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")[1589:1702]
d
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
1589	1589	HIMACHAL PRADESH	1903	76.5	21.4	213.7	25.4	54.7	32.2	157.7	256.5	107.9
1590	1590	HIMACHAL PRADESH	1904	79.3	22.4	131.7	48.0	90.3	33.1	241.1	184.3	56.4
1591	1591	HIMACHAL PRADESH	1905	81.3	76.8	160.2	39.3	50.4	43.6	191.1	132.8	119.1
1592	1592	HIMACHAL PRADESH	1906	44.1	143.9	89.5	5.3	29.9	152.6	168.7	433.7	230.9
1593	1593	HIMACHAL PRADESH	1907	124.2	145.1	144.9	73.0	34.2	23.7	95.7	200.6	18.9
...	...	...	...	...	...	...	...	...	...	...	...	...
1697	1697	HIMACHAL PRADESH	2011	43.9	97.4	49.7	62.4	45.1	118.3	177.7	380.2	120.3
1698	1698	HIMACHAL PRADESH	2012	92.3	51.3	28.4	55.9	9.4	31.1	241.5	280.6	133.1
1699	1699	HIMACHAL PRADESH	2013	79.9	182.6	76.6	28.9	32.6	233.6	208.8	240.0	65.8
1700	1700	HIMACHAL PRADESH	2014	69.6	124.9	125.2	60.6	68.9	51.7	203.6	146.7	84.6
1701	1701	HIMACHAL PRADESH	2015	67.2	156.6	192.5	84.9	45.0	85.8	249.9	195.9	75.5

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
1589	1589	HIMACHAL PRADESH	1903	76.5	21.4	213.7	25.4	54.7	32.2	157.7	256.5	107.9
1590	1590	HIMACHAL PRADESH	1904	79.3	22.4	131.7	48.0	90.3	33.1	241.1	184.3	56.4
1591	1591	HIMACHAL PRADESH	1905	81.3	76.8	160.2	39.3	50.4	43.6	191.1	132.8	119.1
1592	1592	HIMACHAL PRADESH	1906	44.1	143.9	89.5	5.3	29.9	152.6	168.7	433.7	230.9
1593	1593	HIMACHAL PRADESH	1907	124.2	145.1	144.9	73.0	34.2	23.7	95.7	200.6	18.9
...	...	...	...	...	...	...	...	...	...	...	...	...
1697	1697	HIMACHAL PRADESH	2011	43.9	97.4	49.7	62.4	45.1	118.3	177.7	380.2	120.3
1698	1698	HIMACHAL PRADESH	2012	92.3	51.3	28.4	55.9	9.4	31.1	241.5	280.6	133.1
1699	1699	HIMACHAL PRADESH	2013	79.9	182.6	76.6	28.9	32.6	233.6	208.8	240.0	65.8
1700	1700	HIMACHAL PRADESH	2014	69.6	124.9	125.2	60.6	68.9	51.7	203.6	146.7	84.6
1701	1701	HIMACHAL PRADESH	2015	67.2	156.6	192.5	84.9	45.0	85.8	249.9	195.9	75.5

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, 1589 to 1701
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
```

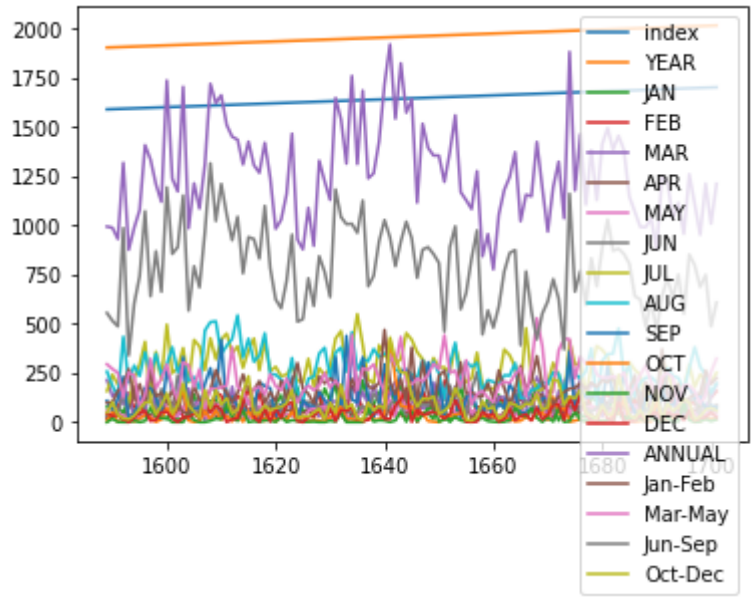
Line Chart

In [6]:

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

Out[6]:

<AxesSubplot:>



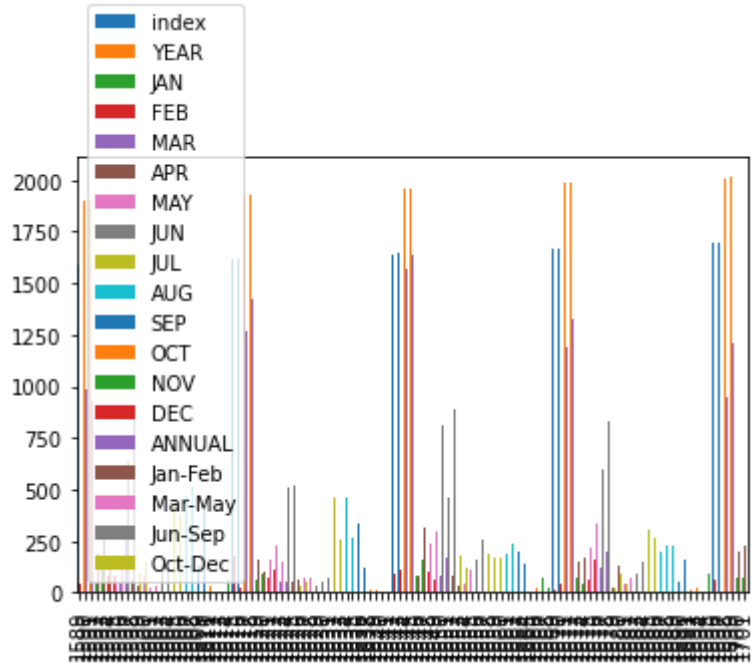
Bar Chart

In [7]:

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

Out[7]:

<AxesSubplot:>



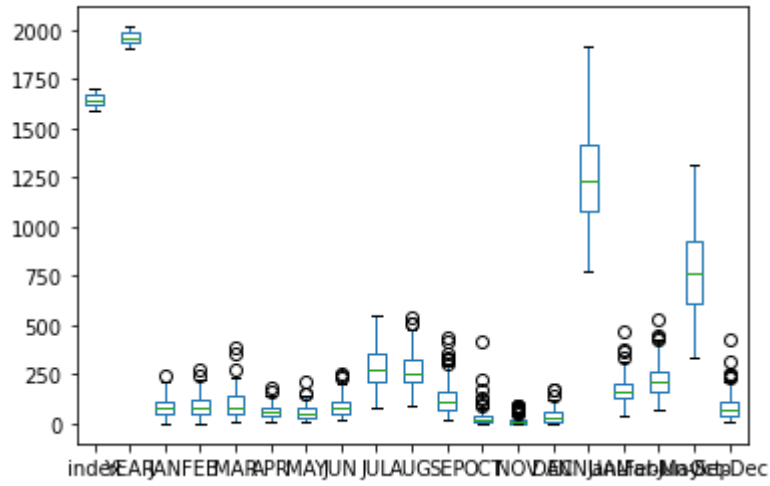
Box Chart

In [8]:

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

Out[8]:

<AxesSubplot:>



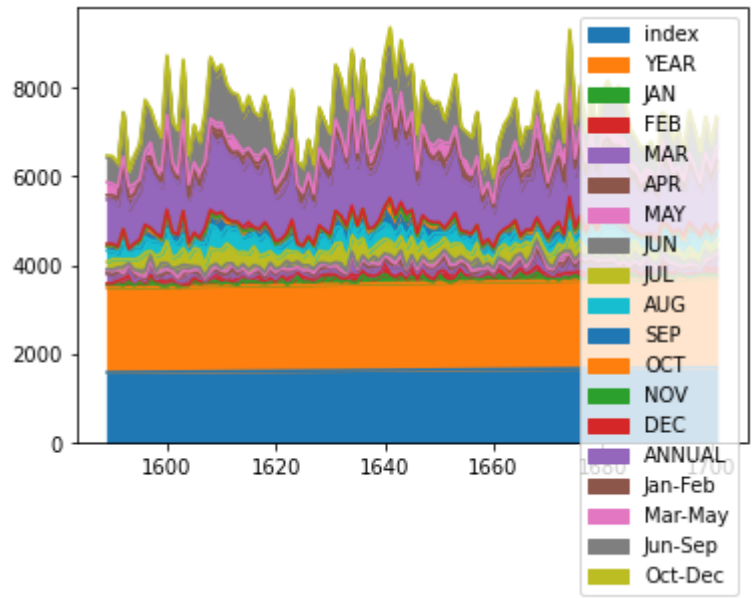
Area Chart

In [9]:

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

Out[9]:

<AxesSubplot:>



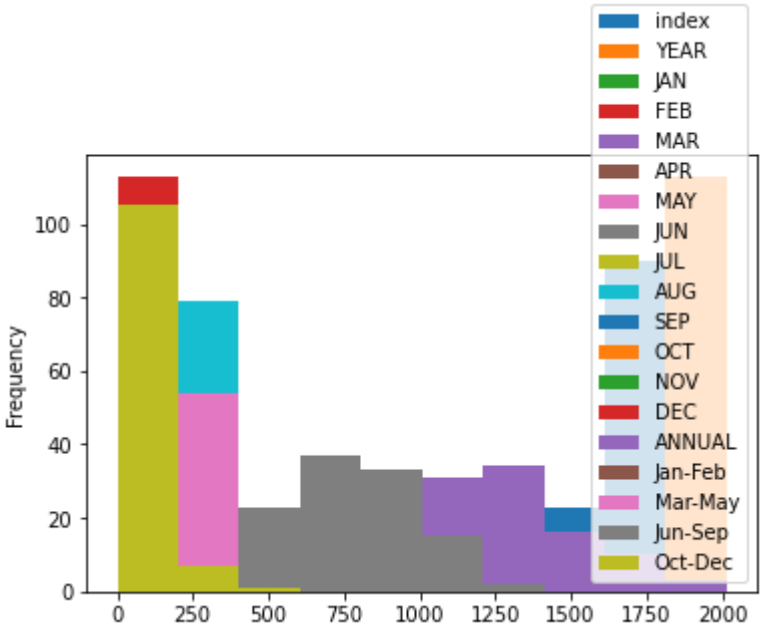
Histogram

In [10]:

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

Out[10]:

```
<AxesSubplot:ylabel='Frequency'>
```



Pie Chart

In [11]:

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

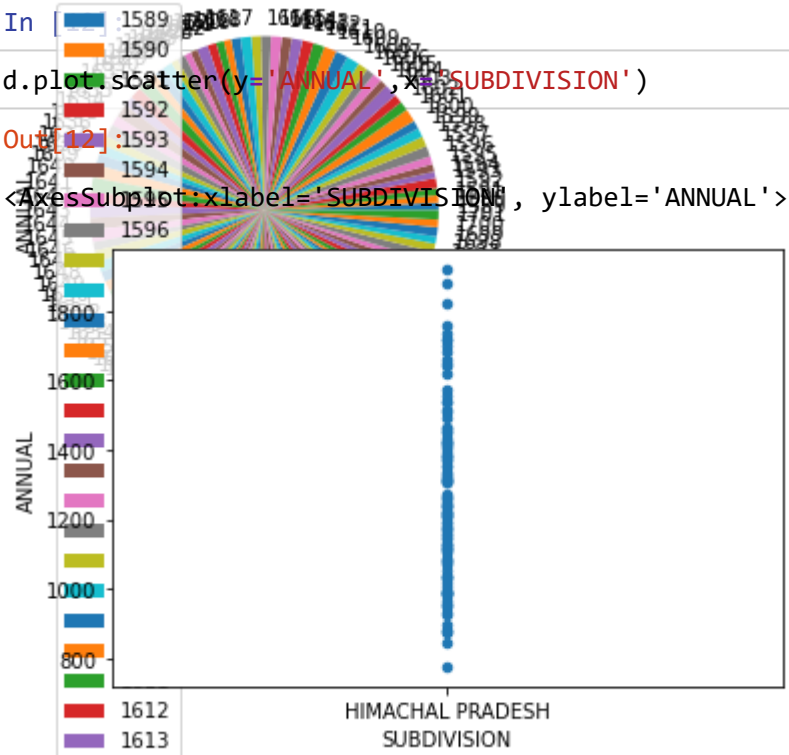
Out[11]:

```
<AxesSubplot:ylabel='ANNUAL'>
```

Scatter Chart

In [ ]:
```
d.plot.scatter(y='ANNUAL',x='SUBDIVISION')
```

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



In [ ]:
```
d.describe()
```

Out[13]:

	index	YEAR	JAN	FEB	MAR	APR	MAY
count	113.00000	113.00000	113.000000	113.000000	113.000000	113.000000	113.000000
mean	1645.00000	1959.00000	84.402655	90.720354	101.348673	62.689381	57.850442
std	32.76431	32.76431	51.615511	51.776172	67.053643	35.948960	34.142776
min	1589.00000	1903.00000	0.300000	0.700000	5.900000	4.500000	8.800000
25%	1617.00000	1931.00000	45.200000	51.300000	53.700000	35.500000	34.200000
50%	1645.00000	1959.00000	78.000000	82.800000	83.600000	55.900000	53.500000
75%	1673.00000	1987.00000	113.400000	124.700000	137.300000	84.900000	78.100000
max	1701.00000	2015.00000	246.300000	271.800000	382.000000	181.700000	214.200000

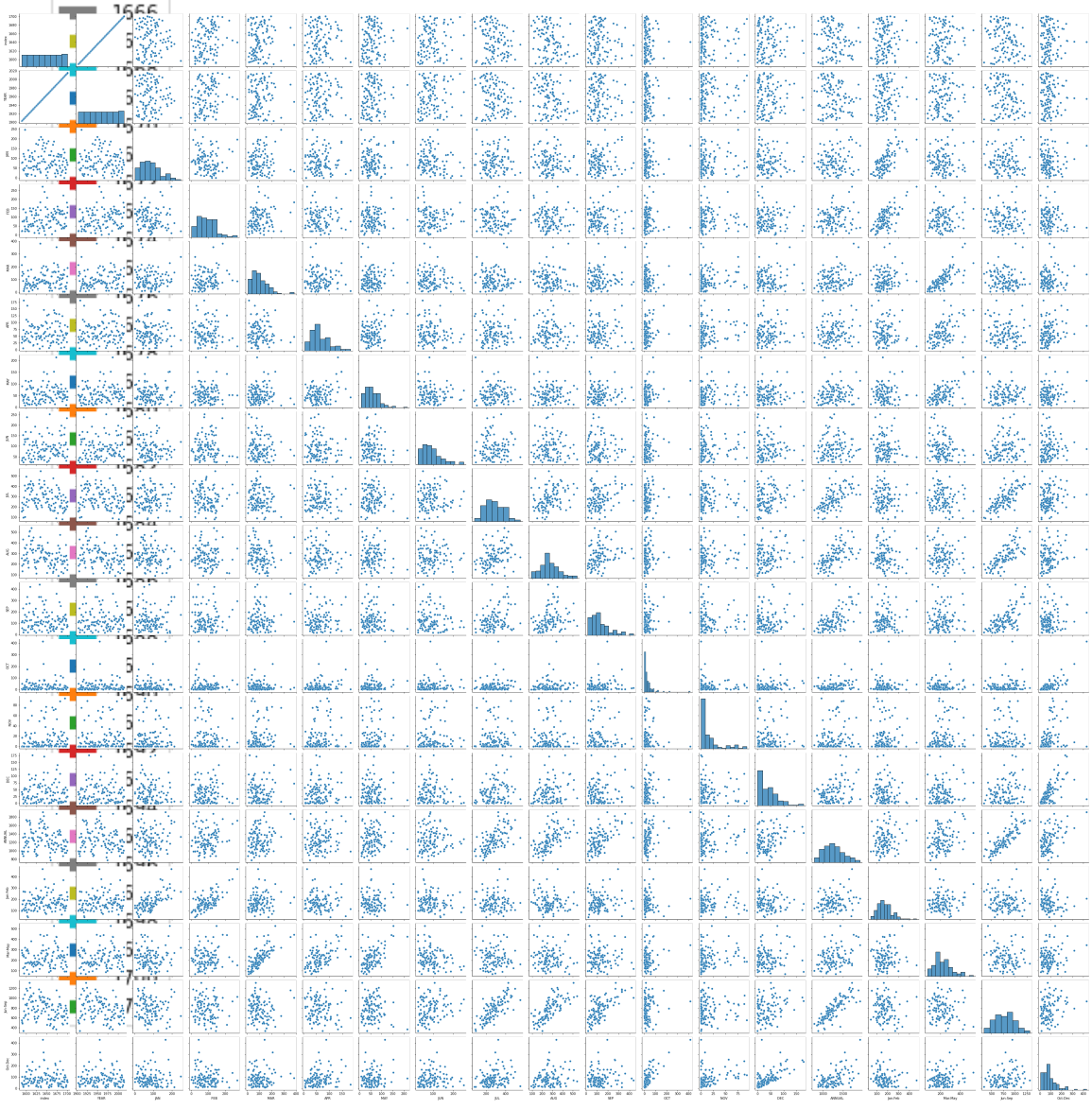
EDA AND VISUALIZATION

In [14]:

```
sns.pairplot(d)
```

Out[14]:

```
<seaborn.axisgrid.PairGrid at 0x200f8bd7910>
```

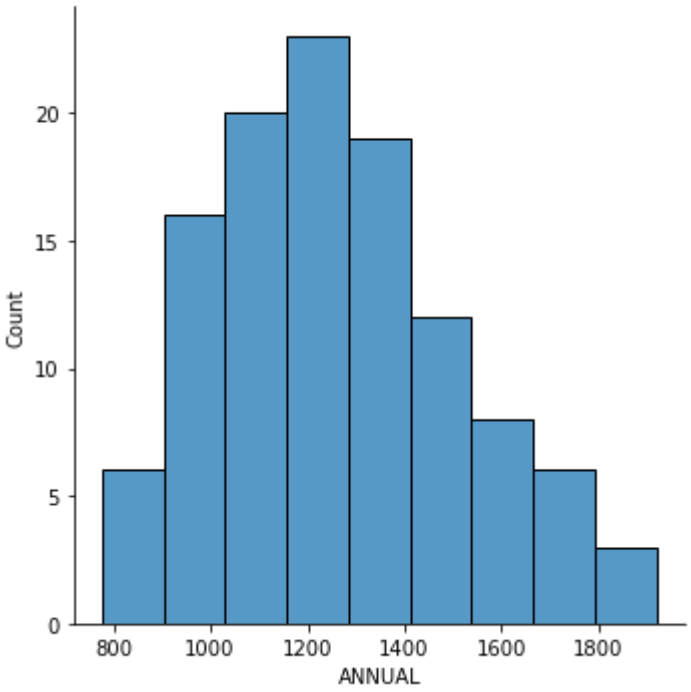


In [15]:

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

Out[15]:

```
<seaborn.axisgrid.FacetGrid at 0x20087cd3310>
```

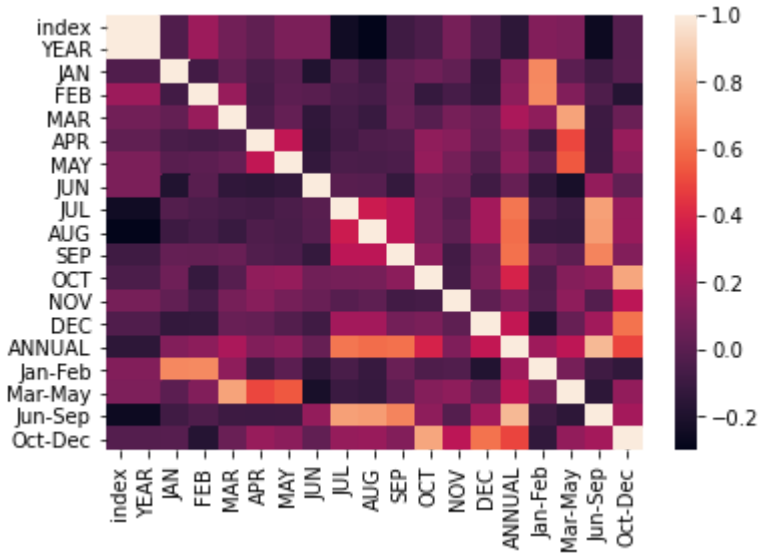


In [16]:

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

Out[16]:

```
<AxesSubplot:>
```



In [ ]: