In [1]:
```
import pandas as pd
import warnings
warnings.filterwarnings("ignore")
```

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

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

Out[3]:

	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

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

Out[4]:

	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.(
mean	1958.218659	18.957320	21.805325	27.359197	43.127432	85.745417	230.234444	347.214334	290.263497	197.361922	95.5
std	33.140898	33.585371	35.909488	46.959424	67.831168	123.234904	234.710758	269.539667	188.770477	135.408345	99.5
min	1901.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.400000	0.000000	0.000000	0.100000	0.(
25%	1930.000000	0.600000	0.600000	1.000000	3.000000	8.600000	70.350000	175.600000	155.975000	100.525000	14.(
50%	1958.000000	6.000000	6.700000	7.800000	15.700000	36.600000	138.700000	284.800000	259.400000	173.900000	65.2
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	948.3

In [5]:
```
data.groupby(['SUBDIVISION']).count()
```

Out[5]:

SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
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
NAGA MANI MIZO TRIPURA	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	115

SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
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 [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.loc[(data.YEAR<=2010)]
```

In [8]:
```
data1.tail(10)
```

Out[8]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oc D
4101	LAKSHADWEEP	2001	4.4	20.4	0.0	104.6	187.3	283.9	198.9	144.3	213.5	105.2	101.5	16.6	1380.6	24.8	291.9	840.6	223
4102	LAKSHADWEEP	2002	10.8	16.8	7.2	23.4	189.8	261.8	81.3	143.9	50.0	178.2	52.9	17.4	1033.5	27.6	220.4	537.0	248
4103	LAKSHADWEEP	2003	11.8	18.2	28.5	18.1	109.6	364.5	400.6	92.1	84.3	191.6	206.1	7.5	1532.9	30.0	156.2	941.5	405
4104	LAKSHADWEEP	2004	7.2	1.5	1.9	7.7	330.2	251.2	280.8	169.5	200.0	193.4	107.6	2.2	1553.2	8.7	339.8	901.5	303
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
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 [9]:
```
data1=data.drop(['Jan-Feb','Mar-May','Jun-Sep','Oct-Dec'],axis=1)
```

In [10]:
```
data1
```

Out[10]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
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 [11]:
```
data1.groupby(['SUBDIVISION']).count()
```

Out[11]:

SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
ANDAMAN & NICOBAR ISLANDS	110	110	110	108	108	109	108	108	108	107	108	108	107	104
ARUNACHAL PRADESH	97	96	96	95	97	97	96	96	97	97	95	95	95	91
ASSAM & MEGHALAYA	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
CHHATTISGARH	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
COASTAL KARNATAKA	115	114	115	115	115	115	115	115	115	115	115	115	115	114
EAST MADHYA PRADESH	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
EAST UTTAR PRADESH	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
GUJARAT REGION	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
HIMACHAL PRADESH	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
JHARKHAND	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
KONKAN & GOA	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
MADHYA MAHARASHTRA	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
NAGA MANI MIZO TRIPURA	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

SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
ORISSA	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
RAYALSEEMA	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
SOUTH INTERIOR KARNATAKA	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
TAMIL NADU	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
UTTARAKHAND	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
WEST MADHYA PRADESH	115	115	114	115	115	115	115	115	115	115	115	115	115	114
WEST RAJASTHAN	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

In [12]:
```
data2=data1.loc[(data1.SUBDIVISION=="TELANGANA")]
```

In [13]:
```
data2
```

Out[13]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
3197	TELANGANA	1901	6.9	41.8	7.8	45.2	22.0	123.6	237.8	177.2	77.7	75.5	12.2	0.0	827.7
3198	TELANGANA	1902	0.0	0.0	0.2	10.7	7.3	52.4	146.3	142.8	190.5	41.7	31.2	7.3	630.4
3199	TELANGANA	1903	12.9	4.6	0.0	9.9	40.7	99.2	505.2	246.7	191.9	155.8	15.5	1.1	1283.4
3200	TELANGANA	1904	0.0	0.0	10.8	0.8	14.7	104.2	139.5	50.0	162.3	44.4	0.0	0.0	526.7
3201	TELANGANA	1905	0.0	4.3	12.8	27.6	32.2	129.5	82.4	237.3	179.1	19.6	0.0	0.0	724.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3307	TELANGANA	2011	0.0	11.9	2.6	25.6	9.3	83.9	268.2	225.9	107.6	13.9	4.2	0.0	753.1
3308	TELANGANA	2012	6.7	0.0	0.2	14.0	8.4	124.4	300.3	229.9	202.4	83.6	38.7	0.0	1008.6
3309	TELANGANA	2013	2.4	29.0	0.2	24.4	8.5	213.4	453.8	230.6	161.4	205.9	16.4	2.7	1348.7
3310	TELANGANA	2014	0.2	2.9	58.3	10.3	73.3	62.3	146.0	205.2	146.8	29.6	10.8	0.7	746.4
3311	TELANGANA	2015	17.5	0.0	43.0	65.7	23.3	266.9	104.4	160.5	158.3	15.6	0.3	1.7	857.3

115 rows × 15 columns

In [14]:
```
data2.isna().sum()
```

Out[14]:
```
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
dtype: int64
```

In [15]:
```
data2.tail(150)
```

Out[15]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
3197	TELANGANA	1901	6.9	41.8	7.8	45.2	22.0	123.6	237.8	177.2	77.7	75.5	12.2	0.0	827.7
3198	TELANGANA	1902	0.0	0.0	0.2	10.7	7.3	52.4	146.3	142.8	190.5	41.7	31.2	7.3	630.4
3199	TELANGANA	1903	12.9	4.6	0.0	9.9	40.7	99.2	505.2	246.7	191.9	155.8	15.5	1.1	1283.4
3200	TELANGANA	1904	0.0	0.0	10.8	0.8	14.7	104.2	139.5	50.0	162.3	44.4	0.0	0.0	526.7
3201	TELANGANA	1905	0.0	4.3	12.8	27.6	32.2	129.5	82.4	237.3	179.1	19.6	0.0	0.0	724.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3307	TELANGANA	2011	0.0	11.9	2.6	25.6	9.3	83.9	268.2	225.9	107.6	13.9	4.2	0.0	753.1
3308	TELANGANA	2012	6.7	0.0	0.2	14.0	8.4	124.4	300.3	229.9	202.4	83.6	38.7	0.0	1008.6
3309	TELANGANA	2013	2.4	29.0	0.2	24.4	8.5	213.4	453.8	230.6	161.4	205.9	16.4	2.7	1348.7
3310	TELANGANA	2014	0.2	2.9	58.3	10.3	73.3	62.3	146.0	205.2	146.8	29.6	10.8	0.7	746.4
3311	TELANGANA	2015	17.5	0.0	43.0	65.7	23.3	266.9	104.4	160.5	158.3	15.6	0.3	1.7	857.3

115 rows × 15 columns

In [16]:
```
data['SUBDIVISION'].unique()
```

Out[16]:
```
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 [17]:
```
data2['SWM-JJAS']=data2.apply(lambda row: row.JUN+row.JUL+row.AUG+row.SEP,axis=1)
data2
```

Out[17]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	SWM-JJAS
3197	TELANGANA	1901	6.9	41.8	7.8	45.2	22.0	123.6	237.8	177.2	77.7	75.5	12.2	0.0	827.7	616.3
3198	TELANGANA	1902	0.0	0.0	0.2	10.7	7.3	52.4	146.3	142.8	190.5	41.7	31.2	7.3	630.4	532.0
3199	TELANGANA	1903	12.9	4.6	0.0	9.9	40.7	99.2	505.2	246.7	191.9	155.8	15.5	1.1	1283.4	1043.0
3200	TELANGANA	1904	0.0	0.0	10.8	0.8	14.7	104.2	139.5	50.0	162.3	44.4	0.0	0.0	526.7	456.0
3201	TELANGANA	1905	0.0	4.3	12.8	27.6	32.2	129.5	82.4	237.3	179.1	19.6	0.0	0.0	724.9	628.3
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3307	TELANGANA	2011	0.0	11.9	2.6	25.6	9.3	83.9	268.2	225.9	107.6	13.9	4.2	0.0	753.1	685.6
3308	TELANGANA	2012	6.7	0.0	0.2	14.0	8.4	124.4	300.3	229.9	202.4	83.6	38.7	0.0	1008.6	857.0
3309	TELANGANA	2013	2.4	29.0	0.2	24.4	8.5	213.4	453.8	230.6	161.4	205.9	16.4	2.7	1348.7	1059.2
3310	TELANGANA	2014	0.2	2.9	58.3	10.3	73.3	62.3	146.0	205.2	146.8	29.6	10.8	0.7	746.4	560.3
3311	TELANGANA	2015	17.5	0.0	43.0	65.7	23.3	266.9	104.4	160.5	158.3	15.6	0.3	1.7	857.3	690.1

115 rows × 16 columns

In [18]:
```
data2['NEM-OND']=data2.apply(lambda row: row.OCT+row.NOV+row.DEC,axis=1)
data2
```

Out[18]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	SWM-JJAS	NEM-OND
3197	TELANGANA	1901	6.9	41.8	7.8	45.2	22.0	123.6	237.8	177.2	77.7	75.5	12.2	0.0	827.7	616.3	87.7
3198	TELANGANA	1902	0.0	0.0	0.2	10.7	7.3	52.4	146.3	142.8	190.5	41.7	31.2	7.3	630.4	532.0	80.2
3199	TELANGANA	1903	12.9	4.6	0.0	9.9	40.7	99.2	505.2	246.7	191.9	155.8	15.5	1.1	1283.4	1043.0	172.4
3200	TELANGANA	1904	0.0	0.0	10.8	0.8	14.7	104.2	139.5	50.0	162.3	44.4	0.0	0.0	526.7	456.0	44.4
3201	TELANGANA	1905	0.0	4.3	12.8	27.6	32.2	129.5	82.4	237.3	179.1	19.6	0.0	0.0	724.9	628.3	19.6
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3307	TELANGANA	2011	0.0	11.9	2.6	25.6	9.3	83.9	268.2	225.9	107.6	13.9	4.2	0.0	753.1	685.6	18.1
3308	TELANGANA	2012	6.7	0.0	0.2	14.0	8.4	124.4	300.3	229.9	202.4	83.6	38.7	0.0	1008.6	857.0	122.3
3309	TELANGANA	2013	2.4	29.0	0.2	24.4	8.5	213.4	453.8	230.6	161.4	205.9	16.4	2.7	1348.7	1059.2	225.0
3310	TELANGANA	2014	0.2	2.9	58.3	10.3	73.3	62.3	146.0	205.2	146.8	29.6	10.8	0.7	746.4	560.3	41.1
3311	TELANGANA	2015	17.5	0.0	43.0	65.7	23.3	266.9	104.4	160.5	158.3	15.6	0.3	1.7	857.3	690.1	17.6

115 rows × 17 columns

In [36]:
```
data2=data2.drop(['SUBDIVISION'],axis=1)
```

In [37]:
```
data2
```

Out[37]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	SWM-JJAS	NEM-OND	ANNUAL RAIN
3197	1901	6.9	41.8	7.8	45.2	22.0	123.6	237.8	177.2	77.7	75.5	12.2	0.0	827.7	616.3	87.7	827.7
3198	1902	0.0	0.0	0.2	10.7	7.3	52.4	146.3	142.8	190.5	41.7	31.2	7.3	630.4	532.0	80.2	630.4
3199	1903	12.9	4.6	0.0	9.9	40.7	99.2	505.2	246.7	191.9	155.8	15.5	1.1	1283.4	1043.0	172.4	1283.5
3200	1904	0.0	0.0	10.8	0.8	14.7	104.2	139.5	50.0	162.3	44.4	0.0	0.0	526.7	456.0	44.4	526.7
3201	1905	0.0	4.3	12.8	27.6	32.2	129.5	82.4	237.3	179.1	19.6	0.0	0.0	724.9	628.3	19.6	724.8
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3307	2011	0.0	11.9	2.6	25.6	9.3	83.9	268.2	225.9	107.6	13.9	4.2	0.0	753.1	685.6	18.1	753.1
3308	2012	6.7	0.0	0.2	14.0	8.4	124.4	300.3	229.9	202.4	83.6	38.7	0.0	1008.6	857.0	122.3	1008.6
3309	2013	2.4	29.0	0.2	24.4	8.5	213.4	453.8	230.6	161.4	205.9	16.4	2.7	1348.7	1059.2	225.0	1348.7
3310	2014	0.2	2.9	58.3	10.3	73.3	62.3	146.0	205.2	146.8	29.6	10.8	0.7	746.4	560.3	41.1	746.4
3311	2015	17.5	0.0	43.0	65.7	23.3	266.9	104.4	160.5	158.3	15.6	0.3	1.7	857.3	690.1	17.6	857.2

115 rows × 17 columns

In [38]:
```
cor=data2.corr()
```

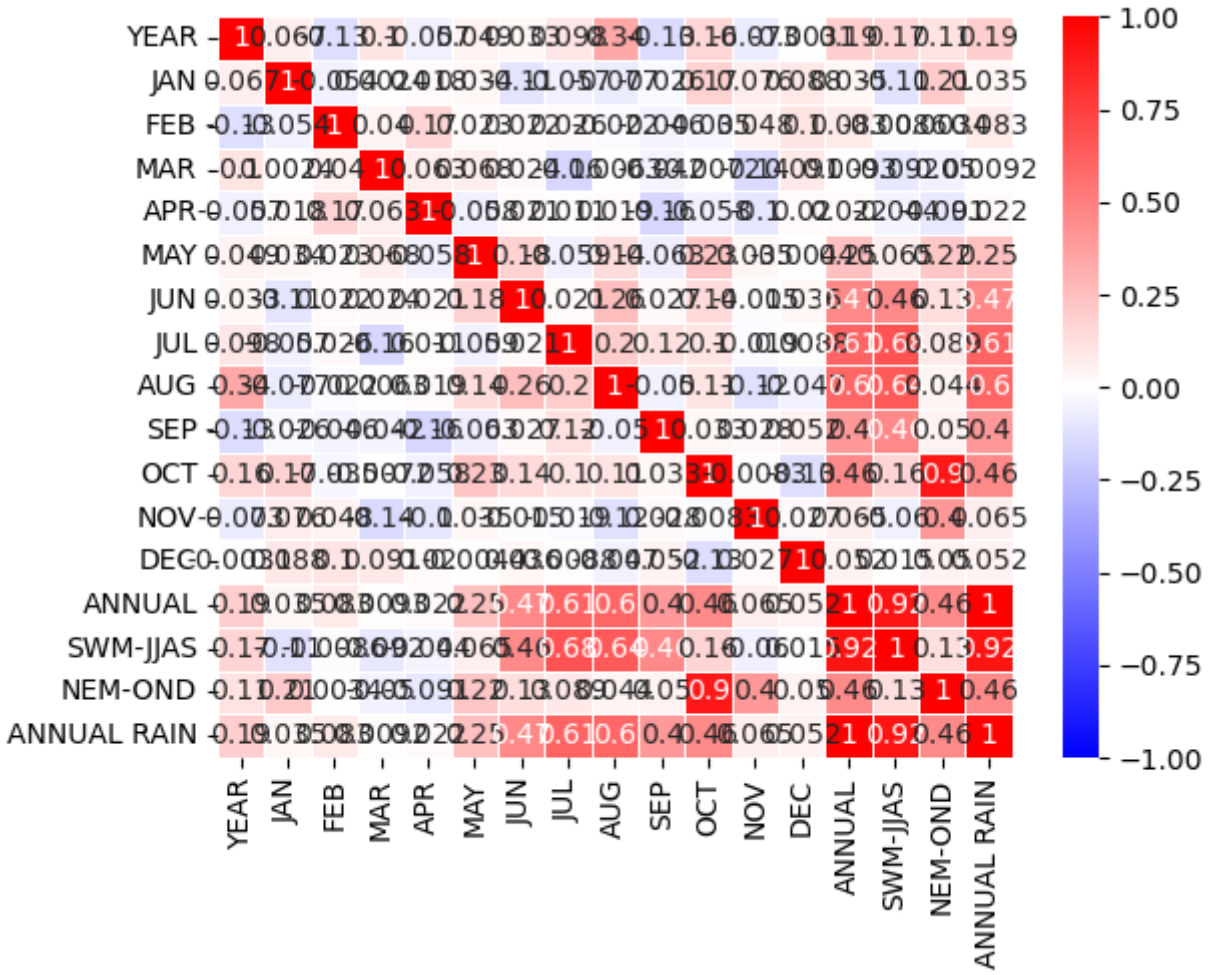
In [39]:
```
cor
```

Out[39]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	D
YEAR	1.000000	0.066679	-0.127789	0.103372	-0.057459	0.049107	0.032522	0.097739	0.343648	-0.134013	0.156396	-0.072673	-0.003
JAN	0.066679	1.000000	-0.054219	0.002439	0.017525	0.033719	-0.107093	-0.056891	-0.076973	-0.026067	0.174812	0.076095	0.088
FEB	-0.127789	-0.054219	1.000000	0.039952	0.166871	0.022623	0.022467	0.026066	-0.021750	-0.045791	-0.035158	0.048416	0.101
MAR	0.103372	0.002439	0.039952	1.000000	0.062920	0.068031	0.024190	-0.162996	0.006257	-0.041883	-0.007192	-0.142749	0.091
APR	-0.057459	0.017525	0.166871	0.062920	1.000000	-0.057561	0.021496	0.011166	0.019299	-0.159302	-0.058436	-0.100563	0.020
MAY	0.049107	0.033719	0.022623	0.068031	-0.057561	1.000000	0.179190	-0.059126	0.137449	-0.063313	0.229135	0.034746	-0.004
JUN	0.032522	-0.107093	0.022467	0.024190	0.021496	0.179190	1.000000	0.021370	0.255568	0.026901	0.141178	-0.015491	0.035
JUL	0.097739	-0.056891	0.026066	-0.162996	0.011166	-0.059126	0.021370	1.000000	0.201899	0.123573	0.103609	-0.018848	0.008
AUG	0.343648	-0.076973	-0.021750	0.006257	0.019299	0.137449	0.255568	0.201899	1.000000	-0.049929	0.108571	-0.119876	-0.047
SEP	-0.134013	-0.026067	-0.045791	-0.041883	-0.159302	-0.063313	0.026901	0.123573	-0.049929	1.000000	0.032922	0.027962	0.051
OCT	0.156396	0.174812	-0.035158	-0.007192	-0.058436	0.229135	0.141178	0.103609	0.108571	0.032922	1.000000	-0.008315	-0.130
NOV	-0.072673	0.076095	0.048416	-0.142749	-0.100563	0.034746	-0.015491	-0.018848	-0.119876	0.027962	-0.008315	1.000000	0.026
DEC	-0.003100	0.088278	0.101628	0.091271	0.020066	-0.004376	0.035836	0.008781	-0.047230	0.051672	-0.130792	0.026896	1.000
ANNUAL	0.192653	0.035032	0.082825	0.009295	0.021866	0.248985	0.468344	0.609029	0.596162	0.400340	0.459625	0.064825	0.051
SWM-JJAS	0.169694	-0.111114	-0.008619	-0.092281	-0.044294	0.064708	0.459444	0.678840	0.642400	0.464608	0.164273	-0.060302	0.014
NEM-OND	0.114384	0.206271	0.003427	-0.049876	-0.091461	0.224828	0.129737	0.089381	0.044067	0.049975	0.898531	0.401799	0.050
ANNUAL RAIN	0.192611	0.035049	0.082843	0.009230	0.021805	0.248903	0.468318	0.609108	0.596173	0.400333	0.459548	0.064896	0.051

In [40]:
```
import seaborn as sns
sns.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidth=.5,cmap='bwr')
```

Out[40]: <Axes: >



In [19]:
```
data3=data.loc[(data.SUBDIVISION =='TAMILNADU')]
```

In [20]:
```
data3
```

Out[20]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec

In [21]:
```
data.isna().sum()
```

Out[21]:
```
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 [22]:
```
data3.fillna(data.mode())
```

Out[22]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec

In [23]:
```
data3
```

Out[23]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec

In [24]:
```
data3.isna().sum()
```

Out[24]:
```
SUBDIVISION    0.0
YEAR           0.0
JAN            0.0
FEB            0.0
MAR            0.0
APR            0.0
MAY            0.0
JUN            0.0
JUL            0.0
AUG            0.0
SEP            0.0
OCT            0.0
NOV            0.0
DEC            0.0
ANNUAL         0.0
Jan-Feb        0.0
Mar-May        0.0
Jun-Sep        0.0
Oct-Dec        0.0
dtype: float64
```

In [25]:
```
data2['ANNUAL RAIN']=data2.apply(lambda row: row.JAN + row.FEB + row.MAR + row.APR + row.MAY + row.JUN + row
```

In [26]:
```
data2
```

Out[26]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	SWM-JJAS	NEM-OND	ANNUAL RAIN
3197	TELANGANA	1901	6.9	41.8	7.8	45.2	22.0	123.6	237.8	177.2	77.7	75.5	12.2	0.0	827.7	616.3	87.7	827.7
3198	TELANGANA	1902	0.0	0.0	0.2	10.7	7.3	52.4	146.3	142.8	190.5	41.7	31.2	7.3	630.4	532.0	80.2	630.4
3199	TELANGANA	1903	12.9	4.6	0.0	9.9	40.7	99.2	505.2	246.7	191.9	155.8	15.5	1.1	1283.4	1043.0	172.4	1283.5
3200	TELANGANA	1904	0.0	0.0	10.8	0.8	14.7	104.2	139.5	50.0	162.3	44.4	0.0	0.0	526.7	456.0	44.4	526.7
3201	TELANGANA	1905	0.0	4.3	12.8	27.6	32.2	129.5	82.4	237.3	179.1	19.6	0.0	0.0	724.9	628.3	19.6	724.8
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3307	TELANGANA	2011	0.0	11.9	2.6	25.6	9.3	83.9	268.2	225.9	107.6	13.9	4.2	0.0	753.1	685.6	18.1	753.1
3308	TELANGANA	2012	6.7	0.0	0.2	14.0	8.4	124.4	300.3	229.9	202.4	83.6	38.7	0.0	1008.6	857.0	122.3	1008.6
3309	TELANGANA	2013	2.4	29.0	0.2	24.4	8.5	213.4	453.8	230.6	161.4	205.9	16.4	2.7	1348.7	1059.2	225.0	1348.7
3310	TELANGANA	2014	0.2	2.9	58.3	10.3	73.3	62.3	146.0	205.2	146.8	29.6	10.8	0.7	746.4	560.3	41.1	746.4
3311	TELANGANA	2015	17.5	0.0	43.0	65.7	23.3	266.9	104.4	160.5	158.3	15.6	0.3	1.7	857.3	690.1	17.6	857.2

115 rows × 18 columns

In [27]:
```
data4=data.drop(['SUBDIVISION'],axis=1)
```

In [28]:
```
data4.corr()
```

Out[28]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
YEAR	1.000000	-0.056235	-0.022144	0.020338	0.008007	0.003594	-0.013594	-0.016240	0.006442	-0.006670	0.002406	-0.018776	-0.019139
JAN	-0.056235	1.000000	0.456183	0.398502	0.209302	0.129622	-0.033725	-0.051642	0.011952	0.024289	0.012374	0.067281	0.219701
FEB	-0.022144	0.456183	1.000000	0.579819	0.367114	0.203062	0.033703	0.016235	0.072159	0.080148	-0.004581	-0.023413	0.132570
MAR	0.020338	0.398502	0.579819	1.000000	0.556856	0.362815	0.165857	0.097334	0.135071	0.178904	0.086187	0.008814	0.136328
APR	0.008007	0.209302	0.367114	0.556856	1.000000	0.650595	0.457091	0.268097	0.256168	0.382525	0.368886	0.165642	0.132892
MAY	0.003594	0.129622	0.203062	0.362815	0.650595	1.000000	0.567618	0.332283	0.329499	0.492378	0.529342	0.351931	0.250112
JUN	-0.013594	-0.033725	0.033703	0.165857	0.457091	0.567618	1.000000	0.741285	0.655142	0.551890	0.490393	0.229718	0.088782
JUL	-0.016240	-0.051642	0.016235	0.097334	0.268097	0.332283	0.741285	1.000000	0.686662	0.513067	0.299221	0.042671	-0.019427
AUG	0.006442	0.011952	0.072159	0.135071	0.256168	0.329499	0.655142	0.686662	1.000000	0.497037	0.250600	0.017488	0.001648
SEP	-0.006670	0.024289	0.080148	0.178904	0.382525	0.492378	0.551890	0.513067	0.497037	1.000000	0.384138	0.153465	0.109457
OCT	0.002406	0.012374	-0.004581	0.086187	0.368886	0.529342	0.490393	0.299221	0.250600	0.384138	1.000000	0.477503	0.281172
NOV	-0.018776	0.067281	-0.023413	0.008814	0.165642	0.351931	0.229718	0.042671	0.017488	0.153465	0.477503	1.000000	0.451407
DEC	-0.019139	0.219701	0.132570	0.136328	0.132892	0.250112	0.088782	-0.019427	0.001648	0.109457	0.281172	0.451407	1.000000
ANNUAL	-0.008044	0.105696	0.181563	0.322199	0.577573	0.698013	0.891303	0.812279	0.759304	0.715135	0.587065	0.308768	0.207176
Jan-Feb	-0.044653	0.842390	0.863815	0.576366	0.340841	0.196168	0.001016	-0.019157	0.050918	0.062131	0.003743	0.022885	0.204848
Mar-May	0.010637	0.242256	0.382620	0.642294	0.864172	0.915019	0.538562	0.313726	0.318347	0.470032	0.468048	0.272268	0.228473
Jun-Sep	-0.009418	-0.022748	0.051066	0.162055	0.394859	0.496164	0.893968	0.907723	0.840352	0.701980	0.416350	0.126338	0.042440
Oct-Dec	-0.010155	0.090932	0.021878	0.090108	0.321407	0.523684	0.409050	0.190400	0.156293	0.319832	0.862761	0.808798	0.606658

In [29]:
```
cor=data4.corr()
```

In [30]:
```
data4
```

Out[30]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
0	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	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	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	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	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4111	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	7.9	196.2	1013.0	316.6
4112	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	19.3	99.6	1119.5	167.1
4113	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	60.6	131.1	1057.0	177.6
4114	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	69.3	76.7	958.5	290.5
4115	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	2.7	223.9	860.9	555.4

4116 rows × 18 columns

In [31]:
```
data4=data4.drop(['JAN','FEB','MAR','APR','MAY','JUN','JUL','AUG','SEP','OCT','NOV','DEC'],axis=1)
```

In [32]:
```
data4
```

Out[32]:

	YEAR	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
0	1901	3373.2	136.3	560.3	1696.3	980.3
1	1902	3520.7	159.8	458.3	2185.9	716.7
2	1903	2957.4	156.7	236.1	1874.0	690.6
3	1904	3079.6	24.1	506.9	1977.6	571.0
4	1905	2566.7	1.3	309.7	1624.9	630.8
...	...	...	...	...	...	...
4111	2011	1533.7	7.9	196.2	1013.0	316.6
4112	2012	1405.5	19.3	99.6	1119.5	167.1
4113	2013	1426.3	60.6	131.1	1057.0	177.6
4114	2014	1395.0	69.3	76.7	958.5	290.5
4115	2015	1642.9	2.7	223.9	860.9	555.4

4116 rows × 6 columns

In [33]:
```
cor=data4.corr()
```

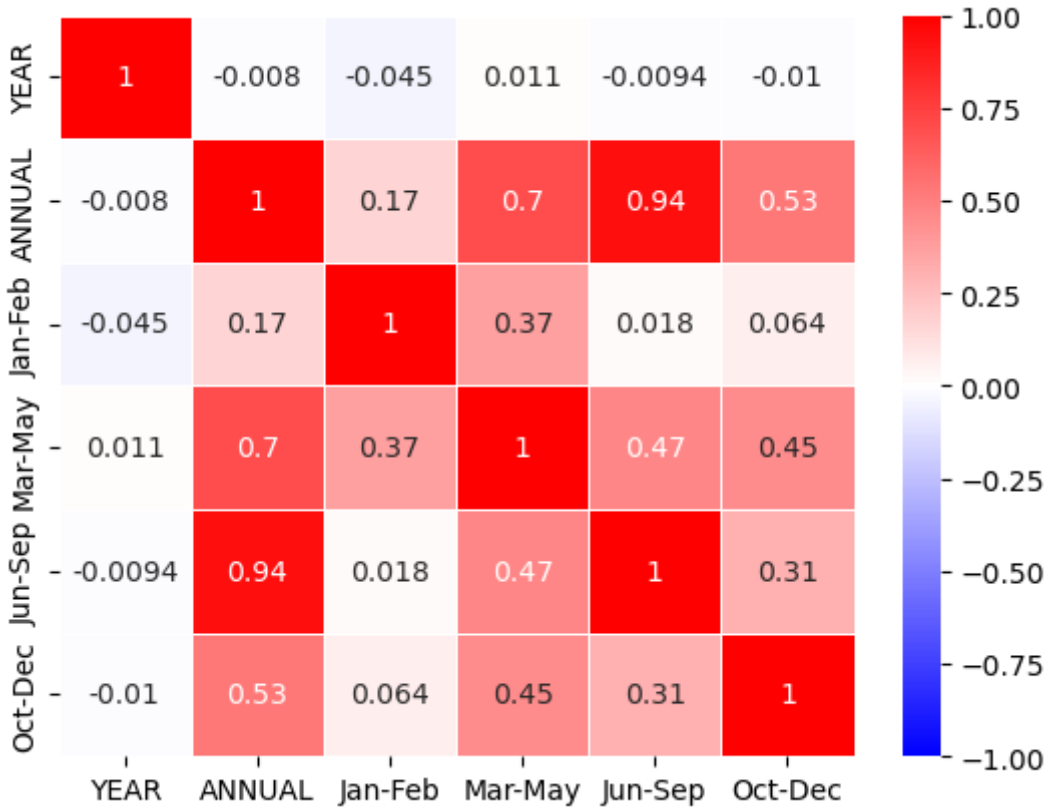
In [34]:
```
cor
```

Out[34]:

	YEAR	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
YEAR	1.000000	-0.008044	-0.044653	0.010637	-0.009418	-0.010155
ANNUAL	-0.008044	1.000000	0.169710	0.696318	0.943661	0.529596
Jan-Feb	-0.044653	0.169710	1.000000	0.368551	0.018126	0.064088
Mar-May	0.010637	0.696318	0.368551	1.000000	0.472508	0.449004
Jun-Sep	-0.009418	0.943661	0.018126	0.472508	1.000000	0.310433
Oct-Dec	-0.010155	0.529596	0.064088	0.449004	0.310433	1.000000

In [35]:
```
import seaborn as sns
sns.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidths=.5,cmap='bwr')
```

Out[35]: <Axes: >



In [ ]: