

```
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")[1014:1126]
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 113 entries, 1014 to 1126
Data columns (total 20 columns):
#   Column          Non-Null Count  Dtype
---  -
0   index           113 non-null    int64
1   SUBDIVISION     113 non-null    object
2   YEAR            113 non-null    int64
3   JAN             113 non-null    float64
4   FEB             113 non-null    float64
5   MAR             113 non-null    float64
6   APR             113 non-null    float64
7   MAY             113 non-null    float64
8   JUN             113 non-null    float64
9   JUL             113 non-null    float64
10  AUG             113 non-null    float64
11  SEP             113 non-null    float64
12  OCT             113 non-null    float64
13  NOV             113 non-null    float64
14  DEC             113 non-null    float64
15  ANNUAL          113 non-null    float64
16  Jan-Feb         113 non-null    float64
17  Mar-May         113 non-null    float64
18  Jun-Sep         113 non-null    float64
19  Oct-Dec         113 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 17.8+ KB
```

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

Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
1014	1014	EAST UTTAR PRADESH	1903	8.2	0.4	1.3	0.7	15.3	71.6	115.3	420.2	258.7	324.7	0.0	0.0	1216.4
1015	1015	EAST UTTAR PRADESH	1904	7.3	1.5	8.3	0.4	28.7	148.0	359.4	328.8	95.0	50.6	17.0	26.3	1071.2
1016	1016	EAST UTTAR PRADESH	1905	16.8	23.6	20.0	5.4	15.4	17.3	302.4	316.2	169.5	3.3	0.0	1.6	891.6
1017	1017	EAST UTTAR PRADESH	1906	5.7	58.2	6.7	0.0	13.5	140.6	341.2	290.9	135.5	5.1	0.0	0.1	997.7
1018	1018	EAST UTTAR PRADESH	1907	2.1	74.9	19.2	20.3	11.1	48.6	173.5	290.9	25.7	0.0	0.0	0.1	666.3

In [5]: data.tail()

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
1122	1122	EAST UTTAR PRADESH	2011	1.0	2.7	1.6	2.9	32.2	163.8	197.9	232.1	146.4	0.6	0.0	0.0	781.2
1123	1123	EAST UTTAR PRADESH	2012	20.3	1.2	3.4	2.8	0.2	18.5	234.2	156.0	164.4	0.7	0.3	0.7	602.7
1124	1124	EAST UTTAR PRADESH	2013	6.1	59.6	2.7	8.7	1.1	309.7	230.0	246.1	78.2	97.4	0.5	1.1	1041.4
1125	1125	EAST UTTAR PRADESH	2014	47.4	25.8	15.4	1.7	10.7	47.8	224.5	138.1	106.7	74.7	0.0	8.4	701.2
1126	1126	EAST UTTAR PRADESH	2015	30.0	4.1	48.2	23.2	8.6	95.3	179.0	175.8	21.9	11.8	0.5	4.9	603.3

In [6]: data.shape

Out[6]: (113, 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
1014	1014	EAST UTTAR PRADESH	1903	8.2	0.4	1.3	0.7	15.3	71.6	115.3	420.2	258.7	324.7	0.0	0.0	1216.4
1015	1015	EAST UTTAR PRADESH	1904	7.3	1.5	8.3	0.4	28.7	148.0	359.4	328.8	95.0	50.6	17.0	26.3	1071.2
1016	1016	EAST UTTAR PRADESH	1905	16.8	23.6	20.0	5.4	15.4	17.3	302.4	316.2	169.5	3.3	0.0	1.6	891.6
1017	1017	EAST UTTAR PRADESH	1906	5.7	58.2	6.7	0.0	13.5	140.6	341.2	290.9	135.5	5.1	0.0	0.1	997.7
1018	1018	EAST UTTAR PRADESH	1907	2.1	74.9	19.2	20.3	11.1	48.6	173.5	290.9	25.7	0.0	0.0	0.1	666.3
...
1122	1122	EAST UTTAR PRADESH	2011	1.0	2.7	1.6	2.9	32.2	163.8	197.9	232.1	146.4	0.6	0.0	0.0	781.2
1123	1123	EAST UTTAR PRADESH	2012	20.3	1.2	3.4	2.8	0.2	18.5	234.2	156.0	164.4	0.7	0.3	0.7	602.7
1124	1124	EAST UTTAR PRADESH	2013	6.1	59.6	2.7	8.7	1.1	309.7	230.0	246.1	78.2	97.4	0.5	1.1	1041.4
1125	1125	EAST UTTAR PRADESH	2014	47.4	25.8	15.4	1.7	10.7	47.8	224.5	138.1	106.7	74.7	0.0	8.4	701.2
1126	1126	EAST UTTAR PRADESH	2015	30.0	4.1	48.2	23.2	8.6	95.3	179.0	175.8	21.9	11.8	0.5	4.9	603.3

113 rows × 20 columns

In [8]: new_data.index

Out[8]: RangeIndex(start=1014, stop=1127, 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='JAN')
```

```
Out[14]: <AxesSubplot:ylabel='JAN'>
```



```
In [15]: new_data.plot.box()
```

```
Out[15]: <AxesSubplot:>
```



```
In [16]: new_data.plot.scatter(x='JAN',y='FEB')
```

```
Out[16]: <AxesSubplot:xlabel='JAN', ylabel='FEB'>
```

```
In [17]: sns.pairplot(new_data)
```

```
Out[17]: <seaborn.axisgrid.PairGrid at 0x29e81e32400>
```

```
In [18]: 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[18]: <AxesSubplot:xlabel='YEAR', ylabel='Density'>
```

```
In [19]: sns.heatmap(new_data.corr())
```

```
Out[19]: <AxesSubplot:>
```