

```
import pandas as pd
a=pd.read_csv('/content/breast_cancer_survival.csv')
print(a)
print(a.head())
print(a.columns)
target_variable = 'Patient_Status'
features = a.columns[a.columns != target_variable]
print('Target Variable:', target_variable)
print('Features:', features)
```

		Surgery_type	Date_of_Surgery	Date_of_Last_Visit	\
0		Other	20-May-18	26-Aug-18	
1		Other	26-Apr-18	25-Jan-19	
2		Lumpectomy	24-Aug-18	08-Apr-20	
3		Other	16-Nov-18	28-Jul-20	
4		Lumpectomy	12-Dec-18	05-Jan-19	
..		
329		Lumpectomy	15-Jan-19	27-Mar-20	
330	Modified Radical	Mastectomy	25-Jul-18	23-Apr-19	
331	Simple	Mastectomy	26-Mar-19	11-Oct-19	
332		Lumpectomy	26-Nov-18	05-Dec-18	
333	Modified Radical	Mastectomy	04-Feb-19	10-Aug-19	

	Patient_Status
0	Alive
1	Dead
2	Alive
3	Alive
4	Alive
..	...
329	Alive
330	Alive
331	Dead
332	Alive
333	Dead

[334 rows x 15 columns]

	Age	Gender	Protein1	Protein2	Protein3	Protein4	Tumour_Stage	\
0	42	FEMALE	0.95256	2.15000	0.007972	-0.048340	II	
1	54	FEMALE	0.00000	1.38020	-0.498030	-0.507320	II	
2	63	FEMALE	-0.52303	1.76400	-0.370190	0.010815	II	
3	78	FEMALE	-0.87618	0.12943	-0.370380	0.132190	I	
4	42	FEMALE	0.22611	1.74910	-0.543970	-0.390210	II	

	Histology	ER status	PR status	HER2 status	Surgery_type	\
0	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Other	
1	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Other	
2	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Lumpectomy	
3	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Other	
4	Infiltrating Ductal Carcinoma	Positive	Positive	Positive	Lumpectomy	

	Date_of_Surgery	Date_of_Last_Visit	Patient_Status
0	20-May-18	26-Aug-18	Alive
1	26-Apr-18	25-Jan-19	Dead
2	24-Aug-18	08-Apr-20	Alive
3	16-Nov-18	28-Jul-20	Alive
4	12-Dec-18	05-Jan-19	Alive

```
Index(['Age', 'Gender', 'Protein1', 'Protein2', 'Protein3', 'Protein4',
      'Tumour_Stage', 'Histology', 'ER status', 'PR status', 'HER2 status',
      'Surgery_type', 'Date_of_Surgery', 'Date_of_Last_Visit',
      'Patient_Status'],
      dtype='object')
Target Variable: Patient_Status
Features: Index(['Age', 'Gender', 'Protein1', 'Protein2', 'Protein3', 'Protein4',
      'Tumour_Stage', 'Histology', 'ER status', 'PR status', 'HER2 status',
      'Surgery_type', 'Date_of_Surgery', 'Date_of_Last_Visit'],
      dtype='object')
```

```
b=a.fillna(0)
b
```

	Age	Gender	Protein1	Protein2	Protein3	Protein4	Tumour_Stage	Histology
0	42	FEMALE	0.952560	2.15000	0.007972	-0.048340	II	Infiltrating Ductal Carcinoma
1	54	FEMALE	0.000000	1.38020	-0.498030	-0.507320	II	Infiltrating Ductal Carcinoma
2	63	FEMALE	-0.523030	1.76400	-0.370190	0.010815	II	Infiltrating Ductal Carcinoma
3	78	FEMALE	-0.876180	0.12943	-0.370380	0.132190	I	Infiltrating Ductal Carcinoma
4	42	FEMALE	0.226110	1.74910	-0.543970	-0.390210	II	Infiltrating Ductal Carcinoma
...
329	59	FEMALE	0.024598	1.40050	0.024751	0.280320	II	Infiltrating Ductal Carcinoma
330	41	FEMALE	0.100120	-0.46547	0.472370	-0.523870	I	Infiltrating Ductal Carcinoma
331	54	FEMALE	0.753820	1.64250	-0.332850	0.857860	II	Infiltrating Ductal Carcinoma
332	74	FEMALE	0.972510	1.42680	-0.366570	-0.107820	II	Infiltrating Lobular Carcinoma
333	66	FEMALE	0.286380	1.39980	0.318830	0.836050	II	Infiltrating Ductal Carcinoma

334 rows × 15 columns

Next steps:

 View recommended plots

```
y=b['Patient_Status']
y

0      Alive
1      Dead
2      Alive
3      Alive
4      Alive
...
329    Alive
330    Alive
331    Dead
332    Alive
333    Dead
Name: Patient_Status, Length: 334, dtype: object
```

```
x=b.drop('Patient_Status',axis=1)
x
```

	Age	Gender	Protein1	Protein2	Protein3	Protein4	Tumour_Stage	Histology
0	42	FEMALE	0.952560	2.15000	0.007972	-0.048340	II	Infiltrating Ductal Carcinoma
1	54	FEMALE	0.000000	1.38020	-0.498030	-0.507320	II	Infiltrating Ductal Carcinoma
2	63	FEMALE	-0.523030	1.76400	-0.370190	0.010815	II	Infiltrating Ductal Carcinoma
3	78	FEMALE	-0.876180	0.12943	-0.370380	0.132190	I	Infiltrating Ductal Carcinoma
4	42	FEMALE	0.226110	1.74910	-0.543970	-0.390210	II	Infiltrating Ductal Carcinoma
...
329	59	FEMALE	0.024598	1.40050	0.024751	0.280320	II	Infiltrating Ductal Carcinoma
330	41	FEMALE	0.100120	-0.46547	0.472370	-0.523870	I	Infiltrating Ductal Carcinoma
331	54	FEMALE	0.753820	1.64250	-0.332850	0.857860	II	Infiltrating Ductal Carcinoma
332	74	FEMALE	0.972510	1.42680	-0.366570	-0.107820	II	Infiltrating Lobular Carcinoma
333	66	FEMALE	0.286380	1.39980	0.318830	0.836050	II	Infiltrating Ductal Carcinoma

334 rows × 14 columns

Next steps:

 View recommended plots

```
x['ER status'] = x['ER status'].replace({'Positive': 1, 'Negative' : 0})
x['HER2 status'] = x['HER2 status'].replace({'Positive': 1, 'Negative' : 0})
x['PR status'] = x['PR status'].replace({'Positive': 1, 'Negative' : 0})
x
```

	Age	Gender	Protein1	Protein2	Protein3	Protein4	Tumour_Stage	Histology
0	42	FEMALE	0.952560	2.15000	0.007972	-0.048340	II	Infiltrating Ductal Carcinoma
1	54	FEMALE	0.000000	1.38020	-0.498030	-0.507320	II	Infiltrating Ductal Carcinoma
2	63	FEMALE	-0.523030	1.76400	-0.370190	0.010815	II	Infiltrating Ductal Carcinoma
3	78	FEMALE	-0.876180	0.12943	-0.370380	0.132190	I	Infiltrating Ductal Carcinoma
4	42	FEMALE	0.226110	1.74910	-0.543970	-0.390210	II	Infiltrating Ductal Carcinoma
...
329	59	FEMALE	0.024598	1.40050	0.024751	0.280320	II	Infiltrating Ductal Carcinoma
330	41	FEMALE	0.100120	-0.46547	0.472370	-0.523870	I	Infiltrating Ductal Carcinoma
331	54	FEMALE	0.753820	1.64250	-0.332850	0.857860	II	Infiltrating Ductal Carcinoma
332	74	FEMALE	0.972510	1.42680	-0.366570	-0.107820	II	Infiltrating Lobular Carcinoma
333	66	FEMALE	0.286380	1.39980	0.318830	0.836050	II	Infiltrating Ductal Carcinoma

334 rows × 14 columns

Next steps:

 View recommended plots

```
x['Gender'] = x['Gender'].replace({'FEMALE': 1, 'MALE':0})

d4=x.drop('Histology',axis=1)
d1=d4.drop('Tumour_Stage',axis=1)
d2=d1.drop('Surgery_type',axis=1)
d3=d2.drop('Date_of_Surgery',axis=1)
d=d3.drop('Date_of_Last_Visit',axis=1)
d
```

	Age	Gender	Protein1	Protein2	Protein3	Protein4	ER status	PR status	HER2 status
0	42	1	0.952560	2.15000	0.007972	-0.048340	1	1	0
1	54	1	0.000000	1.38020	-0.498030	-0.507320	1	1	0
2	63	1	-0.523030	1.76400	-0.370190	0.010815	1	1	0
3	78	1	-0.876180	0.12943	-0.370380	0.132190	1	1	0
4	42	1	0.226110	1.74910	-0.543970	-0.390210	1	1	1

Next steps

[View recommended plots](#)

329	59	1	0.024598	1.40050	0.024751	0.280320	1	1	1
-----	----	---	----------	---------	----------	----------	---	---	---

```
y = y.replace({'Alive':1,'Dead':0})
```

331	54	1	0.753820	1.64250	-0.332850	0.857860	1	1	0
-----	----	---	----------	---------	-----------	----------	---	---	---

```
import numpy as np
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.svm import SVC
```

```
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
```

```
d_train, d_test, y_train, y_test = train_test_split(d, y, test_size=0.2, random_state=42)
```

```
svm_classifier = SVC()
```

```
svm_classifier.fit(d_train, y_train)
```

```
y_pred = svm_classifier.predict(d_test)
```

```
accuracy = accuracy_score(y_test, y_pred)
```

```
precision = precision_score(y_test, y_pred)
```

```
recall = recall_score(y_test, y_pred)
```

```
f1 = f1_score(y_test, y_pred)
```

```
print("Accuracy:", accuracy)
```

```
print("Precision:", precision)
```

```
print("Recall:", recall)
```

```
print("F1-score:", f1)
```

```
Accuracy: 0.7761194029850746
```

```
Precision: 0.7761194029850746
```

```
Recall: 1.0
```

```
F1-score: 0.8739495798319328
```