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")[3658:377
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 3658 to 3771
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
3658	3658	NORTH INTERIOR KARNATAKA	1902	0.0	0.0	0.3	22.5	34.4	111.3	83.2	78.1	146.7	118.8	35.7	85.1	716.1
3659	3659	NORTH INTERIOR KARNATAKA	1903	3.5	0.0	0.1	6.9	53.4	102.8	209.4	146.4	189.3	166.4	34.3	16.0	928.5
3660	3660	NORTH INTERIOR KARNATAKA	1904	0.2	0.3	8.5	11.0	46.3	120.6	91.6	48.5	165.1	86.5	0.0	0.0	578.6
3661	3661	NORTH INTERIOR KARNATAKA	1905	0.0	6.0	2.6	16.0	51.2	99.6	60.1	139.2	42.2	85.0	4.4	0.0	506.2
3662	3662	NORTH INTERIOR KARNATAKA	1906	21.3	0.0	0.2	2.6	30.0	142.0	120.3	182.1	116.0	86.2	24.3	58.9	783.8

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

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
3767	3767	NORTH INTERIOR KARNATAKA	2011	0.5	7.2	7.2	41.2	46.8	101.3	150.8	152.0	69.0	73.4	5.7	0.0	655.2
3768	3768	NORTH INTERIOR KARNATAKA	2012	28.5	6.2	0.4	35.4	19.5	60.0	114.5	105.5	79.2	85.2	46.5	2.9	583.8
3769	3769	NORTH INTERIOR KARNATAKA	2013	1.2	6.1	3.0	25.4	47.4	99.4	160.7	73.9	201.0	101.0	4.2	0.1	723.2
3770	3770	NORTH INTERIOR KARNATAKA	2014	0.0	6.1	29.2	26.4	93.0	50.4	136.8	205.2	90.2	80.3	25.0	14.1	756.8
3771	3771	NORTH INTERIOR KARNATAKA	2015	2.4	0.0	27.5	50.8	45.3	89.6	38.5	78.4	150.8	61.2	5.7	1.7	551.9

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
3658	3658	NORTH INTERIOR KARNATAKA	1902	0.0	0.0	0.3	22.5	34.4	111.3	83.2	78.1	146.7	118.8	35.7	85.1	716.1
3659	3659	NORTH INTERIOR KARNATAKA	1903	3.5	0.0	0.1	6.9	53.4	102.8	209.4	146.4	189.3	166.4	34.3	16.0	928.5
3660	3660	NORTH INTERIOR KARNATAKA	1904	0.2	0.3	8.5	11.0	46.3	120.6	91.6	48.5	165.1	86.5	0.0	0.0	578.6
3661	3661	NORTH INTERIOR KARNATAKA	1905	0.0	6.0	2.6	16.0	51.2	99.6	60.1	139.2	42.2	85.0	4.4	0.0	506.2
3662	3662	NORTH INTERIOR KARNATAKA	1906	21.3	0.0	0.2	2.6	30.0	142.0	120.3	182.1	116.0	86.2	24.3	58.9	783.8
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3767	3767	NORTH INTERIOR KARNATAKA	2011	0.5	7.2	7.2	41.2	46.8	101.3	150.8	152.0	69.0	73.4	5.7	0.0	655.2
3768	3768	NORTH INTERIOR KARNATAKA	2012	28.5	6.2	0.4	35.4	19.5	60.0	114.5	105.5	79.2	85.2	46.5	2.9	583.8
3769	3769	NORTH INTERIOR KARNATAKA	2013	1.2	6.1	3.0	25.4	47.4	99.4	160.7	73.9	201.0	101.0	4.2	0.1	723.2
3770	3770	NORTH INTERIOR KARNATAKA	2014	0.0	6.1	29.2	26.4	93.0	50.4	136.8	205.2	90.2	80.3	25.0	14.1	756.8
3771	3771	NORTH INTERIOR KARNATAKA	2015	2.4	0.0	27.5	50.8	45.3	89.6	38.5	78.4	150.8	61.2	5.7	1.7	551.9

114 rows × 20 columns

In [8]:
```
new_data.index
```

Out[8]:
```
RangeIndex(start=3658, stop=3772, 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 0x26fc2fd78e0>
```

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:>
```