```
In [70]:
         import pandas as pd
          a=pd.read_csv('/content/archive (13).zip')
          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)
              Age
42
                   Gender
                          Protein1
                                     Protein2 Protein3 Protein4 Tumour_Stage
                                      2.15000 0.007972 -0.048340
        0
                   FEMALE
                           0.952560
                   FEMALE
                          0.000000
                                      1.38020 -0.498030 -0.507320
               63
                   FEMALE -0.523030
FEMALE -0.876180
                                      1.76400 -0.370190
0.12943 -0.370380
         2
                                                          0.010815
                                                                             ΙI
         3
               78
                                                          0.132190
         4
                   FEMALE 0.226110
                                      1.74910 -0.543970 -0.390210
                                                                             II
         329
              59
                   FFMALE 0.024598
                                      1.40050 0.024751 0.280320
                                                                             TT
         330
               41
                   FEMALE
                          0.100120
                                     -0.46547
                                              0.472370 -0.523870
                                      1.64250 -0.332850 0.857860
1.42680 -0.366570 -0.107820
                   FFMALE
                          0.753820
                                                                             ΤT
         331
               74
                   FFMALE 0.972510
         332
                                                                             TT
         333
                   FEMALE 0.286380
                                      1.39980 0.318830 0.836050
                                   Histology ER status PR status HER2 status \ \
        0
               Infiltrating Ductal Carcinoma Positive Positive
                                                                     Negative
               Infiltrating Ductal Carcinoma
Infiltrating Ductal Carcinoma
                                              Positive
                                                         Positive
         2
                                              Positive
                                                         Positive
                                                                     Negative
               Infiltrating Ductal Carcinoma
                                              Positive Positive
                                                                     Negative
         4
               Infiltrating Ductal Carcinoma Positive Positive
         329
               Infiltrating Ductal Carcinoma Positive Positive
                                                                     Positive
             Infiltrating Ductal Carcinoma
Infiltrating Ductal Carcinoma
Infiltrating Ductal Carcinoma
Infiltrating Lobular Carcinoma
Positive Positive
Positive
         330
                                                                     Positive
         331
                                                                     Negative
         332
                                                                     Negative
               Infiltrating Ductal Carcinoma Positive Positive
         333
                                                                     Negative
                             Surgery_type Date_of_Surgery Date_of_Last_Visit
        0
                                    Other
                                                 20-May-18
                                                                    26-Aug-18
                                    Other
                                                 26-Apr-18
                                                                    25- Jan-19
                                                 24-Aug-18
                                                                    08-Apr-20
                               Lumpectomy
         3
                                    0ther
                                                 16-Nov-18
                                                                    28-Jul-20
        4
                               Lumpectomy
                                                 12-Dec-18
                                                                    05-Jan-19
         329
                               Lumpectomy
                                                 15-Jan-19
                                                                    27-Mar-20
             Modified Radical Mastectomy
         330
                                                 25- Jul - 18
                                                                    23-Apr-19
                        Simple Mastectomy
                                                 26-Mar-19
                                                                    11-0ct-19
         331
                                                 26-Nov-18
                                                                    05-Dec-18
                               Lumpectomy
        333
            Modified Radical Mastectomy
                                                04-Feb-19
                                                                    10-Aug-19
             Patient Status
                      Alive
        0
                       Dead
                      Alive
         3
                      Alive
         4
                      Alive
         329
                      Alive
         330
                      Alive
                      Alive
         332
         333
                       Dead
         [334 rows x 15 columns]
                Gender
                        Protein1 Protein2 Protein3 Protein4 Tumour Stage
            Age
                 FEMALE
                                    2.15000 0.007972 -0.048340
                                                                            ΙI
             54
                 FFMALE
                          0.00000
                                    1.38020 -0.498030 -0.507320
                                                                           TT
                 FEMALE
                                    1.76400 -0.370190 0.010815
0.12943 -0.370380 0.132190
             63
                         -0.52303
                                                                           ΙI
                         -0.87618
                 FEMALE
         4
             42
                FFMALE
                         0.22611
                                    1.74910 -0.543970 -0.390210
                                                                           TT
                                Histology ER status PR status HER2 status Surgery_type
           Infiltrating Ductal Carcinoma Positive Positive Infiltrating Ductal Carcinoma Positive Positive
                                                                                   n+her
        0
                                                                  Negative
                                                                                  0ther
                                                                  Negative
            Infiltrating Ductal Carcinoma
                                           Positive
                                                      Positive
                                                                  Negative
                                                                             Lumpectomy
            Infiltrating Ductal Carcinoma Positive
                                                      Positive
                                                                  Negative
                                                                                  Other
            Infiltrating Ductal Carcinoma Positive Positive
                                                                  Positive
                                                                             Lumpectomy
           Date_of_Surgery Date_of_Last_Visit Patient_Status
        0
                 20-May-18
                                    26-Aug-18
                                                        Alive
                 26-Apr-18
                                    25-Jan-19
                                                         Dead
        24-Aug-18
                                    08-Apr-20
                                                        Alive
        dtype='object')
In [71]: b=a.fillna(0)
         b
```

										Status	Status	Status	
	0	42	FEMALE	0.952560	2.15000	0.007972	-0.048340	II	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Other 2
	1	54	FEMALE	0.000000	1.38020	-0.498030	-0.507320	II	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Other 2
	2	63	FEMALE	-0.523030	1.76400	-0.370190	0.010815	II	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Lumpectomy 2
	3	78	FEMALE	-0.876180	0.12943	-0.370380	0.132190	I	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Other
	4	42	FEMALE	0.226110	1.74910	-0.543970	-0.390210	II	Infiltrating Ductal Carcinoma	Positive	Positive	Positive	Lumpectomy
	329	59	FEMALE	0.024598	1.40050	0.024751	0.280320	II	Infiltrating Ductal Carcinoma	Positive	Positive	Positive	Lumpectomy
	330	41	FEMALE	0.100120	-0.46547	0.472370	-0.523870	I	Infiltrating Ductal Carcinoma	Positive	Positive	Positive	Modified Radical 2 Mastectomy
	331	54	FEMALE	0.753820	1.64250	-0.332850	0.857860	II	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Simple Mastectomy
	332	74	FEMALE	0.972510	1.42680	-0.366570	-0.107820	II	Infiltrating Lobular Carcinoma	Positive	Positive	Negative	Lumpectomy 2
	333	66	FEMALE	0.286380	1.39980	0.318830	0.836050	II	Infiltrating Ductal Carcinoma	Positive	Positive	Negative	Modified Radical Mastectomy
	334 rc	ows ×	15 columns	S									
In [72]:	y=b[['Pat	ient_Sta	tus']									
Out [72]:	у	Aliv Dea Aliv Aliv	re id re re	tus']									
Out [72]:	y 0 1 2 3 4 329 330 331 332 333	Aliv Dea Aliv Aliv Aliv Dea Aliv Dea	re d d re e e re		RA dtyne:	object							
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name:	Aliv Dea Aliv Aliv Aliv Dea Aliv Dea Patie	re id e re re d e e d d e d e d e nt_Status,	tus'] , Length: 33 t_Status'									
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name:	Aliv Dea Aliv Aliv Aliv Dea Aliv Dea Patie	re id e re re d e e d d e d e d e nt_Status,	, Length: 33									
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name:	Aliv Dea Aliv Aliv Aliv Dea Aliv Dea Patie	re id e re re d e e d d e d e d e nt_Status,	, Length: 33 t_Status'			Protein4	Tumour_Stage	Histology	ER status	PR status	HER2 status	Surgery_type I
Out [72]:	y 0 1 2 3 3 4 329 330 331 332 333 Name:	Aliv Dea Aliv Aliv Aliv Dea Aliv Dea Patie	re d e e e e d d e int_Status, ('Patien	, Length: 33 t_Status'	,axis=1) Protein2	Protein3	Protein4 -0.048340		Histology Infiltrating Ductal Carcinoma	status	status		
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name: x=b.x	Aliv Dea Aliv Aliv Aliv Dea Aliv Dea Patie	re id e re re re d e re d e re d re d re	, Length: 33 t_Status' Protein1 0.952560	, axis=1) Protein2 2.15000	Protein3 0.007972		II	Infiltrating Ductal	Positive	Positive	status	Other 2
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name: x=b. x	Aliv Dea Aliv Aliv Aliv Dea Aliv Dea Aliv Dea Aliv Aliv	re dd ee	Protein1 0.952560 0.000000	,axis=1) Protein2 2.15000 1.38020	Protein3 0.007972 -0.498030	-0.048340		Infiltrating Ductal Carcinoma Infiltrating Ductal	Positive Positive	Positive Positive	status Negative Negative	Other 2
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name: x=b.x	Aliv Dea Aliv Aliv Dea Aliv Dea Aliv Dea Aliv Dea Aliv Dea Aliv	re de de e e e e e e e e e e e e e e e e	Protein1 0.952560 0.000000 -0.523030	,axis=1) Protein2 2.15000 1.38020 1.76400	Protein3 0.007972 -0.498030 -0.370190	-0.048340 -0.507320		Infiltrating Ductal Carcinoma Infiltrating Ductal Carcinoma Infiltrating Ductal	Positive Positive Positive	Positive Positive Positive	status Negative Negative	Other 2 Other 2 Lumpectomy 2
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name: x=b.x	Aliv Dea Aliv Aliv Aliv Dea Aliv Dea Patie	re de de e de e de e de e de e de e de	Protein1 0.952560 0.000000 -0.523030 -0.876180	,axis=1) Protein2 2.15000 1.38020 1.76400 0.12943	Protein3 0.007972 -0.498030 -0.370190 -0.370380	-0.048340 -0.507320 0.010815		Infiltrating Ductal Carcinoma Infiltrating Ductal Carcinoma Infiltrating Ductal Carcinoma Infiltrating Ductal Carcinoma	Positive Positive Positive Positive	Positive Positive Positive Positive	Negative Negative Negative Negative	Other 2 Other 2 Lumpectomy 2
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name: x=b.x 0 1 2 3 4	Alive Dead Alive Alive Alive Dead Alive Dead Patie Alive	re de de e de e de e de e de e de e de	Protein1 0.952560 0.000000 -0.523030 -0.876180	,axis=1) Protein2 2.15000 1.38020 1.76400 0.12943	Protein3 0.007972 -0.498030 -0.370190 -0.370380	-0.048340 -0.507320 0.010815 0.132190		Infiltrating Ductal Carcinoma	Positive Positive Positive Positive	Positive Positive Positive Positive	Negative Negative Negative Negative	Other 2 Other 2 Lumpectomy 2 Other -
Out [72]:	y 0 1 2 3 4 329 330 331 332 333 Name: x=b.x 0 1 2 3 4	Alive Dea Alive Alive Alive Alive Alive Dea Alive Dea Patie Alive Dea Patie Alive Dea Patie Alive Dea Alive Dea Patie Alive De	re de de e e e e e e e e e e e e e e e e	Protein1 0.952560 0.000000 -0.523030 -0.876180 0.226110	,axis=1) Protein2 2.15000 1.38020 1.76400 0.12943 1.74910	Protein3 0.007972 -0.498030 -0.370190 -0.370380 -0.543970	-0.048340 -0.507320 0.010815 0.132190 -0.390210 		Infiltrating Ductal Carcinoma	Positive Positive Positive Positive	Positive Positive Positive Positive Positive	Negative Negative Negative Negative Positive	Other 2 Other 2 Lumpectomy 2 Other 4

Age Gender Protein1 Protein2 Protein3 Protein4 Tumour_Stage Histology ER status status

HER2 Surgery_type

ER

PR

Out [71]:

		Age	Gender	Protein1	Protein2	Protein3	Protein4	Tumour_Sta	age Histology	ER status			HER2 Surge	y_type l
	330	41	FEMALE	0.100120	-0.46547	0.472370	-0.523870	I	Infiltrating Ductal Carcinoma	Positive	Positiv	ve Pos	Modifi itive Radica Maste	ıl 2
	331	54	FEMALE	0.753820	1.64250	-0.332850	0.857860	II	Infiltrating Ductal Carcinoma		Positiv	⁄e Neg	ative Simple Maste	
	332	74	FEMALE	0.972510	1.42680	-0.366570	-0.107820	II	Infiltrating Lobular Carcinoma	Positive	Positiv	/e Neg	ative Lumpe	
	333	66	FEMALE	0.286380	1.39980	0.318830	0.836050	II	Infiltrating Ductal Carcinoma		Positiv	⁄e Neg	Modifi ative Radica Maste	ıl (
	334 rc	ws ×	14 column	S										
In [74]:	x['H	IER2	status']	= x['HEF	R2 status	'].replac	e({'Posit	:ive': 1,	ative' : 0}) 'Negative' : ative' : 0})	0})				
Out [74]:		Age	Gender	Protein1	Protein2	Protein3	Protein4	Tumour_Sta	age Histology	ER status	PR status	HER2 status	Surgery_typ	e Date_c
	0	42	FEMALE	0.952560	2.15000	0.007972	-0.048340	II	Infiltrating Ductal Carcinoma	1	1	0	Other	20-Ma
	1	54	FEMALE	0.000000	1.38020	-0.498030	-0.507320	II	Infiltrating Ductal Carcinoma	1	1	0	Other	26-Apr
	2	63	FEMALE	-0.523030	1.76400	-0.370190	0.010815	II	Infiltrating Ductal Carcinoma	1	1	0	Lumpectom	y 24-Auç
	3	78	FEMALE	-0.876180	0.12943	-0.370380	0.132190	I	Infiltrating Ductal Carcinoma	1	1	0	Other	16-Nov
	4	42	FEMALE	0.226110	1.74910	-0.543970	-0.390210	II	Infiltrating Ductal Carcinoma		1	1	Lumpectom	y 12-Dec
	329	59	FEMALE	0.024598	1.40050	0.024751	0.280320	II	Infiltrating Ductal Carcinoma	1	1	1	Lumpectom	y 15-Jar
	330	41	FEMALE	0.100120	-0.46547	0.472370	-0.523870	I	Infiltrating Ductal Carcinoma	1	1	1	Modified Radical Mastectomy	25-Jul
	331	54	FEMALE	0.753820	1.64250	-0.332850	0.857860	II	Infiltrating Ductal Carcinoma	1	1	0	Simple Mastectomy	26-Ma
	332	74	FEMALE	0.972510	1.42680	-0.366570	-0.107820	II	Infiltrating Lobular Carcinoma	1	1	0	Lumpectom	y 26-Nov
	333	66	FEMALE	0.286380	1.39980	0.318830	0.836050	II	Infiltrating Ductal Carcinoma	1	1	0	Modified Radical Mastectomy	04-Feb
	334 rc	ows × 1	14 columns	S										
In [75]:	x['0	ende	r'] = x['Gender']].replace	({'FEMALE	E': 1,'MAL	.E':0})						
In [76]:	d1=0 d2=0 d3=0	14.dr 11.dr 12.dr	op('Tumo op('Surg op('Date	logy',axi ur_Stage' ery_type' _of_Surge of_Last_\	',axis=1) ',axis=1) ery',axis	=1)								
Out [76]:		Age	Gender	Protein1	Protein2	Protein3	Protein4	ER status P	R 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					

63 1 -0.523030 1.76400 -0.370190 0.010815 1 1 0

		Age	Gender	Protein1	Protein2	Protein3	Protein4	ER status	PR status	HER2 status
	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
3	29	59	1	0.024598	1.40050	0.024751	0.280320	1	1	1
3	30	41	1	0.100120	-0.46547	0.472370	-0.523870	1	1	1
3	31	54	1	0.753820	1.64250	-0.332850	0.857860	1	1	0
3	32	74	1	0.972510	1.42680	-0.366570	-0.107820	1	1	0
3	33	66	1	0.286380	1.39980	0.318830	0.836050	1	1	0

334 rows × 9 columns

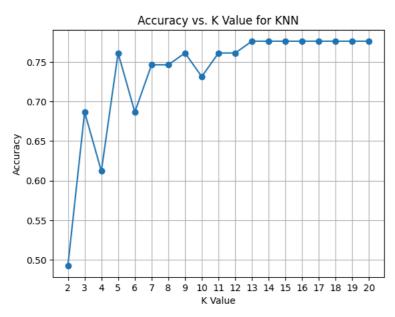
```
In [77]: y = y.replace({'Alive':1,'Dead':0})
```

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
```

```
In [79]: d_train, d_test, y_train, y_test = train_test_split(d, y, test_size=0.2, random_state=42)
```

```
In [80]: k_values = range(2, 21)
    accuracy_scores = []
    for k in k_values:
        knn = KNeighborsClassifier(n_neighbors=k)
        knn.fit(d_train, y_train)
        y_pred = knn.predict(d_test)
        accuracy = accuracy_score(y_test, y_pred)
        accuracy_scores.append(accuracy)
```

```
In [81]: plt.plot(k_values, accuracy_scores, marker='o')
  plt.title('Accuracy vs. K Value for KNN')
  plt.xlabel('K Value')
  plt.ylabel('Accuracy')
  plt.xticks(range(2, 21))
  plt.grid(True)
  plt.show()
```



```
In [82]:
    decision_tree = DecisionTreeClassifier()
    decision_tree.fit(d_train, y_train)
    y_pred_dt = decision_tree.predict(d_test)
```

```
In [83]: accuracy_dt = accuracy_score(y_test, y_pred_dt)
precision_dt = precision_score(y_test, y_pred_dt)
```

```
f1_score_dt = f1_score(y_test, y_pred_dt)

In [84]: print("Decision Tree Metrics:")
    print("Accuracy:", accuracy_dt)
    print("Precision:", precision_dt)
    print("Recall:", recall_dt)
    print("F1 Score:", f1_score_dt)

Decision Tree Metrics:
    Accuracy: 0.6417910447761194
    Precision: 0.8043478260869565
    Recall: 0.7115384615384616
    F1 Score: 0.7551020408163265

In [85]: print("\nComparison:")
    print("KNN Accuracy:", max(accuracy_scores))
    print("Decision Tree Accuracy:", accuracy_dt)
```

recall_dt = recall_score(y_test, y_pred_dt)