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")[3543:365
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 115 entries, 3543 to 3657
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          114 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       114 non-null    float64
 16  Jan-Feb      114 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: 18.1+ KB
```

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

Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUA
3543	3543	COASTAL KARNATAKA	1902	3.2	0.3	4.9	10.2	54.6	698.4	1401.6	454.2	708.4	180.4	50.8	132.2	3699
3544	3544	COASTAL KARNATAKA	1903	0.7	0.0	0.0	4.1	202.8	536.5	1405.5	593.8	304.4	185.0	79.3	5.3	3317
3545	3545	COASTAL KARNATAKA	1904	2.4	0.0	4.8	23.7	93.2	1108.2	1070.0	465.6	245.3	127.2	0.7	0.0	3141
3546	3546	COASTAL KARNATAKA	1905	0.0	0.2	0.0	6.4	83.1	767.3	777.3	586.9	172.9	222.2	36.1	0.0	2652
3547	3547	COASTAL KARNATAKA	1906	23.0	0.0	0.0	0.5	29.8	593.6	1173.4	535.0	273.3	128.5	22.1	56.8	2836

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

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNU
3653	3653	COASTAL KARNATAKA	2012	NaN	11.4	5.1	77.0	22.9	650.9	754.6	1027.6	382.0	115.1	68.0	3.6	N
3654	3654	COASTAL KARNATAKA	2013	2.4	19.6	19.0	28.5	100.4	1153.0	1515.3	680.2	379.1	265.1	56.9	10.0	422
3655	3655	COASTAL KARNATAKA	2014	0.0	0.3	1.9	40.5	181.9	507.0	1155.4	1121.0	379.3	226.4	40.0	30.8	368
3656	3656	COASTAL KARNATAKA	2015	1.4	1.0	32.3	72.2	150.3	735.3	930.9	575.2	260.3	208.5	124.2	14.3	310
3657	3657	NORTH INTERIOR KARNATAKA	1901	3.5	18.8	7.1	67.2	65.5	120.5	151.9	115.1	128.8	80.0	13.6	0.9	77

In [6]:
```
data.shape
```

Out[6]: (115, 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
3543	3543	COASTAL KARNATAKA	1902	3.2	0.3	4.9	10.2	54.6	698.4	1401.6	454.2	708.4	180.4	50.8	132.2	369
3544	3544	COASTAL KARNATAKA	1903	0.7	0.0	0.0	4.1	202.8	536.5	1405.5	593.8	304.4	185.0	79.3	5.3	33
3545	3545	COASTAL KARNATAKA	1904	2.4	0.0	4.8	23.7	93.2	1108.2	1070.0	465.6	245.3	127.2	0.7	0.0	31
3546	3546	COASTAL KARNATAKA	1905	0.0	0.2	0.0	6.4	83.1	767.3	777.3	586.9	172.9	222.2	36.1	0.0	26
3547	3547	COASTAL KARNATAKA	1906	23.0	0.0	0.0	0.5	29.8	593.6	1173.4	535.0	273.3	128.5	22.1	56.8	28
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3653	3653	COASTAL KARNATAKA	2012	1.0	11.4	5.1	77.0	22.9	650.9	754.6	1027.6	382.0	115.1	68.0	3.6	
3654	3654	COASTAL KARNATAKA	2013	2.4	19.6	19.0	28.5	100.4	1153.0	1515.3	680.2	379.1	265.1	56.9	10.0	42
3655	3655	COASTAL KARNATAKA	2014	0.0	0.3	1.9	40.5	181.9	507.0	1155.4	1121.0	379.3	226.4	40.0	30.8	36
3656	3656	COASTAL KARNATAKA	2015	1.4	1.0	32.3	72.2	150.3	735.3	930.9	575.2	260.3	208.5	124.2	14.3	31
3657	3657	NORTH INTERIOR KARNATAKA	1901	3.5	18.8	7.1	67.2	65.5	120.5	151.9	115.1	128.8	80.0	13.6	0.9	7

115 rows × 20 columns

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

Out[8]: RangeIndex(start=3543, stop=3658, 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 0x1c33d41afd0>
```

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:>
```