airvisualizeaf

May 1, 2025

```
[37]: import pandas as pd
      import numpy as np
      import matplotlib.pyplot as plt
      import seaborn as sns
[38]:
     df=pd.read_csv('AirQuality.csv',delimiter=';')
[39]:
      df.shape
[39]: (9471, 17)
[40]:
      df.head()
                          Time CO(GT)
                                                    NMHC(GT) C6H6(GT)
[40]:
               Date
                                       PT08.S1(CO)
                                                                        PT08.S2(NMHC)
      0 10/03/2004 18.00.00
                                  2,6
                                            1360.0
                                                        150.0
                                                                  11,9
                                                                                1046.0
      1 10/03/2004
                     19.00.00
                                    2
                                            1292.0
                                                        112.0
                                                                   9,4
                                                                                 955.0
      2 10/03/2004
                     20.00.00
                                  2,2
                                            1402.0
                                                         88.0
                                                                   9,0
                                                                                 939.0
      3 10/03/2004
                     21.00.00
                                                                   9,2
                                  2,2
                                            1376.0
                                                         80.0
                                                                                 948.0
      4 10/03/2004
                     22.00.00
                                  1,6
                                            1272.0
                                                         51.0
                                                                   6,5
                                                                                 836.0
         NOx(GT) PT08.S3(NOx)
                                 NO2(GT)
                                          PT08.S4(NO2)
                                                        PT08.S5(03)
                                                                              RH \
      0
           166.0
                        1056.0
                                   113.0
                                                1692.0
                                                              1268.0 13,6 48,9
           103.0
                         1174.0
                                    92.0
                                                1559.0
                                                               972.0 13,3 47,7
      1
      2
           131.0
                        1140.0
                                   114.0
                                                1555.0
                                                              1074.0
                                                                     11,9 54,0
      3
           172.0
                                                              1203.0 11,0 60,0
                         1092.0
                                   122.0
                                                1584.0
           131.0
                         1205.0
                                   116.0
                                                1490.0
                                                              1110.0 11,2 59,6
                 Unnamed: 15 Unnamed: 16
             AH
      0 0,7578
                          NaN
                                       NaN
      1 0,7255
                          NaN
                                       NaN
      2 0,7502
                          {\tt NaN}
                                       NaN
      3 0,7867
                          NaN
                                       NaN
      4 0,7888
                          {\tt NaN}
                                       NaN
[41]: | df=df.rename(columns={'T': 'Temperature'})
      df=df.rename(columns={'RH':'Relative Humidity'})
      df=df.rename(columns={'AH':'Absolute Humidity'})
```

df

[41]:		Date	Time	CO(GT)	PT08.S1	(CO) NMH(C(GT)	C6H6(GT)	\
	0	10/03/2004	18.00.00	2,6	136	30.0	150.0	11,9	
	1	10/03/2004	19.00.00	2	129	92.0	112.0	9,4	
	2	10/03/2004	20.00.00	2,2	140	02.0	88.0	9,0	
	3	10/03/2004	21.00.00	2,2	137	76.0	80.0	9,2	
	4	10/03/2004	22.00.00	1,6		2.0	51.0	6,5	
	•••	•••							
	9466	NaN	NaN	NaN		NaN	NaN	NaN	
	9467	NaN	NaN	NaN		NaN	NaN	NaN	
	9468	NaN	NaN	NaN		NaN	NaN	NaN	
	9469	NaN	NaN	NaN		NaN	NaN	NaN	
	9470	NaN	NaN	NaN		NaN	NaN	NaN	
		PT08.S2(NMH	C) NOx(GT	r) PT08	.S3(NOx)	NO2(GT)	PT08	3.S4(NO2)	\
	0	1046			1056.0	113.0		1692.0	•
	1	955			1174.0	92.0		1559.0	
	2	939			1140.0	114.0		1555.0	
	3	948			1092.0	122.0		1584.0	
	4	836			1205.0	116.0		1490.0	
		030		. 0		110.0		1430.0	
	 9466	 N	aN Na	 .N	 NaN	NaN	•••	NaN	
	9467		aN Na		NaN	NaN		NaN	
	9468		an na aN Na		NaN	NaN		NaN	
	9469		an na aN Na		NaN	NaN			
								NaN NaN	
	9470	IV i	aN Na	IIV	NaN	NaN		NaN	
		PT08.S5(03)	Temperatu	ıre Rela	tive Humi	dity Abso	olute	Humidity	\
	0	1268.0	13	3,6		48,9		0,7578	
	1	972.0	13	3,3		47,7		0,7255	
	2	1074.0	11	,9		54,0		0,7502	
	3	1203.0	11	1,0		60,0		0,7867	
	4	1110.0	11	,2		59,6		0,7888	
			•••		•••		•••		
	9466	NaN		NaN 		NaN		NaN	
	9467	NaN		JaN		NaN		NaN	
	9468	NaN		JaN		NaN		NaN	
	9469	NaN		NaN		NaN		NaN	
	9470	NaN	Ŋ	JaN		NaN		NaN	
		Unnamed: 15	Unnamed:	16					
	0	NaN		NaN					
	1	NaN		NaN					
	2	NaN		NaN					
	3	NaN		NaN					
	4	NaN		NaN					

```
9466 NaN NaN 9467 NaN NaN 9468 NaN NaN NaN 9469 NaN NaN NaN 9470 NaN NaN NaN
```

[9471 rows x 17 columns]

```
[42]: df=df.drop(['Unnamed: 15', 'Unnamed: 16'],axis=1) df
```

[42]:		Date	Time	CO(GT)	PT08.S1((CO) NMH	C(GT)	C6H6(GT)	\
	0	10/03/2004	18.00.00	2,6	136	30.0	150.0	11,9	
	1	10/03/2004	19.00.00	2	129	2.0	112.0	9,4	
	2	10/03/2004	20.00.00	2,2	140	2.0	88.0	9,0	
	3	10/03/2004	21.00.00	2,2	137	6.0	80.0	9,2	
	4	10/03/2004	22.00.00	1,6	127	2.0	51.0	6,5	
	•••	•••			•••				
	9466	NaN	NaN	NaN		NaN	NaN	NaN	
	9467	NaN	NaN	NaN		NaN	NaN	NaN	
	9468	NaN	NaN	NaN		NaN	NaN	NaN	
	9469	NaN	NaN	NaN		NaN	NaN	NaN	
	9470	NaN	NaN	NaN		NaN	NaN	NaN	
		PT08.S2(NMH	C) NOx(G	T) PT08	3.S3(NOx)	NO2(GT)	PT08	3.S4(NO2)	\
	0	1046			1056.0	113.0		1692.0	
	1	955	.0 103	.0	1174.0	92.0		1559.0	
	2	939	.0 131	.0	1140.0	114.0		1555.0	
	3	948	.0 172	.0	1092.0	122.0		1584.0	
	4	836	.0 131	.0	1205.0	116.0		1490.0	
	•••		•••	•••			•••		
	9466	Na		aN	NaN	NaN		NaN	
	9467			aN	NaN	NaN		NaN	
	9468			aN	NaN	NaN		NaN	
	9469			aN	NaN	NaN		NaN	
	9470	Na	aN Na	aN	NaN	NaN		NaN	
		PT08.S5(03)	Temperat	ure Rela	tive Humi	dity Abso	olute	Humidity	
	0	1268.0	-	3,6		48,9		0,7578	
	1	972.0		3,3		47,7		0,7255	
	2	1074.0	1:	1,9		54,0		0,7502	
	3	1203.0	1:	1,0		60,0		0,7867	
	4	1110.0	1	1,2		59,6		0,7888	
		***	•••	NT _ NT	•••	N - N	•••	NT . NT	
	9466	NaN N-N		NaN Nan		NaN N-N		NaN N-N	
	9467	NaN]	NaN		NaN		NaN	

```
9469
                     NaN
                                 NaN
                                                    NaN
                                                                        NaN
      9470
                     NaN
                                 NaN
                                                    NaN
                                                                        NaN
      [9471 rows x 15 columns]
[43]: df['CO(GT)']=df['CO(GT)'].str.replace(',','.').astype(float)
      df['C6H6(GT)']=df['C6H6(GT)'].str.replace(',','.').astype(float)
      df['Temperature'] = df['Temperature'].str.replace(',','.').astype(float)
      df['Relative Humidity']=df['Relative Humidity'].str.replace(',','.').
       →astype(float)
      df['Absolute Humidity']=df['Absolute Humidity'].str.replace(',','.').
        →astype(float)
[44]: df=df.drop_duplicates()
[45]: df1=df.iloc[:9466,:]
      df1
[45]:
                                   CO(GT)
                                            PT08.S1(CO)
                                                          NMHC(GT)
                                                                    C6H6(GT)
                  Date
                             Time
      0
            10/03/2004
                         18.00.00
                                       2.6
                                                 1360.0
                                                             150.0
                                                                         11.9
      1
            10/03/2004
                         19.00.00
                                       2.0
                                                 1292.0
                                                             112.0
                                                                          9.4
      2
                         20.00.00
                                       2.2
                                                                          9.0
            10/03/2004
                                                 1402.0
                                                              88.0
      3
            10/03/2004
                         21.00.00
                                       2.2
                                                 1376.0
                                                              80.0
                                                                          9.2
                         22.00.00
      4
            10/03/2004
                                       1.6
                                                 1272.0
                                                                          6.5
                                                              51.0
      9353 04/04/2005
                         11.00.00
                                       2.4
                                                 1163.0
                                                            -200.0
                                                                         11.4
      9354
            04/04/2005
                                       2.4
                                                 1142.0
                                                                         12.4
                         12.00.00
                                                            -200.0
      9355 04/04/2005
                         13.00.00
                                       2.1
                                                 1003.0
                                                            -200.0
                                                                          9.5
      9356
            04/04/2005
                         14.00.00
                                       2.2
                                                 1071.0
                                                            -200.0
                                                                         11.9
      9357
                   NaN
                              NaN
                                       NaN
                                                    NaN
                                                               NaN
                                                                          NaN
            PT08.S2(NMHC)
                            NOx(GT)
                                      PTO8.S3(NOx)
                                                    NO2(GT)
                                                              PT08.S4(NO2)
      0
                    1046.0
                              166.0
                                            1056.0
                                                       113.0
                                                                    1692.0
      1
                     955.0
                              103.0
                                            1174.0
                                                        92.0
                                                                    1559.0
      2
                     939.0
                              131.0
                                            1140.0
                                                       114.0
                                                                    1555.0
      3
                     948.0
                              172.0
                                            1092.0
                                                       122.0
                                                                     1584.0
      4
                     836.0
                              131.0
                                            1205.0
                                                       116.0
                                                                    1490.0
      9353
                    1027.0
                              353.0
                                             604.0
                                                       179.0
                                                                    1264.0
      9354
                                             603.0
                                                       175.0
                                                                    1241.0
                    1063.0
                              293.0
      9355
                     961.0
                              235.0
                                             702.0
                                                       156.0
                                                                    1041.0
      9356
                              265.0
                                             654.0
                                                       168.0
                                                                    1129.0
                    1047.0
      9357
                       NaN
                                NaN
                                               NaN
                                                         NaN
                                                                        NaN
            PT08.S5(03)
                          Temperature Relative Humidity Absolute Humidity
      0
                  1268.0
                                 13.6
                                                      48.9
                                                                        0.7578
```

9468

NaN

NaN

NaN

NaN

1	972.0	13.3	47.7	0.7255
2	1074.0	11.9	54.0	0.7502
3	1203.0	11.0	60.0	0.7867
4	1110.0	11.2	59.6	0.7888
	•••	•••	•••	•••
9353	1269.0	24.3	23.7	0.7119
9354	1092.0	26.9	18.3	0.6406
9355	770.0	28.3	13.5	0.5139
9356	816.0	28.5	13.1	0.5028
9357	NaN	NaN	NaN	NaN

[9358 rows x 15 columns]

```
[46]: df.isna().sum()
```

```
[46]: Date
                            1
      Time
                            1
      CO(GT)
                            1
      PT08.S1(CO)
                            1
      NMHC(GT)
                            1
      C6H6(GT)
                            1
      PT08.S2(NMHC)
                            1
      NOx(GT)
                            1
      PT08.S3(NOx)
                            1
      NO2(GT)
                            1
      PT08.S4(NO2)
                            1
      PT08.S5(03)
                            1
      Temperature
                            1
      Relative Humidity
                            1
      Absolute Humidity
                            1
      dtype: int64
```

[47]: df=df.fillna(df.mean(numeric_only=True)) df=df.dropna()

[48]: df.isna().sum()

[48]: Date 0 Time 0 CO(GT) 0 PT08.S1(CO) 0 0 NMHC(GT) C6H6(GT) 0 PT08.S2(NMHC) 0 NOx(GT) 0 PT08.S3(NOx) 0 NO2(GT) 0

```
0
      PT08.S5(03)
      Temperature
                          0
      Relative Humidity
      Absolute Humidity
                          0
      dtype: int64
[49]: df['Absolute Humidity']=df['Absolute Humidity'].multiply(100)
[50]: def remove outliers(column):
          Q1=column.quantile(0.25)
         Q3=column.quantile(0.75)
         IQR=Q3-Q1
         threshold=1.5*IQR
          outlier_mask=(column<Q1-threshold) | (column>Q3+threshold)
         return column[~outlier_mask]
[51]: df.columns
[51]: Index(['Date', 'Time', 'CO(GT)', 'PT08.S1(CO)', 'NMHC(GT)', 'C6H6(GT)',
             'PT08.S2(NMHC)', 'NOx(GT)', 'PT08.S3(NOx)', 'NO2(GT)', 'PT08.S4(NO2)',
             'PT08.S5(03)', 'Temperature', 'Relative Humidity', 'Absolute Humidity'],
            dtype='object')
[52]: col_name=['PT08.S1(CO)', 'C6H6(GT)', 'PT08.S2(NMHC)', 'NOx(GT)', 'PT08.S3(NOx)',
       →'NO2(GT)', 'PT08.S4(NO2)', 'PT08.S5(O3)', 'Temperature', 'Relative Humidity', □
       for col in col_name:
         df[col]=remove_outliers(df[col])
[53]: df['Date']=pd.to_datetime(df["Date"], format="%d/%m/%Y", errors='coerce')
      df['Year']=df['Date'].dt.year
      df['Month'] = df['Date'].dt.month
[54]: df
[54]:
                          Time CO(GT)
                                        PT08.S1(CO)
                                                     NMHC(GT)
                                                               C6H6(GT) \
                Date
      0
          2004-03-10 18.00.00
                                    2.6
                                             1360.0
                                                        150.0
                                                                    11.9
          2004-03-10 19.00.00
                                    2.0
                                             1292.0
                                                        112.0
                                                                    9.4
      1
          2004-03-10 20.00.00
                                    2.2
                                                                    9.0
      2
                                             1402.0
                                                         88.0
      3
          2004-03-10 21.00.00
                                    2.2
                                             1376.0
                                                         80.0
                                                                    9.2
      4
          2004-03-10 22.00.00
                                    1.6
                                             1272.0
                                                         51.0
                                                                    6.5
      9352 2005-04-04 10.00.00
                                    3.1
                                             1314.0
                                                       -200.0
                                                                   13.5
      9353 2005-04-04 11.00.00
                                    2.4
                                             1163.0
                                                       -200.0
                                                                   11.4
      9354 2005-04-04 12.00.00
                                    2.4
                                             1142.0
                                                       -200.0
                                                                    12.4
      9355 2005-04-04 13.00.00
                                    2.1
                                             1003.0
                                                       -200.0
                                                                    9.5
```

PT08.S4(NO2)

0

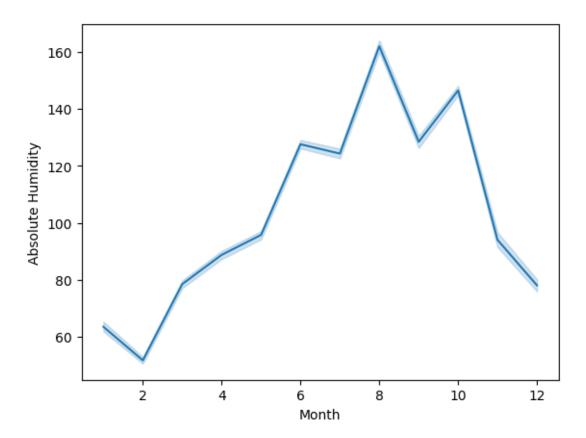
```
9356 2005-04-04 14.00.00
                               2.2
                                         1071.0
                                                    -200.0
                                                                 11.9
                                             NO2(GT) PT08.S4(NO2) \
      PTO8.S2(NMHC)
                     NOx(GT)
                               PTO8.S3(NOx)
0
             1046.0
                        166.0
                                     1056.0
                                                113.0
                                                              1692.0
1
              955.0
                        103.0
                                     1174.0
                                                 92.0
                                                              1559.0
2
              939.0
                        131.0
                                     1140.0
                                                114.0
                                                              1555.0
3
              948.0
                        172.0
                                     1092.0
                                                122.0
                                                              1584.0
4
              836.0
                                                116.0
                        131.0
                                     1205.0
                                                              1490.0
9352
             1101.0
                        472.0
                                       539.0
                                                190.0
                                                              1374.0
9353
             1027.0
                        353.0
                                       604.0
                                                179.0
                                                              1264.0
9354
             1063.0
                        293.0
                                       603.0
                                                175.0
                                                              1241.0
9355
              961.0
                        235.0
                                      702.0
                                                156.0
                                                              1041.0
9356
             1047.0
                        265.0
                                       654.0
                                                168.0
                                                              1129.0
      PT08.S5(03)
                   Temperature Relative Humidity Absolute Humidity Year \
0
           1268.0
                           13.6
                                               48.9
                                                                  75.78
                                                                         2004
1
            972.0
                           13.3
                                               47.7
                                                                  72.55
                                                                         2004
2
                           11.9
                                               54.0
                                                                  75.02
                                                                         2004
           1074.0
           1203.0
                           11.0
                                               60.0
                                                                  78.67
                                                                         2004
4
           1110.0
                           11.2
                                               59.6
                                                                  78.88
                                                                         2004
9352
           1729.0
                           21.9
                                               29.3
                                                                  75.68
                                                                         2005
                           24.3
                                               23.7
9353
           1269.0
                                                                  71.19
                                                                         2005
9354
           1092.0
                           26.9
                                               18.3
                                                                  64.06
                                                                         2005
9355
            770.0
                           28.3
                                               13.5
                                                                  51.39
                                                                         2005
9356
            816.0
                           28.5
                                               13.1
                                                                  50.28
                                                                         2005
      Month
0
          3
1
          3
2
          3
3
          3
4
          3
9352
          4
9353
          4
9354
          4
9355
          4
9356
[9357 rows x 17 columns]
```

[55]: df['yearr']=df.Year.astype(str)
df['month']=df.Month.astype(str)

```
[56]: plt.figure(figsize=(8,6))
         sns.heatmap(df.corr(numeric_only=True), annot=True,cmap='coolwarm',fmt='.2f')
[56]: <Axes: >
                                                                                                                            1.00
                            CO(GT) -1.00 0.11 0.13 0.08 0.08 0.59-0.090.16-0.060.12-0.13-0.02-0.190.20-0.15
                      PT08.S1(CO) -0.11 1.00 0.22 0.87 0.88 0.44-0.80 0.64 0.65 0.89 0.06 0.11 0.13 0.04-0.05
                                                                                                                           - 0.75
                         NMHC(GT) -0.13 0.22 1.00 0.14 0.14 0.05 0.05-0.010.20 0.12-0.07-0.01-0.12-0.16-0.23
                          C6H6(GT) -0.08 0.87 0.14 1.00 0.99 0.41-0.78 0.62 0.73 0.86 0.23-0.100.16-0.12 0.10
                                                                                                                           - 0.50
                   PT08.S2(NMHC) -0.08 0.88 0.14 0.99 1.00 0.41 0.82 0.65 0.76 0.87 0.25 0.100.18 0.14 0.11
                           NOx(GT) -0.59 0.44 0.05 0.41 0.41 1.00-0.46 0.75 0.04 0.49 -0.19 0.04-0.21 0.32 -0.11
                                                                                                                          - 0.25
                     PT08.S3(NOx) -0.09-0.80<mark>0.05-0.78-0.82-0.46</mark> 1.00-0.66-0.51-0.83-0.10-0.08-0.20-0.15-0.09
                           NO2(GT) -0.16 0.64-0.01 0.62 0.65 0.75-0.66 1.00 0.15 0.69-0.17-0.10-0.33 0.36-0.16
                                                                                                                          - 0.00
                     PT08.S4(NO2) -0.06 0.65 0.20 0.73 0.76 0.04-0.51 0.15 1.00 0.57 0.58-0.04 0.64-0.51 0.18
                      PT08.S5(O3) -0.12 0.89 0.12 0.86 0.87 0.49 0.83 0.69 0.57 1.00 -0.020.12 0.08 0.04 0.04
                                                                                                                           - -0.25
                      Temperature -0.130.06-0.070.230.25-0.19-0.10-0.170.58-0.021.00-0.580.66-0.520.28
                 Relative Humidity -0.020.11-0.01-0.10-0.100.04-0.08-0.10-0.040.12-0.581.00 0.17 0.10 0.08
                Absolute Humidity -0.190.13-0.120.16 0.18-0.21-0.20-0.33 0.64 0.08 0.66 0.17 1.00-0.51 0.43
                                                                                                                            -0.50
                               Year -0.20 0.04-0.16-0.12-0.14<mark>0.32</mark>-0.15<mark>0.36</mark>-0.510.04-0.52<mark>0.10</mark>-0.51
                             Month -0.15-0.05-0.23 0.10 0.11-0.11-0.09-0.16 0.18 0.04 0.28 0.08 0.43-0.69 1.00
                                                                NOx(GT)
                                                                          NO2(GT)
                                                                                          Temperature
                                                                                               Relative Humidity
                                                                                                    Absolute Humidity
                                                           7108.S2(NMHC)
                                                                     PT08.53(NOx)
                                                                                PT08.S4(NO2)
                                                                                     PT08.S5(03)
                                            PT08.S1(CO)
                                                 NMHC(GT)
                                                      C6H6(GT)
```

