

In [35]: `import pandas as pd`

In [36]: `data=pd.read_csv("/home/placement/Downloads/arunachal.csv")`

In [37]: `data.describe()`

Out[37]:

	Unnamed: 0	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
count	91.00000	91.000000	90.000000	90.000000	89.000000	91.000000	91.000000	90.000000	90.000000	91.000000	91.000000
mean	155.00000	1962.747253	48.598889	93.966667	154.446067	262.990110	364.651648	659.556667	711.963333	502.163736	433.273626
std	26.41338	27.695003	34.687078	46.258375	87.918484	113.395773	181.095447	311.642230	356.372598	275.716730	204.991358
min	110.00000	1916.000000	1.800000	6.100000	28.500000	94.700000	101.800000	239.400000	233.000000	172.400000	152.500000
25%	132.50000	1938.500000	20.075000	65.625000	101.700000	180.600000	237.150000	425.675000	442.150000	301.100000	282.150000
50%	155.00000	1964.000000	45.400000	87.600000	141.700000	245.400000	314.600000	545.750000	613.000000	411.600000	384.300000
75%	177.50000	1986.500000	65.150000	120.400000	189.600000	335.300000	447.050000	840.400000	922.075000	669.200000	521.150000
max	200.00000	2009.000000	164.500000	208.500000	605.600000	595.100000	1168.600000	1609.900000	2362.800000	1664.600000	1222.000000

```
In [38]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 91 entries, 0 to 90
Data columns (total 20 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Unnamed: 0            91 non-null    int64
1   SUBDIVISION           91 non-null    object
2   YEAR                  91 non-null    int64
3   JAN                   90 non-null    float64
4   FEB                   90 non-null    float64
5   MAR                   89 non-null    float64
6   APR                   91 non-null    float64
7   MAY                   91 non-null    float64
8   JUN                   90 non-null    float64
9   JUL                   90 non-null    float64
10  AUG                   91 non-null    float64
11  SEP                   91 non-null    float64
12  OCT                   89 non-null    float64
13  NOV                   89 non-null    float64
14  DEC                   89 non-null    float64
15  ANNUAL                85 non-null    float64
16  Jan-Feb               90 non-null    float64
17  Mar-May               89 non-null    float64
18  Jun-Sep               89 non-null    float64
19  Oct-Dec               88 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 14.3+ KB
```

In [39]: `data.tail()`

Out[39]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	J
86	196	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	216.0	604.6	134
87	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	2259.6	109.7	587.7	139
88	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	3020.7	110.8	590.9	218
89	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	2244.4	116.4	499.9	150
90	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	1749.9	110.8	394.7	111



```
In [40]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 91 entries, 0 to 90
Data columns (total 20 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Unnamed: 0            91 non-null    int64
1   SUBDIVISION           91 non-null    object
2   YEAR                  91 non-null    int64
3   JAN                   90 non-null    float64
4   FEB                   90 non-null    float64
5   MAR                   89 non-null    float64
6   APR                   91 non-null    float64
7   MAY                   91 non-null    float64
8   JUN                   90 non-null    float64
9   JUL                   90 non-null    float64
10  AUG                   91 non-null    float64
11  SEP                   91 non-null    float64
12  OCT                   89 non-null    float64
13  NOV                   89 non-null    float64
14  DEC                   89 non-null    float64
15  ANNUAL                85 non-null    float64
16  Jan-Feb              90 non-null    float64
17  Mar-May              89 non-null    float64
18  Jun-Sep              89 non-null    float64
19  Oct-Dec              88 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 14.3+ KB
```

```
In [41]: list(data.columns)
```

```
Out[41]: ['Unnamed: 0',  
          'SUBDIVISION',  
          'YEAR',  
          'JAN',  
          'FEB',  
          'MAR',  
          'APR',  
          'MAY',  
          'JUN',  
          'JUL',  
          'AUG',  
          'SEP',  
          'OCT',  
          'NOV',  
          'DEC',  
          'ANNUAL',  
          'Jan-Feb',  
          'Mar-May',  
          'Jun-Sep',  
          'Oct-Dec']
```

```
In [42]: data1=data.drop(["Unnamed: 0"],axis=1)
```

In [43]: data1

Out[43]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
0	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	NaN
1	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.
2	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.
3	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.
4	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.
...
86	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	216.0	604.6	1343.7	171.
87	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	2259.6	109.7	587.7	1393.5	168.
88	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	3020.7	110.8	590.9	2185.1	133.
89	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	2244.4	116.4	499.9	1508.7	119.
90	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	1749.9	110.8	394.7	1111.8	132.

91 rows × 19 columns



```
In [44]: data1.isna().sum()
```

```
Out[44]: 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 [45]: data1=pd.get_dummies(data1)
```

```
In [46]: import warnings
```

```
In [47]: warnings.filterwarnings("ignore")
```

In [48]: data1

Out[48]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec	SUBDIVISION
0	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	NaN	
1	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	262.8	
2	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	146.7	
3	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	997.6	
4	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	103.3	
...	
86	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	216.0	604.6	1343.7	171.2	
87	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	2259.6	109.7	587.7	1393.5	168.7	
88	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	3020.7	110.8	590.9	2185.1	133.9	
89	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	2244.4	116.4	499.9	1508.7	119.4	
90	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	1749.9	110.8	394.7	1111.8	132.7	

91 rows × 19 columns



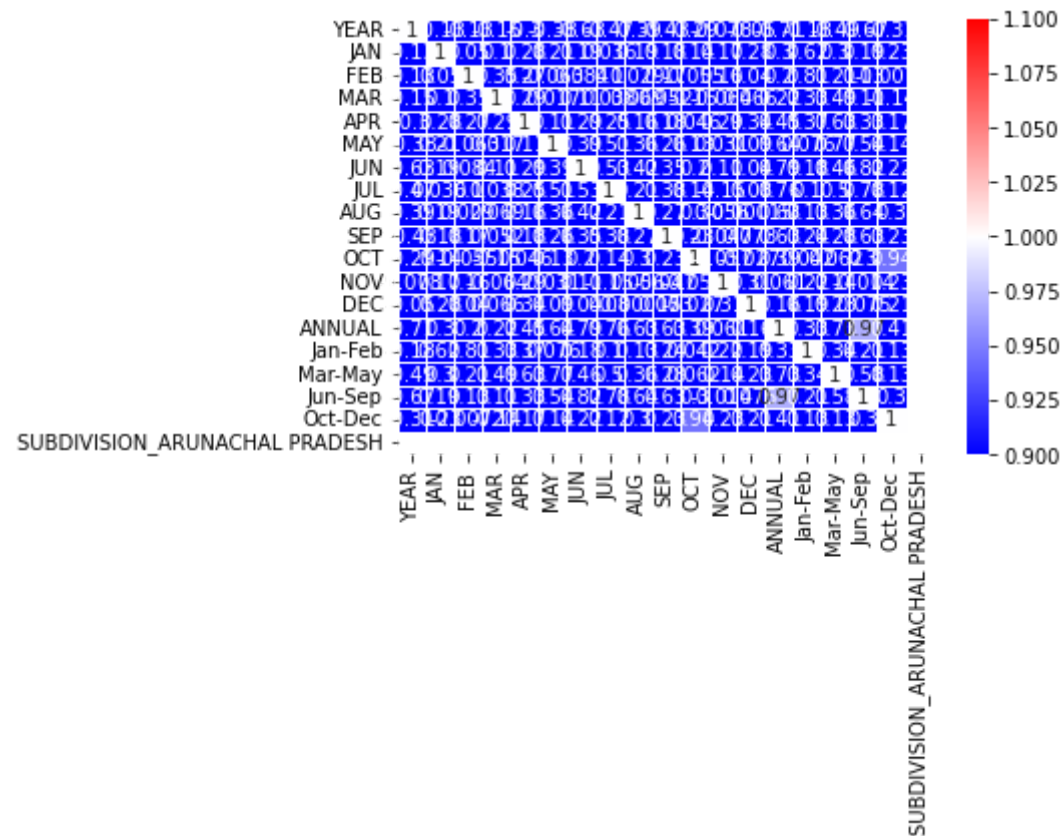

```
In [49]: cor_mat=data1.corr()
cor_mat
```

Out[49]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
YEAR	1.000000	-0.129778	-0.134399	-0.153629	-0.301073	-0.384602	-0.630502	-0.465649	-0.394444	-0.431541	-0.294323
JAN	-0.129778	1.000000	0.049703	0.102652	0.275465	0.213443	0.187899	0.036231	0.186397	0.180209	0.144752
FEB	-0.134399	0.049703	1.000000	0.346448	0.268504	-0.063280	0.084133	0.102494	0.028861	0.168561	-0.054956
MAR	-0.153629	0.102652	0.346448	1.000000	0.292072	0.017182	0.110014	0.038299	0.068542	0.052330	-0.149850
APR	-0.301073	0.275465	0.268504	0.292072	1.000000	0.114128	0.293255	0.252243	0.157620	0.176335	0.046071
MAY	-0.384602	0.213443	-0.063280	0.017182	0.114128	1.000000	0.393818	0.506982	0.363992	0.258744	0.127951
JUN	-0.630502	0.187899	0.084133	0.110014	0.293255	0.393818	1.000000	0.530709	0.417914	0.347814	0.196709
JUL	-0.465649	0.036231	0.102494	0.038299	0.252243	0.506982	0.530709	1.000000	0.210551	0.381137	0.144704
AUG	-0.394444	0.186397	0.028861	0.068542	0.157620	0.363992	0.417914	0.210551	1.000000	0.269123	0.298151
SEP	-0.431541	0.180209	0.168561	0.052330	0.176335	0.258744	0.347814	0.381137	0.269123	1.000000	0.227402
OCT	-0.294323	0.144752	-0.054956	-0.149850	0.046071	0.127951	0.196709	0.144704	0.298151	0.227402	1.000000
NOV	-0.077531	0.165928	0.161269	-0.064477	0.288974	0.031498	0.112004	-0.151322	0.055768	-0.046554	-0.056758
DEC	-0.049529	0.278733	0.040352	0.065763	0.337653	0.090154	0.039772	0.079795	-0.000150	0.073240	-0.026855
ANNUAL	-0.707321	0.302965	0.195393	0.224677	0.452677	0.636692	0.785994	0.755075	0.629581	0.629303	0.386400
Jan-Feb	-0.181229	0.624899	0.810800	0.328052	0.371263	0.075644	0.175961	0.101354	0.131852	0.237410	0.041995
Mar-May	-0.489779	0.304862	0.213342	0.488820	0.634203	0.768722	0.458884	0.499358	0.361450	0.281479	0.062468
Jun-Sep	-0.674418	0.189839	0.129691	0.106128	0.327513	0.542572	0.821251	0.780854	0.635934	0.627832	0.302192
Oct-Dec	-0.310688	0.226917	-0.007222	-0.142865	0.168442	0.141096	0.217790	0.115559	0.303420	0.226139	0.944980
SUBDIVISION_ARUNACHAL PRADESH	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

```
In [50]: import seaborn as sns
sns.heatmap(cor_mat, vmax=1, vmin=1, annot=True, linewidth=.5, cmap='bwr')#plotting of graph using seaborn
```

Out[50]: <Axes: >



```
In [51]: data.isna().sum()
```

```
Out[51]: Unnamed: 0      0  
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 [ ]:
```