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

In [2]:
```
df=pd.read_csv(r"C:\Users\Admin\Downloads\4_drug200 - 4_drug200.csv")
df
```

Out[2]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	drugY
...	...	...	...	...	...	...
195	56	F	LOW	HIGH	11.567	drugC
196	16	M	LOW	HIGH	12.006	drugC
197	52	M	NORMAL	HIGH	9.894	drugX
198	23	M	NORMAL	NORMAL	14.020	drugX
199	40	F	LOW	NORMAL	11.349	drugX

200 rows × 6 columns

In [3]:
```
df.isnull()
```

Out[3]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
...	...	...	...	...	...	...
195	False	False	False	False	False	False
196	False	False	False	False	False	False
197	False	False	False	False	False	False
198	False	False	False	False	False	False
199	False	False	False	False	False	False

200 rows × 6 columns

In [4]:
```
df.fillna(0,inplace=True)
```

In [5]:
```
df.shape
```

Out[5]:
```
(200, 6)
```

In [6]:
```
df.size
```

Out[6]:
```
1200
```

In [7]:
```
df.ndim
```

Out[7]:
```
2
```

In [8]:
```
df.head()
```

Out[8]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	drugY

In [9]:
```
df.tail()
```

Out[9]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
195	56	F	LOW	HIGH	11.567	drugC
196	16	M	LOW	HIGH	12.006	drugC
197	52	M	NORMAL	HIGH	9.894	drugX
198	23	M	NORMAL	NORMAL	14.020	drugX
199	40	F	LOW	NORMAL	11.349	drugX

In [10]:
```
df.dropna(axis=1,how='any')
```

Out[10]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	drugY
...	...	...	...	...	...	...
195	56	F	LOW	HIGH	11.567	drugC
196	16	M	LOW	HIGH	12.006	drugC
197	52	M	NORMAL	HIGH	9.894	drugX
198	23	M	NORMAL	NORMAL	14.020	drugX
199	40	F	LOW	NORMAL	11.349	drugX

200 rows × 6 columns

In [11]:
```
df.loc[10:35]
```

Out[11]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
10	47	F	LOW	HIGH	11.767	drugC
11	34	F	HIGH	NORMAL	19.199	drugY
12	43	M	LOW	HIGH	15.376	drugY
13	74	F	LOW	HIGH	20.942	drugY
14	50	F	NORMAL	HIGH	12.703	drugX
15	16	F	HIGH	NORMAL	15.516	drugY
16	69	M	LOW	NORMAL	11.455	drugX
17	43	M	HIGH	HIGH	13.972	drugA
18	23	M	LOW	HIGH	7.298	drugC
19	32	F	HIGH	NORMAL	25.974	drugY
20	57	M	LOW	NORMAL	19.128	drugY
21	63	M	NORMAL	HIGH	25.917	drugY
22	47	M	LOW	NORMAL	30.568	drugY
23	48	F	LOW	HIGH	15.036	drugY
24	33	F	LOW	HIGH	33.486	drugY
25	28	F	HIGH	NORMAL	18.809	drugY
26	31	M	HIGH	HIGH	30.366	drugY
27	49	F	NORMAL	NORMAL	9.381	drugX
28	39	F	LOW	NORMAL	22.697	drugY
29	45	M	LOW	HIGH	17.951	drugY
30	18	F	NORMAL	NORMAL	8.750	drugX
31	74	M	HIGH	HIGH	9.567	drugB
32	49	M	LOW	NORMAL	11.014	drugX
33	65	F	HIGH	NORMAL	31.876	drugY
34	53	M	NORMAL	HIGH	14.133	drugX
35	46	M	NORMAL	NORMAL	7.285	drugX

In [28]:
```
df2=df[df['Age']<=18]
df2
```

Out[28]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
15	16	F	HIGH	NORMAL	15.516	drugY
30	18	F	NORMAL	NORMAL	8.750	drugX
39	15	M	NORMAL	HIGH	9.084	drugX
67	17	M	NORMAL	NORMAL	10.832	drugX
69	18	F	HIGH	NORMAL	24.276	drugY
99	15	F	HIGH	NORMAL	16.725	drugY
121	15	M	HIGH	NORMAL	17.206	drugY
164	16	M	HIGH	NORMAL	19.007	drugY
184	18	F	HIGH	HIGH	37.188	drugY
196	16	M	LOW	HIGH	12.006	drugC

In [29]:
```
df.iloc[1:3]
```

Out[29]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC

In [30]:
```
df.index
```

Out[30]:
```
RangeIndex(start=0, stop=200, step=1)
```

In [31]:
```
df.columns
```

Out[31]:
```
Index(['Age', 'Sex', 'BP', 'Cholesterol', 'Na_to_K', 'Drug'], dtype='object')
```

In [32]:
```
df.describe()
```

Out[32]:

	Age	Na_to_K
count	200.000000	200.000000
mean	44.315000	16.084485
std	16.544315	7.223956
min	15.000000	6.269000
25%	31.000000	10.445500
50%	45.000000	13.936500
75%	58.000000	19.380000
max	74.000000	38.247000

In [34]:
```
df1=df2[['Age','Na_to_K']]
df1
```

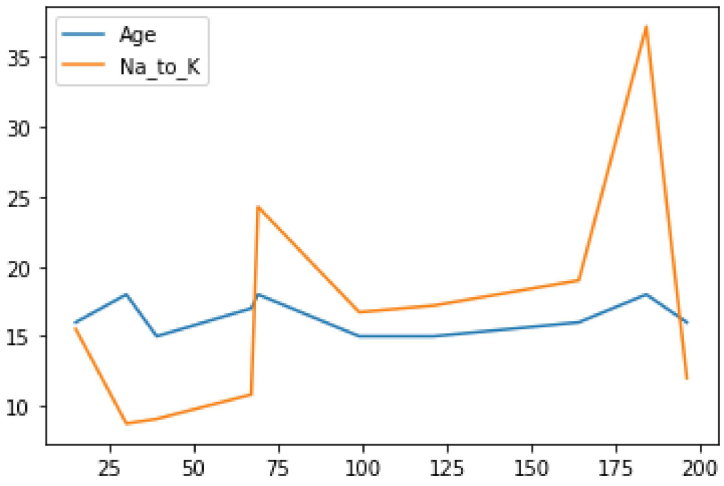
Out[34]:

	Age	Na_to_K
15	16	15.516
30	18	8.750
39	15	9.084
67	17	10.832
69	18	24.276
99	15	16.725
121	15	17.206
164	16	19.007
184	18	37.188
196	16	12.006

In [35]:
```
import matplotlib.pyplot as pp
```

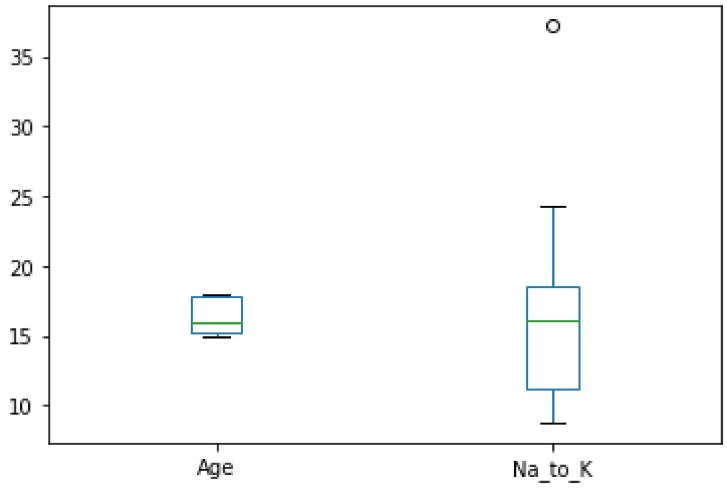
In [36]:
```
df1.plot.line()
```

Out[36]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x142d48d9580>
```



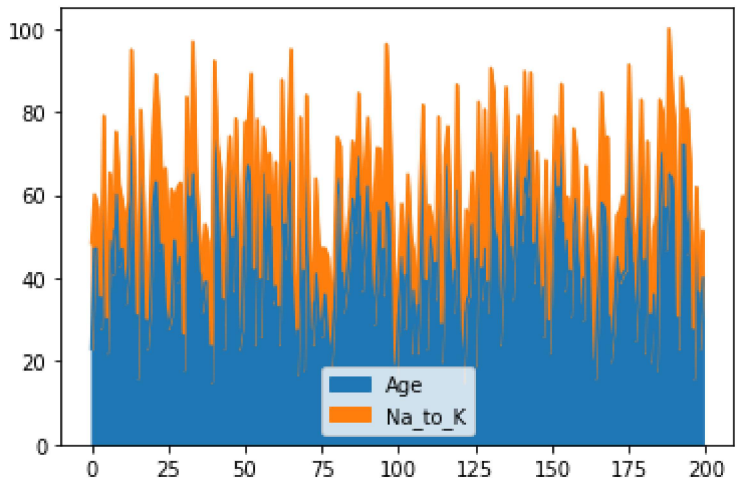
In [37]:
```
df1.plot.box()
```

Out[37]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x142d44a82e0>
```



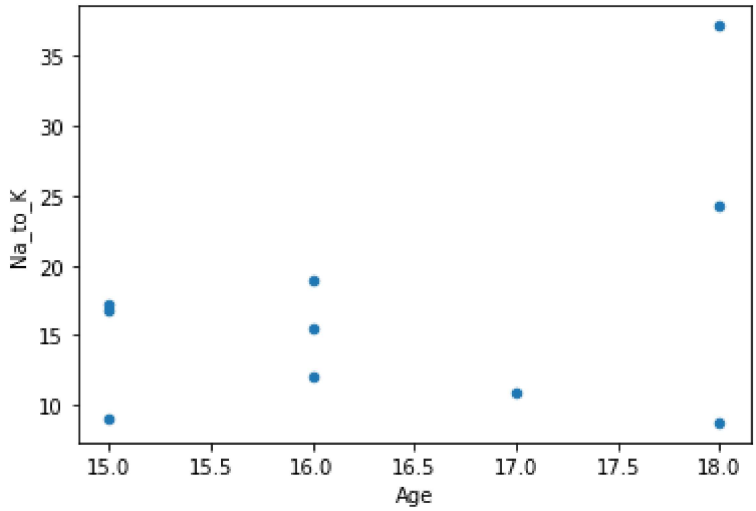
In [38]:
```
df.plot.area()
```

Out[38]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x142d48c9fd0>
```



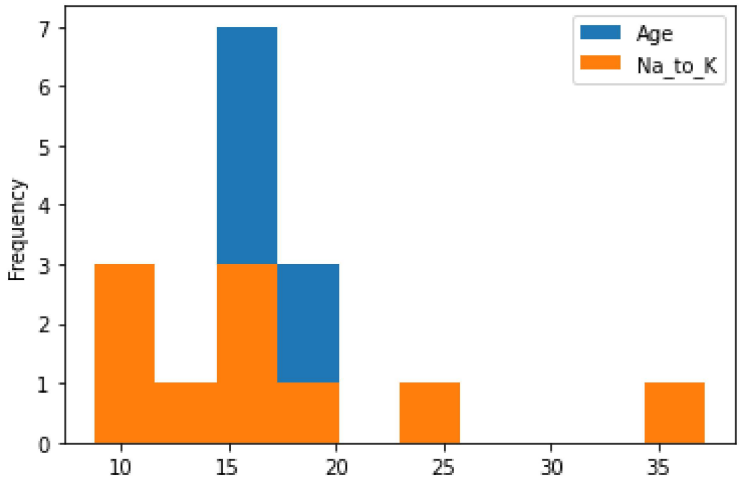
In [39]:
```
df1.plot.scatter(x='Age',y='Na_to_K')
```

Out[39]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x142d4e0d160>
```



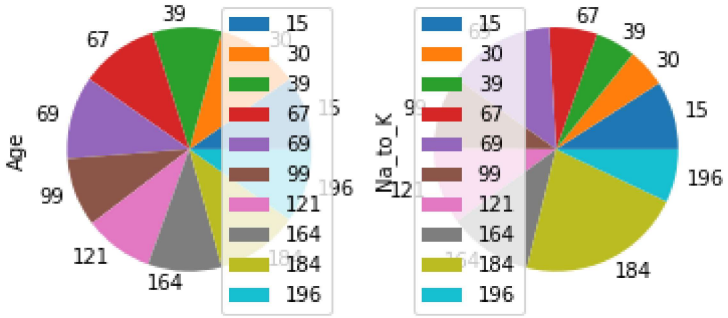
In [40]:
```
df1.plot.hist()
```

Out[40]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x142d4e53760>
```



In [41]:
```
df1.plot.pie(subplots=True)
```

Out[41]:
```
array([<matplotlib.axes._subplots.AxesSubplot object at 0x00000142D4EF4A00>,
       <matplotlib.axes._subplots.AxesSubplot object at 0x00000142D4F1E2B0>],
      dtype=object)
```



In [ ]:

In [ ]: