22/07/2023

```
In [1]: import numpy as np
        import pandas as pd
        import matplotlib.pyplot as pp
```

```
In [32]: x=pd.read_csv(r"C:\Users\user\Downloads\4_drug200 - 4_drug200.csv")
         x
```

Out[32]:

	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 [33]: x.dtypes
```

```
Out[33]: Age              int64
         Sex             object
         BP              object
         Cholesterol     object
         Na_to_K        float64
         Drug            object
         dtype: object
```

In [34]: `x.head()`

Out[34]:

	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 [35]: `x.tail()`

Out[35]:

	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 [36]: `x.columns`

Out[36]: `Index(['Age', 'Sex', 'BP', 'Cholesterol', 'Na_to_K', 'Drug'], dtype='object')`

In [37]: `x.index`

Out[37]: `RangeIndex(start=0, stop=200, step=1)`

In [38]: `x.describe()`

Out[38]:

	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 [59]: `x["Age"]`

Out[59]:
```
0        23
1        47
2        47
3        28
4        61
         ..
195      56
196      16
197      52
198      23
199      40
Name: Age, Length: 200, dtype: int64
```

In [40]: `x[0:2]`

Out[40]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC

In [41]: `x.iloc[0:2]`

Out[41]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC

In [42]: `x.loc[0:3]`

Out[42]:

	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

In [52]: `x.loc["Sex":"Drug"]`

Out[52]:

Age	Sex	BP	Cholesterol	Na_to_K	Drug

In [54]: `x[x["Age"]<=2]`

Out[54]:

Age	Sex	BP	Cholesterol	Na_to_K	Drug

In [45]: `x.fillna(value=5)`

Out[45]:

	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 [46]: `x.dropna()`

Out[46]:

	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

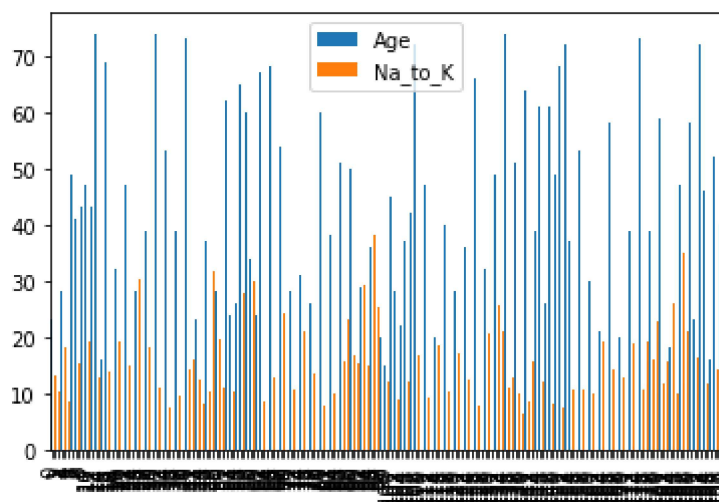In [47]: `x.plot.line()`

Out[47]: `<AxesSubplot:>`
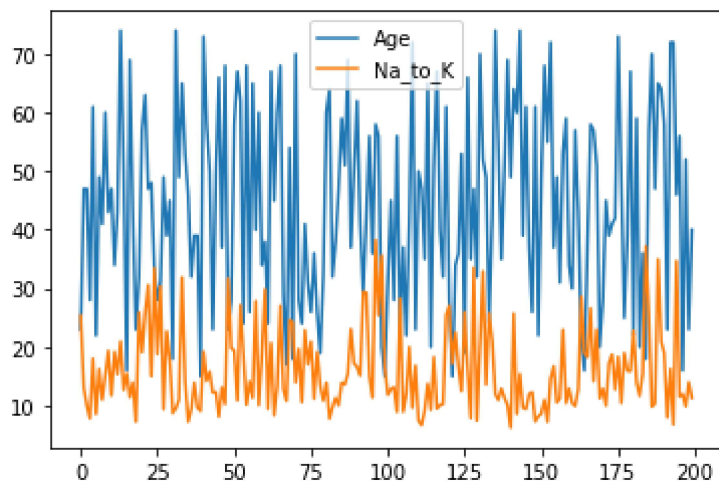
In [48]: `x.plot.bar()`
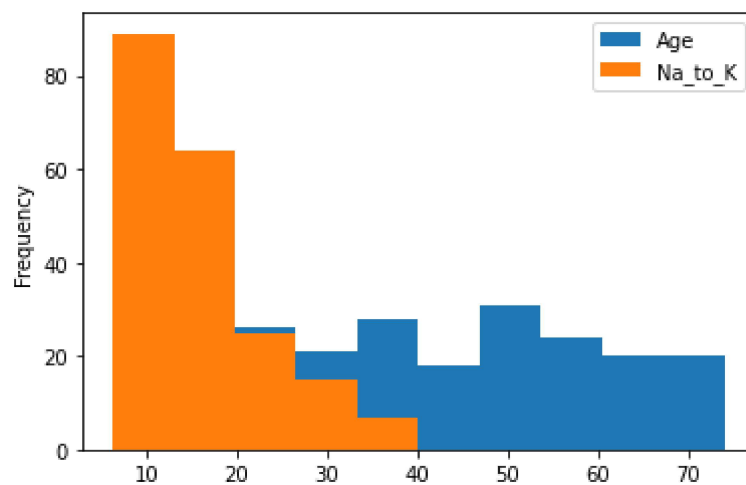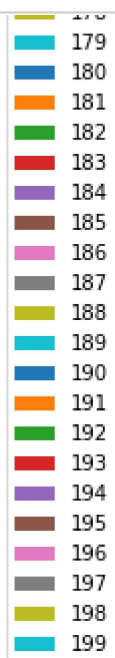
Out[48]: `<AxesSubplot:>`

In [49]:
```python
x.plot.hist()
```
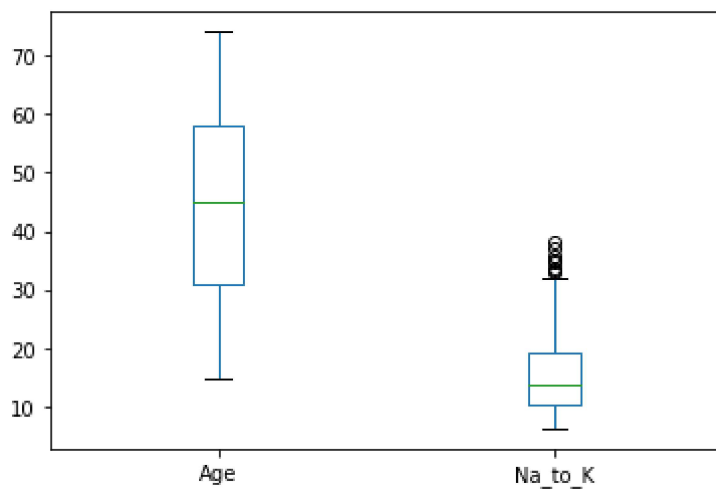
Out[49]: <AxesSubplot:ylabel='Frequency'>

In [56]:
```python
x.plot.pie(y='Age')
```

In [57]:
```python
x.plot.box()
```

Out[57]: <AxesSubplot:>

In [58]:
```python
x.plot.scatter(x='Sex',y='BP')
```

Out[58]: <AxesSubplot:xlabel='Sex', ylabel='BP'>