           ...        ...
2021-09-18 01:16:00        89.216183     2.070006  341.839098
2021-09-18 01:17:00        89.166770     2.079148  343.704447
2021-09-18 01:18:00        89.117356     2.088289  345.569796
2021-09-18 01:19:00        89.066708     2.097659  347.481779
2021-09-18 01:20:00        89.019765     2.106343  349.253860

[313637 rows x 26 columns]
```

1.6.2 2 Minute Resampling

```
CPU times: user 3.18 s, sys: 141 ms, total: 3.32 s
Wall time: 3.32 s
```

	temperature	pressure	humidity	gasResistance
datetime				
2021-02-12 06:04:00	21.527692	777.409231	43.891538	152428.230769
2021-02-12 06:06:00	21.726000	777.409750	43.335750	152099.500000
2021-02-12 06:08:00	21.686250	777.365250	43.291500	147429.200000
2021-02-12 06:10:00	21.499500	777.302000	43.106250	149288.475000
2021-02-12 06:12:00	21.628250	777.279500	42.830750	149325.975000
…	…	…	…	…
2021-09-18 01:12:00	25.913500	782.776500	56.513250	920534.625000
2021-09-18 01:14:00	25.892750	782.793500	56.555500	921024.825000
2021-09-18 01:16:00	25.866000	782.818500	56.597250	921339.325000
2021-09-18 01:18:00	25.842927	782.883415	56.641707	922168.609756
2021-09-18 01:20:00	25.836471	782.932941	56.662941	922768.882353

	IAQ	iaqAccuracy	year	month	day	hour
datetime						
2021-02-12 06:04:00	34.100000	1.0	2021.0	2.0	12.0	6.0
2021-02-12 06:06:00	35.257500	1.0	2021.0	2.0	12.0	6.0
2021-02-12 06:08:00	72.652500	1.0	2021.0	2.0	12.0	6.0
2021-02-12 06:10:00	69.505000	1.0	2021.0	2.0	12.0	6.0
2021-02-12 06:12:00	71.237500	1.0	2021.0	2.0	12.0	6.0
…	…	…	…	…	…	…
2021-09-18 01:12:00	134.397500	1.0	2021.0	9.0	18.0	1.0
2021-09-18 01:14:00	134.380000	1.0	2021.0	9.0	18.0	1.0
2021-09-18 01:16:00	134.377500	1.0	2021.0	9.0	18.0	1.0
2021-09-18 01:18:00	134.024390	1.0	2021.0	9.0	18.0	1.0
2021-09-18 01:20:00	133.876471	1.0	2021.0	9.0	18.0	1.0

	minute	second	NO	CO	NO2
datetime					
2021-02-12 06:04:00	4.615385	34.000000	0.205214	2.201642	0.031022
2021-02-12 06:06:00	6.500000	30.500000	0.206156	2.208896	0.031119
2021-02-12 06:08:00	8.500000	30.500000	0.207580	2.219845	0.031265
2021-02-12 06:10:00	10.500000	29.900000	0.209003	2.230794	0.031411
2021-02-12 06:12:00	12.500000	29.500000	0.210427	2.241743	0.031557
…	…	…	…	…	…
2021-09-18 01:12:00	12.500000	28.500000	0.009174	0.765217	0.027826
2021-09-18 01:14:00	14.500000	28.500000	0.009174	0.765217	0.027826
2021-09-18 01:16:00	16.500000	28.500000	0.009174	0.765217	0.027826
2021-09-18 01:18:00	18.512195	29.073171	0.009174	0.765217	0.027826
2021-09-18 01:20:00	20.000000	26.000000	0.009174	0.765217	0.027826

	NOx	O3	PM10	PM2.5	SO2
datetime					
2021-02-12 06:04:00	0.206989	0.002000	45.065693	22.016423	0.004005
2021-02-12 06:06:00	0.206941	0.002000	45.355839	22.088960	0.004030
2021-02-12 06:08:00	0.206868	0.002000	45.793796	22.198449	0.004066
2021-02-12 06:10:00	0.206795	0.002000	46.231752	22.307938	0.004103
2021-02-12 06:12:00	0.206722	0.002000	46.669708	22.417427	0.004139
…	…	…	…	…	…
2021-09-18 01:12:00	0.039043	0.015304	20.304348	16.391304	0.001304
2021-09-18 01:14:00	0.039043	0.015304	20.304348	16.391304	0.001304
2021-09-18 01:16:00	0.039043	0.015304	20.304348	16.391304	0.001304
2021-09-18 01:18:00	0.039043	0.015304	20.304348	16.391304	0.001304
2021-09-18 01:20:00	0.039043	0.015304	20.304348	16.391304	0.001304

	temperature_outdoor	feels_like	pressure_outdoor
datetime			
2021-02-12 06:04:00	14.924197	13.593759	1020.994526

```
2021-02-12 06:06:00             14.898568    13.566195        1020.970347
2021-02-12 06:08:00             14.859881    13.524589        1020.933850
2021-02-12 06:10:00             14.821195    13.482984        1020.897354
2021-02-12 06:12:00             14.782509    13.441378        1020.860858
...                                    ...          ...                ...
2021-09-18 01:12:00             14.612977    14.464809        1017.708153
2021-09-18 01:14:00             14.530951    14.372159        1017.782273
2021-09-18 01:16:00             14.448925    14.279509        1017.856393
2021-09-18 01:18:00             14.365874    14.185701        1017.931439
2021-09-18 01:20:00             14.306405    14.118530        1017.985176


                     humidity_outdoor   wind_speed     wind_deg
datetime
2021-02-12 06:04:00         43.005474     2.555931   109.397810
2021-02-12 06:06:00         43.029653     2.493791   106.738139
2021-02-12 06:08:00         43.066150     2.399995   102.723540
2021-02-12 06:10:00         43.102646     2.306200    98.708942
2021-02-12 06:12:00         43.139142     2.212404    94.694343
...                               ...          ...          ...
2021-09-18 01:12:00         89.389129     2.038011   335.310377
2021-09-18 01:14:00         89.290303     2.056294   339.041075
2021-09-18 01:16:00         89.191476     2.074577   342.771773
2021-09-18 01:18:00         89.091414     2.093088   346.549104
2021-09-18 01:20:00         89.019765     2.106343   349.253860


[156819 rows x 26 columns]
```

1.6.3 3 Minute Resampling

```
CPU times: user 1.31 s, sys: 313 ms, total: 1.62 s
Wall time: 1.62 s

                     temperature      pressure    humidity   gasResistance  \
datetime
2021-02-12 06:00:00    21.530000    777.410000   43.974000    151849.400000
2021-02-12 06:05:00    21.689773    777.389432   43.361477    150039.409091
2021-02-12 06:10:00    21.538300    777.285200   42.909800    149975.940000
2021-02-12 06:15:00    21.563900    777.269000   42.704100    150897.020000
2021-02-12 06:20:00    21.616931    777.223960   42.695545    149963.910891
...                          ...           ...         ...              ...
2021-09-18 01:00:00    25.987500    782.832000   56.333000    918660.290000
2021-09-18 01:05:00    25.966100    782.800200   56.379800    920302.740000
2021-09-18 01:10:00    25.915100    782.782200   56.511300    920836.000000
2021-09-18 01:15:00    25.860099    782.840594   56.613069    921519.306931
2021-09-18 01:20:00    25.836471    782.932941   56.662941    922768.882353

                         IAQ  iaqAccuracy    year   month    day   hour  \
datetime
2021-02-12 06:00:00    37.200000          1.0  2021.0    2.0   12.0    6.0
2021-02-12 06:05:00    51.973864          1.0  2021.0    2.0   12.0    6.0
2021-02-12 06:10:00    67.172000          1.0  2021.0    2.0   12.0    6.0
2021-02-12 06:15:00    65.798000          1.0  2021.0    2.0   12.0    6.0
2021-02-12 06:20:00    71.275248          1.0  2021.0    2.0   12.0    6.0
...                          ...          ...     ...     ...    ...    ...
2021-09-18 01:00:00   135.811000          1.0  2021.0    9.0   18.0    1.0
2021-09-18 01:05:00   134.243000          1.0  2021.0    9.0   18.0    1.0
2021-09-18 01:10:00   134.184000          1.0  2021.0    9.0   18.0    1.0
2021-09-18 01:15:00   134.372277          1.0  2021.0    9.0   18.0    1.0
2021-09-18 01:20:00   133.876471          1.0  2021.0    9.0   18.0    1.0
```

```
                        minute     second        NO        CO       NO2  \
datetime
2021-02-12 06:00:00   4.000000  15.000000  0.205071  2.200547  0.031007
2021-02-12 06:05:00   7.272727  31.897727  0.206726  2.213276  0.031177
2021-02-12 06:10:00  12.000000  29.660000  0.210071  2.239005  0.031520
2021-02-12 06:15:00  17.000000  28.890000  0.213629  2.266378  0.031885
2021-02-12 06:20:00  22.009901  29.029703  0.217205  2.293887  0.032252
...                        ...        ...       ...       ...       ...
2021-09-18 01:00:00   2.000000  29.710000  0.009174  0.765217  0.027826
2021-09-18 01:05:00   7.000000  29.500000  0.009174  0.765217  0.027826
2021-09-18 01:10:00  12.000000  28.600000  0.009174  0.765217  0.027826
2021-09-18 01:15:00  17.019802  28.732673  0.009174  0.765217  0.027826
2021-09-18 01:20:00  20.000000  26.000000  0.009174  0.765217  0.027826

                          NOx        O3      PM10     PM2.5       SO2  \
datetime
2021-02-12 06:00:00  0.206996  0.002000  45.021898  22.005474  0.004002
2021-02-12 06:05:00  0.206911  0.002000  45.531022  22.132755  0.004044
2021-02-12 06:10:00  0.206740  0.002000  46.560219  22.390055  0.004130
2021-02-12 06:15:00  0.206557  0.002000  47.655109  22.663777  0.004221
2021-02-12 06:20:00  0.206374  0.002000  48.755474  22.938869  0.004313
...                       ...       ...        ...        ...       ...
2021-09-18 01:00:00  0.039043  0.015304  20.304348  16.391304  0.001304
2021-09-18 01:05:00  0.039043  0.015304  20.304348  16.391304  0.001304
2021-09-18 01:10:00  0.039043  0.015304  20.304348  16.391304  0.001304
2021-09-18 01:15:00  0.039043  0.015304  20.304348  16.391304  0.001304
2021-09-18 01:20:00  0.039043  0.015304  20.304348  16.391304  0.001304

                     temperature_outdoor  feels_like  pressure_outdoor  \
datetime
2021-02-12 06:00:00            14.928066   13.597920       1020.998175
2021-02-12 06:05:00            14.883093   13.549553       1020.955748
2021-02-12 06:10:00            14.792181   13.451779       1020.869982
2021-02-12 06:15:00            14.695465   13.347765       1020.778741
2021-02-12 06:20:00            14.598266   13.243230       1020.687044
...                                  ...         ...               ...
2021-09-18 01:00:00            15.043613   14.951220       1017.319024
2021-09-18 01:05:00            14.838548   14.719595       1017.504324
2021-09-18 01:10:00            14.633484   14.487971       1017.689623
2021-09-18 01:15:00            14.427393   14.255188       1017.875849
2021-09-18 01:20:00            14.306405   14.118530       1017.985176

                     humidity_outdoor  wind_speed    wind_deg
datetime
2021-02-12 06:00:00         43.001825    2.565310  109.799270
2021-02-12 06:05:00         43.044252    2.456273  105.132299
2021-02-12 06:10:00         43.130018    2.235853   95.697993
2021-02-12 06:15:00         43.221259    2.001364   85.661496
2021-02-12 06:20:00         43.312956    1.765703   75.574818
...                               ...         ...         ...
2021-09-18 01:00:00         89.907968    1.942026  315.724212
2021-09-18 01:05:00         89.660902    1.987733  325.050957
2021-09-18 01:10:00         89.413836    2.033440  334.377702
2021-09-18 01:15:00         89.165534    2.079376  343.751081
2021-09-18 01:20:00         89.019765    2.106343  349.253860

[62729 rows x 26 columns]
```

1.7 References

- https://scikit-learn.org/stable/modules/linear_model.html#generalized-linear-regression

- https://pythonhealthcare.org/2018/05/03/81-distribution-fitting-to-data/

- https://medium.com/@amirarsalan.rajabi/distribution-fitting-with-python-scipy-bb70a42c0aed

- https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KernelDensity.html?highlight=kernel%20density#sklearn.neighbors.KernelDensity

- https://pandas.pydata.org/pandas-docs/stable/user_guide/timeseries.html#resampling

- https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.merge_asof.html#pandas.merge_asof

- https://openweathermap.org/history-bulk

- https://pandas.pydata.org/pandas-docs/stable/user_guide/missing_data.html#interpolation

- https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.interpolate.html?highlight=interpolate