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

```
In [2]: data=pd.read_csv(r"C:\Users\user\Desktop\vicky\rainfall\rainfall in india 1901-2015.csv")[3428:354:
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 3428 to 3541
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]: data.head()
```

Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
3428	3428	TAMIL NADU	1902	67.2	9.8	25.1	21.9	84.7	39.3	55.1	113.8	98.6	282.2	174.9	165.8	1138.2
3429	3429	TAMIL NADU	1903	19.3	7.8	1.7	18.2	128.5	58.5	72.6	115.0	210.4	128.1	200.5	203.2	1163.9
3430	3430	TAMIL NADU	1904	35.2	0.1	0.7	19.5	121.9	34.9	89.0	40.4	85.7	163.2	23.6	49.1	663.1
3431	3431	TAMIL NADU	1905	6.5	7.5	17.2	64.8	83.7	49.8	39.0	101.8	73.5	250.4	123.7	3.2	821.1
3432	3432	TAMIL NADU	1906	52.4	12.9	17.0	8.5	39.6	43.6	76.0	195.2	65.3	162.8	189.1	127.7	990.2

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

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
3537	3537	TAMIL NADU	2011	4.3	11.2	8.0	91.5	33.4	56.0	45.5	128.9	76.0	200.4	230.5	41.0	926.5
3538	3538	TAMIL NADU	2012	3.0	0.1	2.5	35.5	41.9	30.1	46.5	98.0	84.9	235.2	44.5	14.0	636.1
3539	3539	TAMIL NADU	2013	3.9	30.9	30.0	20.3	42.0	54.6	42.7	110.7	113.5	127.9	112.3	53.2	741.9
3540	3540	TAMIL NADU	2014	7.4	6.1	8.1	8.3	139.1	47.8	50.6	117.7	98.9	252.2	110.8	66.0	913.0
3541	3541	TAMIL NADU	2015	8.3	2.3	21.7	108.8	112.4	62.4	43.5	81.6	98.4	132.6	379.8	152.8	1204.6

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
3428	3428	TAMIL NADU	1902	67.2	9.8	25.1	21.9	84.7	39.3	55.1	113.8	98.6	282.2	174.9	165.8	1138.2
3429	3429	TAMIL NADU	1903	19.3	7.8	1.7	18.2	128.5	58.5	72.6	115.0	210.4	128.1	200.5	203.2	1163.9
3430	3430	TAMIL NADU	1904	35.2	0.1	0.7	19.5	121.9	34.9	89.0	40.4	85.7	163.2	23.6	49.1	663.1
3431	3431	TAMIL NADU	1905	6.5	7.5	17.2	64.8	83.7	49.8	39.0	101.8	73.5	250.4	123.7	3.2	821.1
3432	3432	TAMIL NADU	1906	52.4	12.9	17.0	8.5	39.6	43.6	76.0	195.2	65.3	162.8	189.1	127.7	990.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3537	3537	TAMIL NADU	2011	4.3	11.2	8.0	91.5	33.4	56.0	45.5	128.9	76.0	200.4	230.5	41.0	926.5
3538	3538	TAMIL NADU	2012	3.0	0.1	2.5	35.5	41.9	30.1	46.5	98.0	84.9	235.2	44.5	14.0	636.1
3539	3539	TAMIL NADU	2013	3.9	30.9	30.0	20.3	42.0	54.6	42.7	110.7	113.5	127.9	112.3	53.2	741.9
3540	3540	TAMIL NADU	2014	7.4	6.1	8.1	8.3	139.1	47.8	50.6	117.7	98.9	252.2	110.8	66.0	913.0
3541	3541	TAMIL NADU	2015	8.3	2.3	21.7	108.8	112.4	62.4	43.5	81.6	98.4	132.6	379.8	152.8	1204.6

114 rows × 20 columns

In [8]: `new_data.index`

Out[8]: `RangeIndex(start=3428, stop=3542, 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]:
```python
new_data.plot.line()
```

Out[10]: <AxesSubplot:>

In [11]:
```python
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]:
```python
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]:
```python
sns.pairplot(new_data)
```

Out[16]: <seaborn.axisgrid.PairGrid at 0x1db87552130>

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

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

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

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

Out[18]: `<AxesSubplot:>`