

```
In [1]: import pandas as pd
```

```
In [2]: data=pd.read_csv("/home/placement/Downloads/rainfall in india 1901-2015.csv")
```

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

```
Out[3]:
```

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	
count	4116.000000	4112.000000	4113.000000	4110.000000	4112.000000	4113.000000	4111.000000	4109.000000	4112.000000	4110.000000	4109.0
mean	1958.218659	18.957320	21.805325	27.359197	43.127432	85.745417	230.234444	347.214334	290.263497	197.361922	95.1
std	33.140898	33.585371	35.909488	46.959424	67.831168	123.234904	234.710758	269.539667	188.770477	135.408345	99.1
min	1901.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.400000	0.000000	0.000000	0.100000	0.0
25%	1930.000000	0.600000	0.600000	1.000000	3.000000	8.600000	70.350000	175.600000	155.975000	100.525000	14.0
50%	1958.000000	6.000000	6.700000	7.800000	15.700000	36.600000	138.700000	284.800000	259.400000	173.900000	65.1
75%	1987.000000	22.200000	26.800000	31.300000	49.950000	97.200000	305.150000	418.400000	377.800000	265.800000	148.4
max	2015.000000	583.700000	403.500000	605.600000	595.100000	1168.600000	1609.900000	2362.800000	1664.600000	1222.000000	948.1

In [4]: data.head(10)

Out[4]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
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	33.6	3373.2	136.3	560.3	1696.3	980.3
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	160.5	3520.7	159.8	458.3	2185.9	716.7
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	225.0	2957.4	156.7	236.1	1874.0	690.6
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	40.1	3079.6	24.1	506.9	1977.6	571.0
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	344.7	2566.7	1.3	309.7	1624.9	630.8
5	ANDAMAN & NICOBAR ISLANDS	1906	36.6	0.0	0.0	0.0	556.1	733.3	247.7	320.5	164.3	267.8	128.9	79.2	2534.4	36.6	556.1	1465.8	475.9
6	ANDAMAN & NICOBAR ISLANDS	1907	110.7	0.0	113.3	21.6	616.3	305.2	443.9	377.6	200.4	264.4	648.9	245.6	3347.9	110.7	751.2	1327.1	1158.9
7	ANDAMAN & NICOBAR ISLANDS	1908	20.9	85.1	0.0	29.0	562.0	693.6	481.4	699.9	428.8	170.7	208.1	196.9	3576.4	106.0	591.0	2303.7	575.7
8	ANDAMAN & NICOBAR ISLANDS	1910	26.6	22.7	206.3	89.3	224.5	472.7	264.3	337.4	626.6	208.2	267.3	153.5	2899.4	49.3	520.1	1701.0	629.0
9	ANDAMAN & NICOBAR ISLANDS	1911	0.0	8.4	0.0	122.5	327.3	649.0	253.0	187.1	464.5	333.8	94.5	247.1	2687.2	8.4	449.8	1553.6	675.4

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4116 entries, 0 to 4115
Data columns (total 19 columns):
#   Column          Non-Null Count  Dtype
---  -
0   SUBDIVISION     4116 non-null   object
1   YEAR            4116 non-null   int64
2   JAN             4112 non-null   float64
3   FEB             4113 non-null   float64
4   MAR             4110 non-null   float64
5   APR             4112 non-null   float64
6   MAY             4113 non-null   float64
7   JUN             4111 non-null   float64
8   JUL             4109 non-null   float64
9   AUG             4112 non-null   float64
10  SEP             4110 non-null   float64
11  OCT             4109 non-null   float64
12  NOV             4105 non-null   float64
13  DEC             4106 non-null   float64
14  ANNUAL          4090 non-null   float64
15  Jan-Feb        4110 non-null   float64
16  Mar-May        4107 non-null   float64
17  Jun-Sep        4106 non-null   float64
18  Oct-Dec        4103 non-null   float64
dtypes: float64(17), int64(1), object(1)
memory usage: 611.1+ KB
```

