

Importing Libraries

```
In [1]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

Importing Datasets

```
In [2]: df=pd.read_csv("mm_tripura.csv")
df
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
0	322	NAGA MANI MIZO TRIPURA	1901	11.7	18.1	29.4	206.2	124.0	443.3	331.4	466.0	304.1	166.7	102.4
1	323	NAGA MANI MIZO TRIPURA	1902	4.8	0.5	36.3	297.8	215.5	480.1	392.4	312.8	318.7	102.4	102.4
2	324	NAGA MANI MIZO TRIPURA	1903	6.5	40.5	139.8	45.5	159.9	458.6	300.2	470.6	366.1	166.4	102.4
3	325	NAGA MANI MIZO TRIPURA	1904	2.3	46.9	47.5	290.3	230.5	455.3	423.5	423.6	375.8	128.9	102.4
4	326	NAGA MANI MIZO TRIPURA	1905	9.1	35.3	306.5	161.7	193.6	339.7	450.1	429.9	320.1	246.4	102.4
...
110	432	NAGA MANI MIZO TRIPURA	2011	12.6	3.6	51.4	81.1	334.9	374.2	313.3	367.6	258.3	92.6	102.4
111	433	NAGA MANI MIZO TRIPURA	2012	24.5	10.2	20.3	243.5	163.5	396.2	280.1	342.7	248.7	160.9	102.4
112	434	NAGA MANI MIZO TRIPURA	2013	0.2	5.7	19.7	60.3	348.9	206.6	255.9	291.3	241.4	125.6	102.4
113	435	NAGA MANI MIZO TRIPURA	2014	1.2	21.0	25.4	49.6	192.5	268.3	295.7	372.3	300.9	69.6	102.4
114	436	NAGA MANI MIZO TRIPURA	2015	14.4	14.2	21.6	253.5	198.3	283.9	413.6	334.2	255.9	118.7	102.4

115 rows × 20 columns



Data Cleaning and Data Preprocessing

In [3]:

df=df.dropna()
df

Out[3]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
0	322	NAGA MANI MIZO TRIPURA	1901	11.7	18.1	29.4	206.2	124.0	443.3	331.4	466.0	304.1	166.7	102.4	102.4
1	323	NAGA MANI MIZO TRIPURA	1902	4.8	0.5	36.3	297.8	215.5	480.1	392.4	312.8	318.7	102.4	102.4	102.4
2	324	NAGA MANI MIZO TRIPURA	1903	6.5	40.5	139.8	45.5	159.9	458.6	300.2	470.6	366.1	166.4	102.4	102.4
3	325	NAGA MANI MIZO TRIPURA	1904	2.3	46.9	47.5	290.3	230.5	455.3	423.5	423.6	375.8	128.9	102.4	102.4
4	326	NAGA MANI MIZO TRIPURA	1905	9.1	35.3	306.5	161.7	193.6	339.7	450.1	429.9	320.1	246.4	102.4	102.4
...
110	432	NAGA MANI MIZO TRIPURA	2011	12.6	3.6	51.4	81.1	334.9	374.2	313.3	367.6	258.3	92.6	102.4	102.4
111	433	NAGA MANI MIZO TRIPURA	2012	24.5	10.2	20.3	243.5	163.5	396.2	280.1	342.7	248.7	160.9	102.4	102.4
112	434	NAGA MANI MIZO TRIPURA	2013	0.2	5.7	19.7	60.3	348.9	206.6	255.9	291.3	241.4	125.6	102.4	102.4
113	435	NAGA MANI MIZO TRIPURA	2014	1.2	21.0	25.4	49.6	192.5	268.3	295.7	372.3	300.9	69.6	102.4	102.4
114	436	NAGA MANI MIZO TRIPURA	2015	14.4	14.2	21.6	253.5	198.3	283.9	413.6	334.2	255.9	118.7	102.4	102.4

115 rows × 20 columns

In [4]:

df.columns

Out[4]:

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 [5]:

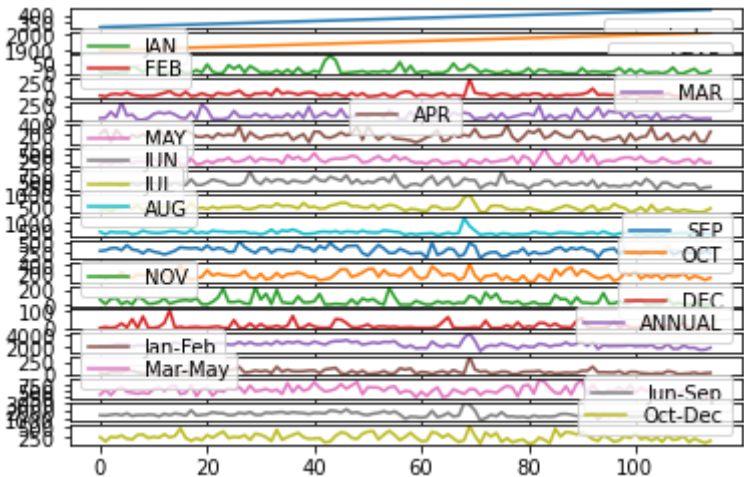
df.info()

```
<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 [6]: df.plot.line(subplots=True)
```

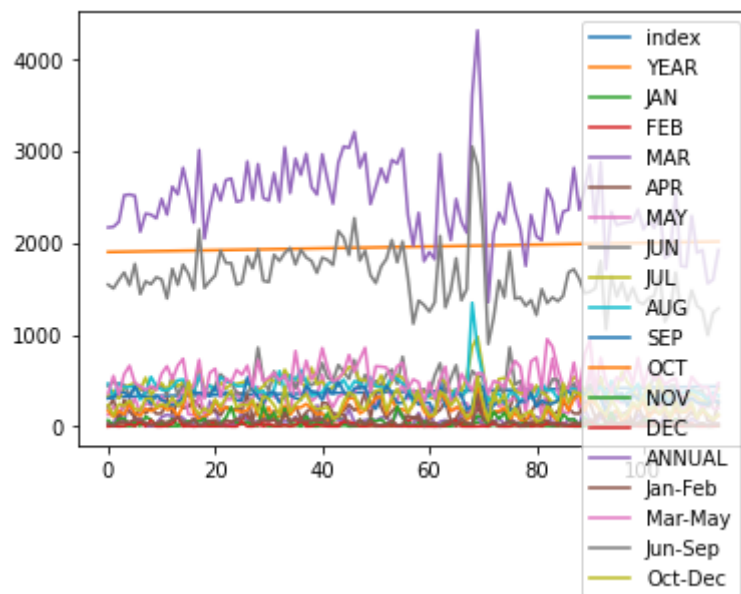
```
Out[6]: array([<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>], dtype=object)
```



Line chart

```
In [7]: df.plot.line()
```

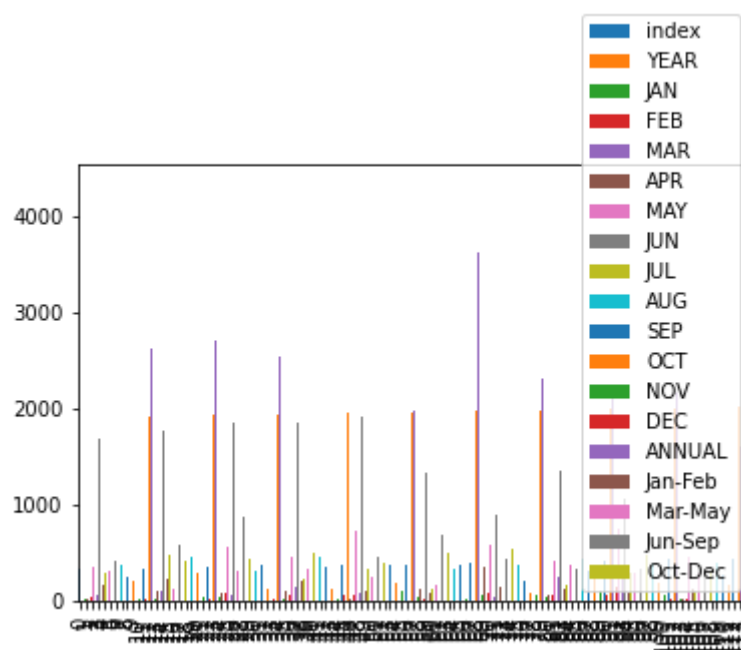
```
Out[7]: <AxesSubplot:~>
```



Bar chart

```
In [8]: df.plot.bar()
```

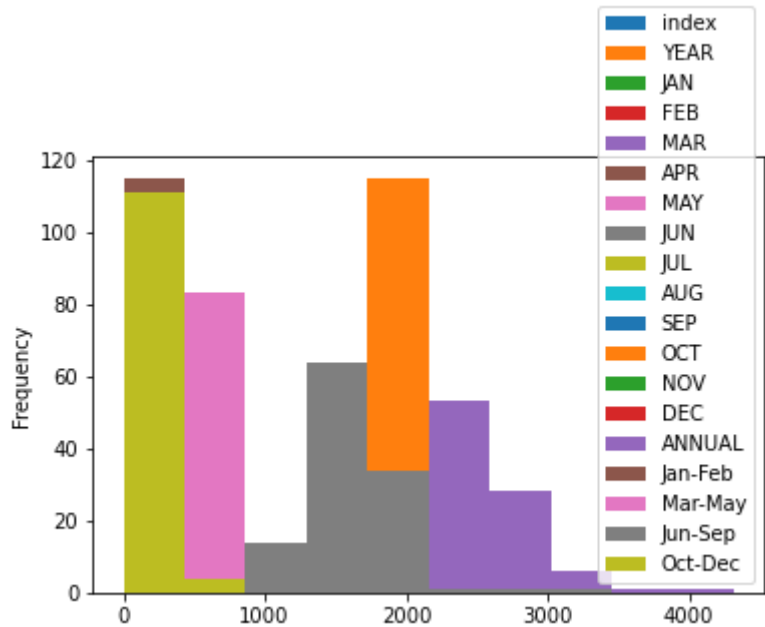
```
Out[8]: <AxesSubplot:>
```



Histogram

```
In [9]: df.plot.hist()
```

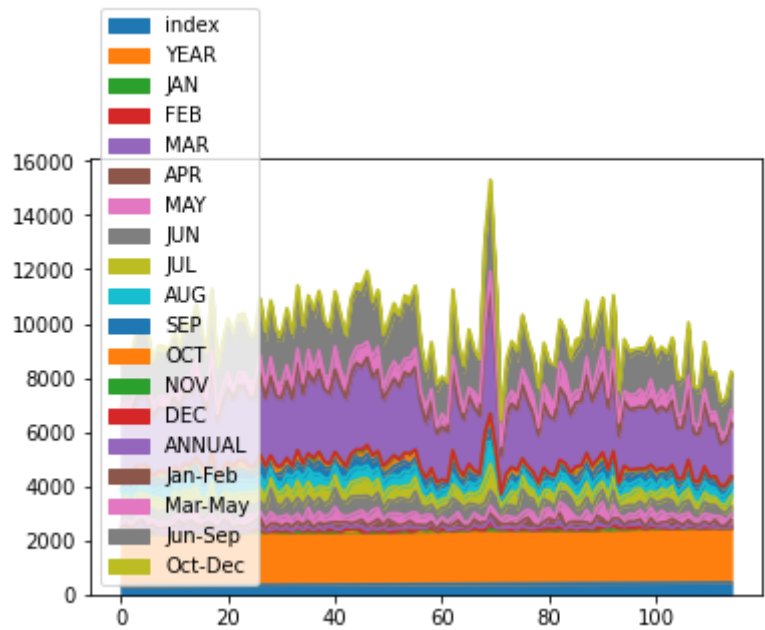
```
Out[9]: <AxesSubplot:ylabel='Frequency'>
```



Area chart

```
In [10]: df.plot.area()
```

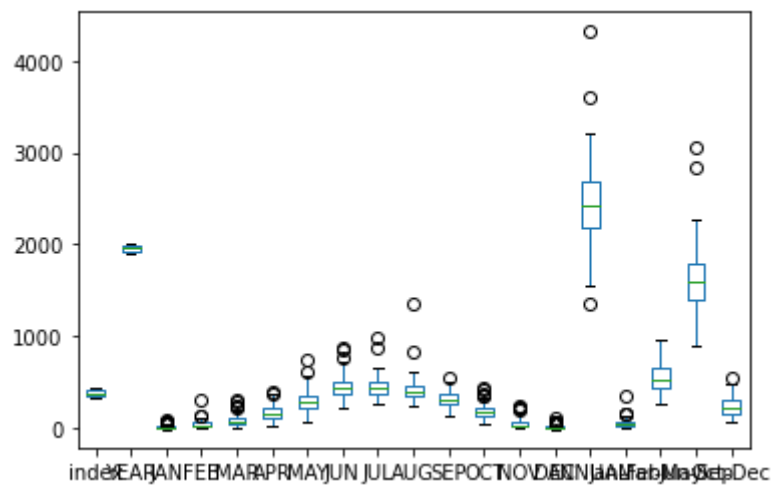
Out[10]: <AxesSubplot:>



Box chart

```
In [11]: df.plot.box()
```

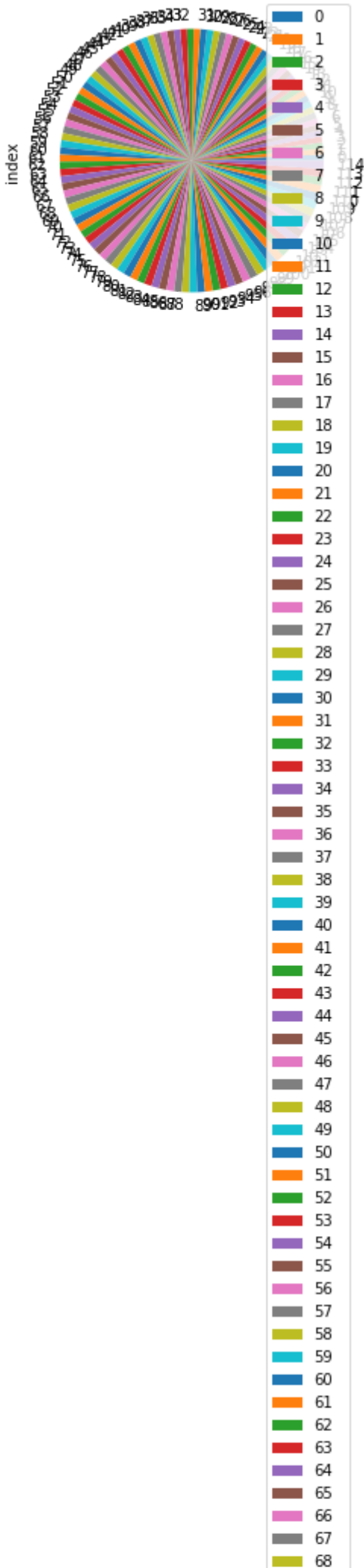
Out[11]: <AxesSubplot:>

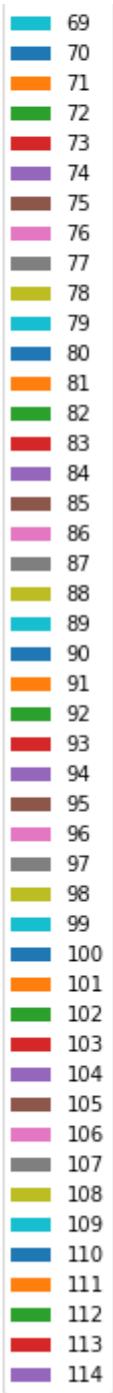


Pie chart

```
In [12]: df.plot.pie(y='index')
```

```
Out[12]: <AxesSubplot:ylabel='index'>
```

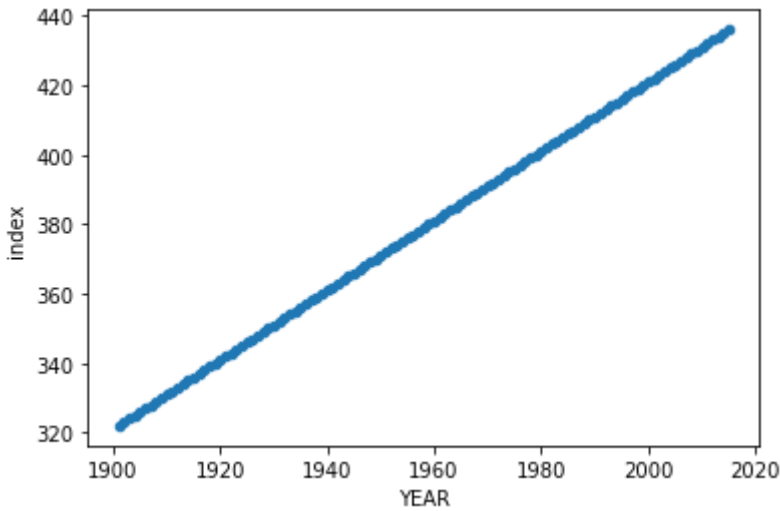




Scatter chart

```
In [13]: df.plot.scatter(x='YEAR' ,y='index')
```

Out[13]: <AxesSubplot:xlabel='YEAR', ylabel='index'>



```
In [14]: 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
```

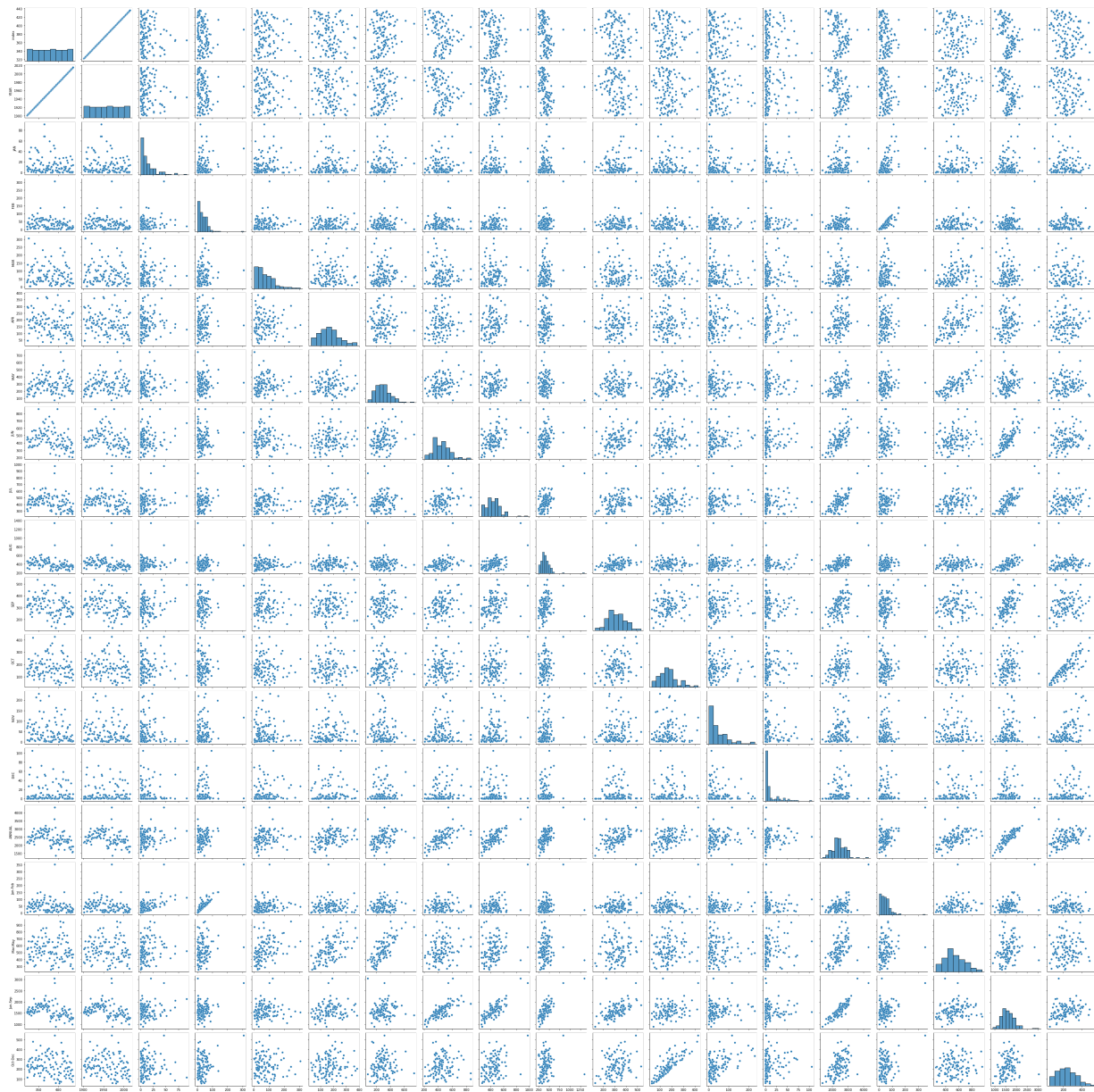
```
In [15]: df.describe()
```

	index	YEAR	JAN	FEB	MAR	APR	MAY	JUN
count	115.000000	115.000000	115.000000	115.000000	115.000000	115.000000	115.000000	115.000000
mean	379.000000	1958.000000	14.025217	36.652174	77.199130	170.733043	290.839130	445.633043
std	33.341666	33.341666	16.627750	37.776876	60.656689	78.559032	112.675514	123.833043
min	322.000000	1901.000000	0.000000	0.000000	3.100000	26.300000	73.500000	206.600000
25%	350.500000	1929.500000	3.150000	11.700000	31.750000	113.750000	210.650000	361.450000
50%	379.000000	1958.000000	7.900000	30.000000	62.700000	161.700000	278.500000	442.600000
75%	407.500000	1986.500000	18.450000	53.300000	105.050000	213.900000	352.300000	511.050000
max	436.000000	2015.000000	91.400000	306.300000	306.500000	383.800000	743.000000	861.100000

EDA AND VISUALIZATION

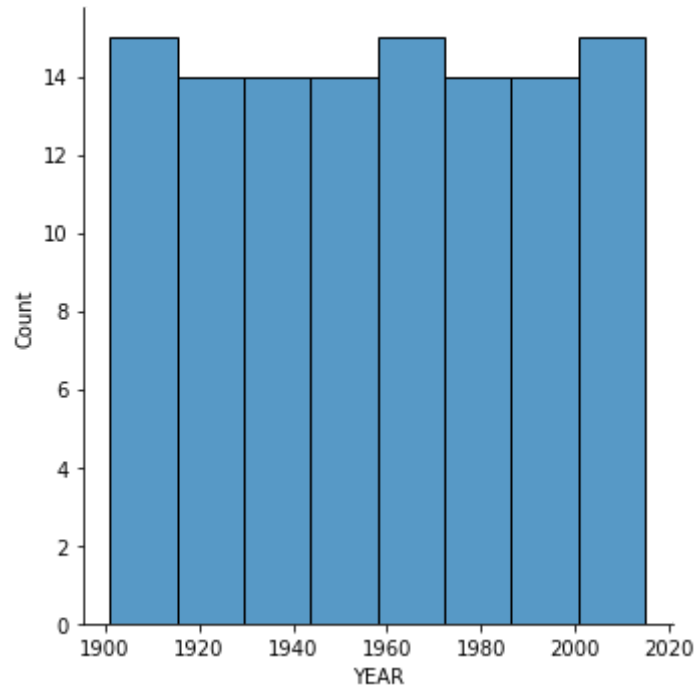
```
In [16]: sns.pairplot(df)
```

```
Out[16]: <seaborn.axisgrid.PairGrid at 0x1b1360c9df0>
```



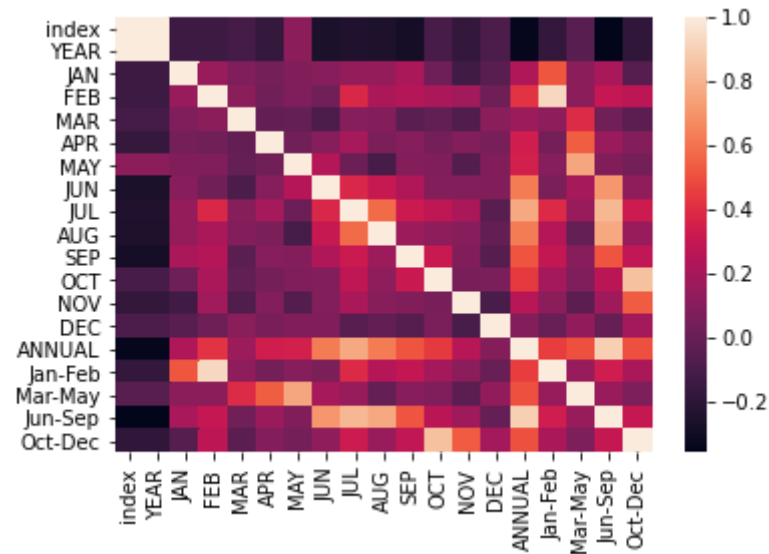
```
In [17]: sns.displot(df['YEAR'])
```

```
Out[17]: <seaborn.axisgrid.FacetGrid at 0x1b1415d08b0>
```



```
In [18]: sns.heatmap(df.corr())
```

Out[18]: <AxesSubplot:>



```
In [ ]:
```