

```
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\NORTH INTERIOR KARNATAKA.csv')
df
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	3658	NORTH INTERIOR KARNATAKA	1902	0.0	0.0	0.3	22.5	34.4	111.3	83.2	78.1	146.7	118.8
1	3659	NORTH INTERIOR KARNATAKA	1903	3.5	0.0	0.1	6.9	53.4	102.8	209.4	146.4	189.3	166.4
2	3660	NORTH INTERIOR KARNATAKA	1904	0.2	0.3	8.5	11.0	46.3	120.6	91.6	48.5	165.1	86.5
3	3661	NORTH INTERIOR KARNATAKA	1905	0.0	6.0	2.6	16.0	51.2	99.6	60.1	139.2	42.2	85.0
4	3662	NORTH INTERIOR KARNATAKA	1906	21.3	0.0	0.2	2.6	30.0	142.0	120.3	182.1	116.0	86.2
...
109	3767	NORTH INTERIOR KARNATAKA	2011	0.5	7.2	7.2	41.2	46.8	101.3	150.8	152.0	69.0	73.4
110	3768	NORTH INTERIOR KARNATAKA	2012	28.5	6.2	0.4	35.4	19.5	60.0	114.5	105.5	79.2	85.2
111	3769	NORTH INTERIOR KARNATAKA	2013	1.2	6.1	3.0	25.4	47.4	99.4	160.7	73.9	201.0	101.0
112	3770	NORTH INTERIOR KARNATAKA	2014	0.0	6.1	29.2	26.4	93.0	50.4	136.8	205.2	90.2	80.3
113	3771	NORTH INTERIOR KARNATAKA	2015	2.4	0.0	27.5	50.8	45.3	89.6	38.5	78.4	150.8	61.2

114 rows × 20 columns

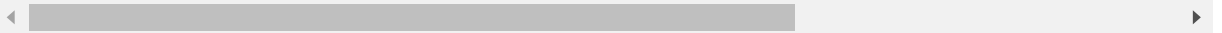


```
In [3]: df=df.dropna()
df
```

Out[3]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	3658	NORTH INTERIOR KARNATAKA	1902	0.0	0.0	0.3	22.5	34.4	111.3	83.2	78.1	146.7	118.8
1	3659	NORTH INTERIOR KARNATAKA	1903	3.5	0.0	0.1	6.9	53.4	102.8	209.4	146.4	189.3	166.4
2	3660	NORTH INTERIOR KARNATAKA	1904	0.2	0.3	8.5	11.0	46.3	120.6	91.6	48.5	165.1	86.5
3	3661	NORTH INTERIOR KARNATAKA	1905	0.0	6.0	2.6	16.0	51.2	99.6	60.1	139.2	42.2	85.0
4	3662	NORTH INTERIOR KARNATAKA	1906	21.3	0.0	0.2	2.6	30.0	142.0	120.3	182.1	116.0	86.2
...
109	3767	NORTH INTERIOR KARNATAKA	2011	0.5	7.2	7.2	41.2	46.8	101.3	150.8	152.0	69.0	73.4
110	3768	NORTH INTERIOR KARNATAKA	2012	28.5	6.2	0.4	35.4	19.5	60.0	114.5	105.5	79.2	85.2
111	3769	NORTH INTERIOR KARNATAKA	2013	1.2	6.1	3.0	25.4	47.4	99.4	160.7	73.9	201.0	101.0
112	3770	NORTH INTERIOR KARNATAKA	2014	0.0	6.1	29.2	26.4	93.0	50.4	136.8	205.2	90.2	80.3
113	3771	NORTH INTERIOR KARNATAKA	2015	2.4	0.0	27.5	50.8	45.3	89.6	38.5	78.4	150.8	61.2

114 rows × 20 columns



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

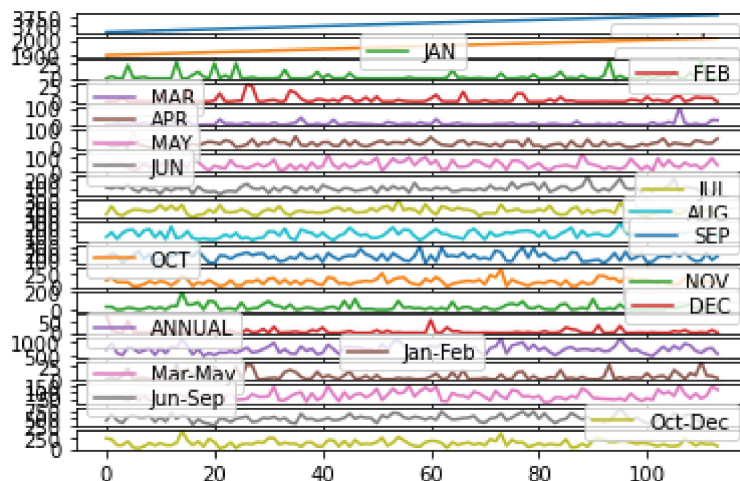
Out[4]: Index(['index', 'SUBDIVISION', 'YEAR', 'JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC', 'ANNUAL', 'Jan-Feb', 'Mar-May', 'Jun-Sep', 'Oct-Dec'], dtype='object')

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
```

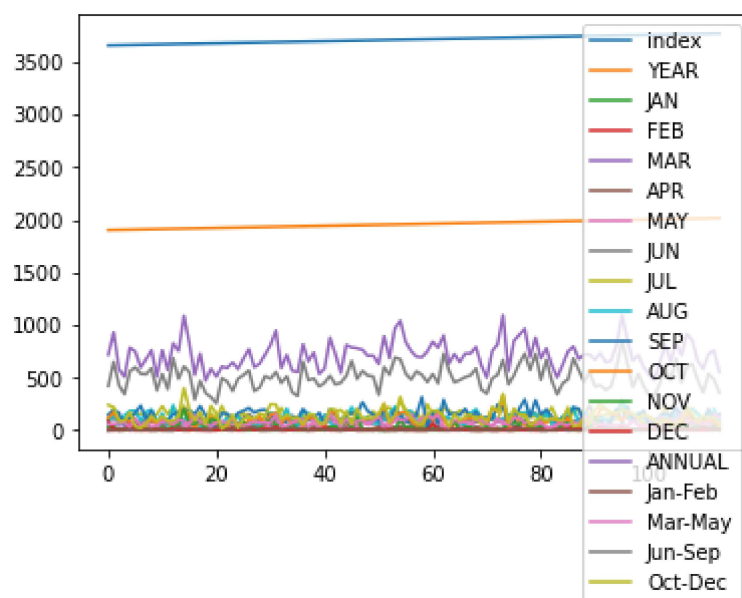
In [6]: df.plot.line(subplots=True)

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



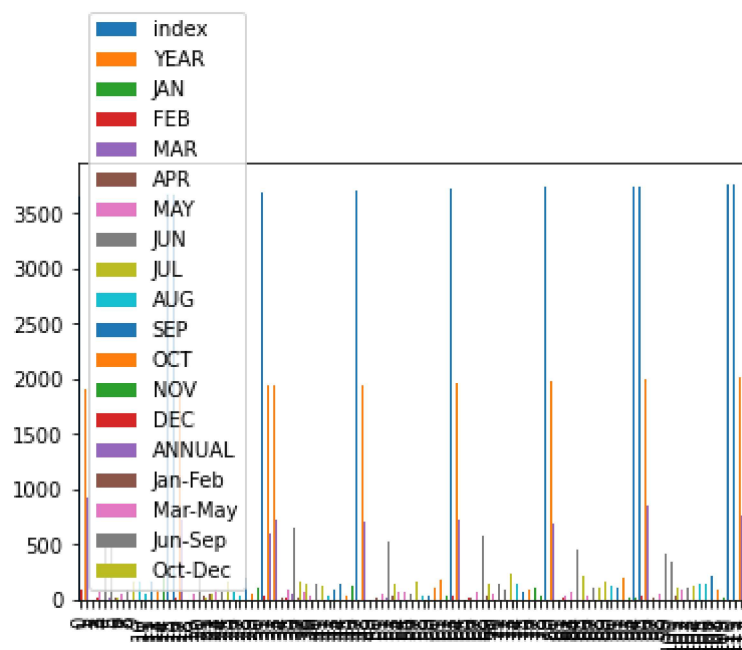
```
In [7]: df.plot.line()
```

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



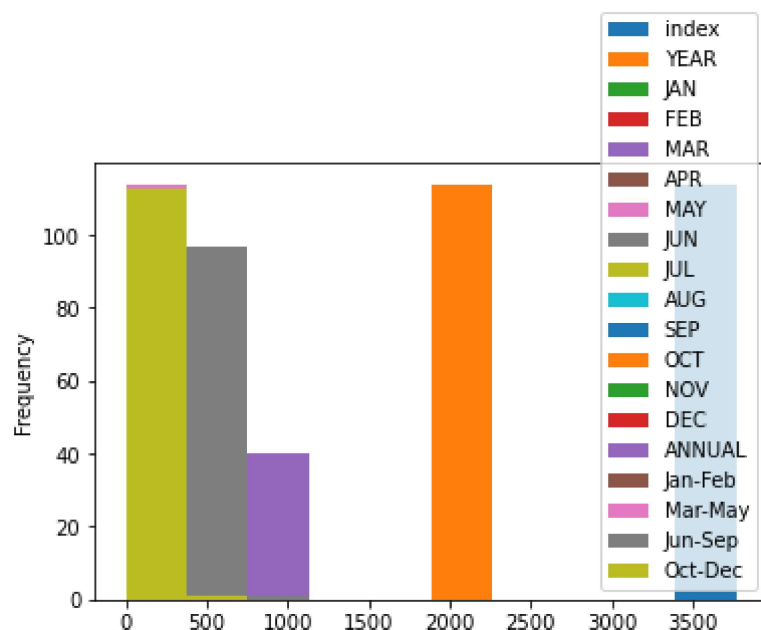
```
In [8]: df.plot.bar()
```

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



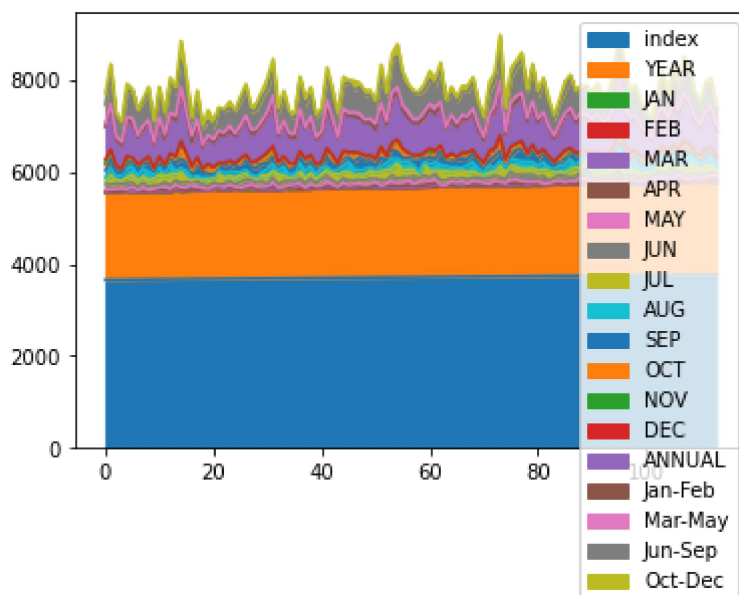
```
In [9]: df.plot.hist()
```

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



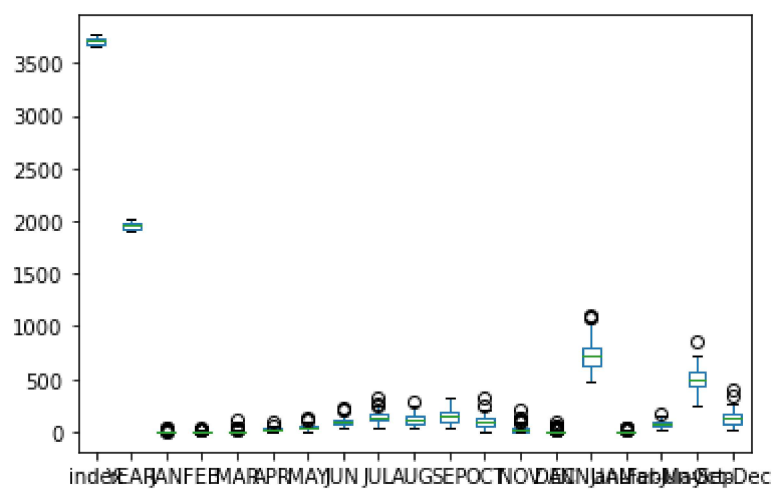
```
In [10]: df.plot.area()
```

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



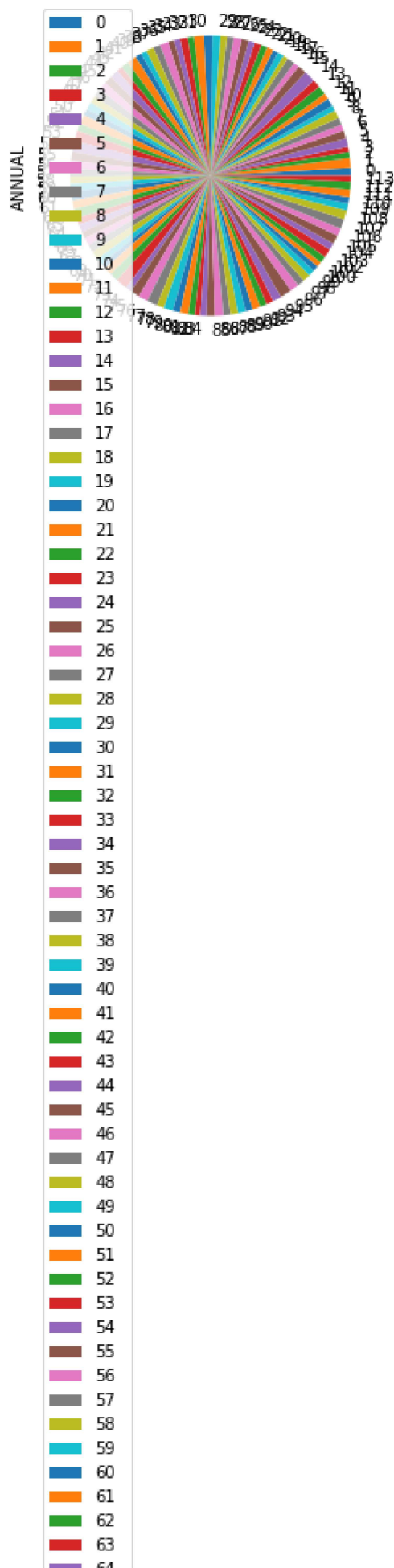
```
In [11]: df.plot.box()
```

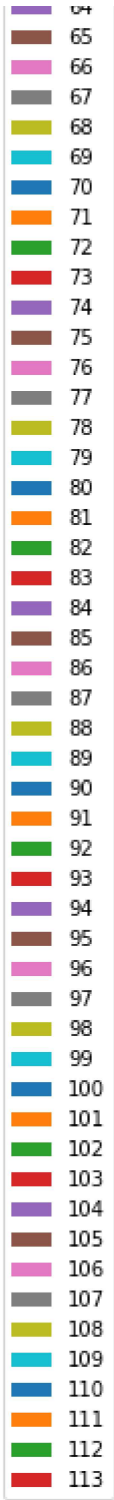
```
Out[11]: <AxesSubplot:>
```



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

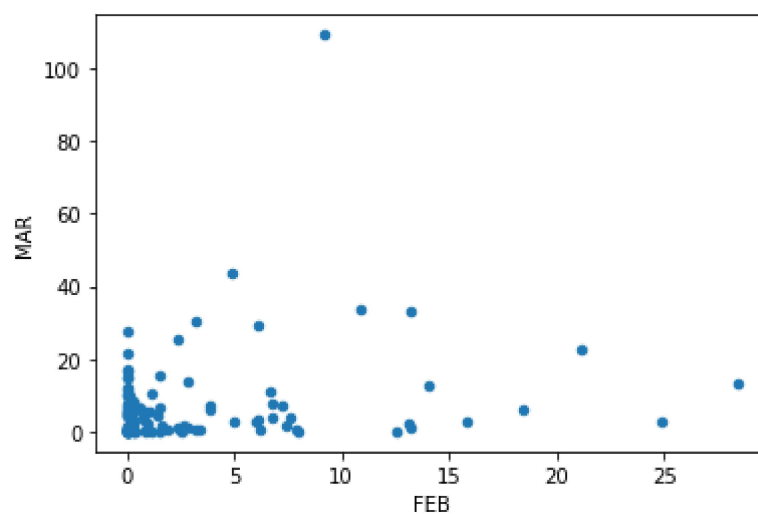
```
Out[12]: <AxesSubplot:ylabel='ANNUAL'>
```



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

```
Out[13]: <AxesSubplot:xlabel='FEB', ylabel='MAR'>
```



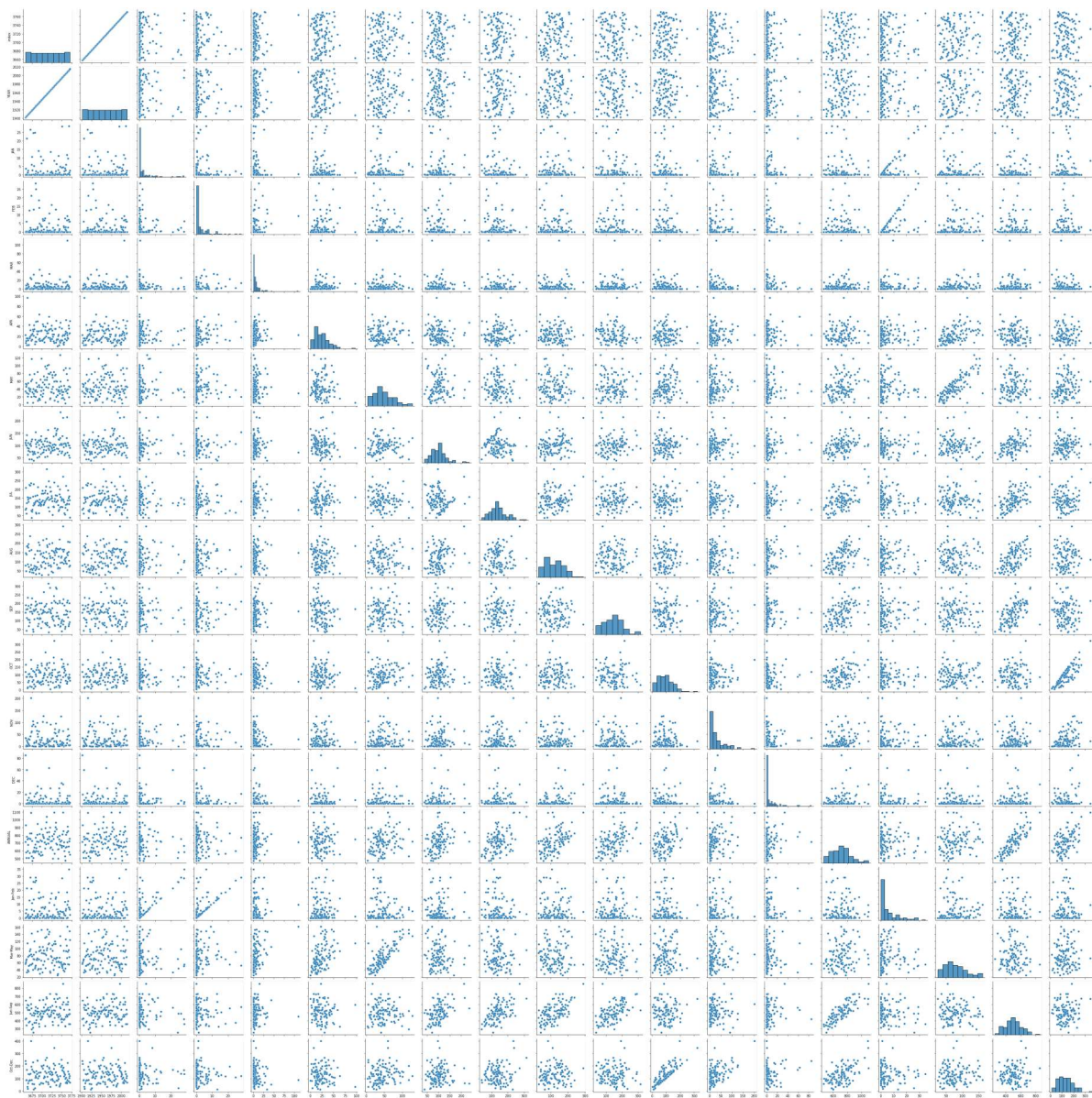
```
In [14]: df.describe()
```

```
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	3714.500000	1958.500000	3.008772	3.035088	7.123684	23.924561	46.873684	10
std	33.052988	33.052988	6.224851	5.360574	12.727298	15.236731	26.636674	3
min	3658.000000	1902.000000	0.000000	0.000000	0.000000	0.200000	3.500000	3
25%	3686.250000	1930.250000	0.000000	0.000000	0.500000	12.150000	29.050000	7
50%	3714.500000	1958.500000	0.200000	0.300000	3.450000	22.300000	40.600000	9
75%	3742.750000	1986.750000	2.400000	3.350000	7.600000	32.100000	62.500000	11
max	3771.000000	2015.000000	28.500000	28.400000	109.200000	96.900000	127.300000	23

```
In [15]: sns.pairplot(df)
```

```
Out[15]: <seaborn.axisgrid.PairGrid at 0x1e4364ffb20>
```

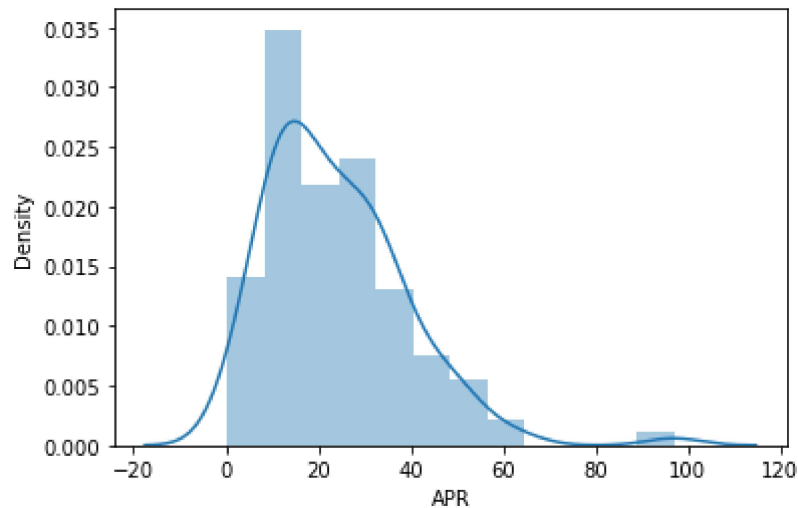


```
In [16]: sns.distplot(df['APR'])
```

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).

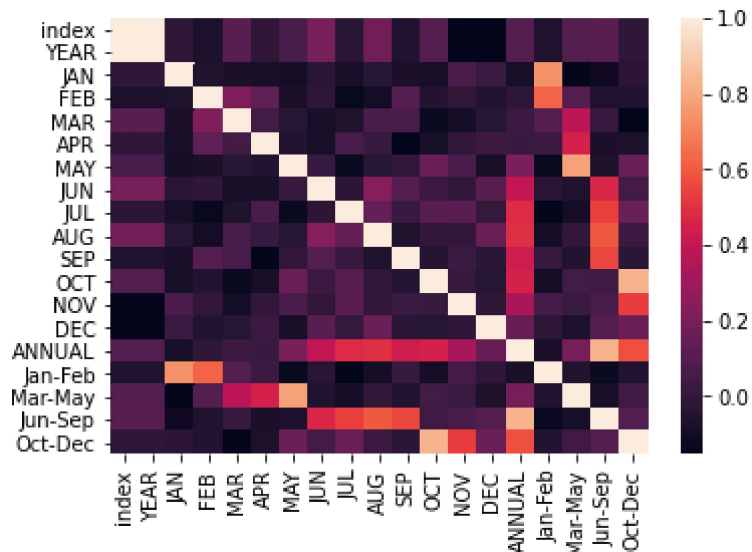
```
warnings.warn(msg, FutureWarning)
```

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



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

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



```
In [ ]:
```

