Importing Libraries

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

Importing Datasets

In [127]:
```
df = pd.read_csv(r"C:\Users\user\Downloads\New folder\COASTAL KARNATAKA.csv")
df
```

Out[127]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	C
0	3543	COASTAL KARNATAKA	1902	3.2	0.3	4.9	10.2	54.6	698.4	1401.6	454.2	708.4	18
1	3544	COASTAL KARNATAKA	1903	0.7	0.0	0.0	4.1	202.8	536.5	1405.5	593.8	304.4	18
2	3545	COASTAL KARNATAKA	1904	2.4	0.0	4.8	23.7	93.2	1108.2	1070.0	465.6	245.3	12
3	3546	COASTAL KARNATAKA	1905	0.0	0.2	0.0	6.4	83.1	767.3	777.3	586.9	172.9	22
4	3547	COASTAL KARNATAKA	1906	23.0	0.0	0.0	0.5	29.8	593.6	1173.4	535.0	273.3	12
...	...	...	...	...	...	...	...	...	...	...	...	...	...
109	3652	COASTAL KARNATAKA	2011	4.8	3.8	8.7	66.1	49.3	1018.4	1080.5	861.3	545.2	17
110	3653	COASTAL KARNATAKA	2012	NaN	11.4	5.1	77.0	22.9	650.9	754.6	1027.6	382.0	1
111	3654	COASTAL KARNATAKA	2013	2.4	19.6	19.0	28.5	100.4	1153.0	1515.3	680.2	379.1	26
112	3655	COASTAL KARNATAKA	2014	0.0	0.3	1.9	40.5	181.9	507.0	1155.4	1121.0	379.3	22
113	3656	COASTAL KARNATAKA	2015	1.4	1.0	32.3	72.2	150.3	735.3	930.9	575.2	260.3	20

114 rows × 20 columns

Data Cleaning and Data Preprocessing

In [128]:
```
df=df.dropna()
```

In [129]:
```
df.columns
```

Out[129]:
```
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')
```

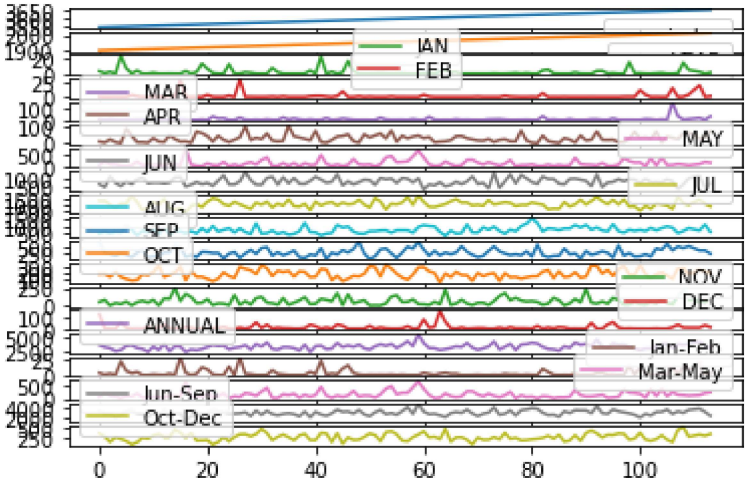
In [130]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 113 entries, 0 to 113
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: 18.5+ KB
```

Line chart

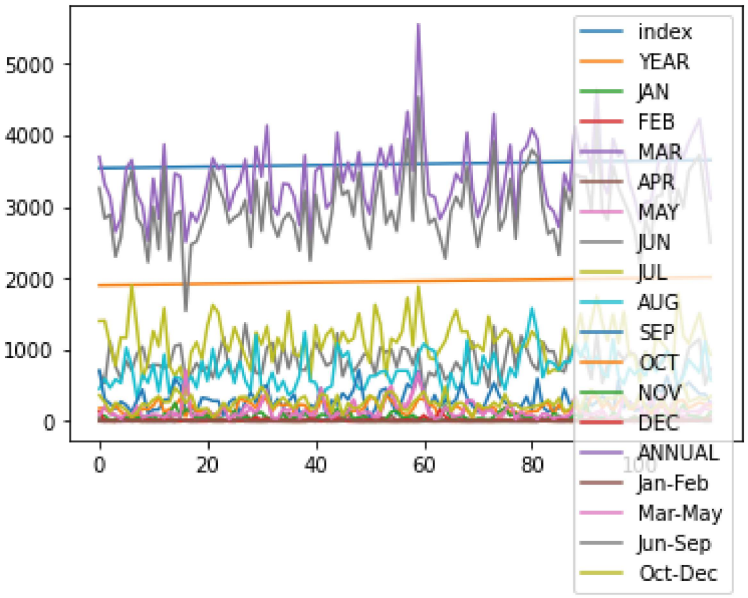
In [131]:
```
df.plot.line(subplots=True)
```

Out[131]:
```
array([<AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>], dtype=object)
```



In [132]:
```
df.plot.line()
```

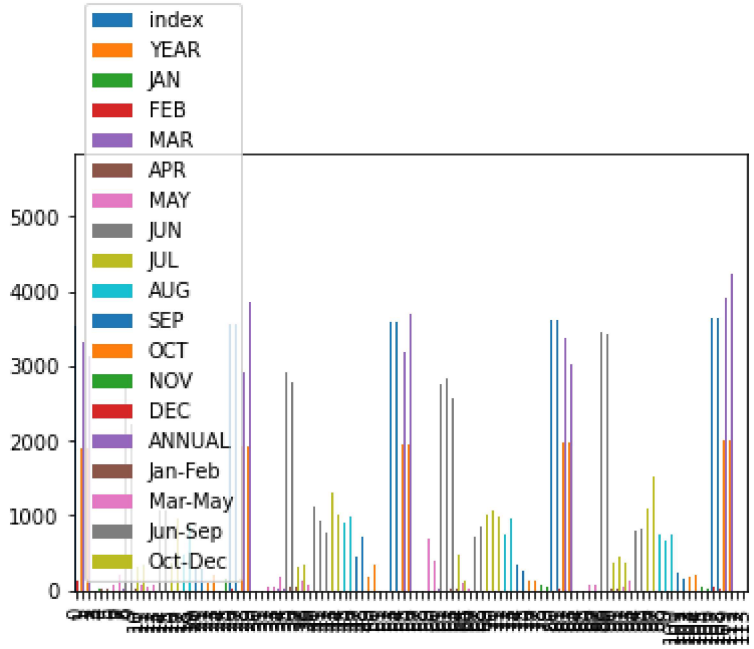
Out[132]:
```
<AxesSubplot:>
```



Bar chart

In [133]:
```
df.plot.bar()
```

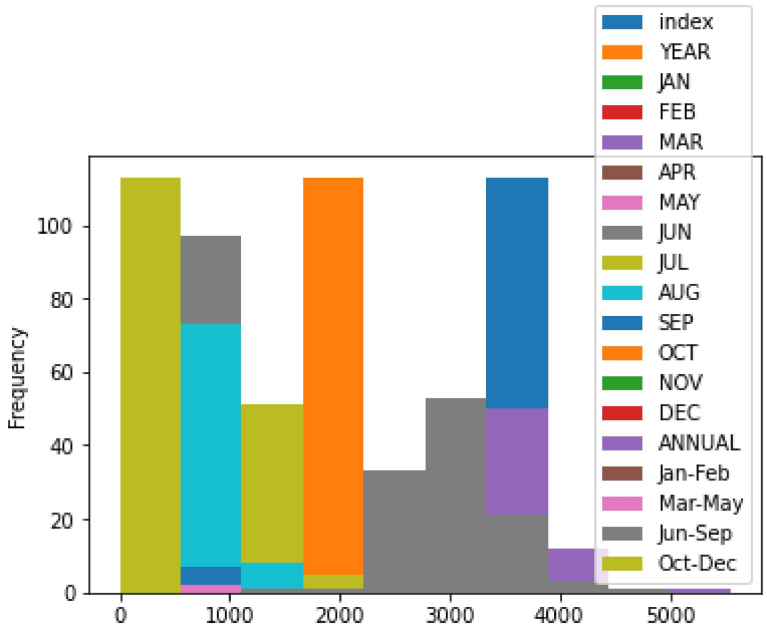
Out[133]: <AxesSubplot:>



Histogram

In [134]:
```
df.plot.hist()
```

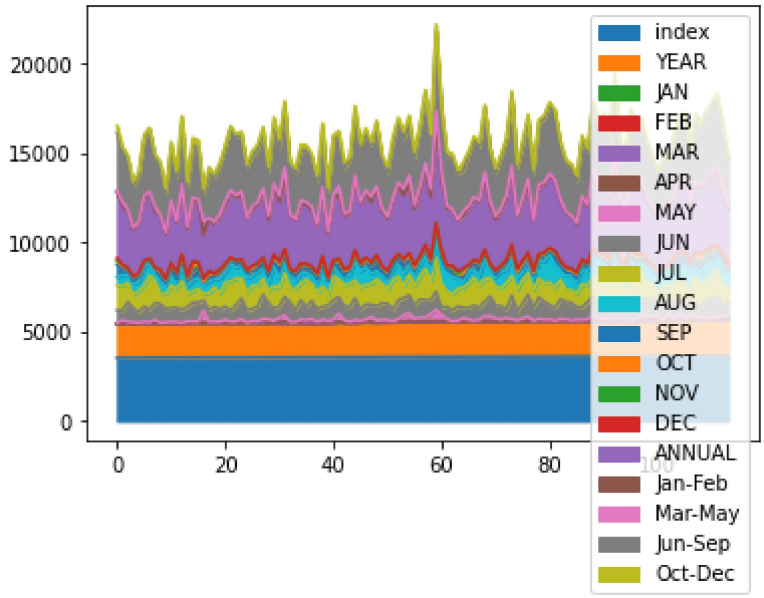
Out[134]: <AxesSubplot:ylabel='Frequency'>



Area chart

In [135]:
```
df.plot.area()
```

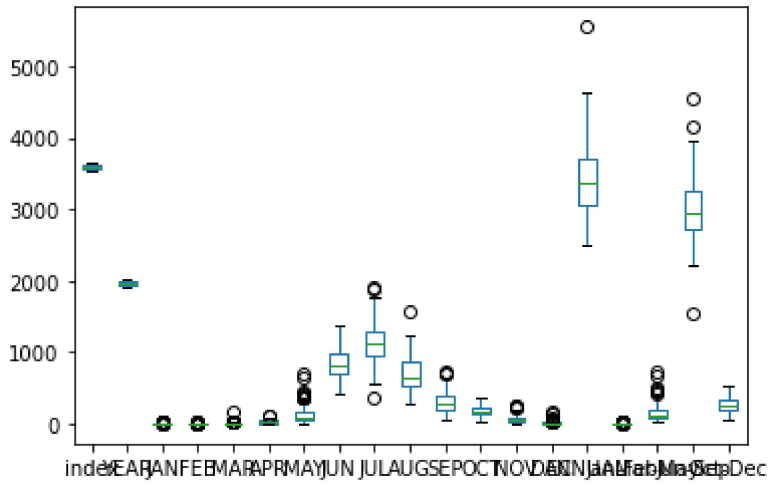
Out[135]: <AxesSubplot:>



Box plot

In [136]:
```
df.plot.box()
```

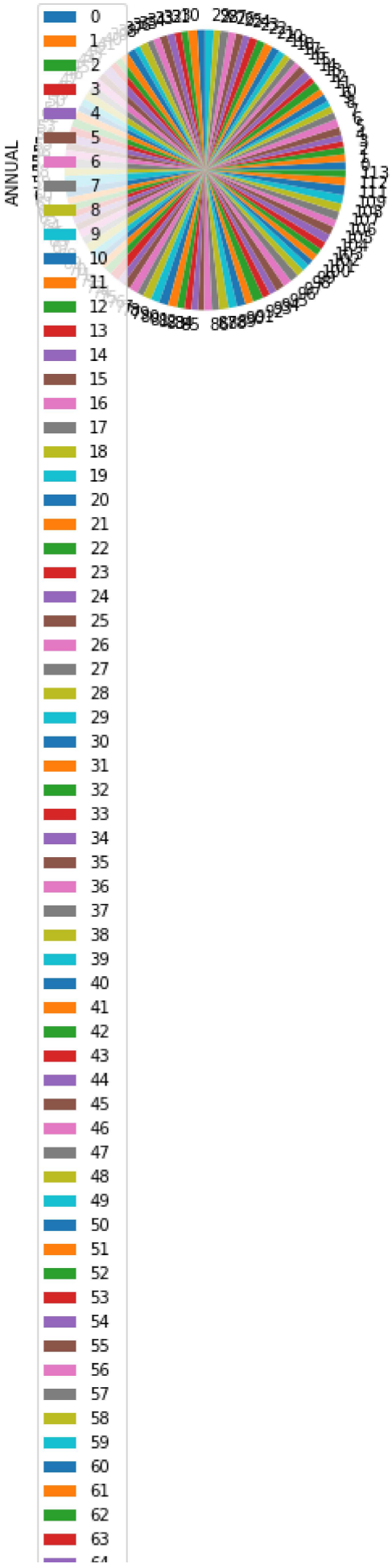
Out[136]: <AxesSubplot:>

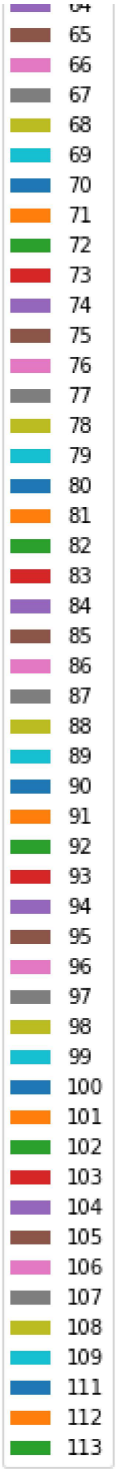


pie chart

In [137]:
```
df.plot.pie(y='ANNUAL')
```

Out[137]:
```
<AxesSubplot:ylabel='ANNUAL'>
```

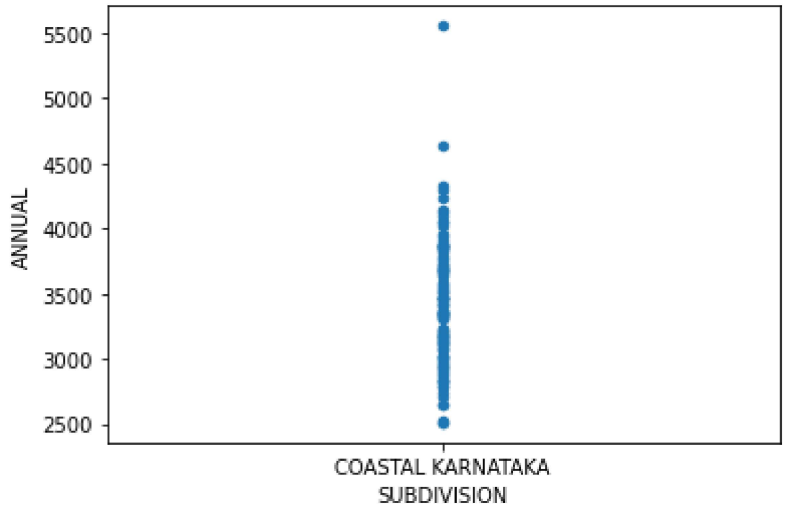

ANNUAL
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63



Scatter chart

In [138]:
```
df.plot.scatter(x='SUBDIVISION',y='ANNUAL')
```

Out[138]:
```
<AxesSubplot:xlabel='SUBDIVISION', ylabel='ANNUAL'>
```



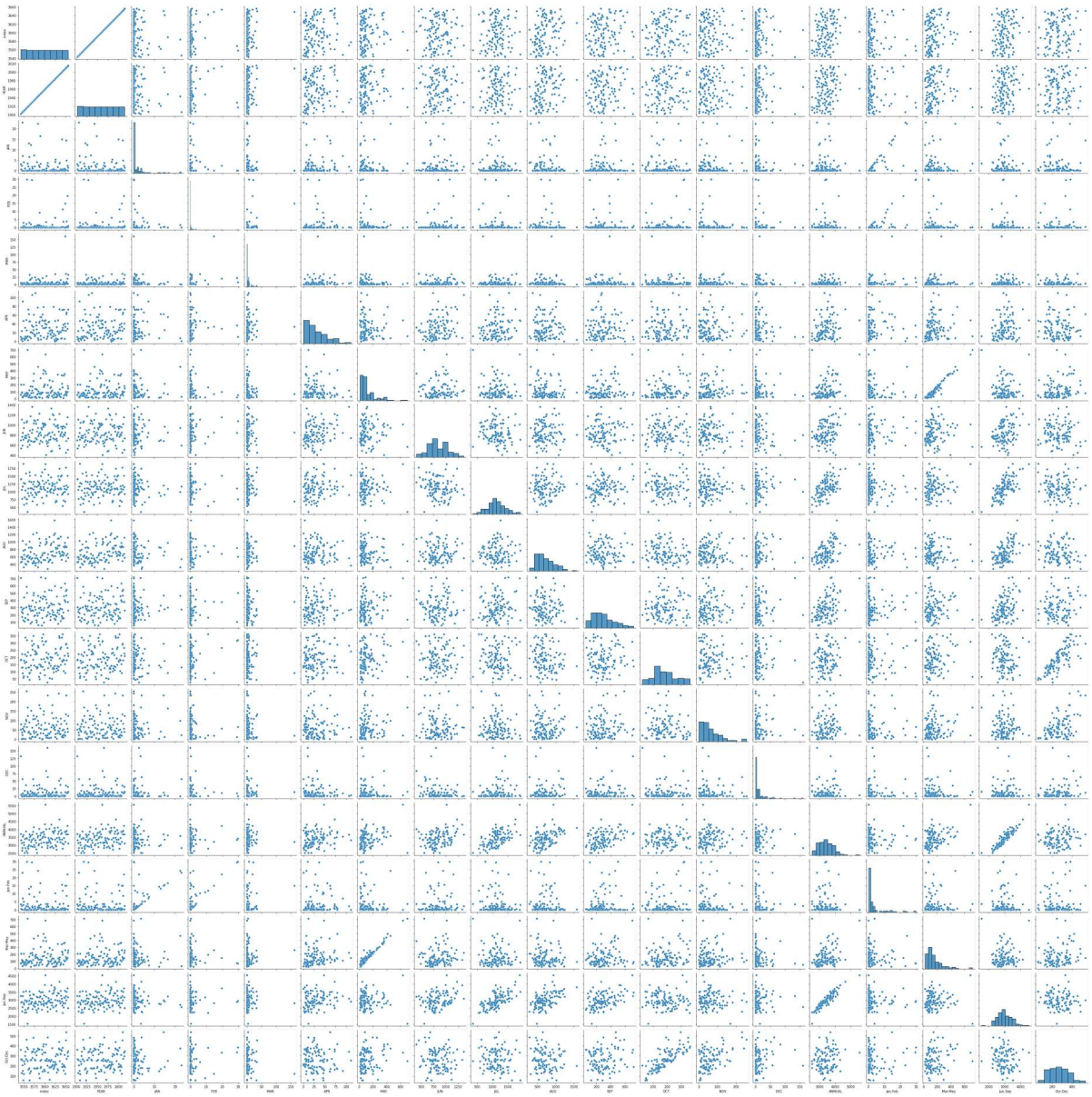
In [139]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 113 entries, 0 to 113
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: 18.5+ KB
```

EDA AND VISUALIZATION

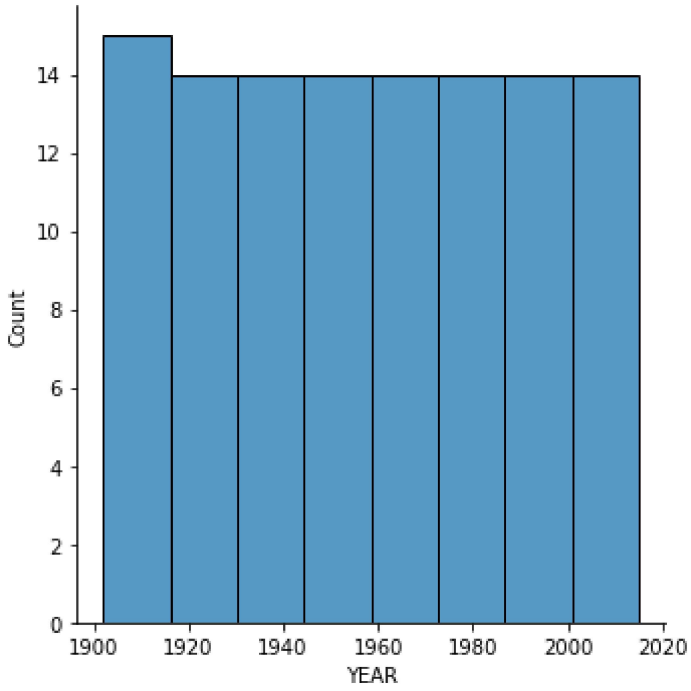
In [140]:
```
sns.pairplot(df)
```

Out[140]:
```
<seaborn.axisgrid.PairGrid at 0x1f59e985f10>
```



In [141]:
```
sns.displot(df['YEAR'])
```

Out[141]: <seaborn.axisgrid.FacetGrid at 0x1f5abe36e50>



In [142]:
```
sns.heatmap(df.corr())
```

Out[142]: <AxesSubplot:>