

```
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\JHARKHAND.csv")
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 115 entries, 0 to 114
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                   115 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                115 non-null   float64
16  Jan-Feb               115 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
0	782	JHARKHAND	1901	92.7	66.6	11.1	18.4	33.5	70.9	269.4	415.1	248.0	37.3
1	783	JHARKHAND	1902	4.2	7.7	13.2	28.5	59.8	89.9	456.1	204.9	306.6	17.6
2	784	JHARKHAND	1903	25.1	19.5	10.7	32.8	56.4	142.1	206.1	280.8	190.2	210.1
3	785	JHARKHAND	1904	2.5	17.0	38.1	9.1	116.1	308.9	494.1	336.1	125.6	30.6
4	786	JHARKHAND	1905	38.4	53.3	61.6	32.9	66.2	41.5	420.3	293.7	322.8	21.3

In [5]: `data.shape`

Out[5]: (115, 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
0	782	JHARKHAND	1901	92.7	66.6	11.1	18.4	33.5	70.9	269.4	415.1	248.0	37.3
1	783	JHARKHAND	1902	4.2	7.7	13.2	28.5	59.8	89.9	456.1	204.9	306.6	17.6
2	784	JHARKHAND	1903	25.1	19.5	10.7	32.8	56.4	142.1	206.1	280.8	190.2	210.1
3	785	JHARKHAND	1904	2.5	17.0	38.1	9.1	116.1	308.9	494.1	336.1	125.6	30.6
4	786	JHARKHAND	1905	38.4	53.3	61.6	32.9	66.2	41.5	420.3	293.7	322.8	21.3
...
110	892	JHARKHAND	2011	3.3	2.5	6.4	25.4	55.0	349.0	181.8	403.2	324.6	23.3
111	893	JHARKHAND	2012	34.6	10.3	1.5	9.6	6.6	121.1	287.2	282.4	217.6	37.8
112	894	JHARKHAND	2013	1.1	17.9	1.6	22.3	85.0	181.5	211.1	278.1	173.8	281.1
113	895	JHARKHAND	2014	9.9	47.5	22.9	1.9	98.2	139.7	321.3	290.9	178.2	44.9
114	896	JHARKHAND	2015	12.2	2.6	21.6	55.5	25.5	183.3	429.7	240.7	85.1	22.7

115 rows × 20 columns

In [7]: `new_data.index`

Out[7]: RangeIndex(start=0, stop=115, 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 0x1cfa0048880>
```

```
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:>
```