Importing Libraries

In [110]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Importing Datasets

In [143]:
```
df = pd.read_csv(r"C:\Users\user\Downloads\New folder\EAST MADHYA PRADESH.csv")
df
```

Out[143]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	2162	EAST MADHYA PRADESH	1901	48.5	38.1	15.7	10.7	6.2	61.0	367.5	589.2	189.9	5.9
1	2163	EAST MADHYA PRADESH	1902	14.9	8.9	0.0	3.6	2.7	28.0	411.9	227.0	236.6	17.0
2	2164	EAST MADHYA PRADESH	1903	5.6	2.9	0.3	0.9	37.5	67.5	261.4	366.7	257.4	177.9
3	2165	EAST MADHYA PRADESH	1904	2.0	15.3	48.2	0.0	8.6	109.9	443.2	316.6	135.6	44.8
4	2166	EAST MADHYA PRADESH	1905	15.9	8.0	14.3	12.3	10.2	34.4	292.4	243.3	250.9	2.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...
110	2272	EAST MADHYA PRADESH	2011	0.6	1.9	0.3	7.1	4.7	332.5	323.6	326.9	276.5	1.1
111	2273	EAST MADHYA PRADESH	2012	39.4	0.7	0.6	1.1	1.2	67.8	398.9	351.7	172.6	12.7
112	2274	EAST MADHYA PRADESH	2013	2.0	43.4	14.1	9.5	0.3	311.9	456.2	480.8	78.0	124.2
113	2275	EAST MADHYA PRADESH	2014	32.1	49.7	17.8	5.1	2.5	91.8	283.4	231.8	139.6	56.4
114	2276	EAST MADHYA PRADESH	2015	37.3	11.0	73.4	25.8	6.3	139.2	262.2	272.1	71.6	38.2

115 rows × 20 columns

Data Cleaning and Data Preprocessing

In [144]:
```
df=df.dropna()
```

In [145]:
```
df.columns
```

Out[145]:
```
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')
```

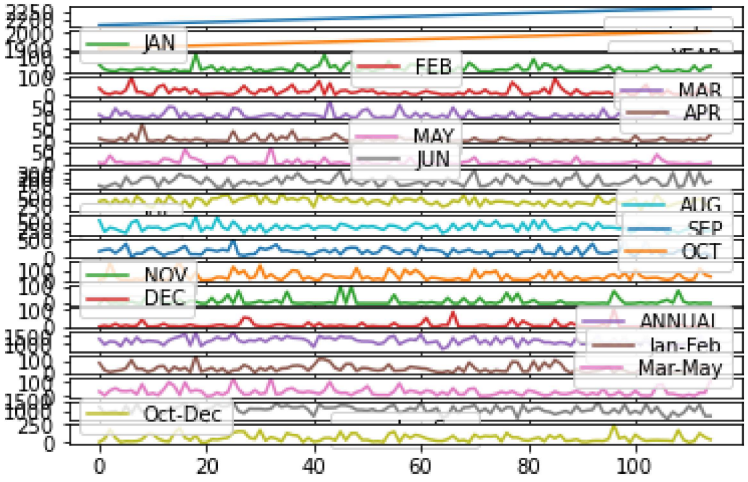
In [146]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 115 entries, 0 to 114
Data columns (total 20 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   index        115 non-null    int64  
 1   SUBDIVISION  115 non-null    object 
 2   YEAR         115 non-null    int64  
 3   JAN          115 non-null    float64
 4   FEB          115 non-null    float64
 5   MAR          115 non-null    float64
 6   APR          115 non-null    float64
 7   MAY          115 non-null    float64
 8   JUN          115 non-null    float64
 9   JUL          115 non-null    float64
 10  AUG          115 non-null    float64
 11  SEP          115 non-null    float64
 12  OCT          115 non-null    float64
 13  NOV          115 non-null    float64
 14  DEC          115 non-null    float64
 15  ANNUAL       115 non-null    float64
 16  Jan-Feb      115 non-null    float64
 17  Mar-May      115 non-null    float64
 18  Jun-Sep      115 non-null    float64
 19  Oct-Dec      115 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 18.9+ KB
```

Line chart

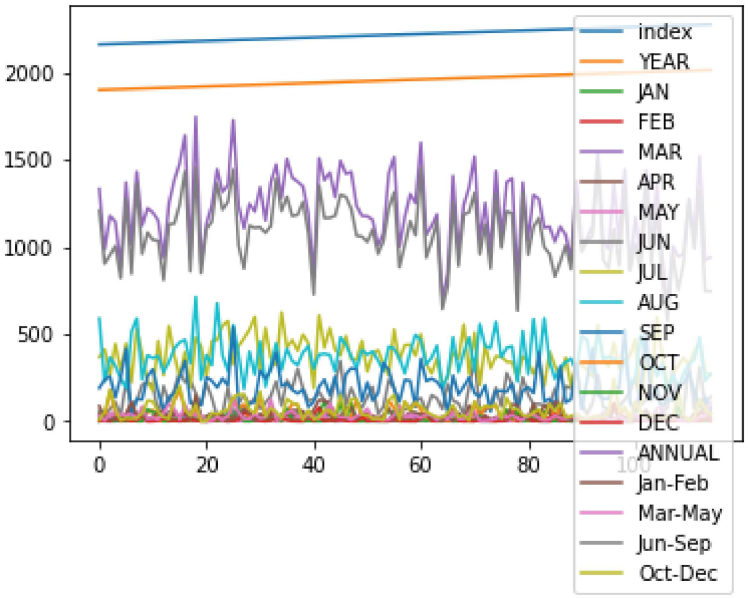
In [147]:
```
df.plot.line(subplots=True)
```

Out[147]:
```
array([<AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>], dtype=object)
```



In [148]:
```
df.plot.line()
```

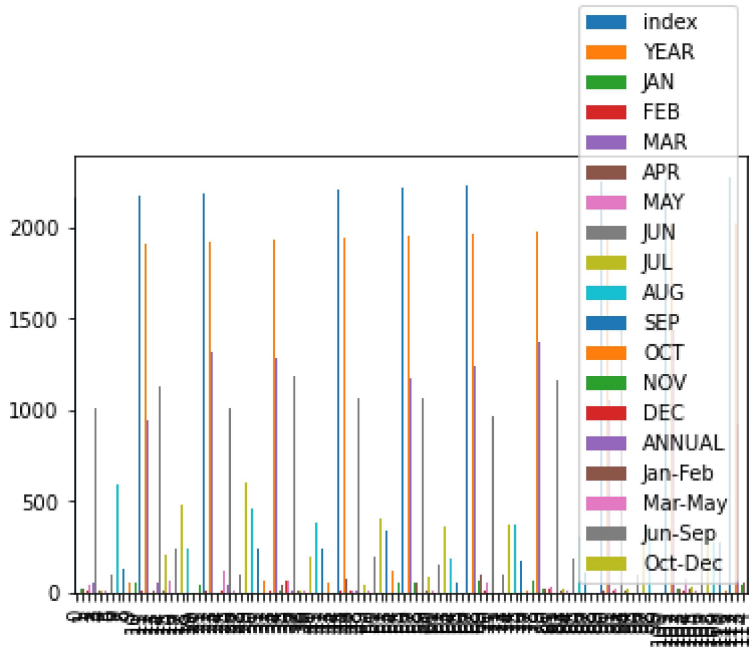
Out[148]:
```
<AxesSubplot:>
```



Bar chart

In [149]:
```
df.plot.bar()
```

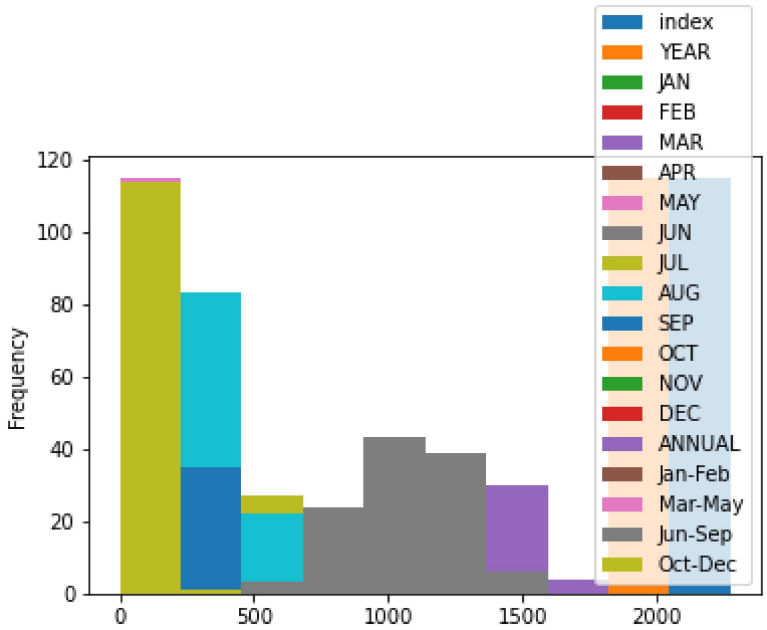
Out[149]: <AxesSubplot:>



Histogram

In [150]:
```
df.plot.hist()
```

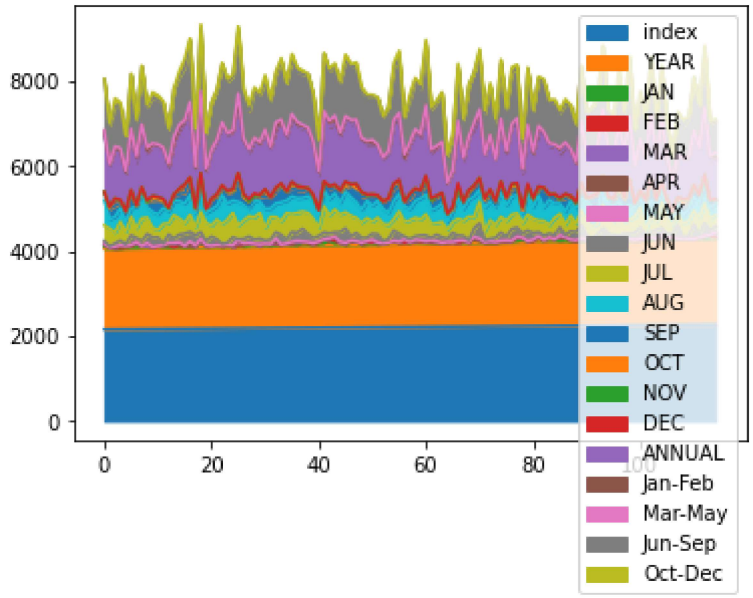
Out[150]: <AxesSubplot:ylabel='Frequency'>



Area chart

In [151]:
```
df.plot.area()
```

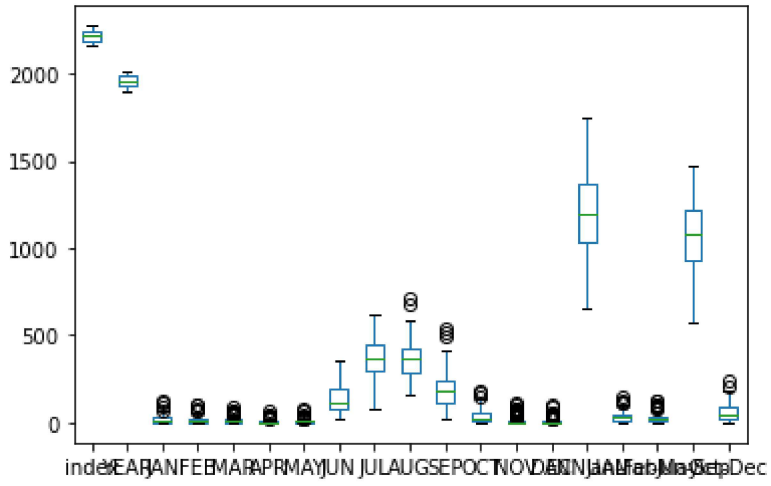
Out[151]: <AxesSubplot:>



Box plot

In [152]:
```
df.plot.box()
```

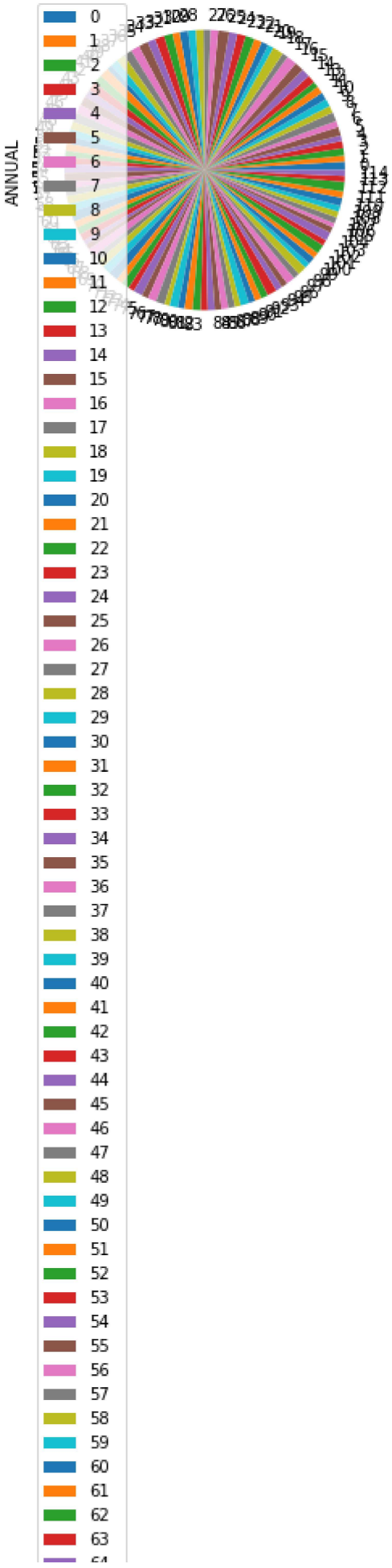
Out[152]: <AxesSubplot:>

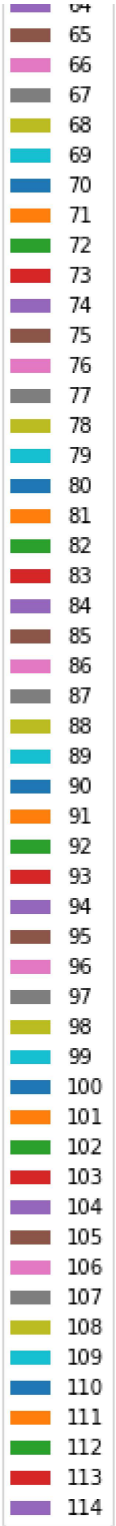


pie chart

In [153]:
```
df.plot.pie(y='ANNUAL')
```

Out[153]:
```
<AxesSubplot:ylabel='ANNUAL'>
```

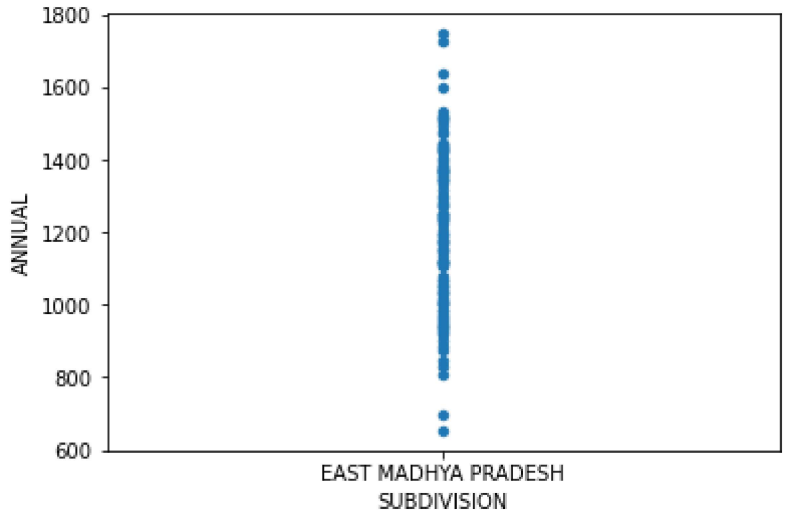

ANNUAL
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63



Scatter chart

In [154]:
```
df.plot.scatter(x='SUBDIVISION',y='ANNUAL')
```

Out[154]:
```
<AxesSubplot:xlabel='SUBDIVISION', ylabel='ANNUAL'>
```



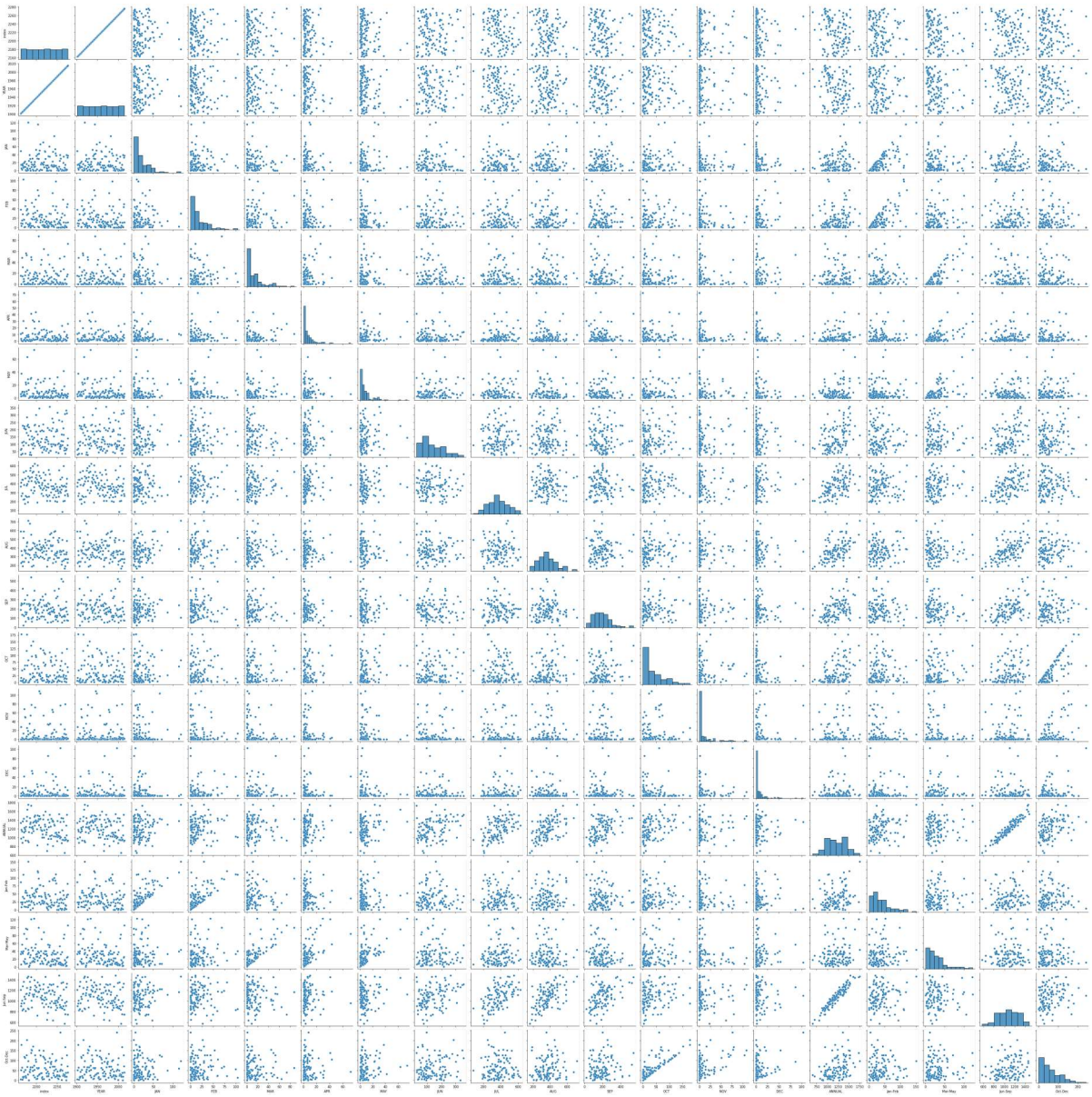
In [155]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 115 entries, 0 to 114
Data columns (total 20 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   index        115 non-null    int64  
 1   SUBDIVISION  115 non-null    object 
 2   YEAR         115 non-null    int64  
 3   JAN          115 non-null    float64
 4   FEB          115 non-null    float64
 5   MAR          115 non-null    float64
 6   APR          115 non-null    float64
 7   MAY          115 non-null    float64
 8   JUN          115 non-null    float64
 9   JUL          115 non-null    float64
 10  AUG          115 non-null    float64
 11  SEP          115 non-null    float64
 12  OCT          115 non-null    float64
 13  NOV          115 non-null    float64
 14  DEC          115 non-null    float64
 15  ANNUAL       115 non-null    float64
 16  Jan-Feb      115 non-null    float64
 17  Mar-May      115 non-null    float64
 18  Jun-Sep      115 non-null    float64
 19  Oct-Dec      115 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 18.9+ KB
```

EDA AND VISUALIZATION

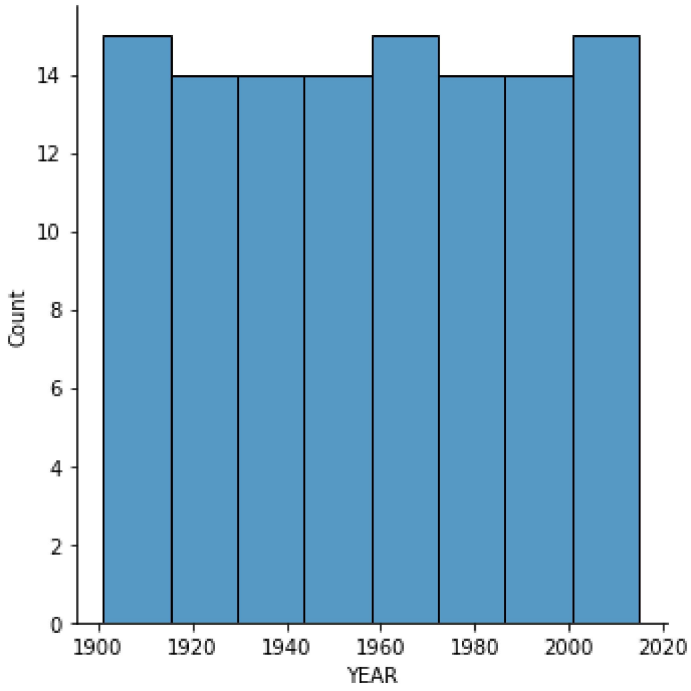
In [156]:
```
sns.pairplot(df)
```

Out[156]:
```
<seaborn.axisgrid.PairGrid at 0x1f5b3fd7670>
```



In [157]:
```
sns.displot(df['YEAR'])
```

Out[157]:
```
<seaborn.axisgrid.FacetGrid at 0x1f5c364ef70>
```



In [158]:
```
sns.heatmap(df.corr())
```

Out[158]:
```
<AxesSubplot:>
```