

```
In [9]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.datasets import load_breast_cancer
```

```
In [39]: breast = load_breast_cancer()
```

```
In [40]: breast_data = breast.data
breast_data.shape
```

Out[40]: (569, 30)

```
In [41]: breast_input = pd.DataFrame(breast_data)
breast_input.head()
```

```
Out[41]:
```

	0	1	2	3	4	5	6	7	8	9	...	20	21	
0	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.3001	0.14710	0.2419	0.07871	...	25.38	17.33	184
1	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.0869	0.07017	0.1812	0.05667	...	24.99	23.41	158
2	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.1974	0.12790	0.2069	0.05999	...	23.57	25.53	152
3	11.42	20.38	77.58	386.1	0.14250	0.28390	0.2414	0.10520	0.2597	0.09744	...	14.91	26.50	98
4	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.1980	0.10430	0.1809	0.05883	...	22.54	16.67	152

5 rows × 30 columns



```
In [42]: breast_labels = breast.target
```

```
In [43]: breast_labels.shape
```

Out[43]: (569,)

```
In [44]: labels = np.reshape(breast_labels,(569,1))
```

```
In [45]: final_breast_data = np.concatenate([breast_data,labels],axis=1)
```

```
In [46]: final_breast_data.shape
```

Out[46]: (569, 31)

```
In [47]: breast_dataset = pd.DataFrame(final_breast_data)
```

```
In [49]: features = breast.feature_names
features
```

```
Out[49]: array(['mean radius', 'mean texture', 'mean perimeter', 'mean area',
               'mean smoothness', 'mean compactness', 'mean concavity',
               'mean concave points', 'mean symmetry', 'mean fractal dimension',
               'radius error', 'texture error', 'perimeter error', 'area error',
               'smoothness error', 'compactness error', 'concavity error',
               'concave points error', 'symmetry error',
               'fractal dimension error', 'worst radius', 'worst texture',
               'worst perimeter', 'worst area', 'worst smoothness',
               'worst compactness', 'worst concavity', 'worst concave points',
               'worst symmetry', 'worst fractal dimension'], dtype='<U23')
```

```
In [50]: features_labels = np.append(features, 'label')
```

```
In [51]: breast_dataset.columns = features_labels
```

```
In [52]: breast_dataset.head()
```

Out[52]:

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	mean fractal dimension
0	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.3001	0.14710	0.2419	0.0766
1	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.0869	0.07017	0.1812	0.0566
2	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.1974	0.12790	0.2069	0.0596
3	11.42	20.38	77.58	386.1	0.14250	0.28390	0.2414	0.10520	0.2597	0.0596
4	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.1980	0.10430	0.1809	0.0596

5 rows × 31 columns



```
In [53]: breast_dataset['label'].replace(0, 'Benign', inplace=True)
breast_dataset['label'].replace(1, 'Malignant', inplace=True)
```

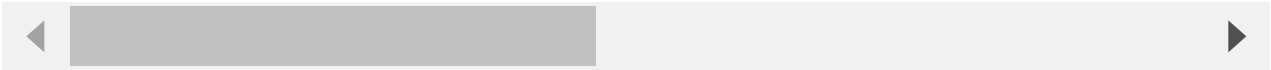
```
In [54]: breast_dataset.tail()
```

Out[54]:

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	mean fractal dimension
564	21.56	22.39	142.00	1479.0	0.11100	0.11590	0.24390	0.13890	0.1726	C
565	20.13	28.25	131.20	1261.0	0.09780	0.10340	0.14400	0.09791	0.1752	C
566	16.60	28.08	108.30	858.1	0.08455	0.10230	0.09251	0.05302	0.1590	C

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	dim
567	20.60	29.33	140.10	1265.0	0.11780	0.27700	0.35140	0.15200	0.2397	C
568	7.76	24.54	47.92	181.0	0.05263	0.04362	0.00000	0.00000	0.1587	C

5 rows × 31 columns



In []: