air-polution-meter-index

August 21, 2023

[55]:
```
import pandas as pd
import numpy as np
```

[56]:
```
air=pd.read_csv(r"C:\Users\DELL\Desktop\air-quality-data.csv")
```

[57]:
```
air
```

[57]:
```
              Timestamp   PM2.5
0       01-01-2018 00:00   90.19
1       01-01-2018 01:00   86.98
2       01-01-2018 02:00   86.06
3       01-01-2018 03:00   94.04
4       01-01-2018 04:00  108.78
…                   …      …
31856   31-12-2021 19:00   72.25
31857   31-12-2021 20:00   71.56
31858   31-12-2021 21:00   70.27
31859   31-12-2021 22:00   66.78
31860   31-12-2021 23:00   61.29

[31861 rows x 2 columns]
```

[58]:
```
air.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 31861 entries, 0 to 31860
Data columns (total 2 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Timestamp  31861 non-null  object
 1   PM2.5      31861 non-null  float64
dtypes: float64(1), object(1)
memory usage: 498.0+ KB
```

[59]:
```
air["Timestamp"]=pd.to_datetime(air["Timestamp"])
```

[60]:
```
air.head()
```

[60]:
```
             Timestamp   PM2.5
0 2018-01-01 00:00:00   90.19
1 2018-01-01 01:00:00   86.98
2 2018-01-01 02:00:00   86.06
3 2018-01-01 03:00:00   94.04
4 2018-01-01 04:00:00  108.78
```

1 SHOW THE YEAR WISE AVERAGE POPULATION WITH A LINE CHART.

[61]:
```
air['year']= air["Timestamp"].dt.year #THESE ARE THE TWO WAY TO ASTRACT YEAR
```

[62]:
```
air['year']=air.Timestamp.dt.year # 2 WAY
```

[63]:
```
air
```

[63]:
```
                Timestamp   PM2.5  year
0     2018-01-01 00:00:00   90.19  2018
1     2018-01-01 01:00:00   86.98  2018
2     2018-01-01 02:00:00   86.06  2018
3     2018-01-01 03:00:00   94.04  2018
4     2018-01-01 04:00:00  108.78  2018
…                     …      …    …
31856 2021-12-31 19:00:00   72.25  2021
31857 2021-12-31 20:00:00   71.56  2021
31858 2021-12-31 21:00:00   70.27  2021
31859 2021-12-31 22:00:00   66.78  2021
31860 2021-12-31 23:00:00   61.29  2021

[31861 rows x 3 columns]
```

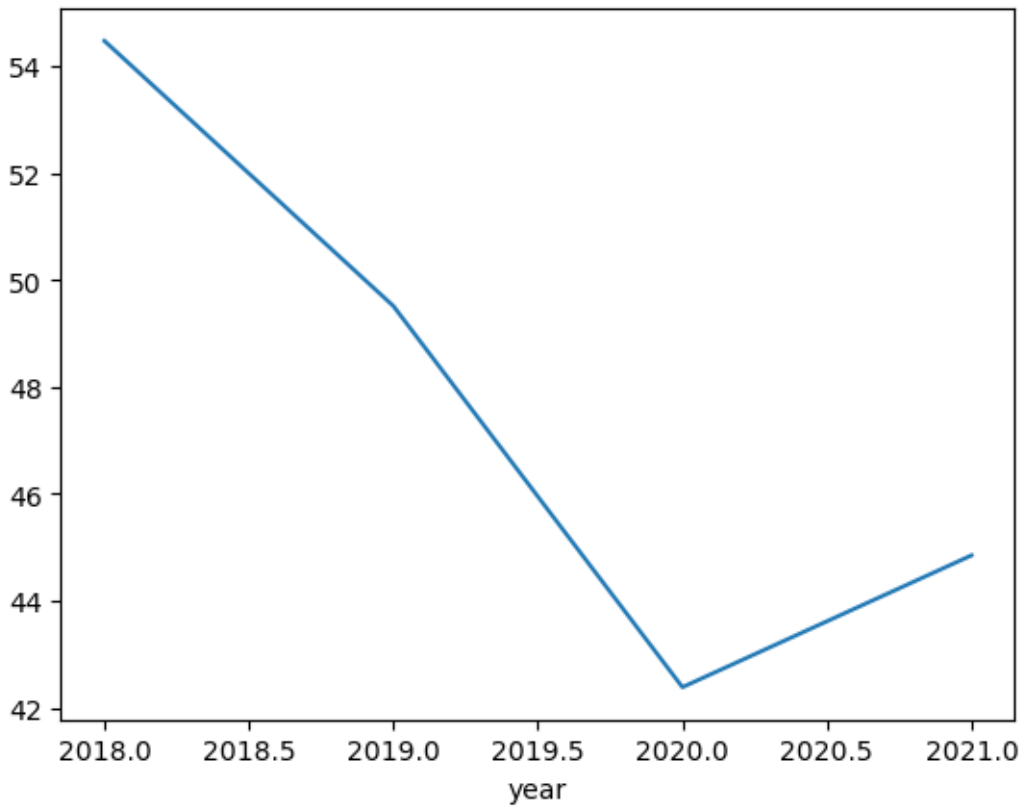
[64]:
```
Year_avg=air.groupby("year")["PM2.5"].mean()
```

[65]:
```
Year_avg
```

[65]:
```
year
2018    54.468449
2019    49.511273
2020    42.386302
2021    44.850633
Name: PM2.5, dtype: float64
```

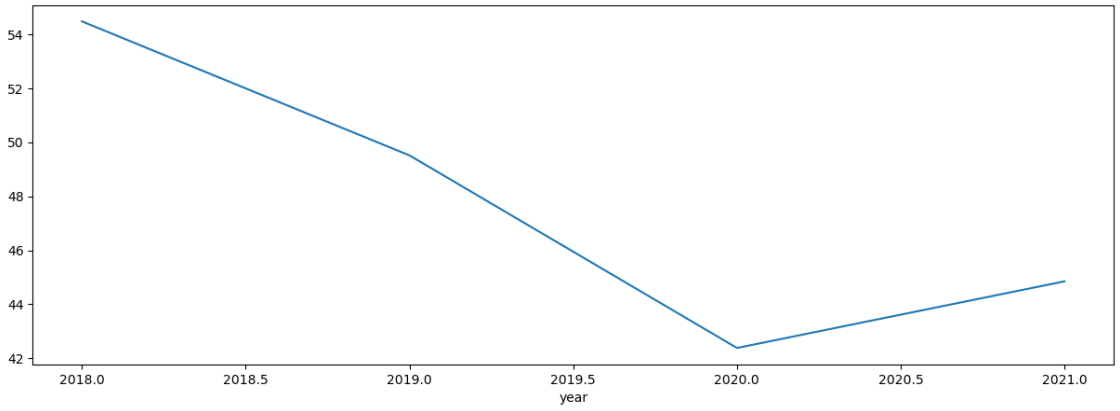
[66]:
```
Year_avg.plot(kind="line")
```

[66]:
```
<Axes: xlabel='year'>
```



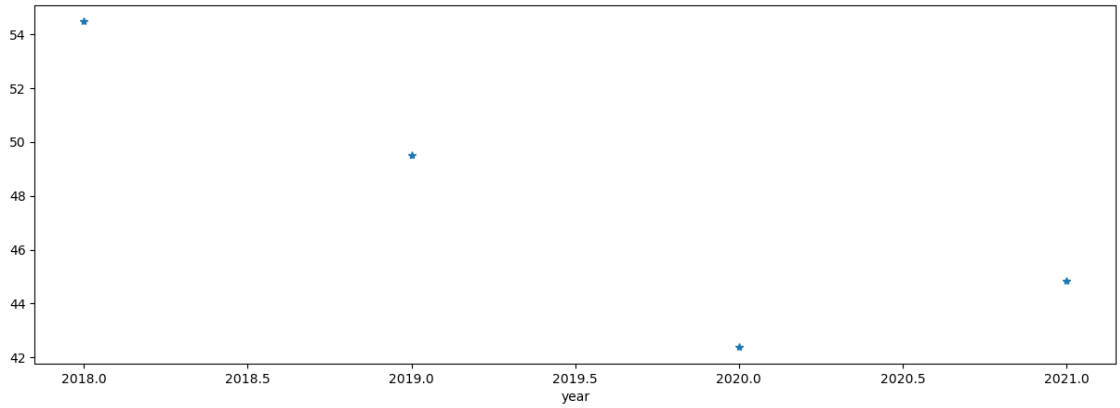
[67]:
```
Year_avg.plot(kind="line",figsize=(15,5))
```

[67]:
```
<Axes: xlabel='year'>
```



[68]:
```
Year_avg.plot(kind="line",figsize=(15,5),style="*")
```

[68]: <Axes: xlabel='year'>



2 DRAW THE AREA PLOT SHOWING THE AVERAGE POLLUTION MONTHS WISE

[69]:
```
air.head(2)
```

[69]:
```
             Timestamp  PM2.5  year
0 2018-01-01 00:00:00  90.19  2018
1 2018-01-01 01:00:00  86.98  2018
```

[70]:
```
air["month"]=air.Timestamp.dt.month
```

[71]:
```
air
```

[71]:
```
                 Timestamp   PM2.5  year  month
0      2018-01-01 00:00:00   90.19  2018      1
1      2018-01-01 01:00:00   86.98  2018      1
2      2018-01-01 02:00:00   86.06  2018      1
3      2018-01-01 03:00:00   94.04  2018      1
4      2018-01-01 04:00:00  108.78  2018      1
…                      …      …    …     …
31856  2021-12-31 19:00:00   72.25  2021     12
31857  2021-12-31 20:00:00   71.56  2021     12
31858  2021-12-31 21:00:00   70.27  2021     12
31859  2021-12-31 22:00:00   66.78  2021     12
31860  2021-12-31 23:00:00   61.29  2021     12

[31861 rows x 4 columns]
```

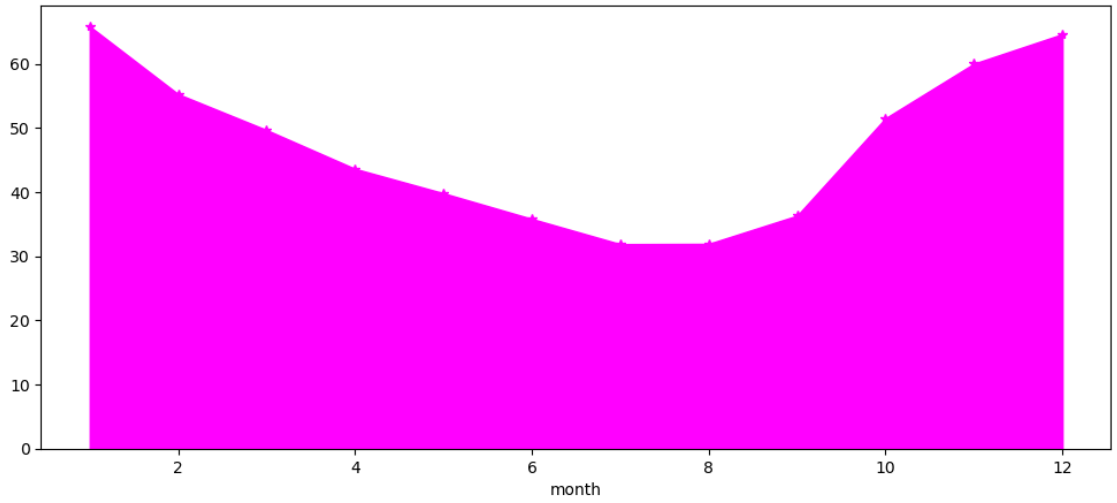
[72]:
```
Month_avg=air.groupby("month")["PM2.5"].mean()
```

[73]:
```
Month_avg
```

[73]:
```
month
1     65.829067
2     55.264291
3     49.666820
4     43.625814
5     39.816285
6     35.817151
7     31.884161
8     31.928300
9     36.324103
10    51.514817
11    60.017498
12    64.614865
Name: PM2.5, dtype: float64
```

[153]:
```
Month_avg.plot(kind="area",figsize=(12,5),style="*",color="magenta")
```

[153]:
```
<Axes: xlabel='month'>
```



3 DRAW THE AREA PLOT SHOWING THE AVERAGE POLLUTION HOURS WISE

[156]:
```
air.head(2)
```

[156]:

	Timestamp	PM2.5	year	month	Hour
0	2018-01-01 00:00:00	90.19	2018	1	0
1	2018-01-01 01:00:00	86.98	2018	1	1

[77]:
```
air["Hour"]=air.Timestamp.dt.hour
```

[78]:
```
air
```

[78]:

	Timestamp	PM2.5	year	month	Hour
0	2018-01-01 00:00:00	90.19	2018	1	0
1	2018-01-01 01:00:00	86.98	2018	1	1
2	2018-01-01 02:00:00	86.06	2018	1	2
3	2018-01-01 03:00:00	94.04	2018	1	3
4	2018-01-01 04:00:00	108.78	2018	1	4
…	…	…	…	…	…
31856	2021-12-31 19:00:00	72.25	2021	12	19
31857	2021-12-31 20:00:00	71.56	2021	12	20
31858	2021-12-31 21:00:00	70.27	2021	12	21
31859	2021-12-31 22:00:00	66.78	2021	12	22
31860	2021-12-31 23:00:00	61.29	2021	12	23

[31861 rows x 5 columns]

[81]:
```
Hour_avg=air.groupby("Hour")["PM2.5"].mean()
```

[82]:
```
Hour_avg
```

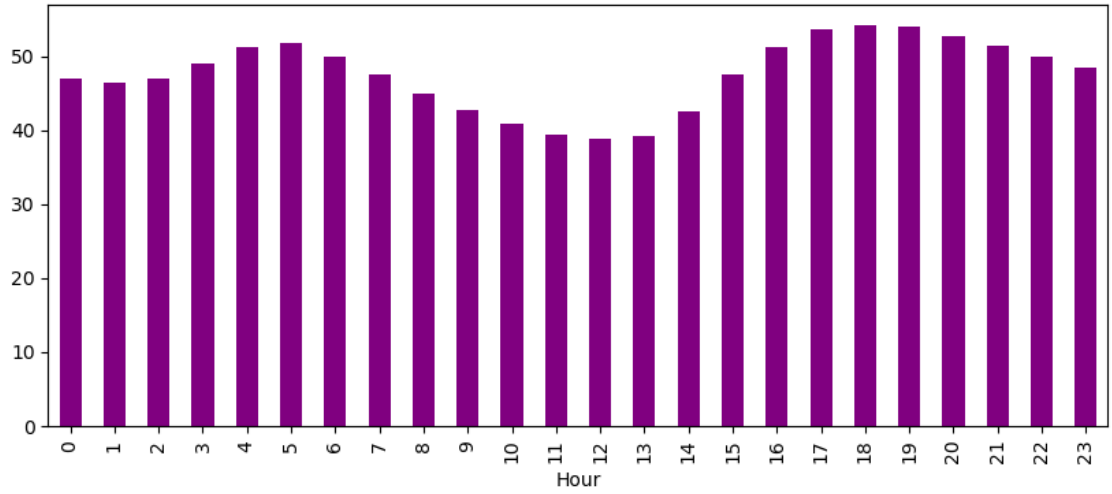
[82]:

Hour	
0	47.068769
1	46.403946
2	46.952281
3	49.150798
4	51.296204
5	51.908798
6	49.907545
7	47.518554
8	44.996127
9	42.681736
10	40.931766
11	39.393820
12	38.789992
13	39.323346
14	42.638162
15	47.500464
16	51.244844
17	53.669757
18	54.317363
19	54.120371
20	52.784130
21	51.383525
22	49.940615

```
23    48.538064
Name: PM2.5, dtype: float64
```

[100]:
```
Hour_avg.plot(kind="bar",figsize=(10,4),color="purple")
```

[100]:
```
<Axes: xlabel='Hour'>
```



4 IN WHICH MONTHS THE AIR WAS RECORDED "VERY UNHEALTHY"

[109]:
```
air.head(2)
```

[109]:

	Timestamp	PM2.5	year	month	Hour
0	2018-01-01 00:00:00	90.19	2018	1	0
1	2018-01-01 01:00:00	86.98	2018	1	1

[112]:
```
x=air[(air["PM2.5"]>=150.5) & (air["PM2.5"]<=250.4)]
```

[113]:
```
x
```

[113]:

	Timestamp	PM2.5	year	month	Hour
6	2018-01-06 17:00:00	150.89	2018	1	17
7	2018-11-07 21:00:00	245.63	2018	11	21
8	2018-11-08 00:00:00	175.19	2018	11	0
31	2018-11-07 22:00:00	232.03	2018	11	22
32	2018-11-08 01:00:00	169.15	2018	11	1
55	2018-11-07 20:00:00	223.21	2018	11	20
56	2018-11-08 03:00:00	160.34	2018	11	3
79	2018-11-07 18:00:00	212.56	2018	11	18

```
103   2018-11-07 19:00:00  207.41  2018     11    19
126   2018-11-07 17:00:00  183.19  2018     11    17
143   2018-11-07 23:00:00  173.67  2018     11    23
284   2018-01-13 16:00:00  165.45  2018      1    16
285   2018-01-13 17:00:00  177.58  2018      1    17
286   2018-01-13 18:00:00  175.32  2018      1    18
287   2018-01-13 19:00:00  166.60  2018      1    19
7578  2019-02-01 17:00:00  165.82  2019      2    17
7579  2019-02-01 18:00:00  157.99  2019      2    18
13697 2019-10-27 22:00:00  151.52  2019     10    22
13698 2019-10-27 23:00:00  219.07  2019     10    23
13699 2019-10-28 00:00:00  234.83  2019     10     0
13700 2019-10-28 01:00:00  230.05  2019     10     1
13701 2019-10-28 02:00:00  176.81  2019     10     2
22450 2020-11-14 18:00:00  183.39  2020     11    18
22451 2020-11-14 19:00:00  196.88  2020     11    19
22452 2020-11-14 20:00:00  178.61  2020     11    20
22455 2020-11-14 23:00:00  181.07  2020     11    23
22456 2020-11-15 00:00:00  193.80  2020     11     0
22457 2020-11-15 01:00:00  172.04  2020     11     1
```

[116]:
```
x.value_counts("month")
```

[116]:
```
month
11    16
1      5
10     5
2      2
dtype: int64
```

5 IN YEAR 2018 , HOW MANY TIMES THE AQI WAS RECORDED “MODERATE”

[118]:
```
air.head(2)
```

[118]:
```
            Timestamp  PM2.5  year  month  Hour
0 2018-01-01 00:00:00  90.19  2018      1     0
1 2018-01-01 01:00:00  86.98  2018      1     1
```

[133]:
```
x=air[(air["year"]==2018)&(air["PM2.5"]>=12.1)&(air["PM2.5"]<=35.4)] # give new␣
↪variable name to this x
```

[134]:
```
x
```

[134]:
```
                Timestamp  PM2.5  year  month  Hour
1820  2018-01-04 04:00:00  33.34  2018      1     4
```

```
2038  2018-12-04 00:00:00  35.39  2018     12     0
2215  2018-04-19 13:00:00  28.90  2018      4    13
2306  2018-04-23 09:00:00  34.83  2018      4     9
2307  2018-04-23 11:00:00  33.50  2018      4    11
…                     …    …    …    …    …
9981  2018-06-28 06:00:00  19.75  2018      6     6
9982  2018-07-17 07:00:00  19.27  2018      7     7
9983  2018-07-16 08:00:00  18.96  2018      7     8
10004 2018-08-25 06:00:00  17.45  2018      8     6
10006 2018-07-17 08:00:00  18.77  2018      7     8

[2239 rows x 5 columns]
```

[135]:
```
x.year.unique()
```

[135]:
```
array([2018], dtype=int64)
```

[137]:
```
x["PM2.5"]
```

[137]:
```
1820     33.34
2038     35.39
2215     28.90
2306     34.83
2307     33.50
         …
9981     19.75
9982     19.27
9983     18.96
10004    17.45
10006    18.77
Name: PM2.5, Length: 2239, dtype: float64
```

6 HOW WAS THE WEATHER IN THE MONTH OF JAN & JULY

[139]:
```
air.head(2)
```

[139]:
```
            Timestamp  PM2.5  year  month  Hour
0 2018-01-01 00:00:00  90.19  2018      1     0
1 2018-01-01 01:00:00  86.98  2018      1     1
```

[143]:
```
air[air["month"]==1]["PM2.5"].mean() # UNHEALTHY
```

[143]:
```
65.82906685236769
```

[145]:
```
air[air["month"]==7]["PM2.5"].mean() #MODERATE
```

[145]: 31.88416146645866

[ ]: