In [279…

```
import pandas as pd
import pickle
import warnings
warnings.filterwarnings('ignore')
```

In [280…

```
a=pd.read_csv("C:\\Users\\reshma_koduri\\OneDrive\\Documents\\rainfall in india 1901
a
```

Out[280…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
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
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
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
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
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
…	…	…	…	…	…	…	…	…	…	…	…	…	…
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
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
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
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
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

4116 rows × 19 columns

In [281…

```
a.head(10)
```

Out[281…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	D
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	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	16
2	ANDAMAN & NICOBAR	1903	12.7	144.0	0.0	1.0	235.1	479.9	728.4	326.7	339.0	181.2	284.4	22

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	D
	ISLANDS													
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	4
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	34
5	ANDAMAN & NICOBAR ISLANDS	1906	36.6	0.0	0.0	0.0	556.1	733.3	247.7	320.5	164.3	267.8	128.9	7
6	ANDAMAN & NICOBAR ISLANDS	1907	110.7	0.0	113.3	21.6	616.3	305.2	443.9	377.6	200.4	264.4	648.9	24
7	ANDAMAN & NICOBAR ISLANDS	1908	20.9	85.1	0.0	29.0	562.0	693.6	481.4	699.9	428.8	170.7	208.1	19
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	15
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	24

In [282…

```
a.tail(10)
```

Out[282…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	
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	
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	1
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	1
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	1
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	
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	
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	
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	
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	
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	1

In [283…

```
a.describe()
```

Out[283…

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	
count	4116.000000	4112.000000	4113.000000	4110.000000	4112.000000	4113.000000	4111.000000	4

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	
mean	1958.218659	18.957320	21.805325	27.359197	43.127432	85.745417	230.234444	
std	33.140898	33.585371	35.909488	46.959424	67.831168	123.234904	234.710758	
min	1901.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.400000	
25%	1930.000000	0.600000	0.600000	1.000000	3.000000	8.600000	70.350000	
50%	1958.000000	6.000000	6.700000	7.800000	15.700000	36.600000	138.700000	
75%	1987.000000	22.200000	26.800000	31.300000	49.950000	97.200000	305.150000	
max	2015.000000	583.700000	403.500000	605.600000	595.100000	1168.600000	1609.900000	2

In [284…

```
a.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4116 entries, 0 to 4115
Data columns (total 19 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   SUBDIVISION  4116 non-null   object 
 1   YEAR         4116 non-null   int64  
 2   JAN          4112 non-null   float64
 3   FEB          4113 non-null   float64
 4   MAR          4110 non-null   float64
 5   APR          4112 non-null   float64
 6   MAY          4113 non-null   float64
 7   JUN          4111 non-null   float64
 8   JUL          4109 non-null   float64
 9   AUG          4112 non-null   float64
 10  SEP          4110 non-null   float64
 11  OCT          4109 non-null   float64
 12  NOV          4105 non-null   float64
 13  DEC          4106 non-null   float64
 14  ANNUAL       4090 non-null   float64
 15  Jan-Feb      4110 non-null   float64
 16  Mar-May      4107 non-null   float64
 17  Jun-Sep      4106 non-null   float64
 18  Oct-Dec      4103 non-null   float64
dtypes: float64(17), int64(1), object(1)
memory usage: 611.1+ KB
```

In [285…

```
a['SUBDIVISION'].unique()
```

Out[285…

```
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 [286…

```
a.nunique()
```

Out[286…
```
SUBDIVISION      36
YEAR            115
JAN             802
FEB             898
MAR             978
APR            1234
MAY            1731
JUN            2722
JUL            3050
AUG            2913
SEP            2632
OCT            1939
NOV            1239
DEC             801
ANNUAL         3712
Jan-Feb        1220
Mar-May        2262
Jun-Sep        3683
Oct-Dec        2389
dtype: int64
```

In [287…
```
b=a.drop(['ANNUAL','Jan-Feb','Mar-May','Jun-Sep','Oct-Dec'],axis=1)
```

In [288…
```
list(b)
```

Out[288…
```
['SUBDIVISION',
 'YEAR',
 'JAN',
 'FEB',
 'MAR',
 'APR',
 'MAY',
 'JUN',
 'JUL',
 'AUG',
 'SEP',
 'OCT',
 'NOV',
 'DEC']
```

In [289…
```
b.isna().sum()
```

Out[289…
```
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
dtype: int64
```

In [290…
```
c=b.loc[(b.SUBDIVISION == "ARUNACHAL PRADESH")]
c
```

Out[290…

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

97 rows × 14 columns

In [291…

```
d=b.loc[(b.SUBDIVISION == "KERALA")]
d
```

Out[291…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
3887	KERALA	1901	28.7	44.7	51.6	160.0	174.7	824.6	743.0	357.5	197.7	266.9	350.8
3888	KERALA	1902	6.7	2.6	57.3	83.9	134.5	390.9	1205.0	315.8	491.6	358.4	158.3
3889	KERALA	1903	3.2	18.6	3.1	83.6	249.7	558.6	1022.5	420.2	341.8	354.1	157.0
3890	KERALA	1904	23.7	3.0	32.2	71.5	235.7	1098.2	725.5	351.8	222.7	328.1	33.9
3891	KERALA	1905	1.2	22.3	9.4	105.9	263.3	850.2	520.5	293.6	217.2	383.5	74.4
…	…	…	…	…	…	…	…	…	…	…	…	…	…
3997	KERALA	2011	20.5	45.7	24.1	165.2	124.2	788.5	536.8	492.7	391.2	227.2	169.7
3998	KERALA	2012	7.4	11.0	21.0	171.1	95.3	430.3	362.6	501.6	241.1	187.5	112.9
3999	KERALA	2013	3.9	40.1	49.9	49.3	119.3	1042.7	830.2	369.7	318.6	259.9	154.9
4000	KERALA	2014	4.6	10.3	17.9	95.7	251.0	454.4	677.8	733.9	298.8	355.5	99.5
4001	KERALA	2015	3.1	5.8	50.1	214.1	201.8	563.6	406.0	252.2	292.9	308.1	223.6

115 rows × 14 columns

In [292…
```
d.isna().sum()
```

Out[292…
```
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 [293…
```
e=b.loc[(b.SUBDIVISION == "EAST RAJASTHAN")]
e
```

Out[293…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DE(
1932	EAST RAJASTHAN	1901	21.6	8.9	2.9	0.7	5.0	15.0	164.8	175.6	7.5	9.8	0.0	0.8
1933	EAST RAJASTHAN	1902	4.1	0.7	0.0	1.8	9.9	34.6	247.6	116.7	145.6	14.4	0.0	2.8
1934	EAST RAJASTHAN	1903	1.9	0.7	1.3	0.1	12.9	15.6	238.2	229.1	168.5	17.8	0.0	0.(
1935	EAST RAJASTHAN	1904	4.3	5.5	21.7	0.2	27.5	49.9	289.7	223.5	50.2	1.5	5.8	14.7
1936	EAST RAJASTHAN	1905	4.1	8.8	3.2	1.6	2.0	14.4	130.5	30.9	83.8	0.0	0.0	0.(
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2042	EAST RAJASTHAN	2011	0.0	11.2	0.2	0.5	5.1	140.9	193.6	284.1	166.4	0.0	0.0	0.(
2043	EAST RAJASTHAN	2012	1.9	0.0	0.0	3.6	9.5	11.2	170.5	365.0	131.3	0.5	0.0	0.7
2044	EAST RAJASTHAN	2013	1.4	21.7	0.4	3.2	1.0	90.6	319.0	278.5	88.0	30.6	1.3	0.3
2045	EAST RAJASTHAN	2014	28.4	10.0	6.4	7.3	8.4	23.5	197.1	261.0	136.9	3.2	0.0	1.7
2046	EAST RAJASTHAN	2015	12.1	0.1	55.9	15.9	3.5	96.4	297.6	142.8	20.1	5.0	0.5	0.8

115 rows × 14 columns

In [294…

```
f=b.loc[(b.SUBDIVISION == "GUJARAT REGION")]
f
```

Out[294…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DE(
2277	GUJARAT REGION	1901	4.2	0.0	0.6	1.6	7.0	60.3	240.2	205.4	18.1	16.6	0.0	0.:
2278	GUJARAT REGION	1902	3.9	0.0	0.0	0.6	1.0	32.8	229.8	299.0	281.2	2.3	1.5	11.:
2279	GUJARAT REGION	1903	0.3	0.1	1.4	0.0	12.3	30.1	452.9	202.0	183.2	5.4	0.0	0.(
2280	GUJARAT REGION	1904	0.8	10.6	16.8	0.2	3.9	48.3	194.8	71.8	138.0	6.1	0.1	1.:
2281	GUJARAT REGION	1905	0.1	0.7	1.1	0.3	0.0	20.1	668.3	37.9	81.3	1.4	0.2	0.:
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
2387	GUJARAT REGION	2011	0.0	0.2	0.0	0.0	0.0	16.3	259.2	451.7	162.5	0.4	0.0	0.(
2388	GUJARAT REGION	2012	0.1	0.0	0.0	0.0	0.0	34.4	178.2	230.3	263.8	7.1	0.0	0.(
2389	GUJARAT REGION	2013	0.0	0.9	0.1	4.6	0.0	155.7	405.4	211.1	287.3	53.2	0.1	0.(
2390	GUJARAT REGION	2014	5.7	0.1	0.2	1.0	1.3	11.6	307.5	138.6	235.1	3.3	1.3	0.(
2391	GUJARAT REGION	2015	1.8	0.0	6.1	5.5	0.9	120.7	354.7	37.4	93.4	2.2	0.3	0.(

115 rows × 14 columns

In [295…

```
b=a.groupby('SUBDIVISION').count()
b
```

Out[295…

SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNU
ANDAMAN & NICOBAR ISLANDS	110	110	110	108	108	109	108	108	108	107	108	108	107	
ARUNACHAL PRADESH	97	96	96	95	97	97	96	96	97	97	95	95	95	
ASSAM & MEGHALAYA	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	
CHHATTISGARH	115	115	115	115	115	115	115	115	115	115	115	115	115	
COASTAL ANDHRA	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	ANNU
PRADESH														
COASTAL KARNATAKA	115	114	115	115	115	115	115	115	115	115	115	115	115	
EAST MADHYA PRADESH	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	
EAST UTTAR PRADESH	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	
GUJARAT REGION	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	
HIMACHAL PRADESH	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	
JHARKHAND	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	
KONKAN & GOA	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	
MADHYA MAHARASHTRA	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	
NAGA MANI MIZO TRIPURA	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	
ORISSA	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	
RAYALSEEMA	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	
SOUTH INTERIOR KARNATAKA	115	115	115	115	115	115	115	115	115	115	115	115	115	
SUB HIMALAYAN	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	ANNU
WEST BENGAL & SIKKIM														
TAMIL NADU	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	
UTTARAKHAND	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	
WEST MADHYA PRADESH	115	115	114	115	115	115	115	115	115	115	115	115	115	
WEST RAJASTHAN	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	

In [311…

```
#b.dropna(inplace=True)b,c,d,e
t=pd.concat([c,d,e,f])
t
```

Out[311…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.100000	316.1	424.6	1124.9	694.544792	629.7	333.9
111	ARUNACHAL PRADESH	1917	21.4	164.5	153.527368	269.6	107.9	823.8	909.100000	628.4	411.5
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.200000	144.6	861.1	1609.9	1303.000000	692.6	515.8
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.500000	256.9	420.6	973.6	999.000000	286.7	628.7
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.600000	364.7	173.6	840.6	535.400000	896.5	376.7
...	...	...	...	...	...	...	...	...	...	...	...
2387	GUJARAT REGION	2011	0.0	0.2	0.000000	0.0	0.0	16.3	259.200000	451.7	162.5
2388	GUJARAT REGION	2012	0.1	0.0	0.000000	0.0	0.0	34.4	178.200000	230.3	263.8
2389	GUJARAT REGION	2013	0.0	0.9	0.100000	4.6	0.0	155.7	405.400000	211.1	287.3
2390	GUJARAT REGION	2014	5.7	0.1	0.200000	1.0	1.3	11.6	307.500000	138.6	235.1
2391	GUJARAT REGION	2015	1.8	0.0	6.100000	5.5	0.9	120.7	354.700000	37.4	93.4

442 rows × 14 columns

In [312…
```
#b.isna().sum()
```

In [313…
```
c['JAN']=c['JAN'].fillna(c['JAN'].mean())
c['FEB']=c['FEB'].fillna(c['FEB'].mean())
c['MAR']=c['MAR'].fillna(c['MAR'].mean())
c['APR']=c['APR'].fillna(c['APR'].mean())
c['MAY']=c['MAY'].fillna(c['MAY'].mean())
c['JUN']=c['JUN'].fillna(c['JUN'].mean())
c['JUL']=c['JUL'].fillna(c['JUL'].mean())
c['AUG']=c['AUG'].fillna(c['AUG'].mean())
c['SEP']=c['SEP'].fillna(c['SEP'].mean())
c['OCT']=c['OCT'].fillna(c['OCT'].mean())
c['NOV']=c['NOV'].fillna(c['NOV'].mean())
c['DEC']=c['DEC'].fillna(c['DEC'].mean())
```

In [314…
```
c.isna().sum()
```

Out[314…
```
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 [315…
```
#b.replace(to_replace="ARUNACHAL PRADESH",
           #value="ap")
```

In [316…
```
#a=b.combine_first(a)
#a
```

In [317…
```
#df1 =pd.DataFrame(a)
#df2=pd.DataFrame(b)
#df1.loc[df1.a.isin(df2.b)]
#print(df1)
```

In [318…
```
t['ANNUAL RAIN']=t.apply(lambda row: row.JAN + row.FEB,axis=1)
t
```

Out[318…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.100000	316.1	424.6	1124.9	694.544792	629.7	333.9
111	ARUNACHAL PRADESH	1917	21.4	164.5	153.527368	269.6	107.9	823.8	909.100000	628.4	411.5

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.200000	144.6	861.1	1609.9	1303.000000	692.6	515.8
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.500000	256.9	420.6	973.6	999.000000	286.7	628.7
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.600000	364.7	173.6	840.6	535.400000	896.5	376.7
...	...	...	...	...	...	...	...	...	...	...	...
2387	GUJARAT REGION	2011	0.0	0.2	0.000000	0.0	0.0	16.3	259.200000	451.7	162.5
2388	GUJARAT REGION	2012	0.1	0.0	0.000000	0.0	0.0	34.4	178.200000	230.3	263.8
2389	GUJARAT REGION	2013	0.0	0.9	0.100000	4.6	0.0	155.7	405.400000	211.1	287.3
2390	GUJARAT REGION	2014	5.7	0.1	0.200000	1.0	1.3	11.6	307.500000	138.6	235.1
2391	GUJARAT REGION	2015	1.8	0.0	6.100000	5.5	0.9	120.7	354.700000	37.4	93.4

442 rows × 15 columns

In [319…

```
cor=t.corr()
cor
```

Out[319…

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AU
YEAR	1.000000	-0.039788	0.027321	0.039578	0.040068	-0.028470	-0.086879	-0.128459	-0.0474(
JAN	-0.039788	1.000000	0.592708	0.591032	0.649345	0.550079	0.468785	0.335847	0.3800{
FEB	0.027321	0.592708	1.000000	0.768436	0.742730	0.555083	0.486546	0.400564	0.3773(
MAR	0.039578	0.591032	0.768436	1.000000	0.768793	0.578756	0.522549	0.408913	0.4001(
APR	0.040068	0.649345	0.742730	0.768793	1.000000	0.685387	0.687044	0.561924	0.4822
MAY	-0.028470	0.550079	0.555083	0.578756	0.685387	1.000000	0.711546	0.623964	0.5079{
JUN	-0.086879	0.468785	0.486546	0.522549	0.687044	0.711546	1.000000	0.711848	0.5602{
JUL	-0.128459	0.335847	0.400564	0.408913	0.561924	0.623964	0.711848	1.000000	0.4495{
AUG	-0.047466	0.380088	0.377364	0.400100	0.482211	0.507921	0.560288	0.449596	1.0000(
SEP	-0.007810	0.483836	0.577966	0.560795	0.637224	0.594587	0.562840	0.547474	0.4608{
OCT	-0.044851	0.305408	0.285336	0.301451	0.501209	0.593225	0.699438	0.576095	0.4403{
NOV	-0.045870	0.049643	0.001481	0.054919	0.244392	0.355528	0.530641	0.393555	0.2473
DEC	-0.056015	0.225000	0.177038	0.242244	0.374718	0.354850	0.455248	0.383294	0.3011(

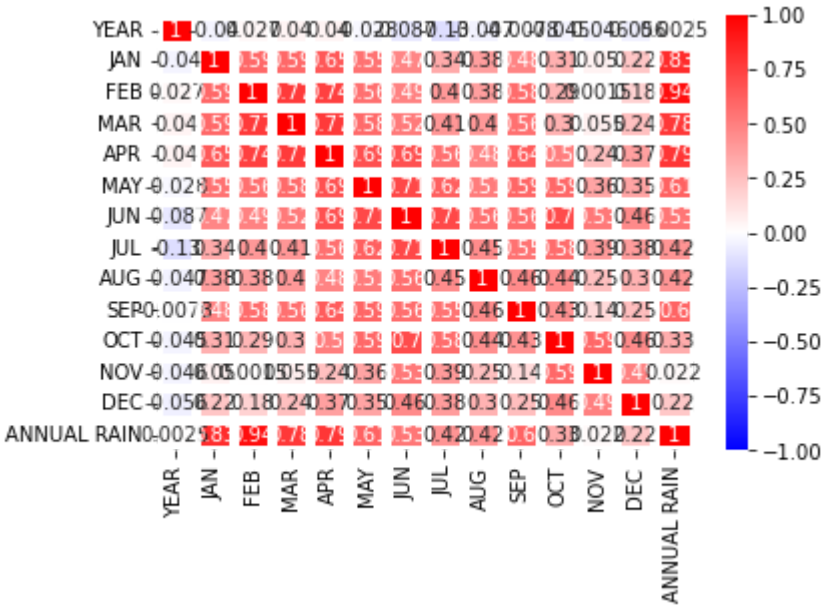
	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AU
ANNUAL RAIN	0.002510	0.827882	0.942453	0.780493	0.786797	0.614912	0.533435	0.418372	0.4205

In [320…

```
import seaborn as sb
sb.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidth=-5,cmap="bwr")
```

Out[320…

<AxesSubplot:>



In [321…

```
t['JJAS']=t.apply(lambda row: row.JUN + row.JUL + row.AUG + row.SEP,axis=1)
t
```

Out[321…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.100000	316.1	424.6	1124.9	694.544792	629.7	333.9
111	ARUNACHAL PRADESH	1917	21.4	164.5	153.527368	269.6	107.9	823.8	909.100000	628.4	411.5
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.200000	144.6	861.1	1609.9	1303.000000	692.6	515.8
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.500000	256.9	420.6	973.6	999.000000	286.7	628.7
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.600000	364.7	173.6	840.6	535.400000	896.5	376.7
…	…	…	…	…	…	…	…	…	…	…	…
2387	GUJARAT REGION	2011	0.0	0.2	0.000000	0.0	0.0	16.3	259.200000	451.7	162.5
2388	GUJARAT REGION	2012	0.1	0.0	0.000000	0.0	0.0	34.4	178.200000	230.3	263.8
2389	GUJARAT REGION	2013	0.0	0.9	0.100000	4.6	0.0	155.7	405.400000	211.1	287.3

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
2390	GUJARAT REGION	2014	5.7	0.1	0.200000	1.0	1.3	11.6	307.500000	138.6	235.1
2391	GUJARAT REGION	2015	1.8	0.0	6.100000	5.5	0.9	120.7	354.700000	37.4	93.4

442 rows × 16 columns

In [322…

```
cor=t.corr()
cor
```

Out[322…

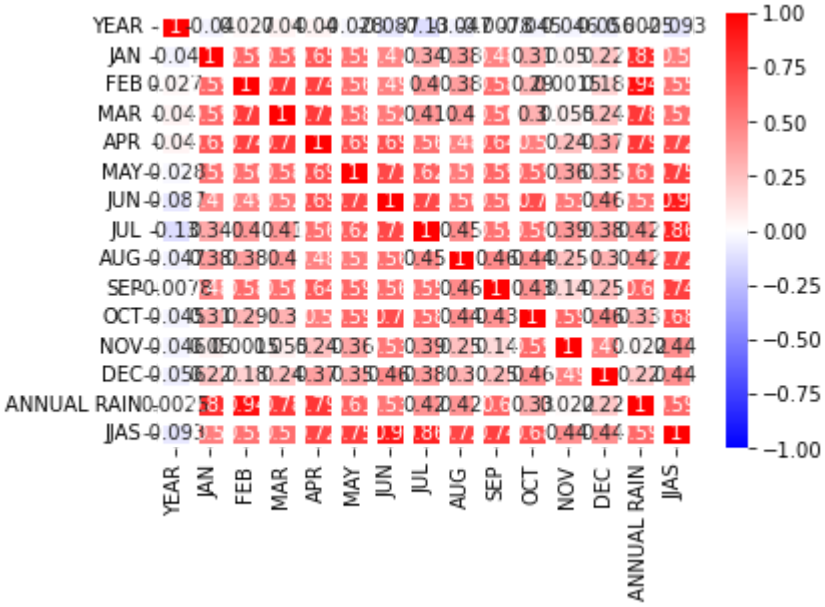
	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AU
YEAR	1.000000	-0.039788	0.027321	0.039578	0.040068	-0.028470	-0.086879	-0.128459	-0.0474
JAN	-0.039788	1.000000	0.592708	0.591032	0.649345	0.550079	0.468785	0.335847	0.3800
FEB	0.027321	0.592708	1.000000	0.768436	0.742730	0.555083	0.486546	0.400564	0.3773
MAR	0.039578	0.591032	0.768436	1.000000	0.768793	0.578756	0.522549	0.408913	0.4001
APR	0.040068	0.649345	0.742730	0.768793	1.000000	0.685387	0.687044	0.561924	0.4822
MAY	-0.028470	0.550079	0.555083	0.578756	0.685387	1.000000	0.711546	0.623964	0.5079
JUN	-0.086879	0.468785	0.486546	0.522549	0.687044	0.711546	1.000000	0.711848	0.5602
JUL	-0.128459	0.335847	0.400564	0.408913	0.561924	0.623964	0.711848	1.000000	0.4495
AUG	-0.047466	0.380088	0.377364	0.400100	0.482211	0.507921	0.560288	0.449596	1.0000
SEP	-0.007810	0.483836	0.577966	0.560795	0.637224	0.594587	0.562840	0.547474	0.4608
OCT	-0.044851	0.305408	0.285336	0.301451	0.501209	0.593225	0.699438	0.576095	0.4403
NOV	-0.045870	0.049643	0.001481	0.054919	0.244392	0.355528	0.530641	0.393555	0.2473
DEC	-0.056015	0.225000	0.177038	0.242244	0.374718	0.354850	0.455248	0.383294	0.3011
ANNUAL RAIN	0.002510	0.827882	0.942453	0.780493	0.786797	0.614912	0.533435	0.418372	0.4205
JJAS	-0.093450	0.500303	0.549931	0.568929	0.724722	0.753784	0.904505	0.864716	0.7230

In [323…

```
import seaborn as sb
sb.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidth=-5,cmap="bwr")
```

Out[323…

```
<AxesSubplot:>
```



In [324…

```
t['OND']=t.apply(lambda row: row.OCT + row.NOV + row.DEC,axis=1)
t
```

Out[324…

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
110	ARUNACHAL PRADESH	1916	48.1	69.8	71.100000	316.1	424.6	1124.9	694.544792	629.7	333.9
111	ARUNACHAL PRADESH	1917	21.4	164.5	153.527368	269.6	107.9	823.8	909.100000	628.4	411.5
112	ARUNACHAL PRADESH	1918	10.4	11.0	191.200000	144.6	861.1	1609.9	1303.000000	692.6	515.8
113	ARUNACHAL PRADESH	1919	34.5	67.8	28.500000	256.9	420.6	973.6	999.000000	286.7	628.7
114	ARUNACHAL PRADESH	1920	14.0	196.3	605.600000	364.7	173.6	840.6	535.400000	896.5	376.7
...	...	...	...	...	...	...	...	...	...	...	...
2387	GUJARAT REGION	2011	0.0	0.2	0.000000	0.0	0.0	16.3	259.200000	451.7	162.5
2388	GUJARAT REGION	2012	0.1	0.0	0.000000	0.0	0.0	34.4	178.200000	230.3	263.8
2389	GUJARAT REGION	2013	0.0	0.9	0.100000	4.6	0.0	155.7	405.400000	211.1	287.3
2390	GUJARAT REGION	2014	5.7	0.1	0.200000	1.0	1.3	11.6	307.500000	138.6	235.1
2391	GUJARAT REGION	2015	1.8	0.0	6.100000	5.5	0.9	120.7	354.700000	37.4	93.4

442 rows × 17 columns

In [325…

```
cor=t.corr()
cor
```

Out[325…

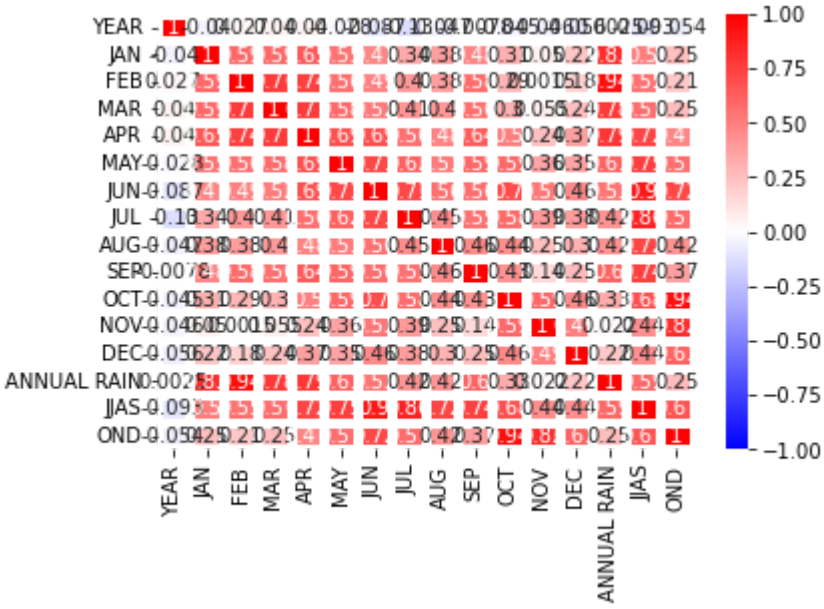
	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AU
YEAR	1.000000	-0.039788	0.027321	0.039578	0.040068	-0.028470	-0.086879	-0.128459	-0.0474
JAN	-0.039788	1.000000	0.592708	0.591032	0.649345	0.550079	0.468785	0.335847	0.3800
FEB	0.027321	0.592708	1.000000	0.768436	0.742730	0.555083	0.486546	0.400564	0.3773
MAR	0.039578	0.591032	0.768436	1.000000	0.768793	0.578756	0.522549	0.408913	0.4001
APR	0.040068	0.649345	0.742730	0.768793	1.000000	0.685387	0.687044	0.561924	0.4822
MAY	-0.028470	0.550079	0.555083	0.578756	0.685387	1.000000	0.711546	0.623964	0.5079
JUN	-0.086879	0.468785	0.486546	0.522549	0.687044	0.711546	1.000000	0.711848	0.5602
JUL	-0.128459	0.335847	0.400564	0.408913	0.561924	0.623964	0.711848	1.000000	0.4495
AUG	-0.047466	0.380088	0.377364	0.400100	0.482211	0.507921	0.560288	0.449596	1.0000
SEP	-0.007810	0.483836	0.577966	0.560795	0.637224	0.594587	0.562840	0.547474	0.4608
OCT	-0.044851	0.305408	0.285336	0.301451	0.501209	0.593225	0.699438	0.576095	0.4403
NOV	-0.045870	0.049643	0.001481	0.054919	0.244392	0.355528	0.530641	0.393555	0.2473
DEC	-0.056015	0.225000	0.177038	0.242244	0.374718	0.354850	0.455248	0.383294	0.3011
ANNUAL RAIN	0.002510	0.827882	0.942453	0.780493	0.786797	0.614912	0.533435	0.418372	0.4205
JJAS	-0.093450	0.500303	0.549931	0.568929	0.724722	0.753784	0.904505	0.864716	0.7230
OND	-0.053536	0.249197	0.212305	0.250634	0.469123	0.568427	0.715475	0.574453	0.4205

In [326…

```
import seaborn as sb
sb.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidth=-5,cmap="bwr")
```

Out[326…

<AxesSubplot:>



In [327…

```
t.to_csv('rainfall2023.csv')
```

In [ ]: