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

```python
In [2]:  data=pd.read_csv(r"C:\Users\user\Desktop\vicky\rainfall\rainfall in india 1901-2015.csv")[3313:342
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 3313 to 3426
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
```

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

Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
3313	3313	RAYALSEEMA	1902	10.0	0.2	1.7	11.0	36.8	73.6	41.3	148.3	181.7	188.5	88.9	36.4	818.4
3314	3314	RAYALSEEMA	1903	30.0	0.1	0.0	3.6	80.5	67.5	127.5	140.6	219.7	95.3	289.4	84.0	1138.2
3315	3315	RAYALSEEMA	1904	14.8	0.0	1.7	7.1	58.8	39.8	75.1	19.4	84.7	111.5	4.4	16.1	433.4
3316	3316	RAYALSEEMA	1905	6.5	6.8	17.0	18.3	44.2	66.1	50.9	219.3	36.5	180.2	55.4	2.0	703.4
3317	3317	RAYALSEEMA	1906	115.3	7.2	6.8	2.1	9.6	84.1	127.9	154.4	130.4	107.7	35.0	132.2	912.6

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

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
3422	3422	RAYALSEEMA	2011	0.8	12.1	0.0	34.6	33.0	44.5	128.9	163.6	71.2	107.5	106.9	35.1	738.0
3423	3423	RAYALSEEMA	2012	2.7	0.0	2.5	32.7	38.8	47.0	139.7	120.0	69.5	113.7	86.6	61.9	715.0
3424	3424	RAYALSEEMA	2013	1.3	30.6	11.5	26.8	38.9	73.8	95.7	110.3	163.2	169.3	38.6	2.6	762.6
3425	3425	RAYALSEEMA	2014	0.2	0.7	12.5	5.1	46.7	66.3	68.7	115.1	81.4	104.6	37.8	12.8	551.8
3426	3426	RAYALSEEMA	2015	1.9	0.0	13.4	73.4	39.7	73.0	43.1	123.6	136.3	106.7	383.8	52.2	1047.1

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	ANNUAL
3313	3313	RAYALSEEMA	1902	10.0	0.2	1.7	11.0	36.8	73.6	41.3	148.3	181.7	188.5	88.9	36.4	818.4
3314	3314	RAYALSEEMA	1903	30.0	0.1	0.0	3.6	80.5	67.5	127.5	140.6	219.7	95.3	289.4	84.0	1138.2
3315	3315	RAYALSEEMA	1904	14.8	0.0	1.7	7.1	58.8	39.8	75.1	19.4	84.7	111.5	4.4	16.1	433.4
3316	3316	RAYALSEEMA	1905	6.5	6.8	17.0	18.3	44.2	66.1	50.9	219.3	36.5	180.2	55.4	2.0	703.4
3317	3317	RAYALSEEMA	1906	115.3	7.2	6.8	2.1	9.6	84.1	127.9	154.4	130.4	107.7	35.0	132.2	912.6
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3422	3422	RAYALSEEMA	2011	0.8	12.1	0.0	34.6	33.0	44.5	128.9	163.6	71.2	107.5	106.9	35.1	738.0
3423	3423	RAYALSEEMA	2012	2.7	0.0	2.5	32.7	38.8	47.0	139.7	120.0	69.5	113.7	86.6	61.9	715.0
3424	3424	RAYALSEEMA	2013	1.3	30.6	11.5	26.8	38.9	73.8	95.7	110.3	163.2	169.3	38.6	2.6	762.6
3425	3425	RAYALSEEMA	2014	0.2	0.7	12.5	5.1	46.7	66.3	68.7	115.1	81.4	104.6	37.8	12.8	551.8
3426	3426	RAYALSEEMA	2015	1.9	0.0	13.4	73.4	39.7	73.0	43.1	123.6	136.3	106.7	383.8	52.2	1047.1

114 rows × 20 columns

In [8]: `new_data.index`

Out[8]: RangeIndex(start=3313, stop=3427, 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]:
```python
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 0x1367c5137c0>`

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:>