

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [2]: df=pd.read_csv(r'C:\Users\user\Desktop\rainfall\SAURASHTRA KUTCH.csv')
df
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	2393	SAURASHTRA & KUTCH	1902	0.1	0.0	0.0	0.5	1.1	14.4	92.9	160.0	123.9	1.5
1	2394	SAURASHTRA & KUTCH	1903	0.5	0.0	1.7	0.0	3.1	10.5	337.9	96.1	61.9	11.1
2	2395	SAURASHTRA & KUTCH	1904	1.4	5.8	17.5	0.0	0.0	9.5	111.2	9.4	28.9	0.3
3	2396	SAURASHTRA & KUTCH	1905	1.5	1.0	0.6	0.4	0.0	6.4	254.5	12.3	12.8	0.4
4	2397	SAURASHTRA & KUTCH	1906	0.9	28.2	0.0	0.0	0.0	126.0	161.0	152.2	56.6	14.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...
109	2502	SAURASHTRA & KUTCH	2011	0.0	1.4	0.0	0.0	0.0	26.0	212.7	290.9	210.1	1.2
110	2503	SAURASHTRA & KUTCH	2012	0.0	0.0	0.0	0.2	0.1	22.4	34.7	34.5	228.5	2.4
111	2504	SAURASHTRA & KUTCH	2013	1.7	0.2	0.1	8.5	0.1	127.7	171.2	83.3	260.2	28.6
112	2505	SAURASHTRA & KUTCH	2014	0.3	0.0	0.1	0.5	2.1	17.3	137.7	118.8	99.2	5.2
113	2506	SAURASHTRA & KUTCH	2015	0.9	0.0	4.4	2.1	0.8	112.6	226.7	10.6	79.9	3.3

114 rows × 20 columns



```
In [3]: df=df.dropna()
df
```

Out[3]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	2393	SAURASHTRA & KUTCH	1902	0.1	0.0	0.0	0.5	1.1	14.4	92.9	160.0	123.9	1.5
1	2394	SAURASHTRA & KUTCH	1903	0.5	0.0	1.7	0.0	3.1	10.5	337.9	96.1	61.9	11.1
2	2395	SAURASHTRA & KUTCH	1904	1.4	5.8	17.5	0.0	0.0	9.5	111.2	9.4	28.9	0.3
3	2396	SAURASHTRA & KUTCH	1905	1.5	1.0	0.6	0.4	0.0	6.4	254.5	12.3	12.8	0.4
4	2397	SAURASHTRA & KUTCH	1906	0.9	28.2	0.0	0.0	0.0	126.0	161.0	152.2	56.6	14.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...
109	2502	SAURASHTRA & KUTCH	2011	0.0	1.4	0.0	0.0	0.0	26.0	212.7	290.9	210.1	1.2
110	2503	SAURASHTRA & KUTCH	2012	0.0	0.0	0.0	0.2	0.1	22.4	34.7	34.5	228.5	2.4
111	2504	SAURASHTRA & KUTCH	2013	1.7	0.2	0.1	8.5	0.1	127.7	171.2	83.3	260.2	28.6
112	2505	SAURASHTRA & KUTCH	2014	0.3	0.0	0.1	0.5	2.1	17.3	137.7	118.8	99.2	5.2
113	2506	SAURASHTRA & KUTCH	2015	0.9	0.0	4.4	2.1	0.8	112.6	226.7	10.6	79.9	3.3

114 rows × 20 columns



```
In [4]: df.columns
```

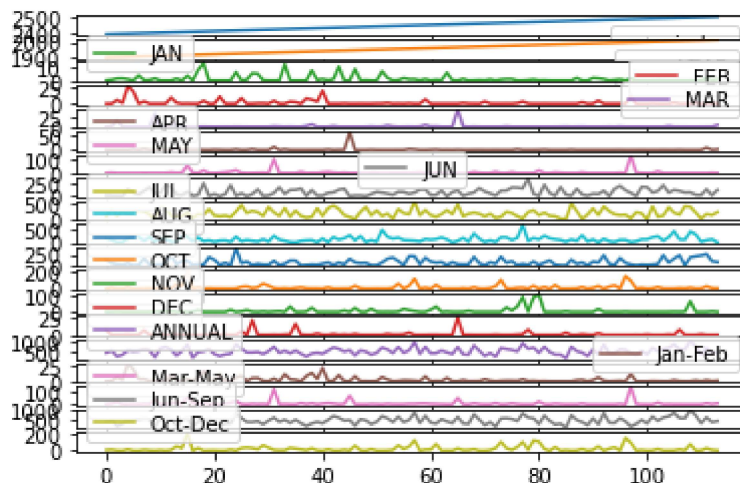
```
Out[4]: Index(['index', 'SUBDIVISION', 'YEAR', 'JAN', 'FEB', 'MAR', 'APR', 'MAY',
              'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC', 'ANNUAL', 'Jan-Feb',
              'Mar-May', 'Jun-Sep', 'Oct-Dec'],
              dtype='object')
```

In [5]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 114 entries, 0 to 113
Data columns (total 20 columns):
#   Column                Non-Null Count  Dtype
---  -
0   index                  114 non-null   int64
1   SUBDIVISION            114 non-null   object
2   YEAR                   114 non-null   int64
3   JAN                    114 non-null   float64
4   FEB                    114 non-null   float64
5   MAR                    114 non-null   float64
6   APR                    114 non-null   float64
7   MAY                    114 non-null   float64
8   JUN                    114 non-null   float64
9   JUL                    114 non-null   float64
10  AUG                    114 non-null   float64
11  SEP                    114 non-null   float64
12  OCT                    114 non-null   float64
13  NOV                    114 non-null   float64
14  DEC                    114 non-null   float64
15  ANNUAL                 114 non-null   float64
16  Jan-Feb               114 non-null   float64
17  Mar-May               114 non-null   float64
18  Jun-Sep               114 non-null   float64
19  Oct-Dec               114 non-null   float64
dtypes: float64(17), int64(2), object(1)
memory usage: 18.7+ KB
```

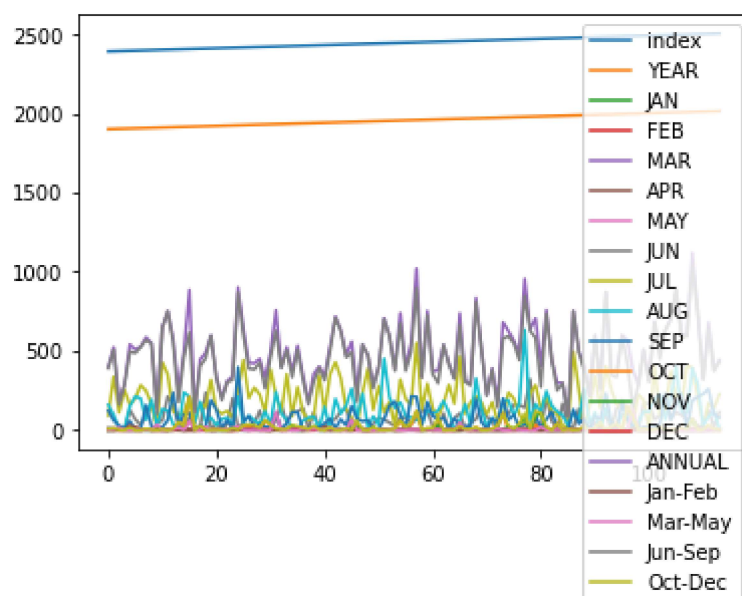
In [6]: df.plot.line(subplots=True)

Out[6]: array([<AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>], dtype=object)



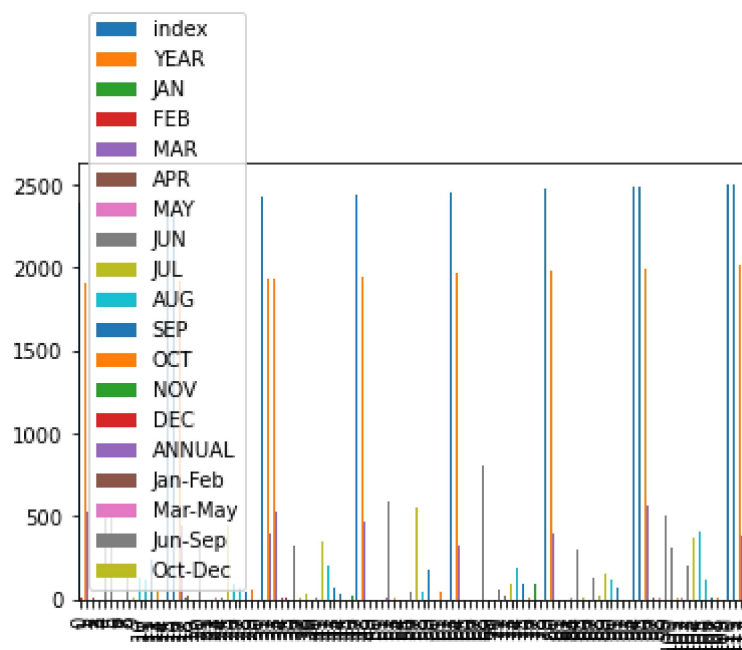
```
In [7]: df.plot.line()
```

```
Out[7]: <AxesSubplot:>
```



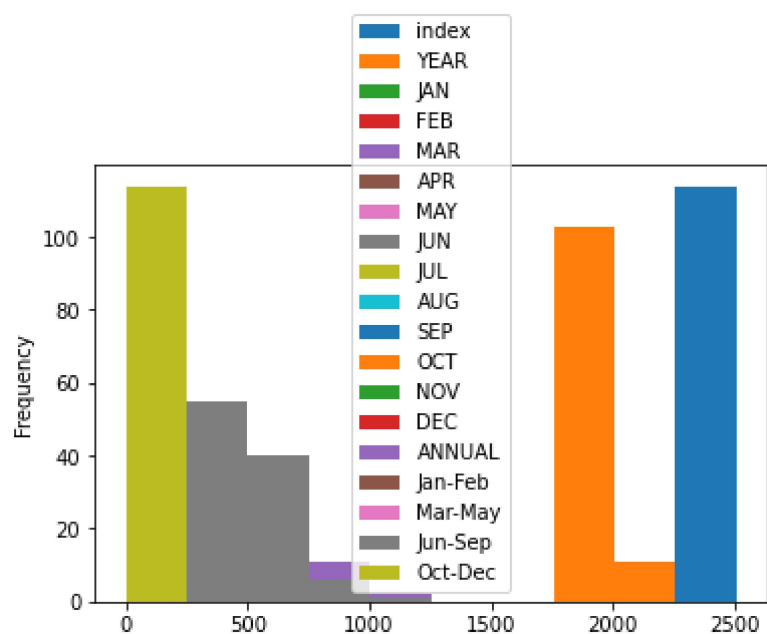
```
In [8]: df.plot.bar()
```

```
Out[8]: <AxesSubplot:>
```



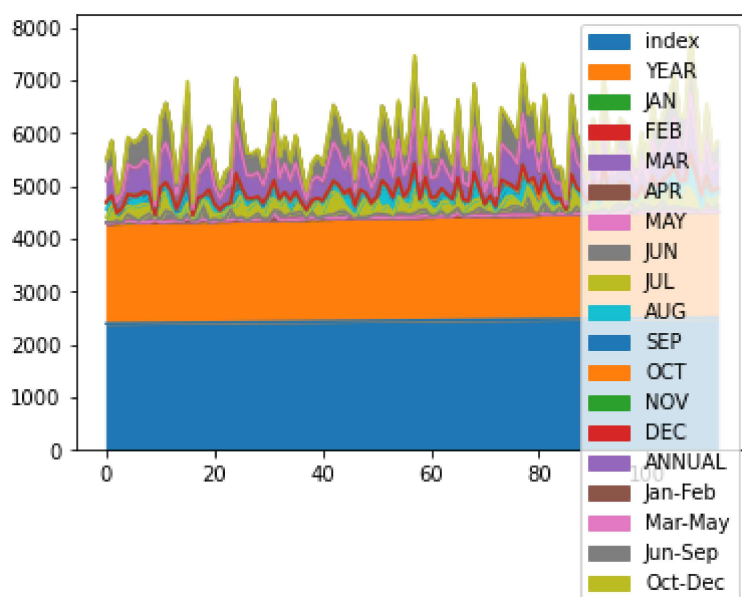
```
In [9]: df.plot.hist()
```

```
Out[9]: <AxesSubplot:ylabel='Frequency'>
```



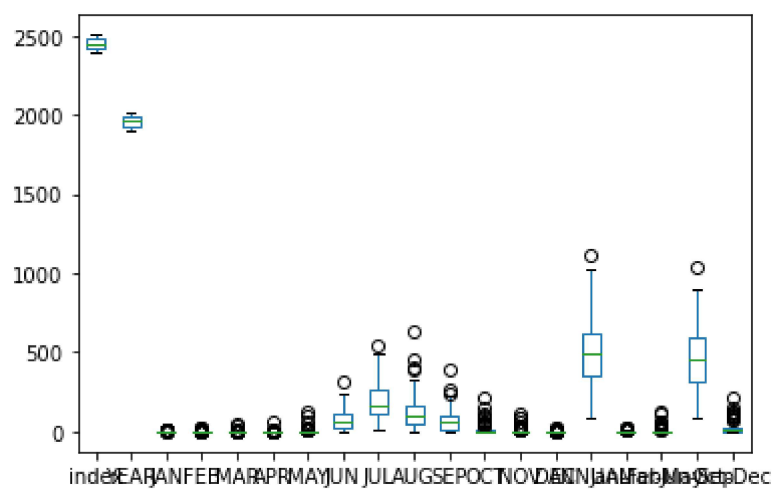
```
In [10]: df.plot.area()
```

```
Out[10]: <AxesSubplot:>
```



```
In [11]: df.plot.box()
```

```
Out[11]: <AxesSubplot:>
```

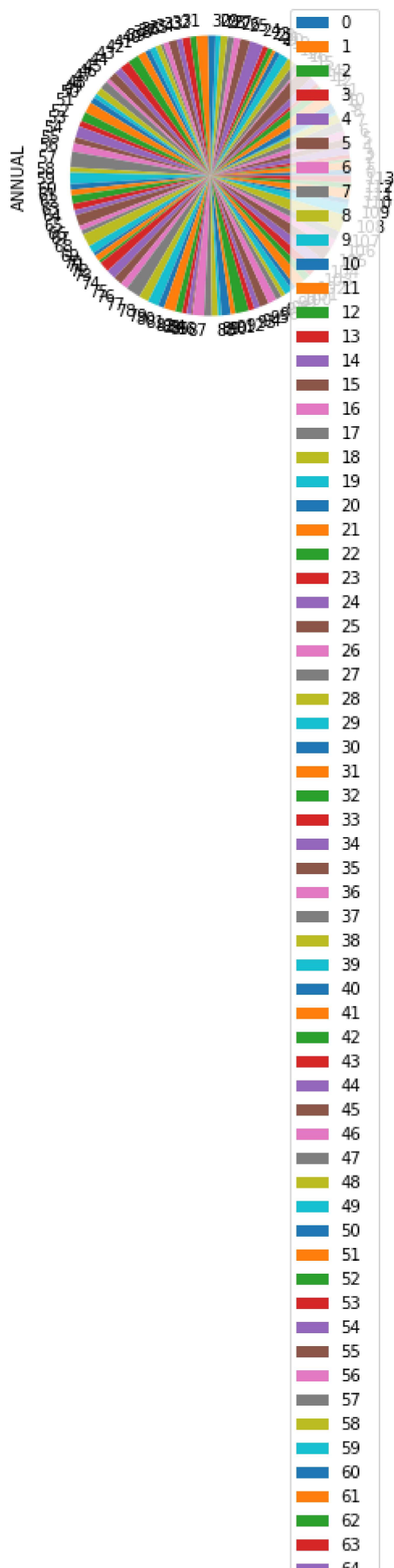




















































```
In [12]: df.plot.pie(y='ANNUAL')
```

```
Out[12]: <AxesSubplot:ylabel='ANNUAL'>
```



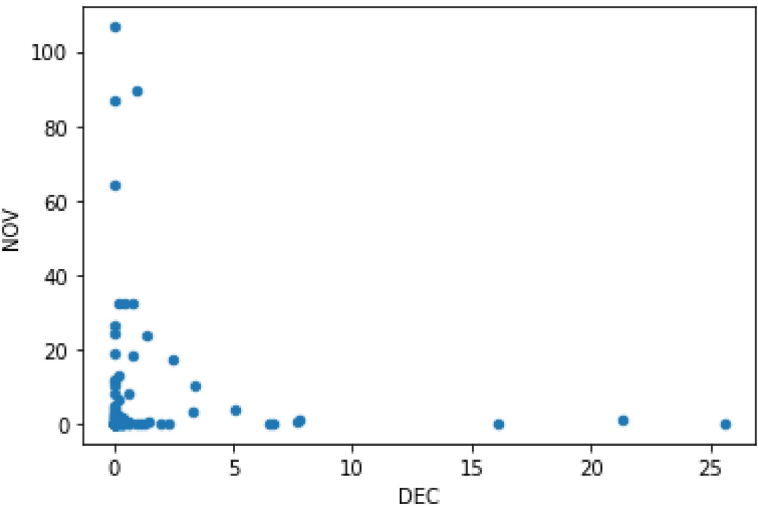




	64
	65
	66
	67
	68
	69
	70
	71
	72
	73
	74
	75
	76
	77
	78
	79
	80
	81
	82
	83
	84
	85
	86
	87
	88
	89
	90
	91
	92
	93
	94
	95
	96
	97
	98
	99
	100
	101
	102
	103
	104
	105
	106
	107
	108
	109
	110
	111
	112
	113

```
In [13]: df.plot.scatter(x='DEC',y='NOV')
```

Out[13]: <AxesSubplot:xlabel='DEC', ylabel='NOV'>



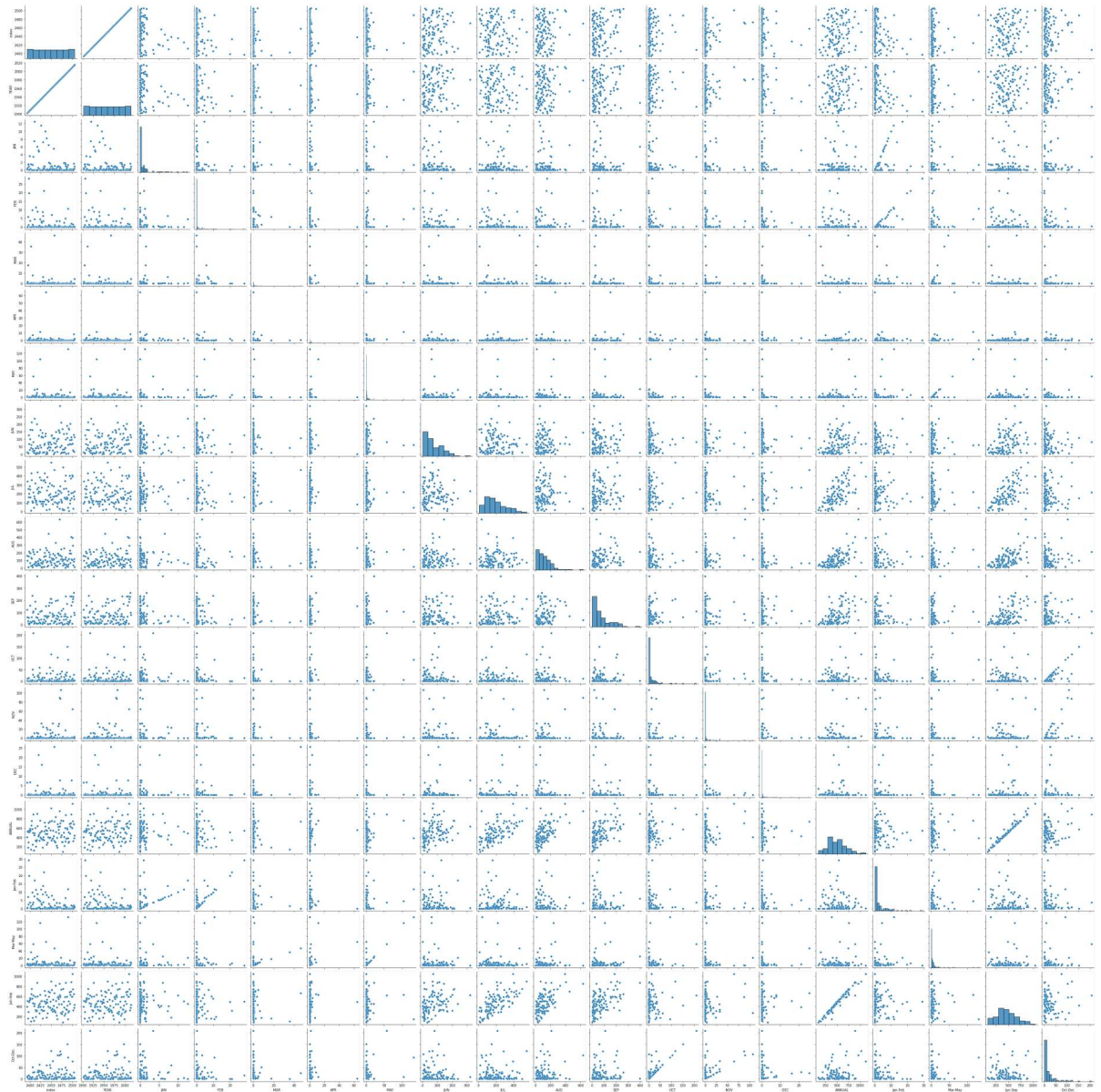
```
In [14]: df.describe()
```

Out[14]:

	index	YEAR	JAN	FEB	MAR	APR	MAY	
count	114.000000	114.000000	114.000000	114.000000	114.000000	114.000000	114.000000	114.000000
mean	2449.500000	1958.500000	1.132456	1.629825	1.307018	1.192105	4.675439	7.142857
std	33.052988	33.052988	2.384110	4.286714	5.715554	6.185340	16.659891	6.144554
min	2393.000000	1902.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	2421.250000	1930.250000	0.000000	0.000000	0.000000	0.000000	0.000000	2.000000
50%	2449.500000	1958.500000	0.150000	0.000000	0.000000	0.000000	0.500000	6.000000
75%	2477.750000	1986.750000	1.000000	0.575000	0.400000	0.500000	2.575000	11.000000
max	2506.000000	2015.000000	12.500000	28.200000	46.200000	64.400000	131.900000	32.000000

```
sns.pairplot(df)
```

```
<seaborn.axisgrid.PairGrid at 0x1b691b133d0>
```

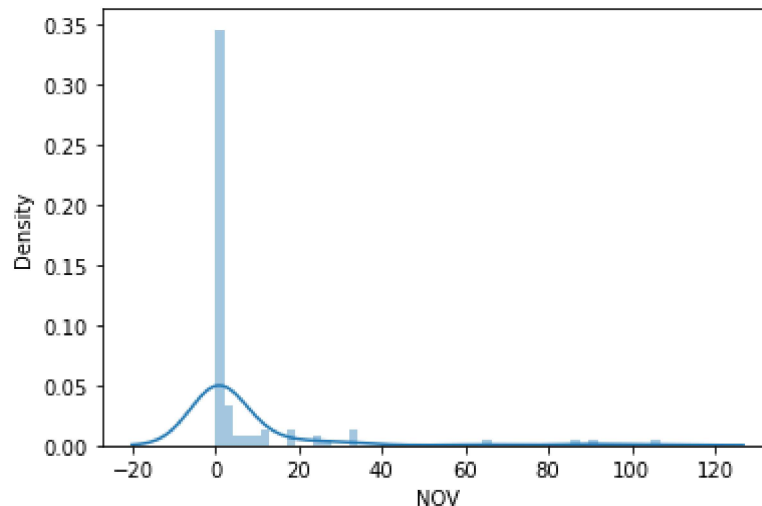


```
In [16]: sns.distplot(df['NOV'])
```

C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

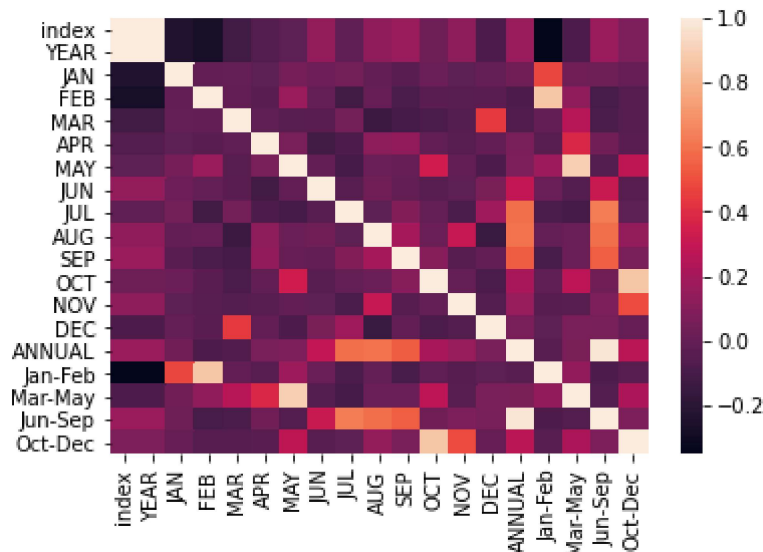
```
warnings.warn(msg, FutureWarning)
```

```
Out[16]: <AxesSubplot:xlabel='NOV', ylabel='Density'>
```



```
In [17]: sns.heatmap(df.corr())
```

```
Out[17]: <AxesSubplot:>
```



```
In [ ]:
```

