Importing Libraries

In [199]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Importing Datasets

In [265]:
```
df = pd.read_csv(r"C:\Users\user\Downloads\New folder\MADHYA MAHARASHTRA.csv")
df
```

Out[265]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OC
0	2623	MADHYA MAHARASHTRA	1902	7.8	0.0	0.1	5.0	9.8	102.6	210.9	114.5	169.5	60.
1	2624	MADHYA MAHARASHTRA	1903	7.6	0.0	0.0	3.2	77.2	86.3	281.8	155.5	142.3	74.
2	2625	MADHYA MAHARASHTRA	1904	0.4	4.7	1.7	3.0	18.7	114.6	126.5	59.5	183.0	91.
3	2626	MADHYA MAHARASHTRA	1905	0.0	1.2	0.0	2.3	23.6	65.0	252.8	79.0	52.6	52.
4	2627	MADHYA MAHARASHTRA	1906	10.5	0.8	0.0	0.1	9.3	184.8	199.3	205.0	88.8	19.
...	...	...	...	...	...	...	...	...	...	...	...	...	...
109	2732	MADHYA MAHARASHTRA	2011	0.0	0.3	0.3	5.0	2.9	133.3	261.4	238.1	148.4	62.
110	2733	MADHYA MAHARASHTRA	2012	0.0	0.0	0.0	3.0	1.4	67.9	203.0	187.8	129.5	95.
111	2734	MADHYA MAHARASHTRA	2013	0.1	5.3	0.8	5.7	6.0	212.4	311.8	147.0	210.3	57.
112	2735	MADHYA MAHARASHTRA	2014	3.1	6.2	24.4	7.5	29.8	44.0	277.9	240.3	120.4	38.
113	2736	MADHYA MAHARASHTRA	2015	1.4	0.8	41.2	9.6	24.4	177.0	111.7	67.2	146.6	48.

114 rows × 20 columns

Data Cleaning and Data Preprocessing

In [266]:
```
df=df.dropna()
```

In [267]:
```
df.columns
```

Out[267]:
```
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')
```

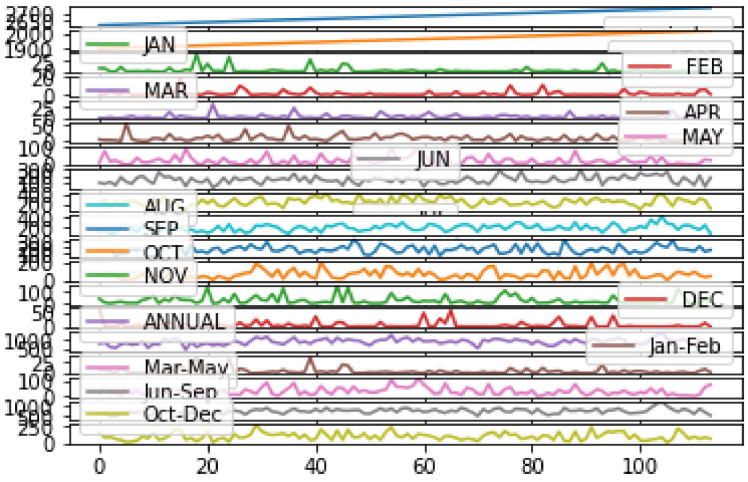
In [268]:
```
df.info()
```

```
<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

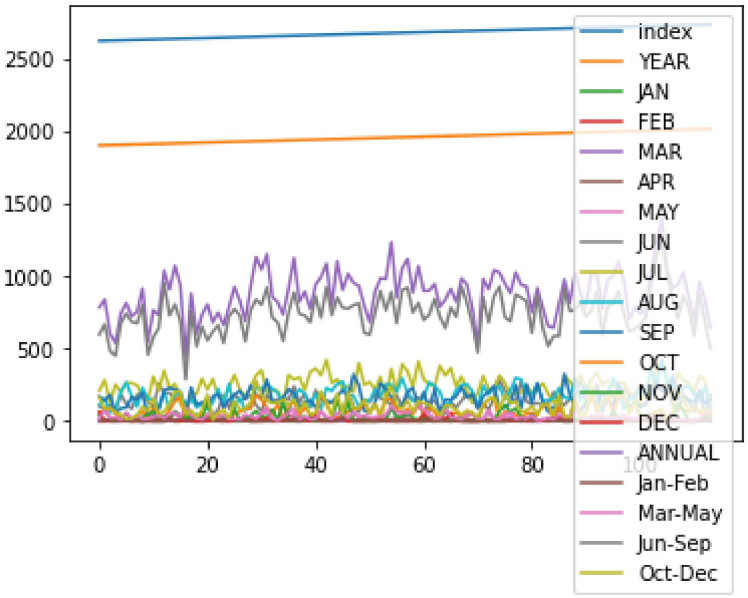
In [269]:
```
df.plot.line(subplots=True)
```

Out[269]:
```
array([<AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>], dtype=object)
```



In [270]:
```
df.plot.line()
```

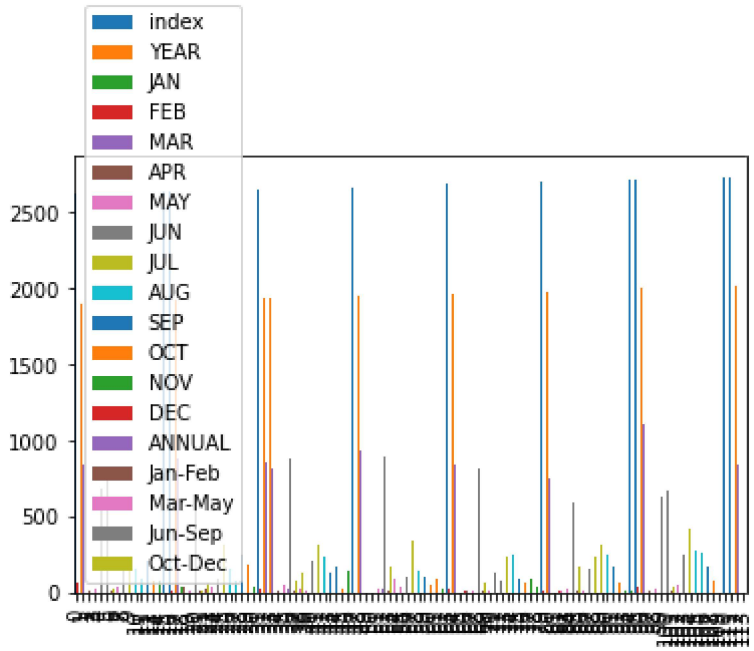
Out[270]:
```
<AxesSubplot:>
```



Bar chart

In [271]:
```
df.plot.bar()
```

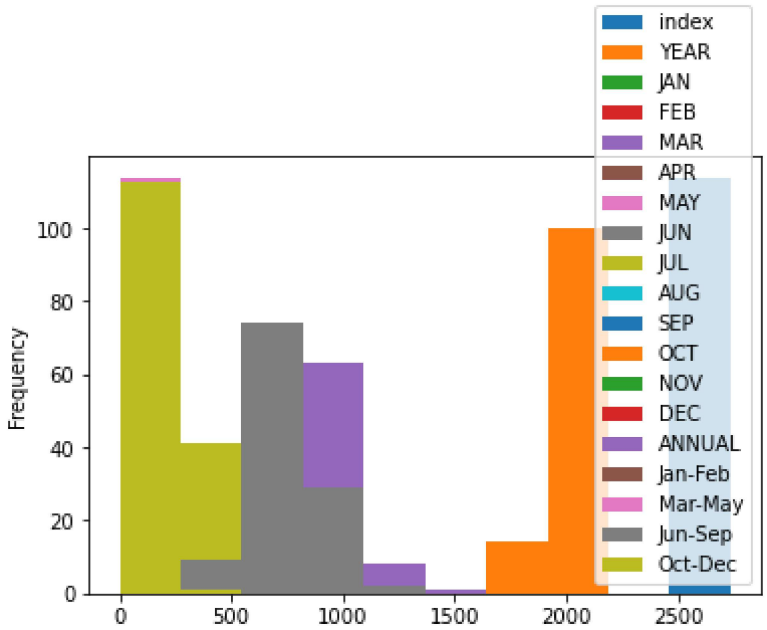
Out[271]: <AxesSubplot:>



Histogram

In [272]:
```
df.plot.hist()
```

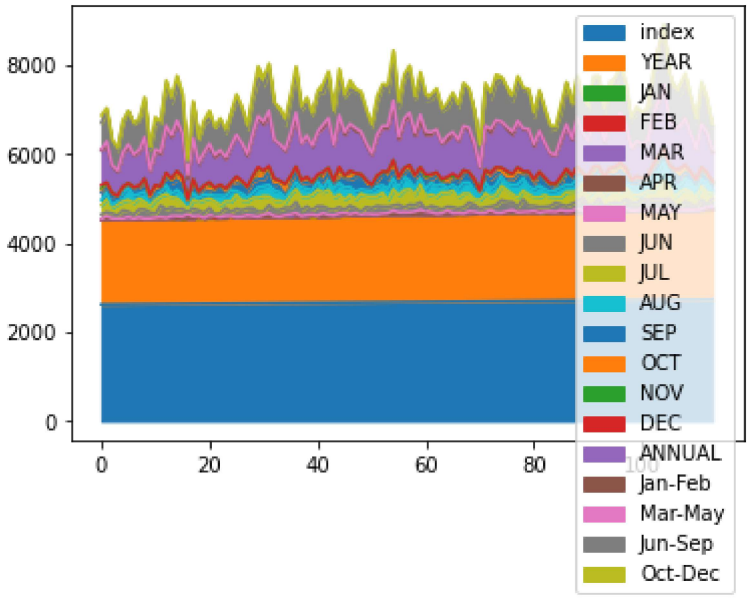
Out[272]: <AxesSubplot:ylabel='Frequency'>



Area chart

In [273]:
```
df.plot.area()
```

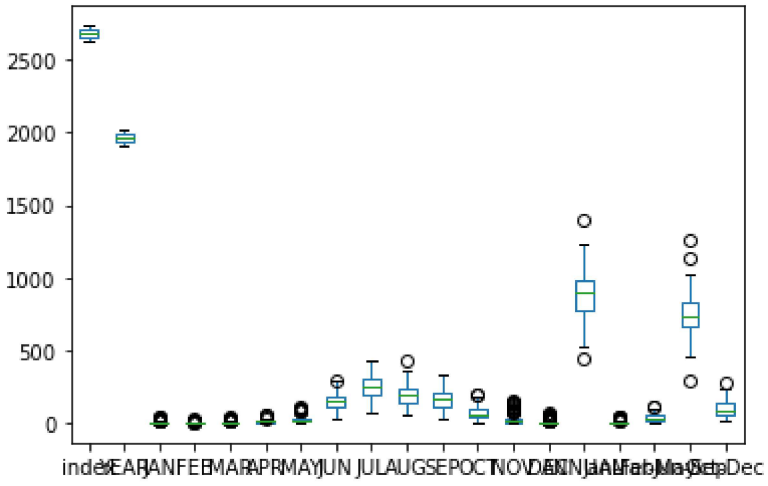
Out[273]: <AxesSubplot:>



Box plot

In [274]:
```
df.plot.box()
```

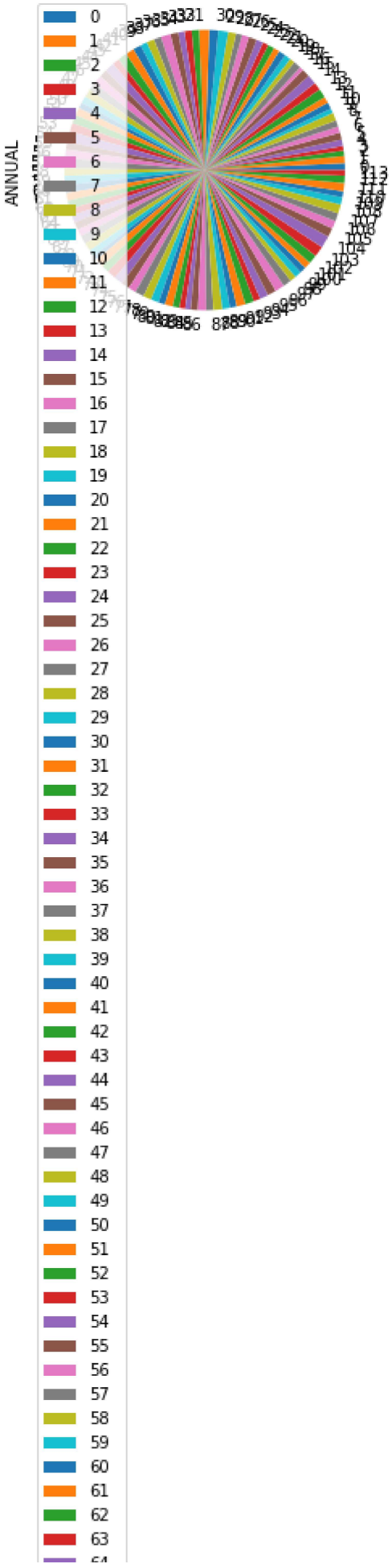
Out[274]: <AxesSubplot:>

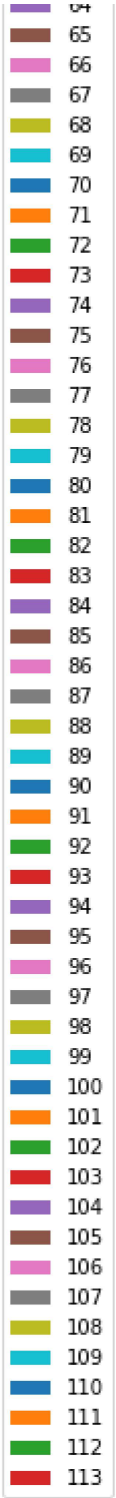


pie chart

In [275]:
```
df.plot.pie(y='ANNUAL')
```

Out[275]:
```
<AxesSubplot:ylabel='ANNUAL'>
```

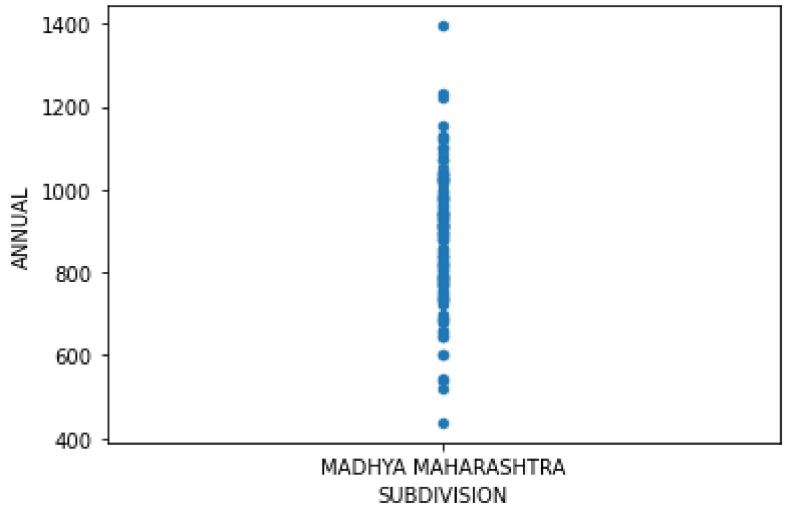

ANNUAL
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63



Scatter chart

In [276]:
```
df.plot.scatter(x='SUBDIVISION',y='ANNUAL')
```

Out[276]:
```
<AxesSubplot:xlabel='SUBDIVISION', ylabel='ANNUAL'>
```



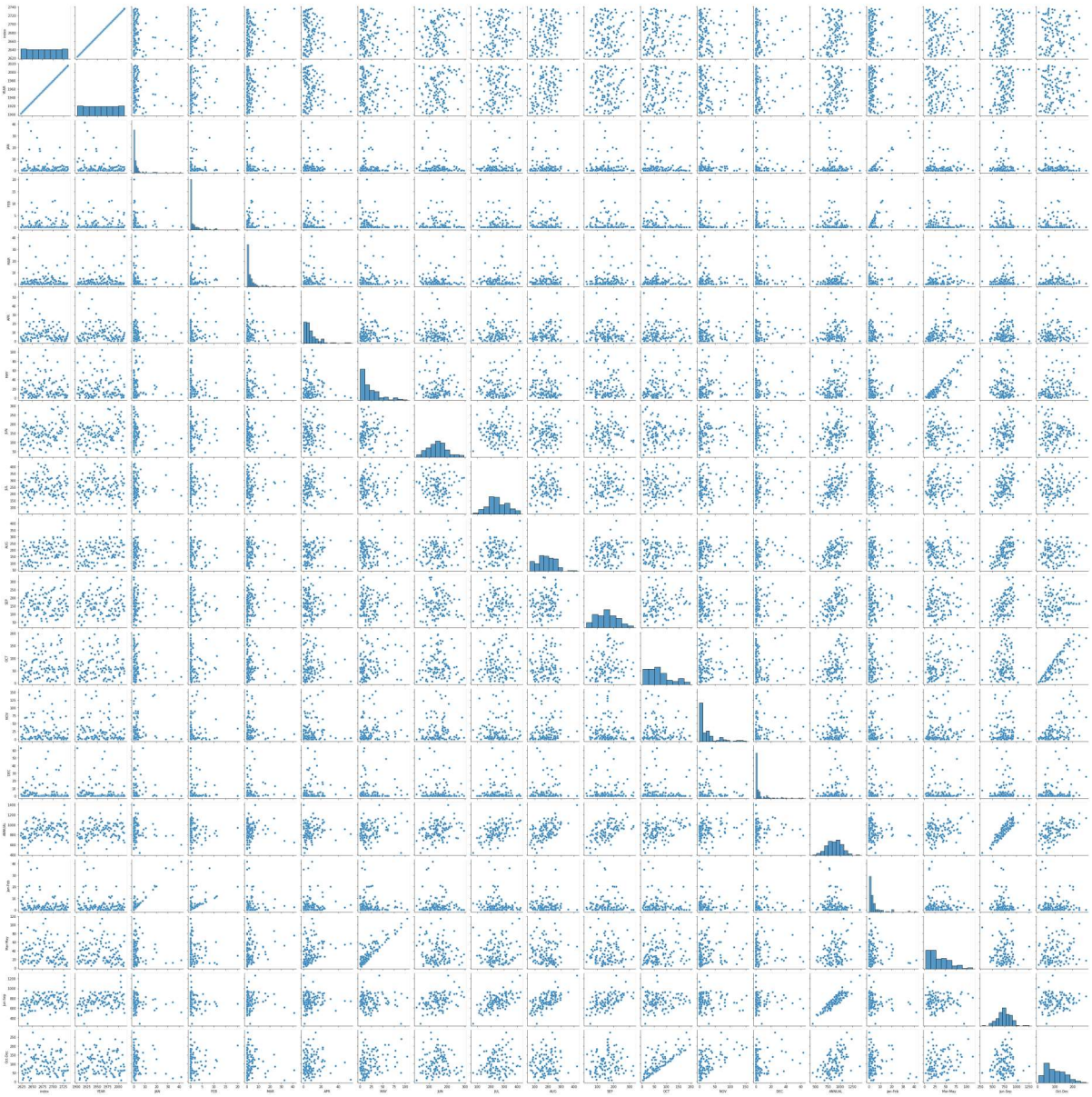
In [277]:
```
df.info()
```

```
<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

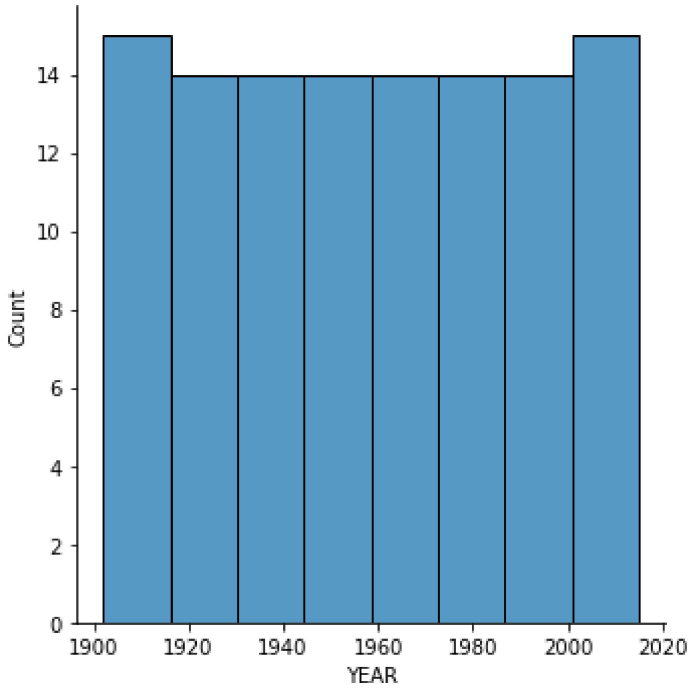
In [278]:
```
sns.pairplot(df)
```

Out[278]:
```
<seaborn.axisgrid.PairGrid at 0x1f65be52280>
```



In [279]:
```
sns.displot(df['YEAR'])
```

Out[279]:
```
<seaborn.axisgrid.FacetGrid at 0x1f6673ec0d0>
```



In [280]:
```
sns.heatmap(df.corr())
```

Out[280]:
```
<AxesSubplot:>
```