paper01_eda

October 9, 2021

1 Exploratory Data Analysis.

Jorge III Altamirano Astorga, Luz Aurora Hernández Martínez, Ita-Andehui Santiago Castillejos.

1.1 Reading Files

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/patsy/constraint.py:13: DeprecationWarning: Using or importing the ABCs
from 'collections' instead of from 'collections.abc' is deprecated since Python
3.3,and in 3.9 it will stop working
  from collections import Mapping
/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
```

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

Listing data files from: /home/jaa6766/Documents/jorge3a/itam/deeplearning/dlfinal/data/airdata

- Loading air-20210212-060408.json.gz
- Loading air-20210212-060529.json.gz
- Loading air-20210212-060534.json.gz
- Loading air-20210212-060545.json.gz
- Loading air-20210212-153737.json.gz
- Loading air-20210212-153808.json.gz
- Loading air-20210212-153921.json.gz
- Loading air-20210218-233000.json.gz
- Loading air-20210308-213952.json.gz
- Loading air-20210308-214511.json.gz
- Loading air-20210326-110304.json.gz
- Loading air-20210520-235254.json.gz
- Loading air-20210617-141714.json.gz

- Loading air-20210415-234052.json.gz

- Loading air-20210627-224727.json.gz

- Loading air-20210723-084450.json.gz

- Loading air-20210812-211858.json.gz

- Loading air-20210822-191723.json.gz

- Loading air-20210827-131714.json.gz

- Loading air-20210828-132805.json.gz

- Loading air-20210830-121252.json.gz

Done!

```
CPU times: user 2min 55s, sys: 5.54 s, total: 3min 1s
Wall time: 3min 1s

   temperature  pressure  humidity  gasResistance   IAQ  iaqAccuracy  \
0        21.54    777.41     43.93         151328  37.5            1
1        21.56    777.41     43.89         152702  35.6            1
2        21.53    777.41     43.97         151328  37.5            1
3        21.51    777.41     44.03         151464  38.5            1
4        21.51    777.41     44.05         152425  36.9            1


                       datetime  year  month  day  hour  minute
0 2021-02-12 06:04:09.089621067  2021      2   12     6       4
1 2021-02-12 06:04:12.087778807  2021      2   12     6       4
2 2021-02-12 06:04:15.072475433  2021      2   12     6       4
3 2021-02-12 06:04:18.070170164  2021      2   12     6       4
4 2021-02-12 06:04:21.061994791  2021      2   12     6       4
```
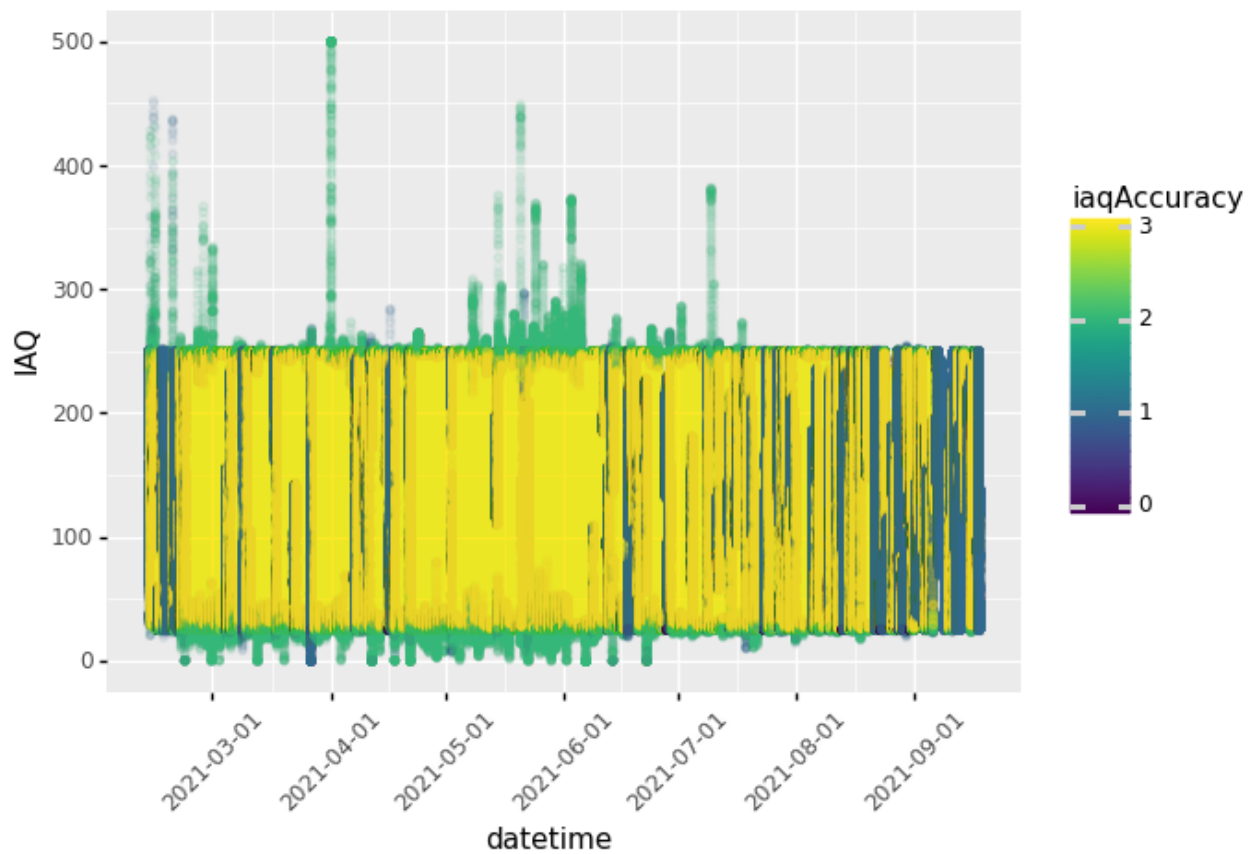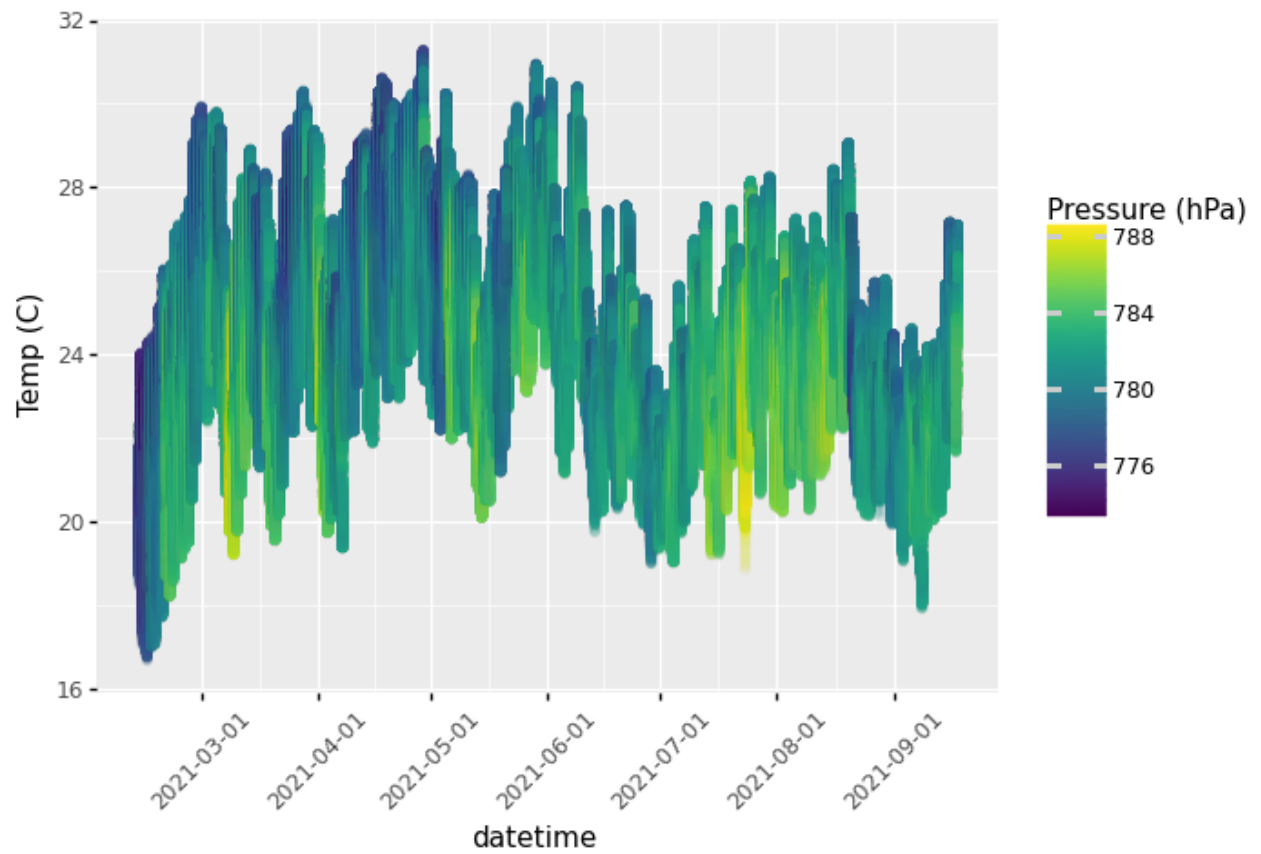
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

```
(6285103, 12)
```

```
<ggplot: (8759878048277)>
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will not call `transform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any exception that happen during thetransform in `preprocessing_exc_tuple` in IPython 7.17 and above.

```
<ggplot: (8759869828161)>
```

1.2   Hourly Plots
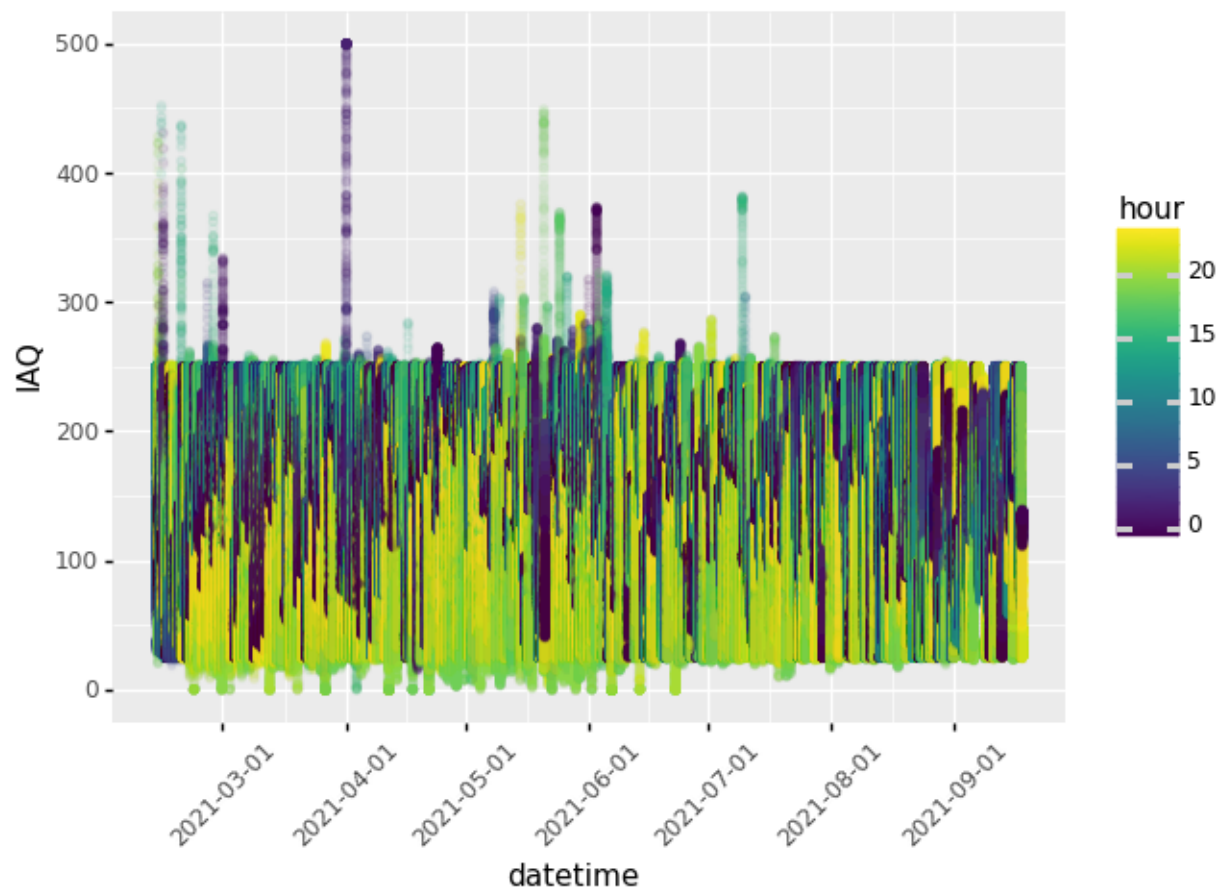
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

```
<ggplot: (8759792563449)>
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
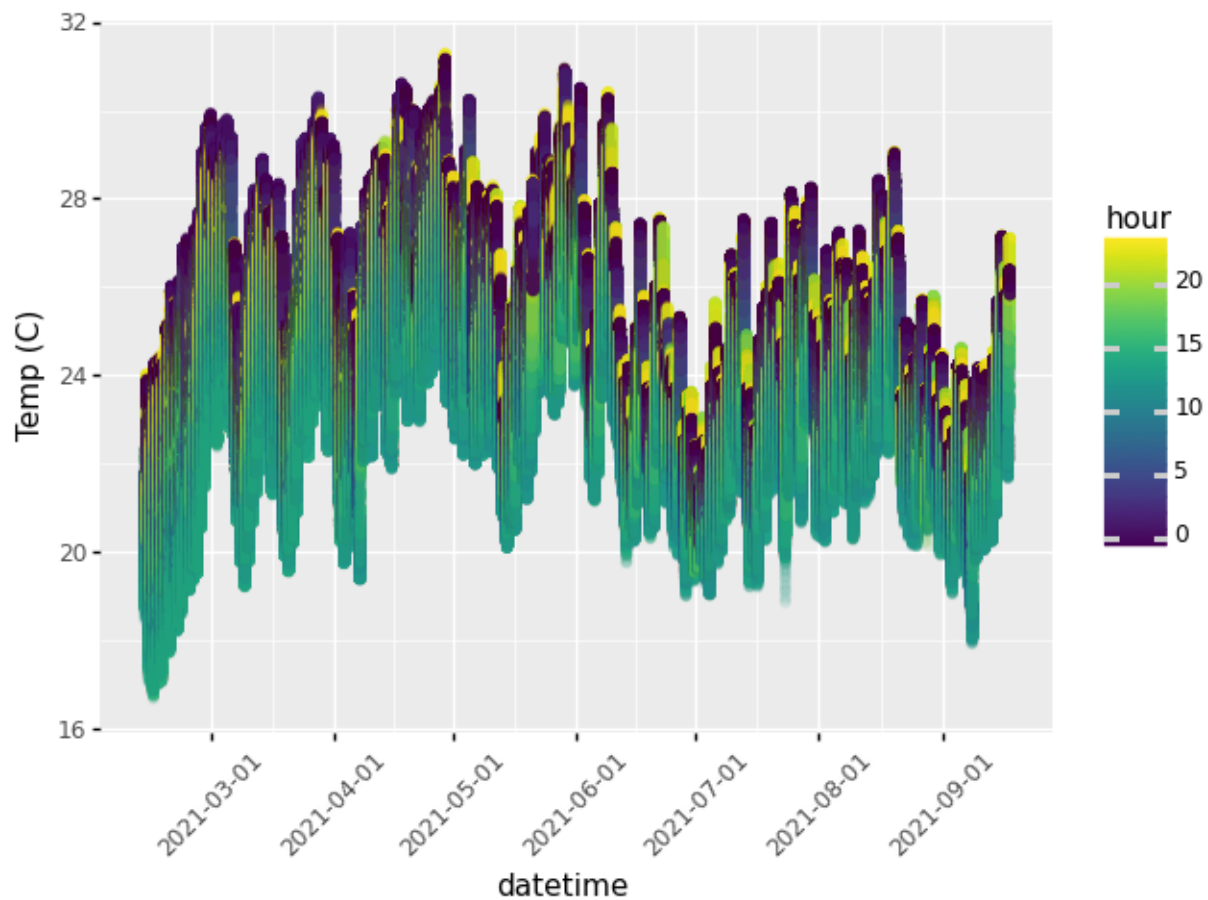`preprocessing_exc_tuple` in IPython 7.17 and above.

```
<ggplot: (8759792572893)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```
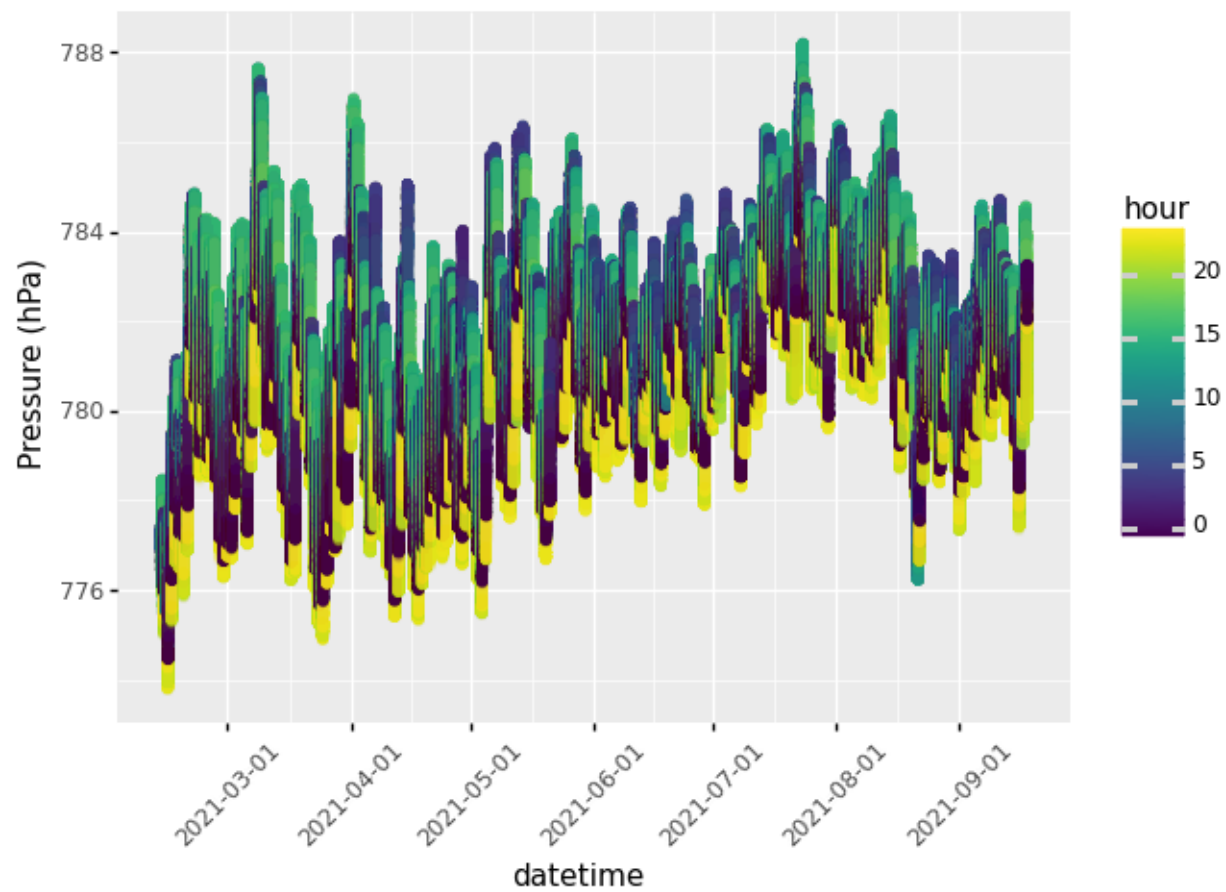
```
<ggplot: (8759792567581)>
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

```
<ggplot: (8759792217949)>
```

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```
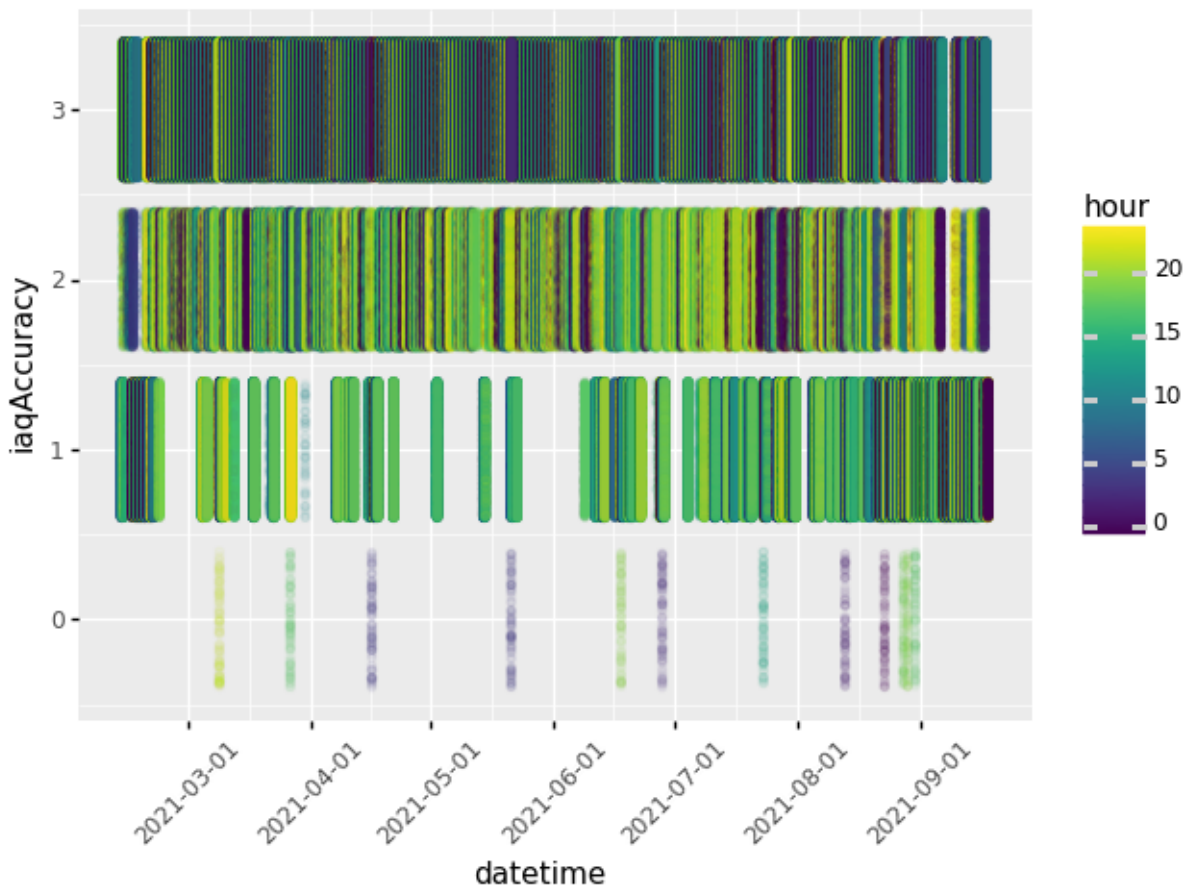
1.3   SINAICA

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

Listing data files from: /home/jaa6766/Documents/jorge3a/itam/deeplearning/dlfinal/data/sinaica/…

Loading pickle prev data…

	Parámetro	Fecha	Valor	Unidad	Estacion
1	CO	2021-01-01	0.600	ppm	Camarones
1	NO	2021-01-01	0.006	ppm	Camarones
1	NO2	2021-01-01	0.029	ppm	Camarones
1	NOx	2021-01-01	0.034	ppm	Camarones
1	O3	2021-01-01	0.011	ppm	Camarones
34	SO2	2021-10-08	0.002	ppm	Merced
35	SO2	2021-10-08	0.001	ppm	Merced
36	SO2	2021-10-08	0.001	ppm	Merced

```
37          SO2 2021-10-08  0.000      ppm      Merced
38          SO2 2021-10-08  0.001      ppm      Merced
```

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

```
    Parámetro        Fecha  Valor Unidad    Estacion
1          CO 2021-01-01  0.600      ppm  Camarones
1          NO 2021-01-01  0.006      ppm  Camarones
1         NO2 2021-01-01  0.029      ppm  Camarones
1         NOx 2021-01-01  0.034      ppm  Camarones
1          O3 2021-01-01  0.011      ppm  Camarones
..         ...        ...    ...       ...        ...
34         SO2 2021-10-08  0.002      ppm      Merced
35         SO2 2021-10-08  0.001      ppm      Merced
36         SO2 2021-10-08  0.001      ppm      Merced
37         SO2 2021-10-08  0.000      ppm      Merced
38         SO2 2021-10-08  0.001      ppm      Merced