```
In [6]: data.isna().sum()
```

```
Out[6]: SUBDIVISION      0  
        YEAR            0  
        JAN             4  
        FEB             3  
        MAR             6  
        APR             4  
        MAY             3  
        JUN             5  
        JUL             7  
        AUG             4  
        SEP             6  
        OCT             7  
        NOV            11  
        DEC            10  
        ANNUAL          26  
        Jan-Feb         6  
        Mar-May         9  
        Jun-Sep         10  
        Oct-Dec         13  
        dtype: int64
```

```
In [7]: data1=data.groupby(['SUBDIVISION']).count()
data1
```

```
Out[7]:
```

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
SUBDIVISION																		
ANDAMAN & NICOBAR ISLANDS	110	110	110	108	108	109	108	108	108	107	108	108	107	104	110	107	107	107
ARUNACHAL PRADESH	97	96	96	95	97	97	96	96	97	97	95	95	95	91	96	95	95	94
ASSAM & MEGHALAYA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
BIHAR	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
CHHATTISGARH	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
COASTAL ANDHRA PRADESH	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
COASTAL KARNATAKA	115	114	115	115	115	115	115	115	115	115	115	115	115	114	114	115	115	115
EAST MADHYA PRADESH	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
EAST RAJASTHAN	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
EAST UTTAR PRADESH	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
GANGETIC WEST BENGAL	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
GUJARAT REGION	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
HARYANA DELHI & CHANDIGARH	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
HIMACHAL PRADESH	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
JAMMU & KASHMIR	115	115	115	115	115	115	115	114	115	115	115	114	114	114	115	115	114	114
JHARKHAND	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
KERALA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
KONKAN & GOA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
LAKSHADWEEP	114	112	113	112	112	112	112	111	112	111	111	108	110	103	111	110	110	108
MADHYA MAHARASHTRA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
MATATHWADA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
SUBDIVISION																		
NAGA MANI MIZO TRIPURA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
NORTH INTERIOR KARNATAKA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
ORISSA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
PUNJAB	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
RAYALSEEMA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
SAURASHTRA & KUTCH	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
SOUTH INTERIOR KARNATAKA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
SUB HIMALAYAN WEST BENGAL & SIKKIM	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
TAMIL NADU	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
TELANGANA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
UTTARAKHAND	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
VIDARBHA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
WEST MADHYA PRADESH	115	115	114	115	115	115	115	115	115	115	115	115	115	114	114	115	115	115
WEST RAJASTHAN	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115
WEST UTTAR PRADESH	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115

In []:

```
In [8]: data.isna().sum()
```

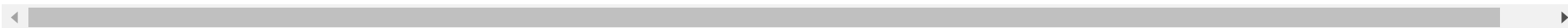
```
Out[8]: SUBDIVISION      0  
        YEAR            0  
        JAN             4  
        FEB             3  
        MAR             6  
        APR             4  
        MAY             3  
        JUN             5  
        JUL             7  
        AUG             4  
        SEP             6  
        OCT             7  
        NOV            11  
        DEC            10  
        ANNUAL          26  
        Jan-Feb         6  
        Mar-May         9  
        Jun-Sep        10  
        Oct-Dec        13  
        dtype: int64
```

```
In [9]: data=data.loc[(data.YEAR<2010)]
data
```

Out[9]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	
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	33.6	3373.2	136.3	560.3	1696.3	9
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	160.5	3520.7	159.8	458.3	2185.9	7
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	225.0	2957.4	156.7	236.1	1874.0	6
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	40.1	3079.6	24.1	506.9	1977.6	5
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	344.7	2566.7	1.3	309.7	1624.9	6
...
4105	LAKSHADWEEP	2005	17.6	11.1	0.0	37.0	92.8	248.5	378.9	102.4	278.0	164.2	218.3	26.6	1575.4	28.7	129.8	1007.8	4
4106	LAKSHADWEEP	2006	20.1	0.0	33.0	0.3	327.9	286.9	172.3	150.7	318.5	119.1	158.9	10.9	1598.6	20.1	361.2	928.4	2
4107	LAKSHADWEEP	2007	2.5	4.2	0.2	22.2	166.2	573.4	427.4	294.7	457.5	256.1	47.6	109.6	2361.6	6.7	188.6	1753.0	4
4108	LAKSHADWEEP	2008	5.5	19.8	120.7	15.8	180.4	254.6	363.9	206.6	108.9	252.9	67.6	130.1	1726.8	25.3	316.9	934.0	4
4109	LAKSHADWEEP	2009	4.7	1.5	0.1	18.1	162.1	401.2	266.4	185.0	145.1	87.4	166.2	132.3	1570.1	6.2	180.3	997.7	3

3900 rows × 19 columns



In [10]: `data.tail(5)`

Out[10]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
4105	LAKSHADWEEP	2005	17.6	11.1	0.0	37.0	92.8	248.5	378.9	102.4	278.0	164.2	218.3	26.6	1575.4	28.7	129.8	1007.8	409.
4106	LAKSHADWEEP	2006	20.1	0.0	33.0	0.3	327.9	286.9	172.3	150.7	318.5	119.1	158.9	10.9	1598.6	20.1	361.2	928.4	288.
4107	LAKSHADWEEP	2007	2.5	4.2	0.2	22.2	166.2	573.4	427.4	294.7	457.5	256.1	47.6	109.6	2361.6	6.7	188.6	1753.0	413.
4108	LAKSHADWEEP	2008	5.5	19.8	120.7	15.8	180.4	254.6	363.9	206.6	108.9	252.9	67.6	130.1	1726.8	25.3	316.9	934.0	450.
4109	LAKSHADWEEP	2009	4.7	1.5	0.1	18.1	162.1	401.2	266.4	185.0	145.1	87.4	166.2	132.3	1570.1	6.2	180.3	997.7	385.

In [11]: `data=data.drop(['ANNUAL', 'Jan-Feb', 'Mar-May', 'Jun-Sep', 'Oct-Dec'],axis=1)`
`data`

Out[11]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
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	33.6
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	160.5
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	225.0
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	40.1
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	344.7
...
4105	LAKSHADWEEP	2005	17.6	11.1	0.0	37.0	92.8	248.5	378.9	102.4	278.0	164.2	218.3	26.6
4106	LAKSHADWEEP	2006	20.1	0.0	33.0	0.3	327.9	286.9	172.3	150.7	318.5	119.1	158.9	10.9
4107	LAKSHADWEEP	2007	2.5	4.2	0.2	22.2	166.2	573.4	427.4	294.7	457.5	256.1	47.6	109.6
4108	LAKSHADWEEP	2008	5.5	19.8	120.7	15.8	180.4	254.6	363.9	206.6	108.9	252.9	67.6	130.1
4109	LAKSHADWEEP	2009	4.7	1.5	0.1	18.1	162.1	401.2	266.4	185.0	145.1	87.4	166.2	132.3

3900 rows × 14 columns

```
In [12]: data['SUBDIVISION'].unique()
```

```
Out[12]: array(['ANDAMAN & NICOBAR ISLANDS', 'ARUNACHAL PRADESH',  
                'ASSAM & MEGHALAYA', 'NAGA MANI MIZO TRIPURA',  
                'SUB HIMALAYAN WEST BENGAL & SIKKIM', 'GANGETIC WEST BENGAL',  
                'ORISSA', 'JHARKHAND', 'BIHAR', 'EAST UTTAR PRADESH',  
                'WEST UTTAR PRADESH', 'UTTARAKHAND', 'HARYANA DELHI & CHANDIGARH',  
                'PUNJAB', 'HIMACHAL PRADESH', 'JAMMU & KASHMIR', 'WEST RAJASTHAN',  
                'EAST RAJASTHAN', 'WEST MADHYA PRADESH', 'EAST MADHYA PRADESH',  
                'GUJARAT REGION', 'SAURASHTRA & KUTCH', 'KONKAN & GOA',  
                'MADHYA MAHARASHTRA', 'MATATHWADA', 'VIDARBHA', 'CHHATTISGARH',  
                'COASTAL ANDHRA PRADESH', 'TELANGANA', 'RAYALSEEMA', 'TAMIL NADU',  
                'COASTAL KARNATAKA', 'NORTH INTERIOR KARNATAKA',  
                'SOUTH INTERIOR KARNATAKA', 'KERALA', 'LAKSHADWEEP'], dtype=object)
```

```
In [13]: [(data.SUBDIVISION=="ARUNACHAL PRADESH")]
```

```
Out[13]: [0      False  
          1      False  
          2      False  
          3      False  
          4      False  
          ...  
         4105     False  
         4106     False  
         4107     False  
         4108     False  
         4109     False  
          Name: SUBDIVISION, Length: 3900, dtype: bool]
```

```
In [14]: data.isna().sum()
```

```
Out[14]: SUBDIVISION    0  
        YEAR          0  
        JAN           3  
        FEB           3  
        MAR           6  
        APR           4  
        MAY           3  
        JUN           5  
        JUL           7  
        AUG           4  
        SEP           6  
        OCT           7  
        NOV          11  
        DEC           10  
        dtype: int64
```

```
In [15]: data['ANNUAL RAIN']=data.apply(lambda row: row.JAN + row.FEB,axis=1)
data
```

Out[15]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN
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	33.6	136.3
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	160.5	159.8
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	225.0	156.7
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	40.1	24.1
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	344.7	1.3
...
4105	LAKSHADWEEP	2005	17.6	11.1	0.0	37.0	92.8	248.5	378.9	102.4	278.0	164.2	218.3	26.6	28.7
4106	LAKSHADWEEP	2006	20.1	0.0	33.0	0.3	327.9	286.9	172.3	150.7	318.5	119.1	158.9	10.9	20.1
4107	LAKSHADWEEP	2007	2.5	4.2	0.2	22.2	166.2	573.4	427.4	294.7	457.5	256.1	47.6	109.6	6.7
4108	LAKSHADWEEP	2008	5.5	19.8	120.7	15.8	180.4	254.6	363.9	206.6	108.9	252.9	67.6	130.1	25.3
4109	LAKSHADWEEP	2009	4.7	1.5	0.1	18.1	162.1	401.2	266.4	185.0	145.1	87.4	166.2	132.3	6.2

3900 rows × 15 columns

```
In [16]: cor=data.corr()  
cor
```

/tmp/ipykernel_6057/4173678507.py:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_only to silence this warning.

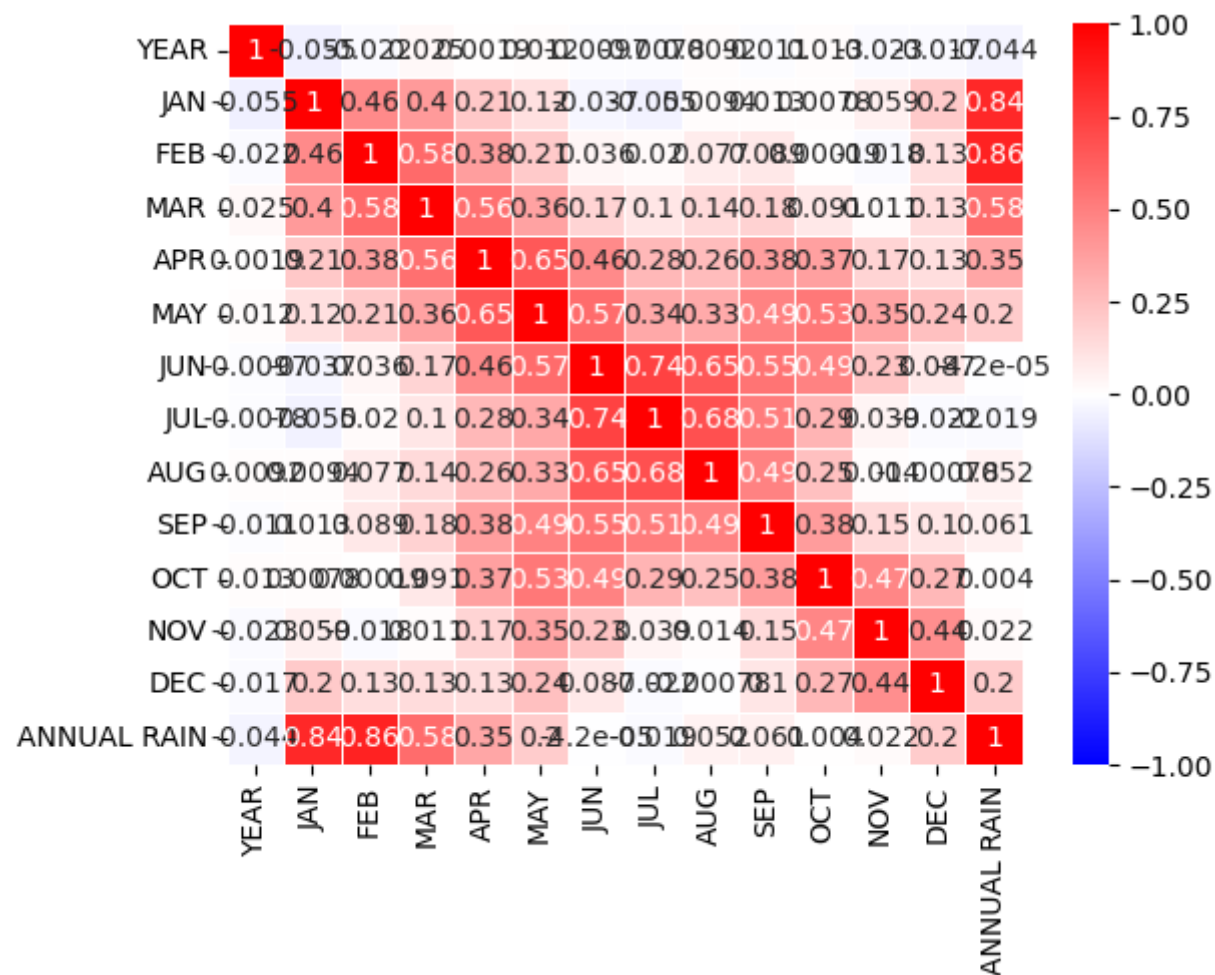
```
cor=data.corr()
```

Out[16]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
YEAR	1.000000	-0.054943	-0.021814	0.024663	0.001913	0.011656	-0.009702	-0.007850	0.009194	-0.011112	0.012884	-0.022893	-0.016688
JAN	-0.054943	1.000000	0.455875	0.396914	0.212725	0.124277	-0.037364	-0.054524	0.009423	0.013421	0.007817	0.059402	0.204835
FEB	-0.021814	0.455875	1.000000	0.582844	0.380745	0.209002	0.035763	0.020171	0.076897	0.088929	0.000194	-0.018420	0.132406
MAR	0.024663	0.396914	0.582844	1.000000	0.557110	0.362911	0.166572	0.100786	0.142062	0.178439	0.090809	0.010802	0.133862
APR	0.001913	0.212725	0.380745	0.557110	1.000000	0.654300	0.458790	0.276126	0.259701	0.381239	0.374406	0.165955	0.132704
MAY	0.011656	0.124277	0.209002	0.362911	0.654300	1.000000	0.571162	0.337946	0.329925	0.489130	0.531166	0.353077	0.242401
JUN	-0.009702	-0.037364	0.035763	0.166572	0.458790	0.571162	1.000000	0.740511	0.653670	0.546800	0.488370	0.228539	0.086661
JUL	-0.007850	-0.054524	0.020171	0.100786	0.276126	0.337946	0.740511	1.000000	0.682643	0.509115	0.294710	0.039455	-0.021688
AUG	0.009194	0.009423	0.076897	0.142062	0.259701	0.329925	0.653670	0.682643	1.000000	0.491390	0.248026	0.014371	-0.000775
SEP	-0.011112	0.013421	0.088929	0.178439	0.381239	0.489130	0.546800	0.509115	0.491390	1.000000	0.382860	0.146776	0.100946
OCT	0.012884	0.007817	0.000194	0.090809	0.374406	0.531166	0.488370	0.294710	0.248026	0.382860	1.000000	0.471935	0.273467
NOV	-0.022893	0.059402	-0.018420	0.010802	0.165955	0.353077	0.228539	0.039455	0.014371	0.146776	0.471935	1.000000	0.439804
DEC	-0.016688	0.204835	0.132406	0.133862	0.132704	0.242401	0.086661	-0.021688	-0.000775	0.100946	0.273467	0.439804	1.000000
ANNUAL RAIN	-0.043822	0.843335	0.862754	0.577039	0.350765	0.196634	-0.000042	-0.018607	0.052000	0.061076	0.004041	0.021506	0.196512

```
In [17]: import seaborn as sns
sns.heatmap(cor, vmax=1, vmin=-1, annot=True, linewidths=.5, cmap='bwr')
```

Out[17]: <Axes: >



```
In [18]: data['SWM']=data.apply(lambda row: row.JUN + row.JUL + row.AUG + row.SEP,axis=1)
data
```

Out[18]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN	SWM
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	33.6	136.3	1696.3
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	160.5	159.8	2185.9
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	225.0	156.7	1874.0
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	40.1	24.1	1977.6
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	344.7	1.3	1624.9
...
4105	LAKSHADWEEP	2005	17.6	11.1	0.0	37.0	92.8	248.5	378.9	102.4	278.0	164.2	218.3	26.6	28.7	1007.8
4106	LAKSHADWEEP	2006	20.1	0.0	33.0	0.3	327.9	286.9	172.3	150.7	318.5	119.1	158.9	10.9	20.1	928.4
4107	LAKSHADWEEP	2007	2.5	4.2	0.2	22.2	166.2	573.4	427.4	294.7	457.5	256.1	47.6	109.6	6.7	1753.0
4108	LAKSHADWEEP	2008	5.5	19.8	120.7	15.8	180.4	254.6	363.9	206.6	108.9	252.9	67.6	130.1	25.3	934.0
4109	LAKSHADWEEP	2009	4.7	1.5	0.1	18.1	162.1	401.2	266.4	185.0	145.1	87.4	166.2	132.3	6.2	997.7

3900 rows × 16 columns

```
In [19]: data['NEM']=data.apply(lambda row: row.OCT + row.NOV,axis=1)
data
```

Out[19]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN	SWM	NEM
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	33.6	136.3	1696.3	946.7
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	160.5	159.8	2185.9	556.2
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	225.0	156.7	1874.0	465.6
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	40.1	24.1	1977.6	530.9
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	344.7	1.3	1624.9	286.1
...
4105	LAKSHADWEEP	2005	17.6	11.1	0.0	37.0	92.8	248.5	378.9	102.4	278.0	164.2	218.3	26.6	28.7	1007.8	382.5
4106	LAKSHADWEEP	2006	20.1	0.0	33.0	0.3	327.9	286.9	172.3	150.7	318.5	119.1	158.9	10.9	20.1	928.4	278.0
4107	LAKSHADWEEP	2007	2.5	4.2	0.2	22.2	166.2	573.4	427.4	294.7	457.5	256.1	47.6	109.6	6.7	1753.0	303.7
4108	LAKSHADWEEP	2008	5.5	19.8	120.7	15.8	180.4	254.6	363.9	206.6	108.9	252.9	67.6	130.1	25.3	934.0	320.5
4109	LAKSHADWEEP	2009	4.7	1.5	0.1	18.1	162.1	401.2	266.4	185.0	145.1	87.4	166.2	132.3	6.2	997.7	253.6

3900 rows × 17 columns


```
In [20]: cor=data.corr()  
cor
```

/tmp/ipykernel_6057/4173678507.py:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_only to silence this warning.

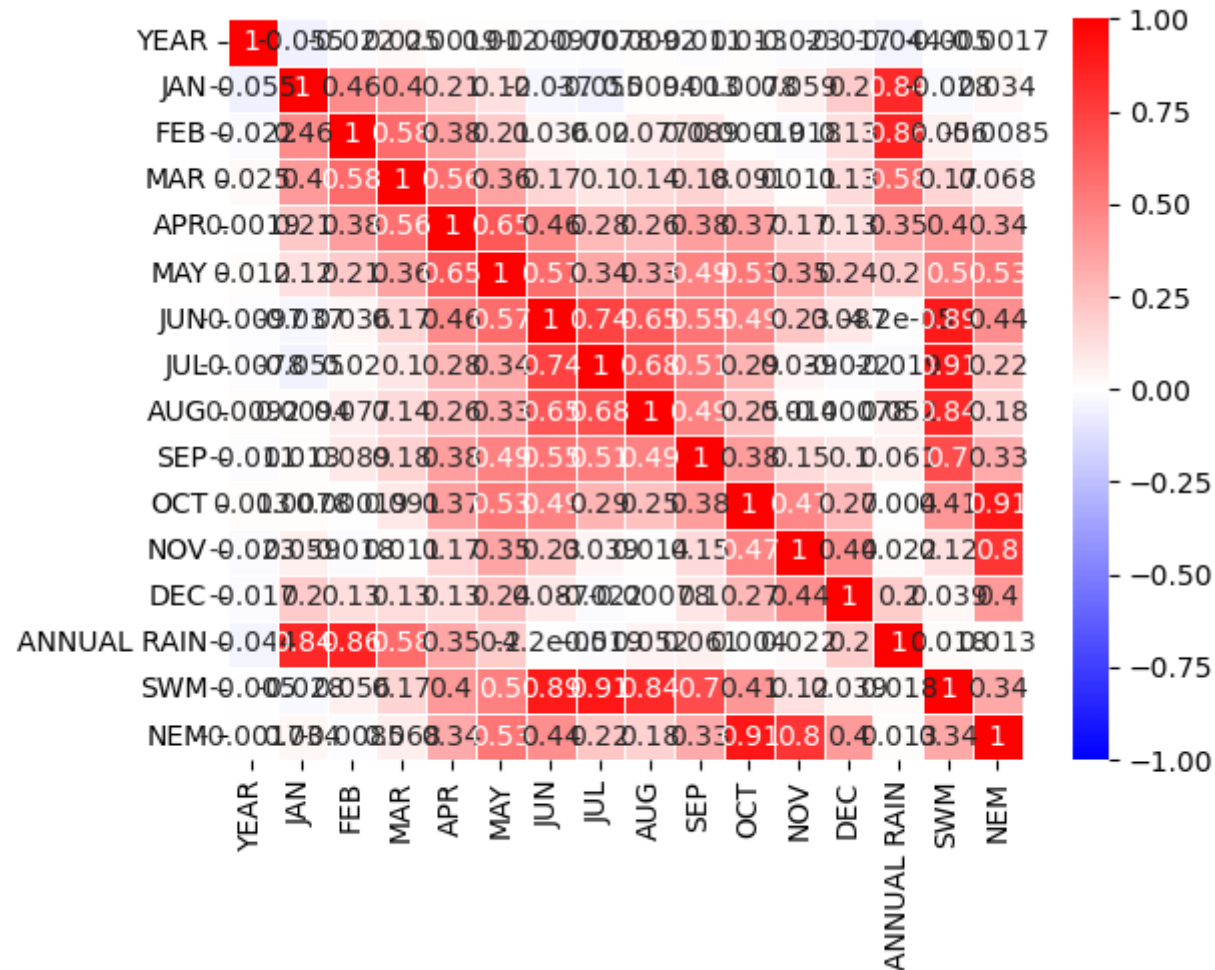
```
cor=data.corr()
```

Out[20]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
YEAR	1.000000	-0.054943	-0.021814	0.024663	0.001913	0.011656	-0.009702	-0.007850	0.009194	-0.011112	0.012884	-0.022893	-0.016688
JAN	-0.054943	1.000000	0.455875	0.396914	0.212725	0.124277	-0.037364	-0.054524	0.009423	0.013421	0.007817	0.059402	0.204835
FEB	-0.021814	0.455875	1.000000	0.582844	0.380745	0.209002	0.035763	0.020171	0.076897	0.088929	0.000194	-0.018420	0.132406
MAR	0.024663	0.396914	0.582844	1.000000	0.557110	0.362911	0.166572	0.100786	0.142062	0.178439	0.090809	0.010802	0.133862
APR	0.001913	0.212725	0.380745	0.557110	1.000000	0.654300	0.458790	0.276126	0.259701	0.381239	0.374406	0.165955	0.132704
MAY	0.011656	0.124277	0.209002	0.362911	0.654300	1.000000	0.571162	0.337946	0.329925	0.489130	0.531166	0.353077	0.242401
JUN	-0.009702	-0.037364	0.035763	0.166572	0.458790	0.571162	1.000000	0.740511	0.653670	0.546800	0.488370	0.228539	0.086661
JUL	-0.007850	-0.054524	0.020171	0.100786	0.276126	0.337946	0.740511	1.000000	0.682643	0.509115	0.294710	0.039455	-0.021688
AUG	0.009194	0.009423	0.076897	0.142062	0.259701	0.329925	0.653670	0.682643	1.000000	0.491390	0.248026	0.014371	-0.000775
SEP	-0.011112	0.013421	0.088929	0.178439	0.381239	0.489130	0.546800	0.509115	0.491390	1.000000	0.382860	0.146776	0.100946
OCT	0.012884	0.007817	0.000194	0.090809	0.374406	0.531166	0.488370	0.294710	0.248026	0.382860	1.000000	0.471935	0.273467
NOV	-0.022893	0.059402	-0.018420	0.010802	0.165955	0.353077	0.228539	0.039455	0.014371	0.146776	0.471935	1.000000	0.439804
DEC	-0.016688	0.204835	0.132406	0.133862	0.132704	0.242401	0.086661	-0.021688	-0.000775	0.100946	0.273467	0.439804	1.000000
ANNUAL RAIN	-0.043822	0.843335	0.862754	0.577039	0.350765	0.196634	-0.000042	-0.018607	0.052000	0.061076	0.004041	0.021506	0.196512
SWM	-0.005001	-0.027833	0.056244	0.165630	0.399782	0.499693	0.893689	0.907044	0.838638	0.698207	0.413606	0.122769	0.038595
NEM	-0.001738	0.033729	-0.008462	0.067830	0.335680	0.531871	0.443082	0.221070	0.176459	0.332487	0.909983	0.795018	0.395008

```
In [21]: import seaborn as sns
sns.heatmap(cor, vmax=1, vmin=-1, annot=True, linewidths=.5, cmap='bwr')
```

Out[21]: <Axes: >



In []:

