In [1]:
```
import pandas as pd
```

In [2]:
```
data=pd.read_csv("/home/placement/Downloads/arunachal.csv")
```

In [3]:
```
data.describe()
```

Out[3]:

	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 [4]:
```
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
```

In [5]:
```
data.groupby(['SUBDIVISION']).count()
```

Out[5]:

SUBDIVISION	Unnamed: 0	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
ARUNACHAL PRADESH	91	91	90	90	89	91	91	90	90	91	91	89	89	89	85	90	89	89	88

In [6]:
```
data.isna().sum()
```

Out[6]:
```
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
```

In [7]:
```
data1=data.loc[(data.YEAR<=2010)]
```

In [8]:
```
data1
```

Out[8]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May
0	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
1	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
2	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
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	102.3	706.0
4	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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

In [9]:
```
data2=data.drop(['ANNUAL','Jan-Feb','Mar-May','Jun-Sep','Oct-Dec'],axis=1)
```

In [10]:
```
data2
```

Out[10]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	NaN	NaN	NaN
1	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
2	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
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
4	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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
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
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
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
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

91 rows × 15 columns

In [11]:
```
data1['SUBDIVISION'].unique()
```

Out[11]:
```
array(['ARUNACHAL PRADESH'], dtype=object)
```

In [27]:
```
data2=data.loc[(data1.SUBDIVISION=='ARUNACHAL PRADESH')]
```

In [28]:
```
data2
```

Out[28]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May
0	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
1	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
2	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
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	102.3	706.0
4	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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

In [29]:
```
data2.isna().sum()
```

Out[29]:
```
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
```

In [34]:
```
data2=data.fillna(data2.mean())
data2
```

Out[34]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.100000	316.1	424.6	1124.9	710.656044	629.7	333.9	198.795556	35.941111	24.638889
1	111	ARUNACHAL PRADESH	1917	21.4	164.5	154.446067	269.6	107.9	823.8	909.100000	628.4	411.5	199.300000	63.500000	0.000000
2	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
3	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
4	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
86	196	ARUNACHAL PRADESH	2005	48.4	167.6	229.500000	195.3	179.8	269.3	430.800000	400.0	243.6	139.300000	28.600000	3.300000
87	197	ARUNACHAL PRADESH	2006	6.0	103.7	63.300000	202.7	321.7	520.4	382.200000	227.6	263.2	77.200000	69.700000	21.700000
88	198	ARUNACHAL PRADESH	2007	13.4	97.4	48.100000	292.4	250.4	530.2	761.000000	364.6	529.3	102.600000	24.300000	6.900000
89	199	ARUNACHAL PRADESH	2008	76.7	39.7	122.600000	192.4	185.0	423.6	456.100000	439.3	189.7	115.100000	1.700000	2.600000
90	200	ARUNACHAL PRADESH	2009	18.0	92.8	72.100000	132.7	189.9	259.1	329.900000	370.3	152.5	82.900000	33.900000	15.900000

91 rows × 20 columns

In [35]:
```
data2['ANNUAL RAIN']=data.apply(lambda row:row.JAN+row.FEB,axis=1)
```

In [36]:
```
data2
```

Out[36]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	...	SEP	OCT	NOV	DEC
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.100000	316.1	424.6	1124.9	710.656044	...	333.9	198.795556	35.941111	24.638889
1	111	ARUNACHAL PRADESH	1917	21.4	164.5	154.446067	269.6	107.9	823.8	909.100000	...	411.5	199.300000	63.500000	0.000000
2	112	ARUNACHAL PRADESH	1918	10.4	11.0	191.200000	144.6	861.1	1609.9	1303.000000	...	515.8	125.200000	7.800000	13.700000
3	113	ARUNACHAL PRADESH	1919	34.5	67.8	28.500000	256.9	420.6	973.6	999.000000	...	628.7	948.300000	40.700000	8.600000
4	114	ARUNACHAL PRADESH	1920	14.0	196.3	605.600000	364.7	173.6	840.6	535.400000	...	376.7	103.300000	0.000000	0.000000
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
86	196	ARUNACHAL PRADESH	2005	48.4	167.6	229.500000	195.3	179.8	269.3	430.800000	...	243.6	139.300000	28.600000	3.300000

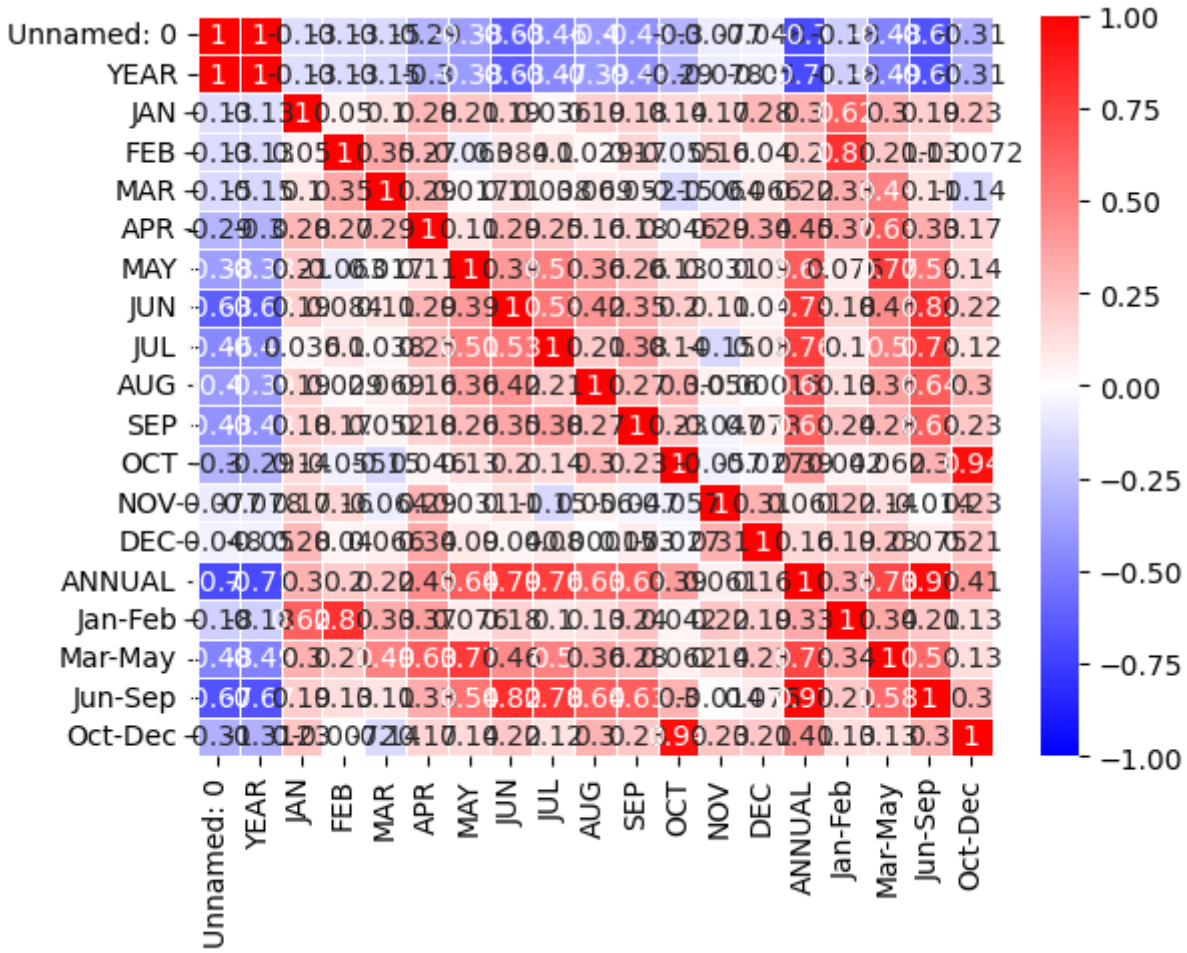
In [37]:
```
cor=data.corr()
cor
```

Out[37]:

