In [2]:
```
import pandas as pd
```

In [4]:
```
df = pd.read_csv("student.csv")
```

In [82]:
```
df.describe
```

Out[82]:
```
<bound method NDFrame.describe of     id         name  class   mark  gender
0    1     John Deo   Four  75.00  female
1    2     Max Ruin  Three  85.00    male
2    3       Arnold  Three  55.00    male
3    4   Krish Star   Four  60.00  female
4    5    John Mike   Four  60.00  female
5    6    Alex John   Four  55.00    male
6    7  My John Rob  Fifth  98.45    male
7    8       Asruid   Five  85.00    male
8    9      Tes Qry    Six  78.00    male
9   10     Big John   Four  55.00  female
10  11       Ronald    Six  89.00  female
11  12        Recky    Six  94.00  female
12  13          Kty  Seven  88.00  female
13  14         Bigy  Seven  88.00  female
14  15     Tade Row   Four  88.00    male
15  16        Gimmy   Four  88.00    male
16  17        Tumyu    Six  54.00    male
17  18        Honny   Five  75.00    male
18  19        Tinny   Nine  18.00    male
19  20       Jackly   Nine  65.00  female
20  21   Babby John   Four  69.00  female
21  22       Reggid  Seven  55.00  female
22  23        Herod  Eight  79.00    male
23  24    Tiddy Now  Seven  78.00    male
24  25     Giff Tow  Seven  88.00    male
25  26       Crelea  Seven  79.00    male
26  27     Big Nose  Three  81.00  female
27  28    Rojj Base  Seven  86.00  female
28  29  Tess Played  Seven  55.00    male
29  30    Reppy Red    Six  79.00  female
30  31  Marry Toeey   Four  88.00    male
31  32    Binn Rott  Seven  90.00  female
32  33    Kenn Rein    Six  96.00  female
33  34     Gain Toe  Seven  69.00    male
34  35   Rows Noump    Six  88.00  female>
```

In [10]:
```
df.shape
```

Out[10]:
```
(35, 5)
```

In [12]:
```
df.head()
```

Out[12]:

	id	name	class	mark	gender
0	1	John Deo	Four	75	female
1	2	Max Ruin	Three	85	male
2	3	Arnold	Three	55	male
3	4	Krish Star	Four	60	female
4	5	John Mike	Four	60	female

In [14]:
```
df.tail()
```

Out[14]:

	id	name	class	mark	gender
30	31	Marry Toeey	Four	88	male
31	32	Binn Rott	Seven	90	female
32	33	Kenn Rein	Six	96	female
33	34	Gain Toe	Seven	69	male
34	35	Rows Noump	Six	88	female

In [17]:
```
df.count()
```

Out[17]:
```
id        35
name      35
class     35
mark      35
gender    35
dtype: int64
```

In [19]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35 entries, 0 to 34
Data columns (total 5 columns):
 #   Column  Non-Null Count  Dtype 
---  ------  --------------  ----- 
 0   id      35 non-null     int64 
 1   name    35 non-null     object
 2   class   35 non-null     object
 3   mark    35 non-null     int64 
 4   gender  35 non-null     object
dtypes: int64(2), object(3)
memory usage: 1.5+ KB
```

In [85]:
```
df.isnull()
```

Out[85]:

	id	name	class	mark	gender
0	False	False	False	False	False
1	False	False	False	False	False
2	False	False	False	False	False
3	False	False	False	False	False
4	False	False	False	False	False
5	False	False	False	False	False
6	False	False	False	False	False
7	False	False	False	False	False
8	False	False	False	False	False
9	False	False	False	False	False
10	False	False	False	False	False
11	False	False	False	False	False
12	False	False	False	False	False
13	False	False	False	False	False
14	False	False	False	False	False
15	False	False	False	False	False
16	False	False	False	False	False
17	False	False	False	False	False
18	False	False	False	False	False
19	False	False	False	False	False
20	False	False	False	False	False
21	False	False	False	False	False
22	False	False	False	False	False
23	False	False	False	False	False
24	False	False	False	False	False
25	False	False	False	False	False
26	False	False	False	False	False
27	False	False	False	False	False
28	False	False	False	False	False
29	False	False	False	False	False
30	False	False	False	False	False
31	False	False	False	False	False
32	False	False	False	False	False
33	False	False	False	False	False
34	False	False	False	False	False

In [22]:
```
df.isnull().sum()
```

Out[22]:
```
id        0
name      0
class     0
mark      0
gender    0
dtype: int64
```

In [24]:
```
df.dropna()
```

Out[24]:

	id	name	class	mark	gender
0	1	John Deo	Four	75	female
1	2	Max Ruin	Three	85	male
2	3	Arnold	Three	55	male
3	4	Krish Star	Four	60	female
4	5	John Mike	Four	60	female
5	6	Alex John	Four	55	male
6	7	My John Rob	Fifth	78	male
7	8	Asruid	Five	85	male
8	9	Tes Qry	Six	78	male
9	10	Big John	Four	55	female
10	11	Ronald	Six	89	female
11	12	Recky	Six	94	female
12	13	Kty	Seven	88	female
13	14	Bigy	Seven	88	female
14	15	Tade Row	Four	88	male
15	16	Gimmy	Four	88	male
16	17	Tumyu	Six	54	male
17	18	Honny	Five	75	male
18	19	Tinny	Nine	18	male
19	20	Jackly	Nine	65	female
20	21	Babby John	Four	69	female
21	22	Reggid	Seven	55	female
22	23	Herod	Eight	79	male
23	24	Tiddy Now	Seven	78	male
24	25	Giff Tow	Seven	88	male
25	26	Crelea	Seven	79	male
26	27	Big Nose	Three	81	female
27	28	Rojj Base	Seven	86	female
28	29	Tess Played	Seven	55	male
29	30	Reppy Red	Six	79	female
30	31	Marry Toeey	Four	88	male
31	32	Binn Rott	Seven	90	female
32	33	Kenn Rein	Six	96	female
33	34	Gain Toe	Seven	69	male
34	35	Rows Noump	Six	88	female

In [26]:
```
df.fillna(0)
```

Out[26]:

	id	name	class	mark	gender
0	1	John Deo	Four	75	female
1	2	Max Ruin	Three	85	male
2	3	Arnold	Three	55	male
3	4	Krish Star	Four	60	female
4	5	John Mike	Four	60	female
5	6	Alex John	Four	55	male
6	7	My John Rob	Fifth	78	male
7	8	Asruid	Five	85	male
8	9	Tes Qry	Six	78	male
9	10	Big John	Four	55	female
10	11	Ronald	Six	89	female
11	12	Recky	Six	94	female
12	13	Kty	Seven	88	female
13	14	Bigy	Seven	88	female
14	15	Tade Row	Four	88	male
15	16	Gimmy	Four	88	male
16	17	Tumyu	Six	54	male
17	18	Honny	Five	75	male
18	19	Tinny	Nine	18	male
19	20	Jackly	Nine	65	female
20	21	Babby John	Four	69	female
21	22	Reggid	Seven	55	female
22	23	Herod	Eight	79	male
23	24	Tiddy Now	Seven	78	male
24	25	Giff Tow	Seven	88	male
25	26	Crelea	Seven	79	male
26	27	Big Nose	Three	81	female
27	28	Rojj Base	Seven	86	female
28	29	Tess Played	Seven	55	male
29	30	Reppy Red	Six	79	female
30	31	Marry Toeey	Four	88	male
31	32	Binn Rott	Seven	90	female
32	33	Kenn Rein	Six	96	female
33	34	Gain Toe	Seven	69	male
34	35	Rows Noump	Six	88	female

In [28]:
```
df['class'].fillna('TE')
```

Out[28]:
```
0      Four
1     Three
2     Three
3      Four
4      Four
5      Four
6     Fifth
7      Five
8       Six
9      Four
10      Six
11      Six
12    Seven
13    Seven
14     Four
15     Four
16      Six
17     Five
18     Nine
19     Nine
20     Four
21    Seven
22    Eight
23    Seven
24    Seven
25    Seven
26    Three
27    Seven
28    Seven
29      Six
30     Four
31    Seven
32      Six
33    Seven
34      Six
Name: class, dtype: object
```

In [32]:
```
df['mark'].fillna(df['mark'].mean())
```

Out[32]:
```
0     75
1     85
2     55
3     60
4     60
5     55
6     78
7     85
8     78
9     55
10    89
11    94
12    88
13    88
14    88
15    88
16    54
17    75
18    18
19    65
20    69
21    55
22    79
23    78
24    88
25    79
26    81
27    86
28    55
29    79
30    88
31    90
32    96
33    69
34    88
Name: mark, dtype: int64
```

In [34]:
```
df['class'].value_counts()
```

Out[34]:
```
Seven    10
Four      9
Six       7
Three     3
Five      2
Nine      2
Fifth     1
Eight     1
Name: class, dtype: int64
```

In [36]:
```
df.fillna(method='backfill')
```

Out[36]:

	id	name	class	mark	gender
0	1	John Deo	Four	75	female
1	2	Max Ruin	Three	85	male
2	3	Arnold	Three	55	male
3	4	Krish Star	Four	60	female
4	5	John Mike	Four	60	female
5	6	Alex John	Four	55	male
6	7	My John Rob	Fifth	78	male
7	8	Asruid	Five	85	male
8	9	Tes Qry	Six	78	male
9	10	Big John	Four	55	female
10	11	Ronald	Six	89	female
11	12	Recky	Six	94	female
12	13	Kty	Seven	88	female
13	14	Bigy	Seven	88	female
14	15	Tade Row	Four	88	male
15	16	Gimmy	Four	88	male
16	17	Tumyu	Six	54	male
17	18	Honny	Five	75	male
18	19	Tinny	Nine	18	male
19	20	Jackly	Nine	65	female
20	21	Babby John	Four	69	female
21	22	Reggid	Seven	55	female
22	23	Herod	Eight	79	male
23	24	Tiddy Now	Seven	78	male
24	25	Giff Tow	Seven	88	male
25	26	Crelea	Seven	79	male
26	27	Big Nose	Three	81	female
27	28	Rojj Base	Seven	86	female
28	29	Tess Played	Seven	55	male
29	30	Reppy Red	Six	79	female
30	31	Marry Toeey	Four	88	male
31	32	Binn Rott	Seven	90	female
32	33	Kenn Rein	Six	96	female
33	34	Gain Toe	Seven	69	male
34	35	Rows Noump	Six	88	female

In [38]:
```
df.fillna(method='pad')
```

Out[38]:

	id	name	class	mark	gender
0	1	John Deo	Four	75	female
1	2	Max Ruin	Three	85	male
2	3	Arnold	Three	55	male
3	4	Krish Star	Four	60	female
4	5	John Mike	Four	60	female
5	6	Alex John	Four	55	male
6	7	My John Rob	Fifth	78	male
7	8	Asruid	Five	85	male
8	9	Tes Qry	Six	78	male
9	10	Big John	Four	55	female
10	11	Ronald	Six	89	female
11	12	Recky	Six	94	female
12	13	Kty	Seven	88	female
13	14	Bigy	Seven	88	female
14	15	Tade Row	Four	88	male
15	16	Gimmy	Four	88	male
16	17	Tumyu	Six	54	male
17	18	Honny	Five	75	male
18	19	Tinny	Nine	18	male
19	20	Jackly	Nine	65	female
20	21	Babby John	Four	69	female
21	22	Reggid	Seven	55	female
22	23	Herod	Eight	79	male
23	24	Tiddy Now	Seven	78	male
24	25	Giff Tow	Seven	88	male
25	26	Crelea	Seven	79	male
26	27	Big Nose	Three	81	female
27	28	Rojj Base	Seven	86	female
28	29	Tess Played	Seven	55	male
29	30	Reppy Red	Six	79	female
30	31	Marry Toeey	Four	88	male
31	32	Binn Rott	Seven	90	female
32	33	Kenn Rein	Six	96	female
33	34	Gain Toe	Seven	69	male
34	35	Rows Noump	Six	88	female

In [40]:
```
df.describe()
```

Out[40]:

	id	mark
count	35.000000	35.000000
mean	18.000000	74.657143
std	10.246951	16.401117
min	1.000000	18.000000
25%	9.500000	62.500000
50%	18.000000	79.000000
75%	26.500000	88.000000
max	35.000000	96.000000

In [43]:
```
import numpy as np
x = np.array([5,4,3,2,7,8,98,28])
```

In [45]:
```
np.mean(x)
```

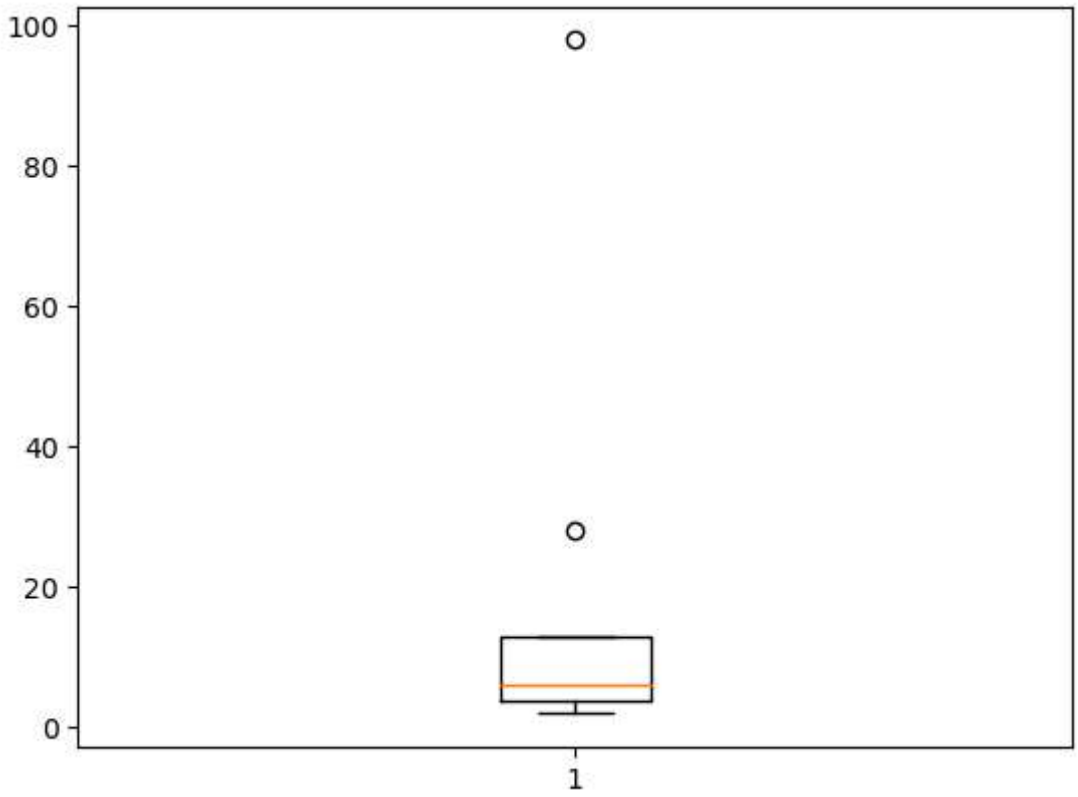
Out[45]:
```
19.375
```

In [47]:
```
np.median(x)
```

Out[47]:
```
6.0
```

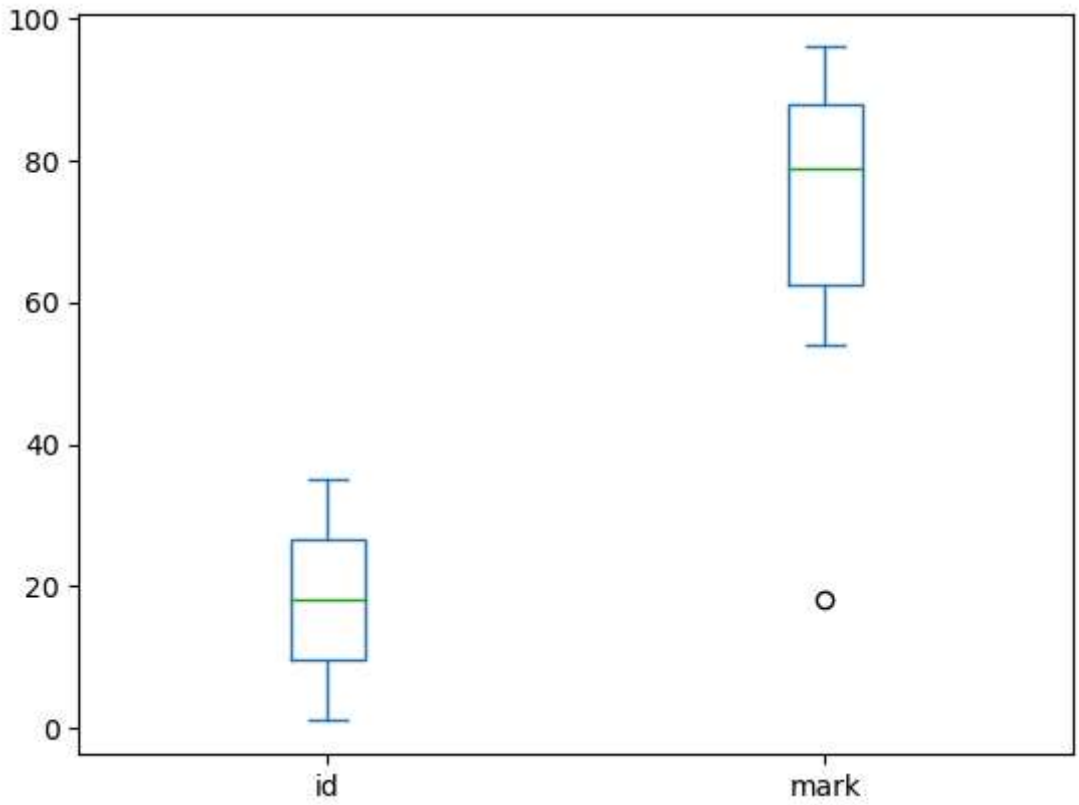
In [49]:
```
import matplotlib.pyplot as plt
```

In [51]:
```
plt.boxplot(x);
```



In [53]:
```
df.plot.box()
```

Out[53]: <Axes: >



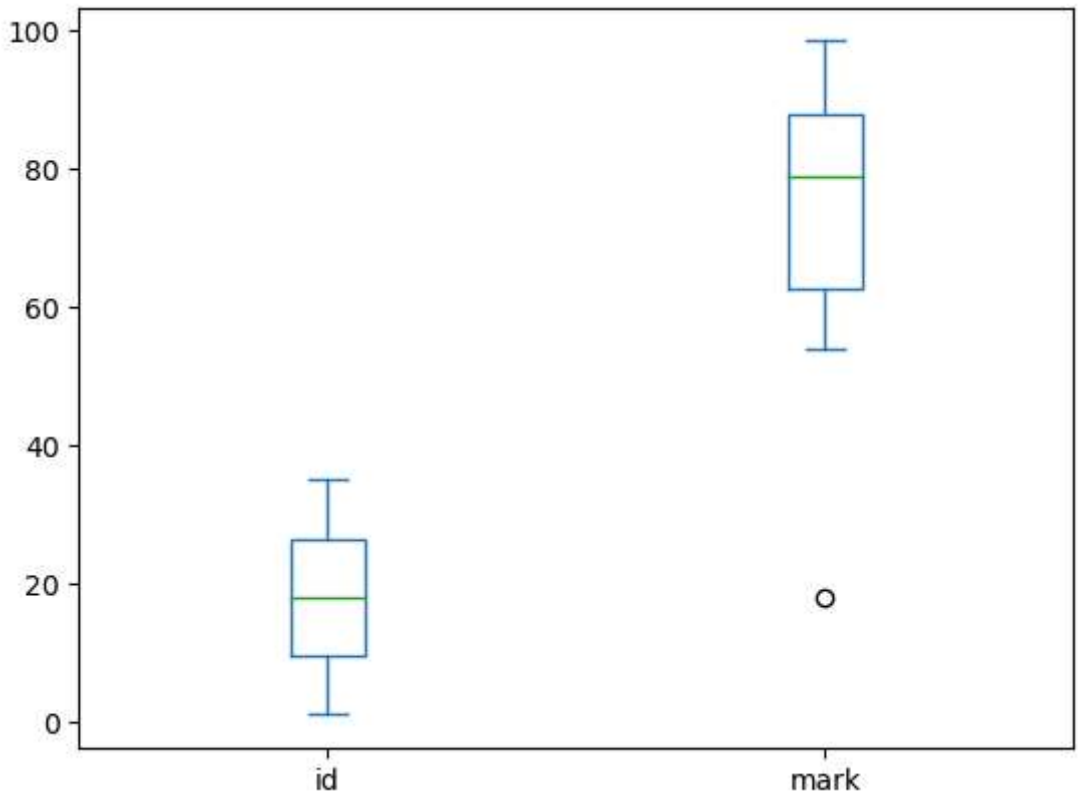
In [55]:
```
df.loc[6,'mark']
```

Out[55]: 78

In [57]:
```
df.loc[6,'mark']=98.45
```

In [59]:
```
df.plot.box()
```

Out[59]: <Axes: >

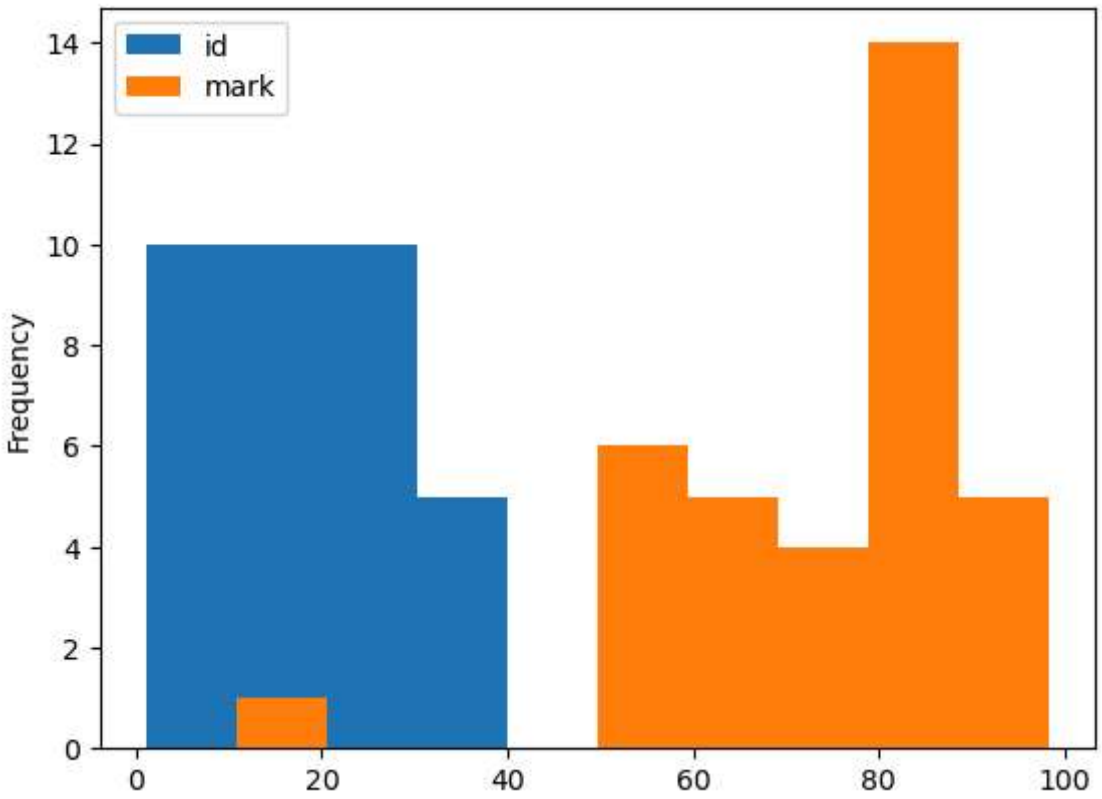


In [61]:
```
df.loc[6,'mark']
```

Out[61]: 98.45

In [63]:
```
df.plot.hist()
```

Out[63]: <Axes: ylabel='Frequency'>



In [71]:
```
x= df[['id','mark']]
```

In [73]:
```
x.describe()
```

Out[73]:

	id	mark
count	35.000000	35.000000
mean	18.000000	75.241429
std	10.246951	16.880952
min	1.000000	18.000000
25%	9.500000	62.500000
50%	18.000000	79.000000
75%	26.500000	88.000000
max	35.000000	98.450000

In [75]:
```
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
x_scaled = scaler.fit_transform(x)
```

In [77]:
```
pd.DataFrame(x_scaled).describe()
```

Out[77]:

	0	1
count	35.000000	35.000000
mean	0.500000	0.711516
std	0.301381	0.209832
min	0.000000	0.000000
25%	0.250000	0.553139
50%	0.500000	0.758235
75%	0.750000	0.870106
max	1.000000	1.000000

In [79]:
```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
x_scaled = scaler.fit_transform(x)
```

In [80]:
```
pd.DataFrame(x_scaled).describe()
```

Out[80]:

	0	1
count	3.500000e+01	3.500000e+01
mean	-1.268826e-17	-1.586033e-17
std	1.014599e+00	1.014599e+00
min	-1.683251e+00	-3.440393e+00
25%	-8.416254e-01	-7.658007e-01
50%	0.000000e+00	2.259022e-01
75%	8.416254e-01	7.668310e-01
max	1.683251e+00	1.394910e+00