24/07/2023

In [40]:
```
import numpy as np
import pandas as pd
```

In [41]:
```
x=pd.read_csv(r"C:\Users\user\Downloads\8_BreastCancerPrediction (1).csv")
x
```

0	842302	M	17.99	10.38	122.80	1001.0
1	842517	M	20.57	17.77	132.90	1326.0
2	84300903	M	19.69	21.25	130.00	1203.0
3	84348301	M	11.42	20.38	77.58	386.1
4	84358402	M	20.29	14.34	135.10	1297.0
...	...	...	...	...	...	...
564	926424	M	21.56	22.39	142.00	1479.0
565	926682	M	20.13	28.25	131.20	1261.0
566	926954	M	16.60	28.08	108.30	858.1
567	927241	M	20.60	29.33	140.10	1265.0
568	92751	B	7.76	24.54	47.92	181.0

569 rows × 33 columns

In [42]:
```
x=x.head(200)
```

In [43]:
```
x.dtypes
```

Out[43]:
```
id                         int64
diagnosis                 object
radius_mean              float64
texture_mean             float64
perimeter_mean           float64
area_mean                float64
smoothness_mean          float64
compactness_mean         float64
concavity_mean           float64
concave points_mean      float64
symmetry_mean            float64
fractal_dimension_mean   float64
radius_se                float64
texture_se               float64
perimeter_se             float64
area_se                  float64
smoothness_se            float64
compactness_se           float64
concavity_se             float64
concave points_se        float64
symmetry_se              float64
fractal_dimension_se     float64
radius_worst             float64
texture_worst            float64
perimeter_worst          float64
area_worst               float64
smoothness_worst         float64
compactness_worst        float64
concavity_worst          float64
concave points_worst     float64
symmetry_worst           float64
fractal_dimension_worst  float64
Unnamed: 32              float64
dtype: object
```

In [44]:
```
x.head()
```

Out[44]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_m
0	842302	M	17.99	10.38	122.80	1001.0	0.11
1	842517	M	20.57	17.77	132.90	1326.0	0.08
2	84300903	M	19.69	21.25	130.00	1203.0	0.10
3	84348301	M	11.42	20.38	77.58	386.1	0.14
4	84358402	M	20.29	14.34	135.10	1297.0	0.10

5 rows × 33 columns

In [45]:
```
x.tail()
```

Out[45]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_m
195	875878	B	12.91	16.33	82.53	516.4	0.07
196	875938	M	13.77	22.29	90.63	588.9	0.12
197	877159	M	18.08	21.84	117.40	1024.0	0.07
198	877486	M	19.18	22.49	127.50	1148.0	0.08
199	877500	M	14.45	20.22	94.49	642.7	0.09

5 rows × 33 columns

In [46]:
```
x.columns
```

Out[46]:
```
Index(['id', 'diagnosis', 'radius_mean', 'texture_mean', 'perimeter_mean',
       'area_mean', 'smoothness_mean', 'compactness_mean', 'concavity_mean',
       'concave points_mean', 'symmetry_mean', 'fractal_dimension_mean',
       'radius_se', 'texture_se', 'perimeter_se', 'area_se', 'smoothness_se',
       'compactness_se', 'concavity_se', 'concave points_se', 'symmetry_se',
       'fractal_dimension_se', 'radius_worst', 'texture_worst',
       'perimeter_worst', 'area_worst', 'smoothness_worst',
       'compactness_worst', 'concavity_worst', 'concave points_worst',
       'symmetry_worst', 'fractal_dimension_worst', 'Unnamed: 32'],
      dtype='object')
```

In [47]:
```
x.index
```

Out[47]:
```
RangeIndex(start=0, stop=200, step=1)
```

In [48]:
```
x.describe()
```

Out[48]:

	id	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mea
count	2.000000e+02	200.000000	200.000000	200.000000	200.000000	200.00000
mean	1.992890e+07	14.292755	18.953050	93.390200	668.167500	0.09999
std	8.970791e+07	3.496191	3.778076	24.166686	338.981755	0.01403
min	8.670000e+03	6.981000	9.710000	43.790000	143.500000	0.06251
25%	8.574265e+05	11.885000	15.935000	76.922500	436.400000	0.09055
50%	8.657755e+05	13.610000	19.165000	87.980000	572.200000	0.09861
75%	8.510696e+06	16.137500	21.630000	107.275000	814.175000	0.10920
max	8.710015e+08	27.220000	27.540000	182.100000	2250.000000	0.14470

8 rows × 32 columns

In [50]:
```
x["texture_mean"]
```

Out[50]:
```
0      10.38
1      17.77
2      21.25
3      20.38
4      14.34
       ...
195    16.33
196    22.29
197    21.84
198    22.49
199    20.22
Name: texture_mean, Length: 200, dtype: float64
```

In [52]:
```
x[0:2]
```

Out[52]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mea
0	842302	M	17.99	10.38	122.8	1001.0	0.1184
1	842517	M	20.57	17.77	132.9	1326.0	0.0847

2 rows × 33 columns

In [53]:
```
x.iloc[0:9]
```

Out[53]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_m
0	842302	M	17.99	10.38	122.80	1001.0	0.11
1	842517	M	20.57	17.77	132.90	1326.0	0.08
2	84300903	M	19.69	21.25	130.00	1203.0	0.10
3	84348301	M	11.42	20.38	77.58	386.1	0.14
4	84358402	M	20.29	14.34	135.10	1297.0	0.10
5	843786	M	12.45	15.70	82.57	477.1	0.12
6	844359	M	18.25	19.98	119.60	1040.0	0.09
7	84458202	M	13.71	20.83	90.20	577.9	0.11
8	844981	M	13.00	21.82	87.50	519.8	0.12

9 rows × 33 columns

In [54]:
```
x.loc[0:80]
```

Out[54]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_r
0	842302	M	17.99	10.38	122.80	1001.0	0.1
1	842517	M	20.57	17.77	132.90	1326.0	0.0
2	84300903	M	19.69	21.25	130.00	1203.0	0.1
3	84348301	M	11.42	20.38	77.58	386.1	0.1
4	84358402	M	20.29	14.34	135.10	1297.0	0.1
...	...	...	...	...	...	...	
76	8610629	B	13.53	10.94	87.91	559.2	0.1
77	8610637	M	18.05	16.15	120.20	1006.0	0.1
78	8610862	M	20.18	23.97	143.70	1245.0	0.1
79	8610908	B	12.86	18.00	83.19	506.3	0.0
80	861103	B	11.45	20.97	73.81	401.5	0.1

81 rows × 33 columns

In [55]:
```
x.loc["perimeter_mean":"compactness_mean"]
```

Out[55]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	cor

0 rows × 33 columns

In [56]:
```
x[x["compactness_mean"]<=2]
```

Out[56]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_
0	842302	M	17.99	10.38	122.80	1001.0	0.
1	842517	M	20.57	17.77	132.90	1326.0	0.
2	84300903	M	19.69	21.25	130.00	1203.0	0.
3	84348301	M	11.42	20.38	77.58	386.1	0.
4	84358402	M	20.29	14.34	135.10	1297.0	0.
...	...	...	...	...	...	...	
195	875878	B	12.91	16.33	82.53	516.4	0.
196	875938	M	13.77	22.29	90.63	588.9	0.
197	877159	M	18.08	21.84	117.40	1024.0	0.
198	877486	M	19.18	22.49	127.50	1148.0	0.
199	877500	M	14.45	20.22	94.49	642.7	0.

200 rows × 33 columns

In [58]:
```
x.fillna(value=5)
```

Out[58]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_
0	842302	M	17.99	10.38	122.80	1001.0	0.
1	842517	M	20.57	17.77	132.90	1326.0	0.
2	84300903	M	19.69	21.25	130.00	1203.0	0.
3	84348301	M	11.42	20.38	77.58	386.1	0.
4	84358402	M	20.29	14.34	135.10	1297.0	0.
...	...	...	...	...	...	...	
195	875878	B	12.91	16.33	82.53	516.4	0.
196	875938	M	13.77	22.29	90.63	588.9	0.
197	877159	M	18.08	21.84	117.40	1024.0	0.
198	877486	M	19.18	22.49	127.50	1148.0	0.
199	877500	M	14.45	20.22	94.49	642.7	0.

200 rows × 33 columns

In [59]:
```
x.dropna()
```

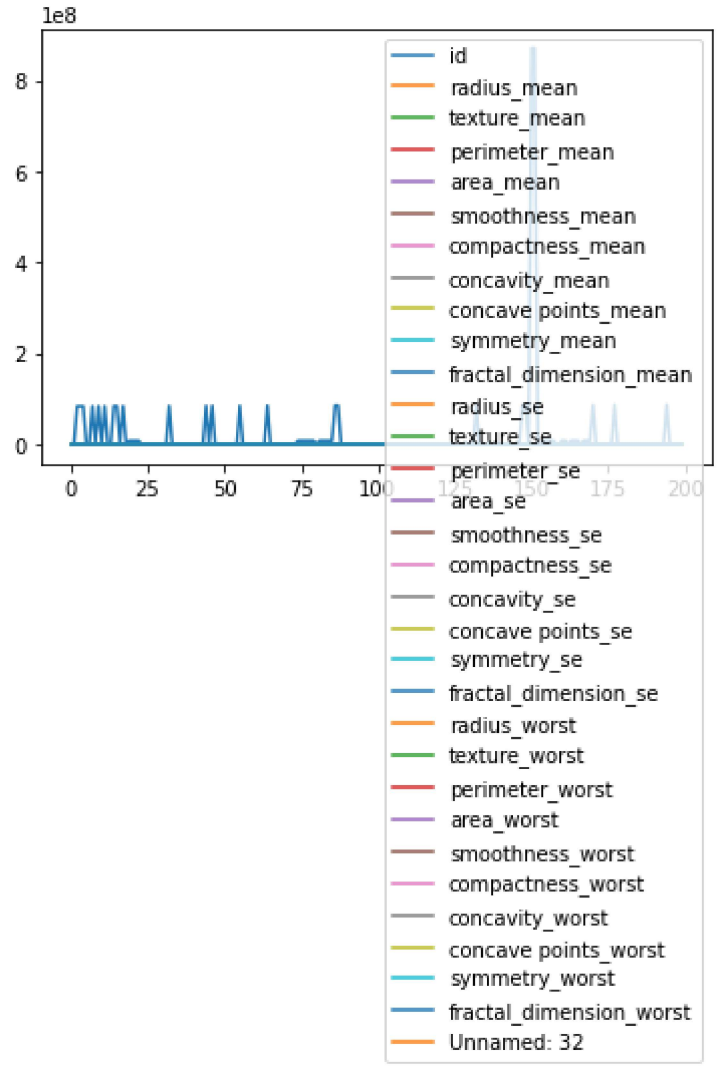
Out[59]:

id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	cor

0 rows × 33 columns

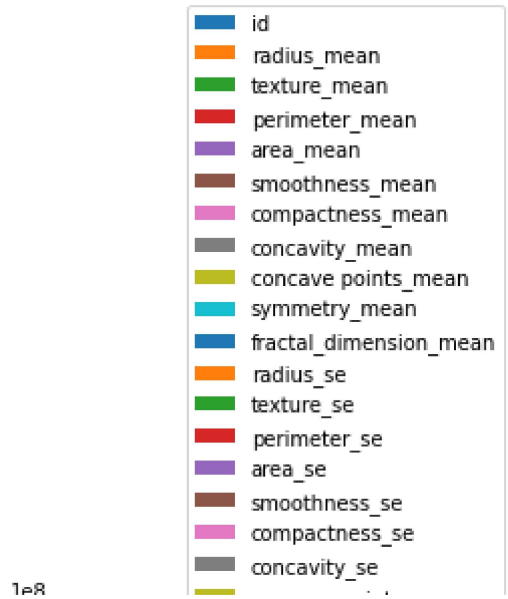
In [60]:
```
x.plot.line()
```

Out[60]: <AxesSubplot:>



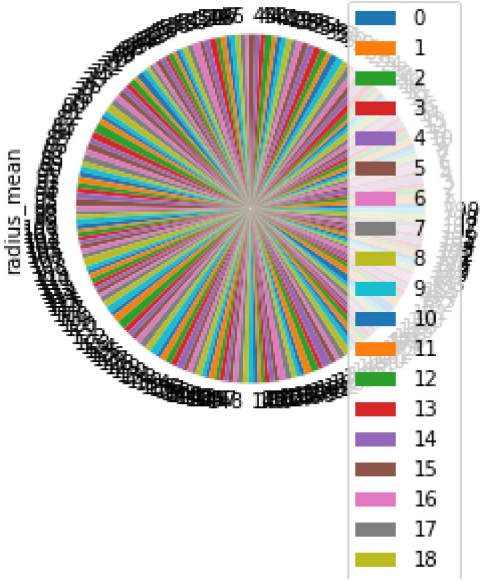
In [61]:
```
x.plot.bar()
```

Out[61]: <AxesSubplot:>



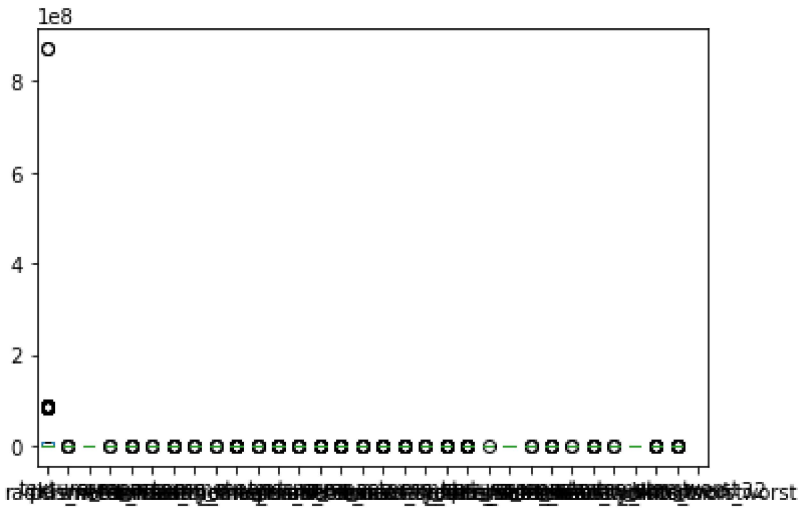
In [62]:
```
x.plot.pie(y='radius_mean')
```

Out[62]: <AxesSubplot:ylabel='radius_mean'>



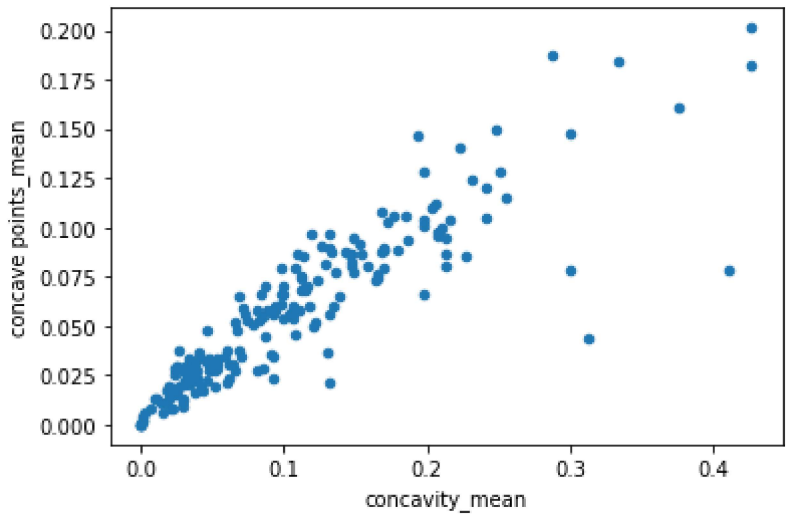
In [65]:
```
x.plot.box()
```

Out[65]: <AxesSubplot:>



In [64]:
```
x.plot.scatter(x='concavity_mean',y='concave points_mean')
```

Out[64]: <AxesSubplot:xlabel='concavity_mean', ylabel='concave points_mean'>



In [ ]: