

```
In [47]: import pandas as pd
```

```
In [48]: data=pd.read_csv("/home/placement/Downloads/rainfall in india 1901-2015.csv")
```

```
In [49]: data.isna().sum()
```

```
Out[49]: 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 [50]: data1=data.loc[(data.YEAR<=2010)]
```

In [51]: `data1.tail(5)`

Out[51]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
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.
4110	LAKSHADWEEP	2010	18.8	0.0	1.2	35.6	79.0	318.9	336.7	335.1	161.5	155.4	201.5	81.5	1725.2	18.8	115.8	1152.2	438.

In [52]: `data2=data.drop(['ANNUAL','Jan-Feb','Mar-May','Jun-Sep','Oct-Dec'],axis=1)`

In [53]: data2

Out[53]:

	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
...
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	14.9
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	8.8
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	26.7
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	62.3
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	159.0

4116 rows × 14 columns

In [54]: data2['ANNUAL RAIN']=data2.apply(lambda row:row.JAN+row.FEB,axis=1)

In [55]: data2

Out[55]:

	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
...
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	14.9	7.9
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	8.8	19.3
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	26.7	60.6
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	62.3	69.3
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	159.0	2.7

4116 rows × 15 columns

In [56]: data2['ANNUAL RAIN']=data2.apply(lambda row:row.JAN+row.FEB+row.MAR+row.APR+row.MAY+row.JUN+row.JUL+row.AUG+

In [57]: data2

Out[57]:

	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	3373.2
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
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
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
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
...
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	14.9	1533.7
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	8.8	1405.5
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	26.7	1426.3
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	62.3	1395.0
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	159.0	1642.9

4116 rows × 15 columns

In [58]: data3=data.loc[(data.SUBDIVISION=='ARUNACHAL PRADESH')]

In [59]: data4=data3.drop(['SUBDIVISION'],axis=1)

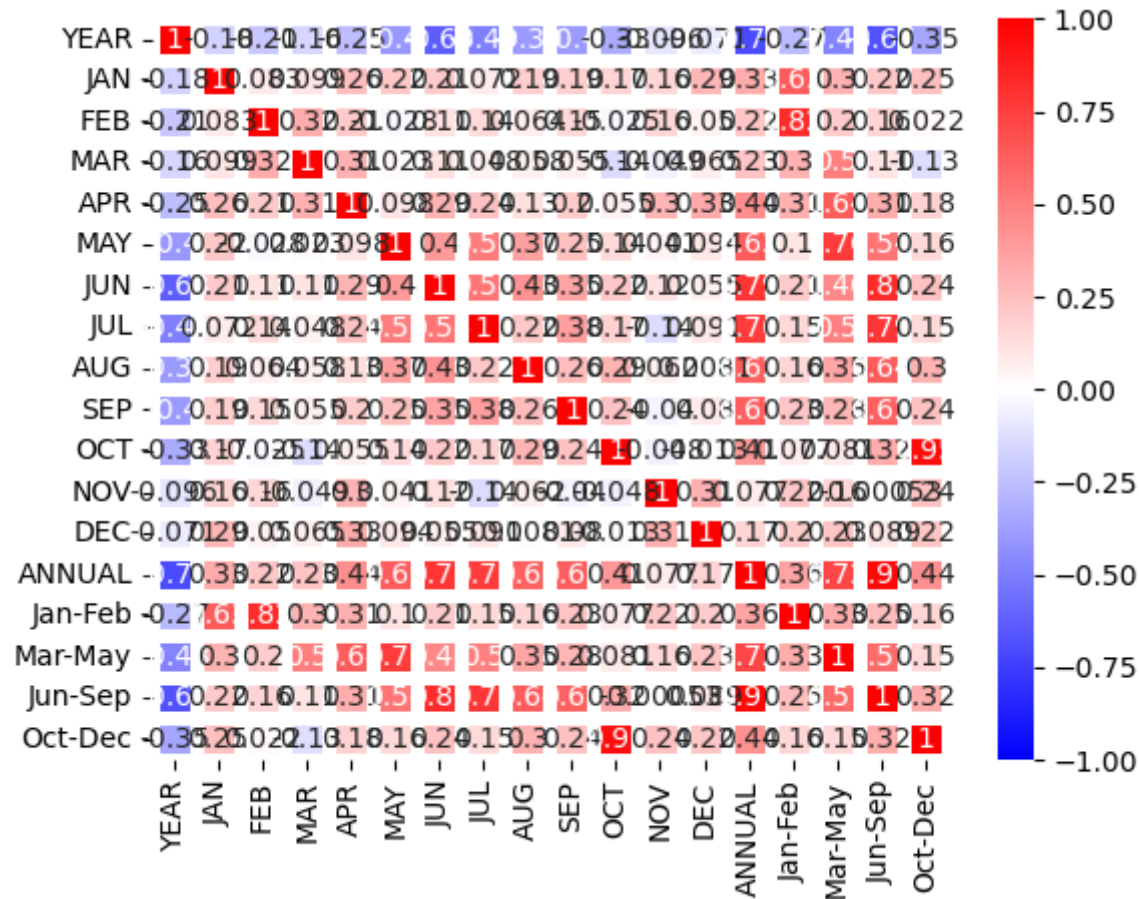
```
In [61]: cor=data4.corr()  
cor
```

Out[61]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	D
YEAR	1.000000	-0.176300	-0.213918	-0.157302	-0.253946	-0.401934	-0.626889	-0.494607	-0.394066	-0.396480	-0.325208	-0.096053	-0.070
JAN	-0.176300	1.000000	0.083391	0.099054	0.256921	0.217524	0.208187	0.071617	0.193102	0.189419	0.169379	0.162395	0.286
FEB	-0.213918	0.083391	1.000000	0.321564	0.205643	-0.027577	0.111802	0.140842	0.063751	0.154883	-0.024632	0.159914	0.050
MAR	-0.157302	0.099054	0.321564	1.000000	0.307354	0.023099	0.111475	0.047547	0.058362	0.054761	-0.137731	-0.048569	0.065
APR	-0.253946	0.256921	0.205643	0.307354	1.000000	0.097526	0.290472	0.238319	0.132668	0.198362	0.054994	0.295455	0.329
MAY	-0.401934	0.217524	-0.027577	0.023099	0.097526	1.000000	0.398268	0.510852	0.367445	0.246939	0.141720	0.040734	0.093
JUN	-0.626889	0.208187	0.111802	0.111475	0.290472	0.398268	1.000000	0.540408	0.426753	0.354854	0.217141	0.124429	0.054
JUL	-0.494607	0.071617	0.140842	0.047547	0.238319	0.510852	0.540408	1.000000	0.218016	0.380741	0.173107	-0.137416	0.091
AUG	-0.394066	0.193102	0.063751	0.058362	0.132668	0.367445	0.426753	0.218016	1.000000	0.259420	0.293511	0.062165	0.008
SEP	-0.396480	0.189419	0.154883	0.054761	0.198362	0.246939	0.354854	0.380741	0.259420	1.000000	0.241075	-0.040257	0.080
OCT	-0.325208	0.169379	-0.024632	-0.137731	0.054994	0.141720	0.217141	0.173107	0.293511	0.241075	1.000000	-0.047687	-0.013
NOV	-0.096053	0.162395	0.159914	-0.048569	0.295455	0.040734	0.124429	-0.137416	0.062165	-0.040257	-0.047687	1.000000	0.312
DEC	-0.070899	0.286771	0.050085	0.065364	0.329066	0.093530	0.054968	0.091248	0.008145	0.080062	-0.013078	0.312240	1.000
ANNUAL	-0.709544	0.325117	0.223998	0.227488	0.439047	0.634682	0.793095	0.762012	0.624845	0.626636	0.406220	0.076944	0.168
Jan-Feb	-0.266442	0.634518	0.823129	0.303319	0.305895	0.102570	0.205393	0.150032	0.159517	0.228104	0.077493	0.216565	0.202
Mar-May	-0.480430	0.300178	0.202870	0.500652	0.636367	0.759252	0.462125	0.498620	0.348350	0.284368	0.080929	0.159040	0.229
Jun-Sep	-0.673985	0.216207	0.164077	0.107055	0.314862	0.542457	0.827003	0.785435	0.636887	0.624762	0.321843	-0.000532	0.089
Oct-Dec	-0.346587	0.249172	0.022492	-0.126796	0.175082	0.156031	0.241761	0.147782	0.299315	0.241389	0.946660	0.236445	0.223

```
In [63]: import seaborn as sns
sns.heatmap(cor, vmax=1, vmin=-1, annot=True, linewidth=5, cmap='bwr')
```

Out[63]: <Axes: >



In []:

