

```
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")[2049:2162]
```

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

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 2049 to 2162
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             113 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          113 non-null   float64 
16  Jan-Feb         113 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 |
|------|-------|---------------------|------|-----|------|------|------|------|-------|-------|-------|-------|------|-----|-----|--------|
| 2049 | 2049  | WEST MADHYA PRADESH | 1903 | 5.3 | 0.0  | 0.0  | 0.0  | 22.3 | 50.6  | 304.9 | 261.1 | 250.2 | 55.1 | 0.0 | 0.0 | 949.6  |
| 2050 | 2050  | WEST MADHYA PRADESH | 1904 | 3.2 | 15.5 | 14.8 | 0.0  | 12.0 | 96.6  | 273.0 | 218.6 | 125.9 | 3.3  | 1.8 | 9.6 | 774.4  |
| 2051 | 2051  | WEST MADHYA PRADESH | 1905 | 3.5 | 4.4  | 1.1  | 0.8  | 3.0  | 36.1  | 326.3 | 137.6 | 183.5 | 0.3  | 0.0 | 0.0 | 696.5  |
| 2052 | 2052  | WEST MADHYA PRADESH | 1906 | 0.0 | 11.0 | 6.8  | 0.0  | 0.5  | 180.0 | 344.5 | 198.6 | 266.2 | 1.5  | 0.2 | 0.9 | 1010.2 |
| 2053 | 2053  | WEST MADHYA PRADESH | 1907 | 5.2 | 25.1 | 0.6  | 12.3 | 2.8  | 48.7  | 202.2 | 328.5 | 17.3  | 0.0  | 7.8 | 0.0 | 650.4  |

```
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 |  |
|------|-------|---------------------------|------|------|------|------|------|------|-------|-------|-------|-------|------|-----|------|--------|--|
| 2049 | 2049  | WEST<br>MADHYA<br>PRADESH | 1903 | 5.3  | 0.0  | 0.0  | 0.0  | 22.3 | 50.6  | 304.9 | 261.1 | 250.2 | 55.1 | 0.0 | 0.0  | 949.6  |  |
| 2050 | 2050  | WEST<br>MADHYA<br>PRADESH | 1904 | 3.2  | 15.5 | 14.8 | 0.0  | 12.0 | 96.6  | 273.0 | 218.6 | 125.9 | 3.3  | 1.8 | 9.6  | 774.4  |  |
| 2051 | 2051  | WEST<br>MADHYA<br>PRADESH | 1905 | 3.5  | 4.4  | 1.1  | 0.8  | 3.0  | 36.1  | 326.3 | 137.6 | 183.5 | 0.3  | 0.0 | 0.0  | 696.5  |  |
| 2052 | 2052  | WEST<br>MADHYA<br>PRADESH | 1906 | 0.0  | 11.0 | 6.8  | 0.0  | 0.5  | 180.0 | 344.5 | 198.6 | 266.2 | 1.5  | 0.2 | 0.9  | 1010.2 |  |
| 2053 | 2053  | WEST<br>MADHYA<br>PRADESH | 1907 | 5.2  | 25.1 | 0.6  | 12.3 | 2.8  | 48.7  | 202.2 | 328.5 | 17.3  | 0.0  | 7.8 | 0.0  | 650.4  |  |
| ...  | ...   | ...                       | ...  | ...  | ...  | ...  | ...  | ...  | ...   | ...   | ...   | ...   | ...  | ... | ...  | ...    |  |
| 2158 | 2158  | WEST<br>MADHYA<br>PRADESH | 2012 | 6.2  | 0.0  | 0.0  | 0.9  | 3.1  | 48.2  | 439.2 | 341.2 | 194.3 | 2.1  | 0.0 | 0.0  | 1035.2 |  |
| 2159 | 2159  | WEST<br>MADHYA<br>PRADESH | 2013 | 1.7  | 31.1 | 8.5  | 2.8  | 0.4  | 263.7 | 485.1 | 432.6 | 98.9  | 68.7 | 0.3 | 2.4  | 1396.3 |  |
| 2160 | 2160  | WEST<br>MADHYA<br>PRADESH | 2014 | 25.6 | 34.4 | 4.6  | 1.4  | 1.4  | 30.6  | 337.4 | 211.0 | 192.6 | 7.0  | 3.0 | 15.8 | 864.9  |  |
| 2161 | 2161  | WEST<br>MADHYA<br>PRADESH | 2015 | 40.2 | 6.4  | 53.5 | 13.3 | 2.0  | 154.1 | 428.2 | 276.6 | 55.6  | 11.0 | 0.3 | 1.0  | 1042.3 |  |
| 2162 | 2162  | EAST<br>MADHYA<br>PRADESH | 1901 | 48.5 | 38.1 | 15.7 | 10.7 | 6.2  | 61.0  | 367.5 | 589.2 | 189.9 | 5.9  | 0.0 | 0.0  | 1332.7 |  |

114 rows × 20 columns

In [7]:

new\_data.index

Out[7]:

RangeIndex(start=2049, stop=2163, 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 0x20a5f250df0>
```

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

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

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