In [5]:
```
import pandas as pd
```

In [25]:
```
data=pd.read_csv("/home/palcement/Downloads/nikhil/rainfall in india 1901-2015.csv")
data
```

Out[25]:

	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
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	7.9	196.2	1013.0	3
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	19.3	99.6	1119.5	1
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	60.6	131.1	1057.0	1
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	69.3	76.7	958.5	2
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	2.7	223.9	860.9	5

4116 rows × 19 columns

In [7]:
```
data.describe()
```

Out[7]:

	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 [8]:
```
data1=data.fillna(data.mode())
```

In [9]:
```
data1
```

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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
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	7.9	196.2	1013.0	3
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	19.3	99.6	1119.5	1
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	60.6	131.1	1057.0	1
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	69.3	76.7	958.5	2
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	2.7	223.9	860.9	5

4116 rows × 19 columns

In [10]:
```
data2=data.groupby(['SUBDIVISION']).count()
data2
```

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
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

SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
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 [14]:
```
datac=data['SUBDIVISION'].unique()
datac
```

Out[14]:
```
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 [16]:
```
data.isna().count()
```

Out[16]:
```
SUBDIVISION    4116
YEAR           4116
JAN            4116
FEB            4116
MAR            4116
APR            4116
MAY            4116
JUN            4116
JUL            4116
AUG            4116
SEP            4116
OCT            4116
NOV            4116
DEC            4116
ANNUAL         4116
Jan-Feb        4116
Mar-May        4116
Jun-Sep        4116
Oct-Dec        4116
dtype: int64
```

In [18]:
```
data.isna().sum()
```

Out[18]:
```
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 [20]:
```
# data3=data.loc[(data.YEAR<=2010)]
# data3
data3=data.loc[(data.JAN>40) & (data.FEB>40)]
data3
```

Out[20]:

	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
13	ANDAMAN & NICOBAR ISLANDS	1915	45.0	56.7	33.3	40.9	170.2	334.7	269.0	317.2	429.8	468.1	258.4	318.0	2741.3	101.7	244.4	1350.7
26	ANDAMAN & NICOBAR ISLANDS	1928	50.9	67.6	80.7	129.3	499.5	410.2	406.3	391.5	404.8	444.5	99.5	13.5	2998.3	118.5	709.5	1612.8
27	ANDAMAN & NICOBAR ISLANDS	1929	74.2	118.4	129.2	69.8	316.6	588.8	134.0	644.7	172.9	413.0	251.5	13.5	2926.6	192.6	515.6	1540.4
28	ANDAMAN & NICOBAR ISLANDS	1930	87.4	105.4	131.2	10.9	231.5	533.6	317.9	446.7	677.2	82.3	249.4	201.6	3075.1	192.8	373.6	1975.4
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4028	LAKSHADWEEP	1928	42.5	66.6	NaN	1.4	212.3	206.9	268.8	174.9	53.8	124.1	116.3	139.3	NaN	109.1	NaN	704.4
4043	LAKSHADWEEP	1943	112.6	57.4	0.0	NaN	146.0	422.3	346.0	203.1	154.8	309.3	142.8	55.5	NaN	170.0	NaN	1126.2
4061	LAKSHADWEEP	1961	60.3	47.4	0.0	31.3	421.1	593.2	297.3	228.2	215.8	148.1	98.8	93.8	2235.3	107.7	452.4	1334.5
4063	LAKSHADWEEP	1963	126.7	112.9	6.5	72.6	204.6	233.0	367.0	249.9	121.8	113.5	211.5	82.1	1902.1	239.6	283.7	971.7
4084	LAKSHADWEEP	1984	71.1	114.9	96.7	65.1	46.7	318.9	271.5	124.9	83.1	218.3	155.4	25.0	1591.6	186.0	208.5	798.4

293 rows × 19 columns

In [33]:
```
data4=data.drop(['ANNUAL','Jan-Feb','Mar-May','Jun-Sep','Oct-Dec'],axis=1)
```

In [34]:
```
data4
```

Out[34]:

	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 [38]:
```
data4['YEAR'].unique()
```

Out[38]:
```
array([1901, 1902, 1903, 1904, 1905, 1906, 1907, 1908, 1910, 1911, 1912,
       1913, 1914, 1915, 1916, 1917, 1918, 1919, 1920, 1921, 1922, 1923,
       1924, 1925, 1926, 1927, 1928, 1929, 1930, 1931, 1932, 1933, 1934,
       1935, 1936, 1937, 1938, 1939, 1940, 1941, 1942, 1946, 1947, 1949,
       1950, 1951, 1952, 1953, 1954, 1955, 1956, 1957, 1958, 1959, 1960,
       1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970, 1971,
       1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981, 1982,
       1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993,
       1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004,
       2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015,
       1943, 1944, 1945, 1948, 1909])
```

In [97]:
```
data6=data1.loc[(data1.SUBDIVISION == "ARUNACHAL PRADESH")]
data6
```

Out[97]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	O D
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	NaN	N
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	2772.8	262
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	4121.3	146
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	2888.0	997
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	2649.2	103
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
202	ARUNACHAL PRADESH	2011	40.0	51.3	174.5	240.8	219.6	288.4	531.4	277.6	286.7	51.9	16.2	15.2	2193.7	91.4	634.9	1384.2	83
203	ARUNACHAL PRADESH	2012	57.8	35.8	134.2	403.4	187.4	645.8	638.9	316.0	724.9	248.1	22.0	26.2	3440.3	93.6	724.9	2325.6	296
204	ARUNACHAL PRADESH	2013	18.5	40.5	115.1	175.1	335.8	290.0	329.6	230.2	316.1	164.1	13.3	14.6	2042.9	59.0	626.0	1165.9	192
205	ARUNACHAL PRADESH	2014	19.0	101.9	80.3	86.7	299.0	415.8	392.4	599.6	343.0	35.1	20.1	10.2	2403.2	120.9	466.0	1750.8	65
206	ARUNACHAL PRADESH	2015	30.8	47.5	97.5	287.1	238.9	637.9	329.3	595.5	374.2	65.2	33.8	29.8	2767.5	78.3	623.5	1936.9	128

97 rows × 19 columns

In [41]:
```
data5.isna().sum()
```

Out[41]:
```
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
```

In [42]:
```
data6=data5.fillna(data5.mean())
data6
```

```
/tmp/ipykernel_4143/1463832817.py:1: FutureWarning: The default value of numeric_only in DataFrame.mean is
deprecated. In a future version, it will default to False. In addition, specifying 'numeric_only=None' is d
eprecated. Select only valid columns or specify the value of numeric_only to silence this warning.
  data6=data5.fillna(data5.mean())
```

Out[42]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUA
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.100000	316.1	424.6	1124.9	694.544792	629.7	333.9	194.686316	35.696842	24.502105	3418.8571
111	ARUNACHAL PRADESH	1917	21.4	164.5	153.527368	269.6	107.9	823.8	909.100000	628.4	411.5	199.300000	63.500000	0.000000	3418.8571
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.200000	144.6	861.1	1609.9	1303.000000	692.6	515.8	125.200000	7.800000	13.700000	5486.3000
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.500000	256.9	420.6	973.6	999.000000	286.7	628.7	948.300000	40.700000	8.600000	4693.9000
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.600000	364.7	173.6	840.6	535.400000	896.5	376.7	103.300000	0.000000	0.000000	4106.7000
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
202	ARUNACHAL PRADESH	2011	40.0	51.3	174.500000	240.8	219.6	288.4	531.400000	277.6	286.7	51.900000	16.200000	15.200000	2193.7000
203	ARUNACHAL PRADESH	2012	57.8	35.8	134.200000	403.4	187.4	645.8	638.900000	316.0	724.9	248.100000	22.000000	26.200000	3440.3000
204	ARUNACHAL PRADESH	2013	18.5	40.5	115.100000	175.1	335.8	290.0	329.600000	230.2	316.1	164.100000	13.300000	14.600000	2042.9000
205	ARUNACHAL PRADESH	2014	19.0	101.9	80.300000	86.7	299.0	415.8	392.400000	599.6	343.0	35.100000	20.100000	10.200000	2403.2000
206	ARUNACHAL PRADESH	2015	30.8	47.5	97.500000	287.1	238.9	637.9	329.300000	595.5	374.2	65.200000	33.800000	29.800000	2767.5000

97 rows × 19 columns

In [47]:
```
# data7=data5.fillna(data5.mode())
# data7

data7=data5.fillna(data5.median())
data7
```

```
/tmp/ipykernel_4143/4005783182.py:4: FutureWarning: The default value of numeric_only in DataFrame.median i
s deprecated. In a future version, it will default to False. In addition, specifying 'numeric_only=None' is
deprecated. Select only valid columns or specify the value of numeric_only to silence this warning.
  data7=data5.fillna(data5.median())
```

Out[47]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	609.35	629.7	333.9	164.1	25.6	15.9	3354.2	117.9	811.8	2185.1	22
111	ARUNACHAL PRADESH	1917	21.4	164.5	139.9	269.6	107.9	823.8	909.10	628.4	411.5	199.3	63.5	0.0	3354.2	185.9	727.9	2772.8	26
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.00	692.6	515.8	125.2	7.8	13.7	5486.3	21.4	1196.9	4121.3	14
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.00	286.7	628.7	948.3	40.7	8.6	4693.9	102.3	706.0	2888.0	99
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.40	896.5	376.7	103.3	0.0	0.0	4106.7	210.3	1143.9	2649.2	10
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
202	ARUNACHAL PRADESH	2011	40.0	51.3	174.5	240.8	219.6	288.4	531.40	277.6	286.7	51.9	16.2	15.2	2193.7	91.4	634.9	1384.2	8
203	ARUNACHAL PRADESH	2012	57.8	35.8	134.2	403.4	187.4	645.8	638.90	316.0	724.9	248.1	22.0	26.2	3440.3	93.6	724.9	2325.6	29
204	ARUNACHAL PRADESH	2013	18.5	40.5	115.1	175.1	335.8	290.0	329.60	230.2	316.1	164.1	13.3	14.6	2042.9	59.0	626.0	1165.9	19
205	ARUNACHAL PRADESH	2014	19.0	101.9	80.3	86.7	299.0	415.8	392.40	599.6	343.0	35.1	20.1	10.2	2403.2	120.9	466.0	1750.8	6
206	ARUNACHAL PRADESH	2015	30.8	47.5	97.5	287.1	238.9	637.9	329.30	595.5	374.2	65.2	33.8	29.8	2767.5	78.3	623.5	1936.9	12

97 rows × 19 columns

In [50]:
```
data4
```

Out[50]:

	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 [55]:
```
data4['ANNUAL RAIN']=data4.apply(lambda row:row.JAN+row.FEB+row.MAR+row.APR+row.MAY+row.JUN+row.JUL+row.AUG+
```

In [56]:
```
data4
```

Out[56]:

	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 [57]:
```
data5=data4.drop(columns='SUBDIVISION')
```

In [58]:
```
data5
```

Out[58]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN
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
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
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
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
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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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
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
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
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
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

4116 rows × 14 columns

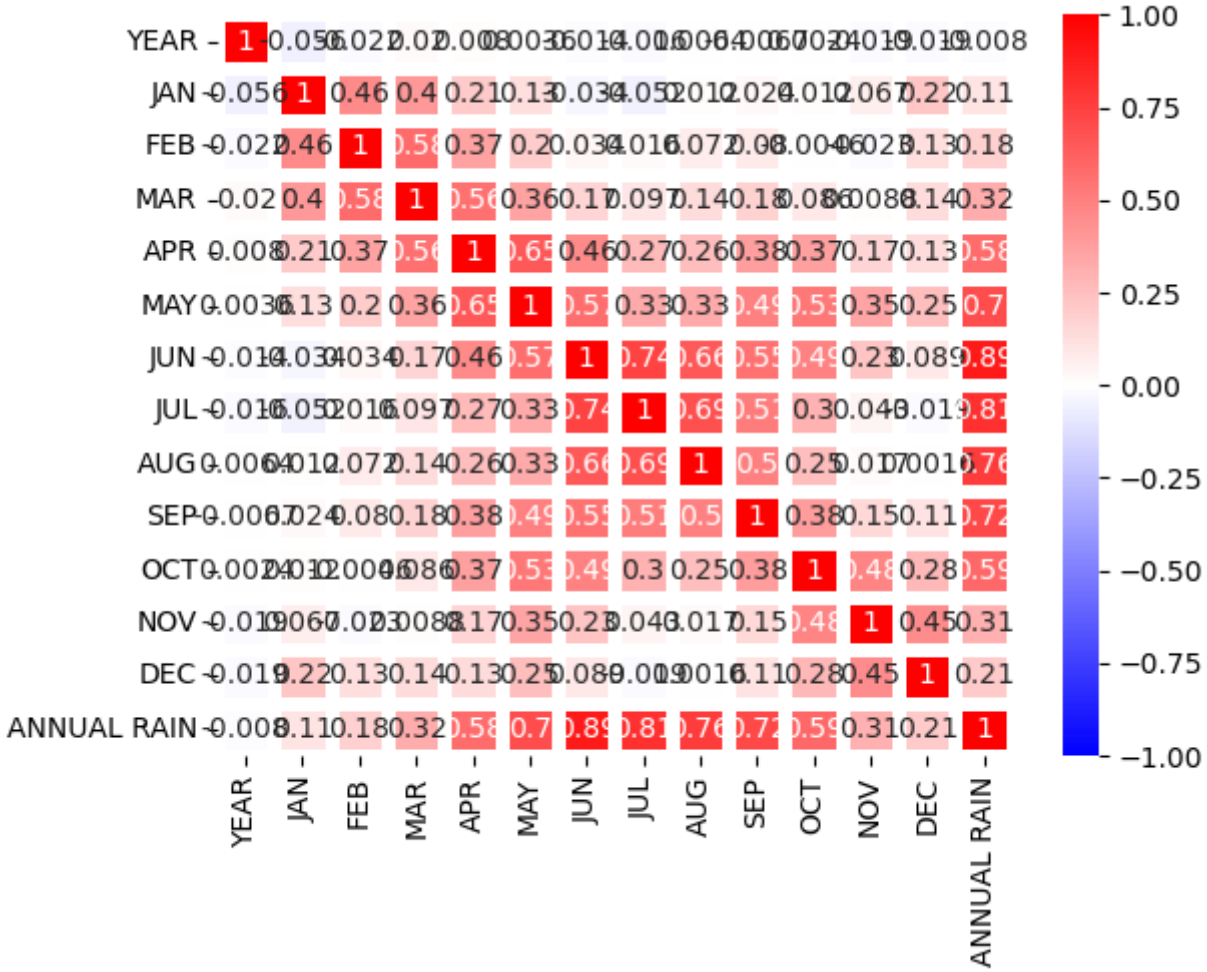
In [59]:
```
cor=data5.corr()
cor
```

Out[59]:

	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 RAIN	-0.008041	0.105697	0.181561	0.322201	0.577574	0.698014	0.891302	0.812279	0.759303	0.715133	0.587066	0.308771	0.207178

In [70]:
```
import seaborn as sns
sns.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidths=5,cmap='bwr')
```

Out[70]: <Axes: >



In [71]:
```
data5
```

Out[71]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN
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
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
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
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
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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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
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
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
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
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

4116 rows × 14 columns

In [77]:
```
data5['SWM']=data5.apply(lambda row:row.JUN+row.JUL+row.AUG+row.SEP,axis=1)
```

In [78]:
```
data5
```

Out[78]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN	SWM
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	1696.3
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	2185.9
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	1874.0
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	1977.6
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	1624.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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	1013.0
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	1119.5
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	1057.0
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	958.5
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	860.9

4116 rows × 15 columns

In [79]:
```
data5['NEM']=data5['OCT']+data5['NOV']+data5['DEC']   #column addition
data5
```

Out[79]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN	SWM	NEM
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	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	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	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	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	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	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	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	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	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	860.9	555.4

4116 rows × 16 columns

In [92]:
```
data6=data5.drop(['JAN','FEB','MAR','APR','MAY','JUN','JUL','AUG','SEP','OCT','NOV','DEC'],axis=1)
data6
```

Out[92]:

	YEAR	ANNUAL RAIN	SWM	NEM
0	1901	3373.2	1696.3	980.3
1	1902	3520.7	2185.9	716.7
2	1903	2957.4	1874.0	690.6
3	1904	3079.6	1977.6	571.0
4	1905	2566.7	1624.9	630.8
...	...	...	...	...
4111	2011	1533.7	1013.0	316.6
4112	2012	1405.5	1119.5	167.1
4113	2013	1426.3	1057.0	177.6
4114	2014	1395.0	958.5	290.5
4115	2015	1642.9	860.9	555.4

4116 rows × 4 columns

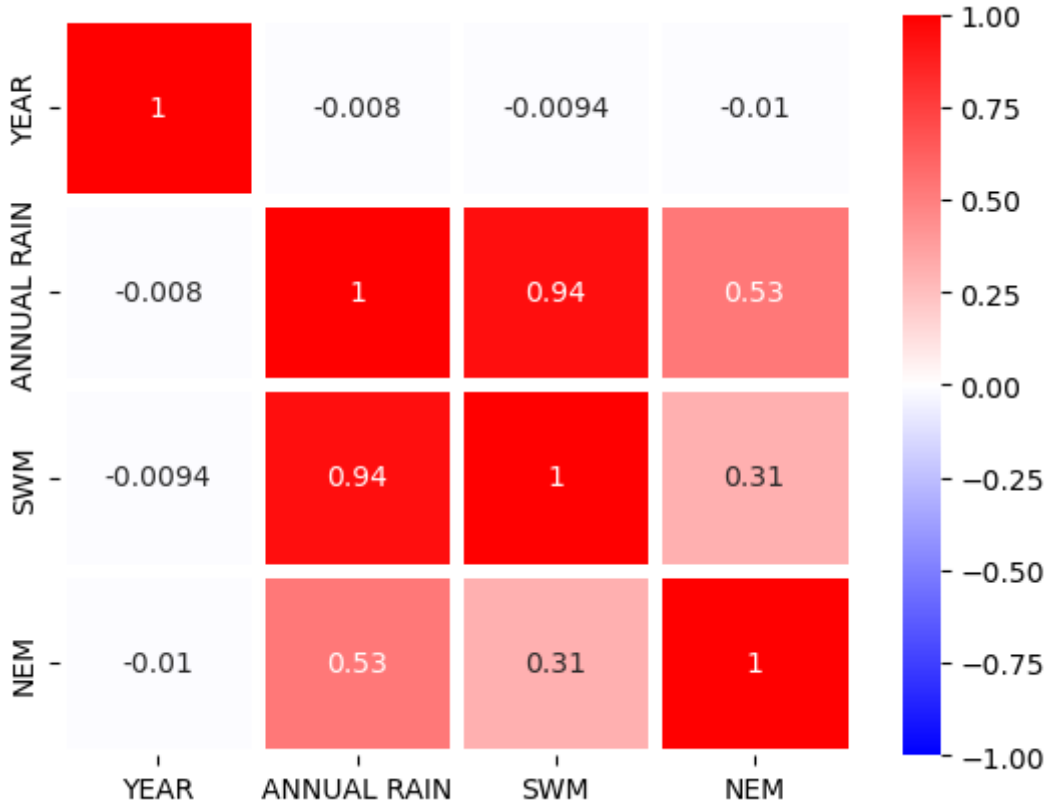
In [93]:
```
cor1=data6.corr()
cor1
```

Out[93]:

	YEAR	ANNUAL RAIN	SWM	NEM
YEAR	1.000000	-0.008041	-0.009414	-0.010152
ANNUAL RAIN	-0.008041	1.000000	0.943660	0.529596
SWM	-0.009414	0.943660	1.000000	0.310432
NEM	-0.010152	0.529596	0.310432	1.000000

In [94]:
```
import seaborn as sns
sns.heatmap(cor1,vmax=1,vmin=-1,annot=True,linewidths=5,cmap='bwr')
```

Out[94]: <Axes: >



In [96]:
```
data4
```

Out[96]:

	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 [98]:
```
data6
```

Out[98]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	O D
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	NaN	N
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	2772.8	262
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	4121.3	146
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	2888.0	997
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	2649.2	103
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
202	ARUNACHAL PRADESH	2011	40.0	51.3	174.5	240.8	219.6	288.4	531.4	277.6	286.7	51.9	16.2	15.2	2193.7	91.4	634.9	1384.2	83
203	ARUNACHAL PRADESH	2012	57.8	35.8	134.2	403.4	187.4	645.8	638.9	316.0	724.9	248.1	22.0	26.2	3440.3	93.6	724.9	2325.6	296
204	ARUNACHAL PRADESH	2013	18.5	40.5	115.1	175.1	335.8	290.0	329.6	230.2	316.1	164.1	13.3	14.6	2042.9	59.0	626.0	1165.9	192
205	ARUNACHAL PRADESH	2014	19.0	101.9	80.3	86.7	299.0	415.8	392.4	599.6	343.0	35.1	20.1	10.2	2403.2	120.9	466.0	1750.8	65
206	ARUNACHAL PRADESH	2015	30.8	47.5	97.5	287.1	238.9	637.9	329.3	595.5	374.2	65.2	33.8	29.8	2767.5	78.3	623.5	1936.9	128

97 rows × 19 columns

In [99]:
```
data7
```

Out[99]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	609.35	629.7	333.9	164.1	25.6	15.9	3354.2	117.9	811.8	2185.1	22
111	ARUNACHAL PRADESH	1917	21.4	164.5	139.9	269.6	107.9	823.8	909.10	628.4	411.5	199.3	63.5	0.0	3354.2	185.9	727.9	2772.8	26
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.00	692.6	515.8	125.2	7.8	13.7	5486.3	21.4	1196.9	4121.3	14
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.00	286.7	628.7	948.3	40.7	8.6	4693.9	102.3	706.0	2888.0	99
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.40	896.5	376.7	103.3	0.0	0.0	4106.7	210.3	1143.9	2649.2	10
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
202	ARUNACHAL PRADESH	2011	40.0	51.3	174.5	240.8	219.6	288.4	531.40	277.6	286.7	51.9	16.2	15.2	2193.7	91.4	634.9	1384.2	8
203	ARUNACHAL PRADESH	2012	57.8	35.8	134.2	403.4	187.4	645.8	638.90	316.0	724.9	248.1	22.0	26.2	3440.3	93.6	724.9	2325.6	29
204	ARUNACHAL PRADESH	2013	18.5	40.5	115.1	175.1	335.8	290.0	329.60	230.2	316.1	164.1	13.3	14.6	2042.9	59.0	626.0	1165.9	19
205	ARUNACHAL PRADESH	2014	19.0	101.9	80.3	86.7	299.0	415.8	392.40	599.6	343.0	35.1	20.1	10.2	2403.2	120.9	466.0	1750.8	6
206	ARUNACHAL PRADESH	2015	30.8	47.5	97.5	287.1	238.9	637.9	329.30	595.5	374.2	65.2	33.8	29.8	2767.5	78.3	623.5	1936.9	12

97 rows × 19 columns

In [102]:
```
data7=data7.drop(columns=['Jan-Feb','Mar-May','Jun-Sep','Oct-Dec'])
```

In [103]:
```
data7
```

Out[103]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	609.35	629.7	333.9	164.1	25.6	15.9	3354.2
111	ARUNACHAL PRADESH	1917	21.4	164.5	139.9	269.6	107.9	823.8	909.10	628.4	411.5	199.3	63.5	0.0	3354.2
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.00	692.6	515.8	125.2	7.8	13.7	5486.3
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.00	286.7	628.7	948.3	40.7	8.6	4693.9
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.40	896.5	376.7	103.3	0.0	0.0	4106.7
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
202	ARUNACHAL PRADESH	2011	40.0	51.3	174.5	240.8	219.6	288.4	531.40	277.6	286.7	51.9	16.2	15.2	2193.7
203	ARUNACHAL PRADESH	2012	57.8	35.8	134.2	403.4	187.4	645.8	638.90	316.0	724.9	248.1	22.0	26.2	3440.3
204	ARUNACHAL PRADESH	2013	18.5	40.5	115.1	175.1	335.8	290.0	329.60	230.2	316.1	164.1	13.3	14.6	2042.9
205	ARUNACHAL PRADESH	2014	19.0	101.9	80.3	86.7	299.0	415.8	392.40	599.6	343.0	35.1	20.1	10.2	2403.2
206	ARUNACHAL PRADESH	2015	30.8	47.5	97.5	287.1	238.9	637.9	329.30	595.5	374.2	65.2	33.8	29.8	2767.5

97 rows × 15 columns

In [104]:
```
data7['SWM']=data7.apply(lambda row:row.JUN+row.JUL+row.AUG+row.SEP,axis=1)
```

In [105]:
```
data7['NEM']=data7['OCT']+data7['NOV']+data7['DEC']
data7
```

Out[105]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	SWM	NEM
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	609.35	629.7	333.9	164.1	25.6	15.9	3354.2	2697.85	205.6
111	ARUNACHAL PRADESH	1917	21.4	164.5	139.9	269.6	107.9	823.8	909.10	628.4	411.5	199.3	63.5	0.0	3354.2	2772.80	262.8
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.00	692.6	515.8	125.2	7.8	13.7	5486.3	4121.30	146.7
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.00	286.7	628.7	948.3	40.7	8.6	4693.9	2888.00	997.6
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.40	896.5	376.7	103.3	0.0	0.0	4106.7	2649.20	103.3
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
202	ARUNACHAL PRADESH	2011	40.0	51.3	174.5	240.8	219.6	288.4	531.40	277.6	286.7	51.9	16.2	15.2	2193.7	1384.10	83.3
203	ARUNACHAL PRADESH	2012	57.8	35.8	134.2	403.4	187.4	645.8	638.90	316.0	724.9	248.1	22.0	26.2	3440.3	2325.60	296.3
204	ARUNACHAL PRADESH	2013	18.5	40.5	115.1	175.1	335.8	290.0	329.60	230.2	316.1	164.1	13.3	14.6	2042.9	1165.90	192.0
205	ARUNACHAL PRADESH	2014	19.0	101.9	80.3	86.7	299.0	415.8	392.40	599.6	343.0	35.1	20.1	10.2	2403.2	1750.80	65.4
206	ARUNACHAL PRADESH	2015	30.8	47.5	97.5	287.1	238.9	637.9	329.30	595.5	374.2	65.2	33.8	29.8	2767.5	1936.90	128.8

97 rows × 17 columns

In [106]:
```
data8=data7.drop(['SUBDIVISION','JAN','FEB','MAR','APR','MAY','JUN','JUL','AUG','SEP','OCT','NOV','DEC'],axi
data8
```

Out[106]:

	YEAR	ANNUAL	SWM	NEM
110	1916	3354.2	2697.85	205.6
111	1917	3354.2	2772.80	262.8
112	1918	5486.3	4121.30	146.7
113	1919	4693.9	2888.00	997.6
114	1920	4106.7	2649.20	103.3
...	...	...	...	...
202	2011	2193.7	1384.10	83.3
203	2012	3440.3	2325.60	296.3
204	2013	2042.9	1165.90	192.0
205	2014	2403.2	1750.80	65.4
206	2015	2767.5	1936.90	128.8

97 rows × 4 columns

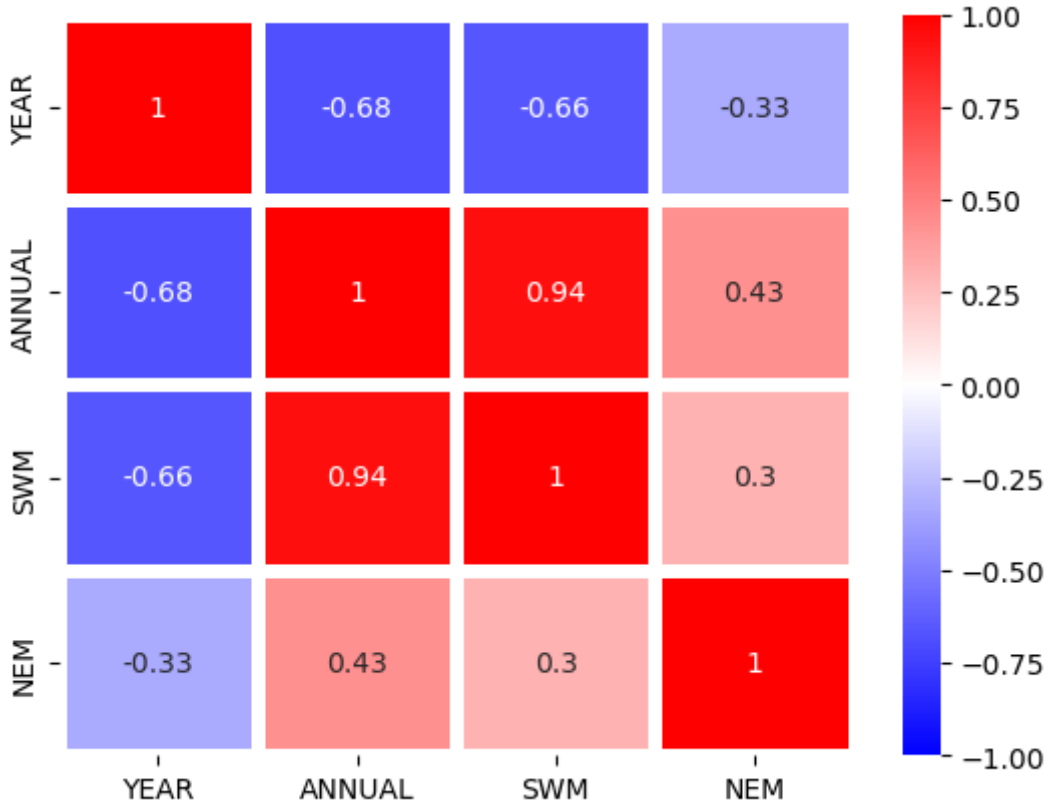
In [107]:
```
cor=data8.corr()
cor
```

Out[107]:

	YEAR	ANNUAL	SWM	NEM
YEAR	1.000000	-0.682087	-0.661699	-0.333263
ANNUAL	-0.682087	1.000000	0.941847	0.430912
SWM	-0.661699	0.941847	1.000000	0.302315
NEM	-0.333263	0.430912	0.302315	1.000000

In [108]:
```
import seaborn as sns
sns.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidths=5,cmap='bwr')
```

Out[108]: <Axes: >



In [ ]:
```
-
```