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

IMPORT AND PRINT DATA SET

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

Out[2]:

	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

SHAPE

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

Out[3]: (569, 32)

SIZE

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

Out[4]: 18208

PRINT FIRST 10 VALUES

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

Out[5]:

	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 × 32 columns

PRINT LAST 7 VALUES

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

Out[6]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_m
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

5 rows × 32 columns

DESCRIPTION OF TABLE

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

Out[7]:

	id	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mea
count	5.690000e+02	569.000000	569.000000	569.000000	569.000000	569.0000(
mean	3.037183e+07	14.127292	19.289649	91.969033	654.889104	0.0963(
std	1.250206e+08	3.524049	4.301036	24.298981	351.914129	0.0140(
min	8.670000e+03	6.981000	9.710000	43.790000	143.500000	0.0526(
25%	8.692180e+05	11.700000	16.170000	75.170000	420.300000	0.0863(
50%	9.060240e+05	13.370000	18.840000	86.240000	551.100000	0.0958(
75%	8.813129e+06	15.780000	21.800000	104.100000	782.700000	0.1053(
max	9.113205e+08	28.110000	39.280000	188.500000	2501.000000	0.1634(

8 rows × 31 columns

FIND NULL VALUES

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

Out[8]:

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

569 rows × 32 columns

FILL NULL VALUES

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

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 × 32 columns

In [10]:
```
data.columns
```

Out[10]:
```
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'],
      dtype='object')
```

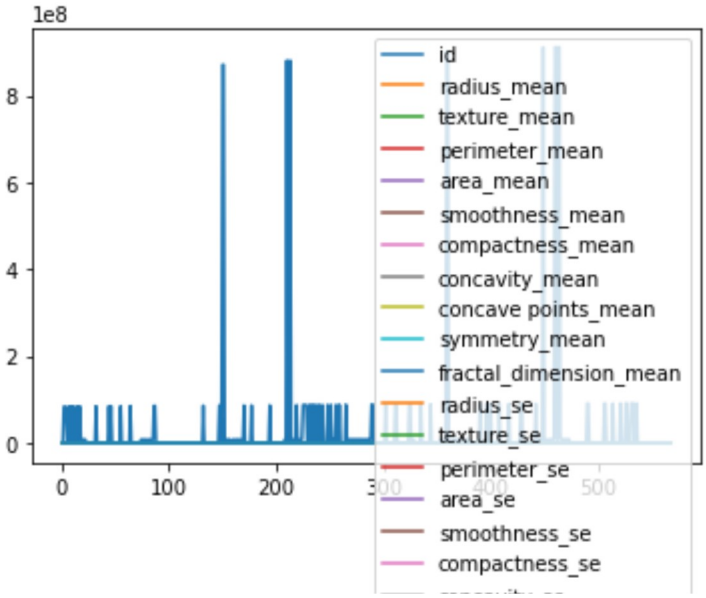
In [11]:
```
data.index
```

Out[11]:
```
RangeIndex(start=0, stop=569, step=1)
```

LINE PLOT

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

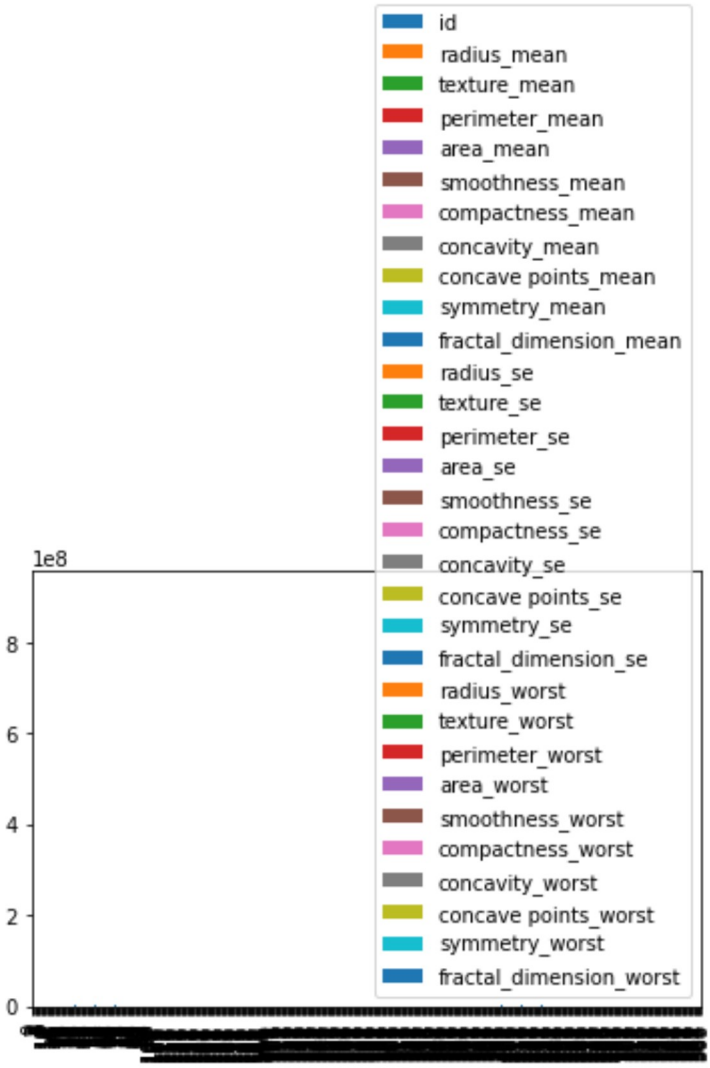
Out[12]: <AxesSubplot:>



BAR CHART

In [22]:
```
data.plot.bar()
```

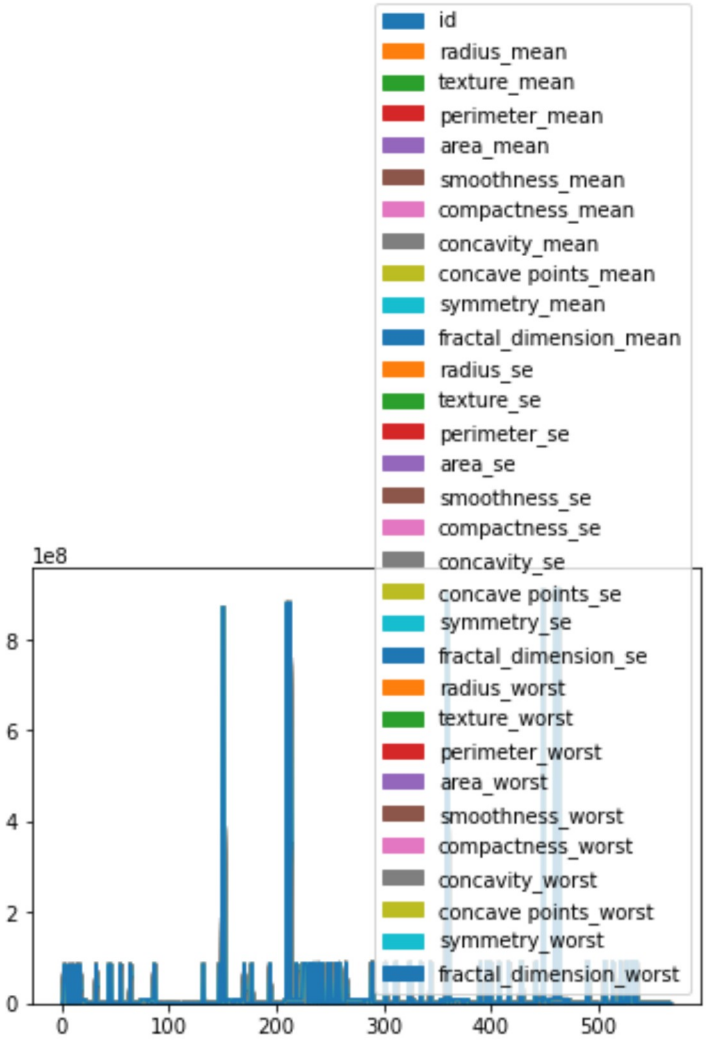
Out[22]: <AxesSubplot:>



AREA CHART

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

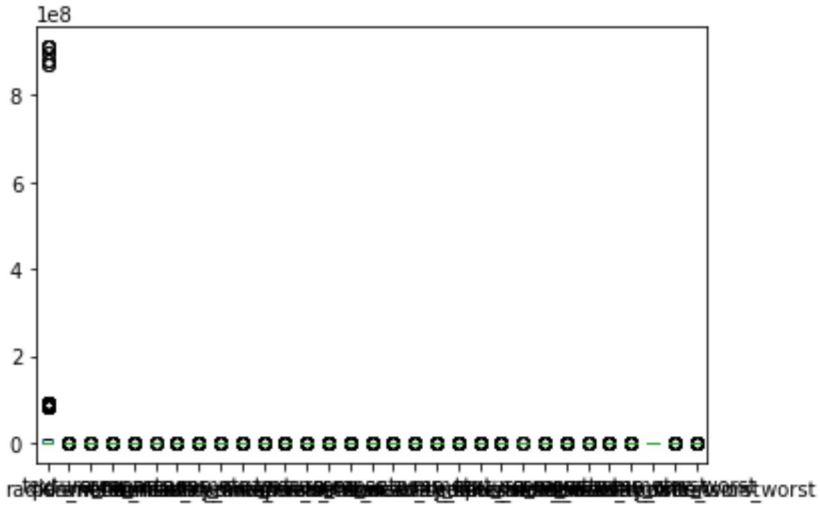
Out[14]: <AxesSubplot:>



BOX PLOT

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

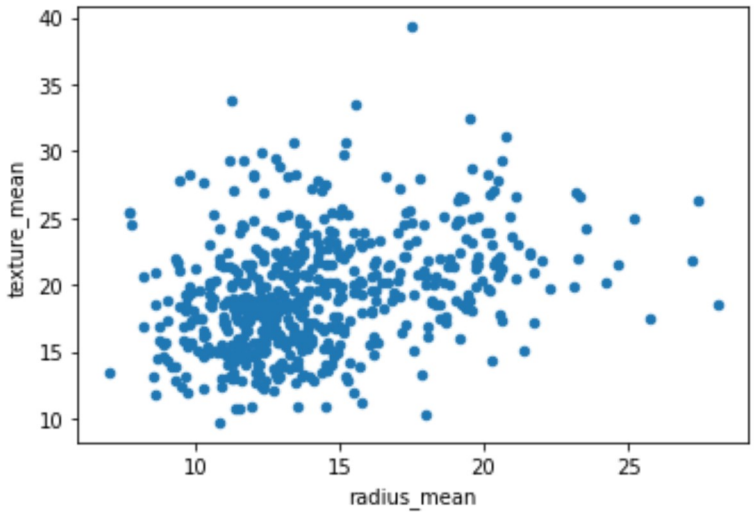
Out[15]: <AxesSubplot:>



SCATTER PLOT

In [19]:
```
data.plot.scatter(x = "radius_mean", y = "texture_mean" )
```

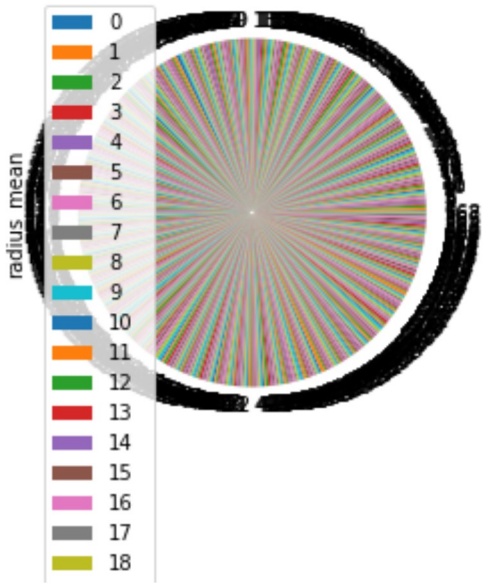
Out[19]: <AxesSubplot:xlabel='radius_mean', ylabel='texture_mean'>



PIE CHART

In [20]:
```
data.plot.pie(y = "radius_mean")
```

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



In [ ]: