

```
In [108]: import pandas as pd  
import numpy as np  
import warnings  
warnings.filterwarnings("ignore")
```

Reading the data

```
In [109]: data=pd.read_csv('/home/placement/Desktop/arunachal.csv')
```

In [110]: data

Out[110]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	NaN	NaN	NaN	NaN	117.9	811.8
1	111	ARUNACHAL PRADESH	1917	21.4	164.5	NaN	269.6	107.9	823.8	909.1	628.4	411.5	199.3	63.5	0.0	NaN	185.9	NaN
2	112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	125.2	7.8	13.7	5486.3	21.4	1196.9
3	113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6	4693.9	102.3	706.0
4	114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	103.3	0.0	0.0	4106.7	210.3	1143.9
...
86	196	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	216.0	604.6
87	197	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7	2259.6	109.7	587.7
88	198	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9	3020.7	110.8	590.9
89	199	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6	2244.4	116.4	499.9
90	200	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9	1749.9	110.8	394.7

91 rows × 20 columns



In [111]: data.describe()

Out[111]:

	Unnamed: 0	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SE
count	91.00000	91.000000	90.000000	90.000000	89.000000	91.000000	91.000000	90.000000	90.000000	91.000000	91.000000
mean	155.00000	1962.747253	48.598889	93.966667	154.446067	262.990110	364.651648	659.556667	711.963333	502.163736	433.27362
std	26.41338	27.695003	34.687078	46.258375	87.918484	113.395773	181.095447	311.642230	356.372598	275.716730	204.99135
min	110.00000	1916.000000	1.800000	6.100000	28.500000	94.700000	101.800000	239.400000	233.000000	172.400000	152.50000
25%	132.50000	1938.500000	20.075000	65.625000	101.700000	180.600000	237.150000	425.675000	442.150000	301.100000	282.15000
50%	155.00000	1964.000000	45.400000	87.600000	141.700000	245.400000	314.600000	545.750000	613.000000	411.600000	384.30000
75%	177.50000	1986.500000	65.150000	120.400000	189.600000	335.300000	447.050000	840.400000	922.075000	669.200000	521.15000
max	200.00000	2009.000000	164.500000	208.500000	605.600000	595.100000	1168.600000	1609.900000	2362.800000	1664.600000	1222.00000

In [112]: list(data)

Out[112]: ['Unnamed: 0',
'SUBDIVISION',
'YEAR',
'JAN',
'FEB',
'MAR',
'APR',
'MAY',
'JUN',
'JUL',
'AUG',
'SEP',
'OCT',
'NOV',
'DEC',
'ANNUAL',
'Jan-Feb',
'Mar-May',
'Jun-Sep',
'Oct-Dec']

drop()

```
In [113]: data1=data.drop(columns="Unnamed: 0",axis=1)
data1
```

Out[113]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
0	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	NaN	NaN	NaN	NaN	117.9	811.8	NaN	NaN
1	ARUNACHAL PRADESH	1917	21.4	164.5	NaN	269.6	107.9	823.8	909.1	628.4	411.5	199.3	63.5	0.0	NaN	185.9	NaN	2772.8	262.
2	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	125.2	7.8	13.7	5486.3	21.4	1196.9	4121.3	146.
3	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6	4693.9	102.3	706.0	2888.0	997.
4	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	103.3	0.0	0.0	4106.7	210.3	1143.9	2649.2	103.
...
86	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	216.0	604.6	1343.7	171.
87	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7	2259.6	109.7	587.7	1393.5	168.
88	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9	3020.7	110.8	590.9	2185.1	133.
89	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6	2244.4	116.4	499.9	1508.7	119.
90	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9	1749.9	110.8	394.7	1111.8	132.

91 rows × 19 columns



isna()

```
In [114]: data1.isna().sum()
```

```
Out[114]: SUBDIVISION    0  
YEAR                0  
JAN                 1  
FEB                 1  
MAR                 2  
APR                 0  
MAY                 0  
JUN                 1  
JUL                 1  
AUG                 0  
SEP                 0  
OCT                 2  
NOV                 2  
DEC                 2  
ANNUAL              6  
Jan-Feb             1  
Mar-May             2  
Jun-Sep             2  
Oct-Dec             3  
dtype: int64
```

fillna()

```
In [115]: data2=data1.fillna(data.median())
```

```
In [116]: data2.isna().sum()
```

```
Out[116]: SUBDIVISION      0  
          YEAR            0  
          JAN             0  
          FEB             0  
          MAR             0  
          APR             0  
          MAY             0  
          JUN             0  
          JUL             0  
          AUG             0  
          SEP             0  
          OCT             0  
          NOV             0  
          DEC             0  
          ANNUAL          0  
          Jan-Feb         0  
          Mar-May         0  
          Jun-Sep         0  
          Oct-Dec         0  
          dtype: int64
```

In [117]: data2

Out[117]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	O D
0	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	613.0	629.7	333.9	176.6	25.9	16.2	3433.2	117.9	811.8	2224.6	224.
1	ARUNACHAL PRADESH	1917	21.4	164.5	141.7	269.6	107.9	823.8	909.1	628.4	411.5	199.3	63.5	0.0	3433.2	185.9	733.1	2772.8	262.
2	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	125.2	7.8	13.7	5486.3	21.4	1196.9	4121.3	146.
3	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6	4693.9	102.3	706.0	2888.0	997.
4	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	103.3	0.0	0.0	4106.7	210.3	1143.9	2649.2	103.
...
86	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	216.0	604.6	1343.7	171.
87	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7	2259.6	109.7	587.7	1393.5	168.
88	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9	3020.7	110.8	590.9	2185.1	133.
89	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6	2244.4	116.4	499.9	1508.7	119.
90	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9	1749.9	110.8	394.7	1111.8	132.

91 rows × 19 columns



```
In [118]: data3=data2.drop(columns=['ANNUAL', 'Jan-Feb', 'Mar-May', 'Jun-Sep', 'Oct-Dec'])
data3
```

```
Out[118]:
```

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
0	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	613.0	629.7	333.9	176.6	25.9	16.2
1	ARUNACHAL PRADESH	1917	21.4	164.5	141.7	269.6	107.9	823.8	909.1	628.4	411.5	199.3	63.5	0.0
2	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	125.2	7.8	13.7
3	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6
4	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	103.3	0.0	0.0
...
86	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3
87	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7
88	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9
89	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6
90	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9

91 rows × 14 columns


```
In [121]: cor=data3.corr()  
cor
```

Out[121]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
YEAR	1.000000	-0.129527	-0.134034	-0.148148	-0.301073	-0.384602	-0.627409	-0.452705	-0.394444	-0.431541	-0.285585	-0.070690	-0.042148
JAN	-0.129527	1.000000	0.049837	0.103696	0.275567	0.212694	0.188988	0.034359	0.186214	0.179709	0.144535	0.167078	0.278241
FEB	-0.134034	0.049837	1.000000	0.339449	0.268663	-0.063912	0.084401	0.100425	0.028629	0.167881	-0.053775	0.163604	0.041696
MAR	-0.148148	0.103696	0.339449	1.000000	0.292102	0.018514	0.106516	0.037049	0.067452	0.051875	-0.147519	-0.056494	0.073333
APR	-0.301073	0.275567	0.268663	0.292102	1.000000	0.114128	0.283818	0.250371	0.157620	0.176335	0.045876	0.287188	0.335431
MAY	-0.384602	0.212694	-0.063912	0.018514	0.114128	1.000000	0.395012	0.505427	0.363992	0.258744	0.127682	0.025861	0.082868
JUN	-0.627409	0.188988	0.084401	0.106516	0.283818	0.395012	1.000000	0.519710	0.419358	0.349478	0.186572	0.098365	0.024855
JUL	-0.452705	0.034359	0.100425	0.037049	0.250371	0.505427	0.519710	1.000000	0.208773	0.381974	0.145849	-0.150645	0.080195
AUG	-0.394444	0.186214	0.028629	0.067452	0.157620	0.363992	0.419358	0.208773	1.000000	0.269123	0.297110	0.040123	-0.014359
SEP	-0.431541	0.179709	0.167881	0.051875	0.176335	0.258744	0.349478	0.381974	0.269123	1.000000	0.227757	-0.041297	0.078362
OCT	-0.285585	0.144535	-0.053775	-0.147519	0.045876	0.127682	0.186572	0.145849	0.297110	0.227757	1.000000	-0.055923	-0.027530
NOV	-0.070690	0.167078	0.163604	-0.056494	0.287188	0.025861	0.098365	-0.150645	0.040123	-0.041297	-0.055923	1.000000	0.313143
DEC	-0.042148	0.278241	0.041696	0.073333	0.335431	0.082868	0.024855	0.080195	-0.014359	0.078362	-0.027530	0.313143	1.000000

creating a new column from the data by the lambda function

In [122]: `data3['ANNUAL_RAIN']=data3.apply(lambda row:row.JAN+row.FEB+row.MAR+row.APR+row.MAY+row.JUN+row.JUL+row.AUG+row.SEP+row.OCT+row.NOV+row.DEC,axis=1)`
data3

Out[122]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL_RAIN
0	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	613.0	629.7	333.9	176.6	25.9	16.2	3849.9
1	ARUNACHAL PRADESH	1917	21.4	164.5	141.7	269.6	107.9	823.8	909.1	628.4	411.5	199.3	63.5	0.0	3740.7
2	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	125.2	7.8	13.7	5486.3
3	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6	4693.9
4	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	103.3	0.0	0.0	4106.7
...
86	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5
87	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7	2259.4
88	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9	3020.6
89	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6	2244.5
90	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9	1750.0

91 rows × 15 columns

creating new column swmonsoon by adding some months

In [123]: `data3['SWMONSOON']=data3.apply(lambda row:row.JUN+row.JUL+row.AUG+row.SEP,axis=1)`

In [124]: data3

Out[124]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL_RAIN	SWMONSOON
0	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	613.0	629.7	333.9	176.6	25.9	16.2	3849.9	2701.5
1	ARUNACHAL PRADESH	1917	21.4	164.5	141.7	269.6	107.9	823.8	909.1	628.4	411.5	199.3	63.5	0.0	3740.7	2772.8
2	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	125.2	7.8	13.7	5486.3	4121.3
3	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6	4693.9	2888.0
4	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	103.3	0.0	0.0	4106.7	2649.2
...
86	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	1343.7
87	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7	2259.4	1393.4
88	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9	3020.6	2185.1
89	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6	2244.5	1508.7
90	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9	1750.0	1111.8

91 rows × 16 columns

creating new column nwmonsoon by adding some months

In [129]: data3['NWMONSOON']=data3.apply(lambda row:row.OCT+row.NOV+row.DEC ,axis=1)

In [130]: data3

Out[130]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL_RAIN	SWMONSOON	NWM
0	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	613.0	629.7	333.9	176.6	25.9	16.2	3849.9	2701.5	
1	ARUNACHAL PRADESH	1917	21.4	164.5	141.7	269.6	107.9	823.8	909.1	628.4	411.5	199.3	63.5	0.0	3740.7	2772.8	
2	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	125.2	7.8	13.7	5486.3	4121.3	
3	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6	4693.9	2888.0	
4	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	103.3	0.0	0.0	4106.7	2649.2	
...
86	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	1343.7	
87	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7	2259.4	1393.4	
88	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9	3020.6	2185.1	
89	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6	2244.5	1508.7	
90	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9	1750.0	1111.8	

91 rows × 17 columns



dropping the unwanted columns

In [131]: data4=data3.drop(columns=['JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC'])

In [132]: data4

Out[132]:

	SUBDIVISION	YEAR	ANNUAL_RAIN	SWMONSOON	NWMONSOON
0	ARUNACHAL PRADESH	1916	3849.9	2701.5	218.7
1	ARUNACHAL PRADESH	1917	3740.7	2772.8	262.8
2	ARUNACHAL PRADESH	1918	5486.3	4121.3	146.7
3	ARUNACHAL PRADESH	1919	4693.9	2888.0	997.6
4	ARUNACHAL PRADESH	1920	4106.7	2649.2	103.3
...
86	ARUNACHAL PRADESH	2005	2335.5	1343.7	171.2
87	ARUNACHAL PRADESH	2006	2259.4	1393.4	168.6
88	ARUNACHAL PRADESH	2007	3020.6	2185.1	133.8
89	ARUNACHAL PRADESH	2008	2244.5	1508.7	119.4
90	ARUNACHAL PRADESH	2009	1750.0	1111.8	132.7

91 rows × 5 columns

Calculating Corelation

In [133]: corr=data4.corr()
corr

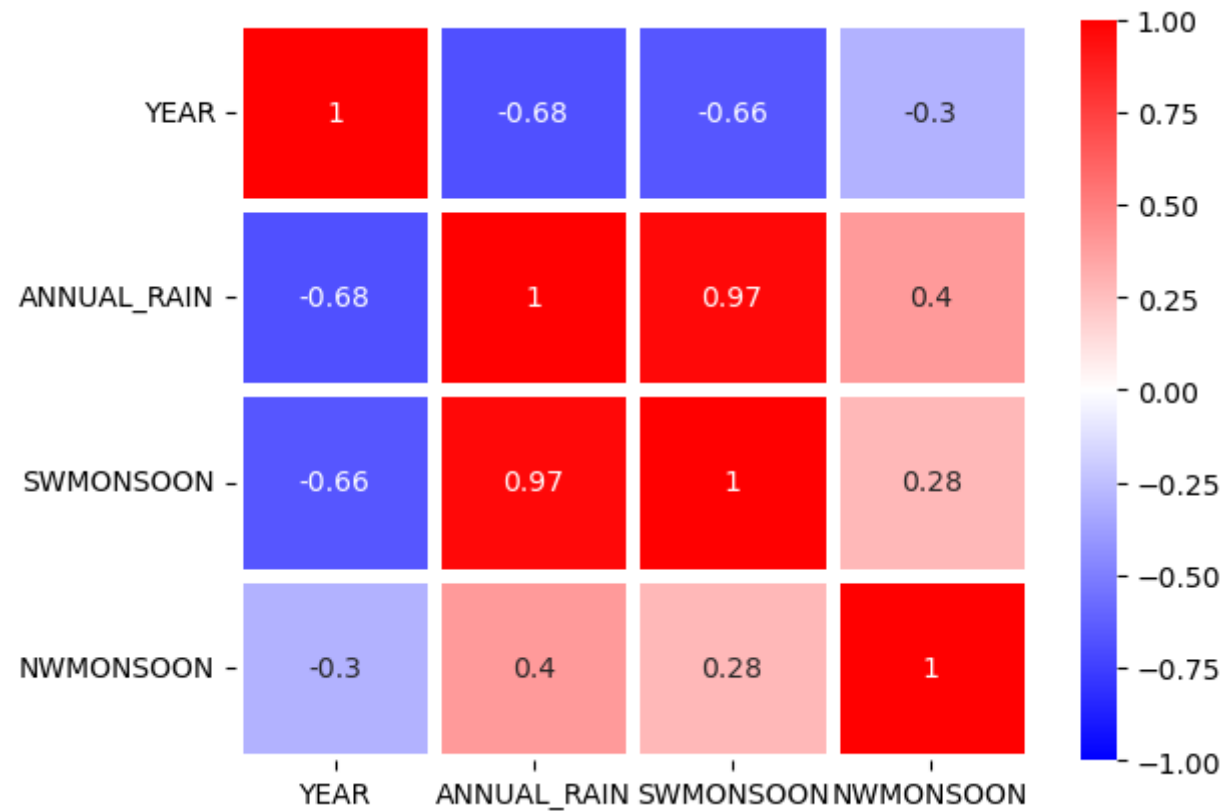
Out[133]:

	YEAR	ANNUAL_RAIN	SWMONSOON	NWMONSOON
YEAR	1.000000	-0.679822	-0.661484	-0.298527
ANNUAL_RAIN	-0.679822	1.000000	0.966705	0.398233
SWMONSOON	-0.661484	0.966705	1.000000	0.278525
NWMONSOON	-0.298527	0.398233	0.278525	1.000000

getting heatmap from the correalted data

```
In [135]: import seaborn as sns  
sns.heatmap(corr,vmax=1,vmin=-1,annot=True,linewidth=5,cmap='bwr')
```

Out[135]: <Axes: >



In []:

