In [124]:
```
import pandas as pd
```

Reading csv file

In [125]:
```
data=pd.read_csv("/home/placement/Downloads/rainfall in india 1901-2015.csv")
```

In [126]:
```
data.describe()
```

Out[126]:

	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

Display top 5 rows default

In [127]:
```
data.head()
```

Out[127]:

	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 [128]:
```
data
```

Out[128]:

	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 [129]:
```
data.groupby(["SUBDIVISION"]).count()#count the subdivision
```

Out[129]:

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 [130]:
```
data.isna().sum()#duplicate values  are displayed
```

Out[130]:
```
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 [131]:
```
data1=data.loc[(data.YEAR<=2022)]#getting data greater than 2022
```

In [132]:
```
data1
```

Out[132]:

	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 [133]:
```
data1.tail(60)#bottom below 60 rows
```

Out[133]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep
4056	LAKSHADWEEP	1956	13.6	27.7	1.3	40.8	269.5	356.1	109.8	91.6	148.0	175.6	48.5	4.9	1287.4	41.3	311.6	705.5
4057	LAKSHADWEEP	1957	0.9	11.4	34.6	23.7	195.6	309.9	129.1	149.6	47.5	215.2	198.5	31.2	1347.2	12.3	253.9	636.1
4058	LAKSHADWEEP	1958	0.2	1.6	0.0	8.0	116.0	269.4	84.5	238.2	54.0	133.4	51.1	36.2	992.6	1.8	124.0	646.1
4059	LAKSHADWEEP	1959	15.0	0.0	8.0	51.6	320.5	337.7	205.8	151.2	183.2	183.8	97.8	29.0	1583.6	15.0	380.1	877.9
4060	LAKSHADWEEP	1960	29.6	40.0	3.3	85.4	497.1	176.8	257.3	168.4	221.3	169.7	235.9	17.0	1901.8	69.6	585.8	823.8
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
4062	LAKSHADWEEP	1962	11.1	0.0	0.0	81.5	331.2	149.3	204.5	286.7	201.8	210.9	33.5	76.1	1586.6	11.1	412.7	842.3
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
4064	LAKSHADWEEP	1964	2.5	3.5	5.2	40.4	64.2	376.1	407.0	174.6	342.5	120.9	90.8	24.0	1651.7	6.0	109.8	1300.2
4065	LAKSHADWEEP	1965	0.2	7.3	8.3	35.0	162.5	200.6	164.2	267.7	92.4	69.2	77.9	320.6	1405.9	7.5	205.8	724.9
4066	LAKSHADWEEP	1966	21.3	0.2	6.2	4.4	34.9	376.7	421.9	158.9	225.4	266.1	158.0	67.5	1741.5	21.5	45.5	1182.9
4067	LAKSHADWEEP	1967	7.8	2.7	0.0	0.0	148.3	301.2	380.1	196.9	100.4	92.3	24.5	33.1	1287.3	10.5	148.3	978.6
4068	LAKSHADWEEP	1968	8.7	1.8	30.6	45.4	16.6	393.8	439.8	84.1	202.8	46.5	31.4	42.9	1344.4	10.5	92.6	1120.5
4069	LAKSHADWEEP	1969	30.0	4.1	0.0	63.3	237.9	125.6	149.0	225.3	95.9	162.3	157.5	165.2	1416.1	34.1	301.2	595.8
4070	LAKSHADWEEP	1970	18.7	6.3	13.8	32.3	164.9	322.1	484.4	284.6	228.9	126.9	81.2	18.9	1783.0	25.0	211.0	1320.0
4071	LAKSHADWEEP	1971	3.8	12.3	5.4	25.0	271.6	333.9	308.1	349.6	281.8	113.8	63.0	133.1	1901.4	16.1	302.0	1273.4
4072	LAKSHADWEEP	1972	0.0	1.1	0.0	11.4	81.3	349.2	197.3	172.5	152.0	219.2	94.2	9.4	1287.6	1.1	92.7	871.0
4073	LAKSHADWEEP	1973	0.3	5.8	0.0	40.7	96.3	260.5	299.9	437.9	57.3	150.9	108.1	99.7	1557.4	6.1	137.0	1055.6
4074	LAKSHADWEEP	1974	0.0	16.8	0.8	35.4	171.9	277.6	491.6	165.7	258.3	65.9	20.1	7.2	1511.3	16.8	208.1	1193.2
4075	LAKSHADWEEP	1975	8.6	0.7	18.4	92.9	282.2	250.0	278.0	348.4	299.5	111.6	155.2	45.0	1890.5	9.3	393.5	1175.9
4076	LAKSHADWEEP	1976	10.3	0.0	1.1	72.9	85.3	260.7	329.2	344.1	36.6	189.2	161.6	11.1	1502.1	10.3	159.3	970.6
4077	LAKSHADWEEP	1977	1.4	38.2	16.5	27.6	256.8	460.7	337.4	125.3	90.7	248.4	243.8	1.4	1848.2	39.6	300.9	1014.1
4078	LAKSHADWEEP	1978	0.2	6.5	0.6	44.1	309.5	568.3	224.7	303.5	105.1	92.1	294.2	18.6	1967.4	6.7	354.2	1201.6
4079	LAKSHADWEEP	1979	14.9	1.3	10.0	15.4	23.1	367.1	451.3	127.3	221.1	173.6	378.1	42.2	1825.4	16.2	48.5	1166.8

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep
4080	LAKSHADWEEP	1980	7.5	10.5	13.1	19.7	123.5	233.6	185.0	217.6	67.0	122.1	98.5	113.6	1211.7	18.0	156.3	703.2
4081	LAKSHADWEEP	1981	1.8	20.4	0.6	22.3	208.6	421.5	168.9	213.3	281.4	167.4	76.6	34.6	1617.4	22.2	231.5	1085.1
4082	LAKSHADWEEP	1982	1.6	0.4	0.6	0.4	106.8	321.6	254.5	326.1	121.3	70.7	178.6	9.4	1392.0	2.0	107.8	1023.5
4083	LAKSHADWEEP	1983	1.4	0.0	0.0	0.2	113.5	215.0	277.0	375.0	219.8	77.6	71.2	52.5	1403.2	1.4	113.7	1086.8
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
4085	LAKSHADWEEP	1985	6.4	0.0	27.4	11.1	182.9	326.1	152.1	203.0	206.4	42.0	62.3	83.7	1303.4	6.4	221.4	887.6
4086	LAKSHADWEEP	1986	12.4	3.0	25.3	6.5	102.1	399.2	193.3	218.3	179.9	128.6	147.8	49.5	1465.9	15.4	133.9	990.7
4087	LAKSHADWEEP	1987	1.9	1.5	0.0	16.5	72.9	389.3	29.4	284.7	183.4	133.0	47.7	128.3	1288.6	3.4	89.4	886.8
4088	LAKSHADWEEP	1988	0.8	7.0	0.9	54.0	142.8	352.5	293.6	240.9	240.9	28.9	82.7	5.8	1450.8	7.8	197.7	1127.9
4089	LAKSHADWEEP	1989	20.6	0.0	4.4	93.0	106.7	473.5	387.5	167.3	216.3	172.4	69.2	3.5	1714.4	20.6	204.1	1244.6
4090	LAKSHADWEEP	1990	38.9	0.3	21.4	0.0	191.1	181.6	334.0	123.5	98.9	160.4	155.4	5.0	1310.5	39.2	212.5	738.0
4091	LAKSHADWEEP	1991	12.3	0.0	18.7	12.3	68.0	604.3	241.1	253.8	45.7	222.7	75.5	28.6	1583.0	12.3	99.0	1144.9
4092	LAKSHADWEEP	1992	4.0	0.1	0.0	3.3	128.1	346.6	363.0	373.3	95.1	120.4	69.3	31.7	1534.9	4.1	131.4	1178.0
4093	LAKSHADWEEP	1993	1.2	0.5	0.2	0.2	56.5	276.1	346.7	154.4	161.0	131.6	280.5	40.8	1449.7	1.7	56.9	938.2
4094	LAKSHADWEEP	1994	12.4	66.6	34.8	88.8	78.9	361.1	240.2	219.7	76.2	213.5	153.7	1.2	1547.1	79.0	202.5	897.2
4095	LAKSHADWEEP	1995	131.3	18.5	0.3	315.4	179.6	286.0	486.7	384.8	71.8	81.0	72.2	3.3	2030.9	149.8	495.3	1229.3
4096	LAKSHADWEEP	1996	44.7	1.1	1.6	17.4	50.0	427.1	335.3	197.3	230.4	109.0	60.5	131.6	1606.0	45.8	69.0	1190.1
4097	LAKSHADWEEP	1997	2.2	0.1	4.9	33.8	62.3	307.0	459.6	216.8	144.0	213.5	200.8	119.7	1764.7	2.3	101.0	1127.4
4098	LAKSHADWEEP	1998	52.0	0.0	1.8	40.3	68.2	382.0	388.8	196.7	274.7	184.8	144.1	253.5	1986.9	52.0	110.3	1242.2
4099	LAKSHADWEEP	1999	47.8	2.5	18.3	20.6	416.7	279.6	459.4	133.8	73.4	305.0	51.2	49.0	1857.3	50.3	455.6	946.2
4100	LAKSHADWEEP	2000	83.3	18.9	3.4	47.9	204.6	225.4	95.5	319.9	164.5	141.4	56.3	11.0	1372.1	102.2	255.9	805.3
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
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
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
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

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

In [134]:
```
data2=data1.drop(["ANNUAL","Jan-Feb","Mar-May","Jun-Sep","Oct-Dec"],axis=1)#delete the columns
```

In [135]:
```
data2
```

Out[135]:

	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 [136]:
```
data2['SUBDIVISION'].unique()#unique subdivision names can be printed in array
```

Out[136]:
```
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 [137]:
```
data2=data2.loc[(data2.SUBDIVISION=="ARUNACHAL PRADESH")]#only arunachal data can be printed
```

In [138]:
```
data2
```

Out[138]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	NaN	NaN	NaN
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
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
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
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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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
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
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
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
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

97 rows × 14 columns

In [139]:
```
data3.isna().sum()#duplicate values  are displayed
```

Out[139]:
```
SUBDIVISION    0
YEAR           0
ANNUAL RAIN    6
dtype: int64
```

In [140]:
```
data2['ANNUAL RAIN']=data2.apply(lambda row:row.JAN + row.FEB+ row.MAR+ row.APR+ row.JUN+ row.JUL+ row.AUG+
```

In [141]:
```
data2.head(10)
```

Out[141]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN
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
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
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	4625.2
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	4273.3
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	3933.1
115	ARUNACHAL PRADESH	1921	78.9	54.3	180.3	358.0	598.0	1233.2	1433.0	885.9	603.4	246.3	4.6	15.5	5093.4
116	ARUNACHAL PRADESH	1922	50.7	59.4	170.4	299.5	350.5	1109.3	918.7	488.3	207.6	483.5	30.3	19.0	3836.7
117	ARUNACHAL PRADESH	1923	9.4	160.8	34.0	240.9	445.4	408.6	1278.5	251.3	617.3	50.1	8.4	2.8	3062.1
118	ARUNACHAL PRADESH	1924	85.7	45.1	74.1	162.4	515.7	1165.0	942.7	713.8	410.8	303.3	31.9	0.0	3934.8
119	ARUNACHAL PRADESH	1925	80.6	114.0	143.3	223.0	587.2	611.6	611.0	684.7	1222.0	153.2	5.1	4.1	3852.6

In [142]:
```
data3=data2.drop(["SUBDIVISION"],axis=1)#delete a column
```

In [143]:
```
data3
```

Out[143]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN
110	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	NaN	NaN	NaN	NaN
111	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
112	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	4625.2
113	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	4273.3
114	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	3933.1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
202	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	1974.0
203	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	3253.1
204	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	1707.1
205	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	2104.1
206	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	2528.6

97 rows × 14 columns

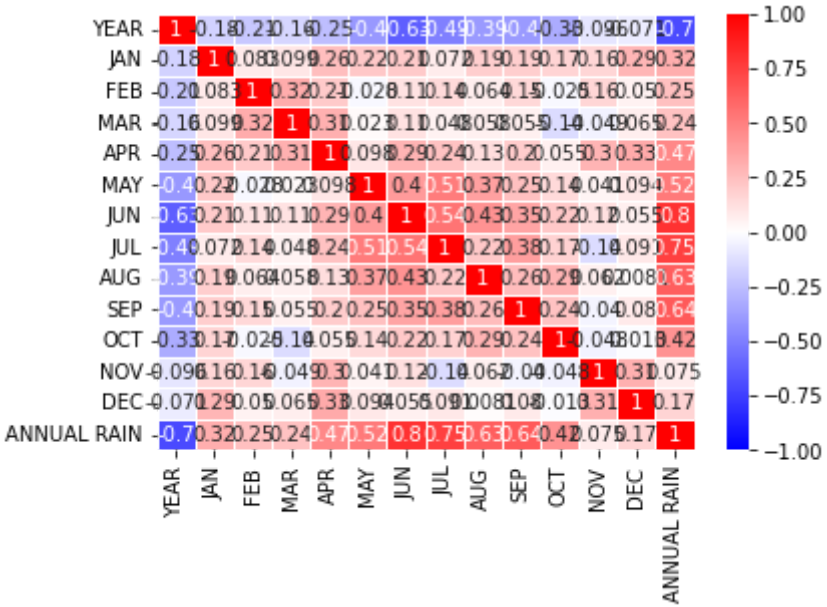
In [144]:
```
cor_mat=data3.corr()#correlation
cor_mat
```

Out[144]:

	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 RAIN	-0.703229	0.320120	0.247846	0.242045	0.465778	0.518707	0.804169	0.746564	0.627006	0.644468	0.422273	0.074708	0.171

In [145]:
```
import seaborn as reddy#correlation by using graph
reddy.heatmap(cor_mat,vmax=1,vmin=-1,annot=True,linewidth=.5,cmap='bwr')
```

Out[145]: <AxesSubplot:>



In [146]:
```
data3['nem']=data2.apply(lambda row:row.OCT+ row.NOV+ row.DEC,axis=1)#adding rows
```

In [147]:
```
data3
```

Out[147]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN	nem
110	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	NaN	NaN	NaN	NaN	NaN
111	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	262.8
112	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	4625.2	146.7
113	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	4273.3	997.6
114	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	3933.1	103.3
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
202	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	1974.0	83.3
203	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	3253.1	296.3
204	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	1707.1	192.0
205	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	2104.1	65.4
206	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	2528.6	128.8

97 rows × 15 columns

In [148]:
```
data3['swm']=data2.apply(lambda row: row.JUN+ row.JUL+ row.AUG+ row.SEP,axis=1)#adding columns
```

In [149]:
```
data3
```

Out[149]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN	nem	swm
110	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	NaN	NaN	NaN	NaN	NaN	NaN
111	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	262.8	2772.8
112	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	4625.2	146.7	4121.3
113	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	4273.3	997.6	2888.0
114	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	3933.1	103.3	2649.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
202	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	1974.0	83.3	1384.1
203	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	3253.1	296.3	2325.6
204	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	1707.1	192.0	1165.9
205	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	2104.1	65.4	1750.8
206	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	2528.6	128.8	1936.9

97 rows × 16 columns

In [150]:
```
data3=data3.drop(["JAN","FEB","MAR","APR","MAY","JUN","JUL","AUG","SEP","OCT","NOV","DEC"],axis=1)#drop colu
```

In [151]:
```
data3
```

Out[151]:

	YEAR	ANNUAL RAIN	nem	swm
110	1916	NaN	NaN	NaN
111	1917	NaN	262.8	2772.8
112	1918	4625.2	146.7	4121.3
113	1919	4273.3	997.6	2888.0
114	1920	3933.1	103.3	2649.2
...	...	...	...	...
202	2011	1974.0	83.3	1384.1
203	2012	3253.1	296.3	2325.6
204	2013	1707.1	192.0	1165.9
205	2014	2104.1	65.4	1750.8
206	2015	2528.6	128.8	1936.9

97 rows × 4 columns

In [157]:
```
data4=data3.drop(["YEAR"],axis=1)#drop column
```

In [158]:
```
data4
```

Out[158]:

	ANNUAL RAIN	nem	swm
110	NaN	NaN	NaN
111	NaN	262.8	2772.8
112	4625.2	146.7	4121.3
113	4273.3	997.6	2888.0
114	3933.1	103.3	2649.2
...	...	...	...
202	1974.0	83.3	1384.1
203	3253.1	296.3	2325.6
204	1707.1	192.0	1165.9
205	2104.1	65.4	1750.8
206	2528.6	128.8	1936.9

97 rows × 3 columns

In [159]:
```
cor_mat=data4.corr()#correlation
cor_mat
```

Out[159]:

	ANNUAL RAIN	nem	swm
ANNUAL RAIN	1.000000	0.451615	0.970880
nem	0.451615	1.000000	0.320171
swm	0.970880	0.320171	1.000000

In [160]:
```
import seaborn as reddy#correlation by using graph
reddy.heatmap(cor_mat,vmax=1,vmin=-1,annot=True,linewidth=.5,cmap='bwr')
```

Out[160]: <AxesSubplot:>