

Import Libraries

In [1]:

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

In [2]:

```
d=pd.read_csv(r"C:\Users\user\Downloads\FP2_RainFall\rain.csv")[3659:3772]
d
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
324	324	NAGA MANI MIZO TRIPURA	1903	6.5	40.5	139.8	45.5	159.9	458.6	300.2	470.6	366.1
325	325	NAGA MANI MIZO TRIPURA	1904	2.3	46.9	47.5	290.3	230.5	455.3	423.5	423.6	375.8
326	326	NAGA MANI MIZO TRIPURA	1905	9.1	35.3	306.5	161.7	193.6	339.7	450.1	429.9	320.1
327	327	NAGA MANI MIZO TRIPURA	1906	7.0	71.5	72.5	99.0	302.7	417.4	475.2	439.2	439.1
328	328	NAGA MANI MIZO TRIPURA	1907	39.1	30.4	92.8	189.1	156.0	414.1	400.9	291.4	336.1
...
432	432	NAGA MANI MIZO TRIPURA	2011	12.6	3.6	51.4	81.1	334.9	374.2	313.3	367.6	258.3
433	433	NAGA MANI MIZO TRIPURA	2012	24.5	10.2	20.3	243.5	163.5	396.2	280.1	342.7	248.7
434	434	NAGA MANI MIZO TRIPURA	2013	0.2	5.7	19.7	60.3	348.9	206.6	255.9	291.3	241.4
435	435	NAGA MANI MIZO TRIPURA	2014	1.2	21.0	25.4	49.6	192.5	268.3	295.7	372.3	300.9
436	436	NAGA MANI MIZO TRIPURA	2015	14.4	14.2	21.6	253.5	198.3	283.9	413.6	334.2	255.9

113 rows × 20 columns



Data Cleaning and preprocessing

In [3]:

```
d.dropna()
```

Out[3]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
324	324	NAGA MANI MIZO TRIPURA	1903	6.5	40.5	139.8	45.5	159.9	458.6	300.2	470.6	366.1
325	325	NAGA MANI MIZO TRIPURA	1904	2.3	46.9	47.5	290.3	230.5	455.3	423.5	423.6	375.8
326	326	NAGA MANI MIZO TRIPURA	1905	9.1	35.3	306.5	161.7	193.6	339.7	450.1	429.9	320.1
327	327	NAGA MANI MIZO TRIPURA	1906	7.0	71.5	72.5	99.0	302.7	417.4	475.2	439.2	439.1
328	328	NAGA MANI MIZO TRIPURA	1907	39.1	30.4	92.8	189.1	156.0	414.1	400.9	291.4	336.1
...
432	432	NAGA MANI MIZO TRIPURA	2011	12.6	3.6	51.4	81.1	334.9	374.2	313.3	367.6	258.3
433	433	NAGA MANI MIZO TRIPURA	2012	24.5	10.2	20.3	243.5	163.5	396.2	280.1	342.7	248.7
434	434	NAGA MANI MIZO TRIPURA	2013	0.2	5.7	19.7	60.3	348.9	206.6	255.9	291.3	241.4
435	435	NAGA MANI MIZO TRIPURA	2014	1.2	21.0	25.4	49.6	192.5	268.3	295.7	372.3	300.9
436	436	NAGA MANI MIZO TRIPURA	2015	14.4	14.2	21.6	253.5	198.3	283.9	413.6	334.2	255.9

113 rows × 20 columns



In [4]:

```
d.columns
```

Out[4]:

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

In [5]:

```
d.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 113 entries, 324 to 436
Data columns (total 20 columns):
#   Column          Non-Null Count  Dtype
---  -
0   index           113 non-null    int64
1   SUBDIVISION     113 non-null    object
2   YEAR            113 non-null    int64
3   JAN             113 non-null    float64
4   FEB             113 non-null    float64
5   MAR             113 non-null    float64
6   APR             113 non-null    float64
7   MAY             113 non-null    float64
8   JUN             113 non-null    float64
9   JUL             113 non-null    float64
10  AUG             113 non-null    float64
11  SEP             113 non-null    float64
12  OCT             113 non-null    float64
13  NOV             113 non-null    float64
14  DEC             113 non-null    float64
15  ANNUAL          113 non-null    float64
16  Jan-Feb        113 non-null    float64
17  Mar-May        113 non-null    float64
18  Jun-Sep        113 non-null    float64
19  Oct-Dec        113 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 17.8+ KB
```

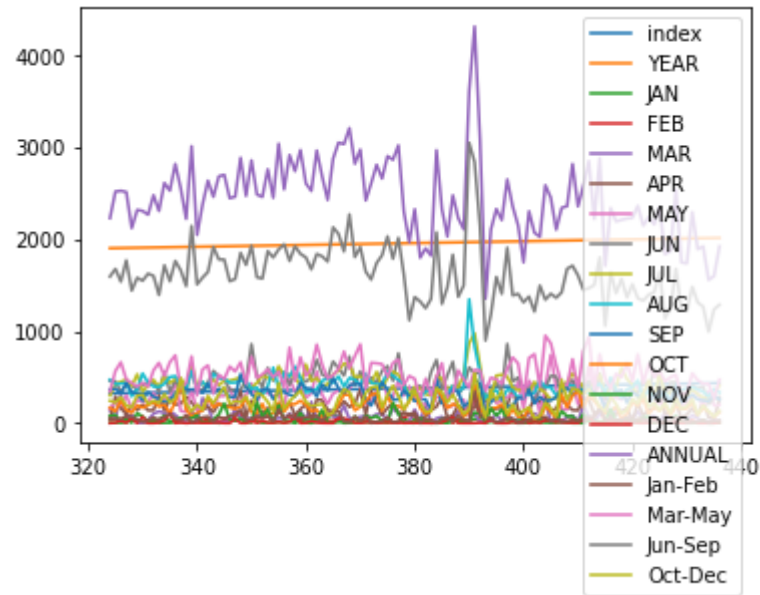
Line Chart

In [6]:

```
d.plot.line()
```

Out[6]:

<AxesSubplot:>



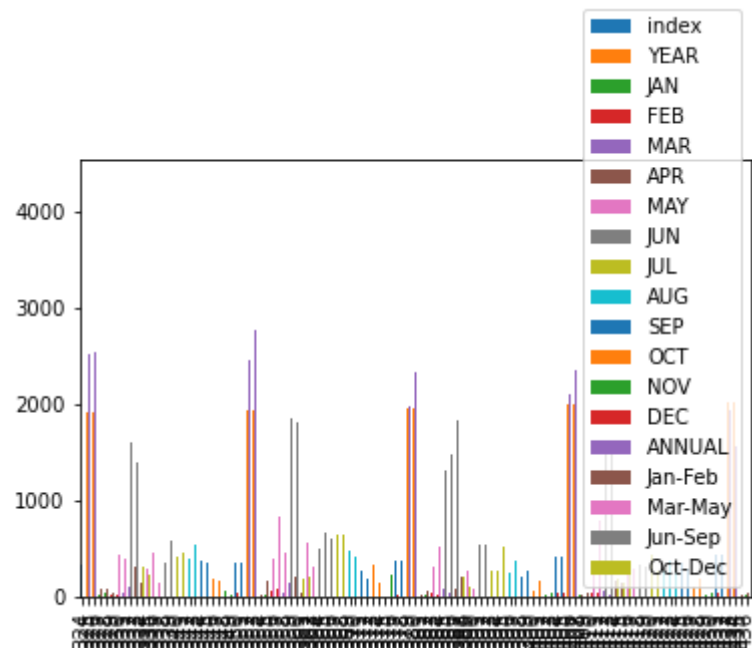
Bar Chart

In [7]:

```
d.plot.bar()
```

Out[7]:

<AxesSubplot:>



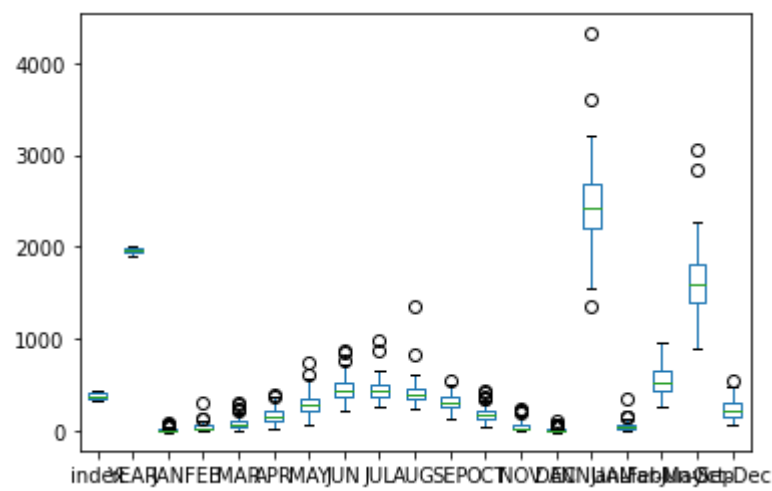
Box Chart

In [8]:

```
d.plot.box()
```

Out[8]:

<AxesSubplot:>



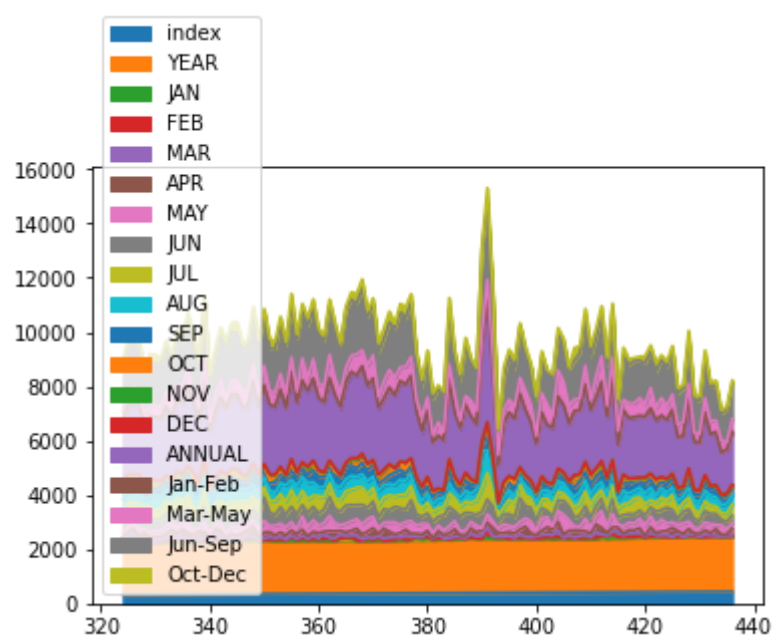
Area Chart

In [9]:

```
d.plot.area()
```

Out[9]:

<AxesSubplot:>



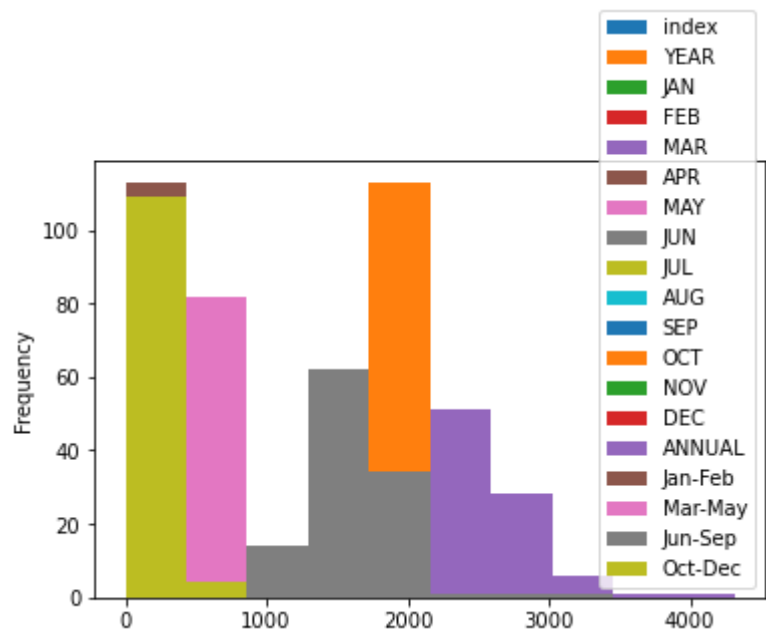
Histogram

In [10]:

```
d.plot.hist()
```

Out[10]:

<AxesSubplot:ylabel='Frequency'>



Pie Chart

In [11]:

```
d.plot.pie(y='ANNUAL')
```

Out[11]:

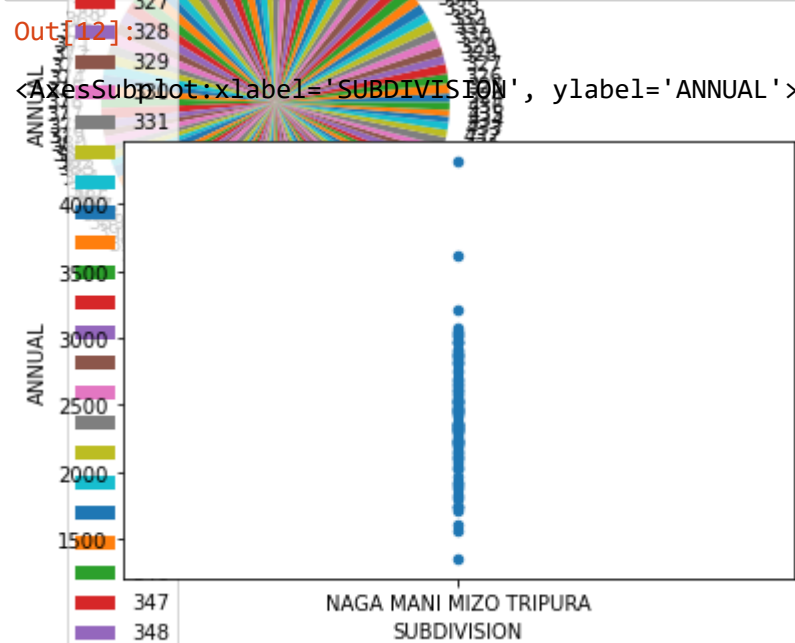
<AxesSubplot:ylabel='ANNUAL'>

Scatter Chart

In []

```
d.plot.scatter(y='ANNUAL', x='SUBDIVISION')
```

Out [12]:



In []

```
d.describe()
```

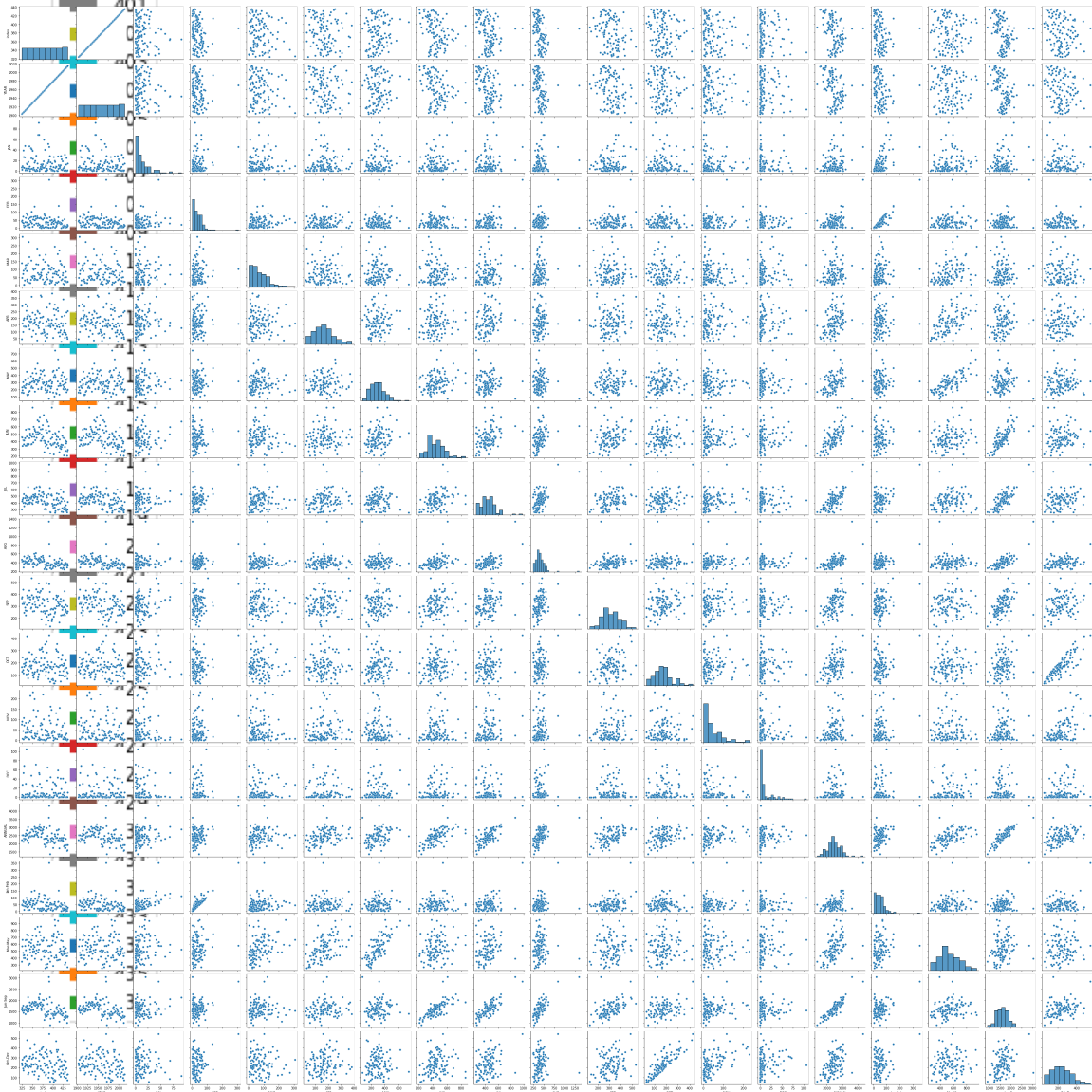
Out [13]:

	index	YEAR	JAN	FEB	MAR	APR	MAY
count	113.00000	113.00000	113.00000	113.00000	113.00000	113.00000	113.00000
mean	380.00000	1959.00000	14.127434	37.136283	77.984071	169.294690	292.982301
std	32.76431	32.76431	16.751137	37.915658	60.901380	78.257574	112.332719
min	324.00000	1903.00000	0.000000	0.000000	3.100000	26.300000	73.500000
25%	352.00000	1931.00000	3.000000	11.900000	32.800000	112.000000	211.000000
50%	380.00000	1959.00000	7.900000	30.100000	64.100000	161.500000	280.600000
75%	408.00000	1987.00000	18.600000	53.600000	105.500000	213.600000	352.400000
max	436.00000	2015.00000	91.400000	306.300000	306.500000	383.800000	743.000000

EDA AND VISUALIZATION

```
In [ ]: sns.pairplot(d)
```

```
Out[14]: <seaborn.axisgrid.PairGrid at 0x21dc1f3bbe0>
```

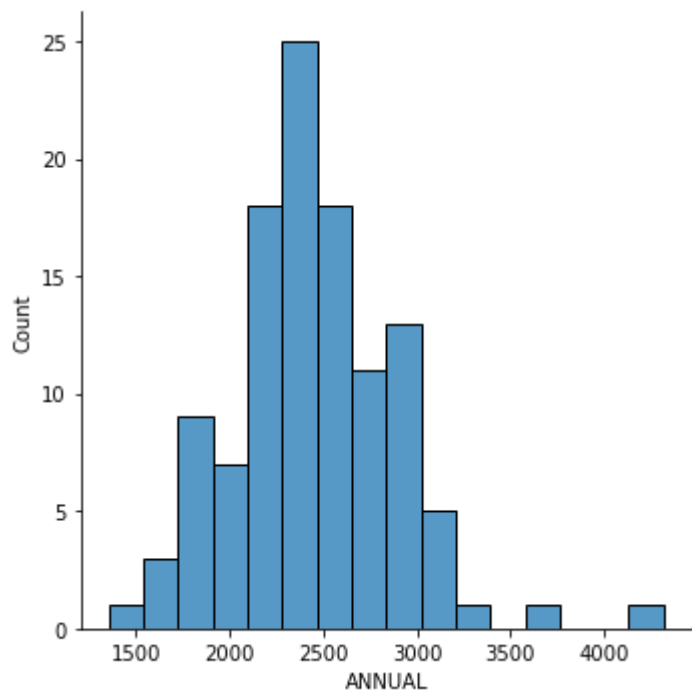


In [15]:

```
sns.displot(d['ANNUAL'])
```

Out[15]:

```
<seaborn.axisgrid.FacetGrid at 0x21dd0ecff70>
```

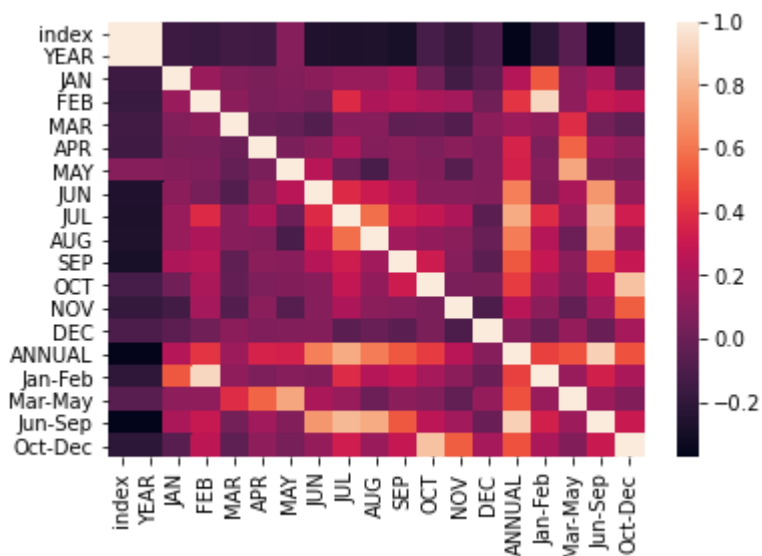


In [16]:

```
sns.heatmap(d.corr())
```

Out[16]:

```
<AxesSubplot:>
```



In []:

