

## Others

July 5, 2023

```
[26]: import pandas as pd
import matplotlib.pyplot as plt
```

```
[27]: air_quality=pd.read_csv("air_quality_no2_long.csv")
```

```
[28]: air_quality.head()
```

```
[28]:
```

	city	country	date.utc	location	parameter	value	\
0	Paris	FR	2019-06-21 00:00:00+00:00	FR04014	no2	20.0	
1	Paris	FR	2019-06-20 23:00:00+00:00	FR04014	no2	21.8	
2	Paris	FR	2019-06-20 22:00:00+00:00	FR04014	no2	26.5	
3	Paris	FR	2019-06-20 21:00:00+00:00	FR04014	no2	24.9	
4	Paris	FR	2019-06-20 20:00:00+00:00	FR04014	no2	21.4	

```
unit
0 g/m
1 g/m
2 g/m
3 g/m
4 g/m
```

```
[29]: air_quality=air_quality.rename(columns={"date.utc": "datetime"})
```

```
[30]: air_quality.head()
```

```
[30]:
```

	city	country	datetime	location	parameter	value	\
0	Paris	FR	2019-06-21 00:00:00+00:00	FR04014	no2	20.0	
1	Paris	FR	2019-06-20 23:00:00+00:00	FR04014	no2	21.8	
2	Paris	FR	2019-06-20 22:00:00+00:00	FR04014	no2	26.5	
3	Paris	FR	2019-06-20 21:00:00+00:00	FR04014	no2	24.9	
4	Paris	FR	2019-06-20 20:00:00+00:00	FR04014	no2	21.4	

```
unit
0 g/m
1 g/m
2 g/m
3 g/m
4 g/m
```

```
[31]: air_quality.columns
```

```
[31]: Index(['city', 'country', 'datetime', 'location', 'parameter', 'value',  
        'unit'],  
        dtype='object')
```

```
[32]: air_quality.columns=['city ', 'country ', 'datetime ', 'location ', 'parameter_'  
    ↪, 'value ',  
        'unit']
```

```
[33]: air_quality
```

```
[33]:
```

	city	country	datetime	location	\
0	Paris	FR	2019-06-21 00:00:00+00:00	FR04014	
1	Paris	FR	2019-06-20 23:00:00+00:00	FR04014	
2	Paris	FR	2019-06-20 22:00:00+00:00	FR04014	
3	Paris	FR	2019-06-20 21:00:00+00:00	FR04014	
4	Paris	FR	2019-06-20 20:00:00+00:00	FR04014	
...	...	...	...	...	
2063	London	GB	2019-05-07 06:00:00+00:00	London Westminster	
2064	London	GB	2019-05-07 04:00:00+00:00	London Westminster	
2065	London	GB	2019-05-07 03:00:00+00:00	London Westminster	
2066	London	GB	2019-05-07 02:00:00+00:00	London Westminster	
2067	London	GB	2019-05-07 01:00:00+00:00	London Westminster	

	parameter	value	unit
0	no2	20.0	g/m
1	no2	21.8	g/m
2	no2	26.5	g/m
3	no2	24.9	g/m
4	no2	21.4	g/m
...	...	...	...
2063	no2	26.0	g/m
2064	no2	16.0	g/m
2065	no2	19.0	g/m
2066	no2	19.0	g/m
2067	no2	23.0	g/m

```
[2068 rows x 7 columns]
```

```
[34]: air_quality["city "]
```

```
[34]:
```

0	Paris
1	Paris
2	Paris
3	Paris
4	Paris

```

...
2063    London
2064    London
2065    London
2066    London
2067    London
Name: city , Length: 2068, dtype: object

```

```
[35]: air_quality["datetime"]=pd.to_datetime(air_quality["datetime "])
```

```
[36]: air_quality["datetime"]
```

```

[36]: 0      2019-06-21 00:00:00+00:00
      1      2019-06-20 23:00:00+00:00
      2      2019-06-20 22:00:00+00:00
      3      2019-06-20 21:00:00+00:00
      4      2019-06-20 20:00:00+00:00
      ...
2063    2019-05-07 06:00:00+00:00
2064    2019-05-07 04:00:00+00:00
2065    2019-05-07 03:00:00+00:00
2066    2019-05-07 02:00:00+00:00
2067    2019-05-07 01:00:00+00:00
Name: datetime, Length: 2068, dtype: datetime64[ns, UTC]

```

```
[37]: air_quality
```

```

[37]:      city  country      datetime      location \
0      Paris      FR  2019-06-21 00:00:00+00:00      FR04014
1      Paris      FR  2019-06-20 23:00:00+00:00      FR04014
2      Paris      FR  2019-06-20 22:00:00+00:00      FR04014
3      Paris      FR  2019-06-20 21:00:00+00:00      FR04014
4      Paris      FR  2019-06-20 20:00:00+00:00      FR04014
...      ...      ...      ...      ...
2063    London      GB  2019-05-07 06:00:00+00:00  London Westminster
2064    London      GB  2019-05-07 04:00:00+00:00  London Westminster
2065    London      GB  2019-05-07 03:00:00+00:00  London Westminster
2066    London      GB  2019-05-07 02:00:00+00:00  London Westminster
2067    London      GB  2019-05-07 01:00:00+00:00  London Westminster

```

```

      parameter  value  unit      datetime
0          no2    20.0  g/m  2019-06-21 00:00:00+00:00
1          no2    21.8  g/m  2019-06-20 23:00:00+00:00
2          no2    26.5  g/m  2019-06-20 22:00:00+00:00
3          no2    24.9  g/m  2019-06-20 21:00:00+00:00
4          no2    21.4  g/m  2019-06-20 20:00:00+00:00
...      ...      ...      ...

```

```

2063      no2      26.0   g/m  2019-05-07 06:00:00+00:00
2064      no2      16.0   g/m  2019-05-07 04:00:00+00:00
2065      no2      19.0   g/m  2019-05-07 03:00:00+00:00
2066      no2      19.0   g/m  2019-05-07 02:00:00+00:00
2067      no2      23.0   g/m  2019-05-07 01:00:00+00:00

```

[2068 rows x 8 columns]

```
[38]: air_quality["datetime"]=pd.to_datetime(air_quality["datetime"])
```

```
[39]: air_quality["datetime"]
```

```

[39]: 0      2019-06-21 00:00:00+00:00
      1      2019-06-20 23:00:00+00:00
      2      2019-06-20 22:00:00+00:00
      3      2019-06-20 21:00:00+00:00
      4      2019-06-20 20:00:00+00:00
      ...
2063   2019-05-07 06:00:00+00:00
2064   2019-05-07 04:00:00+00:00
2065   2019-05-07 03:00:00+00:00
2066   2019-05-07 02:00:00+00:00
2067   2019-05-07 01:00:00+00:00
Name: datetime, Length: 2068, dtype: datetime64[ns, UTC]

```

```
[40]: air_quality.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2068 entries, 0 to 2067
Data columns (total 8 columns):
#   Column      Non-Null Count  Dtype
---  -
0   city        2068 non-null  object
1   country     2068 non-null  object
2   datetime    2068 non-null  object
3   location    2068 non-null  object
4   parameter   2068 non-null  object
5   value       2068 non-null  float64
6   unit        2068 non-null  object
7   datetime    2068 non-null  datetime64[ns, UTC]
dtypes: datetime64[ns, UTC](1), float64(1), object(6)
memory usage: 129.4+ KB

```

```
[41]: pd.read_csv("air_quality_no2_long.csv", parse_dates=["date.utc"])
```

```

[41]:      city country      date.utc      location \
0     Paris    FR  2019-06-21 00:00:00+00:00  FR04014
1     Paris    FR  2019-06-20 23:00:00+00:00  FR04014

```

2	Paris	FR	2019-06-20 22:00:00+00:00	FR04014
3	Paris	FR	2019-06-20 21:00:00+00:00	FR04014
4	Paris	FR	2019-06-20 20:00:00+00:00	FR04014
...	...	...	...	...
2063	London	GB	2019-05-07 06:00:00+00:00	London Westminster
2064	London	GB	2019-05-07 04:00:00+00:00	London Westminster
2065	London	GB	2019-05-07 03:00:00+00:00	London Westminster
2066	London	GB	2019-05-07 02:00:00+00:00	London Westminster
2067	London	GB	2019-05-07 01:00:00+00:00	London Westminster

	parameter	value	unit
0	no2	20.0	g/m
1	no2	21.8	g/m
2	no2	26.5	g/m
3	no2	24.9	g/m
4	no2	21.4	g/m
...	...	...	...
2063	no2	26.0	g/m
2064	no2	16.0	g/m
2065	no2	19.0	g/m
2066	no2	19.0	g/m
2067	no2	23.0	g/m

[2068 rows x 7 columns]

```
[42]: air_quality["datetime"].min(),air_quality["datetime"].max()
```

```
[42]: (Timestamp('2019-05-07 01:00:00+0000', tz='UTC'),
      Timestamp('2019-06-21 00:00:00+0000', tz='UTC'))
```

```
[43]: air_quality["datetime"].max()-air_quality["datetime"].min()
```

```
[43]: Timedelta('44 days 23:00:00')
```

```
[44]: air_quality["month"]=air_quality["datetime"].dt.month
```

```
[45]: air_quality.head()
```

```
[45]:
```

	city	country	datetime	location	parameter	value	\
0	Paris	FR	2019-06-21 00:00:00+00:00	FR04014	no2	20.0	
1	Paris	FR	2019-06-20 23:00:00+00:00	FR04014	no2	21.8	
2	Paris	FR	2019-06-20 22:00:00+00:00	FR04014	no2	26.5	
3	Paris	FR	2019-06-20 21:00:00+00:00	FR04014	no2	24.9	
4	Paris	FR	2019-06-20 20:00:00+00:00	FR04014	no2	21.4	

	unit	datetime	month
0	g/m	2019-06-21 00:00:00+00:00	6

```

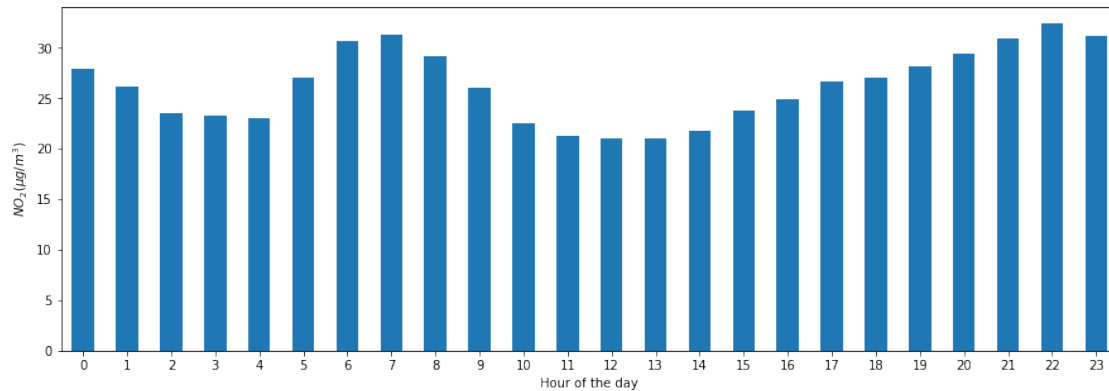
1  g/m  2019-06-20 23:00:00+00:00      6
2  g/m  2019-06-20 22:00:00+00:00      6
3  g/m  2019-06-20 21:00:00+00:00      6
4  g/m  2019-06-20 20:00:00+00:00      6

```

```
[46]: air_quality.groupby([air_quality["datetime"].dt.weekday,"location "])["value "].
      ↪mean()
```

```
[46]: datetime  location
0          BETR801      27.875000
          FR04014      24.856250
          London Westminster  23.969697
1          BETR801      22.214286
          FR04014      30.999359
          London Westminster  24.885714
2          BETR801      21.125000
          FR04014      29.165753
          London Westminster  23.460432
3          BETR801      27.500000
          FR04014      28.600690
          London Westminster  24.780142
4          BETR801      28.400000
          FR04014      31.617986
          London Westminster  26.446809
5          BETR801      33.500000
          FR04014      25.266154
          London Westminster  24.977612
6          BETR801      21.896552
          FR04014      23.274306
          London Westminster  24.859155
Name: value , dtype: float64
```

```
[47]: fig, axs=plt.subplots(figsize=(15,5))
      air_quality.groupby(air_quality["datetime"].dt.hour)["value "].mean().
      ↪plot(kind="bar",rot=0,ax=axs)
      plt.xlabel("Hour of the day");
      plt.ylabel("$NO_2$ (µg/m³)");
```



```
[50]: no_2=air_quality.pivot(index="datetime",columns="location ",values="value ")
```

```
[51]: no_2
```

```
[51]: location          BETR801    FR04014    London Westminster
datetime
2019-05-07 01:00:00+00:00      50.5      25.0              23.0
2019-05-07 02:00:00+00:00      45.0      27.7              19.0
2019-05-07 03:00:00+00:00       NaN      50.4              19.0
2019-05-07 04:00:00+00:00       NaN      61.9              16.0
2019-05-07 05:00:00+00:00       NaN      72.4              NaN
...
2019-06-20 20:00:00+00:00       NaN      21.4              NaN
2019-06-20 21:00:00+00:00       NaN      24.9              NaN
2019-06-20 22:00:00+00:00       NaN      26.5              NaN
2019-06-20 23:00:00+00:00       NaN      21.8              NaN
2019-06-21 00:00:00+00:00       NaN      20.0              NaN
```

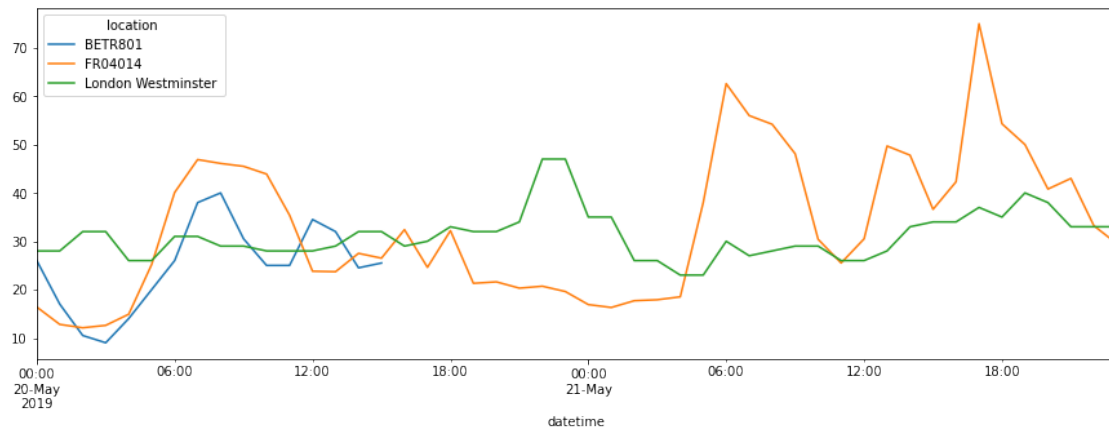
```
[1033 rows x 3 columns]
```

```
[54]: no_2.index.year, no_2.index.weekday
```

```
[54]: (Int64Index([2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019,
...
2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019],
dtype='int64', name='datetime', length=1033),
Int64Index([1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
...
3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4],
dtype='int64', name='datetime', length=1033))
```

```
[58]: fig, axs=plt.subplots(figsize=(15,5))
no_2["2019-05-20":"2019-05-21"].plot(ax=axs)
```

```
[58]: <AxesSubplot:xlabel='datetime'>
```



```
[59]: monthly_max=no_2.resample("M").max()
```

```
[60]: monthly_max
```

```
[60]: location          BETR801  FR04014  London Westminster
datetime
2019-05-31 00:00:00+00:00    74.5    97.0                97.0
2019-06-30 00:00:00+00:00    52.5    84.7                52.0
```

```
[ ]:
```