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")[2509:2622]
d
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
2509	2509	KONKAN & GOA	1903	0.0	0.0	0.1	0.0	201.1	470.5	1298.6	673.9	285.1
2510	2510	KONKAN & GOA	1904	0.0	0.1	6.6	6.3	4.6	975.8	771.7	321.3	217.0
2511	2511	KONKAN & GOA	1905	0.1	0.1	0.0	0.4	8.6	293.7	770.6	305.5	208.3
2512	2512	KONKAN & GOA	1906	5.0	0.9	0.0	0.0	2.9	547.4	1090.9	506.7	222.5
2513	2513	KONKAN & GOA	1907	1.7	1.0	0.0	20.1	0.0	583.1	1170.9	811.5	164.0
...	...	...	...	...	...	...	...	...	...	...	...	..
2617	2617	KONKAN & GOA	2011	0.0	0.0	0.0	3.4	1.1	857.0	1384.1	987.9	468.3
2618	2618	KONKAN & GOA	2012	0.0	0.0	0.0	0.6	1.1	633.0	928.5	762.5	515.3
2619	2619	KONKAN & GOA	2013	1.8	5.4	0.1	0.1	18.5	1028.3	1478.5	497.6	340.7
2620	2620	KONKAN & GOA	2014	1.3	5.3	1.8	0.7	21.3	238.2	1293.2	658.0	419.5
2621	2621	KONKAN & GOA	2015	2.7	0.0	36.8	3.6	11.3	764.0	526.5	377.3	240.9

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
2509	2509	KONKAN & GOA	1903	0.0	0.0	0.1	0.0	201.1	470.5	1298.6	673.9	285.1
2510	2510	KONKAN & GOA	1904	0.0	0.1	6.6	6.3	4.6	975.8	771.7	321.3	217.0
2511	2511	KONKAN & GOA	1905	0.1	0.1	0.0	0.4	8.6	293.7	770.6	305.5	208.3
2512	2512	KONKAN & GOA	1906	5.0	0.9	0.0	0.0	2.9	547.4	1090.9	506.7	222.5
2513	2513	KONKAN & GOA	1907	1.7	1.0	0.0	20.1	0.0	583.1	1170.9	811.5	164.0
...	...	...	...	...	...	...	...	...	...	...	...	..
2617	2617	KONKAN & GOA	2011	0.0	0.0	0.0	3.4	1.1	857.0	1384.1	987.9	468.3
2618	2618	KONKAN & GOA	2012	0.0	0.0	0.0	0.6	1.1	633.0	928.5	762.5	515.3
2619	2619	KONKAN & GOA	2013	1.8	5.4	0.1	0.1	18.5	1028.3	1478.5	497.6	340.7
2620	2620	KONKAN & GOA	2014	1.3	5.3	1.8	0.7	21.3	238.2	1293.2	658.0	419.5
2621	2621	KONKAN & GOA	2015	2.7	0.0	36.8	3.6	11.3	764.0	526.5	377.3	240.9

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, 2509 to 2621
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
```

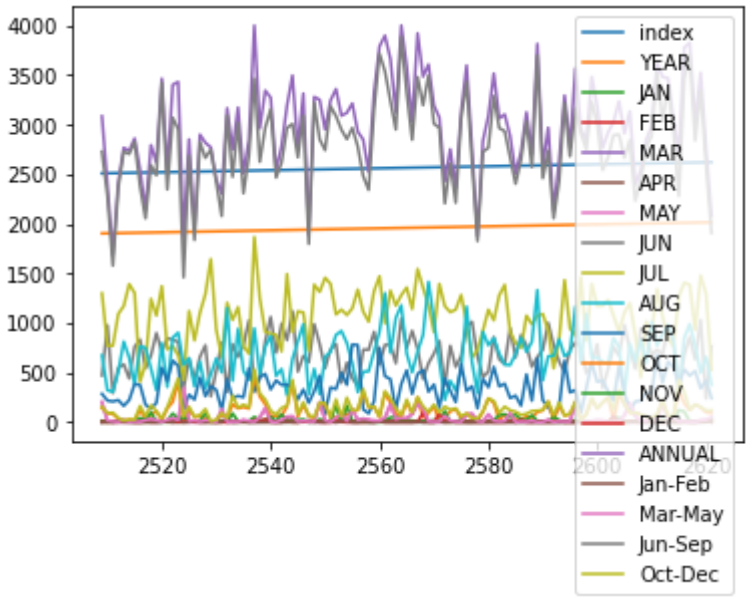
Line Chart

In [6]:

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

Out[6]:

<AxesSubplot:>



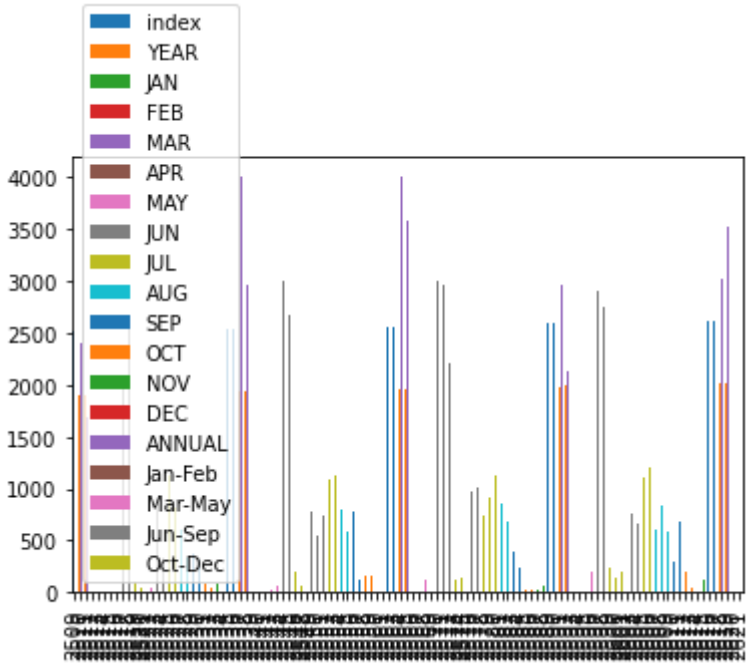
Bar Chart

In [7]:

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

Out[7]:

<AxesSubplot:>



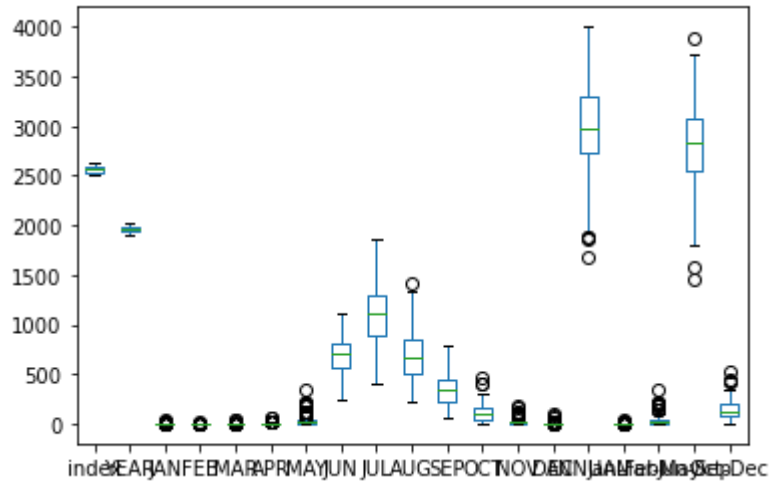
Box Chart

In [8]:

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

Out[8]:

<AxesSubplot:>



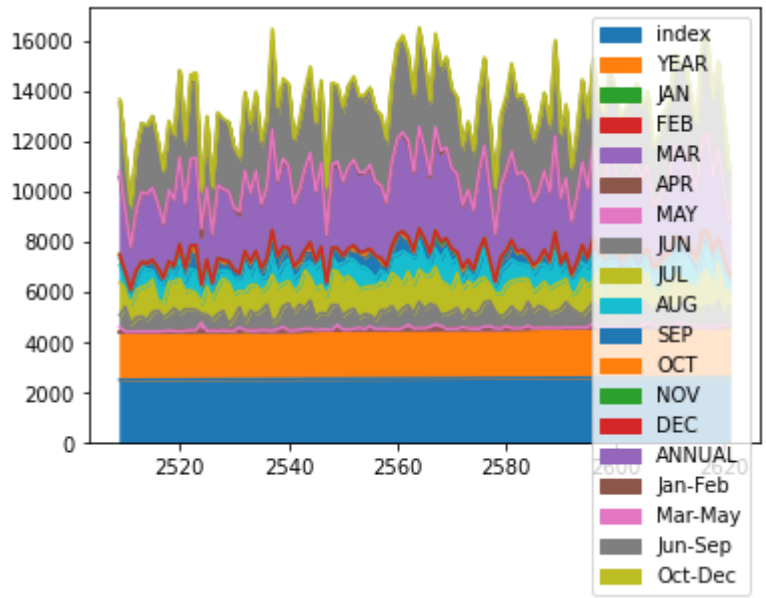
Area Chart

In [9]:

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

Out[9]:

<AxesSubplot:>



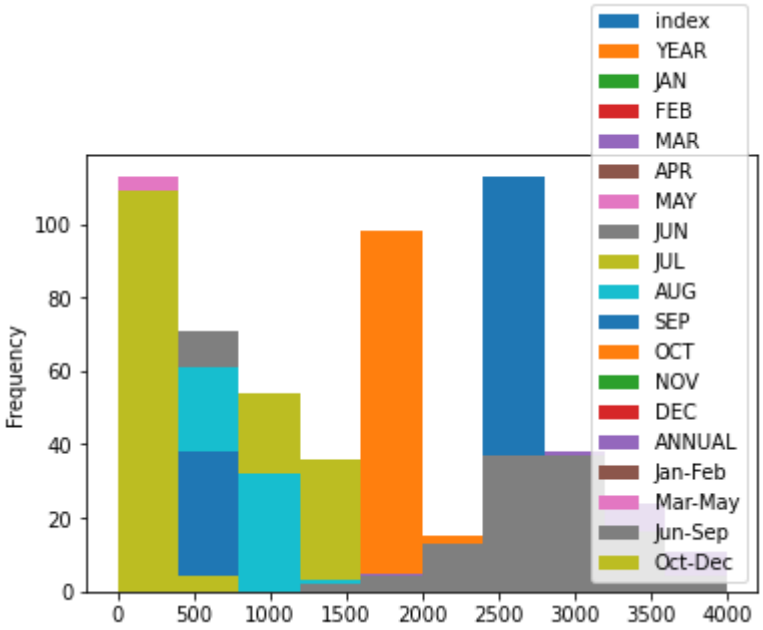
Histogram

In [10]:

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

Out[10]:

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



Pie Chart

In [11]:

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

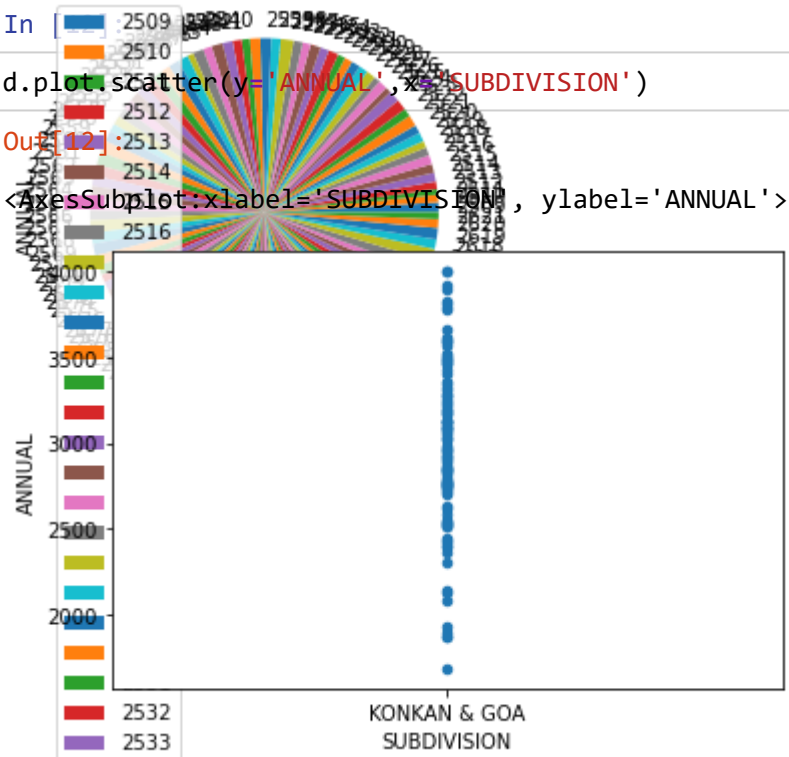
Out[11]:

```
<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	2565.00000	1959.00000	1.232743	0.555752	1.395575	4.022124	33.865487
std	32.76431	32.76431	3.896093	2.065142	4.788834	8.679355	58.779635
min	2509.00000	1903.00000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	2537.00000	1931.00000	0.000000	0.000000	0.000000	0.300000	2.900000
50%	2565.00000	1959.00000	0.000000	0.000000	0.100000	1.300000	9.700000
75%	2593.00000	1987.00000	0.500000	0.100000	0.400000	4.100000	30.600000
max	2621.00000	2015.00000	31.800000	18.400000	36.800000	67.300000	345.400000

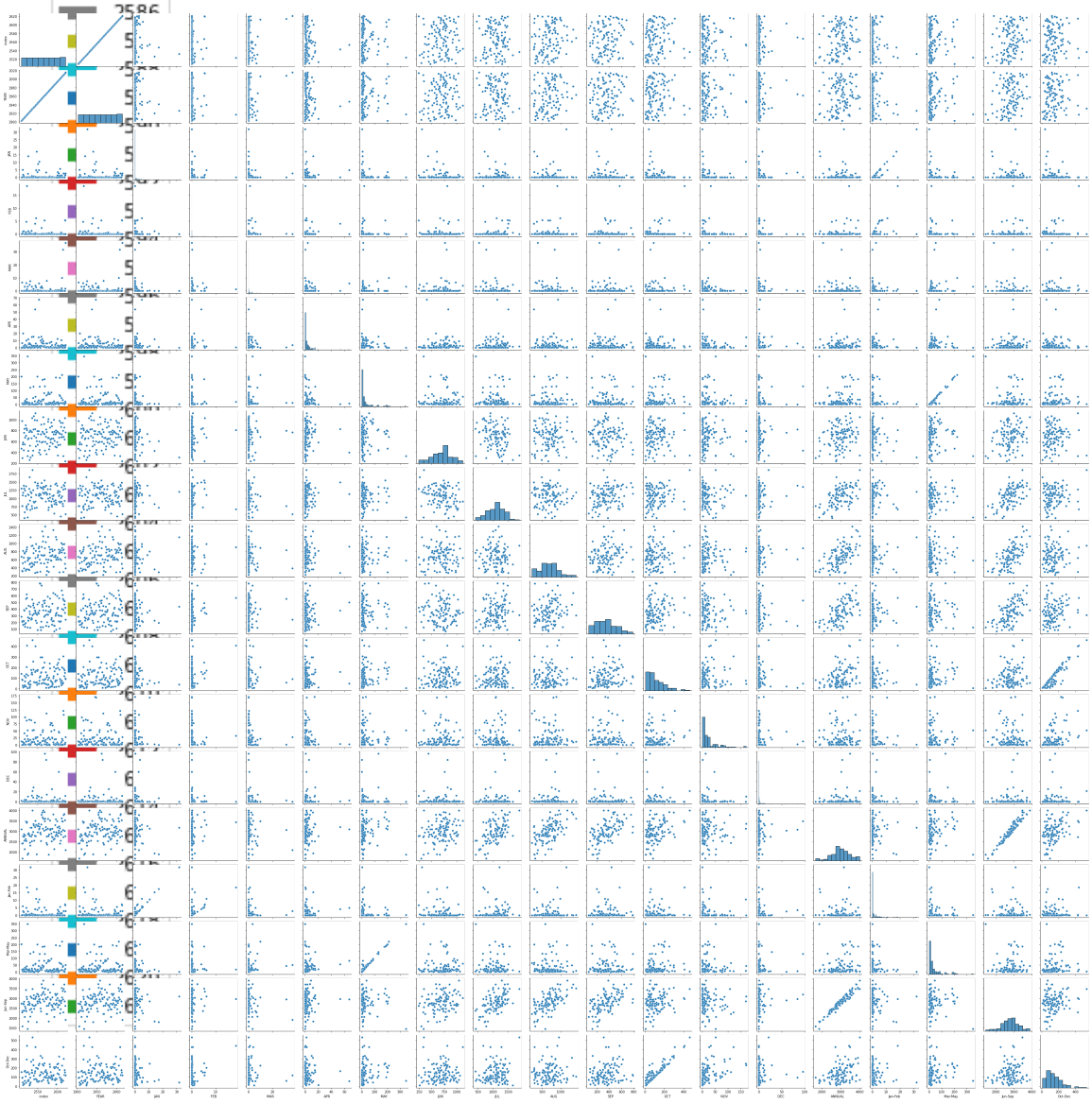
EDA AND VISUALIZATION

In [14]:

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

Out[14]:

```
<seaborn.axisgrid.PairGrid at 0x2ce691bb610>
```

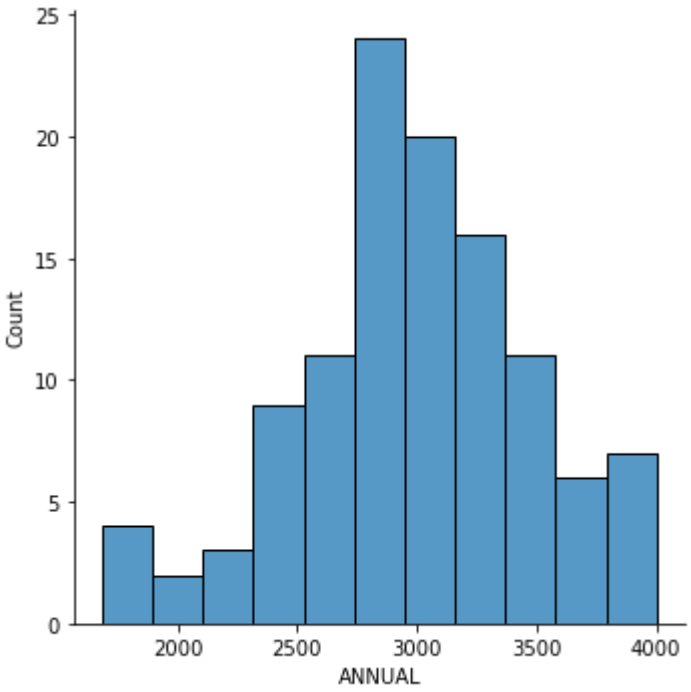


In [15]:

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

Out[15]:

```
<seaborn.axisgrid.FacetGrid at 0x2ce7564ef70>
```

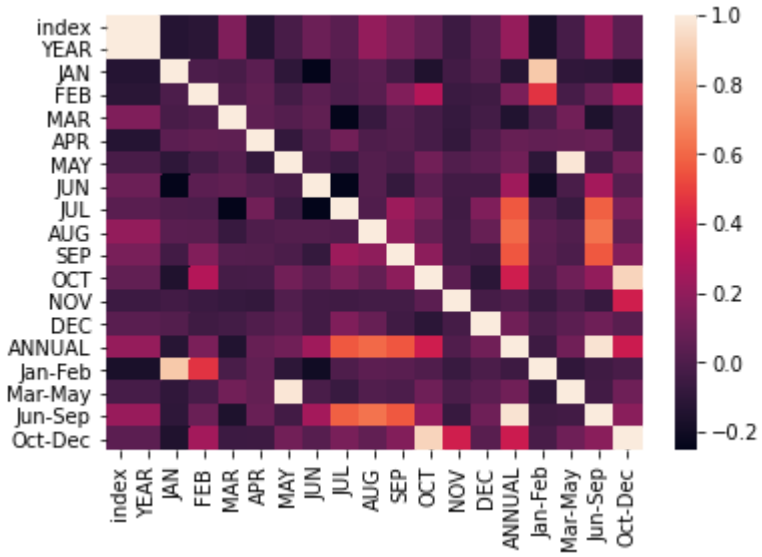


In [16]:

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

Out[16]:

```
<AxesSubplot:>
```



In [ ]: