

```
In [3]: import pandas as pd
import numpy as np
import warnings
warnings.filterwarnings("ignore")# ignore the warnings
```

```
In [4]: data=pd.read_csv("/home/placement/Desktop/nio/rainfall in india 1901-2015.csv")# read thed data
```

In [15]: data # describing the total data

Out[15]:

	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

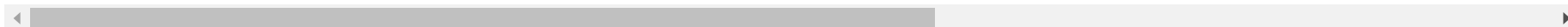
4116 rows × 19 columns



In [16]: data.describe()*# describing the count and VALUES*

Out[16]:

	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	4109.0
mean	1958.218659	18.957320	21.805325	27.359197	43.127432	85.745417	230.234444	347.214334	290.263497	197.361922	95.1
std	33.140898	33.585371	35.909488	46.959424	67.831168	123.234904	234.710758	269.539667	188.770477	135.408345	99.1
min	1901.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.400000	0.000000	0.000000	0.100000	0.0
25%	1930.000000	0.600000	0.600000	1.000000	3.000000	8.600000	70.350000	175.600000	155.975000	100.525000	14.0
50%	1958.000000	6.000000	6.700000	7.800000	15.700000	36.600000	138.700000	284.800000	259.400000	173.900000	65.1
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	948.1



In []:

In [17]: `data.head()` *# PRINTING THE FIRST 5 VALUES*

Out[17]:

	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 [18]: data.isna().sum()# PRINT THE NUMBER OF NULL VALUES IN ANY COL
```

```
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 [19]: `data.groupby(['SUBDIVISION']).count()` *# it tells us about the count of the years the data is being taken*

Out[19]:

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

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
SUBDIVISION																		
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]: data1=data.loc[(data.YEAR>=1910)&(data.YEAR<=2010)] # MODIFY THE DTAT ACCORDING TO OUR CONDITION
```

In [7]: data1

Out[7]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	
8	ANDAMAN & NICOBAR ISLANDS	1910	26.6	22.7	206.3	89.3	224.5	472.7	264.3	337.4	626.6	208.2	267.3	153.5	2899.4	49.3	520.1	1701.0	6
9	ANDAMAN & NICOBAR ISLANDS	1911	0.0	8.4	0.0	122.5	327.3	649.0	253.0	187.1	464.5	333.8	94.5	247.1	2687.2	8.4	449.8	1553.6	6
10	ANDAMAN & NICOBAR ISLANDS	1912	583.7	0.8	0.0	21.9	140.7	549.8	468.9	370.3	386.2	318.7	117.2	2.3	2960.5	584.5	162.6	1775.2	4
11	ANDAMAN & NICOBAR ISLANDS	1913	84.8	0.5	1.3	2.5	190.7	530.0	280.8	205.8	580.1	288.8	133.0	67.5	2365.8	85.3	194.5	1596.7	4
12	ANDAMAN & NICOBAR ISLANDS	1914	0.0	0.0	0.0	37.7	298.8	383.3	792.8	520.5	310.8	139.8	184.4	289.7	2957.8	0.0	336.5	2007.4	6
...
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	2
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	4
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	4
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	3
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	4

3622 rows × 19 columns




```
In [8]: data0=data1.drop(['ANNUAL',  
    'Jan-Feb',  
    'Mar-May',  
    'Jun-Sep',  
    'Oct-Dec'],axis=1)# REMOVING UNWANTED COLS
```

```
In [1]: data0
```

```
-----  
NameError                                Traceback (most recent call last)  
Cell In[1], line 1  
----> 1 data0  
  
NameError: name 'data0' is not defined
```

```
In [ ]:
```

```
In [36]: data['SUBDIVISION'].unique()
```

```
Out[36]: 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 [37]: data.isna().count() # COUNTING THE NUMBER OF NULL VALUES
```

```
Out[37]: SUBDIVISION    3622  
YEAR                3622  
JAN                 3622  
FEB                 3622  
MAR                 3622  
APR                 3622  
MAY                 3622  
JUN                 3622  
JUL                 3622  
AUG                 3622  
SEP                 3622  
OCT                 3622  
NOV                 3622  
DEC                 3622  
dtype: int64
```

```
In [38]: data.isna().sum() # PRINTING THE NULL VALUES PRESENT IN DATA
```

```
Out[38]: SUBDIVISION    0  
YEAR                0  
JAN                 3  
FEB                 3  
MAR                 6  
APR                 4  
MAY                 3  
JUN                 5  
JUL                 7  
AUG                 4  
SEP                 6  
OCT                 7  
NOV                10  
DEC                 9  
dtype: int64
```

```
In [66]: data2=data.loc[(data.SUBDIVISION=="COASTAL ANDHRA PRADESH")] # MODIFYING THE DATA ACCORDING TO OUR CONDITION
```

In [67]: data2

Out[67]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
3091	COASTAL ANDHRA PRADESH	1910	0.9	2.4	0.2	31.9	22.6	179.6	222.2	204.5	216.8	286.4	87.2	0.0
3092	COASTAL ANDHRA PRADESH	1911	0.0	0.0	10.3	14.4	31.2	150.5	158.1	76.2	160.4	119.7	77.6	21.0
3093	COASTAL ANDHRA PRADESH	1912	0.0	9.6	4.8	15.2	33.6	55.9	217.9	230.1	188.7	86.1	61.7	0.0
3094	COASTAL ANDHRA PRADESH	1913	0.0	12.9	0.5	3.3	68.2	129.0	166.9	124.3	119.5	206.7	7.2	2.4
3095	COASTAL ANDHRA PRADESH	1914	0.0	10.9	5.8	73.0	77.7	171.7	157.4	178.5	292.4	22.8	17.8	5.1
...
3187	COASTAL ANDHRA PRADESH	2006	0.0	0.0	30.7	58.0	67.6	132.0	126.8	279.7	224.1	185.4	55.3	0.0
3188	COASTAL ANDHRA PRADESH	2007	0.0	5.1	0.1	8.5	13.0	300.1	141.0	164.4	287.1	167.7	11.9	0.5
3189	COASTAL ANDHRA PRADESH	2008	3.2	61.6	93.8	18.5	34.9	119.8	208.5	276.2	166.8	51.6	71.4	1.2
3190	COASTAL ANDHRA PRADESH	2009	0.0	0.0	5.7	6.4	53.0	72.6	140.9	163.5	151.9	92.6	102.6	1.3
3191	COASTAL ANDHRA PRADESH	2010	21.8	2.3	4.4	14.8	162.0	156.1	318.9	248.6	230.5	204.0	210.9	138.2

101 rows × 14 columns

```
In [41]: data2.isna().sum()
```

```
Out[41]: 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  
        dtype: int64
```

```
In [52]: data3=data.loc[(data.SUBDIVISION=="ARUNACHAL PRADESH")]
```

In [53]: data3

Out[53]:

	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
...
197	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7
198	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9
199	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6
200	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9
201	ARUNACHAL PRADESH	2010	0.6	13.2	237.8	466.9	312.7	509.9	378.0	321.5	444.2	97.7	58.9	14.2

92 rows × 14 columns

In [54]: data3.isna().sum()

```
Out[54]: 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
dtype: int64
```

```
In [55]: data3["SUBDIVISION"].unique()
```

```
Out[55]: array(['ARUNACHAL PRADESH'], dtype=object)
```

```
In [62]: data3['annual rain']=data3.apply(lambda row:row.JAN+row.FEB+row.MAR+row.APR+row.MAY+row.JUN+row.JUL+row.AUG+
```

```
In [63]: data3
```

```
Out[63]:
```

	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	5486.3
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6	4693.9
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
...
197	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7	2259.4
198	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9	3020.6
199	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6	2244.5
200	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9	1750.0
201	ARUNACHAL PRADESH	2010	0.6	13.2	237.8	466.9	312.7	509.9	378.0	321.5	444.2	97.7	58.9	14.2	2855.6

92 rows × 15 columns

```
In [70]: data2['ANNUAL RAIN']=data2.apply(lambda row:row.JAN+row.FEB+row.MAR+row.APR+row.MAY+row.JUN+row.JUL+row.AUG+
# ADDING NEW COLUMN
```

In [71]: data2

Out[71]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	annual rain	ANNUAL RAIN
3091	COASTAL ANDHRA PRADESH	1910	0.9	2.4	0.2	31.9	22.6	179.6	222.2	204.5	216.8	286.4	87.2	0.0	1254.7	1254.7
3092	COASTAL ANDHRA PRADESH	1911	0.0	0.0	10.3	14.4	31.2	150.5	158.1	76.2	160.4	119.7	77.6	21.0	819.4	819.4
3093	COASTAL ANDHRA PRADESH	1912	0.0	9.6	4.8	15.2	33.6	55.9	217.9	230.1	188.7	86.1	61.7	0.0	903.6	903.6
3094	COASTAL ANDHRA PRADESH	1913	0.0	12.9	0.5	3.3	68.2	129.0	166.9	124.3	119.5	206.7	7.2	2.4	840.9	840.9
3095	COASTAL ANDHRA PRADESH	1914	0.0	10.9	5.8	73.0	77.7	171.7	157.4	178.5	292.4	22.8	17.8	5.1	1013.1	1013.1
...
3187	COASTAL ANDHRA PRADESH	2006	0.0	0.0	30.7	58.0	67.6	132.0	126.8	279.7	224.1	185.4	55.3	0.0	1159.6	1159.6
3188	COASTAL ANDHRA PRADESH	2007	0.0	5.1	0.1	8.5	13.0	300.1	141.0	164.4	287.1	167.7	11.9	0.5	1099.4	1099.4
3189	COASTAL ANDHRA PRADESH	2008	3.2	61.6	93.8	18.5	34.9	119.8	208.5	276.2	166.8	51.6	71.4	1.2	1107.5	1107.5
3190	COASTAL ANDHRA PRADESH	2009	0.0	0.0	5.7	6.4	53.0	72.6	140.9	163.5	151.9	92.6	102.6	1.3	790.5	790.5
3191	COASTAL ANDHRA PRADESH	2010	21.8	2.3	4.4	14.8	162.0	156.1	318.9	248.6	230.5	204.0	210.9	138.2	1712.5	1712.5

101 rows × 16 columns

In [72]: data2=data2.drop(['annual rain'],axis=1)

In [73]: data2

Out[73]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN
3091	COASTAL ANDHRA PRADESH	1910	0.9	2.4	0.2	31.9	22.6	179.6	222.2	204.5	216.8	286.4	87.2	0.0	1254.7
3092	COASTAL ANDHRA PRADESH	1911	0.0	0.0	10.3	14.4	31.2	150.5	158.1	76.2	160.4	119.7	77.6	21.0	819.4
3093	COASTAL ANDHRA PRADESH	1912	0.0	9.6	4.8	15.2	33.6	55.9	217.9	230.1	188.7	86.1	61.7	0.0	903.6
3094	COASTAL ANDHRA PRADESH	1913	0.0	12.9	0.5	3.3	68.2	129.0	166.9	124.3	119.5	206.7	7.2	2.4	840.9
3095	COASTAL ANDHRA PRADESH	1914	0.0	10.9	5.8	73.0	77.7	171.7	157.4	178.5	292.4	22.8	17.8	5.1	1013.1
...
3187	COASTAL ANDHRA PRADESH	2006	0.0	0.0	30.7	58.0	67.6	132.0	126.8	279.7	224.1	185.4	55.3	0.0	1159.6
3188	COASTAL ANDHRA PRADESH	2007	0.0	5.1	0.1	8.5	13.0	300.1	141.0	164.4	287.1	167.7	11.9	0.5	1099.4
3189	COASTAL ANDHRA PRADESH	2008	3.2	61.6	93.8	18.5	34.9	119.8	208.5	276.2	166.8	51.6	71.4	1.2	1107.5
3190	COASTAL ANDHRA PRADESH	2009	0.0	0.0	5.7	6.4	53.0	72.6	140.9	163.5	151.9	92.6	102.6	1.3	790.5
3191	COASTAL ANDHRA PRADESH	2010	21.8	2.3	4.4	14.8	162.0	156.1	318.9	248.6	230.5	204.0	210.9	138.2	1712.5

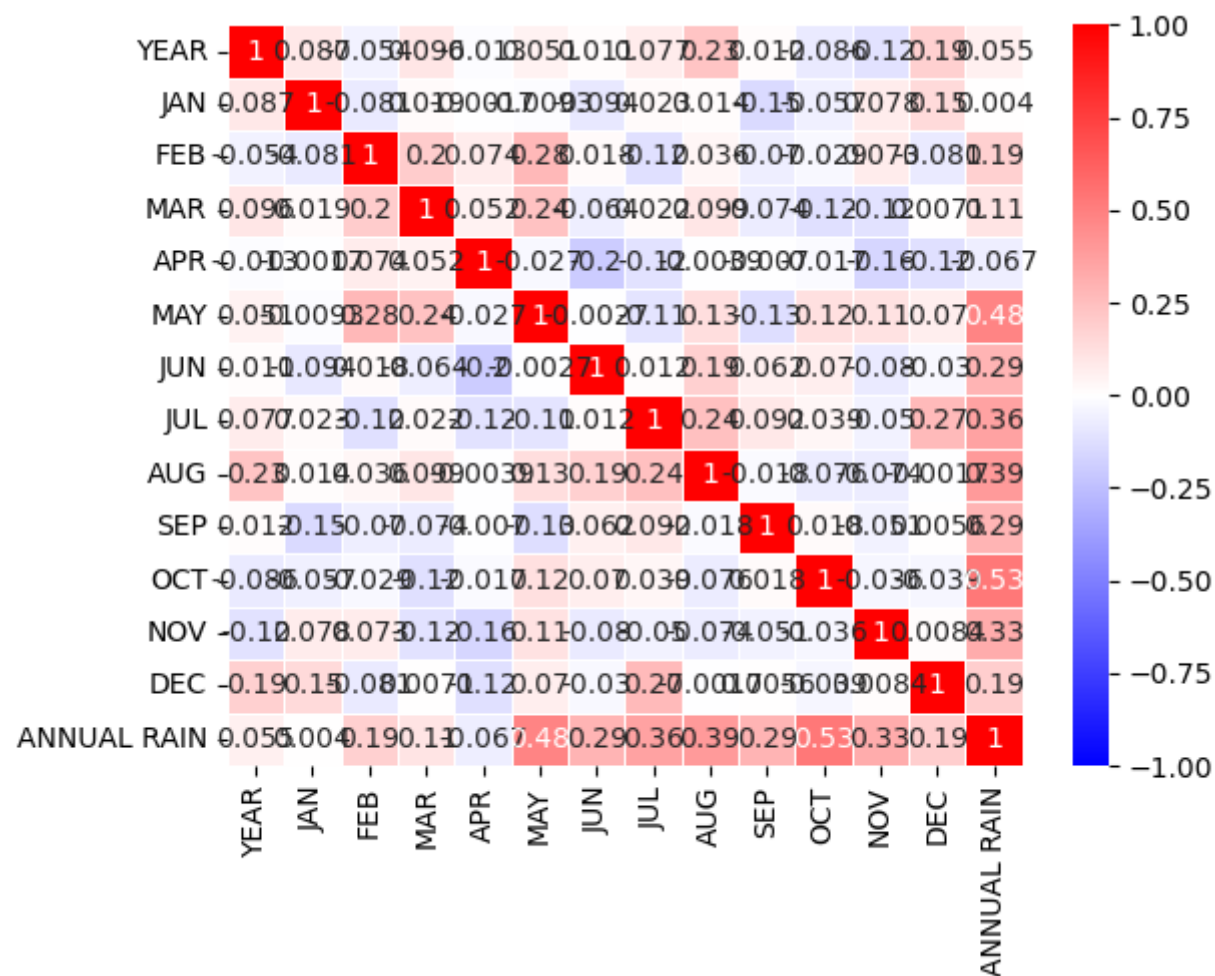
101 rows × 15 columns

```
In [86]: import seaborn as sns
import matplotlib.pyplot as plt
cor=data2.corr()# USING CORRELATION FUNCTION RELATING THE DATA
```



```
In [90]: sns.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidths=.5,cmap="bwr") # SHOWING THE HEATMAP
```

```
Out[90]: <Axes: >
```



```
In [88]: data2['SWM']=data2.apply(lambda row:row.JUN+row.JUL+row.AUG+row.SEP,axis=1)
data2['NSM']=data2.apply(lambda row:row.OCT+row.NOV+row.DEC,axis=1)
#ADDING THE SWM AND NSM COLUMNS
```

```
In [89]: data2
```

```
Out[89]:
```

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL RAIN	SWM	NSM
3091	COASTAL ANDHRA PRADESH	1910	0.9	2.4	0.2	31.9	22.6	179.6	222.2	204.5	216.8	286.4	87.2	0.0	1254.7	823.1	373.6
3092	COASTAL ANDHRA PRADESH	1911	0.0	0.0	10.3	14.4	31.2	150.5	158.1	76.2	160.4	119.7	77.6	21.0	819.4	545.2	218.3
3093	COASTAL ANDHRA PRADESH	1912	0.0	9.6	4.8	15.2	33.6	55.9	217.9	230.1	188.7	86.1	61.7	0.0	903.6	692.6	147.8
3094	COASTAL ANDHRA PRADESH	1913	0.0	12.9	0.5	3.3	68.2	129.0	166.9	124.3	119.5	206.7	7.2	2.4	840.9	539.7	216.3
3095	COASTAL ANDHRA PRADESH	1914	0.0	10.9	5.8	73.0	77.7	171.7	157.4	178.5	292.4	22.8	17.8	5.1	1013.1	800.0	45.7
...
3187	COASTAL ANDHRA PRADESH	2006	0.0	0.0	30.7	58.0	67.6	132.0	126.8	279.7	224.1	185.4	55.3	0.0	1159.6	762.6	240.7
3188	COASTAL ANDHRA PRADESH	2007	0.0	5.1	0.1	8.5	13.0	300.1	141.0	164.4	287.1	167.7	11.9	0.5	1099.4	892.6	180.1
3189	COASTAL ANDHRA PRADESH	2008	3.2	61.6	93.8	18.5	34.9	119.8	208.5	276.2	166.8	51.6	71.4	1.2	1107.5	771.3	124.2
3190	COASTAL ANDHRA PRADESH	2009	0.0	0.0	5.7	6.4	53.0	72.6	140.9	163.5	151.9	92.6	102.6	1.3	790.5	528.9	196.5
3191	COASTAL ANDHRA PRADESH	2010	21.8	2.3	4.4	14.8	162.0	156.1	318.9	248.6	230.5	204.0	210.9	138.2	1712.5	954.1	553.1

101 rows × 17 columns

```
In [104]: data3=data2.drop(["ANNUAL RAIN"],axis=1)
```

In [105]: data3

Out[105]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	SWM	NSM
3091	COASTAL ANDHRA PRADESH	1910	0.9	2.4	0.2	31.9	22.6	179.6	222.2	204.5	216.8	286.4	87.2	0.0	823.1	373.6
3092	COASTAL ANDHRA PRADESH	1911	0.0	0.0	10.3	14.4	31.2	150.5	158.1	76.2	160.4	119.7	77.6	21.0	545.2	218.3
3093	COASTAL ANDHRA PRADESH	1912	0.0	9.6	4.8	15.2	33.6	55.9	217.9	230.1	188.7	86.1	61.7	0.0	692.6	147.8
3094	COASTAL ANDHRA PRADESH	1913	0.0	12.9	0.5	3.3	68.2	129.0	166.9	124.3	119.5	206.7	7.2	2.4	539.7	216.3
3095	COASTAL ANDHRA PRADESH	1914	0.0	10.9	5.8	73.0	77.7	171.7	157.4	178.5	292.4	22.8	17.8	5.1	800.0	45.7
...
3187	COASTAL ANDHRA PRADESH	2006	0.0	0.0	30.7	58.0	67.6	132.0	126.8	279.7	224.1	185.4	55.3	0.0	762.6	240.7
3188	COASTAL ANDHRA PRADESH	2007	0.0	5.1	0.1	8.5	13.0	300.1	141.0	164.4	287.1	167.7	11.9	0.5	892.6	180.1
3189	COASTAL ANDHRA PRADESH	2008	3.2	61.6	93.8	18.5	34.9	119.8	208.5	276.2	166.8	51.6	71.4	1.2	771.3	124.2
3190	COASTAL ANDHRA PRADESH	2009	0.0	0.0	5.7	6.4	53.0	72.6	140.9	163.5	151.9	92.6	102.6	1.3	528.9	196.5
3191	COASTAL ANDHRA PRADESH	2010	21.8	2.3	4.4	14.8	162.0	156.1	318.9	248.6	230.5	204.0	210.9	138.2	954.1	553.1

101 rows × 16 columns

In [106]: data3=data3.drop(['JAN', 'FEB', 'MAR', 'APR', 'MAY'],axis=1)

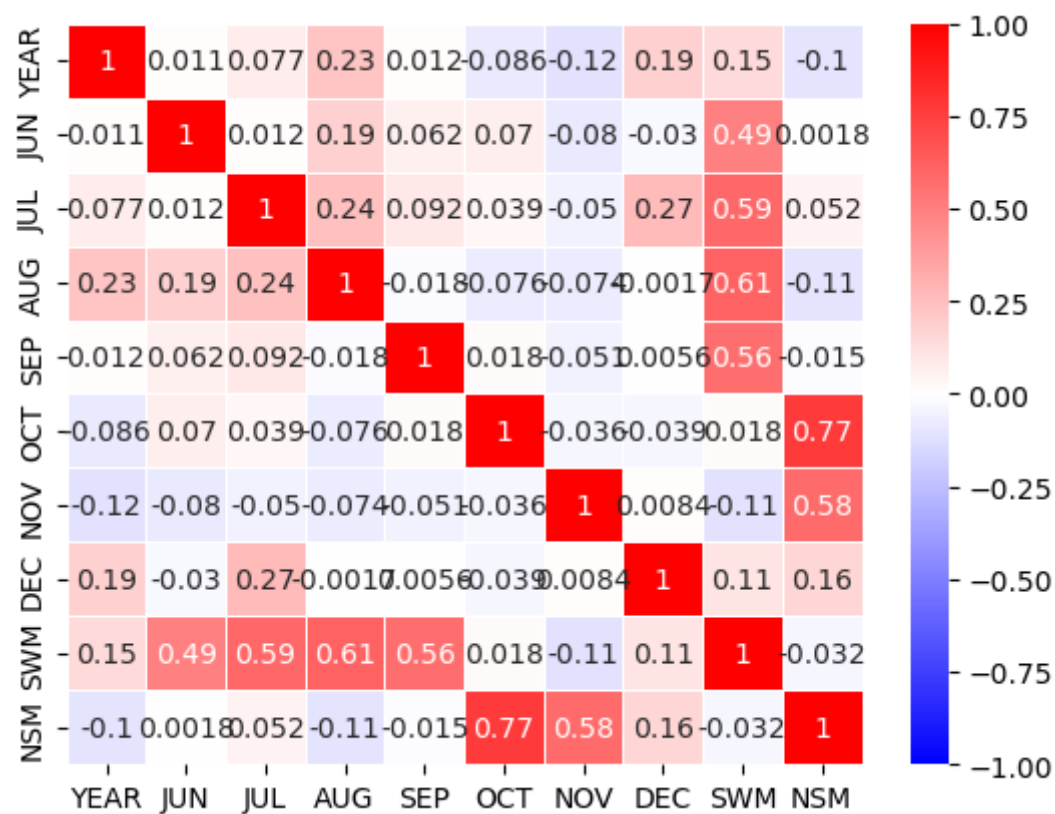
In [107]: `import seaborn as sns`
`import matplotlib.pyplot as plt`
`cor=data3.corr()`

```
In [98]: list(data3)
```

```
Out[98]: ['SUBDIVISION',  
          'YEAR',  
          'JAN',  
          'FEB',  
          'MAR',  
          'APR',  
          'MAY',  
          'JUN',  
          'JUL',  
          'AUG',  
          'SEP',  
          'OCT',  
          'NOV',  
          'DEC',  
          'SWM',  
          'NSM']
```

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

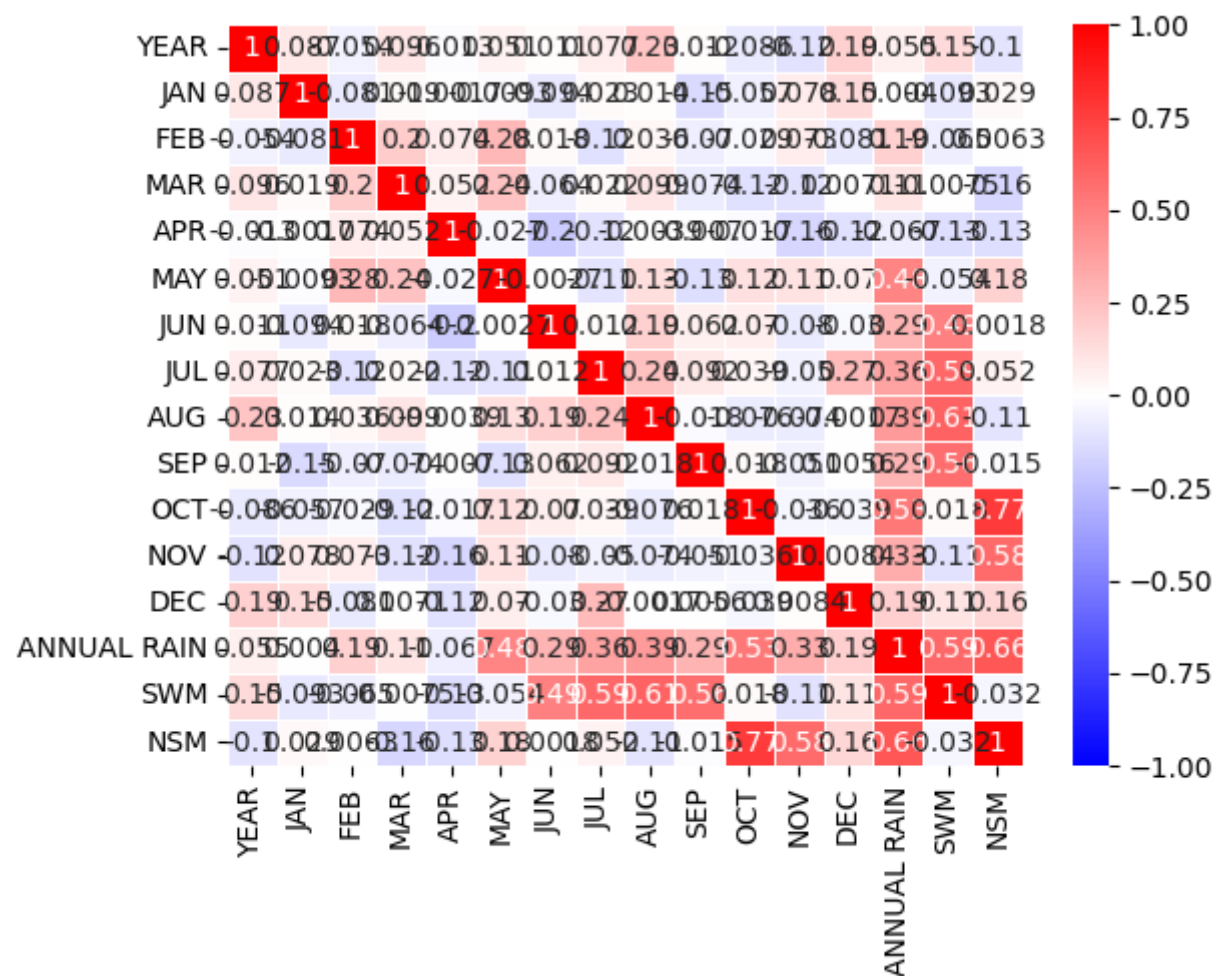
```
Out[108]: <Axes: >
```



```
In [96]: import seaborn as sns  
import matplotlib.pyplot as plt  
cor=data2.corr()
```

```
In [97]: sns.heatmap(cor, vmax=1, vmin=-1, annot=True, linewidths=.5, cmap="bwr")
```

```
Out[97]: <Axes: >
```



In []: