

```
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")[3083:3196]
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 3083 to 3196
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
3083	3083	COASTAL ANDHRA PRADESH	1902	2.0	0.0	2.8	23.9	37.6	72.6	144.5	236.1	204.5	262.0	50.4	27.1	1063.6
3084	3084	COASTAL ANDHRA PRADESH	1903	0.8	13.3	0.2	6.2	73.4	154.0	248.6	258.0	216.5	159.1	173.9	12.1	1316.2
3085	3085	COASTAL ANDHRA PRADESH	1904	1.3	0.0	5.4	3.0	136.3	107.8	120.2	117.7	116.8	240.9	0.0	10.7	860.2
3086	3086	COASTAL ANDHRA PRADESH	1905	1.1	16.7	68.0	37.0	68.8	84.4	64.6	210.8	170.2	66.0	7.4	0.0	795.2
3087	3087	COASTAL ANDHRA PRADESH	1906	3.9	23.5	9.9	2.3	11.0	252.6	155.8	241.1	126.9	92.1	21.5	80.8	1021.5

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

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
3192	3192	COASTAL ANDHRA PRADESH	2011	0.0	17.9	0.9	62.3	67.9	86.8	196.0	215.8	129.7	74.6	4.9	5.0	861.9
3193	3193	COASTAL ANDHRA PRADESH	2012	37.6	0.0	2.7	24.0	39.3	95.4	221.9	221.2	246.5	140.0	289.7	0.0	1318.4
3194	3194	COASTAL ANDHRA PRADESH	2013	2.0	29.6	0.2	48.0	28.2	127.5	162.4	123.1	132.0	411.5	53.1	2.8	1120.5
3195	3195	COASTAL ANDHRA PRADESH	2014	0.4	1.2	9.1	6.0	112.9	45.7	151.8	177.8	144.5	195.6	23.7	6.4	874.9
3196	3196	COASTAL ANDHRA PRADESH	2015	2.0	0.6	5.5	32.3	34.1	283.8	116.0	192.0	201.8	59.7	81.2	2.0	1010.9

```
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
3083	3083	COASTAL ANDHRA PRADESH	1902	2.0	0.0	2.8	23.9	37.6	72.6	144.5	236.1	204.5	262.0	50.4	27.1	1063.6
3084	3084	COASTAL ANDHRA PRADESH	1903	0.8	13.3	0.2	6.2	73.4	154.0	248.6	258.0	216.5	159.1	173.9	12.1	1316.2
3085	3085	COASTAL ANDHRA PRADESH	1904	1.3	0.0	5.4	3.0	136.3	107.8	120.2	117.7	116.8	240.9	0.0	10.7	860.2
3086	3086	COASTAL ANDHRA PRADESH	1905	1.1	16.7	68.0	37.0	68.8	84.4	64.6	210.8	170.2	66.0	7.4	0.0	795.2
3087	3087	COASTAL ANDHRA PRADESH	1906	3.9	23.5	9.9	2.3	11.0	252.6	155.8	241.1	126.9	92.1	21.5	80.8	1021.5
...
3192	3192	COASTAL ANDHRA PRADESH	2011	0.0	17.9	0.9	62.3	67.9	86.8	196.0	215.8	129.7	74.6	4.9	5.0	861.9
3193	3193	COASTAL ANDHRA PRADESH	2012	37.6	0.0	2.7	24.0	39.3	95.4	221.9	221.2	246.5	140.0	289.7	0.0	1318.4
3194	3194	COASTAL ANDHRA PRADESH	2013	2.0	29.6	0.2	48.0	28.2	127.5	162.4	123.1	132.0	411.5	53.1	2.8	1120.5
3195	3195	COASTAL ANDHRA PRADESH	2014	0.4	1.2	9.1	6.0	112.9	45.7	151.8	177.8	144.5	195.6	23.7	6.4	874.9
3196	3196	COASTAL ANDHRA PRADESH	2015	2.0	0.6	5.5	32.3	34.1	283.8	116.0	192.0	201.8	59.7	81.2	2.0	1010.9

114 rows × 20 columns

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

Out[8]: RangeIndex(start=3083, stop=3197, 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 0x21870454760>
```

```
In [17]: 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[17]: <AxesSubplot:xlabel='YEAR', ylabel='Density'>
```

```
In [18]: sns.heatmap(new_data.corr())
```

```
Out[18]: <AxesSubplot:>
```