[196289 rows x 5 columns]
```

1.3.1 Nearby Air Quality Monitoring Stations

These are the air quality monitoring stations that are close to "Camarones", which is the one nearby to our sensor:

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```
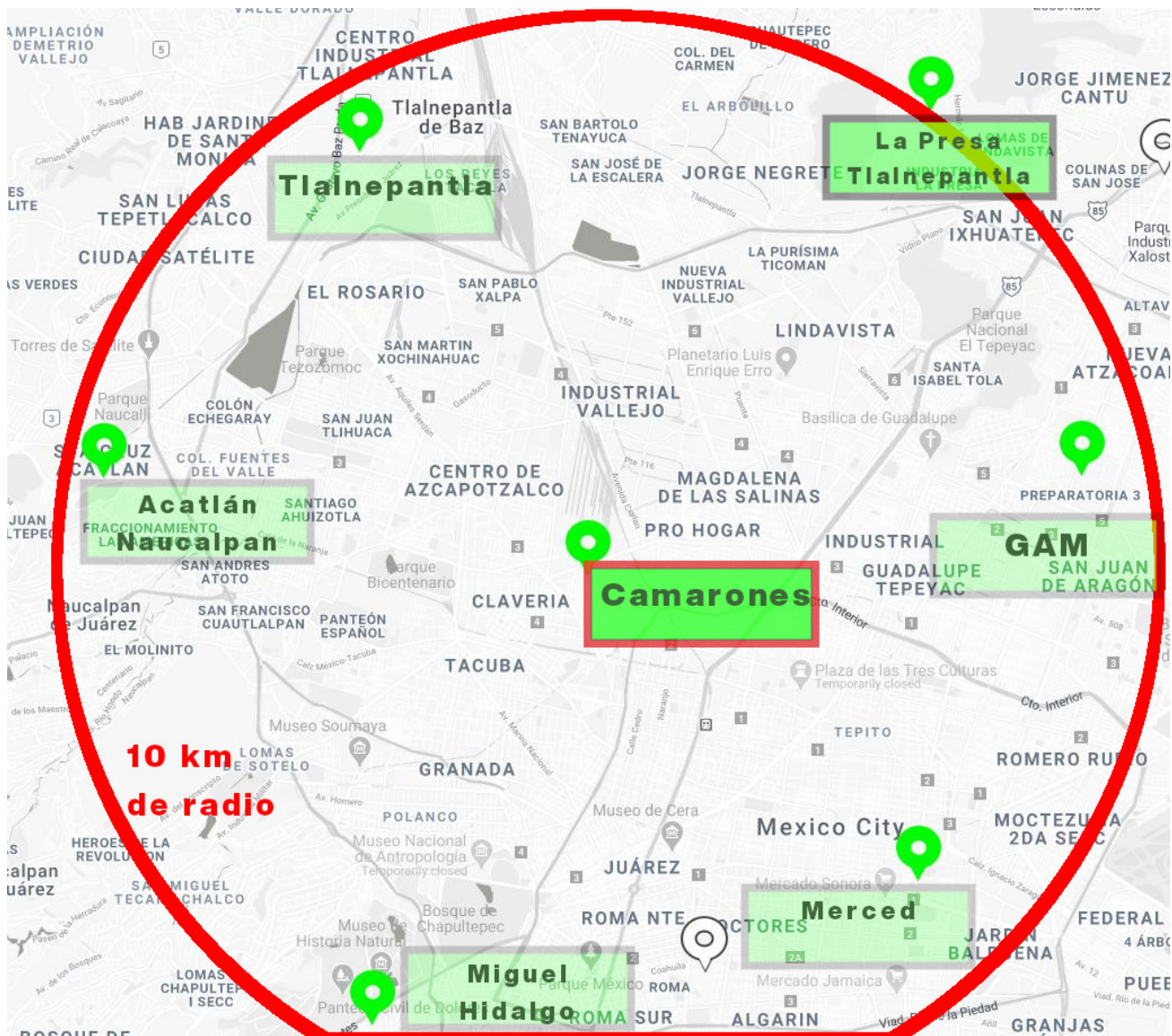
Visualización general de las Variables de Contaminantes

```
<ggplot: (8759791742085)>
```

1.3.2 Camarones Air Quality Monitoring Station

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

Estación Camarones

```
<ggplot: (8759877268541)>
```
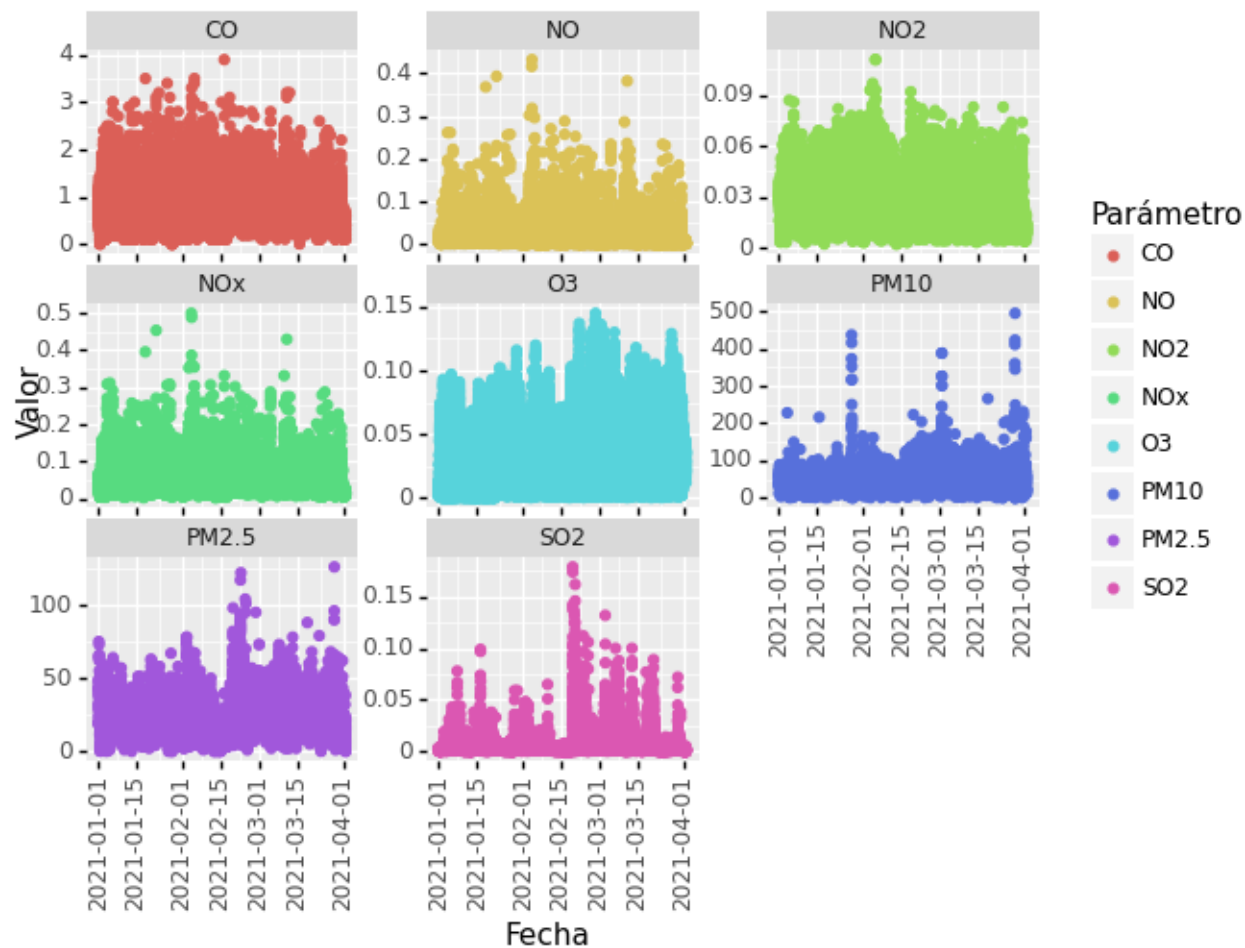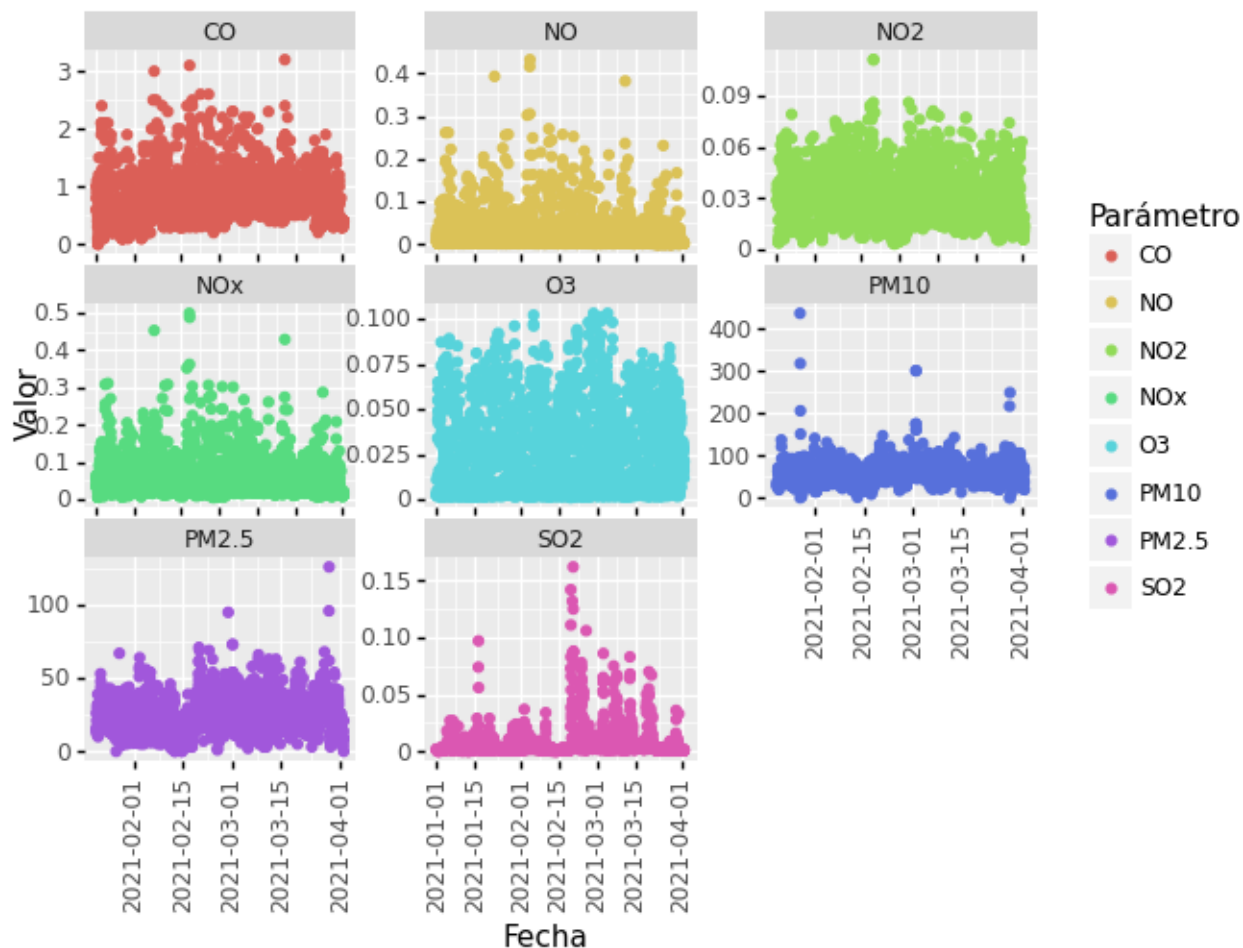
```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

1.4 Sensor Data

These are the hourly averages of the sensor in order to make them match the government air quality monitoring stations that report hourly vs every 3 seconds.

```
   temperature  pressure  humidity  gasResistance   IAQ  iaqAccuracy  \
0        21.54    777.41     43.93         151328  37.5            1
1        21.56    777.41     43.89         152702  35.6            1
2        21.53    777.41     43.97         151328  37.5            1
3        21.51    777.41     44.03         151464  38.5            1
4        21.51    777.41     44.05         152425  36.9            1

                      datetime  year  month  day  hour  minute
0 2021-02-12 06:04:09.089621067  2021      2   12     6       4
1 2021-02-12 06:04:12.087778807  2021      2   12     6       4
2 2021-02-12 06:04:15.072475433  2021      2   12     6       4
```

```
   3 2021-02-12 06:04:18.070170164  2021      2   12    6      4
   4 2021-02-12 06:04:21.061994791  2021      2   12    6      4
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/dplython/dplython.py:196: DeprecationWarning: 'dfilter' is deprecated.
Please use 'sift' instead.

```
         temperature  pressure  humidity  gasResistance    IAQ  iaqAccuracy  \
490473         28.06     780.7     30.37         175863  198.1            3
490474         28.05     780.7     30.38         176417  197.7            3
490475         28.05     780.7     30.41         175313  198.0            3

                         datetime  year  month  day  hour  minute
490473 2021-03-01 06:00:01.807887316  2021      3    1     6       0
490474 2021-03-01 06:00:04.803511858  2021      3    1     6       0
490475 2021-03-01 06:00:07.798833609  2021      3    1     6       0
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

CPU times: user 31.8 s, sys: 7.06 s, total: 38.9 s
Wall time: 38.9 s

```
      year  month  day  hour  temperature     pressure   humidity  \
0     2021      2   12     6    21.557391   777.271496  44.289745
1     2021      2   12     7    21.153699   777.077872  43.183375
2     2021      2   12     8    20.653242   776.620657  42.604564
3     2021      2   12     9    20.406470   776.213214  42.223995
4     2021      2   12    10    20.051380   776.202968  42.269584
...    ...    ...  ...   ...          ...          ...        ...
5223  2021      9   17    21    26.476714   780.191339  50.048186
5224  2021      9   17    22    26.849135   780.496165  50.588394
5225  2021      9   17    23    26.281820   782.067298  54.032219
5226  2021      9   18     0    26.222995   782.853860  55.496814
5227  2021      9   18     1    25.928134   782.818660  56.467943

      gasResistance         IAQ  iaqAccuracy
0      1.439648e+05   90.755292            1
1      1.497397e+05   81.831588            1
2      1.537118e+05   86.220615            1
3      1.491061e+05  138.266030            1
4      1.428894e+05  198.164339            1
...             ...         ...          ...
5223   1.043343e+06   37.047504            1
5224   1.050633e+06   38.850749            1
5225   9.918547e+05   81.164589            1
5226   9.266802e+05  128.776123            1
5227   9.204316e+05  134.620335            1

[5228 rows x 10 columns]
```

Estadísticas de los valores de las lecturas

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-

```
              year       month          day         hour   temperature  \
count   5228.0   5228.000000   5228.000000   5228.000000   5228.000000
mean    2021.0      5.509946     15.633512     11.506121     24.349677
std        0.0      2.093404      8.655932      6.921888      2.489229
min     2021.0      2.000000      1.000000      0.000000     17.282542
25%     2021.0      4.000000      8.000000      6.000000     22.518177
50%     2021.0      6.000000     16.000000     12.000000     24.198175
75%     2021.0      7.000000     23.000000     18.000000     26.116874
max     2021.0      9.000000     31.000000     23.000000     31.066215

            pressure      humidity   gasResistance          IAQ   iaqAccuracy
count    5228.000000   5228.000000    5.228000e+03  5228.000000   5228.000000
mean      781.624137     43.442600    6.950943e+05   157.429759      2.551071
std         2.187106     12.554327    3.111298e+05    69.918421      0.817749
min       774.004780      8.750125    9.540458e+04    22.751331      1.000000
25%       780.216749     32.662378    5.044701e+05    98.951158      3.000000
50%       781.726205     43.999339    6.915713e+05   171.485025      3.000000
75%       783.156527     54.322785    8.638399e+05   219.580799      3.000000
max       787.963968     70.166841    2.716493e+06   255.292928      3.000000
```

1.5 Air Quality Government Data

```
     Parámetro                 Fecha   Valor Unidad        Estacion
1           CO  2021-01-01 00:00:00   0.600    ppm        Camarones
1           NO  2021-01-01 00:00:00   0.006    ppm        Camarones
1          NO2  2021-01-01 00:00:00   0.029    ppm        Camarones
1          NOx  2021-01-01 00:00:00   0.034    ppm        Camarones
1           O3  2021-01-01 00:00:00   0.011    ppm        Camarones
..         ...                   ...     ...    ...              ...
425        NOx  2021-04-01 23:00:00   0.014    ppm      Tlalnepantla
701         O3  2021-04-01 23:00:00   0.029    ppm      Tlalnepantla
681       PM10  2021-04-01 23:00:00  55.000   µg/m³    Tlalnepantla
679      PM2.5  2021-04-01 23:00:00  22.000   µg/m³    Tlalnepantla
696        SO2  2021-04-01 23:00:00   0.002    ppm      Tlalnepantla

[88034 rows x 5 columns]
```

```
                   Fecha  Camarones_CO  Camarones_NO  Camarones_NO2  \
0    2021-01-01 00:00:00           0.6         0.006          0.029
1    2021-01-01 01:00:00           1.0         0.021          0.038
2    2021-01-01 02:00:00           0.8         0.013          0.035
3    2021-01-01 03:00:00           1.0         0.031          0.034
4    2021-01-01 04:00:00           0.6         0.005          0.029
```

```
...              ...          ...       ...        ...        ...
2166 2021-04-01 19:00:00             0.4       0.003      0.011
2167 2021-04-01 20:00:00             0.4       0.002      0.011
2168 2021-04-01 21:00:00             0.4       0.002      0.013
2169 2021-04-01 22:00:00             0.4       0.002      0.019
2170 2021-04-01 23:00:00             0.4       0.001      0.014


      Camarones_NOx  Camarones_O3  Camarones_PM10  Camarones_PM2.5  \
0             0.034         0.011             NaN              NaN
1             0.059         0.002             NaN              NaN
2             0.049         0.003             NaN              NaN
3             0.065         0.002             NaN              NaN
4             0.034         0.005             NaN              NaN
...             ...           ...             ...              ...
2166          0.013         0.016            69.0              7.0
2167          0.012         0.018            71.0              9.0
2168          0.015         0.016            37.0              9.0
2169          0.021         0.012            19.0              0.0
2170          0.015         0.021            61.0             21.0


      Camarones_SO2  FES Acatlán_CO  …  Miguel Hidalgo_O3  \
0             0.002             0.4  …              0.009
1             0.002             0.6  …              0.006
2             0.001             0.9  …              0.003
3             0.001             0.8  …              0.004
4             0.001             1.0  …              0.006
...             ...             ...  …                ...
2166          0.001             0.3  …              0.023
2167          0.001             0.2  …              0.024
2168          0.001             0.2  …              0.023
2169          0.001             0.1  …              0.027
2170          0.001             0.2  …              0.033


      Miguel Hidalgo_SO2  Tlalnepantla_CO  Tlalnepantla_NO  Tlalnepantla_NO2  \
0                  0.003              0.6              NaN             0.030
1                  0.003              0.6              NaN             0.026
2                  0.002              0.7              NaN             0.032
3                  0.002              0.7              NaN             0.033
4                  0.002              0.7              NaN             0.032
...                  ...              ...              ...               ...
2166               0.000              0.3            0.009             0.017
2167               0.000              0.4            0.004             0.015
2168               0.000              0.3            0.002             0.014
2169               0.000              0.3            0.002             0.011
2170               0.001              0.3            0.001             0.014


      Tlalnepantla_NOx  Tlalnepantla_O3  Tlalnepantla_PM10  \
0                0.034            0.012               37.0
1                0.029            0.013               42.0
2                0.036            0.006               58.0
3                0.039            0.004               59.0
4                0.038            0.004               64.0
...                ...              ...                ...
2166             0.025            0.017               52.0
2167             0.019            0.020               22.0
2168             0.017            0.021               21.0
2169             0.013            0.022               11.0
2170             0.014            0.029               55.0
```

```
      Tlalnepantla_PM2.5  Tlalnepantla_SO2
0                   19.0             0.002
1                   29.0             0.003
2                   43.0             0.002
3                   41.0             0.002
4                   46.0             0.002
...                  ...               ...
2166                24.0             0.003
2167                10.0             0.005
2168                14.0             0.002
2169                 7.0             0.001
2170                22.0             0.002

[2171 rows x 45 columns]

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

	Estacion	count	mean
Camarones_CO	Camarones_CO	2050.0	0.771805
Camarones_NO	Camarones_NO	2042.0	0.025181
Camarones_NO2	Camarones_NO2	2042.0	0.031787
Camarones_NOx	Camarones_NOx	2042.0	0.056961
Camarones_O3	Camarones_O3	2042.0	0.025693
Camarones_PM10	Camarones_PM10	1616.0	58.066213
Camarones_PM2.5	Camarones_PM2.5	1604.0	25.158978
Camarones_SO2	Camarones_SO2	2042.0	0.006598
FES Acatlán_CO	FES Acatlán_CO	2020.0	0.634901
FES Acatlán_NO	FES Acatlán_NO	1304.0	0.014301
FES Acatlán_NO2	FES Acatlán_NO2	2015.0	0.025607
FES Acatlán_NOx	FES Acatlán_NOx	2019.0	0.041234
FES Acatlán_O3	FES Acatlán_O3	2015.0	0.033841
FES Acatlán_PM10	FES Acatlán_PM10	1991.0	47.245605
FES Acatlán_SO2	FES Acatlán_SO2	2015.0	0.006546
Gustavo A. Madero_NO2	Gustavo A. Madero_NO2	2085.0	0.025859
Gustavo A. Madero_O3	Gustavo A. Madero_O3	2076.0	0.031555
Gustavo A. Madero_PM10	Gustavo A. Madero_PM10	2035.0	52.755283
Gustavo A. Madero_PM2.5	Gustavo A. Madero_PM2.5	2026.0	23.785291
La Presa_CO	La Presa_CO	2034.0	0.933628
La Presa_O3	La Presa_O3	1863.0	0.029086
La Presa_SO2	La Presa_SO2	2017.0	0.005163
Merced_CO	Merced_CO	2087.0	1.151749
Merced_NO	Merced_NO	1381.0	0.021361
Merced_NO2	Merced_NO2	2070.0	0.032723
Merced_NOx	Merced_NOx	2070.0	0.056423
Merced_O3	Merced_O3	2076.0	0.028389
Merced_PM10	Merced_PM10	2142.0	54.485061
Merced_PM2.5	Merced_PM2.5	2137.0	26.730463
Merced_SO2	Merced_SO2	2094.0	0.006420
Miguel Hidalgo_CO	Miguel Hidalgo_CO	2090.0	0.544785
Miguel Hidalgo_NO	Miguel Hidalgo_NO	2077.0	0.021710
Miguel Hidalgo_NO2	Miguel Hidalgo_NO2	2078.0	0.029400
Miguel Hidalgo_NOx	Miguel Hidalgo_NOx	2078.0	0.051113
Miguel Hidalgo_O3	Miguel Hidalgo_O3	2082.0	0.033624
Miguel Hidalgo_SO2	Miguel Hidalgo_SO2	2081.0	0.005228
Tlalnepantla_CO	Tlalnepantla_CO	2023.0	0.740287
Tlalnepantla_NO	Tlalnepantla_NO	1039.0	0.020687

```
Tlalnepantla_NO2                   Tlalnepantla_NO2    1751.0    0.031346
Tlalnepantla_NOx                   Tlalnepantla_NOx    1752.0    0.052864
Tlalnepantla_O3                     Tlalnepantla_O3    2077.0    0.027766
Tlalnepantla_PM10               Tlalnepantla_PM10    1989.0   48.649573
Tlalnepantla_PM2.5             Tlalnepantla_PM2.5    1973.0   22.273188
Tlalnepantla_SO2                   Tlalnepantla_SO2    2066.0    0.008569
Fecha                                         Fecha       NaN         NaN
```

	std	min	25%	50%	75%	max	\
Camarones_CO	0.430131	0.000	0.500	0.700	1.0000	3.200	
Camarones_NO	0.045487	0.000	0.002	0.006	0.0260	0.432	
Camarones_NO2	0.015682	0.003	0.020	0.031	0.0420	0.111	
Camarones_NOx	0.054418	0.004	0.022	0.039	0.0710	0.499	
Camarones_O3	0.024107	0.001	0.004	0.018	0.0420	0.103	
Camarones_PM10	26.906212	0.000	41.000	55.000	71.0000	437.000	
Camarones_PM2.5	12.946419	0.000	16.000	24.000	33.0000	126.000	
Camarones_SO2	0.013053	0.000	0.002	0.003	0.0050	0.162	
FES Acatlán_CO	0.370808	0.100	0.400	0.500	0.8000	2.900	
FES Acatlán_NO	0.024926	0.000	0.002	0.004	0.0130	0.215	
FES Acatlán_NO2	0.014263	0.002	0.015	0.023	0.0330	0.092	
FES Acatlán_NOx	0.035480	0.002	0.018	0.028	0.0520	0.260	
FES Acatlán_O3	0.026037	0.003	0.013	0.027	0.0480	0.137	
FES Acatlán_PM10	32.511497	0.000	27.000	41.000	61.0000	388.000	
FES Acatlán_SO2	0.010034	0.000	0.002	0.004	0.0070	0.136	
Gustavo A. Madero_NO2	0.014361	0.003	0.013	0.026	0.0360	0.081	
Gustavo A. Madero_O3	0.029610	0.001	0.004	0.023	0.0510	0.128	
Gustavo A. Madero_PM10	27.911189	0.000	34.500	50.000	67.0000	495.000	
Gustavo A. Madero_PM2.5	13.182547	0.000	14.000	22.000	31.0000	117.000	
La Presa_CO	0.513861	0.100	0.600	0.800	1.1000	3.500	
La Presa_O3	0.025911	0.001	0.006	0.024	0.0455	0.118	
La Presa_SO2	0.009749	0.000	0.001	0.002	0.0050	0.118	
Merced_CO	0.413130	0.500	0.900	1.000	1.3000	3.900	
Merced_NO	0.034552	0.000	0.003	0.008	0.0220	0.318	
Merced_NO2	0.014052	0.005	0.022	0.032	0.0410	0.087	
Merced_NOx	0.044741	0.006	0.027	0.042	0.0710	0.386	
Merced_O3	0.029109	0.000	0.003	0.018	0.0460	0.140	
Merced_PM10	23.727621	0.000	39.000	53.000	67.0000	411.000	
Merced_PM2.5	12.805604	0.000	18.000	25.000	33.0000	122.000	
Merced_SO2	0.010345	0.000	0.002	0.003	0.0060	0.146	
Miguel Hidalgo_CO	0.343697	0.000	0.300	0.500	0.7000	2.600	
Miguel Hidalgo_NO	0.038798	0.000	0.002	0.005	0.0220	0.368	
Miguel Hidalgo_NO2	0.013166	0.004	0.019	0.028	0.0390	0.086	
Miguel Hidalgo_NOx	0.046729	0.005	0.022	0.034	0.0620	0.395	
Miguel Hidalgo_O3	0.028588	0.002	0.009	0.027	0.0500	0.145	
Miguel Hidalgo_SO2	0.008925	0.000	0.001	0.002	0.0050	0.099	
Tlalnepantla_CO	0.360615	0.100	0.500	0.600	0.9000	2.900	
Tlalnepantla_NO	0.031055	0.000	0.003	0.007	0.0220	0.219	
Tlalnepantla_NO2	0.014525	0.004	0.021	0.030	0.0390	0.097	
Tlalnepantla_NOx	0.040767	0.005	0.026	0.039	0.0660	0.274	
Tlalnepantla_O3	0.025335	0.000	0.007	0.020	0.0430	0.125	
Tlalnepantla_PM10	27.959538	0.000	33.000	45.000	59.0000	423.000	
Tlalnepantla_PM2.5	12.075484	0.000	14.000	21.000	29.0000	90.000	
Tlalnepantla_SO2	0.014821	0.001	0.002	0.004	0.0080	0.179	
Fecha	NaN	NaN	NaN	NaN	NaN	NaN	

```
                     NAs
Camarones_CO        121.0
Camarones_NO        129.0
Camarones_NO2       129.0
```

```
Camarones_NOx           129.0
Camarones_O3            129.0
Camarones_PM10          555.0
Camarones_PM2.5         567.0
Camarones_SO2           129.0
FES Acatlán_CO          151.0
FES Acatlán_NO          867.0
FES Acatlán_NO2         156.0
FES Acatlán_NOx         152.0
FES Acatlán_O3          156.0
FES Acatlán_PM10        180.0
FES Acatlán_SO2         156.0
Gustavo A. Madero_NO2    86.0
Gustavo A. Madero_O3     95.0
Gustavo A. Madero_PM10  136.0
Gustavo A. Madero_PM2.5 145.0
La Presa_CO             137.0
La Presa_O3             308.0
La Presa_SO2            154.0
Merced_CO                84.0
Merced_NO               790.0
Merced_NO2              101.0
Merced_NOx              101.0
Merced_O3                95.0
Merced_PM10              29.0
Merced_PM2.5             34.0
Merced_SO2               77.0
Miguel Hidalgo_CO        81.0
Miguel Hidalgo_NO        94.0
Miguel Hidalgo_NO2       93.0
Miguel Hidalgo_NOx       93.0
Miguel Hidalgo_O3        89.0
Miguel Hidalgo_SO2       90.0
Tlalnepantla_CO         148.0
Tlalnepantla_NO        1132.0
Tlalnepantla_NO2        420.0
Tlalnepantla_NOx        419.0
Tlalnepantla_O3          94.0
Tlalnepantla_PM10       182.0
Tlalnepantla_PM2.5      198.0
Tlalnepantla_SO2        105.0
Fecha                     0.0
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will not call `transform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any exception that happen during thetransform in `preprocessing_exc_tuple` in IPython 7.17 and above.
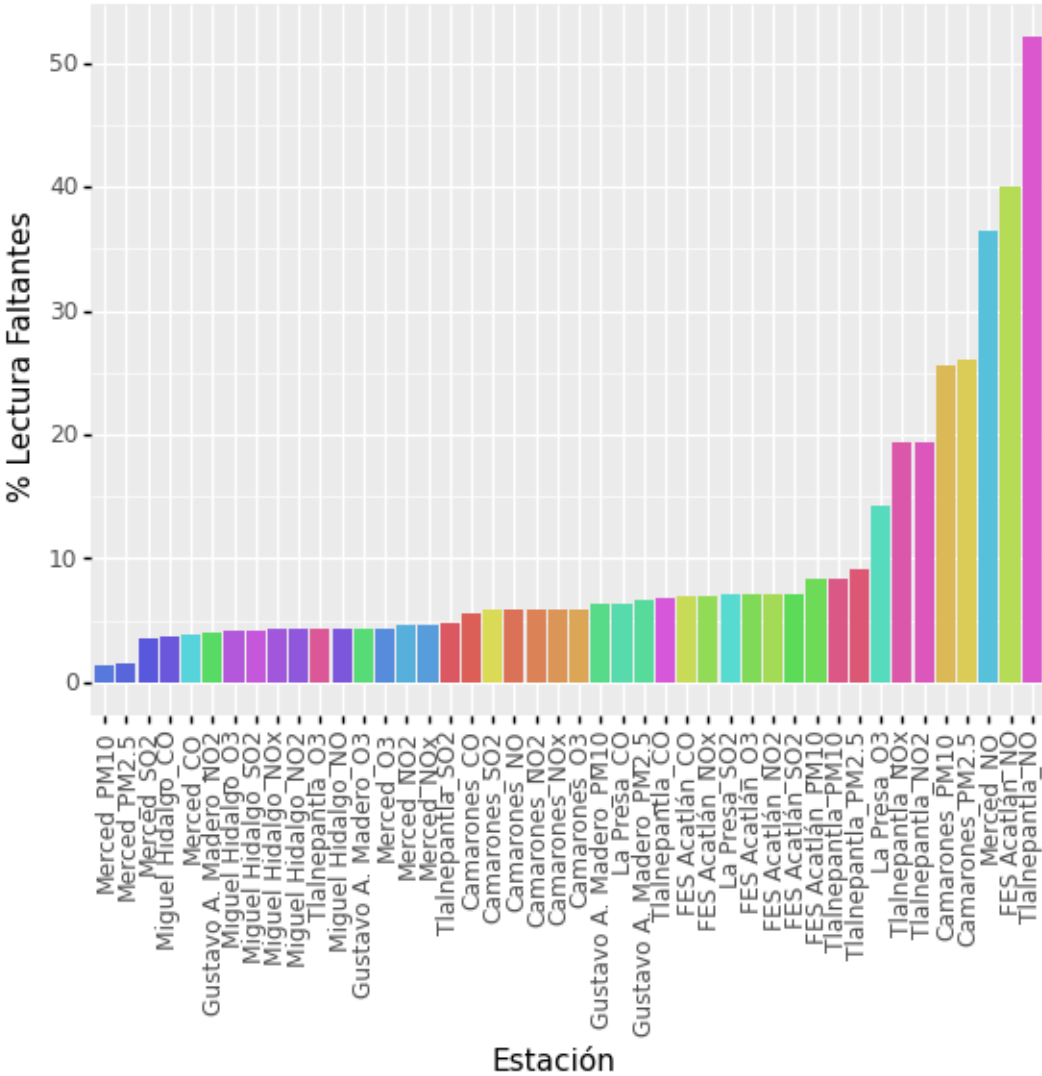/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-packages/ipykernel_launcher.py:8: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

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

	Estacion	NAs
0	Camarones_CO	Camarones_CO	121.0
1	Camarones_NO	Camarones_NO	129.0
2	Camarones_NO2	Camarones_NO2	129.0
3	Camarones_NOx	Camarones_NOx	129.0

```
4              Camarones_O3         Camarones_O3   129.0
5            Camarones_PM10       Camarones_PM10   555.0
6           Camarones_PM2.5      Camarones_PM2.5   567.0
7             Camarones_SO2        Camarones_SO2   129.0
8             FES Acatlán_CO       FES Acatlán_CO   151.0
9             FES Acatlán_NO       FES Acatlán_NO   867.0
10           FES Acatlán_NO2      FES Acatlán_NO2   156.0
11           FES Acatlán_NOx      FES Acatlán_NOx   152.0
12            FES Acatlán_O3       FES Acatlán_O3   156.0
13          FES Acatlán_PM10     FES Acatlán_PM10   180.0
14           FES Acatlán_SO2      FES Acatlán_SO2   156.0
15      Gustavo A. Madero_NO2  Gustavo A. Madero_NO2    86.0
16       Gustavo A. Madero_O3   Gustavo A. Madero_O3    95.0
17     Gustavo A. Madero_PM10  Gustavo A. Madero_PM10   136.0
18    Gustavo A. Madero_PM2.5  Gustavo A. Madero_PM2.5   145.0
19               La Presa_CO           La Presa_CO   137.0
20               La Presa_O3           La Presa_O3   308.0
21              La Presa_SO2          La Presa_SO2   154.0
22                 Merced_CO             Merced_CO    84.0
23                 Merced_NO             Merced_NO   790.0
24                Merced_NO2            Merced_NO2   101.0
25                Merced_NOx            Merced_NOx   101.0
26                 Merced_O3             Merced_O3    95.0
27               Merced_PM10           Merced_PM10    29.0
28              Merced_PM2.5          Merced_PM2.5    34.0
29                Merced_SO2            Merced_SO2    77.0
30          Miguel Hidalgo_CO     Miguel Hidalgo_CO    81.0
31          Miguel Hidalgo_NO     Miguel Hidalgo_NO    94.0
32         Miguel Hidalgo_NO2    Miguel Hidalgo_NO2    93.0
33         Miguel Hidalgo_NOx    Miguel Hidalgo_NOx    93.0
34          Miguel Hidalgo_O3     Miguel Hidalgo_O3    89.0
35         Miguel Hidalgo_SO2    Miguel Hidalgo_SO2    90.0
36            Tlalnepantla_CO       Tlalnepantla_CO   148.0
37            Tlalnepantla_NO       Tlalnepantla_NO  1132.0
38           Tlalnepantla_NO2      Tlalnepantla_NO2   420.0
39           Tlalnepantla_NOx      Tlalnepantla_NOx   419.0
40            Tlalnepantla_O3       Tlalnepantla_O3    94.0
41          Tlalnepantla_PM10     Tlalnepantla_PM10   182.0
42         Tlalnepantla_PM2.5    Tlalnepantla_PM2.5   198.0
43           Tlalnepantla_SO2      Tlalnepantla_SO2   105.0
```

Histograma de Lecturas Faltantes
por Contaminante-Estacion de Monitoreo

<ggplot: (8759875176429)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
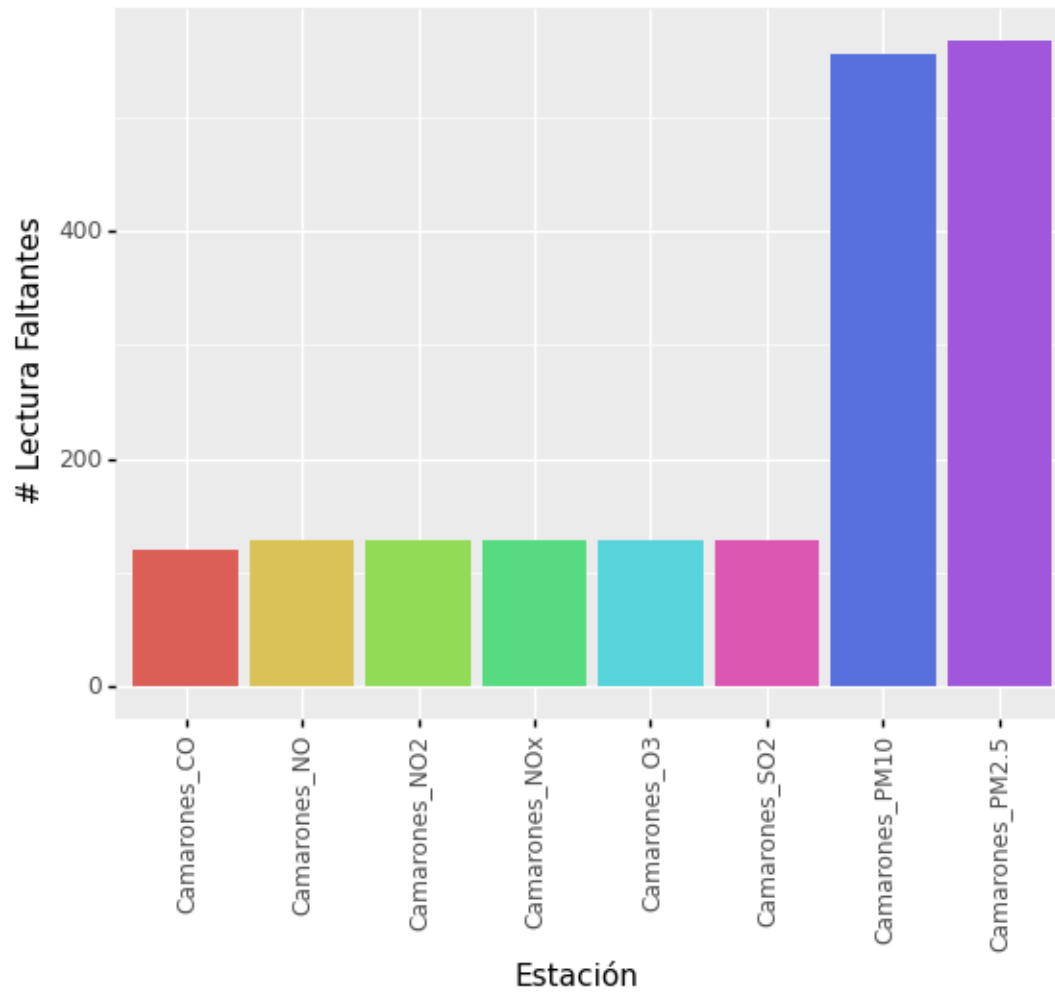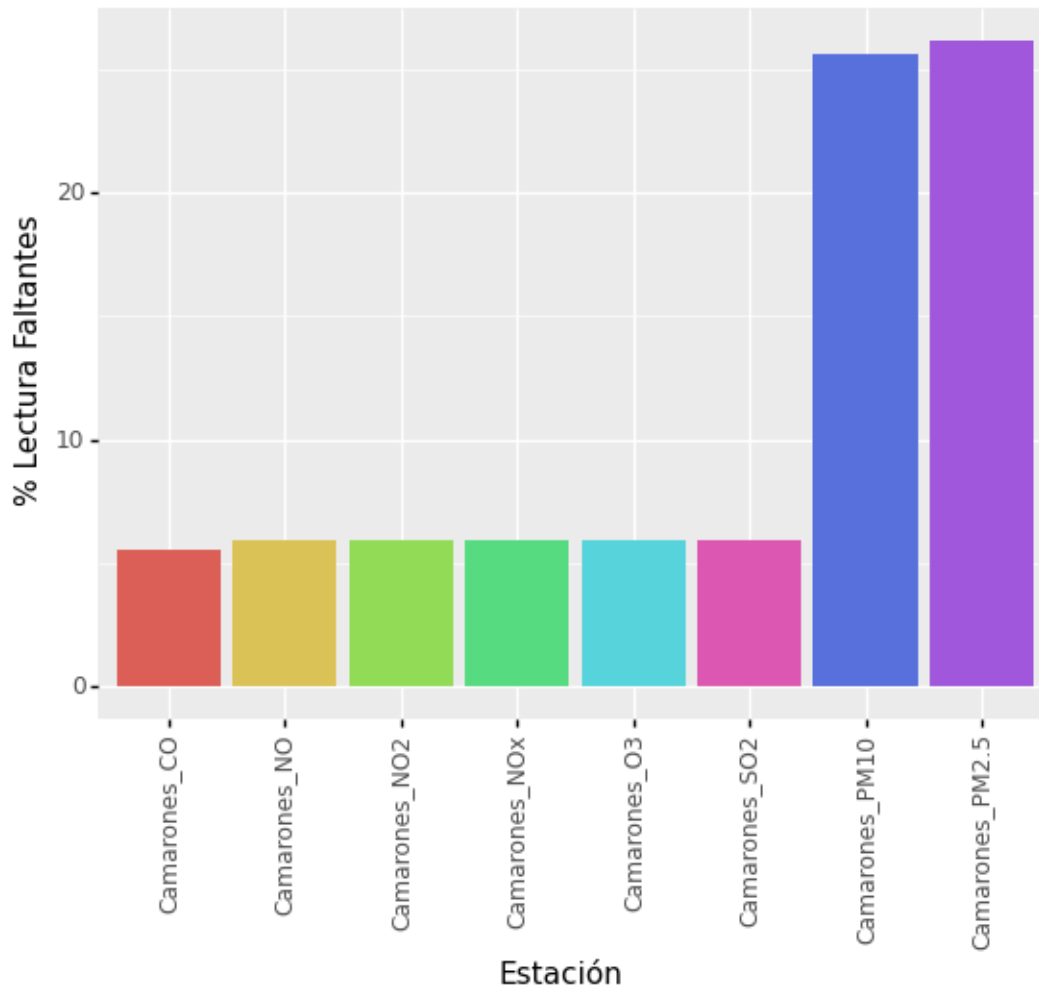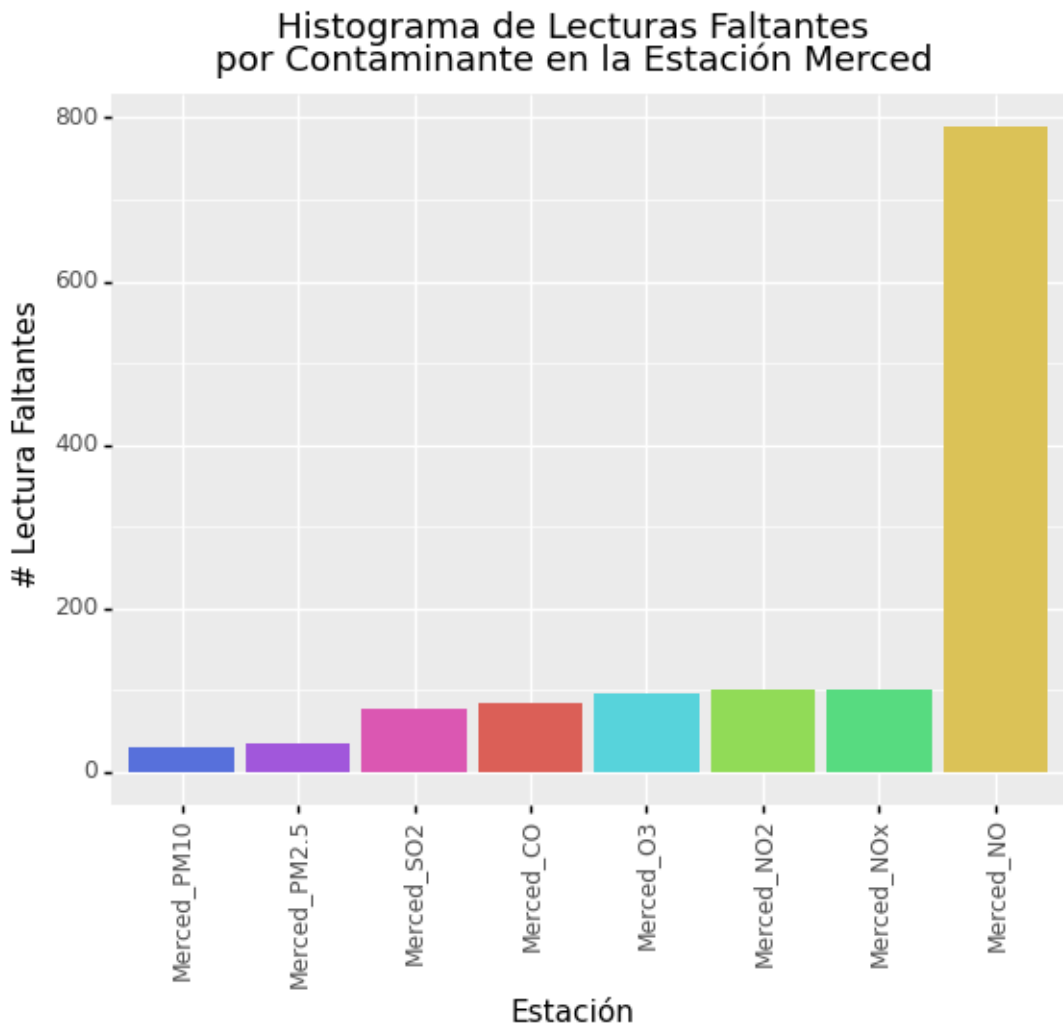`preprocessing_exc_tuple` in IPython 7.17 and above.

Porcentaje de Lecturas Faltantes
por Contaminante-Estacion de Monitoreo

<ggplot: (8759685738189)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

Histograma de Lecturas Faltantes
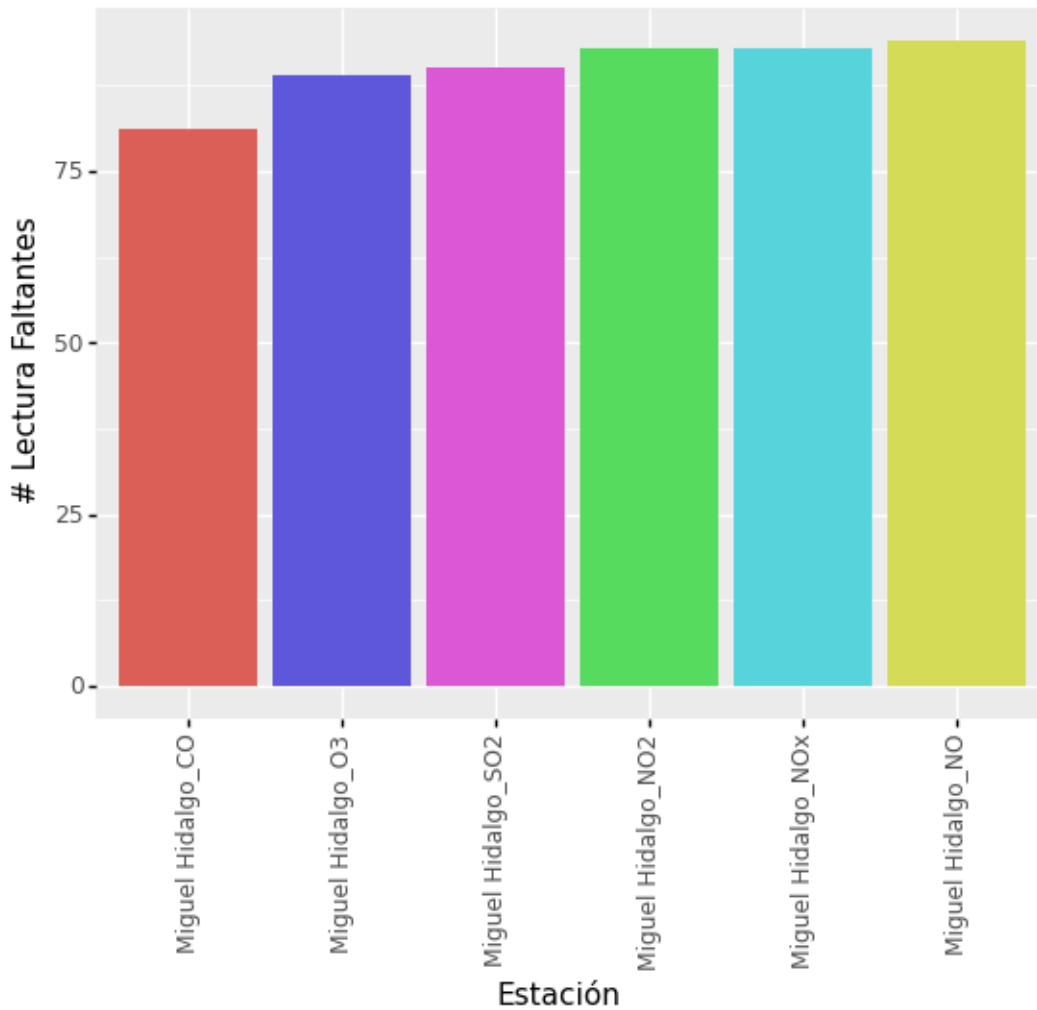por Contaminante en la Estación Camarones

```
<ggplot: (8759791694833)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

Porcentaje de Lecturas Faltantes
por Contaminante en Camarones

```
<ggplot: (8759792099749)>
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

```
                Estacion     NAs
0     Camarones_CO     Camarones_CO   121.0
1     Camarones_NO     Camarones_NO   129.0
2    Camarones_NO2    Camarones_NO2   129.0
3    Camarones_NOx    Camarones_NOx   129.0
4     Camarones_O3     Camarones_O3   129.0
5   Camarones_PM10   Camarones_PM10   555.0
6  Camarones_PM2.5  Camarones_PM2.5   567.0
7    Camarones_SO2    Camarones_SO2   129.0
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

```
         Estacion     NAs
22     Merced_CO     Merced_CO    84.0
23     Merced_NO     Merced_NO   790.0
24     Merced_NO2    Merced_NO2  101.0
25     Merced_NOx    Merced_NOx  101.0
26     Merced_O3     Merced_O3    95.0
27    Merced_PM10   Merced_PM10   29.0
28   Merced_PM2.5  Merced_PM2.5   34.0
29     Merced_SO2    Merced_SO2   77.0
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

Histograma de Lecturas Faltantes por Contaminante en la Estación Merced
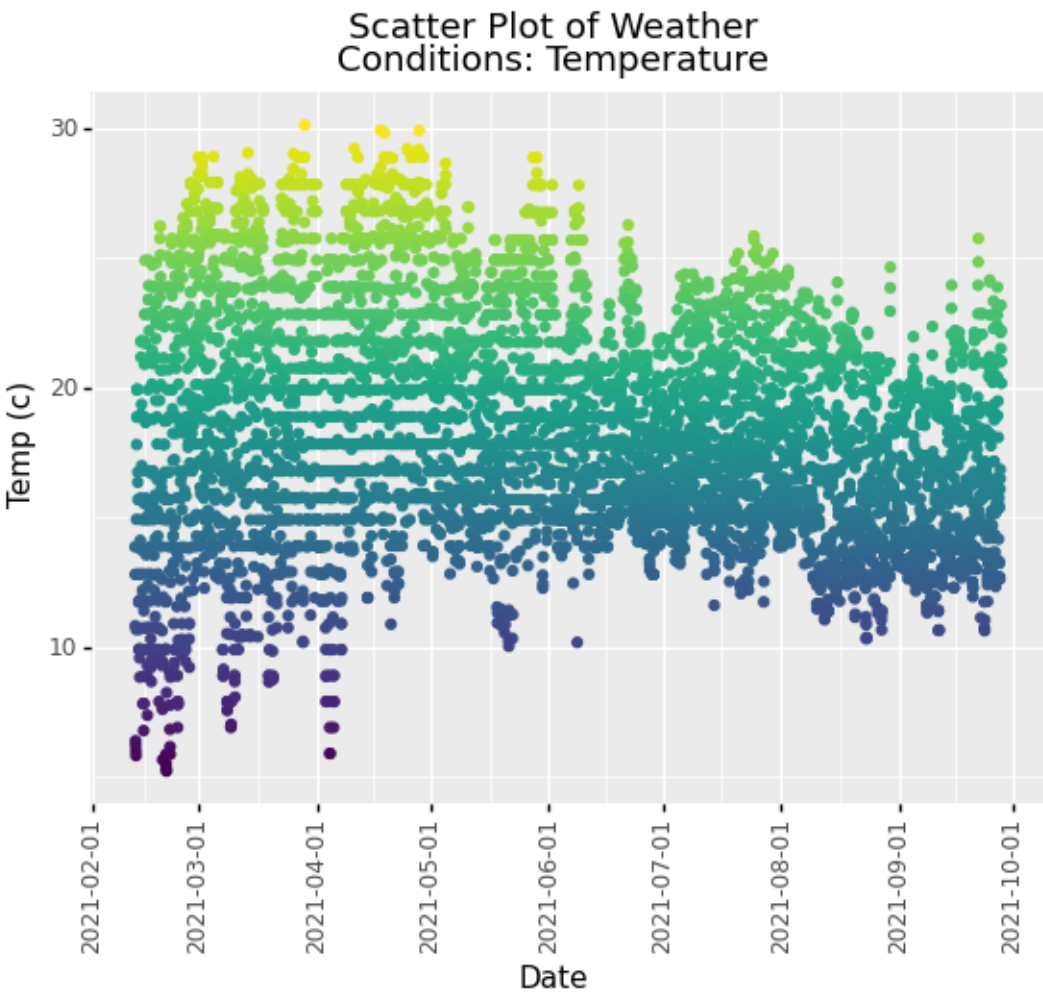
<ggplot: (8759792245861)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

Histograma de Lecturas Faltantes
por Contaminante en la Estación Miguel Hidalgo

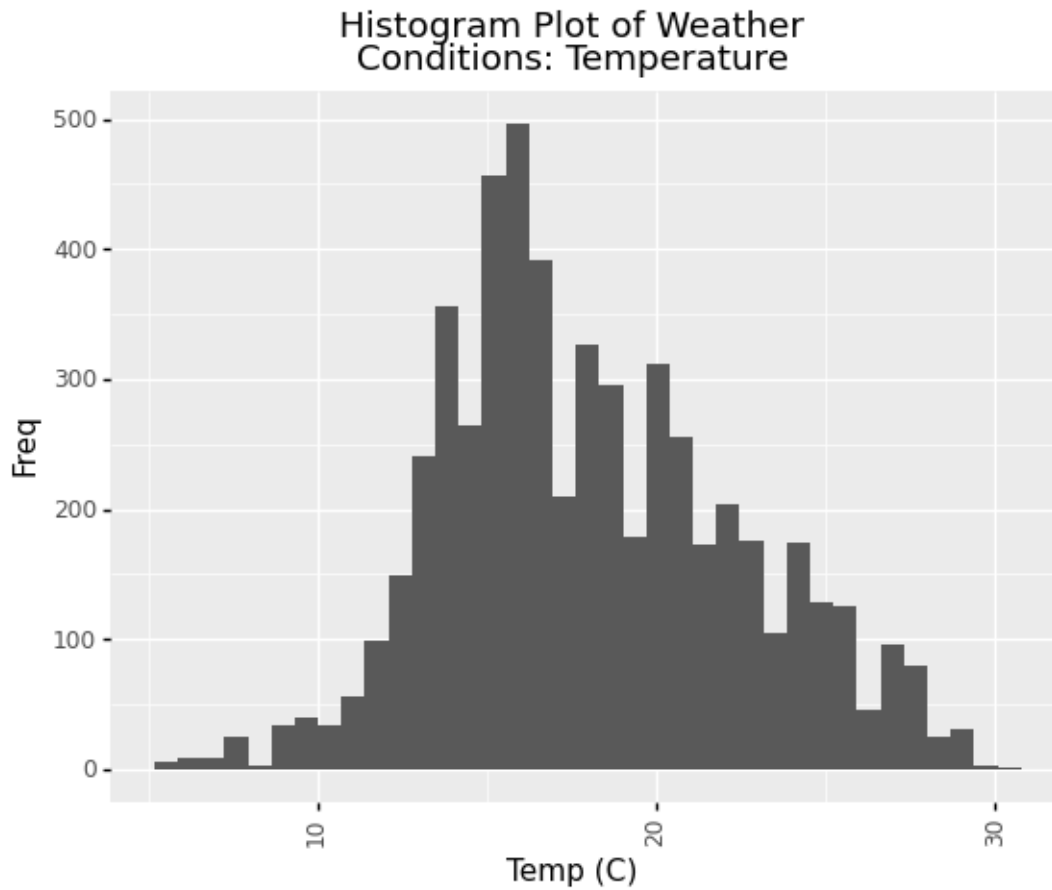```
<ggplot: (8759791713777)>
```

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

1.6 Weather Data

OpenWeatherMap Data

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```
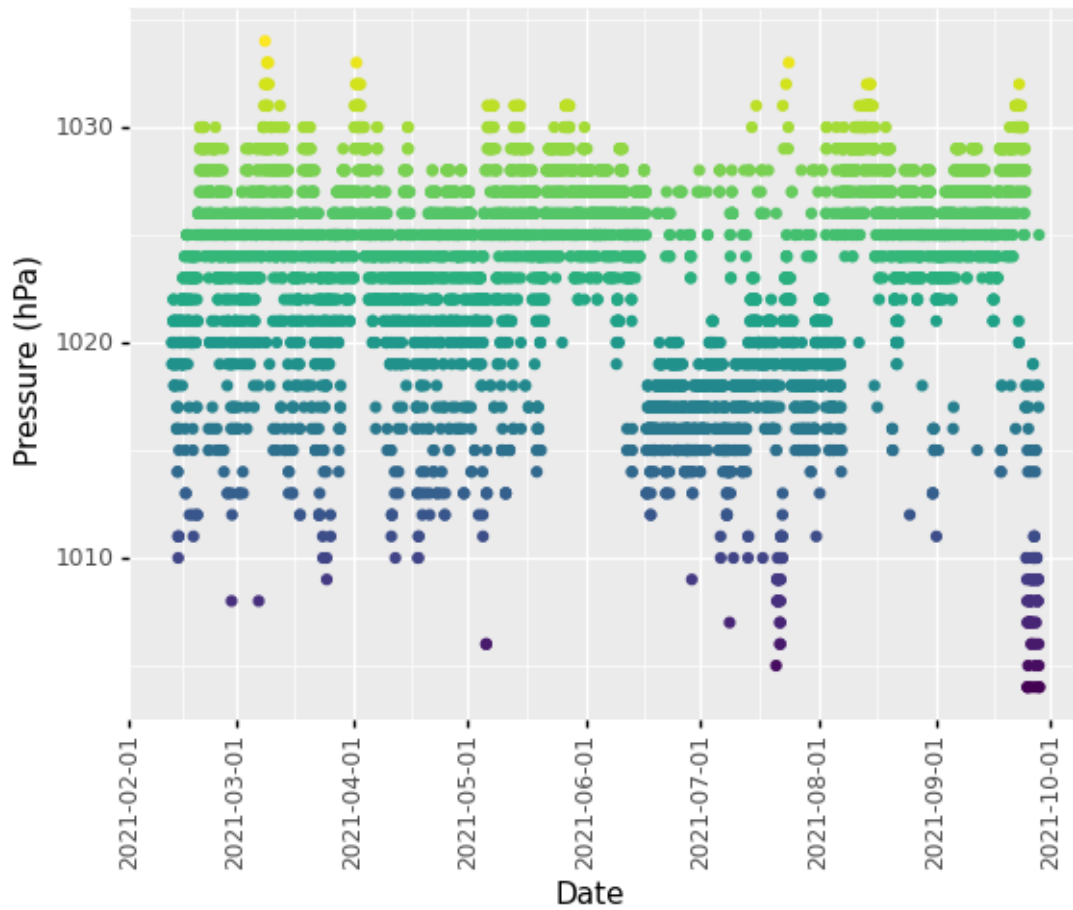
```
CPU times: user 1.06 s, sys: 118 ms, total: 1.17 s
Wall time: 1.17 s
```

```
                   dt   temp  feels_like  temp_min  temp_max  pressure  \
0  2021-02-12 07:00:00  13.87       12.46      5.21     13.92      1020
1  2021-02-12 08:00:00  12.81       11.37      4.21     12.92      1020

   humidity  wind_speed  wind_deg  rain_1h  rain_3h  clouds_all  weather_id  \
0        44         0.0         0      0.0      0.0           1         800
1        47         0.0         0      0.0      0.0           1         800

  weather_main
0        Clear
1        Clear
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
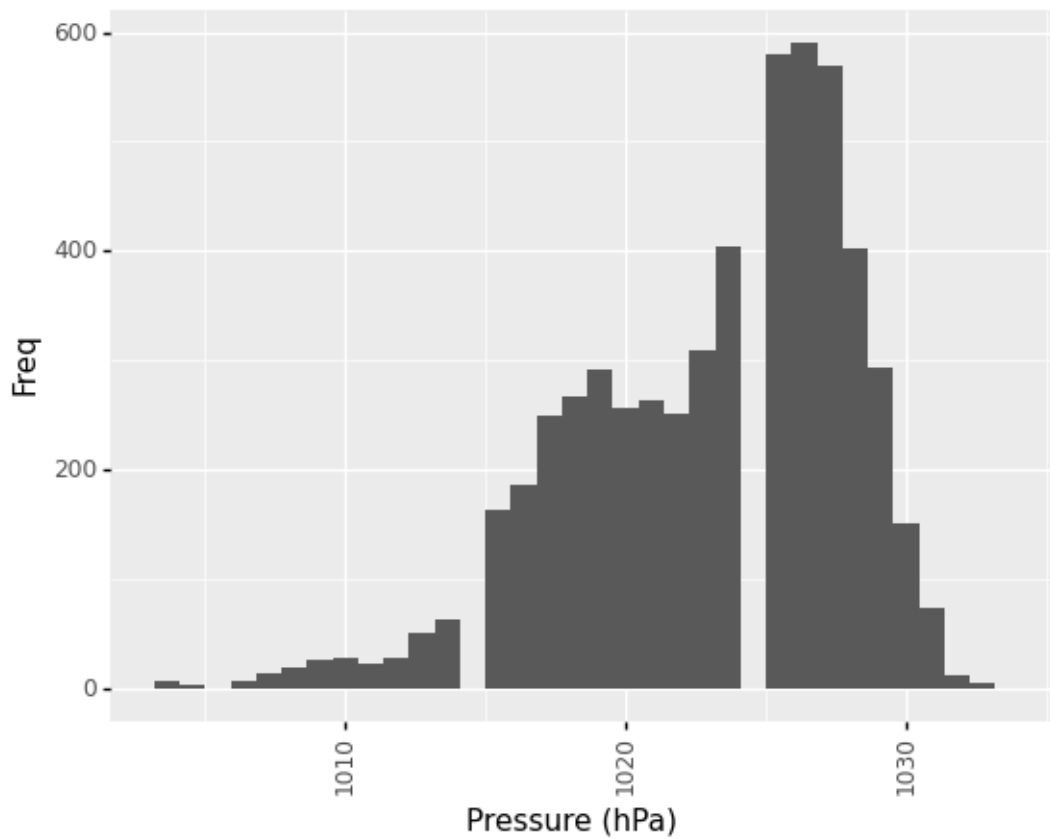`preprocessing_exc_tuple` in IPython 7.17 and above.

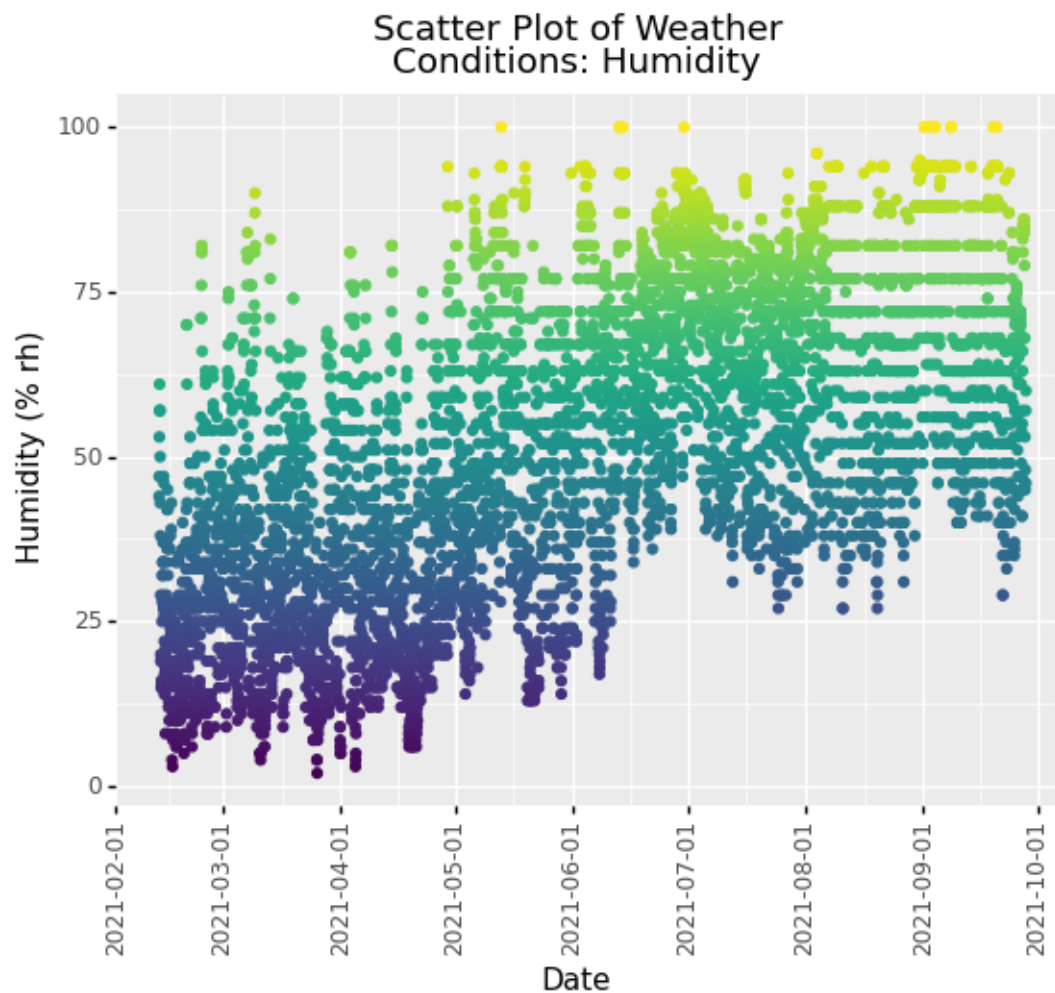Scatter Plot of Weather Conditions: Temperature

<ggplot: (8759871460657)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/plotnine/stats/stat_bin.py:93: PlotnineWarning: 'stat_bin()' using
'bins = 37'. Pick better value with 'binwidth'.
```


Histogram Plot of Weather Conditions: Temperature
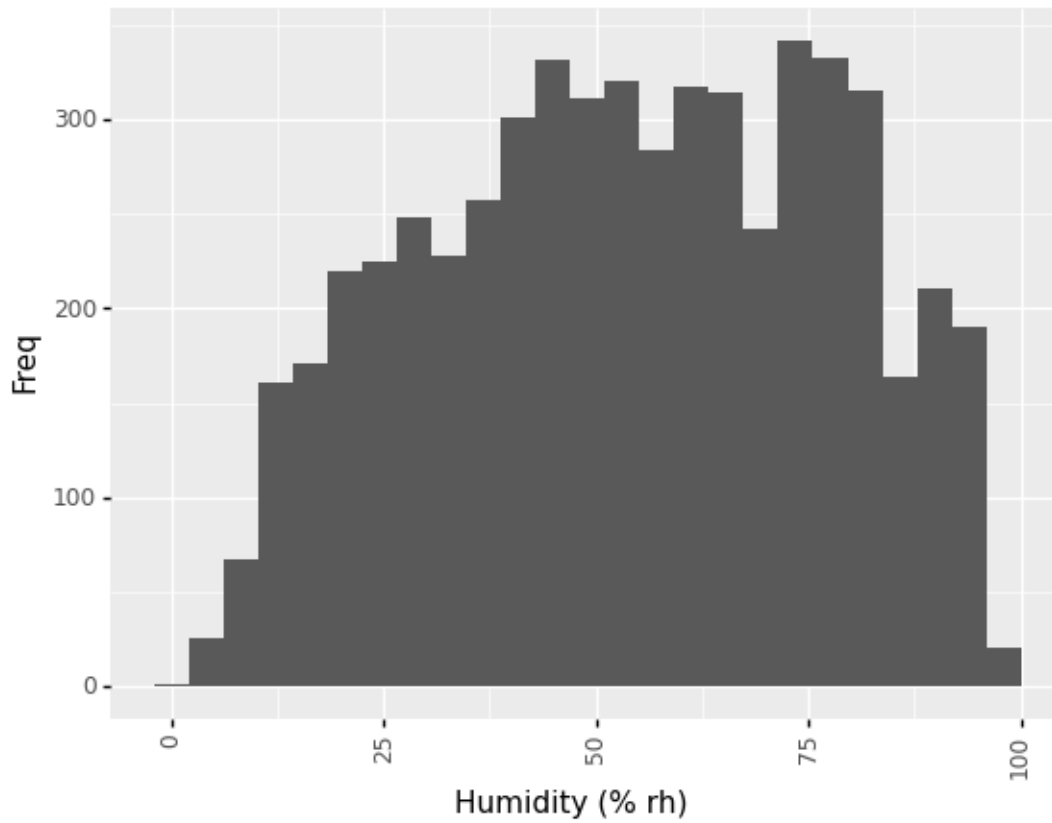
```
<ggplot: (8759871700921)>
```

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```

Scatter Plot of Weather Conditions: Pressure

```
<ggplot: (8759871516885)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/plotnine/stats/stat_bin.py:93: PlotnineWarning: 'stat_bin()' using
'bins = 34'. Pick better value with 'binwidth'.
```
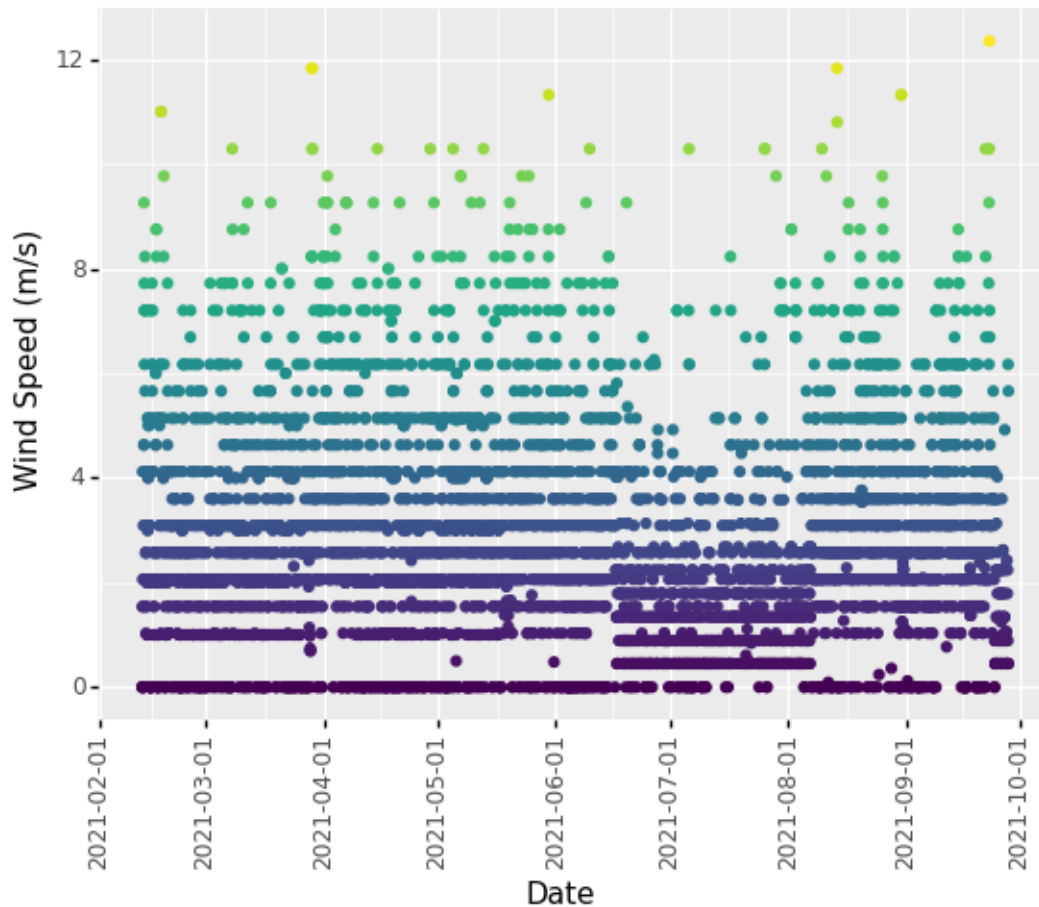
Histogram Plot of Weather
Conditions: Pressure


```
<ggplot: (8759871192141)>
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.

Scatter Plot of Weather Conditions: Humidity

```
<ggplot: (8759872460273)>
```

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/plotnine/stats/stat_bin.py:93: PlotnineWarning: 'stat_bin()' using
'bins = 25'. Pick better value with 'binwidth'.

Histogram Plot of Weather
Conditions: Humidity

```
 <ggplot: (8759872421517)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
```
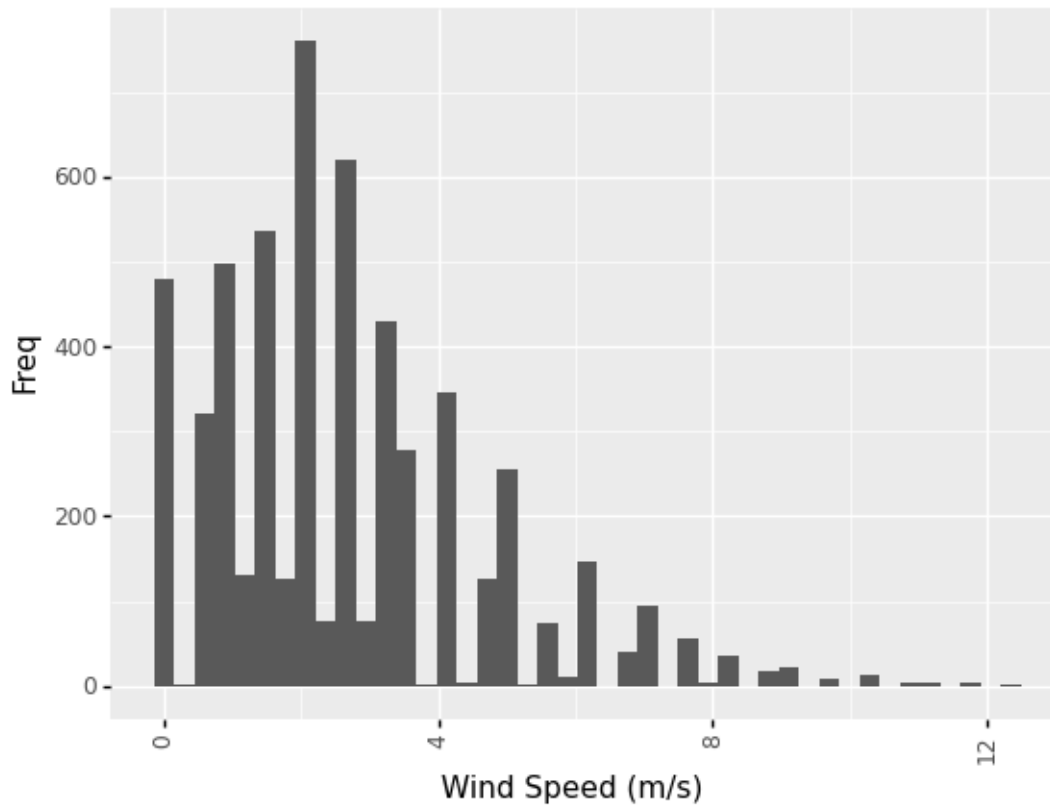
Scatter Plot of Weather
Conditions: Wind Speeed

```
<ggplot: (8759871546769)>
```

```
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/plotnine/stats/stat_bin.py:93: PlotnineWarning: 'stat_bin()' using
'bins = 43'. Pick better value with 'binwidth'.
```

Histogram Plot of Weather Conditions: Wind Speed

```
<ggplot: (8759871509709)>

/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will
not call `transform_cell` automatically in the future. Please pass the result to
`transformed_cell` argument and any exception that happen during thetransform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
/home/jaa6766/.conda/envs/cuda/lib/python3.7/site-
packages/plotnine/stats/stat_bin.py:93: PlotnineWarning: 'stat_bin()' using
'bins = 15'. Pick better value with 'binwidth'.
```

Histogram Plot of Weather Conditions: General Weather Conditions


```
<ggplot: (8759872072773)>
```
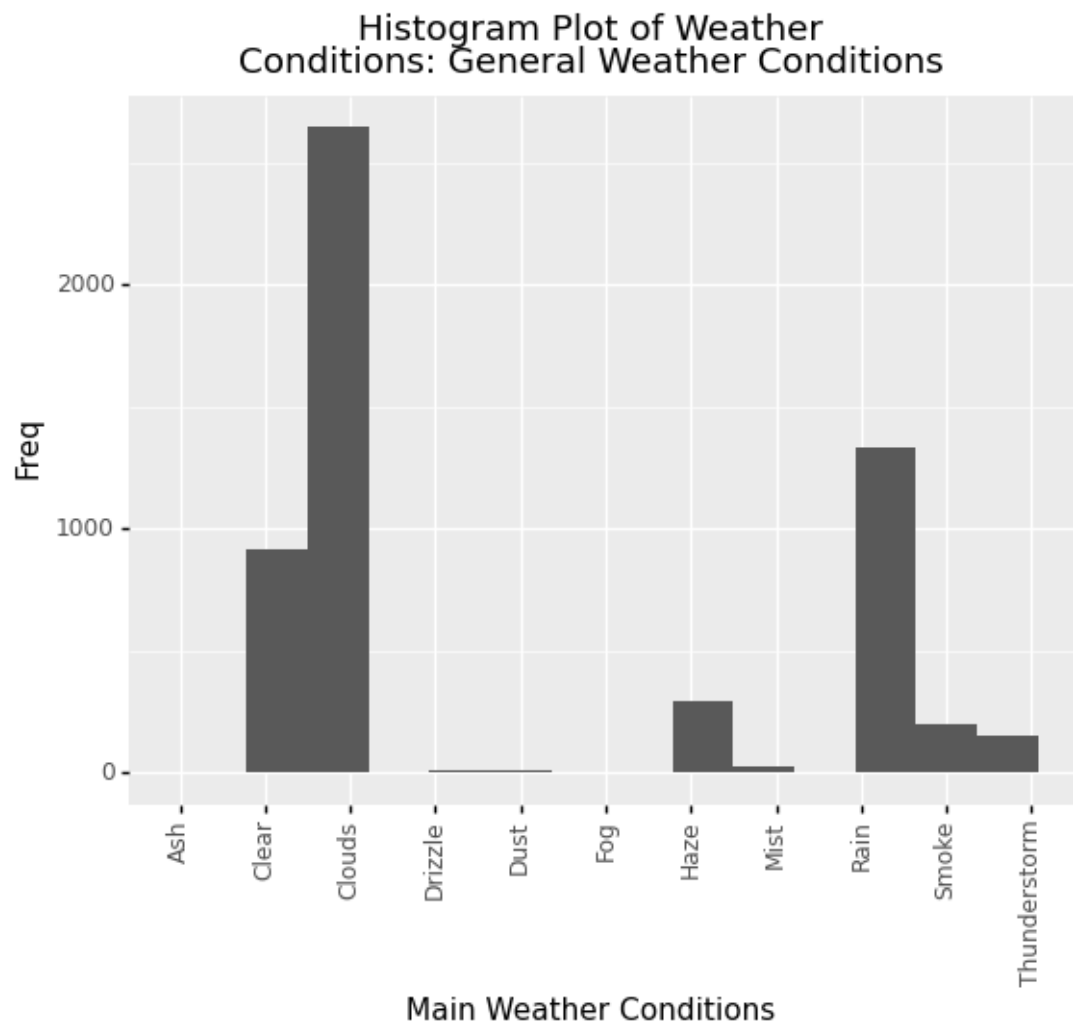
1.7 References

- Bosch BME680 Datasheet. 2021.

- Mancuso, Daniel. Indoor Air Quality Monitor | Hackster.io. 2019.

- OpenWeatherData: History Bulk weather data Documentation