

In [1]:

```
import numpy as np
import pandas as pd
```

In [2]:

```
a=pd.read_csv(r"C:\Users\user\Downloads\4_drug200 - 4_drug200.csv")
```

To print the first 1000 rows

In [3]:

```
a.head(100)
```

Out[3]:

	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
...	...	...	...	...	...	...
95	36	M	LOW	NORMAL	11.424	drugX
96	58	F	LOW	HIGH	38.247	drugY
97	56	F	HIGH	HIGH	25.395	drugY
98	20	M	HIGH	NORMAL	35.639	drugY
99	15	F	HIGH	NORMAL	16.725	drugY

100 rows × 6 columns

To print the last 5 rows

In [4]:

```
a.tail(5)
```

Out[4]:

	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

To print the size function

In [5]:

```
print(np.size(a))
```

1200

To print shape function

In [6]:

```
print(np.shape(a))
```

(200, 6)

To print the na function

In [7]:

```
pd.isna(a)
```

Out[7]:

	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

To print the na function

In [8]:

```
pd.isna(a)
```

Out[8]:

	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

To print the na function

In [9]:

```
pd.isna(a)
```

Out[9]:

	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

To print the drop function

In [10]:

```
a.dropna()
```

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

To print the fill function

In [11]:

```
a.fillna(value=10)
```

Out[11]:

	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

To describe the function

In [12]:

```
a.describe()
```

Out[12]:

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

In [13]:

```
conda install matplotlib
```

```
Collecting package metadata (current_repodata.json): ...working... done  
Solving environment: ...working... done
```

```
# All requested packages already installed.
```

Note: you may need to restart the kernel to use updated packages.

```
==> WARNING: A newer version of conda exists. <==  
current version: 4.10.1  
latest version: 23.5.2
```

Please update conda by running

```
$ conda update -n base -c defaults conda
```

In [14]:

```
import matplotlib.pyplot as pp
```

In [23]:

```
b=a[["Age","Na_to_K"]]  
b
```

Out[23]:

	Age	Na_to_K
0	23	25.355
1	47	13.093
2	47	10.114
3	28	7.798
4	61	18.043
...	...	...
195	56	11.567
196	16	12.006
197	52	9.894
198	23	14.020
199	40	11.349

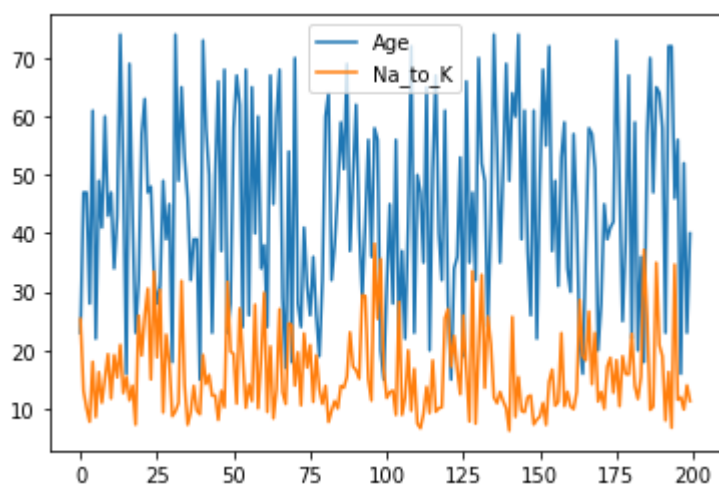
200 rows × 2 columns

In [24]:

```
b.plot.line()
```

Out[24]:

&lt;AxesSubplot:&gt;

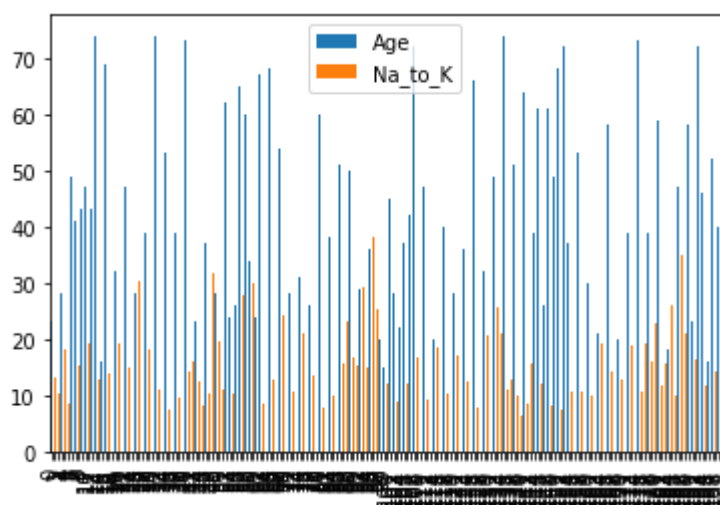


In [25]:

```
b.plot.bar()
```

Out[25]:

<AxesSubplot:>

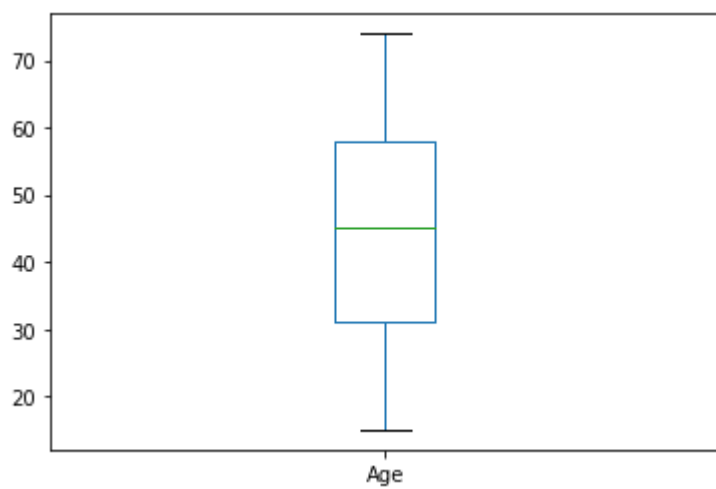


In [19]:

```
b.plot.box()
```

Out[19]:

<AxesSubplot:>

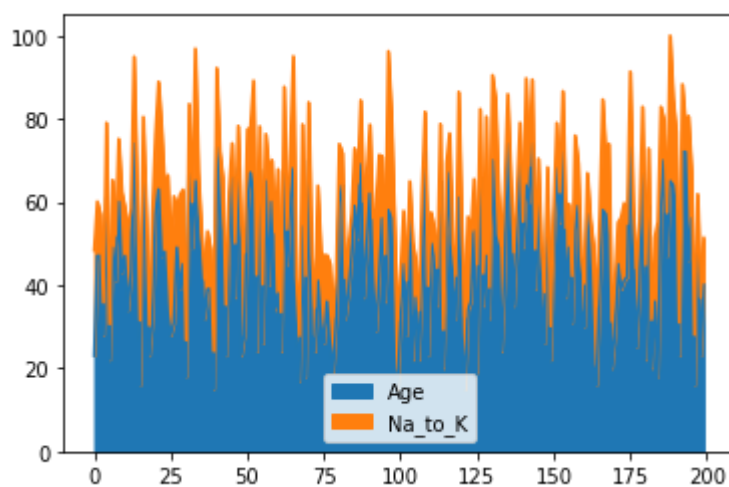


In [26]:

```
b.plot.area()
```

Out[26]:

<AxesSubplot:>

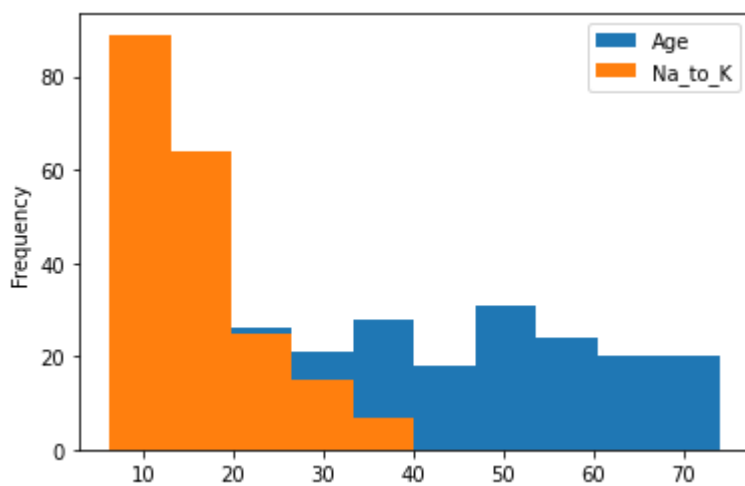


In [27]:

```
b.plot.hist()
```

Out[27]:

<AxesSubplot:ylabel='Frequency'>



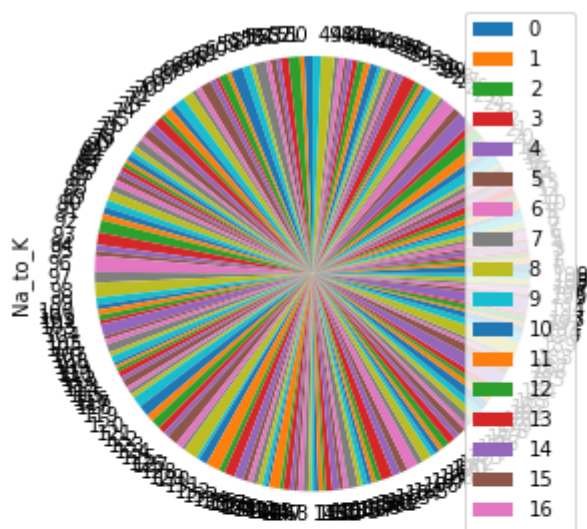


In [28]:

```
b.plot.pie(y='Na_to_K',figsize=(5,5))
```

Out[28]:

<AxesSubplot:ylabel='Na\_to\_K'>

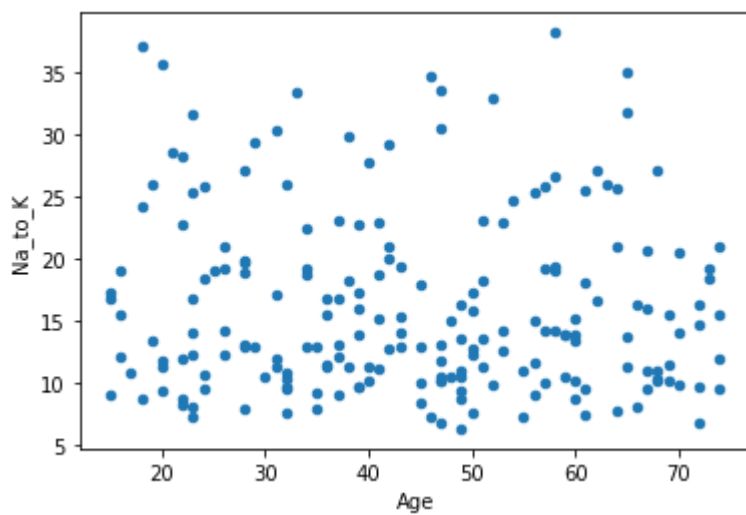


In [29]:

```
b.plot.scatter(x='Age',y='Na_to_K')
```

Out[29]:

<AxesSubplot:xlabel='Age', ylabel='Na\_to\_K'>



In [ ]: