In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from numpy import cov
```

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

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

Head

In [3]:
```
a.head(50)
```

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
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
10	47	F	LOW	HIGH	11.767	drugC

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
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
36	32	M	HIGH	NORMAL	9.445	drugA
37	39	M	LOW	NORMAL	13.938	drugX
38	39	F	NORMAL	NORMAL	9.709	drugX
39	15	M	NORMAL	HIGH	9.084	drugX
40	73	F	NORMAL	HIGH	19.221	drugY
41	58	F	HIGH	NORMAL	14.239	drugB
42	50	M	NORMAL	NORMAL	15.790	drugY
43	23	M	NORMAL	HIGH	12.260	drugX
44	50	F	NORMAL	NORMAL	12.295	drugX
45	66	F	NORMAL	NORMAL	8.107	drugX
46	37	F	HIGH	HIGH	13.091	drugA

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
47	68	M	LOW	HIGH	10.291	drugC
48	23	M	NORMAL	HIGH	31.686	drugY
49	28	F	LOW	HIGH	19.796	drugY

Tail

In [4]:
```
a.tail(50)
```

Out[4]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
150	49	M	HIGH	NORMAL	8.700	drugA
151	68	M	HIGH	HIGH	11.009	drugB
152	55	M	NORMAL	NORMAL	7.261	drugX
153	72	F	LOW	NORMAL	14.642	drugX
154	37	M	LOW	NORMAL	16.724	drugY
155	49	M	LOW	HIGH	10.537	drugC
156	31	M	HIGH	NORMAL	11.227	drugA
157	53	M	LOW	HIGH	22.963	drugY
158	59	F	LOW	HIGH	10.444	drugC
159	34	F	LOW	NORMAL	12.923	drugX
160	30	F	NORMAL	HIGH	10.443	drugX
161	57	F	HIGH	NORMAL	9.945	drugB
162	43	M	NORMAL	NORMAL	12.859	drugX
163	21	F	HIGH	NORMAL	28.632	drugY
164	16	M	HIGH	NORMAL	19.007	drugY
165	38	M	LOW	HIGH	18.295	drugY
166	58	F	LOW	HIGH	26.645	drugY
167	57	F	NORMAL	HIGH	14.216	drugX
168	51	F	LOW	NORMAL	23.003	drugY
169	20	F	HIGH	HIGH	11.262	drugA
170	28	F	NORMAL	HIGH	12.879	drugX
171	45	M	LOW	NORMAL	10.017	drugX
172	39	F	NORMAL	NORMAL	17.225	drugY
173	41	F	LOW	NORMAL	18.739	drugY
174	42	M	HIGH	NORMAL	12.766	drugA
175	73	F	HIGH	HIGH	18.348	drugY
176	48	M	HIGH	NORMAL	10.446	drugA

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
177	25	M	NORMAL	HIGH	19.011	drugY
178	39	M	NORMAL	HIGH	15.969	drugY
179	67	F	NORMAL	HIGH	15.891	drugY
180	22	F	HIGH	NORMAL	22.818	drugY
181	59	F	NORMAL	HIGH	13.884	drugX
182	20	F	LOW	NORMAL	11.686	drugX
183	36	F	HIGH	NORMAL	15.490	drugY
184	18	F	HIGH	HIGH	37.188	drugY
185	57	F	NORMAL	NORMAL	25.893	drugY
186	70	M	HIGH	HIGH	9.849	drugB
187	47	M	HIGH	HIGH	10.403	drugA
188	65	M	HIGH	NORMAL	34.997	drugY
189	64	M	HIGH	NORMAL	20.932	drugY
190	58	M	HIGH	HIGH	18.991	drugY
191	23	M	HIGH	HIGH	8.011	drugA
192	72	M	LOW	HIGH	16.310	drugY
193	72	M	LOW	HIGH	6.769	drugC
194	46	F	HIGH	HIGH	34.686	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

Describe

In [5]:
```
a.describe()
```

Out[5]:

	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

	Age	Na_to_K
max	74.000000	38.247000

Shape

In [6]:
```
a.shape
```

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

Size

In [7]:
```
a.size
```

Out[7]:
```
1200
```

isna

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

fillna

In [9]:
```
a.fillna(10)
```

Out[9]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
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

dropna

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

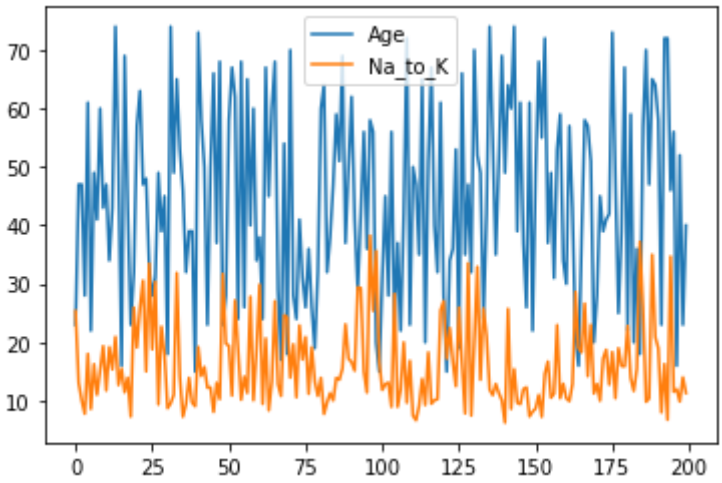
200 rows × 6 columns

Plots

Line plot

In [11]:
```
a.plot.line()
```

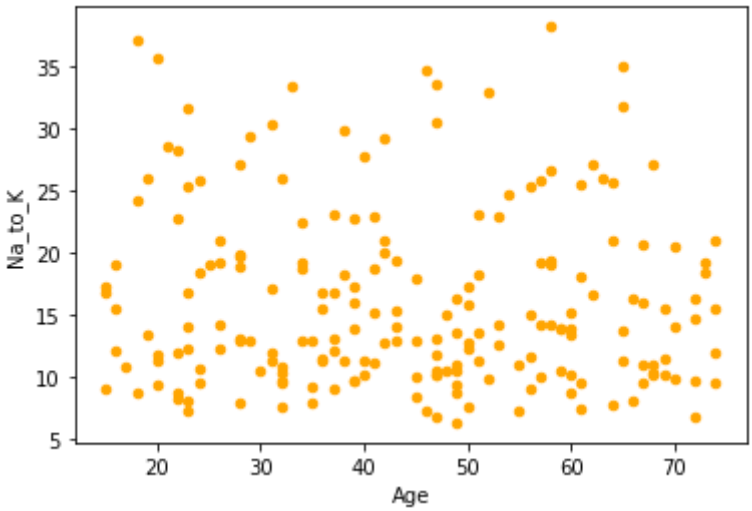
Out[11]: <AxesSubplot:>



Scatter Plot

In [14]:
```
a.plot.scatter(x="Age",y="Na_to_K",color="orange")
```

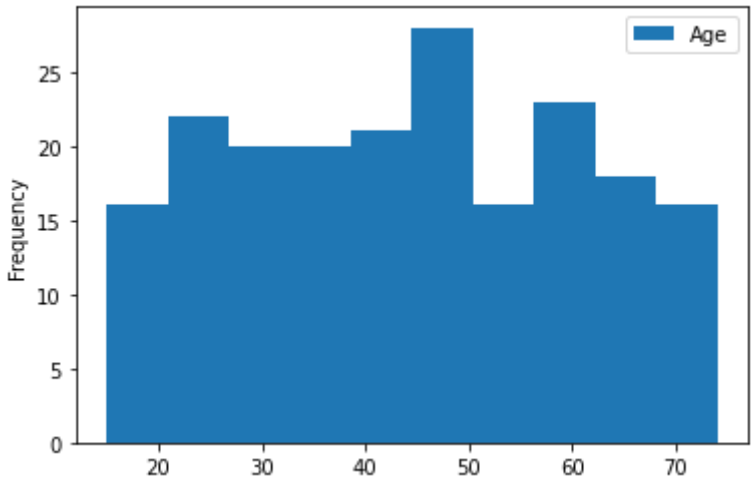
Out[14]: <AxesSubplot:xlabel='Age', ylabel='Na_to_K'>



Histagram

In [16]:
```
a.plot.hist(x="Na_to_K")
```

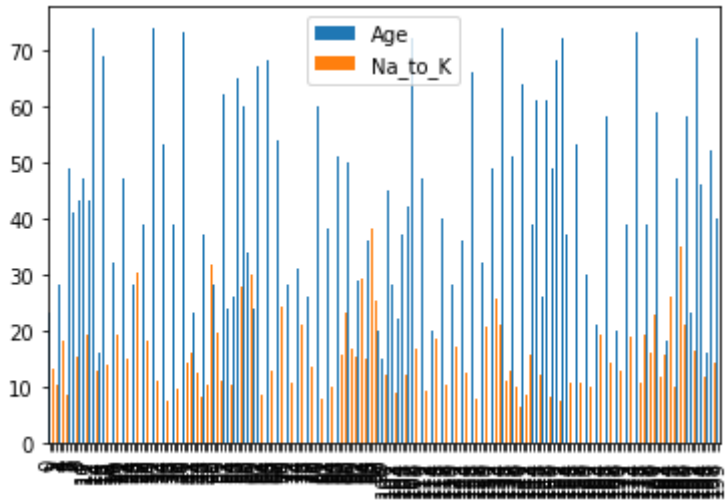
Out[16]: <AxesSubplot:ylabel='Frequency'>



Bar Chart

In [17]:
```
a.plot.bar()
```

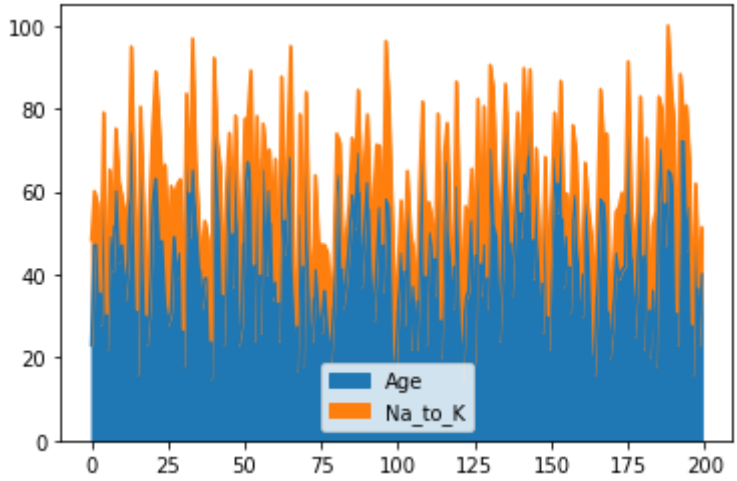
Out[17]: <AxesSubplot:>



Area Plot

In [18]:
```
a.plot.area()
```

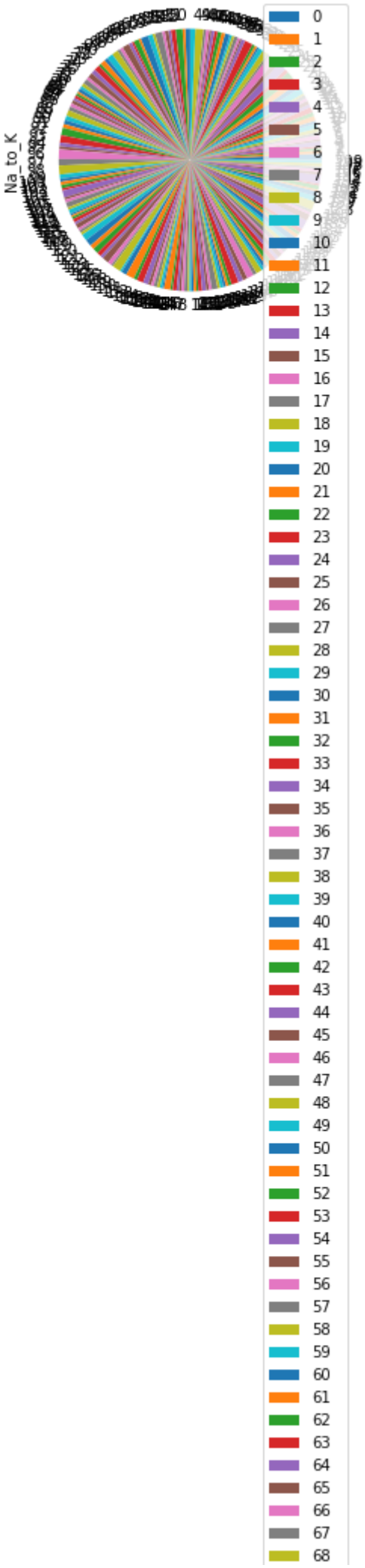
Out[18]: <AxesSubplot:>

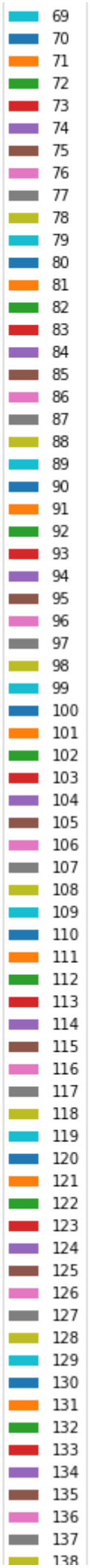


Pie Chart

In [19]:
```
a.plot.pie(y="Na_to_K")
```

Out[19]: <AxesSubplot:ylabel='Na_to_K'>


Na_to_K
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

Sum

In [20]:
```
print(a.sum())
```

```
Age                                                          8863
Sex            FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...
BP             HIGHLOWLOWNORMALLOWNORMALNORMALLOWNORMALLOWLOW...
Cholesterol    HIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
Na_to_K                                                3216.897
Drug           drugYdrugCdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
dtype: object
```

Mean

In [21]:
```
print(a.mean())
```

```
Age        44.315000
Na_to_K    16.084485
dtype: float64
```

Median

In [22]:
```
print(a.median())
```

```
Age        45.0000
Na_to_K    13.9365
dtype: float64
```

Mode

In [23]:
```
print(a.mode())
```

```
    Age  Sex    BP Cholesterol  Na_to_K   Drug
0  47.0    M  HIGH        HIGH   12.006  drugY
1   NaN  NaN   NaN         NaN   18.295    NaN
```

Standard Deviation

In [24]:
```
print(a.std())
```

```
Age        16.544315
Na_to_K     7.223956
dtype: float64
```

Variance

In [25]:
```
print(a.var())
```

```
Age        273.714347
Na_to_K     52.185533
dtype: float64
```

min and max

In [26]:
```
print(a.min())
```

```
Age               15
Sex                F
BP              HIGH
Cholesterol     HIGH
Na_to_K        6.269
Drug           drugA
dtype: object
```

In [27]:
```
print(a.max())
```

```
Age               74
Sex                M
BP            NORMAL
Cholesterol   NORMAL
Na_to_K       38.247
Drug           drugY
dtype: object
```

Correlation and covariance

In [28]:
```
print(a.cov())
```

```
                Age    Na_to_K
Age      273.714347  -7.543752
Na_to_K   -7.543752  52.185533
```

Pearson

In [29]:
```
from scipy.stats import pearsonr
print(pearsonr(a["Na_to_K"],a["Age"]))
```

```
(-0.06311949726772592, 0.3745756399034559)
```

Spearman

In [30]:
```
from scipy.stats import spearmanr
print(spearmanr(a["Na_to_K"],a["Age"]))
```

```
SpearmanrResult(correlation=-0.047273882688479915, pvalue=0.5062200581387418)
```

Count

In [31]:
```
a.count()
```

Out[31]:
```
Age            200
Sex            200
BP             200
Cholesterol    200
Na_to_K        200
```

```
Drug           200
dtype: int64
```

cumsum

In [32]:
```
a.cumsum()
```

Out[32]:

	Age	Sex
0	23	F
1	70	FM
2	117	FMM
3	145	FMMF
4	206	FMMFF
...	...	...
195	8732	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...
196	8748	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...
197	8800	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...
198	8823	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...
199	8863	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...

200 rows × 6 columns

In [ ]:
```
```