

# Importing Libraries

In [199]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

# Importing Datasets

In [249]:

```
df = pd.read_csv(r"C:\Users\user\Downloads\New folder\HIMACHAL PRADESH.csv")
df
```

Out[249]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	1587	HIMACHAL PRADESH	1901	137.8	174.5	75.0	19.2	89.6	32.7	280.5	459.7	53.0	3.0
1	1588	HIMACHAL PRADESH	1902	6.5	27.0	104.4	76.2	61.3	78.8	258.6	199.3	113.4	23.0
2	1589	HIMACHAL PRADESH	1903	76.5	21.4	213.7	25.4	54.7	32.2	157.7	256.5	107.9	5.0
3	1590	HIMACHAL PRADESH	1904	79.3	22.4	131.7	48.0	90.3	33.1	241.1	184.3	56.4	51.0
4	1591	HIMACHAL PRADESH	1905	81.3	76.8	160.2	39.3	50.4	43.6	191.1	132.8	119.1	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
110	1697	HIMACHAL PRADESH	2011	43.9	97.4	49.7	62.4	45.1	118.3	177.7	380.2	120.3	6.0
111	1698	HIMACHAL PRADESH	2012	92.3	51.3	28.4	55.9	9.4	31.1	241.5	280.6	133.1	3.0
112	1699	HIMACHAL PRADESH	2013	79.9	182.6	76.6	28.9	32.6	233.6	208.8	240.0	65.8	21.0
113	1700	HIMACHAL PRADESH	2014	69.6	124.9	125.2	60.6	68.9	51.7	203.6	146.7	84.6	19.0
114	1701	HIMACHAL PRADESH	2015	67.2	156.6	192.5	84.9	45.0	85.8	249.9	195.9	75.5	17.0

115 rows × 20 columns

# Data Cleaning and Data Preprocessing

```
In [250]: df=df.dropna()
```

```
In [251]: df.columns
```

```
Out[251]: 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')
```

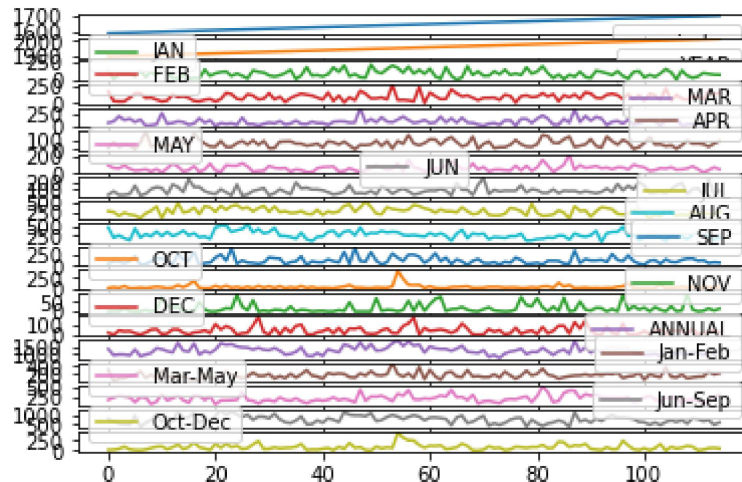
```
In [252]: df.info()
```

```
<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

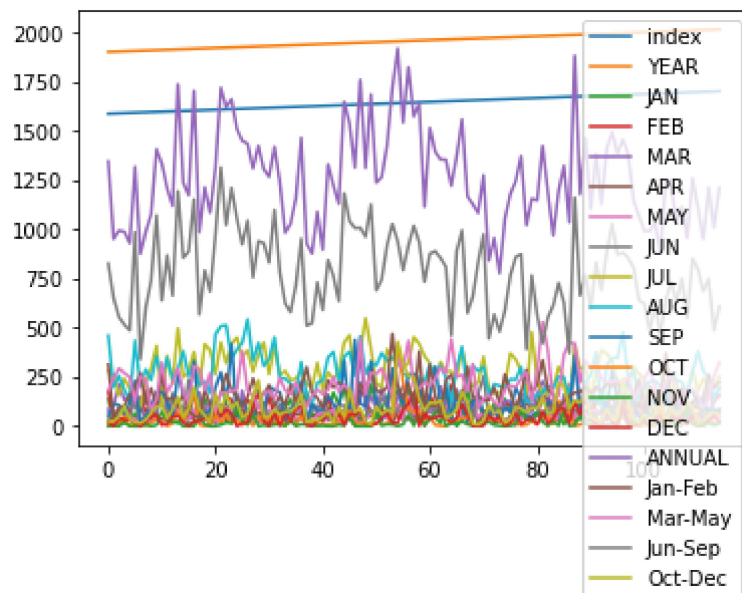
```
In [253]: df.plot.line(subplots=True)
```

```
Out[253]: array([<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>], dtype=object)
```



```
In [254]: df.plot.line()
```

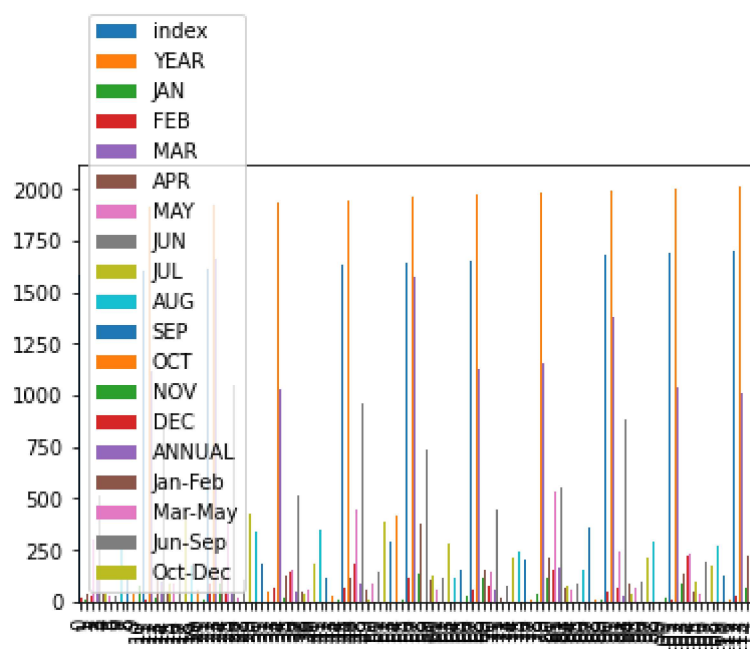
```
Out[254]: <AxesSubplot:~>
```



## Bar chart

```
In [255]: df.plot.bar()
```

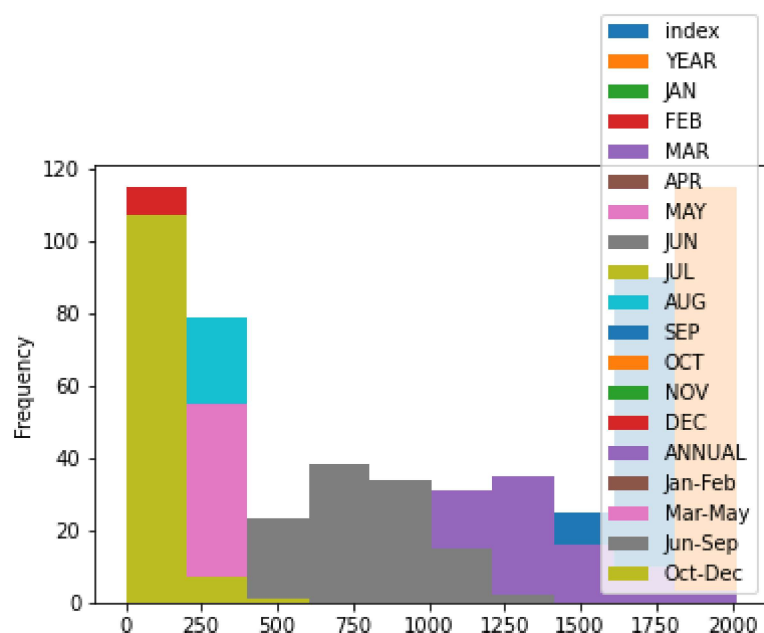
```
Out[255]: <AxesSubplot:>
```



## Histogram

```
In [256]: df.plot.hist()
```

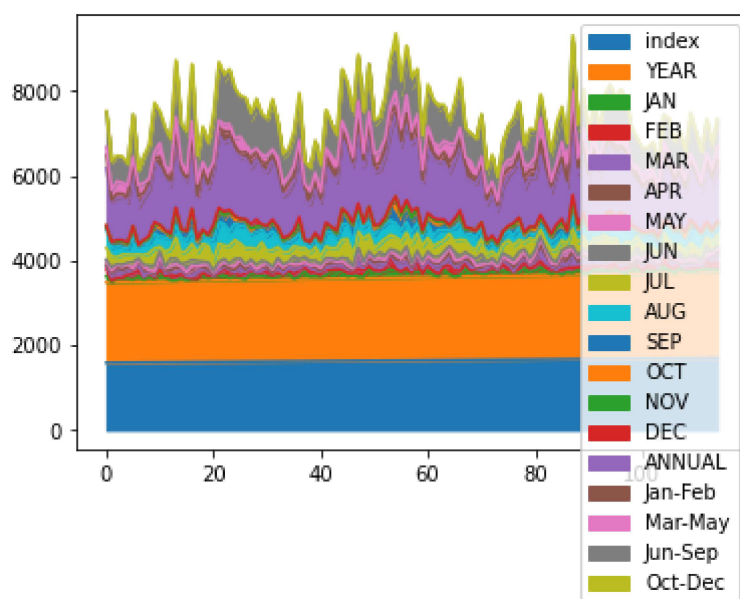
```
Out[256]: <AxesSubplot:ylabel='Frequency'>
```



## Area chart

In [257]: `df.plot.area()`

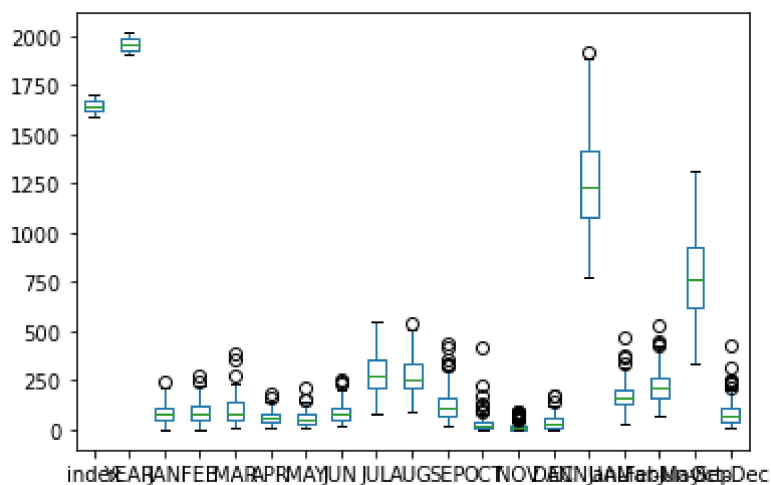
Out[257]: `<AxesSubplot:>`



## Box plot

In [258]: `df.plot.box()`

Out[258]: `<AxesSubplot:>`

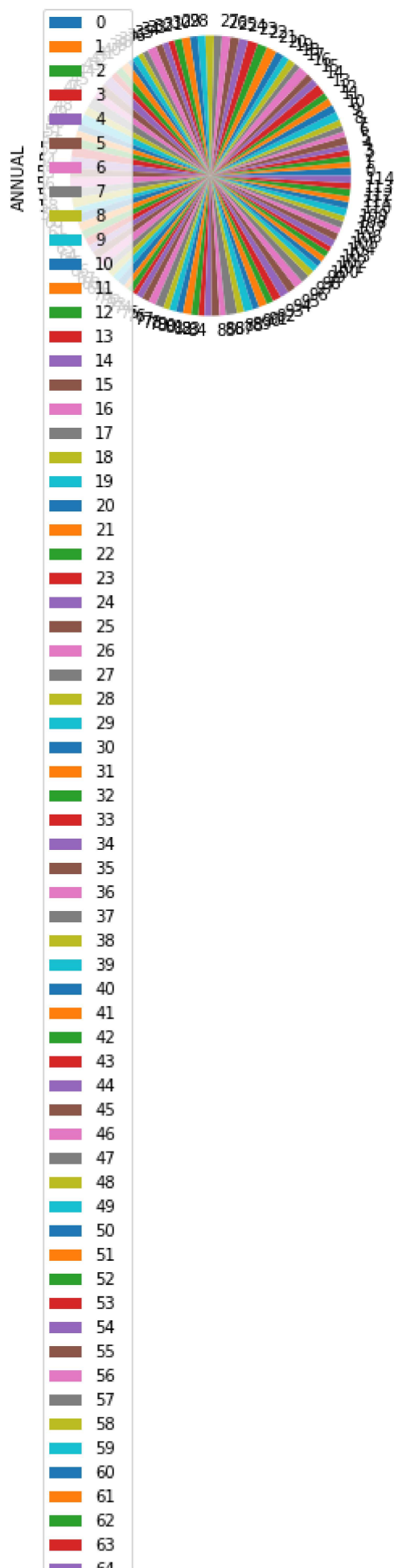


## pie chart

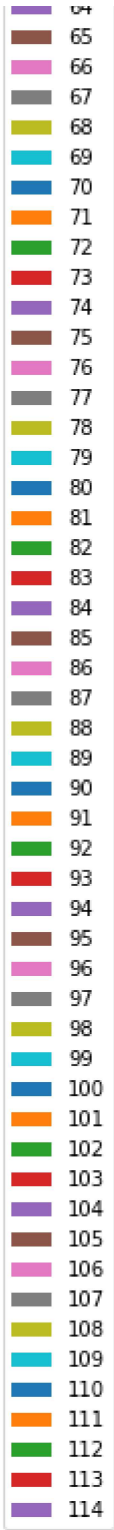
```
In [259]: df.plot.pie(y='ANNUAL')
```

```
Out[259]: <AxesSubplot:ylabel='ANNUAL'>
```





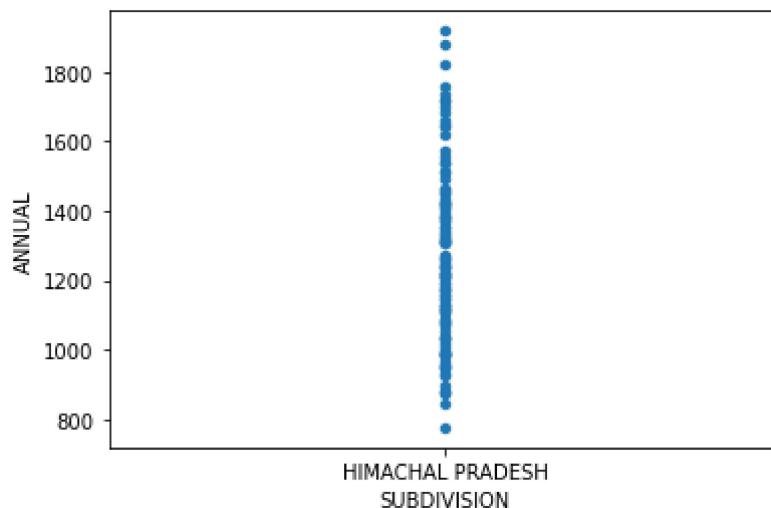




Scatter chart

```
In [260]: df.plot.scatter(x='SUBDIVISION',y='ANNUAL')
```

```
Out[260]: <AxesSubplot:xlabel='SUBDIVISION', ylabel='ANNUAL'>
```



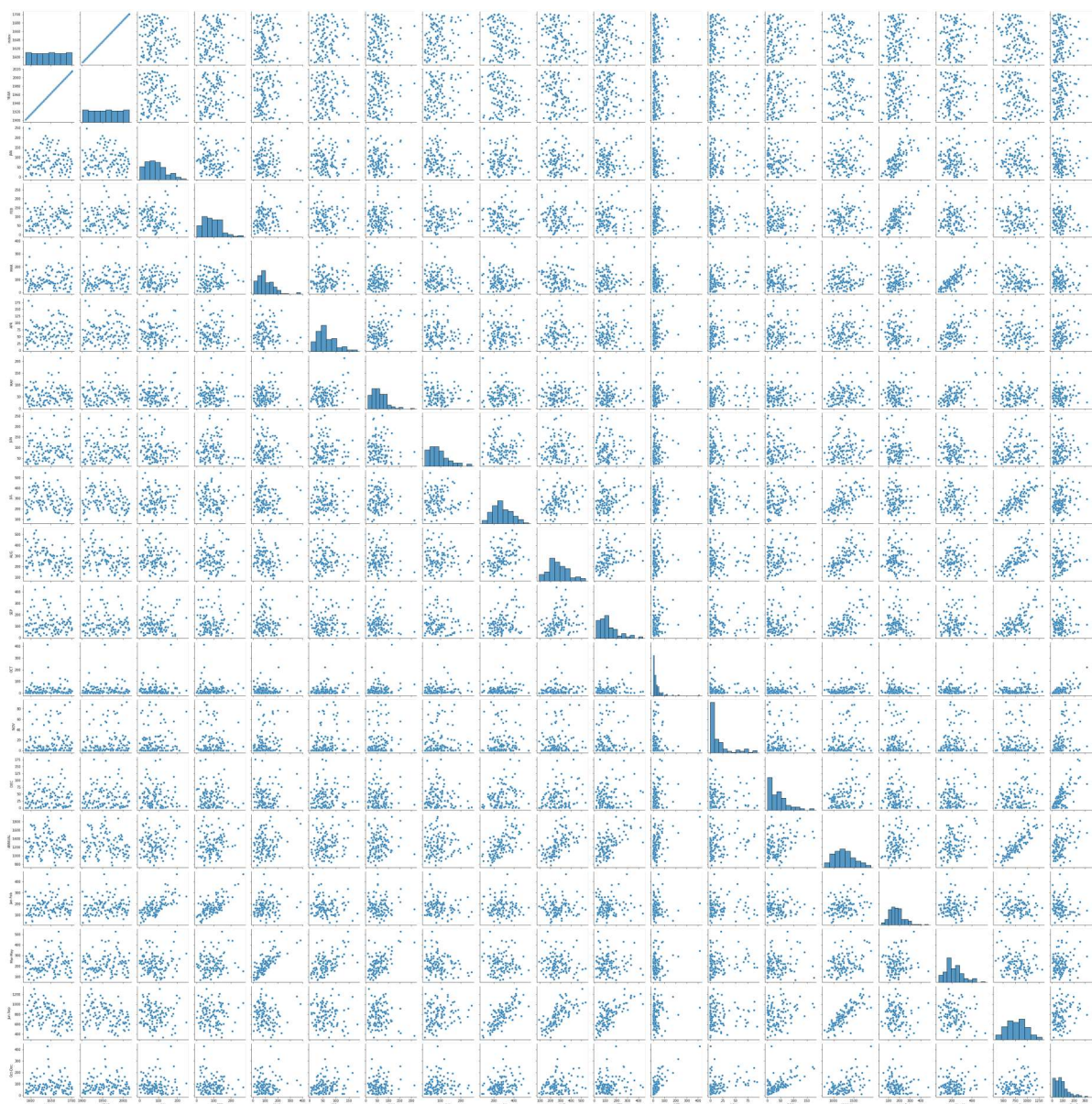
```
In [261]: df.info()
```

```
<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

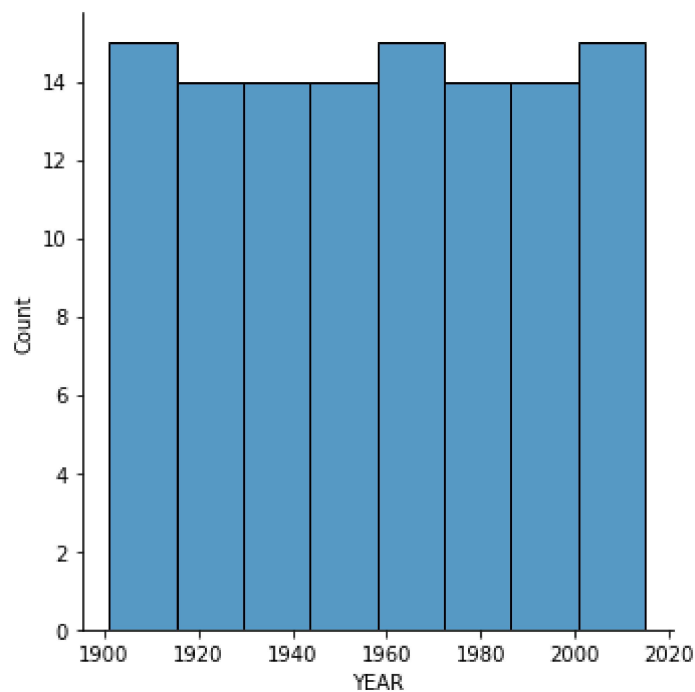
```
In [262]: sns.pairplot(df)
```

```
Out[262]: <seaborn.axisgrid.PairGrid at 0x1f646d70730>
```



```
In [263]: sns.displot(df['YEAR'])
```

```
Out[263]: <seaborn.axisgrid.FacetGrid at 0x1f652241ee0>
```



```
In [264]: sns.heatmap(df.corr())
```

```
Out[264]: <AxesSubplot:>
```

