Importing Libraries

In [93]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Importing Datasets

In [94]:
```
df = pd.read_csv(r"C:\Users\user\Downloads\New folder\CHHATTISGARH.csv")
df
```

Out[94]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OC
0	2968	CHHATTISGARH	1902	0.6	6.5	0.4	13.9	10.3	37.2	403.8	236.6	198.1	4
1	2969	CHHATTISGARH	1903	6.2	13.9	0.4	6.8	51.1	110.7	365.9	396.0	212.0	168
2	2970	CHHATTISGARH	1904	0.0	8.6	32.3	0.2	77.5	369.5	303.6	483.6	86.8	129
3	2971	CHHATTISGARH	1905	50.3	22.6	19.0	24.6	31.8	40.4	443.7	270.8	338.8	8
4	2972	CHHATTISGARH	1906	25.0	91.0	52.5	0.0	4.1	210.1	445.2	258.3	242.3	41
...	...	...	...	...	...	...	...	...	...	...	...	...	...
109	3077	CHHATTISGARH	2011	0.3	11.5	2.6	35.0	16.8	183.5	272.6	379.8	382.2	15
110	3078	CHHATTISGARH	2012	36.6	4.8	1.1	14.9	9.4	147.3	430.6	442.2	245.3	19
111	3079	CHHATTISGARH	2013	2.8	19.7	4.9	45.8	5.7	263.6	418.8	336.6	140.9	180
112	3080	CHHATTISGARH	2014	2.3	29.0	21.4	17.3	25.0	104.9	416.7	327.7	252.7	77
113	3081	CHHATTISGARH	2015	15.8	1.2	21.2	37.0	13.0	257.6	248.6	286.6	216.9	17

114 rows × 20 columns

Data Cleaning and Data Preprocessing

In [95]:
```
df=df.dropna()
```

In [96]:
```
df.columns
```

Out[96]:
```
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')
```

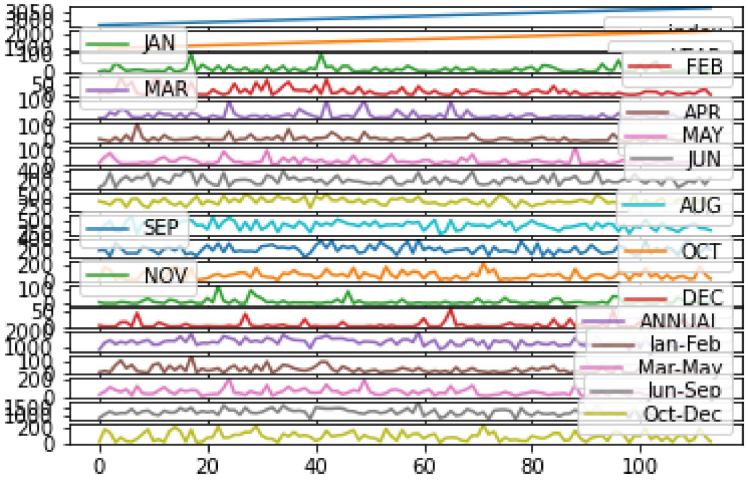
In [97]:
```
df.info()
```

```
<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
```

Line chart

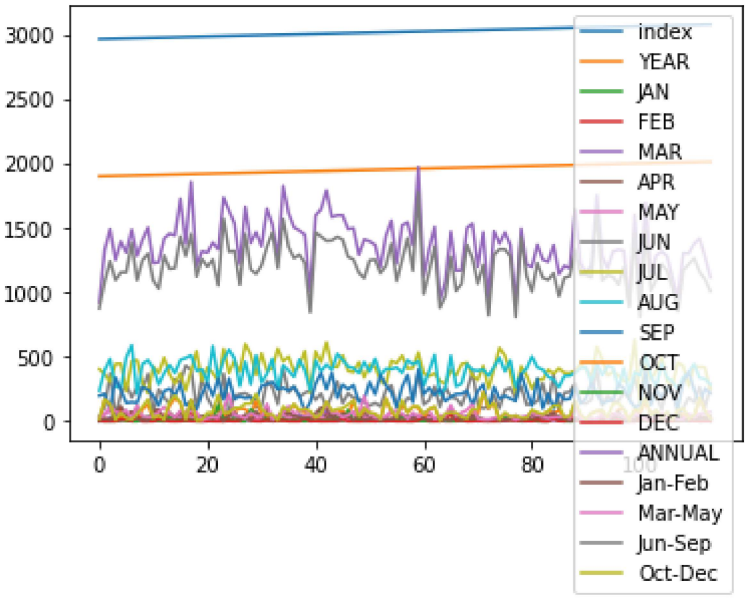
In [98]:
```
df.plot.line(subplots=True)
```

Out[98]:
```
array([<AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>], dtype=object)
```



In [99]:
```
df.plot.line()
```

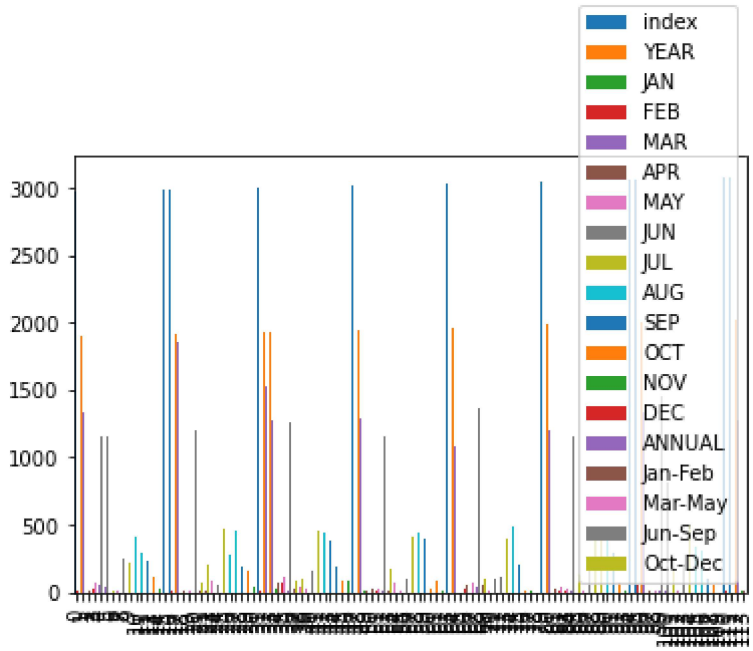
Out[99]:
```
<AxesSubplot:>
```



Bar chart

In [100]:
```
df.plot.bar()
```

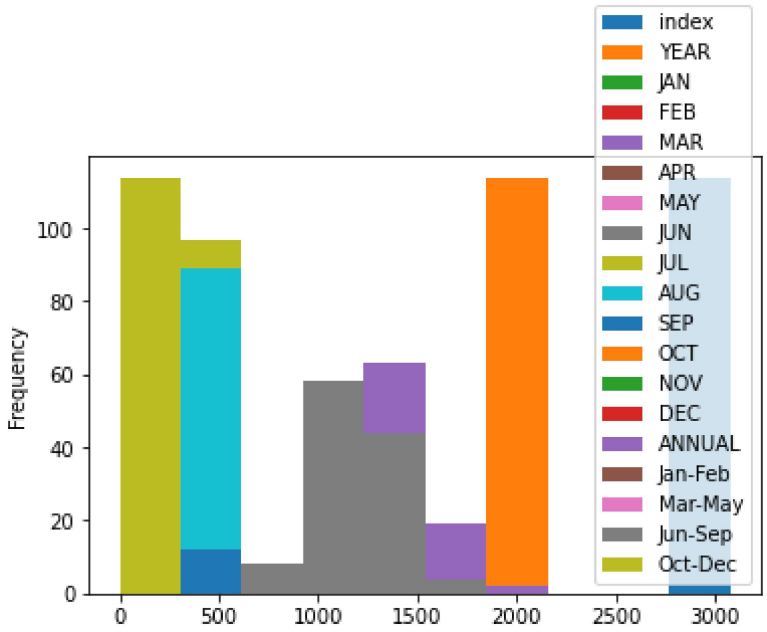
Out[100]: <AxesSubplot:>



Histogram

In [101]:
```
df.plot.hist()
```

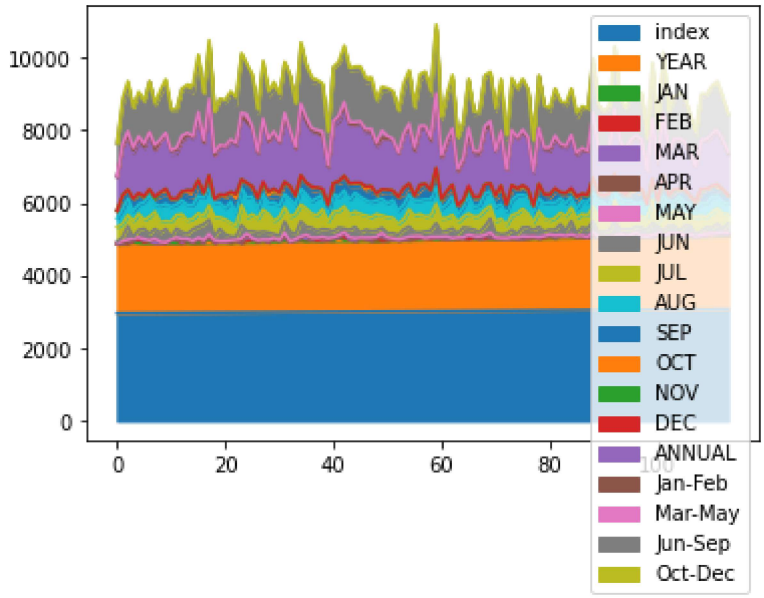
Out[101]: <AxesSubplot:ylabel='Frequency'>



Area chart

In [102]:
```
df.plot.area()
```

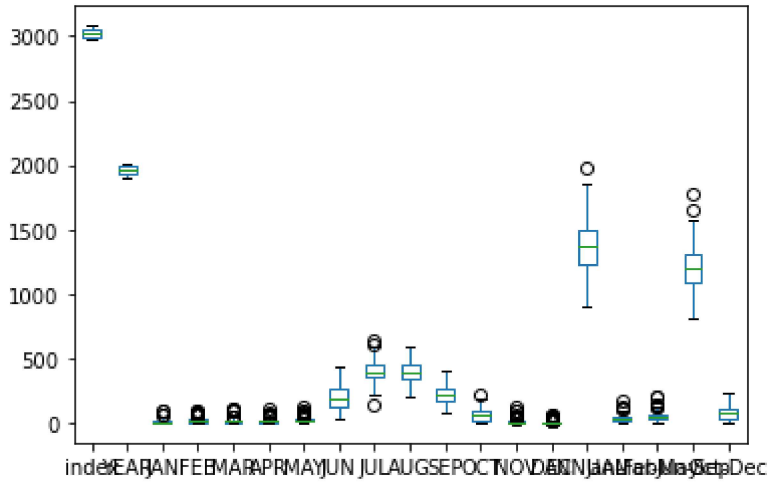
Out[102]: <AxesSubplot:>



Box plot

In [103]:
```
df.plot.box()
```

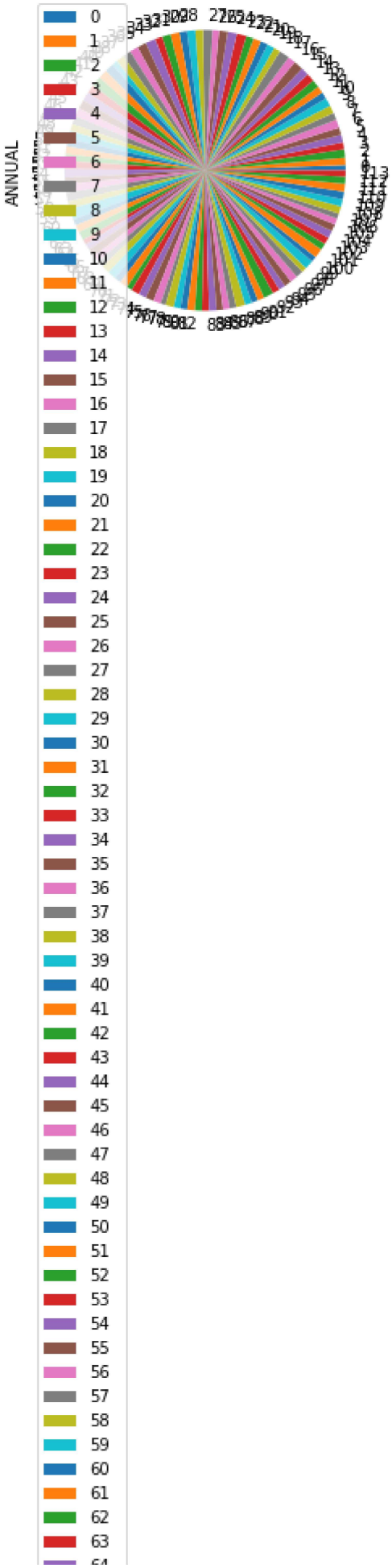
Out[103]: <AxesSubplot:>

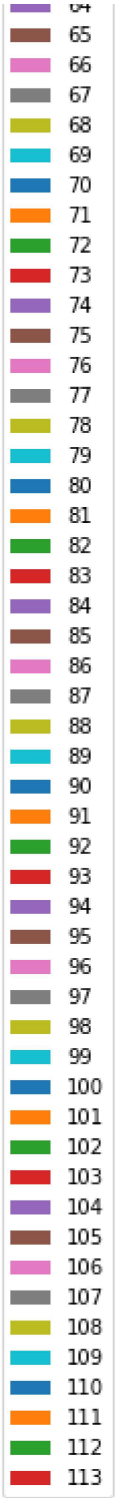


pie chart

In [104]:
```
df.plot.pie(y='ANNUAL')
```

Out[104]:
```
<AxesSubplot:ylabel='ANNUAL'>
```

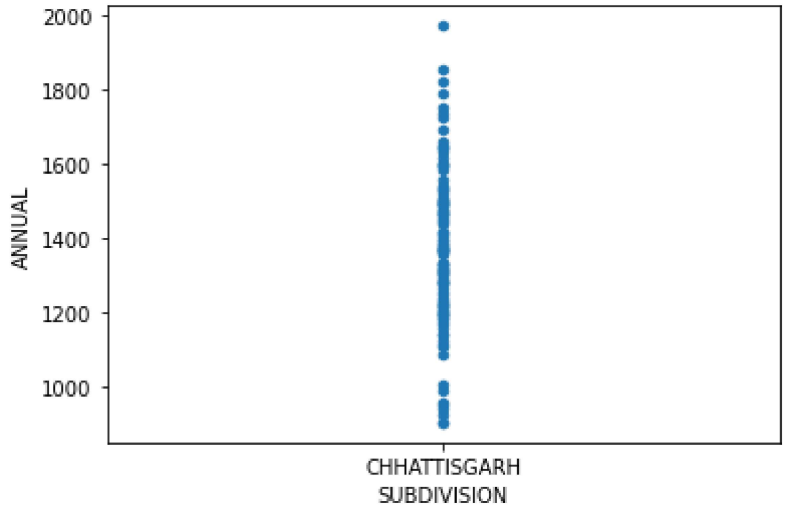

ANNUAL
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63



Scatter chart

In [105]:
```
df.plot.scatter(x='SUBDIVISION',y='ANNUAL')
```

Out[105]:
```
<AxesSubplot:xlabel='SUBDIVISION', ylabel='ANNUAL'>
```



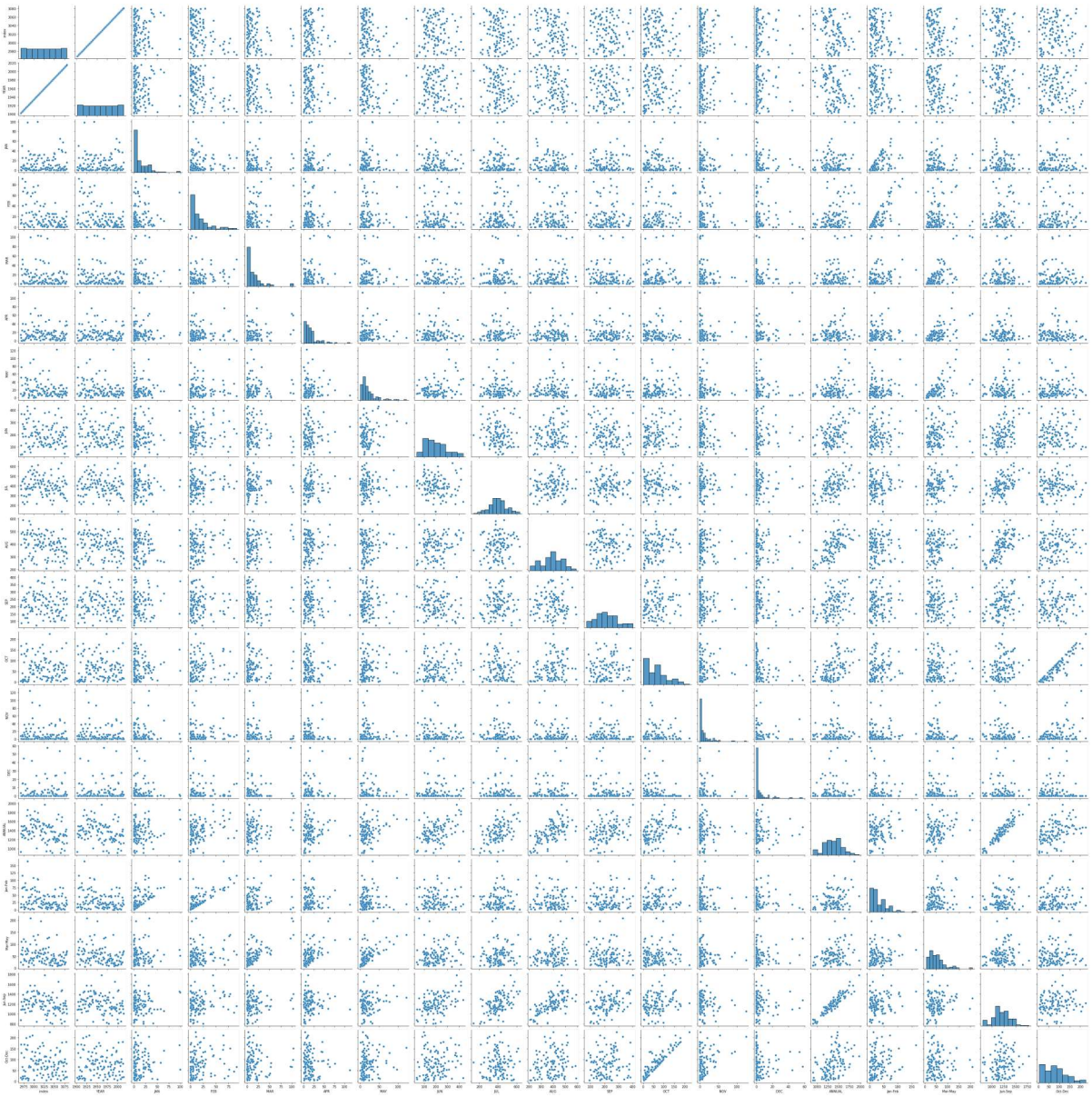
In [106]:
```
df.info()
```

```
<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
```

EDA AND VISUALIZATION

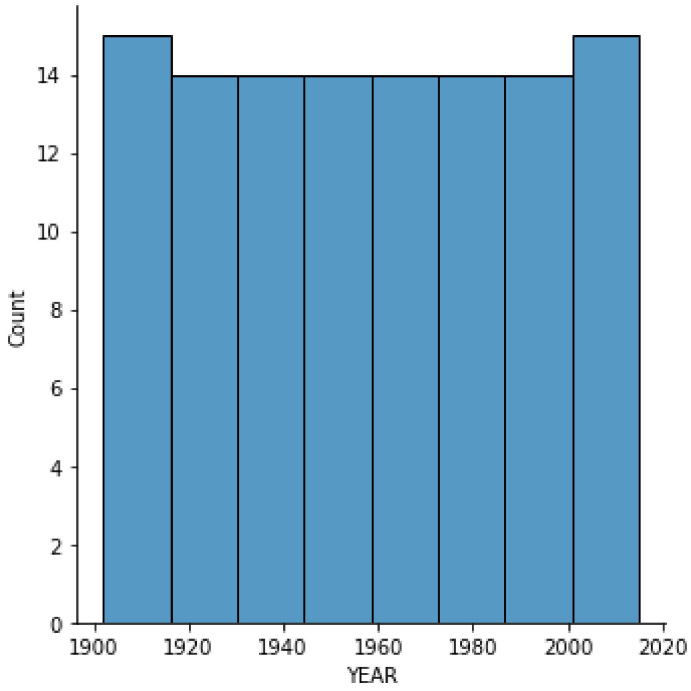
In [107]:
```
sns.pairplot(df)
```

Out[107]:
```
<seaborn.axisgrid.PairGrid at 0x1f51a6bc310>
```



In [108]:
```
sns.displot(df['YEAR'])
```

Out[108]: <seaborn.axisgrid.FacetGrid at 0x1f58152ea00>



In [109]:
```
sns.heatmap(df.corr())
```

Out[109]: <AxesSubplot:>