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")[1819:193
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 1819 to 1932
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 | Jan Fe |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1819** | 1819 | WEST RAJASTHAN | 1903 | 1.7 | 1.3 | 5.5 | 0.0 | 4.2 | 2.7 | 154.8 | 87.1 | 49.3 | 0.1 | 0.0 | 0.5 | 307.0 | 3 |
| **1820** | 1820 | WEST RAJASTHAN | 1904 | 3.8 | 2.9 | 16.3 | 0.7 | 11.4 | 14.6 | 39.8 | 45.6 | 21.4 | 1.4 | 2.9 | 7.1 | 167.9 | 6 |
| **1821** | 1821 | WEST RAJASTHAN | 1905 | 6.3 | 4.8 | 0.7 | 1.3 | 0.3 | 4.9 | 30.1 | 0.6 | 64.5 | 0.0 | 0.0 | 0.9 | 114.4 | 11 |
| **1822** | 1822 | WEST RAJASTHAN | 1906 | 0.1 | 39.1 | 7.8 | 0.0 | 0.2 | 19.9 | 53.7 | 58.6 | 87.1 | 0.7 | 0.0 | 2.4 | 269.5 | 39 |
| **1823** | 1823 | WEST RAJASTHAN | 1907 | 0.6 | 24.5 | 6.3 | 7.3 | 6.8 | 12.8 | 52.2 | 242.5 | 0.2 | 0.1 | 0.0 | 0.0 | 353.2 | 25 |

In [5]:
```
data.shape
```

Out[5]:
```
(114, 20)
```

In [6]:
```
new_data=data.fillna(value=1)
new_data
```

Out[6]:

| | index | SUBDIVISION | YEAR | JAN | FEB | MAR | APR | MAY | JUN | JUL | AUG | SEP | OCT | NOV | DEC | ANNUAL | Jan Fe |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1819** | 1819 | WEST RAJASTHAN | 1903 | 1.7 | 1.3 | 5.5 | 0.0 | 4.2 | 2.7 | 154.8 | 87.1 | 49.3 | 0.1 | 0.0 | 0.5 | 307.0 | 3 |
| **1820** | 1820 | WEST RAJASTHAN | 1904 | 3.8 | 2.9 | 16.3 | 0.7 | 11.4 | 14.6 | 39.8 | 45.6 | 21.4 | 1.4 | 2.9 | 7.1 | 167.9 | 6 |
| **1821** | 1821 | WEST RAJASTHAN | 1905 | 6.3 | 4.8 | 0.7 | 1.3 | 0.3 | 4.9 | 30.1 | 0.6 | 64.5 | 0.0 | 0.0 | 0.9 | 114.4 | 11 |
| **1822** | 1822 | WEST RAJASTHAN | 1906 | 0.1 | 39.1 | 7.8 | 0.0 | 0.2 | 19.9 | 53.7 | 58.6 | 87.1 | 0.7 | 0.0 | 2.4 | 269.5 | 39 |
| **1823** | 1823 | WEST RAJASTHAN | 1907 | 0.6 | 24.5 | 6.3 | 7.3 | 6.8 | 12.8 | 52.2 | 242.5 | 0.2 | 0.1 | 0.0 | 0.0 | 353.2 | 25 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **1928** | 1928 | WEST RAJASTHAN | 2012 | 0.5 | 0.0 | 0.0 | 9.5 | 10.4 | 5.3 | 40.4 | 166.7 | 92.0 | 1.9 | 0.0 | 0.6 | 327.3 | 0 |
| **1929** | 1929 | WEST RAJASTHAN | 2013 | 8.6 | 21.8 | 4.2 | 3.1 | 1.7 | 37.6 | 104.5 | 138.2 | 58.7 | 10.1 | 1.0 | 0.0 | 389.4 | 30 |
| **1930** | 1930 | WEST RAJASTHAN | 2014 | 0.8 | 2.2 | 4.7 | 8.4 | 23.0 | 13.8 | 94.3 | 69.6 | 84.9 | 0.5 | 0.2 | 0.0 | 302.4 | 3 |
| **1931** | 1931 | WEST RAJASTHAN | 2015 | 1.4 | 0.9 | 30.3 | 25.2 | 15.5 | 53.2 | 234.6 | 60.5 | 35.7 | 1.1 | 0.1 | 0.0 | 458.4 | 2 |
| **1932** | 1932 | EAST RAJASTHAN | 1901 | 21.6 | 8.9 | 2.9 | 0.7 | 5.0 | 15.0 | 164.8 | 175.6 | 7.5 | 9.8 | 0.0 | 0.8 | 412.5 | 30 |

114 rows × 20 columns

In [7]:
```
new_data.index
```

Out[7]:
```
RangeIndex(start=1819, stop=1933, step=1)
```

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

Out[8]:
```
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 [9]:
```
new_data.plot.line()
```

Out[9]: <AxesSubplot:>

In [10]:
```
new_data.plot.bar()
```

Out[10]: <AxesSubplot:>

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

Out[11]:
```
<AxesSubplot:>
```

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

Out[12]:
```
<AxesSubplot:ylabel='Frequency'>
```

In [13]:
```
new_data.plot.pie(y='ANNUAL')
```

Out[13]:
```
<AxesSubplot:ylabel='ANNUAL'>
```

In [14]:
```
new_data.plot.scatter(x='YEAR',y='ANNUAL')
```

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

In [15]:
```
sns.pairplot(new_data)
```

Out[15]:
```
<seaborn.axisgrid.PairGrid at 0x1fb5fc72df0>
```

In [16]:
```
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[16]:
```
<AxesSubplot:xlabel='YEAR', ylabel='Density'>
```

In [17]:
```
sns.heatmap(new_data.corr())
```

Out[17]:
```
<AxesSubplot:>
```