	Unnamed: 0	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	
Unnamed: 0	1.000000	0.999610	-0.130520	-0.134521	-0.146730	-0.293626	-0.381310	-0.629492	-0.464201	-0.397151	-0.429514	-0.295578	-0
YEAR	0.999610	1.000000	-0.129778	-0.134399	-0.153629	-0.301073	-0.384602	-0.630502	-0.465649	-0.394444	-0.431541	-0.294323	-0
JAN	-0.130520	-0.129778	1.000000	0.049703	0.102652	0.275465	0.213443	0.187899	0.036231	0.186397	0.180209	0.144752	0
FEB	-0.134521	-0.134399	0.049703	1.000000	0.346448	0.268504	-0.063280	0.084133	0.102494	0.028861	0.168561	-0.054956	0
MAR	-0.146730	-0.153629	0.102652	0.346448	1.000000	0.292072	0.017182	0.110014	0.038299	0.068542	0.052330	-0.149850	-0
APR	-0.293626	-0.301073	0.275465	0.268504	0.292072	1.000000	0.114128	0.293255	0.252243	0.157620	0.176335	0.046071	0
MAY	-0.381310	-0.384602	0.213443	-0.063280	0.017182	0.114128	1.000000	0.393818	0.506982	0.363992	0.258744	0.127951	0
JUN	-0.629492	-0.630502	0.187899	0.084133	0.110014	0.293255	0.393818	1.000000	0.530709	0.417914	0.347814	0.196709	0
JUL	-0.464201	-0.465649	0.036231	0.102494	0.038299	0.252243	0.506982	0.530709	1.000000	0.210551	0.381137	0.144704	-0
AUG	-0.397151	-0.394444	0.186397	0.028861	0.068542	0.157620	0.363992	0.417914	0.210551	1.000000	0.269123	0.298151	0

In [21]:
```
import seaborn as sns
sns.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidth=.5,cmap='bwr')
```

Out[21]: <Axes: >



In [48]:
```
data1=data1.drop(['SUBDIVISION','JAN','FEB','MAR','APR','MAY','JUN','JUL','AUG','SEP','OCT','NOV','DEC','ANN
```

In [49]:
```
data2
```

Out[49]:

	Unnamed: 0	YEAR	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	M
0	110	1916	71.100000	316.1	424.6	1124.9	710.656044	629.7	333.9	198.795556	35.941111	24.638889	3475.443529	117.9	811
1	111	1917	154.446067	269.6	107.9	823.8	909.100000	628.4	411.5	199.300000	63.500000	0.000000	3475.443529	185.9	784
2	112	1918	191.200000	144.6	861.1	1609.9	1303.000000	692.6	515.8	125.200000	7.800000	13.700000	5486.300000	21.4	1196
3	113	1919	28.500000	256.9	420.6	973.6	999.000000	286.7	628.7	948.300000	40.700000	8.600000	4693.900000	102.3	706
4	114	1920	605.600000	364.7	173.6	840.6	535.400000	896.5	376.7	103.300000	0.000000	0.000000	4106.700000	210.3	1143
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
86	196	2005	229.500000	195.3	179.8	269.3	430.800000	400.0	243.6	139.300000	28.600000	3.300000	2335.500000	216.0	604
87	197	2006	63.300000	202.7	321.7	520.4	382.200000	227.6	263.2	77.200000	69.700000	21.700000	2259.600000	109.7	587
88	198	2007	48.100000	292.4	250.4	530.2	761.000000	364.6	529.3	102.600000	24.300000	6.900000	3020.700000	110.8	590
89	199	2008	122.600000	192.4	185.0	423.6	456.100000	439.3	189.7	115.100000	1.700000	2.600000	2244.400000	116.4	499
90	200	2009	72.100000	132.7	189.9	259.1	329.900000	370.3	152.5	82.900000	33.900000	15.900000	1749.900000	110.8	394

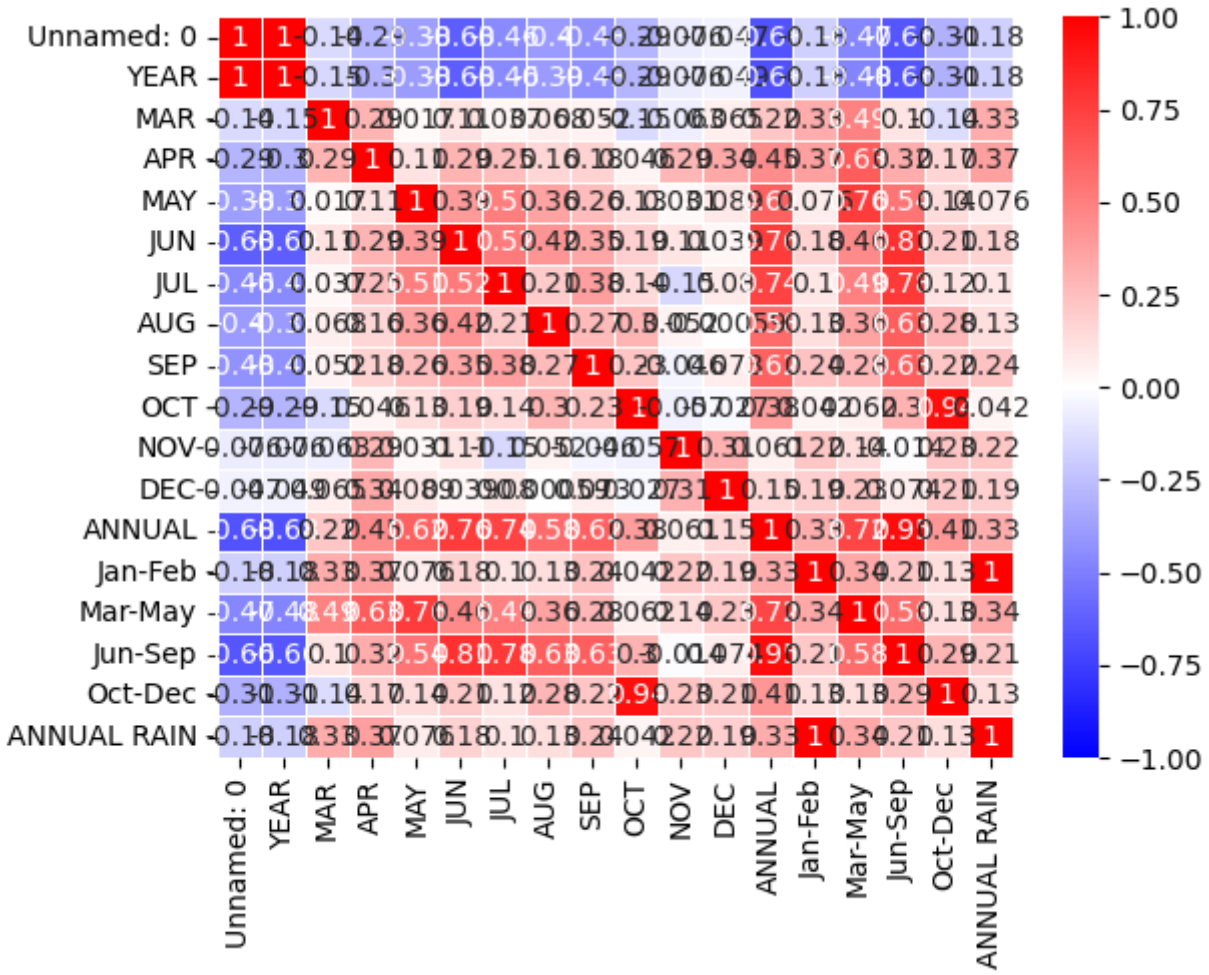
In [50]:
```
cor=data2.corr()
cor
```

Out[50]:

	Unnamed: 0	YEAR	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	A
Unnamed: 0	1.000000	0.999610	-0.144357	-0.293626	-0.381310	-0.628853	-0.456500	-0.397151	-0.429514	-0.290304	-0.075613	-0.047054	-0
YEAR	0.999610	1.000000	-0.151211	-0.301073	-0.384602	-0.629724	-0.458067	-0.394444	-0.431541	-0.289099	-0.076111	-0.048522	-0
MAR	-0.144357	-0.151211	1.000000	0.292034	0.016967	0.109600	0.037387	0.068452	0.052290	-0.148228	-0.063092	0.064708	0
APR	-0.293626	-0.301073	0.292034	1.000000	0.114128	0.289778	0.251912	0.157620	0.176335	0.045964	0.288564	0.337172	0
MAY	-0.381310	-0.384602	0.016967	0.114128	1.000000	0.393296	0.506657	0.363992	0.258744	0.127813	0.031010	0.089021	0
JUN	-0.628853	-0.629724	0.109600	0.289778	0.393296	1.000000	0.521135	0.415872	0.346845	0.192219	0.109060	0.038574	0
JUL	-0.456500	-0.458067	0.037387	0.251912	0.506657	0.521135	1.000000	0.210279	0.380653	0.144514	-0.151320	0.079772	0
AUG	-0.397151	-0.394444	0.068452	0.157620	0.363992	0.415872	0.210279	1.000000	0.269123	0.296405	0.051843	-0.000592	0
SEP	-0.429514	-0.431541	0.052290	0.176335	0.258744	0.346845	0.380653	0.269123	1.000000	0.227142	-0.046062	0.072884	0
OCT	-0.290304	-0.289099	-0.148228	0.045964	0.127813	0.192219	0.144514	0.296405	0.227142	1.000000	-0.056529	-0.026873	0

In [51]:
```
import seaborn as sns
sns.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidth=.5,cmap='bwr')
```

Out[51]: <Axes: >



In [ ]: