```
import pandas as pd
data = pd.read_excel('export.xlsx')
```

```
data
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	8.8	NaN	1010.8	NaN
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	10.0	NaN	1010.6	NaN
2	2023-06-03	26.2	21.1	32.6	0.7	NaN	272	10.7	NaN	1010.4	NaN
3	2023-06-04	26.4	23.0	33.4	0.5	NaN	258	12.5	NaN	1010.8	NaN
4	2023-06-05	25.7	20.9	31.0	4.2	NaN	260	12.9	NaN	1010.0	NaN
...	...	...	...	...	...	...	...	...	...	...	...
380	2024-06-15	25.4	21.1	29.9	0.0	NaN	264	15.7	NaN	1009.0	NaN
381	2024-06-16	24.0	19.7	29.3	0.5	NaN	254	15.1	NaN	1009.6	NaN
382	2024-06-17	23.9	20.5	29.9	0.7	NaN	255	14.2	NaN	1009.7	NaN
383	2024-06-18	23.4	20.1	29.0	5.9	NaN	258	15.7	NaN	1009.6	NaN
384	2024-06-19	23.4	20.1	28.7	0.7	NaN	259	18.4	NaN	1008.5	NaN

385 rows × 11 columns

HEAD

```
data.head(5)
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	8.8	NaN	1010.8	NaN
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	10.0	NaN	1010.6	NaN
2	2023-06-03	26.2	21.1	32.6	0.7	NaN	272	10.7	NaN	1010.4	NaN
3	2023-06-04	26.4	23.0	33.4	0.5	NaN	258	12.5	NaN	1010.8	NaN
4	2023-06-05	25.7	20.9	31.0	4.2	NaN	260	12.9	NaN	1010.0	NaN

SHAPE

```
data.shape
```

```
(385, 11)
```

COLUMNS

```
data.columns
```

```
Index(['date', 'tavg', 'tmin', 'tmax', 'prcp', 'snow', 'wdir', 'wspd', 'wpgt',
       'pres', 'tsun'],
      dtype='object')
```

INFO

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 385 entries, 0 to 384
Data columns (total 11 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   date    385 non-null    object 
 1   tavg    385 non-null    float64
 2   tmin    385 non-null    float64
 3   tmax    385 non-null    float64
 4   prcp    385 non-null    float64
 5   snow    0 non-null      float64
 6   wdir    385 non-null    int64  
 7   wspd    385 non-null    float64
 8   wpgt    0 non-null      float64
 9   pres    385 non-null    float64
 10  tsun    0 non-null      float64
dtypes: float64(9), int64(1), object(1)
memory usage: 33.2+ KB
```

DESCRIBE

```
data.describe()
```

	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
count	385.000000	385.000000	385.000000	385.000000	0.0	385.000000	385.000000	0.0	385.000000	0.0
mean	24.622338	19.736104	30.498961	2.969091	NaN	175.974026	11.845974	NaN	1012.015844	NaN
std	2.305320	2.093733	2.959686	9.979808	NaN	84.035548	3.614819	NaN	2.689947	NaN
min	20.500000	13.600000	23.600000	0.000000	NaN	25.000000	5.300000	NaN	1004.500000	NaN
25%	23.000000	18.500000	28.500000	0.000000	NaN	95.000000	9.400000	NaN	1010.000000	NaN
50%	24.200000	20.200000	29.900000	0.000000	NaN	147.000000	11.000000	NaN	1011.600000	NaN
75%	25.900000	21.200000	32.000000	1.300000	NaN	260.000000	13.900000	NaN	1014.200000	NaN
max	31.100000	24.600000	39.500000	111.000000	NaN	350.000000	24.300000	NaN	1018.100000	NaN

NULL VALUES

```
data.isnull()
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	False	False	False	False	False	True	False	False	True	False	True
1	False	False	False	False	False	True	False	False	True	False	True
2	False	False	False	False	False	True	False	False	True	False	True
3	False	False	False	False	False	True	False	False	True	False	True
4	False	False	False	False	False	True	False	False	True	False	True
...	...	...	...	...	...	...	...	...	...	...	...
380	False	False	False	False	False	True	False	False	True	False	True
381	False	False	False	False	False	True	False	False	True	False	True
382	False	False	False	False	False	True	False	False	True	False	True
383	False	False	False	False	False	True	False	False	True	False	True
384	False	False	False	False	False	True	False	False	True	False	True

385 rows × 11 columns

SUM OF NULL VALUES

```
data.isnull().sum()
```

```
date      0
tavg      0
tmin      0
tmax      0
```

```
prcp      0
snow    385
wdir      0
wspd      0
wpgt    385
pres      0
tsun    385
dtype: int64
```

NOT NULL VALUES

```
data.notnull()
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	True	True	True	True	True	False	True	True	False	True	False
1	True	True	True	True	True	False	True	True	False	True	False
2	True	True	True	True	True	False	True	True	False	True	False
3	True	True	True	True	True	False	True	True	False	True	False
4	True	True	True	True	True	False	True	True	False	True	False
...	...	...	...	...	...	...	...	...	...	...	...
380	True	True	True	True	True	False	True	True	False	True	False
381	True	True	True	True	True	False	True	True	False	True	False
382	True	True	True	True	True	False	True	True	False	True	False
383	True	True	True	True	True	False	True	True	False	True	False
384	True	True	True	True	True	False	True	True	False	True	False

385 rows × 11 columns

SUM OF NOT NULL VALUE

```
data.notnull().sum()
```

```
date    385
tavg    385
tmin    385
tmax    385
prcp    385
snow      0
wdir    385
wspd    385
wpgt      0
pres    385
tsun      0
dtype: int64
```

VALUE_COUNT

```
data.value_counts
```

```
<bound method DataFrame.value_counts of           date  tavg  tmin  tmax  prcp  snow  wdir  wspd  wpgt    pres  tsun
0    2023-06-01  26.3  21.3  31.4   0.0   NaN   248   8.8   NaN  1010.8   NaN
1    2023-06-02  26.4  22.6  31.4   0.0   NaN   264  10.0   NaN  1010.6   NaN
2    2023-06-03  26.2  21.1  32.6   0.7   NaN   272  10.7   NaN  1010.4   NaN
3    2023-06-04  26.4  23.0  33.4   0.5   NaN   258  12.5   NaN  1010.8   NaN
4    2023-06-05  25.7  20.9  31.0   4.2   NaN   260  12.9   NaN  1010.0   NaN
..          ...   ...   ...   ...   ...   ...   ...   ...   ...     ...   ...
380  2024-06-15  25.4  21.1  29.9   0.0   NaN   264  15.7   NaN  1009.0   NaN
381  2024-06-16  24.0  19.7  29.3   0.5   NaN   254  15.1   NaN  1009.6   NaN
382  2024-06-17  23.9  20.5  29.9   0.7   NaN   255  14.2   NaN  1009.7   NaN
383  2024-06-18  23.4  20.1  29.0   5.9   NaN   258  15.7   NaN  1009.6   NaN
384  2024-06-19  23.4  20.1  28.7   0.7   NaN   259  18.4   NaN  1008.5   NaN

[385 rows x 11 columns]>
```

MEANS OF DATASET

```
data.mean()
```

```
C:\Users\SANTHOSHRAJ E\AppData\Local\Temp\ipykernel_6184\531903386.py:1: FutureWarning: The default value of numeric_only in DataFrame.mean is deprecated. In a future version, it will default to F
  data.mean()
tavg      24.622338
tmin      19.736104
tmax      30.498961
prcp       2.969091
snow            NaN
wdir     175.974026
wspd      11.845974
wpgt            NaN
pres    1012.015844
tsun            NaN
dtype: float64
```

STANDARD DEVIATION OF DATASET

```
data.std()
```

```
C:\Users\SANTHOSHRAJ E\AppData\Local\Temp\ipykernel_6184\2723740006.py:1: FutureWarning: The default value of numeric_only in DataFrame.std is deprecated. In a future version, it will default to F
  data.std()
tavg     2.305320
tmin     2.093733
tmax     2.959686
prcp     9.979808
snow          NaN
wdir    84.035548
```

```
wspd     3.614819
wpgt          NaN
pres     2.689947
tsun          NaN
dtype: float64
```

```
data.head(2)
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	8.8	NaN	1010.8	NaN
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	10.0	NaN	1010.6	NaN

FIND THE RECORD OF THE DATE IS '2024-01-01'

```
data[data['date'] == '2024-01-01']
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
214	2024-01-01	21.7	15.9	28.2	0.0	NaN	100	10.2	NaN	1015.3	NaN

FIND THE RECORDS OF TMAX IS GREATER THAN 30 AND TOTAL VALUES

```
data[data['tmax'] > 30]
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	8.8	NaN	1010.8	NaN
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	10.0	NaN	1010.6	NaN
2	2023-06-03	26.2	21.1	32.6	0.7	NaN	272	10.7	NaN	1010.4	NaN
3	2023-06-04	26.4	23.0	33.4	0.5	NaN	258	12.5	NaN	1010.8	NaN
4	2023-06-05	25.7	20.9	31.0	4.2	NaN	260	12.9	NaN	1010.0	NaN
...	...	...	...	...	...	...	...	...	...	...	...
366	2024-06-01	25.4	22.4	32.8	8.1	NaN	278	12.3	NaN	1008.8	NaN
367	2024-06-02	24.3	20.4	33.8	30.0	NaN	281	9.5	NaN	1009.3	NaN
368	2024-06-03	24.9	19.3	31.8	111.0	NaN	280	10.1	NaN	1009.9	NaN
370	2024-06-05	25.1	20.9	31.4	0.5	NaN	247	9.3	NaN	1011.0	NaN
371	2024-06-06	24.3	21.6	31.2	6.1	NaN	251	10.8	NaN	1011.4	NaN

174 rows × 11 columns

```
data[data['tmax'] > 30].shape
```

```
(174, 11)
```

FIND THE RECORDS OF TMIN IS LESS THAN 20 AND TOTAL VALUES

```
data[data['tmin'] < 20]
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
42	2023-07-13	22.6	19.7	28.2	11.9	NaN	248	13.8	NaN	1010.1	NaN
53	2023-07-24	21.2	19.9	26.6	7.1	NaN	258	20.4	NaN	1009.9	NaN
55	2023-07-26	21.1	19.4	24.5	12.2	NaN	256	16.6	NaN	1011.4	NaN
56	2023-07-27	23.4	18.8	29.2	6.9	NaN	258	19.0	NaN	1011.2	NaN
57	2023-07-28	23.2	19.2	27.5	0.0	NaN	266	18.1	NaN	1011.0	NaN
...	...	...	...	...	...	...	...	...	...	...	...
301	2024-03-28	27.7	19.2	34.9	0.0	NaN	120	12.9	NaN	1013.3	NaN
314	2024-04-10	28.1	19.2	35.3	0.0	NaN	109	11.8	NaN	1011.4	NaN
320	2024-04-16	28.4	19.7	35.8	0.0	NaN	122	12.7	NaN	1011.7	NaN
368	2024-06-03	24.9	19.3	31.8	111.0	NaN	280	10.1	NaN	1009.9	NaN
381	2024-06-16	24.0	19.7	29.3	0.5	NaN	254	15.1	NaN	1009.6	NaN

175 rows × 11 columns

```
data[data['tmin'] < 20].shape
```

```
(175, 11)
```

FIND THE RECORDS OF TAVG IS GREATER THAN 25 AND TAVG IS EQUAL TO 25

```
data[(data['tavg'] > 25) | (data['tavg'] == 25)]
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	8.8	NaN	1010.8	NaN
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	10.0	NaN	1010.6	NaN
2	2023-06-03	26.2	21.1	32.6	0.7	NaN	272	10.7	NaN	1010.4	NaN
3	2023-06-04	26.4	23.0	33.4	0.5	NaN	258	12.5	NaN	1010.8	NaN
4	2023-06-05	25.7	20.9	31.0	4.2	NaN	260	12.9	NaN	1010.0	NaN
...	...	...	...	...	...	...	...	...	...	...	...
365	2024-05-31	26.7	21.7	33.0	1.2	NaN	300	12.9	NaN	1008.1	NaN
366	2024-06-01	25.4	22.4	32.8	8.1	NaN	278	12.3	NaN	1008.8	NaN
369	2024-06-04	26.2	21.9	29.4	0.0	NaN	246	9.1	NaN	1010.6	NaN
370	2024-06-05	25.1	20.9	31.4	0.5	NaN	247	9.3	NaN	1011.0	NaN
380	2024-06-15	25.4	21.1	29.9	0.0	NaN	264	15.7	NaN	1009.0	NaN

141 rows × 11 columns

FIND THE RECORDS OF TAVG IS GREATER THAN 25, TMAX IS GREATER THAN 25 AND TMIN IS LESS THAN 25

```
data[(data['tavg'] > 25) & (data['tmax'] > 25) & (data['tmin'] < 25)]
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	8.8	NaN	1010.8	NaN
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	10.0	NaN	1010.6	NaN
2	2023-06-03	26.2	21.1	32.6	0.7	NaN	272	10.7	NaN	1010.4	NaN
3	2023-06-04	26.4	23.0	33.4	0.5	NaN	258	12.5	NaN	1010.8	NaN
4	2023-06-05	25.7	20.9	31.0	4.2	NaN	260	12.9	NaN	1010.0	NaN
...	...	...	...	...	...	...	...	...	...	...	...
365	2024-05-31	26.7	21.7	33.0	1.2	NaN	300	12.9	NaN	1008.1	NaN
366	2024-06-01	25.4	22.4	32.8	8.1	NaN	278	12.3	NaN	1008.8	NaN
369	2024-06-04	26.2	21.9	29.4	0.0	NaN	246	9.1	NaN	1010.6	NaN
370	2024-06-05	25.1	20.9	31.4	0.5	NaN	247	9.3	NaN	1011.0	NaN
380	2024-06-15	25.4	21.1	29.9	0.0	NaN	264	15.7	NaN	1009.0	NaN

138 rows × 11 columns

FIND THE RECORDS OF WSPD IS LESS THAN 10.

```
data[data['wspd'] < 10]
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	8.8	NaN	1010.8	NaN
70	2023-08-10	25.4	21.8	30.4	0.0	NaN	286	9.1	NaN	1013.0	NaN
71	2023-08-11	25.6	22.0	30.0	0.1	NaN	297	8.7	NaN	1011.6	NaN
74	2023-08-14	25.4	19.8	30.4	0.5	NaN	277	9.2	NaN	1011.3	NaN
82	2023-08-22	26.0	21.8	30.8	1.1	NaN	269	7.9	NaN	1009.5	NaN
...	...	...	...	...	...	...	...	...	...	...	...
355	2024-05-21	25.7	21.7	29.6	2.0	NaN	113	9.0	NaN	1006.8	NaN
357	2024-05-23	27.6	23.2	31.4	0.3	NaN	34	9.1	NaN	1004.5	NaN
367	2024-06-02	24.3	20.4	33.8	30.0	NaN	281	9.5	NaN	1009.3	NaN
369	2024-06-04	26.2	21.9	29.4	0.0	NaN	246	9.1	NaN	1010.6	NaN
370	2024-06-05	25.1	20.9	31.4	0.5	NaN	247	9.3	NaN	1011.0	NaN

125 rows × 11 columns

FIND THE RECORDS OF PRCP IS EQUAL TO ZERO AND WDIR IS GREATER THAN 250 OR WSPD IS LESS THAN 10 AND PRES IS GREATER THAN 1015 OF THEM.

```
data[(data['prcp'] == 0) & (data['wdir'] > 250) | (data['wspd'] < 10) & (data['pres'] > 1015)]
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	10.0	NaN	1010.6	NaN
10	2023-06-11	25.4	21.2	31.0	0.0	NaN	258	16.0	NaN	1008.4	NaN
13	2023-06-14	25.8	21.6	31.6	0.0	NaN	261	14.8	NaN	1010.9	NaN
15	2023-06-16	26.5	21.8	31.5	0.0	NaN	257	15.3	NaN	1009.8	NaN
18	2023-06-19	24.9	21.7	28.4	0.0	NaN	274	12.0	NaN	1008.8	NaN
...	...	...	...	...	...	...	...	...	...	...	...
361	2024-05-27	25.4	20.7	30.7	0.0	NaN	267	19.9	NaN	1006.8	NaN
362	2024-05-28	25.7	21.4	32.1	0.0	NaN	264	18.8	NaN	1007.3	NaN
363	2024-05-29	25.8	20.3	31.4	0.0	NaN	268	17.1	NaN	1007.1	NaN
378	2024-06-13	23.5	21.2	29.0	0.0	NaN	263	17.9	NaN	1011.3	NaN
380	2024-06-15	25.4	21.1	29.9	0.0	NaN	264	15.7	NaN	1009.0	NaN

65 rows × 11 columns

FIND THE RECORDS OF TAVG IS GREATER THAN 30 AND WSPD IS GREATER THAN 10

```
data[(data['tavg'] > 30) & (data['wspd'] > 10)]
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
322	2024-04-18	30.5	22.9	37.7	0.0	NaN	139	11.6	NaN	1009.4	NaN
329	2024-04-25	30.1	22.4	36.9	0.0	NaN	102	10.3	NaN	1008.5	NaN
330	2024-04-26	30.2	22.9	36.6	0.0	NaN	122	10.8	NaN	1008.5	NaN
335	2024-05-01	31.1	23.0	39.5	0.0	NaN	225	10.6	NaN	1006.9	NaN
339	2024-05-05	30.3	24.1	37.6	0.0	NaN	226	14.4	NaN	1007.9	NaN

CALCULATE THE VALUES ARE PRES CAN BE ADDED BY 5.

```
data['pres'] = data['pres'].apply(lambda x : x+5)
```

```
data.head(2)
```

	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	8.8	NaN	1015.8	NaN
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	10.0	NaN	1015.6	NaN

CALCULATE THE VALUES ARE WSPD CAN BE MULTIPLY BY 2

```
data['wspd'] = data['wspd'].apply(lambda x : x*2)
```

```
data.head(2)
```

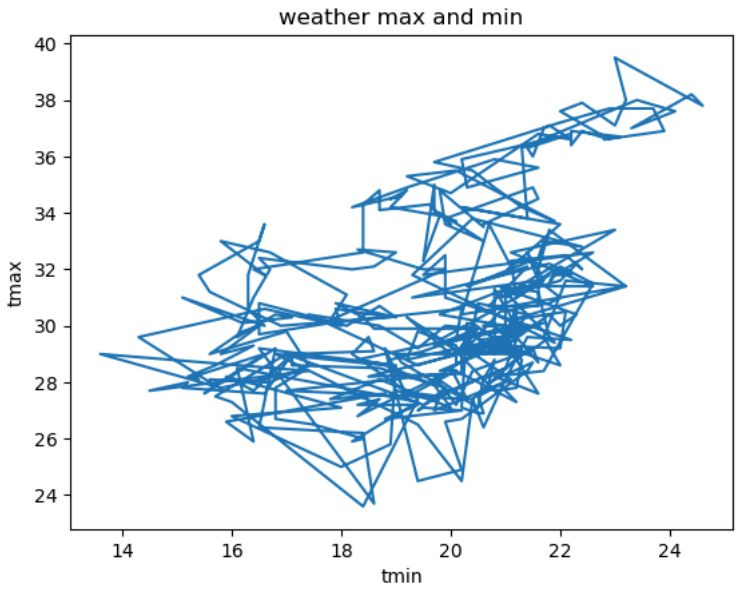
	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
0	2023-06-01	26.3	21.3	31.4	0.0	NaN	248	17.6	NaN	1015.8	NaN
1	2023-06-02	26.4	22.6	31.4	0.0	NaN	264	20.0	NaN	1015.6	NaN

```
data[data['pres'] == 1015]
```

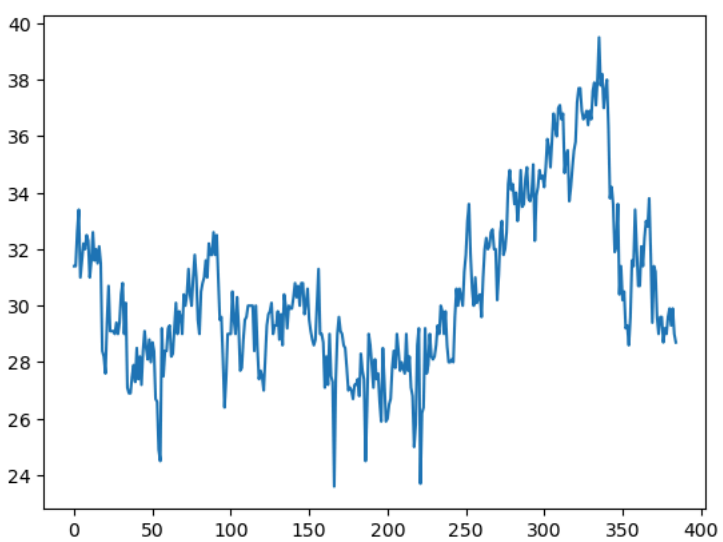
	date	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
4	2023-06-05	25.7	20.9	31.0	4.2	NaN	260	25.8	NaN	1015.0	NaN
118	2023-09-27	24.1	19.6	27.4	1.8	NaN	252	25.2	NaN	1015.0	NaN
306	2024-04-02	29.0	22.2	36.8	0.0	NaN	130	23.4	NaN	1015.0	NaN

```
Start coding or generate with AI.
```

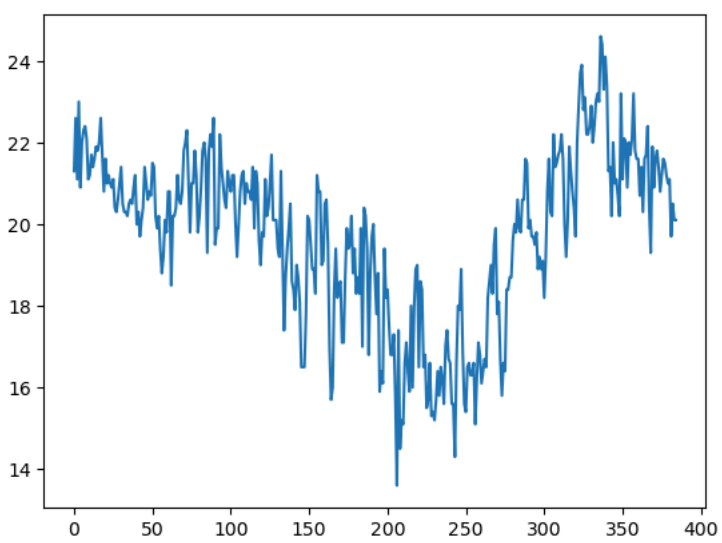
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

data = pd.read_excel('export.xlsx')
plt.plot(data['tmin'], data['tmax'])
plt.xlabel('tmin')
plt.ylabel('tmax')
plt.title('weather max and min')
plt.show()
```



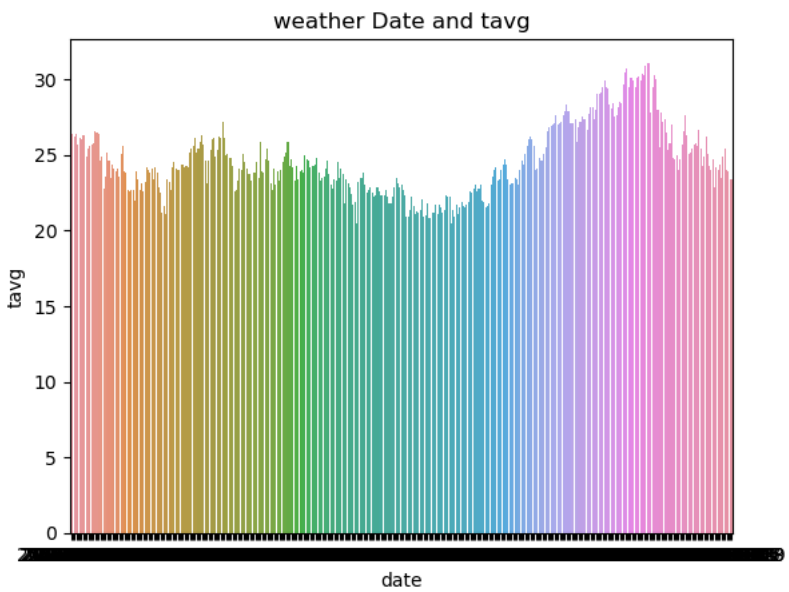

```
plt.plot(data['tmax'])
plt.show()
```



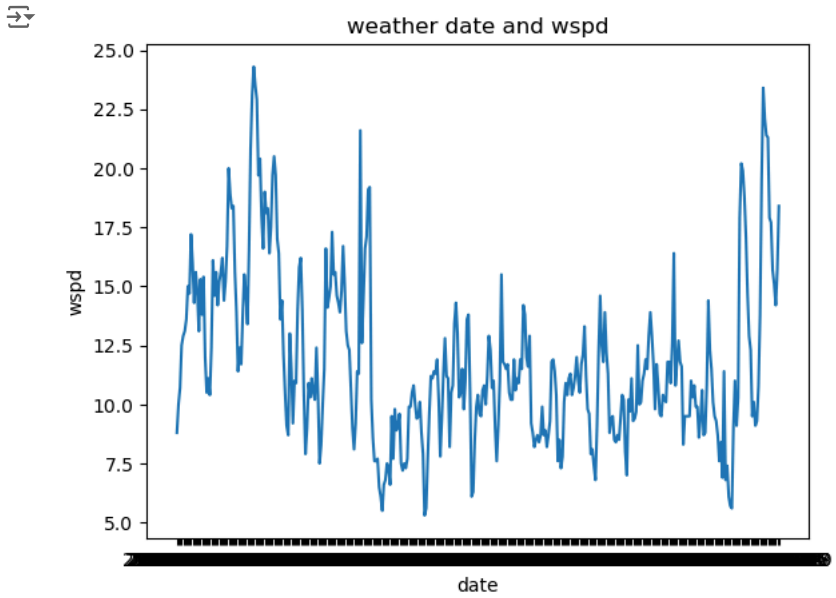

```
plt.plot(data['tmin'])
plt.show()
```



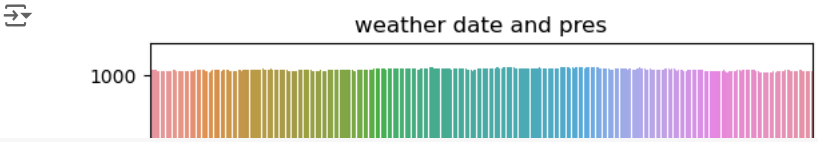

```
sns.barplot(x = 'date', y = 'tavg', data = data)
plt.xlabel('date')
plt.ylabel('tavg')
plt.title('weather Date and tavg')
plt.show()
```



```
sns.lineplot(x = 'date', y = 'wspd', data = data)
plt.xlabel('date')
plt.ylabel('wspd')
plt.title('weather date and wspd')
plt.show()
```




```
sns.barplot(x = 'date', y = 'pres', data = data)
plt.xlabel('date')
plt.ylabel('pres')
plt.title('weather date and pres')
plt.show()
```




```
sns.scatterplot(x = 'date', y = 'wdir', data = data)
plt.xlabel('date')
plt.ylabel('wdir')
plt.title('weather date and wdir')
plt.show()
```