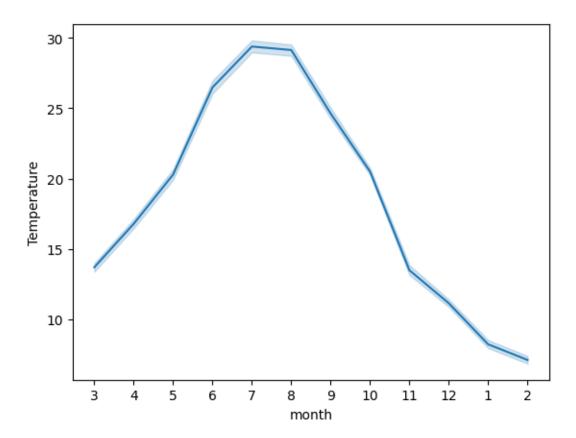
```
[57]: sns.lineplot(df,x="Month",y="Absolute Humidity")
```

[57]: <Axes: xlabel='Month', ylabel='Absolute Humidity'>



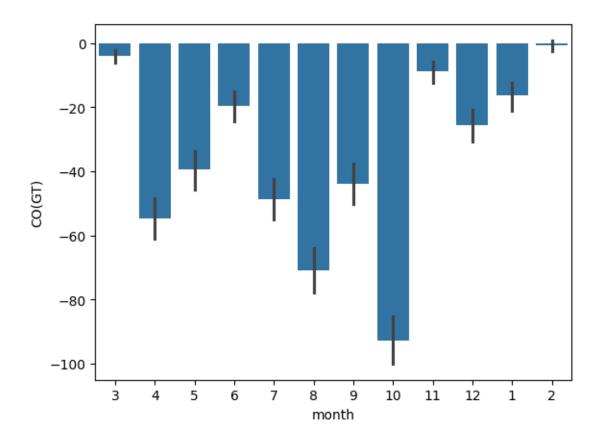
```
[58]: sns.lineplot(df,x='month',y='Temperature')
```

[58]: <Axes: xlabel='month', ylabel='Temperature'>



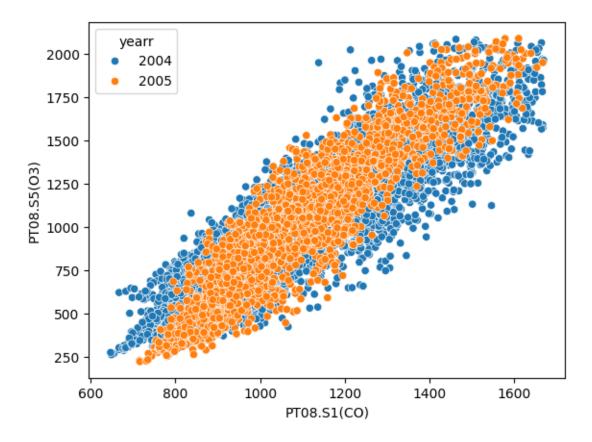
```
[68]: sns.barplot(df,x=df.month,y=df['CO(GT)'])
```

[68]: <Axes: xlabel='month', ylabel='CO(GT)'>



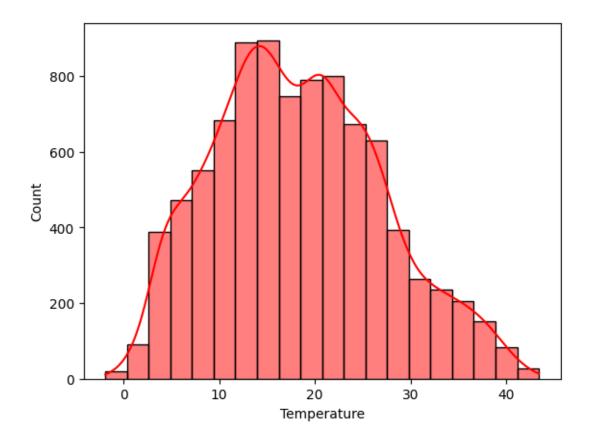
```
[60]: sns.scatterplot(df,x='PT08.S1(C0)',y='PT08.S5(03)',hue='yearr')
```

[60]: <Axes: xlabel='PT08.S1(C0)', ylabel='PT08.S5(03)'>



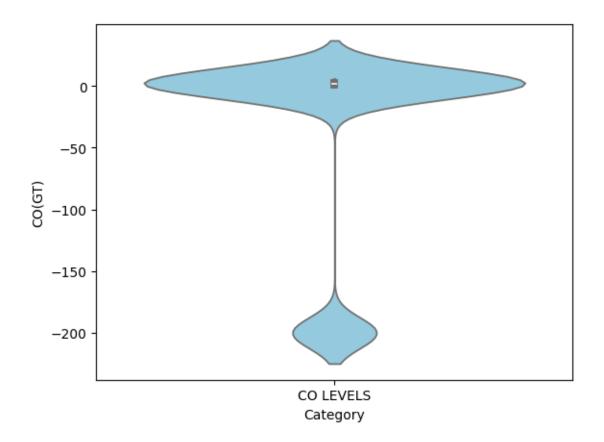
```
[61]: sns.histplot(df["Temperature"], bins=20,kde=True,color="red")
```

[61]: <Axes: xlabel='Temperature', ylabel='Count'>



```
[62]: df['Category']='CO LEVELS'
sns.violinplot(x=df['Category'],y=df['CO(GT)'],color='skyblue')
```

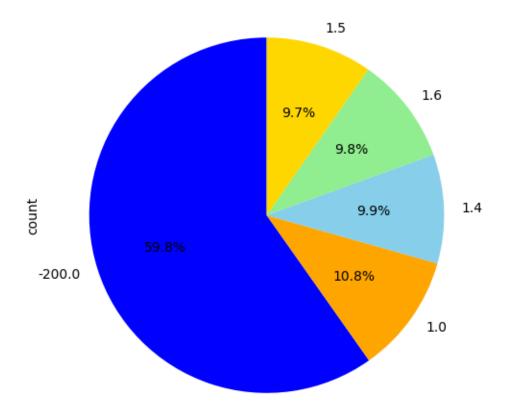
[62]: <Axes: xlabel='Category', ylabel='CO(GT)'>



```
[66]: co_counts=df['CO(GT)'].value_counts().nlargest(5)
plt.figure(figsize=(8,6))
co_counts.plot.pie(autopct='%1.1f%%',labels=co_counts.

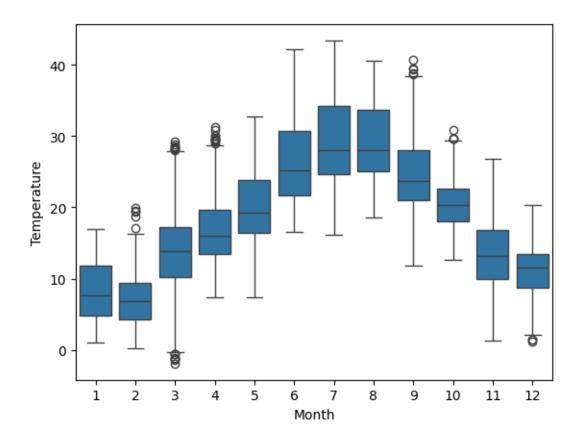
→index,startangle=90,colors=['blue','orange','skyblue','lightgreen','gold'])
```

[66]: <Axes: ylabel='count'>



```
[67]: sns.boxplot(df,x="Month",y='Temperature')
```

[67]: <Axes: xlabel='Month', ylabel='Temperature'>



[]: