

```
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\MATATHWADA.csv')
df
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	2738	MATATHWADA	1902	1.3	0.0	0.4	7.2	0.8	52.4	120.9	85.2	273.3	61.3
1	2739	MATATHWADA	1903	2.6	0.8	0.0	1.7	58.3	104.4	264.2	281.9	173.3	139.9
2	2740	MATATHWADA	1904	0.0	0.9	12.1	0.3	7.2	79.2	118.4	57.3	339.0	76.2
3	2741	MATATHWADA	1905	1.3	2.0	0.0	6.6	4.8	84.6	94.8	137.6	157.8	15.4
4	2742	MATATHWADA	1906	19.8	0.0	0.1	0.0	0.2	220.6	254.9	156.9	82.1	19.8
...
109	2847	MATATHWADA	2011	0.0	3.8	0.7	3.5	3.1	79.2	230.1	228.5	90.0	24.8
110	2848	MATATHWADA	2012	0.0	0.0	0.0	0.6	2.3	72.2	161.1	101.4	120.0	68.8
111	2849	MATATHWADA	2013	1.5	9.4	2.6	7.9	6.4	160.9	293.4	136.9	154.1	94.3
112	2850	MATATHWADA	2014	1.4	13.4	79.0	11.9	7.0	30.4	105.0	178.9	84.5	14.2
113	2851	MATATHWADA	2015	10.1	1.6	32.0	39.6	12.3	118.3	27.4	112.2	154.3	19.5

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	2738	MATATHWADA	1902	1.3	0.0	0.4	7.2	0.8	52.4	120.9	85.2	273.3	61.3
1	2739	MATATHWADA	1903	2.6	0.8	0.0	1.7	58.3	104.4	264.2	281.9	173.3	139.9
2	2740	MATATHWADA	1904	0.0	0.9	12.1	0.3	7.2	79.2	118.4	57.3	339.0	76.2
3	2741	MATATHWADA	1905	1.3	2.0	0.0	6.6	4.8	84.6	94.8	137.6	157.8	15.4
4	2742	MATATHWADA	1906	19.8	0.0	0.1	0.0	0.2	220.6	254.9	156.9	82.1	19.8
...
109	2847	MATATHWADA	2011	0.0	3.8	0.7	3.5	3.1	79.2	230.1	228.5	90.0	24.8
110	2848	MATATHWADA	2012	0.0	0.0	0.0	0.6	2.3	72.2	161.1	101.4	120.0	68.8
111	2849	MATATHWADA	2013	1.5	9.4	2.6	7.9	6.4	160.9	293.4	136.9	154.1	94.3
112	2850	MATATHWADA	2014	1.4	13.4	79.0	11.9	7.0	30.4	105.0	178.9	84.5	14.2
113	2851	MATATHWADA	2015	10.1	1.6	32.0	39.6	12.3	118.3	27.4	112.2	154.3	19.5

114 rows × 20 columns



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

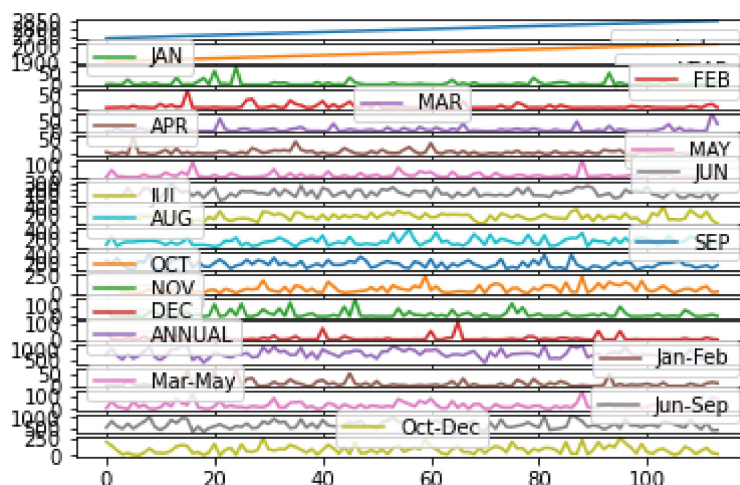
Out[4]: Index(['index', 'SUBDIVISION', 'YEAR', 'JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC', 'ANNUAL', 'Jan-Feb', 'Mar-May', 'Jun-Sep', 'Oct-Dec'], dtype='object')

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
```

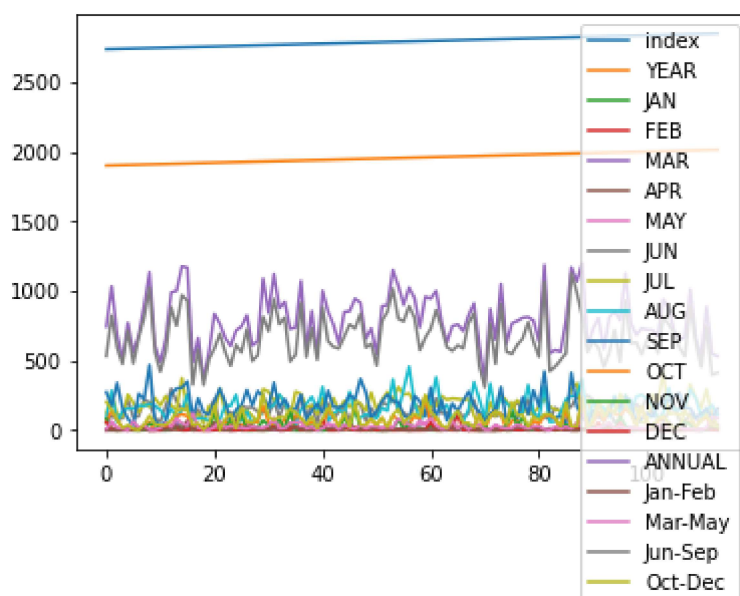
In [6]: `df.plot.line(subplots=True)`

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



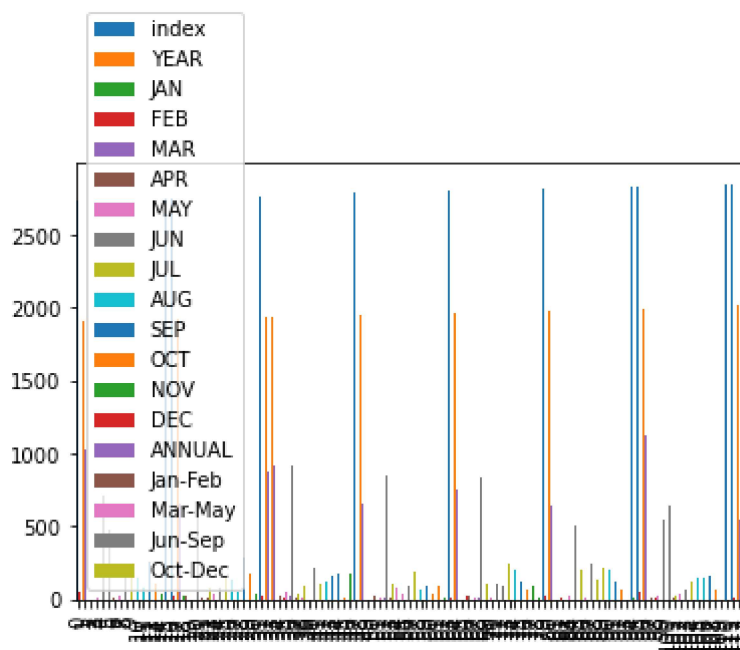
```
In [7]: df.plot.line()
```

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



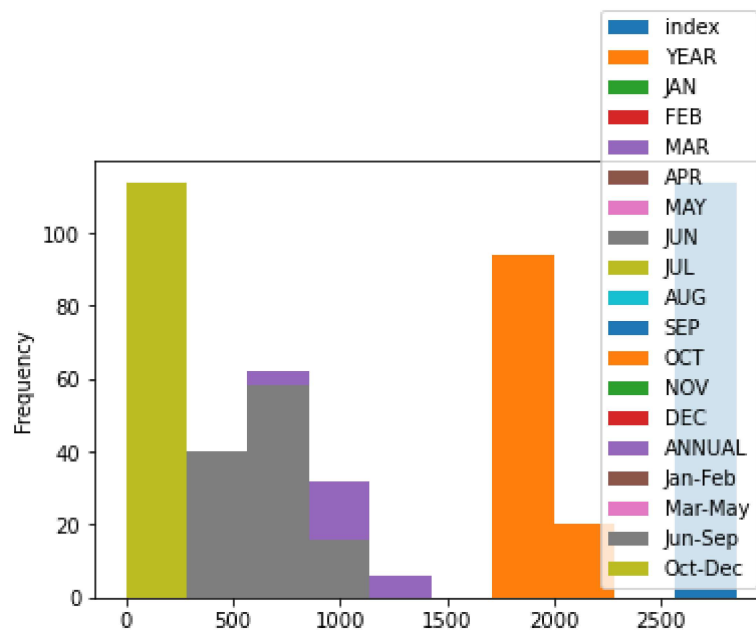
```
In [8]: df.plot.bar()
```

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



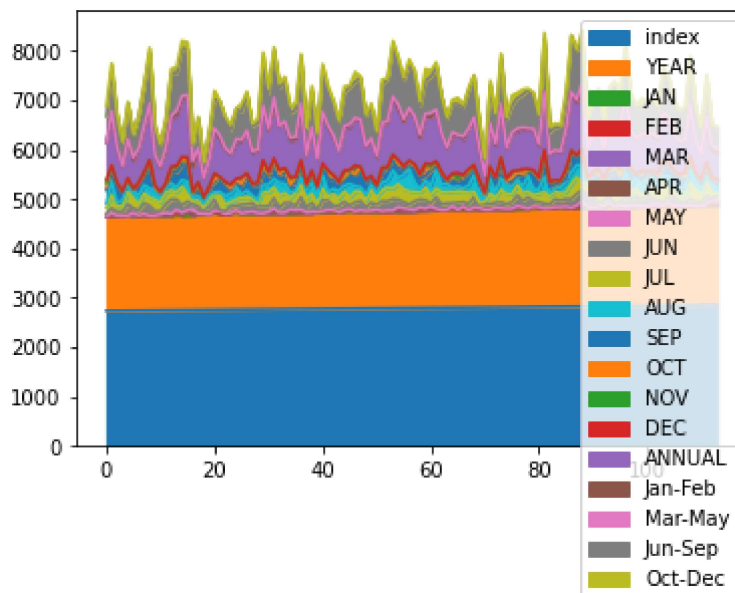
```
In [9]: df.plot.hist()
```

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



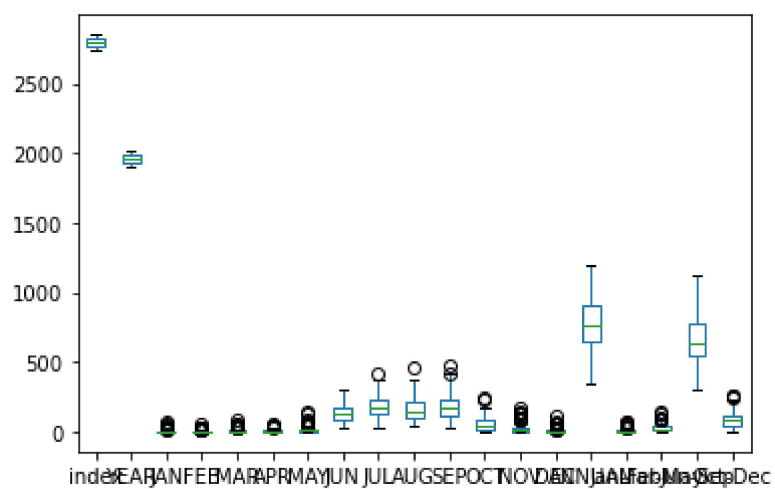
```
In [10]: df.plot.area()
```

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



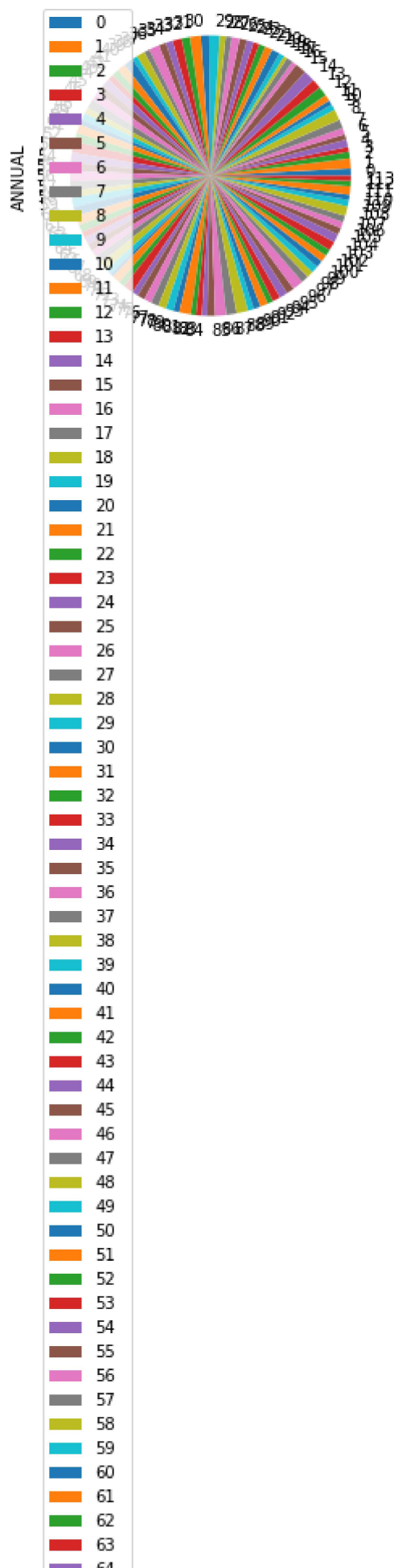
```
In [11]: df.plot.box()
```

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



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

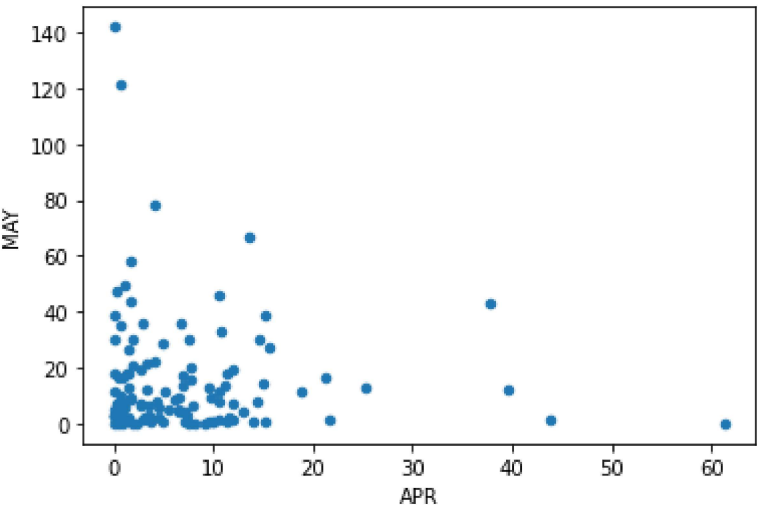
```
Out[12]: <AxesSubplot:ylabel='ANNUAL'>
```

	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='APR',y='MAY')
```

Out[13]: <AxesSubplot:xlabel='APR', ylabel='MAY'>



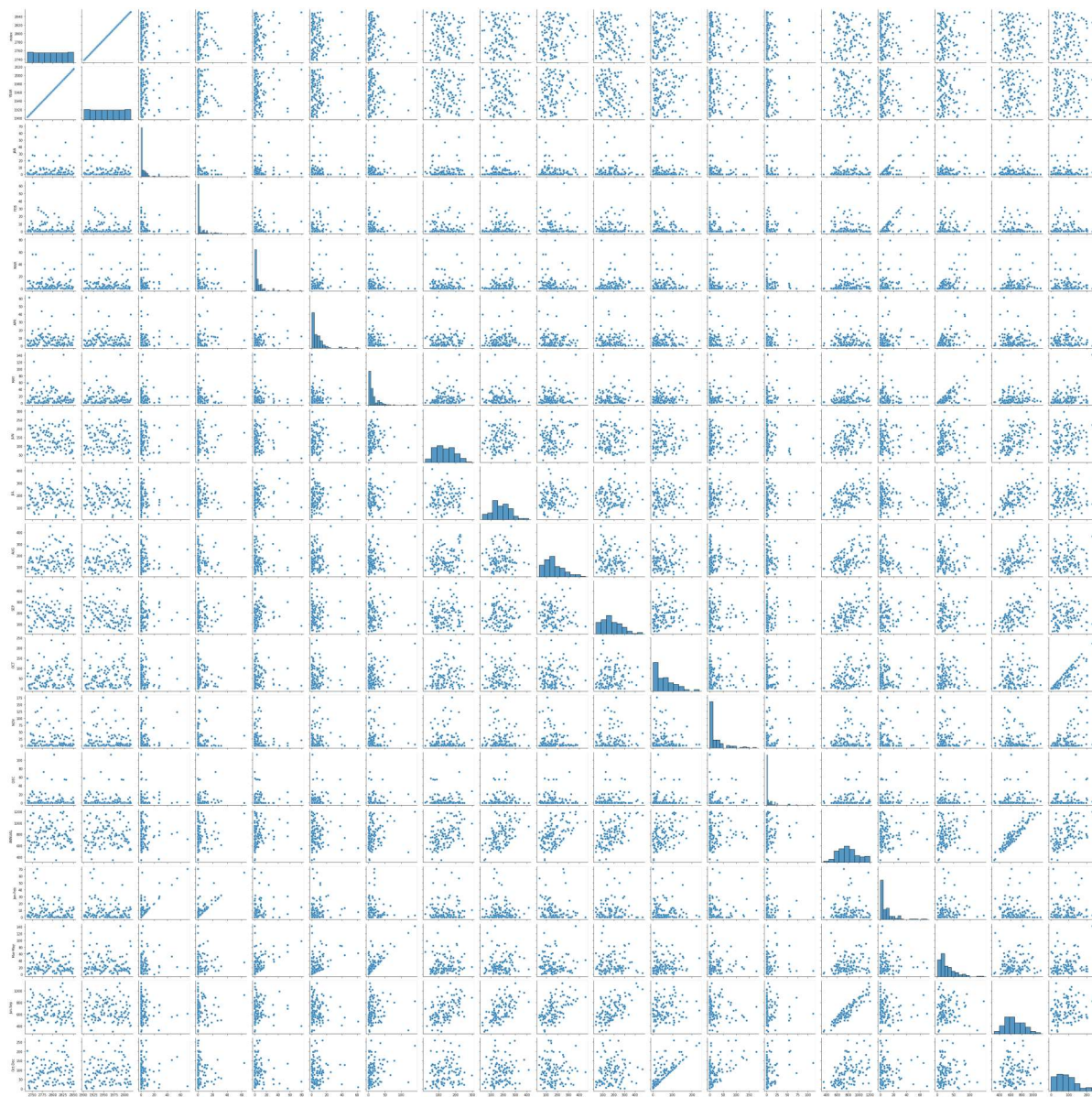
```
In [14]: df.describe()
```

Out[14]:

	index	YEAR	JAN	FEB	MAR	APR	MAY	
count	114.000000	114.000000	114.000000	114.000000	114.000000	114.000000	114.000000	11
mean	2794.500000	1958.500000	4.906140	4.453509	6.885965	7.235965	15.551754	13
std	33.052988	33.052988	10.642994	8.764552	12.374309	9.301890	21.899150	5
min	2738.000000	1902.000000	0.000000	0.000000	0.000000	0.000000	0.000000	2
25%	2766.250000	1930.250000	0.000000	0.000000	0.200000	1.500000	2.150000	9
50%	2794.500000	1958.500000	0.800000	0.700000	2.500000	4.500000	7.850000	12
75%	2822.750000	1986.750000	5.500000	4.775000	7.775000	10.350000	18.850000	17
max	2851.000000	2015.000000	70.400000	63.500000	79.000000	61.300000	142.100000	29

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

```
<seaborn.axisgrid.PairGrid at 0x1af8d7d1a00>
```

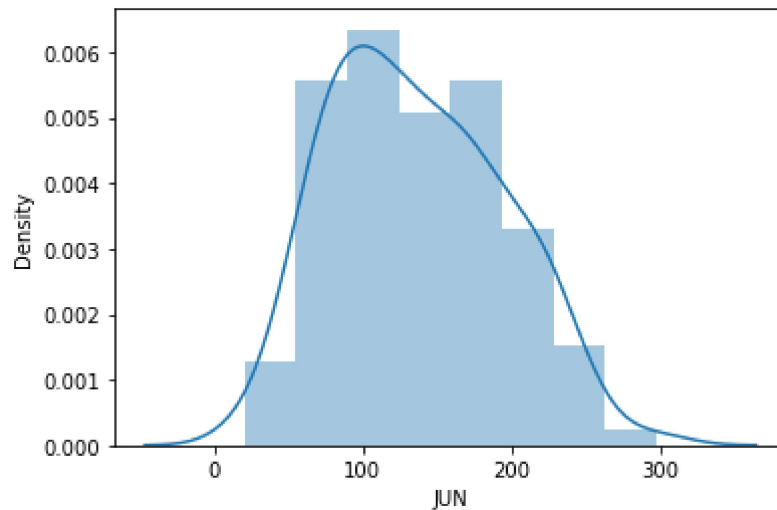


```
In [16]: sns.distplot(df['JUN'])
```

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).

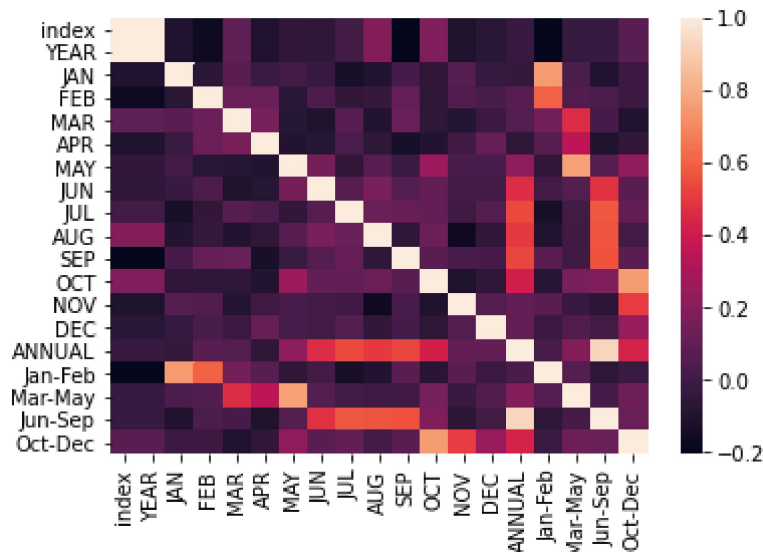
```
warnings.warn(msg, FutureWarning)
```

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



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

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



```
In [ ]:
```

