In [12]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
from scipy.stats import skew
from scipy.stats import kurtosis
```

In [13]:
```
mba=pd.read_csv("C:/Users/Gopi/Desktop/machine learning/csv files/mba.csv")
del mba['Datasrno']
mba.head()
```

Out[13]:

	workex	gmat
0	21	720
1	107	640
2	57	740
3	99	690
4	208	710

In [14]:
```
len(mba)
```

Out[14]: 773

In [15]:
```
len(mba.columns)
```

Out[15]: 2

In [16]:
```
mba.describe()
```

Out[16]:

	workex	gmat
count	773.000000	773.000000
mean	57.50194	711.164295
std	27.38682	29.339714
min	9.00000	600.000000
25%	41.00000	690.000000
50%	52.00000	710.000000
75%	69.00000	730.000000
max	279.00000	780.000000

In [17]:
```
mba[45:47]
```

Out[17]:

	workex	gmat
45	77	700
46	91	730

In [18]:
```
mba.mean()
```

Out[18]:
```
workex     57.501940
gmat      711.164295
dtype: float64
```

In [19]:
```
mba.std()
```

Out[19]:
```
workex    27.386820
gmat      29.339714
dtype: float64
```

In [20]:
```
mba['gmat'].mean()
```

Out[20]: 711.1642949547219

In [21]:
```
std=mba['gmat'].std()
std
```

Out[21]: 29.339714162029753

In [22]:
```
std=mba['gmat'].mode()
std
```

Out[22]:
```
0    710
dtype: int64
```

In [23]:
```
mba.skew()
```

Out[23]:
```
workex    2.608537
gmat     -0.595477
dtype: float64
```

In [24]:
```
mba.kurt()
```

Out[24]:
```
workex    13.404732
gmat       1.167164
dtype: float64
```

In [25]:
```
from scipy.stats import skew
skew(mba['gmat'])
```

Out[25]: -0.5943203842382888

In [26]:
```
from scipy.stats import kurtosis
mba.kurtosis()
```

Out[26]:
```
workex    13.404732
gmat       1.167164
dtype: float64
```

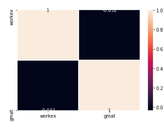
In [27]:
```
plot = mba1.plot(kind='hist')
plot
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
<ipython-input-27-4d905f281622> in <module>
----> 1 plot = mba1.plot(kind='hist')
      2 plot

NameError: name 'mba1' is not defined
```

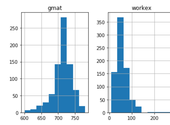
In [28]:
```
sns.heatmap(mba.corr(),annot=True,linewidth=0.5)
```

Out[28]: <matplotlib.axes._subplots.AxesSubplot at 0x69148c8>



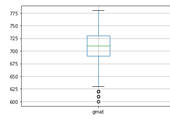
In [30]:
```
mba.hist()
```

Out[30]:
```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x0000000008B21FD8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000000006BF9188>]],
      dtype=object)
```



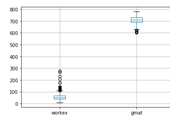
In [32]:
```
mba.boxplot(column ='gmat')
```

Out[32]: <matplotlib.axes._subplots.AxesSubplot at 0xb43c3c8>



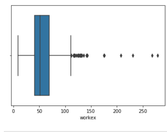
In [33]:
```
mba.boxplot()
```

Out[33]: <matplotlib.axes._subplots.AxesSubplot at 0xb497d88>



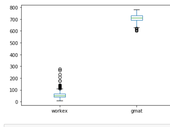
In [34]:
```
sns.boxplot(mba['workex'])
```

Out[34]: <matplotlib.axes._subplots.AxesSubplot at 0xb721f48>



In [36]:
```
plot = mba.plot(kind='box')
plot
```

Out[36]: <matplotlib.axes._subplots.AxesSubplot at 0xb7ad208>



In [37]:
```
from sklearn import preprocessing
```

In [38]:
```
scale=preprocessing.StandardScaler()
```

In [45]:
```
from sklearn.preprocessing import scale
a=scale(mba)
a
```

Out[45]:
```
array([[-1.3336917 ,  0.30134669],
       [ 1.80853813, -2.42709832],
       [-0.01833968,  0.98345794],
       ...,
       [-1.07792881, -3.4502652 ],
       [-1.73560482, -3.4502652 ],
       [-0.20102746, -3.10920957]])
```

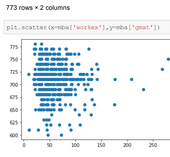
In [54]:
```
s=pd.DataFrame(a,columns=['workex1','gmat1'])
s
```

Out[54]:

	workex1	gmat1
0	-1.333692	0.301347
1	1.808538	-2.427098
2	-0.018340	0.983458
3	1.516238	-0.721820
4	5.498831	-0.039709
...	...	...
768	1.114325	-3.109210
769	2.721977	-1.403931
770	-1.077929	-3.450265
771	-1.735605	-3.450265
772	-0.201027	-3.109210

773 rows × 2 columns

In [52]:
```
plt.scatter(x=mba['workex'],y=mba['gmat'])
```



In [55]:
```
plt.scatter(x=s['workex1'],y=s['gmat1'])
```

Out[55]: <matplotlib.collections.PathCollection at 0x5542908>