

importing libraries

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

```
In [2]: df=pd.read_csv(r"C:\Users\user\Desktop\8_BreastCancerPrediction.csv")
```

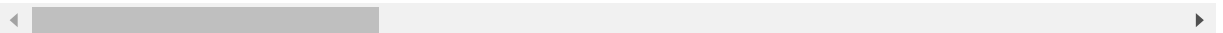
To show the first 10 values

```
In [3]: df.head(10)
```

```
Out[3]:
```

	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	84501001	M	12.46	24.04	83.97	475.9	0.11

10 rows × 33 columns



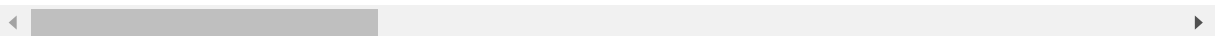
To show the last 11 values

In [4]: `df.tail(11)`

Out[4]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_m
558	925277	B	14.59	22.68	96.39	657.1	0.08
559	925291	B	11.51	23.93	74.52	403.5	0.09
560	925292	B	14.05	27.15	91.38	600.4	0.09
561	925311	B	11.20	29.37	70.67	386.0	0.07
562	925622	M	15.22	30.62	103.40	716.9	0.10
563	926125	M	20.92	25.09	143.00	1347.0	0.10
564	926424	M	21.56	22.39	142.00	1479.0	0.11
565	926682	M	20.13	28.25	131.20	1261.0	0.09
566	926954	M	16.60	28.08	108.30	858.1	0.08
567	927241	M	20.60	29.33	140.10	1265.0	0.11
568	92751	B	7.76	24.54	47.92	181.0	0.05

11 rows × 33 columns



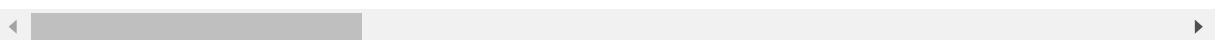
To show statistical data

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

Out[5]:

	id	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mea
count	5.690000e+02	569.000000	569.000000	569.000000	569.000000	569.000000
mean	3.037183e+07	14.127292	19.289649	91.969033	654.889104	0.09636
std	1.250206e+08	3.524049	4.301036	24.298981	351.914129	0.01406
min	8.670000e+03	6.981000	9.710000	43.790000	143.500000	0.05263
25%	8.692180e+05	11.700000	16.170000	75.170000	420.300000	0.08637
50%	9.060240e+05	13.370000	18.840000	86.240000	551.100000	0.09587
75%	8.813129e+06	15.780000	21.800000	104.100000	782.700000	0.10530
max	9.113205e+08	28.110000	39.280000	188.500000	2501.000000	0.16340

8 rows × 32 columns



shape()

```
In [6]: np.shape(df)
```

Out[6]: (569, 33)

size()

```
In [7]: np.size(df)
```

Out[7]: 18777

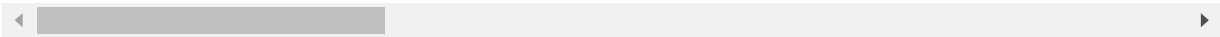
isna()

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

Out[8]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean
0	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...
564	False	False	False	False	False	False	False
565	False	False	False	False	False	False	False
566	False	False	False	False	False	False	False
567	False	False	False	False	False	False	False
568	False	False	False	False	False	False	False

569 rows × 33 columns



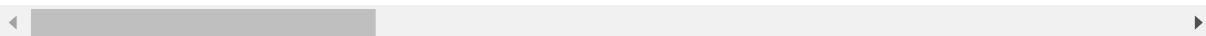
fillna()

```
In [9]: df.fillna(value='8')
```

```
Out[9]:
```

	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.
...
564	926424	M	21.56	22.39	142.00	1479.0	0.
565	926682	M	20.13	28.25	131.20	1261.0	0.
566	926954	M	16.60	28.08	108.30	858.1	0.
567	927241	M	20.60	29.33	140.10	1265.0	0.
568	92751	B	7.76	24.54	47.92	181.0	0.

569 rows × 33 columns

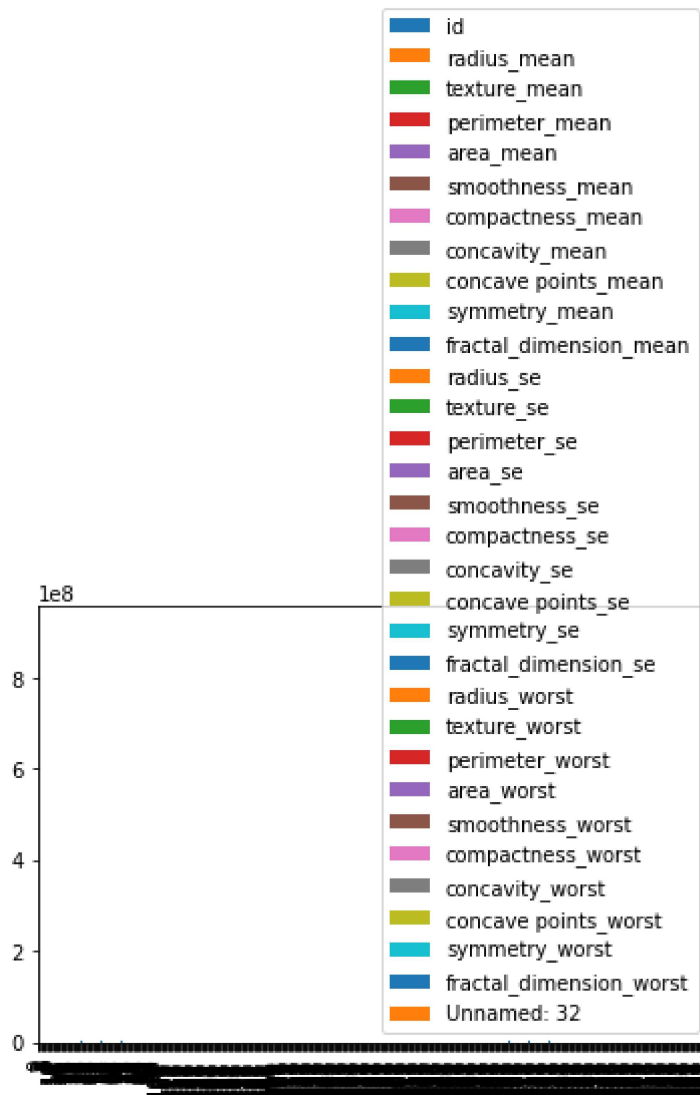


Visualization

```
In [10]: import matplotlib.pyplot as pp
```

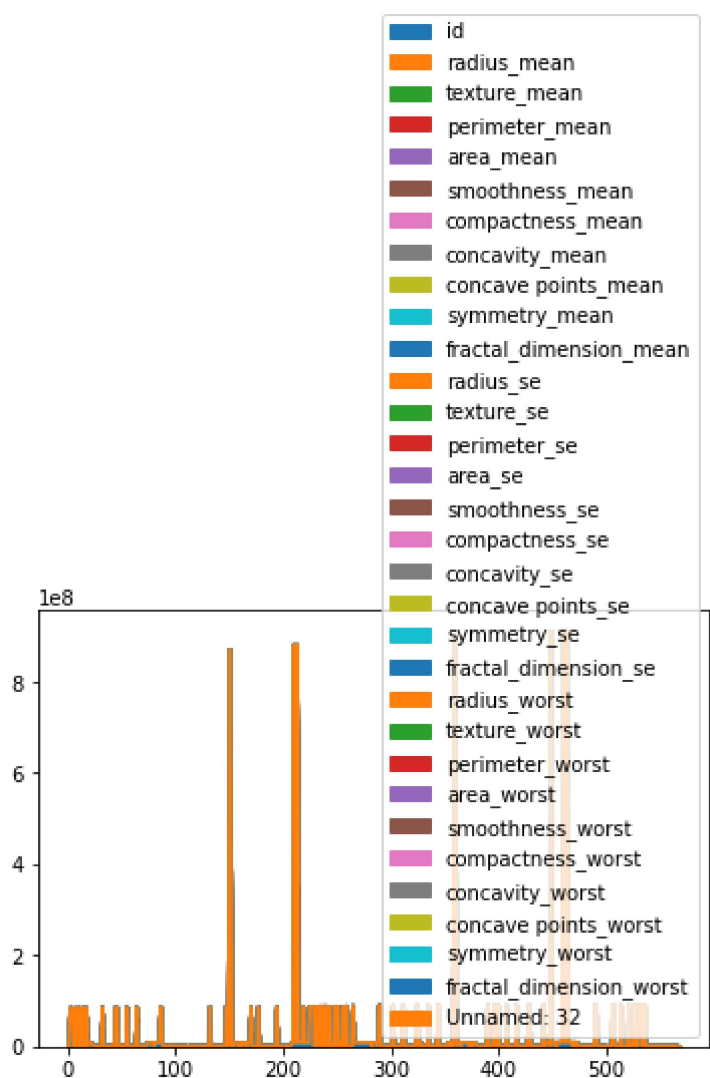
```
In [11]: df.plot.bar()
```

```
Out[11]: <AxesSubplot:>
```



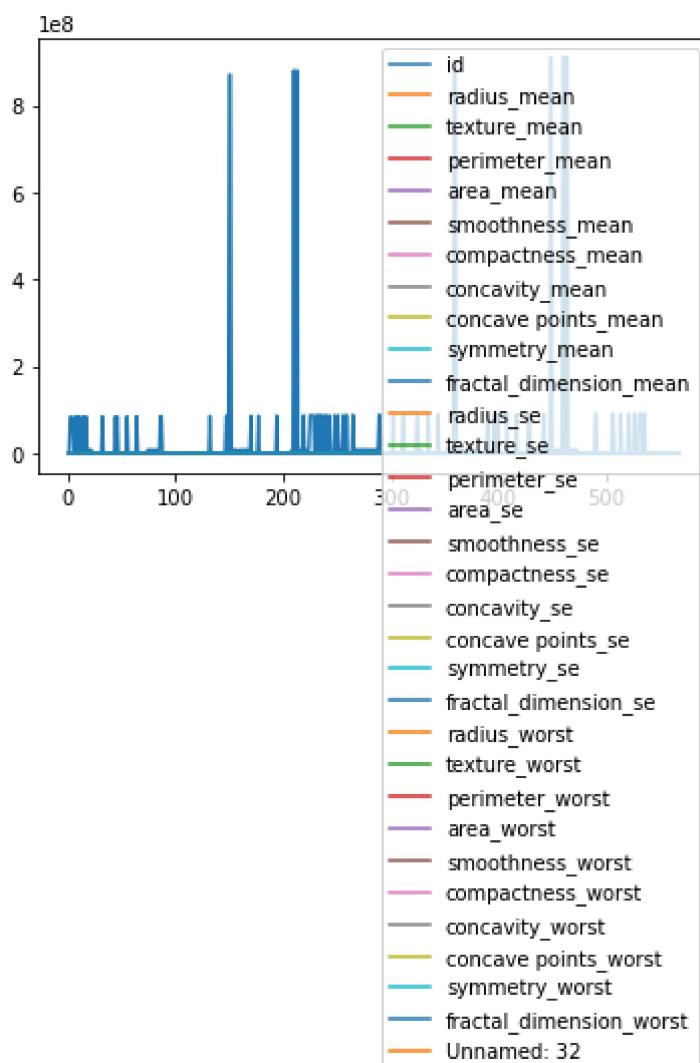
```
In [12]: df.plot.area()
```

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



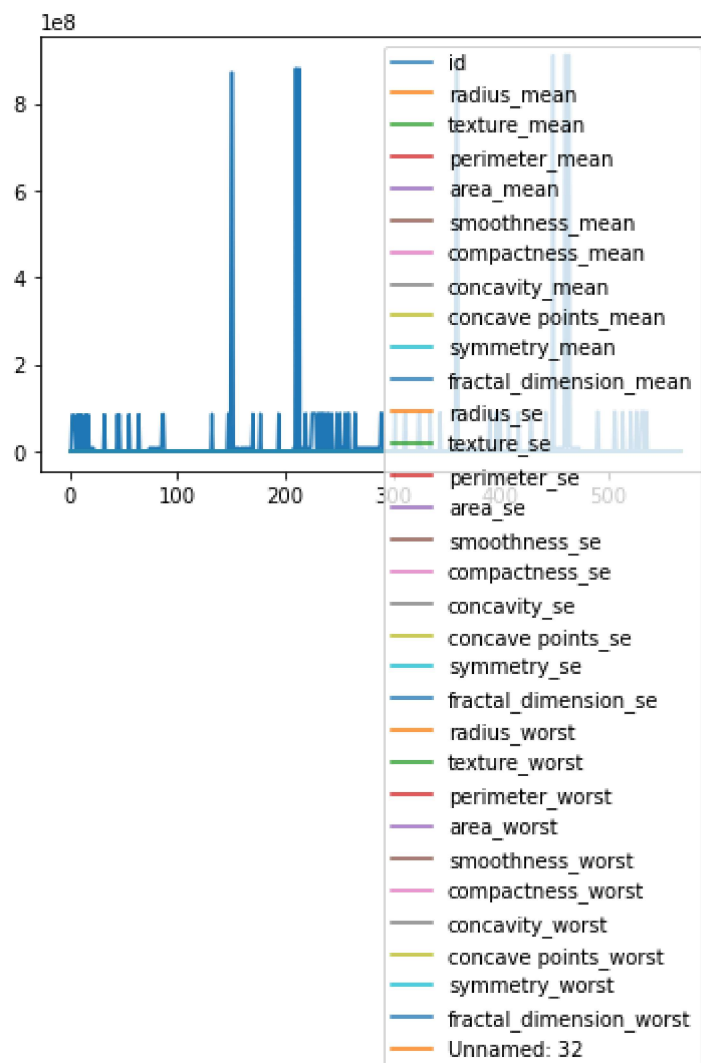
```
In [13]: df.plot.line()
```

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



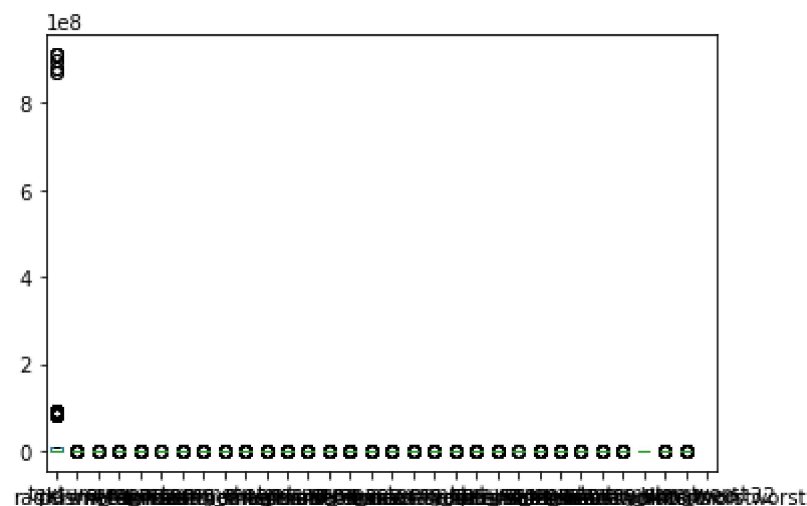
In [14]: `df.plot.line()`

Out[14]: `<AxesSubplot:>`



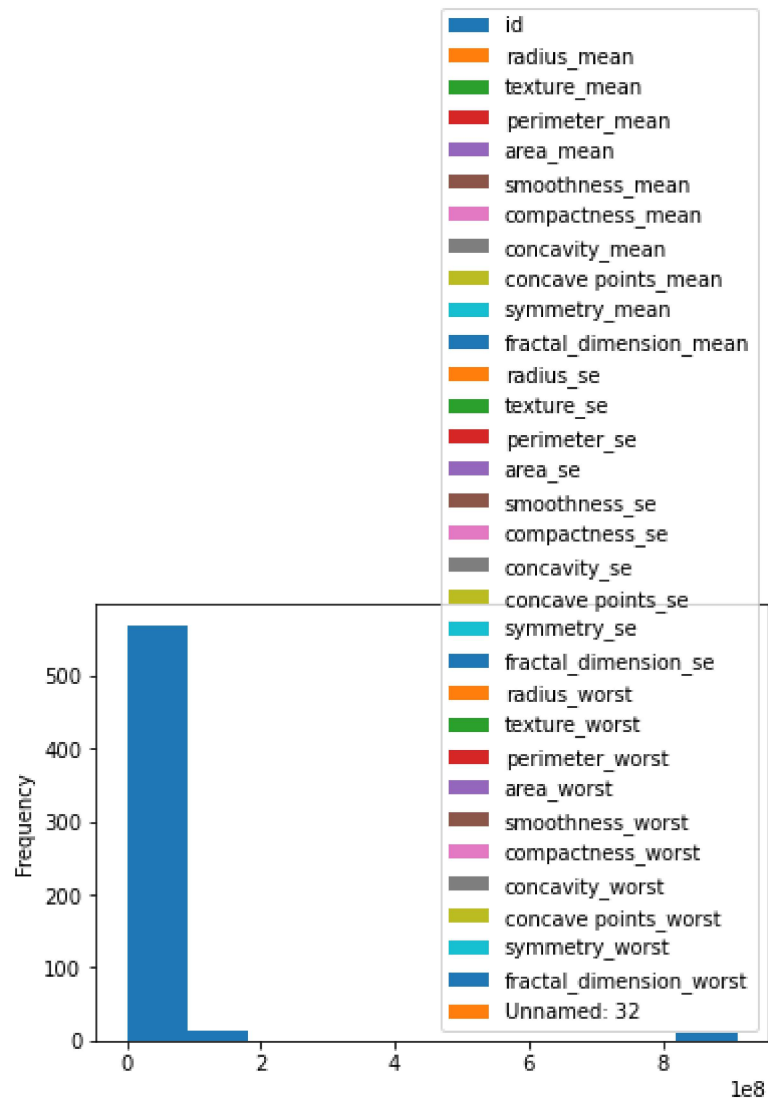
In [15]: `df.plot.box()`

Out[15]: `<AxesSubplot:>`



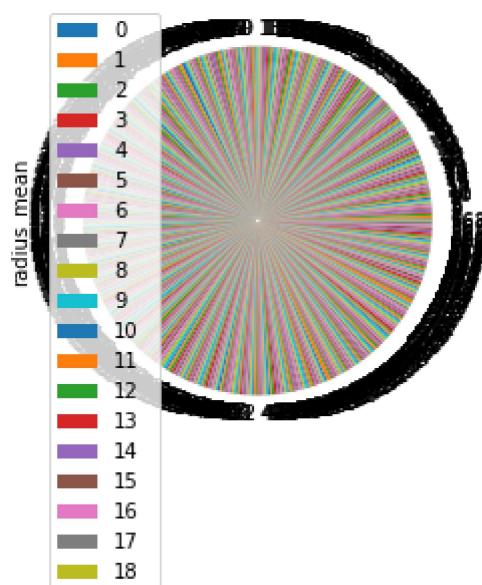

```
In [16]: df.plot.hist()
```

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



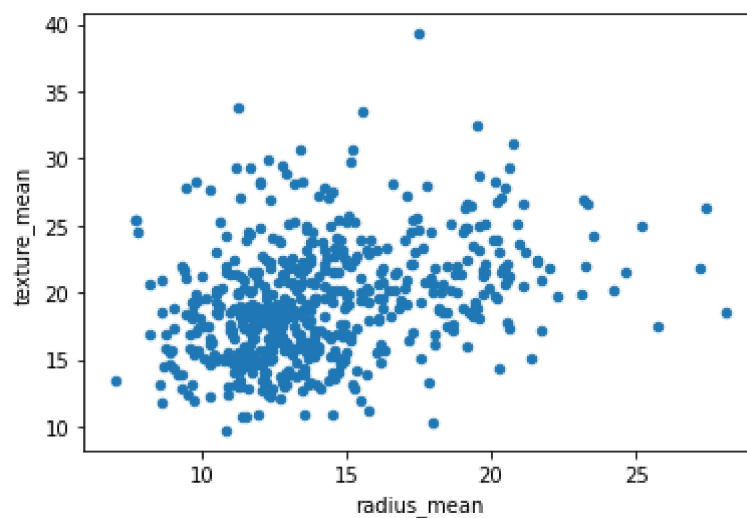
```
In [20]: df.plot.pie(y='radius_mean')
```

```
Out[20]: <AxesSubplot:ylabel='radius_mean'>
```



```
In [22]: df.plot.scatter(x='radius_mean',y='texture_mean')
```

```
Out[22]: <AxesSubplot:xlabel='radius_mean', ylabel='texture_mean'>
```



```
In [ ]:
```