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

IMPORT AND PRINT DATA SET

In [2]:
```
data = pd.read_csv("drug.csv")
data
```

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

SHAPE

In [3]:
```
np.shape(data)
```

Out[3]: (200, 6)

SIZE

In [4]:
```
np.size(data)
```

Out[4]: 1200

PRINT FIRST 10 VALUES

In [5]:
```
data.head(10)
```

Out[5]:

	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
5	22	F	NORMAL	HIGH	8.607	drugX
6	49	F	NORMAL	HIGH	16.275	drugY
7	41	M	LOW	HIGH	11.037	drugC
8	60	M	NORMAL	HIGH	15.171	drugY
9	43	M	LOW	NORMAL	19.368	drugY

PRINT LAST 7 VALUES

In [6]:
```
data.tail(5)
```

Out[6]:

	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

DESCRIPTION OF TABLE

In [7]:
```
data.describe()
```

Out[7]:

	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

	Age	Na_to_K
50%	45.000000	13.936500
75%	58.000000	19.380000

FIND NULL VALUES

In [8]:
```
data.isna()
```

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

FILL NULL VALUES

In [9]:
```
data.fillna(1)
```

Out[9]:

	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

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
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]:
```
data.columns
```

Out[10]:
```
Index(['Age', 'Sex', 'BP', 'Cholesterol', 'Na_to_K', 'Drug'], dtype='object')
```

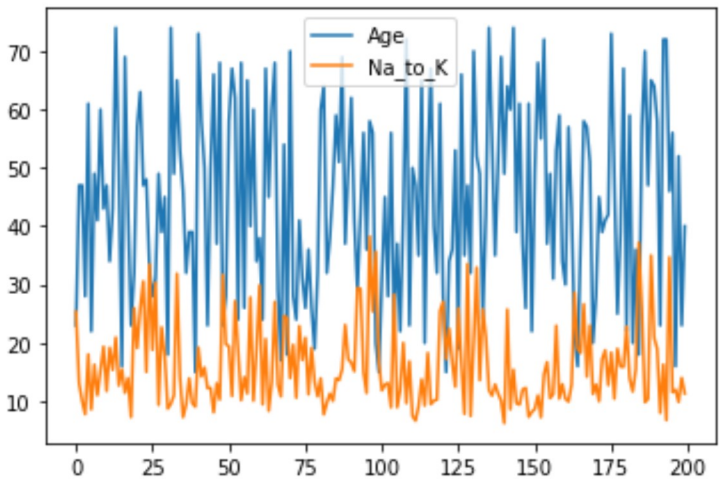
In [11]:
```
data.index
```

Out[11]:
```
RangeIndex(start=0, stop=200, step=1)
```

LINE PLOT

In [12]:
```
data.plot.line()
```

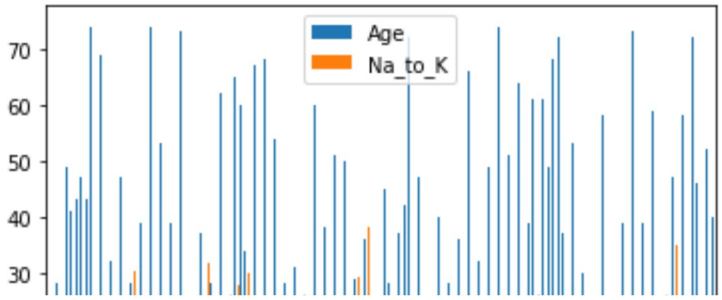
Out[12]:
```
<AxesSubplot:>
```



BAR CHART

In [13]:
```
data.plot.bar()
```

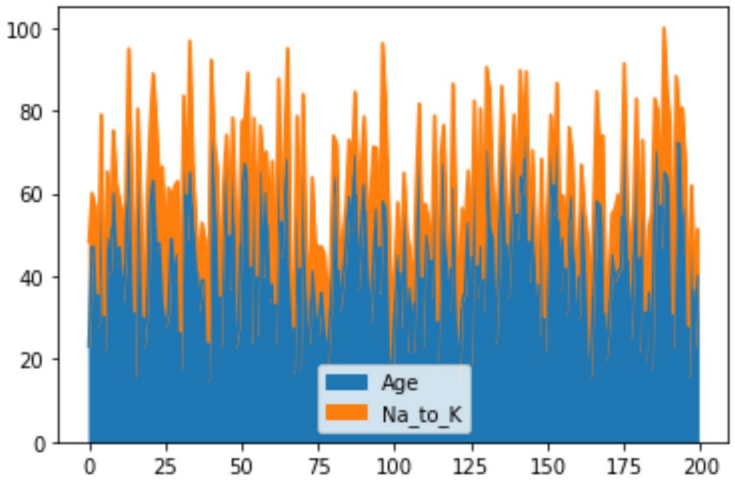
Out[13]:
```
<AxesSubplot:>
```



AREA CHART

In [14]:
```
data.plot.area()
```

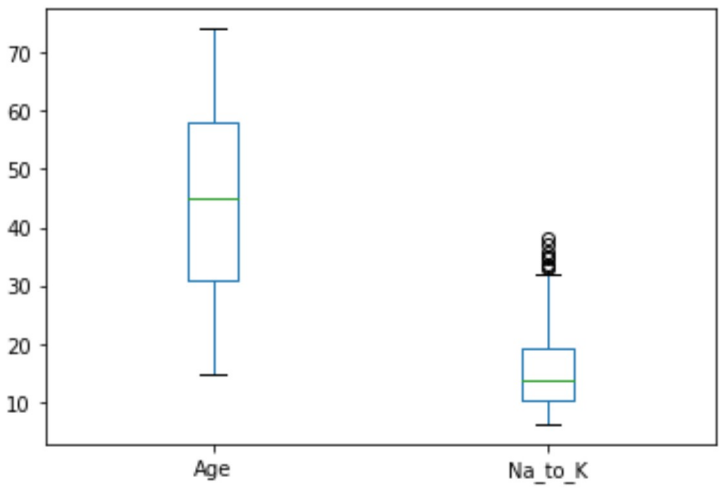
Out[14]: <AxesSubplot:>



BOX PLOT

In [15]:
```
data.plot.box()
```

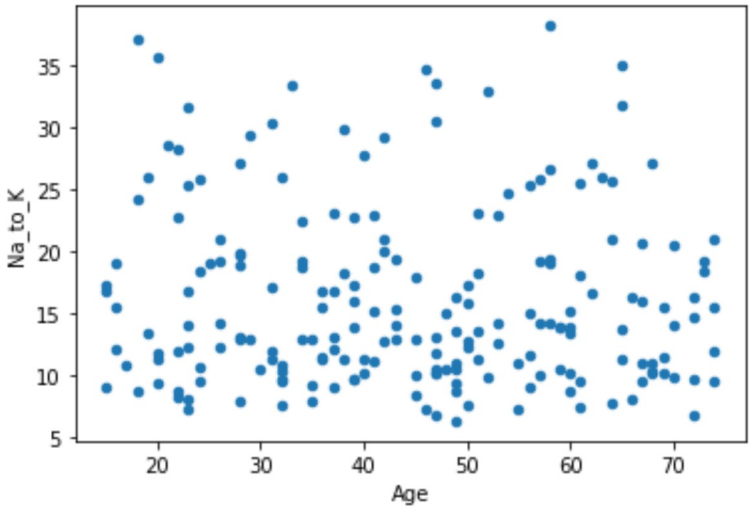
Out[15]: <AxesSubplot:>



SCATTER PLOT

In [20]:
```
data.plot.scatter(x = "Age", y = "Na_to_K" )
```

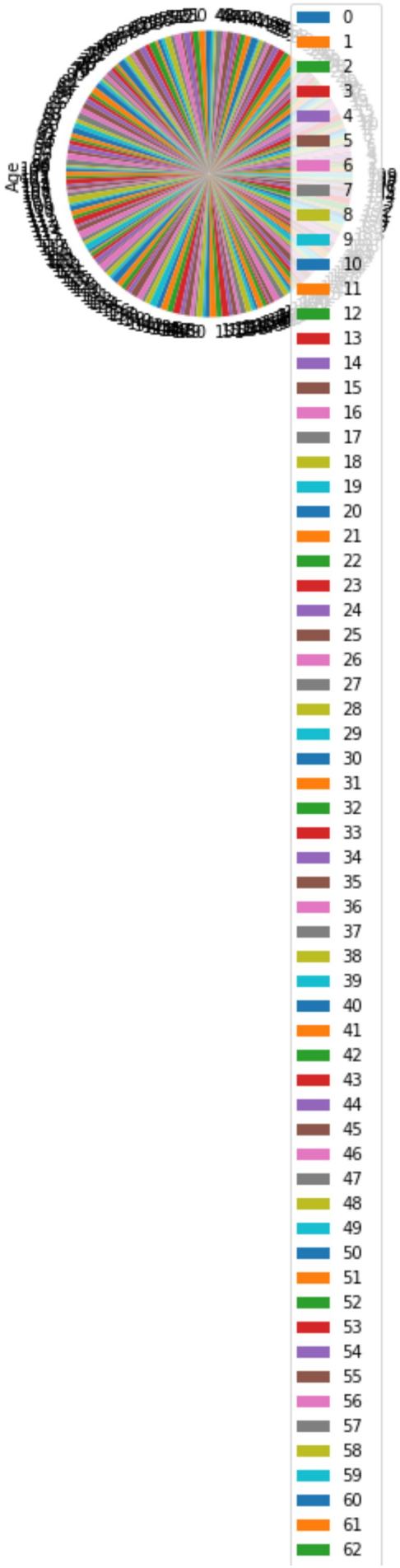
Out[20]: <AxesSubplot:xlabel='Age', ylabel='Na_to_K'>

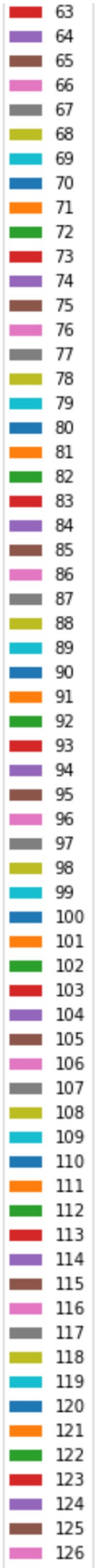


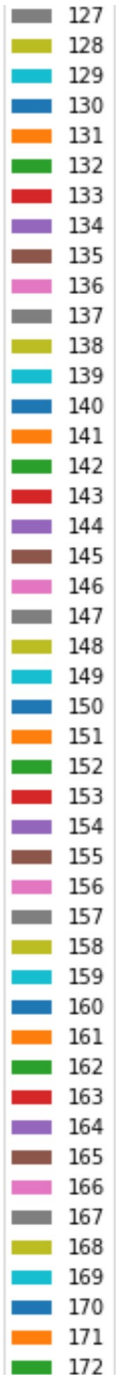
PIE CHART

In [21]:
```
data.plot.pie(y = "Age")
```

Out[21]: <AxesSubplot:ylabel='Age'>


Age
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62


63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126


127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172