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

In [2]:
```python
data=pd.read_csv(r"C:\Users\user\Desktop\vicky\rainfall\rainfall in india 1901-2015.csv")[3888:400
```

In [3]:
```python
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 3888 to 4001
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: 17.9+ KB
```

In [4]:
```python
data.head()
```

Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNU
3888	3888	KERALA	1902	6.7	2.6	57.3	83.9	134.5	390.9	1205.0	315.8	491.6	358.4	158.3	121.5	332
3889	3889	KERALA	1903	3.2	18.6	3.1	83.6	249.7	558.6	1022.5	420.2	341.8	354.1	157.0	59.0	327
3890	3890	KERALA	1904	23.7	3.0	32.2	71.5	235.7	1098.2	725.5	351.8	222.7	328.1	33.9	3.3	312
3891	3891	KERALA	1905	1.2	22.3	9.4	105.9	263.3	850.2	520.5	293.6	217.2	383.5	74.4	0.2	274
3892	3892	KERALA	1906	26.7	7.4	9.9	59.4	160.8	414.9	954.2	442.8	131.2	251.7	163.1	86.0	270

In [5]: `data.tail()`

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUA
3997	3997	KERALA	2011	20.5	45.7	24.1	165.2	124.2	788.5	536.8	492.7	391.2	227.2	169.7	49.5	3035
3998	3998	KERALA	2012	7.4	11.0	21.0	171.1	95.3	430.3	362.6	501.6	241.1	187.5	112.9	9.4	2151
3999	3999	KERALA	2013	3.9	40.1	49.9	49.3	119.3	1042.7	830.2	369.7	318.6	259.9	154.9	17.0	3255
4000	4000	KERALA	2014	4.6	10.3	17.9	95.7	251.0	454.4	677.8	733.9	298.8	355.5	99.5	47.2	3046
4001	4001	KERALA	2015	3.1	5.8	50.1	214.1	201.8	563.6	406.0	252.2	292.9	308.1	223.6	79.4	2600

In [6]: `data.shape`

Out[6]: `(114, 20)`

In [7]:
```
new_data=data.fillna(value=1)
new_data
```

Out[7]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNU
3888	3888	KERALA	1902	6.7	2.6	57.3	83.9	134.5	390.9	1205.0	315.8	491.6	358.4	158.3	121.5	332
3889	3889	KERALA	1903	3.2	18.6	3.1	83.6	249.7	558.6	1022.5	420.2	341.8	354.1	157.0	59.0	327
3890	3890	KERALA	1904	23.7	3.0	32.2	71.5	235.7	1098.2	725.5	351.8	222.7	328.1	33.9	3.3	312
3891	3891	KERALA	1905	1.2	22.3	9.4	105.9	263.3	850.2	520.5	293.6	217.2	383.5	74.4	0.2	274
3892	3892	KERALA	1906	26.7	7.4	9.9	59.4	160.8	414.9	954.2	442.8	131.2	251.7	163.1	86.0	270
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
3997	3997	KERALA	2011	20.5	45.7	24.1	165.2	124.2	788.5	536.8	492.7	391.2	227.2	169.7	49.5	303
3998	3998	KERALA	2012	7.4	11.0	21.0	171.1	95.3	430.3	362.6	501.6	241.1	187.5	112.9	9.4	215
3999	3999	KERALA	2013	3.9	40.1	49.9	49.3	119.3	1042.7	830.2	369.7	318.6	259.9	154.9	17.0	325
4000	4000	KERALA	2014	4.6	10.3	17.9	95.7	251.0	454.4	677.8	733.9	298.8	355.5	99.5	47.2	304
4001	4001	KERALA	2015	3.1	5.8	50.1	214.1	201.8	563.6	406.0	252.2	292.9	308.1	223.6	79.4	260

114 rows × 20 columns

In [8]: `new_data.index`

Out[8]: `RangeIndex(start=3888, stop=4002, step=1)`

In [9]: `new_data.columns`

Out[9]:
```
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 [10]:
```
new_data.plot.line()
```

Out[10]: <AxesSubplot:>

In [11]:
```
new_data.plot.bar()
```

Out[11]: <AxesSubplot:>

In [12]:
```
new_data.plot.area()
```

Out[12]: <AxesSubplot:>

In [13]:
```
new_data.plot.hist()
```

Out[13]: <AxesSubplot:ylabel='Frequency'>

In [14]: `new_data.plot.pie(y='ANNUAL')`

Out[14]: `<AxesSubplot:ylabel='ANNUAL'>`

In [15]: `new_data.plot.scatter(x='YEAR',y='ANNUAL')`

Out[15]: `<AxesSubplot:xlabel='YEAR', ylabel='ANNUAL'>`

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

Out[16]: `<seaborn.axisgrid.PairGrid at 0x16548ad0df0>`

In [17]:
```python
sns.distplot(data['YEAR'])
```

C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)

Out[17]: <AxesSubplot:xlabel='YEAR', ylabel='Density'>

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

Out[18]: <AxesSubplot:>