```python
In [1]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
```

```python
In [2]: data=pd.read_csv(r"C:\Users\user\Desktop\vicky\rainfall\rainfall in india 1901-2015.csv"
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4116 entries, 0 to 4115
Data columns (total 20 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   index        4116 non-null   int64
 1   SUBDIVISION  4116 non-null   object
 2   YEAR         4116 non-null   int64
 3   JAN          4112 non-null   float64
 4   FEB          4113 non-null   float64
 5   MAR          4110 non-null   float64
 6   APR          4112 non-null   float64
 7   MAY          4113 non-null   float64
 8   JUN          4111 non-null   float64
 9   JUL          4109 non-null   float64
 10  AUG          4112 non-null   float64
 11  SEP          4110 non-null   float64
 12  OCT          4109 non-null   float64
 13  NOV          4105 non-null   float64
 14  DEC          4106 non-null   float64
 15  ANNUAL       4090 non-null   float64
 16  Jan-Feb      4110 non-null   float64
 17  Mar-May      4107 non-null   float64
 18  Jun-Sep      4106 non-null   float64
 19  Oct-Dec      4103 non-null   float64
dtypes: float64(17), int64(2), object(1)
memory usage: 643.2+ KB
```

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

Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DE(
0	0	ANDAMAN & NICOBAR ISLANDS	1901	49.2	87.1	29.2	2.3	528.8	517.5	365.1	481.1	332.6	388.5	558.2	33.
1	1	ANDAMAN & NICOBAR ISLANDS	1902	0.0	159.8	12.2	0.0	446.1	537.1	228.9	753.7	666.2	197.2	359.0	160.
2	2	ANDAMAN & NICOBAR ISLANDS	1903	12.7	144.0	0.0	1.0	235.1	479.9	728.4	326.7	339.0	181.2	284.4	225.
3	3	ANDAMAN & NICOBAR ISLANDS	1904	9.4	14.7	0.0	202.4	304.5	495.1	502.0	160.1	820.4	222.2	308.7	40.
4	4	ANDAMAN & NICOBAR ISLANDS	1905	1.3	0.0	3.3	26.9	279.5	628.7	368.7	330.5	297.0	260.7	25.4	344.

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

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
4111	4111	LAKSHADWEEP	2011	5.1	2.8	3.1	85.9	107.2	153.6	350.2	254.0	255.2	117.4	184.3
4112	4112	LAKSHADWEEP	2012	19.2	0.1	1.6	76.8	21.2	327.0	231.5	381.2	179.8	145.9	12.4
4113	4113	LAKSHADWEEP	2013	26.2	34.4	37.5	5.3	88.3	426.2	296.4	154.4	180.0	72.8	78.1
4114	4114	LAKSHADWEEP	2014	53.2	16.1	4.4	14.9	57.4	244.1	116.1	466.1	132.2	169.2	59.0
4115	4115	LAKSHADWEEP	2015	2.2	0.5	3.7	87.1	133.1	296.6	257.5	146.4	160.4	165.4	231.0

In [6]: `data.shape`

Out[6]: (4116, 20)

In [7]:
```python
new_data=data.fillna(value=1)
new_data
```

Out[7]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
0	0	ANDAMAN & NICOBAR ISLANDS	1901	49.2	87.1	29.2	2.3	528.8	517.5	365.1	481.1	332.6	388.5	558.2
1	1	ANDAMAN & NICOBAR ISLANDS	1902	0.0	159.8	12.2	0.0	446.1	537.1	228.9	753.7	666.2	197.2	359.0
2	2	ANDAMAN & NICOBAR ISLANDS	1903	12.7	144.0	0.0	1.0	235.1	479.9	728.4	326.7	339.0	181.2	284.4
3	3	ANDAMAN & NICOBAR ISLANDS	1904	9.4	14.7	0.0	202.4	304.5	495.1	502.0	160.1	820.4	222.2	308.7
4	4	ANDAMAN & NICOBAR ISLANDS	1905	1.3	0.0	3.3	26.9	279.5	628.7	368.7	330.5	297.0	260.7	25.4
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4111	4111	LAKSHADWEEP	2011	5.1	2.8	3.1	85.9	107.2	153.6	350.2	254.0	255.2	117.4	184.3
4112	4112	LAKSHADWEEP	2012	19.2	0.1	1.6	76.8	21.2	327.0	231.5	381.2	179.8	145.9	12.4
4113	4113	LAKSHADWEEP	2013	26.2	34.4	37.5	5.3	88.3	426.2	296.4	154.4	180.0	72.8	78.1
4114	4114	LAKSHADWEEP	2014	53.2	16.1	4.4	14.9	57.4	244.1	116.1	466.1	132.2	169.2	59.0
4115	4115	LAKSHADWEEP	2015	2.2	0.5	3.7	87.1	133.1	296.6	257.5	146.4	160.4	165.4	231.0

4116 rows × 20 columns
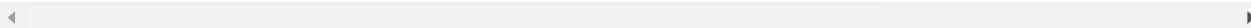
In [8]:
```python
new_data.index
```

Out[8]:
```
RangeIndex(start=0, stop=4116, step=1)
```

In [9]:
```python
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]:
```python
new_data.plot.line()
```

Out[10]: <AxesSubplot:>

In [11]:
```python
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 [20]: new_data.plot.pie(y='JAN')
```

```
In [24]: new_data.plot.box()
```

Out[24]: <AxesSubplot:>

In [25]:
```python
new_data.plot.scatter(x='JAN',y='FEB')
```

Out[25]: <AxesSubplot:xlabel='JAN', ylabel='FEB'>

In [26]:   `sns.pairplot(new_data)`

Out[26]:   `<seaborn.axisgrid.PairGrid at 0x19f9bdd6f70>`

In [28]: `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[28]: `<AxesSubplot:xlabel='YEAR', ylabel='Density'>`

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

Out[18]: `<AxesSubplot:>`