In [1]:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

In [2]:

```
df=pd.read_csv('titanic_train.csv')
```

In [3]:

```
df.head()
```

Out[3]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

In [4]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   PassengerId  891 non-null    int64  
 1   Survived     891 non-null    int64  
 2   Pclass       891 non-null    int64  
 3   Name         891 non-null    object 
 4   Sex          891 non-null    object 
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64  
 7   Parch        891 non-null    int64  
 8   Ticket       891 non-null    object 
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object 
 11  Embarked     889 non-null    object 
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

```
In df.info we can see that this data has total 891 records and total 12 columns.
From which Cabin column has null values more than 30% so we will drop that.
Age and Embarked columns too has null values but we will replace those will mean and mode respectively.
All the data is in proper datatype.
```

In [5]:

```
df.describe()
```

Out[5]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

```
In df.describe() we can see count mean standard deviation min 1st quartile 2nd quartile and 3rd quartile.
We can see all the numerical columns here,
we can see the correlation of all columns by comparing mean and median .
And also we can see the value of standard deviation.
```

In [6]:

```
df.drop('Cabin',axis=1,inplace=True)
```

In [7]:

```
from sklearn.impute import SimpleImputer
si= SimpleImputer(missing_values=np.nan,strategy='mean')
df[['Age']]=si.fit_transform(df[['Age']])
```

In [8]:

```
from sklearn.impute import SimpleImputer
si= SimpleImputer(missing_values=np.nan,strategy='most_frequent')
df[['Embarked']]=si.fit_transform(df[['Embarked']])
```

In [9]:

```
df['Embarked'].unique()
```

Out[9]:

```
array(['S', 'C', 'Q'], dtype=object)
```

In [10]:

```
features=df.iloc[:,2:]
```

In [11]:

```
target=df.iloc[:,[1]]
target
```

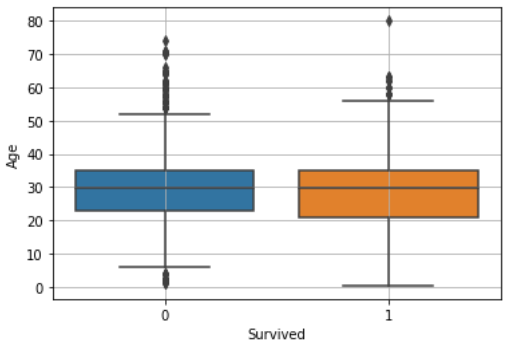
Out[11]:

	Survived
0	0
1	1
2	1
3	1
4	0
...	...
886	0
887	1
888	0
889	1
890	0

891 rows × 1 columns

In [12]:

```
sns.boxplot(data=features,x=target['Survived'],y=features['Age'])
plt.grid()
plt.show()
```



In [13]:

```
features[(features.Age>51)&(target.Survived==0)]
```

Out[13]:

	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
6	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	S
33	2	Wheadon, Mr. Edward H	male	66.0	0	0	C.A. 24579	10.5000	S
54	1	Ostby, Mr. Engelhart Cornelius	male	65.0	0	1	113509	61.9792	C
94	3	Coxon, Mr. Daniel	male	59.0	0	0	364500	7.2500	S
96	1	Goldschmidt, Mr. George B	male	71.0	0	0	PC 17754	34.6542	C
116	3	Connors, Mr. Patrick	male	70.5	0	0	370369	7.7500	Q
124	1	White, Mr. Percival Wayland	male	54.0	0	1	35281	77.2875	S
152	3	Meo, Mr. Alfonzo	male	55.5	0	0	A.5. 11206	8.0500	S
170	1	Van der hoef, Mr. Wyckoff	male	61.0	0	0	111240	33.5000	S
174	1	Smith, Mr. James Clinch	male	56.0	0	0	17764	30.6958	C
232	2	Sjostedt, Mr. Ernst Adolf	male	59.0	0	0	237442	13.5000	S
249	2	Carter, Rev. Ernest Courtenay	male	54.0	1	0	244252	26.0000	S
252	1	Stead, Mr. William Thomas	male	62.0	0	0	113514	26.5500	S
262	1	Taussig, Mr. Emil	male	52.0	1	1	110413	79.6500	S
280	3	Duane, Mr. Frank	male	65.0	0	0	336439	7.7500	Q
317	2	Moraweck, Dr. Ernest	male	54.0	0	0	29011	14.0000	S
326	3	Nysveen, Mr. Johan Hansen	male	61.0	0	0	345364	6.2375	S
438	1	Fortune, Mr. Mark	male	64.0	1	4	19950	263.0000	S
456	1	Millet, Mr. Francis Davis	male	65.0	0	0	13509	26.5500	S
467	1	Smart, Mr. John Montgomery	male	56.0	0	0	113792	26.5500	S
487	1	Kent, Mr. Edward Austin	male	58.0	0	0	11771	29.7000	C
492	1	Molson, Mr. Harry Markland	male	55.0	0	0	113787	30.5000	S
493	1	Artagaveytia, Mr. Ramon	male	71.0	0	0	PC 17609	49.5042	C
545	1	Nicholson, Mr. Arthur Ernest	male	64.0	0	0	693	26.0000	S
555	1	Wright, Mr. George	male	62.0	0	0	113807	26.5500	S
582	2	Downton, Mr. William James	male	54.0	0	0	28403	26.0000	S
625	1	Sutton, Mr. Frederick	male	61.0	0	0	36963	32.3208	S
626	2	Kirkland, Rev. Charles Leonard	male	57.0	0	0	219533	12.3500	Q
659	1	Newell, Mr. Arthur Webster	male	58.0	0	2	35273	113.2750	C
672	2	Mitchell, Mr. Henry Michael	male	70.0	0	0	C.A. 24580	10.5000	S
684	2	Brown, Mr. Thomas William Solomon	male	60.0	1	1	29750	39.0000	S
694	1	Weir, Col. John	male	60.0	0	0	113800	26.5500	S
695	2	Chapman, Mr. Charles Henry	male	52.0	0	0	248731	13.5000	S
714	2	Greenberg, Mr. Samuel	male	52.0	0	0	250647	13.0000	S
745	1	Crosby, Capt. Edward Gifford	male	70.0	1	1	WE/P 5735	71.0000	S
772	2	Mack, Mrs. (Mary)	female	57.0	0	0	S.O./P.P. 3	10.5000	S
851	3	Svensson, Mr. Johan	male	74.0	0	0	347060	7.7750	S

In [14]:

```
features.drop([6,33,54,94,96,116,124,152,170,174,232,249,252,262,280,317,326,438,456,467,487,432,492,493,545,555,582,625,626,659,672,684,
```

In [15]:

```
features
```

Out[15]:

	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
0	3	Braund, Mr. Owen Harris	male	22.000000	1	0	A/5 21171	7.2500	S
1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.000000	1	0	PC 17599	71.2833	C
2	3	Heikkinen, Miss. Laina	female	26.000000	0	0	STON/O2. 3101282	7.9250	S
3	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.000000	1	0	113803	53.1000	S
4	3	Allen, Mr. William Henry	male	35.000000	0	0	373450	8.0500	S
...	...	...	...	...	...	...	...	...	...
886	2	Montvila, Rev. Juozas	male	27.000000	0	0	211536	13.0000	S
887	1	Graham, Miss. Margaret Edith	female	19.000000	0	0	112053	30.0000	S
888	3	Johnston, Miss. Catherine Helen "Carrie"	female	29.699118	1	2	W./C. 6607	23.4500	S
889	1	Behr, Mr. Karl Howell	male	26.000000	0	0	111369	30.0000	C
890	3	Dooley, Mr. Patrick	male	32.000000	0	0	370376	7.7500	Q

853 rows × 9 columns

In [16]:

```
features[(features.Age<8)&(target.Survived==0)]
```

Out[16]:

	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
7	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	S
16	3	Rice, Master. Eugene	male	2.0	4	1	382652	29.1250	Q
50	3	Panula, Master. Juha Niilo	male	7.0	4	1	3101295	39.6875	S
63	3	Skoog, Master. Harald	male	4.0	3	2	347088	27.9000	S
119	3	Andersson, Miss. Ellis Anna Maria	female	2.0	4	2	347082	31.2750	S
164	3	Panula, Master. Eino Viljami	male	1.0	4	1	3101295	39.6875	S
171	3	Rice, Master. Arthur	male	4.0	4	1	382652	29.1250	Q
205	3	Strom, Miss. Telma Matilda	female	2.0	0	1	347054	10.4625	S
278	3	Rice, Master. Eric	male	7.0	4	1	382652	29.1250	Q
297	1	Allison, Miss. Helen Loraine	female	2.0	1	2	113781	151.5500	S
374	3	Palsson, Miss. Stina Viola	female	3.0	3	1	349909	21.0750	S
386	3	Goodwin, Master. Sidney Leonard	male	1.0	5	2	CA 2144	46.9000	S
642	3	Skoog, Miss. Margit Elizabeth	female	2.0	3	2	347088	27.9000	S
813	3	Andersson, Miss. Ebba Iris Alfrida	female	6.0	4	2	347082	31.2750	S
824	3	Panula, Master. Urho Abraham	male	2.0	4	1	3101295	39.6875	S
850	3	Andersson, Master. Sigvard Harald Elias	male	4.0	4	2	347082	31.2750	S

In [17]:

```
features.drop([7,16,50,63,119,164,171,205,278,297,374,386,642,813,824,850],inplace=True)
```

In [18]:

```
features[(features.Age>56)&(target.Survived==1)]
```

Out[18]:

	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
11	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	S
195	1	Lurette, Miss. Elise	female	58.0	0	0	PC 17569	146.5208	C
268	1	Graham, Mrs. William Thompson (Edith Junkins)	female	58.0	0	1	PC 17582	153.4625	S
275	1	Andrews, Miss. Kornelia Theodosia	female	63.0	1	0	13502	77.9583	S
366	1	Warren, Mrs. Frank Manley (Anna Sophia Atkinson)	female	60.0	1	0	110813	75.2500	C
483	3	Turkula, Mrs. (Hedwig)	female	63.0	0	0	4134	9.5875	S
570	2	Harris, Mr. George	male	62.0	0	0	S.W./PP 752	10.5000	S
587	1	Frolicher-Stehli, Mr. Maxmillian	male	60.0	1	1	13567	79.2000	C
630	1	Barkworth, Mr. Algernon Henry Wilson	male	80.0	0	0	27042	30.0000	S
829	1	Stone, Mrs. George Nelson (Martha Evelyn)	female	62.0	0	0	113572	80.0000	S

In [19]:

```
features.drop([11,195,268,275,366,483,570,587,630,823],inplace=True)
```

In [20]:

```
features
```

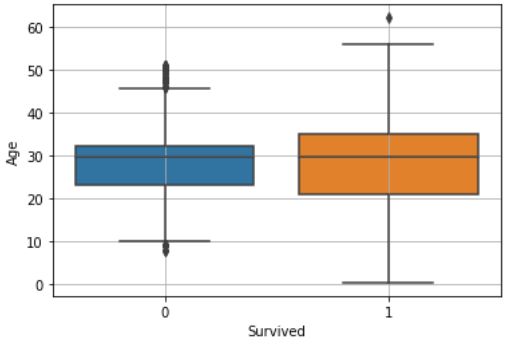
Out[20]:

	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
0	3	Braund, Mr. Owen Harris	male	22.000000	1	0	A/5 21171	7.2500	S
1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.000000	1	0	PC 17599	71.2833	C
2	3	Heikkinen, Miss. Laina	female	26.000000	0	0	STON/O2. 3101282	7.9250	S
3	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.000000	1	0	113803	53.1000	S
4	3	Allen, Mr. William Henry	male	35.000000	0	0	373450	8.0500	S
...	...	...	...	...	...	...	...	...	...
886	2	Montvila, Rev. Juozas	male	27.000000	0	0	211536	13.0000	S
887	1	Graham, Miss. Margaret Edith	female	19.000000	0	0	112053	30.0000	S
888	3	Johnston, Miss. Catherine Helen "Carrie"	female	29.699118	1	2	W./C. 6607	23.4500	S
889	1	Behr, Mr. Karl Howell	male	26.000000	0	0	111369	30.0000	C
890	3	Dooley, Mr. Patrick	male	32.000000	0	0	370376	7.7500	Q

827 rows × 9 columns

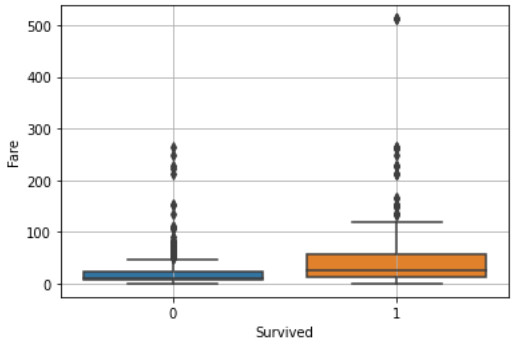
In [21]:

```
sns.boxplot(data=features,x=target['Survived'],y=features['Age'])
plt.grid()
plt.show()
```



In [22]:

```
sns.boxplot(data=features,x=target['Survived'],y=features['Fare'])
plt.grid()
plt.show()
```



In [23]:

```
features[(features.Fare>50)&(target.Survived==0)]
```

Out[23]:

	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
27	1	Fortune, Mr. Charles Alexander	male	19.000000	3	2	19950	263.0000	S
34	1	Meyer, Mr. Edgar Joseph	male	28.000000	1	0	PC 17604	82.1708	C
35	1	Holverson, Mr. Alexander Oskar	male	42.000000	1	0	113789	52.0000	S
62	1	Harris, Mr. Henry Birkhardt	male	45.000000	1	0	36973	83.4750	S
72	2	Hood, Mr. Ambrose Jr	male	21.000000	0	0	S.O.C. 14879	73.5000	S
92	1	Chaffee, Mr. Herbert Fuller	male	46.000000	1	0	W.E.P. 5734	61.1750	S
102	1	White, Mr. Richard Frasar	male	21.000000	0	1	35281	77.2875	S
110	1	Porter, Mr. Walter Chamberlain	male	47.000000	0	0	110465	52.0000	S
118	1	Baxter, Mr. Quigg Edmond	male	24.000000	0	1	PC 17558	247.5208	C
120	2	Hickman, Mr. Stanley George	male	21.000000	2	0	S.O.C. 14879	73.5000	S
137	1	Futrelle, Mr. Jacques Heath	male	37.000000	1	0	113803	53.1000	S
139	1	Giglio, Mr. Victor	male	24.000000	0	0	PC 17593	79.2000	C
155	1	Williams, Mr. Charles Duane	male	51.000000	0	1	PC 17597	61.3792	C
159	3	Sage, Master. Thomas Henry	male	29.699118	8	2	CA. 2343	69.5500	S
169	3	Ling, Mr. Lee	male	28.000000	0	0	1601	56.4958	S
180	3	Sage, Miss. Constance Gladys	female	29.699118	8	2	CA. 2343	69.5500	S
201	3	Sage, Mr. Frederick	male	29.699118	8	2	CA. 2343	69.5500	S
245	1	Minahan, Dr. William Edward	male	44.000000	2	0	19928	90.0000	Q
324	3	Sage, Mr. George John Jr	male	29.699118	8	2	CA. 2343	69.5500	S
332	1	Graham, Mr. George Edward	male	38.000000	0	1	PC 17582	153.4625	S
336	1	Pears, Mr. Thomas Clinton	male	29.000000	1	0	113776	66.6000	S
373	1	Ringhini, Mr. Sante	male	22.000000	0	0	PC 17760	135.6333	C
377	1	Widener, Mr. Harry Elkins	male	27.000000	0	2	113503	211.5000	C
385	2	Davies, Mr. Charles Henry	male	18.000000	0	0	S.O.C. 14879	73.5000	S
434	1	Silvey, Mr. William Baird	male	50.000000	1	0	13507	55.9000	S
475	1	Clifford, Mr. George Quincy	male	29.699118	0	0	110465	52.0000	S
498	1	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	25.000000	1	2	113781	151.5500	S
505	1	Penasco y Castellana, Mr. Victor de Satode	male	18.000000	1	0	PC 17758	108.9000	C
527	1	Farthing, Mr. John	male	29.699118	0	0	PC 17483	221.7792	S
544	1	Douglas, Mr. Walter Donald	male	50.000000	1	0	PC 17761	106.4250	C
557	1	Robbins, Mr. Victor	male	29.699118	0	0	PC 17757	227.5250	C
655	2	Hickman, Mr. Leonard Mark	male	24.000000	2	0	S.O.C. 14879	73.5000	S
665	2	Hickman, Mr. Lewis	male	32.000000	2	0	S.O.C. 14879	73.5000	S
671	1	Davidson, Mr. Thornton	male	31.000000	1	0	F.C. 12750	52.0000	S
698	1	Thayer, Mr. John Borland	male	49.000000	1	1	17421	110.8833	C
741	1	Cavendish, Mr. Tyrell William	male	36.000000	1	0	19877	78.8500	S
748	1	Marvin, Mr. Daniel Warner	male	19.000000	1	0	113773	53.1000	S
789	1	Guggenheim, Mr. Benjamin	male	46.000000	0	0	PC 17593	79.2000	C
792	3	Sage, Miss. Stella Anna	female	29.699118	8	2	CA. 2343	69.5500	S
826	3	Lam, Mr. Len	male	29.699118	0	0	1601	56.4958	S
846	3	Sage, Mr. Douglas Bullen	male	29.699118	8	2	CA. 2343	69.5500	S
863	3	Sage, Miss. Dorothy Edith "Dolly"	female	29.699118	8	2	CA. 2343	69.5500	S
867	1	Roebling, Mr. Washington Augustus II	male	31.000000	0	0	PC 17590	50.4958	S

In [24]:

```
df.drop([27,34,35,62,72,92,102,110,118,120,137,139,155,159,169,180,201,245,324,332,336,373,377,385,434,475,498,505,527,544,557,655,665,67
```

In [25]:

```
features[(features.Fare>110)&(target.Survived==1)]
```

Out[25]:

	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
31	1	Spencer, Mrs. William Augustus (Marie Eugenie)	female	29.699118	1	0	PC 17569	146.5208	C
88	1	Fortune, Miss. Mabel Helen	female	23.000000	3	2	19950	263.0000	S
215	1	Newell, Miss. Madeleine	female	31.000000	1	0	35273	113.2750	C
258	1	Ward, Miss. Anna	female	35.000000	0	0	PC 17755	512.3292	C
269	1	Bissette, Miss. Amelia	female	35.000000	0	0	PC 17760	135.6333	S
299	1	Baxter, Mrs. James (Helene DeLaudeniere Chaput)	female	50.000000	0	1	PC 17558	247.5208	C
305	1	Allison, Master. Hudson Trevor	male	0.920000	1	2	113781	151.5500	S
306	1	Fleming, Miss. Margaret	female	29.699118	0	0	17421	110.8833	C
311	1	Ryerson, Miss. Emily Borie	female	18.000000	2	2	PC 17608	262.3750	C
318	1	Wick, Miss. Mary Natalie	female	31.000000	0	2	36928	164.8667	S
319	1	Spedden, Mrs. Frederic Oakley (Margaretta Corn...	female	40.000000	1	1	16966	134.5000	C
325	1	Young, Miss. Marie Grice	female	36.000000	0	0	PC 17760	135.6333	C
334	1	Frauenthal, Mrs. Henry William (Clara Heinshei...	female	29.699118	1	0	PC 17611	133.6500	S
337	1	Burns, Miss. Elizabeth Margaret	female	41.000000	0	0	16966	134.5000	C
341	1	Fortune, Miss. Alice Elizabeth	female	24.000000	3	2	19950	263.0000	S
380	1	Bidois, Miss. Rosalie	female	42.000000	0	0	PC 17757	227.5250	C
390	1	Carter, Mr. William Ernest	male	36.000000	1	2	113760	120.0000	S
393	1	Newell, Miss. Marjorie	female	23.000000	1	0	35273	113.2750	C
435	1	Carter, Miss. Lucile Polk	female	14.000000	1	2	113760	120.0000	S
550	1	Thayer, Mr. John Borland Jr	male	17.000000	0	2	17421	110.8833	C
581	1	Thayer, Mrs. John Borland (Marian Longstreth M...	female	39.000000	1	1	17421	110.8833	C
609	1	Shutes, Miss. Elizabeth W	female	40.000000	0	0	PC 17582	153.4625	S
660	1	Frauenthal, Dr. Henry William	male	50.000000	2	0	PC 17611	133.6500	S
679	1	Cardeza, Mr. Thomas Drake Martinez	male	36.000000	0	1	PC 17755	512.3292	C
689	1	Madill, Miss. Georgette Alexandra	female	15.000000	0	1	24160	211.3375	S
700	1	Astor, Mrs. John Jacob (Madeleine Talmadge Force)	female	18.000000	1	0	PC 17757	227.5250	C
708	1	Cleaver, Miss. Alice	female	22.000000	0	0	113781	151.5500	S
716	1	Endres, Miss. Caroline Louise	female	38.000000	0	0	PC 17757	227.5250	C
730	1	Allen, Miss. Elisabeth Walton	female	29.000000	0	0	24160	211.3375	S
737	1	Lesurer, Mr. Gustave J	male	35.000000	0	0	PC 17755	512.3292	C
742	1	Ryerson, Miss. Susan Parker "Suzette"	female	21.000000	2	2	PC 17608	262.3750	C
763	1	Carter, Mrs. William Ernest (Lucile Polk)	female	36.000000	1	2	113760	120.0000	S
779	1	Robert, Mrs. Edward Scott (Elisabeth Walton Mc...	female	43.000000	0	1	24160	211.3375	S
802	1	Carter, Master. William Thornton II	male	11.000000	1	2	113760	120.0000	S
856	1	Wick, Mrs. George Dennick (Mary Hitchcock)	female	45.000000	1	1	36928	164.8667	S

In [26]:

```
df.drop([31,88,215,258,269,299,305,306,311,318,319,325,334,337,341,380,390,393,435,550,581,609,660,679,689,700,708,716,730,737,763,779,80
```

In [27]:

```
from sklearn.preprocessing import OrdinalEncoder
oe=OrdinalEncoder()
features['Sex']=oe.fit_transform(features[['Sex']])
features['Embarked']=oe.fit_transform(features[['Embarked']])
features['Name']=oe.fit_transform(features[['Name']])
features['Ticket']=oe.fit_transform(features[['Ticket']])
```

In [28]:

```
features
```

Out[28]:

	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
0	3	100.0	1.0	22.000000	1	0	495.0	7.2500	2.0
1	1	176.0	0.0	38.000000	1	0	566.0	71.2833	0.0
2	3	330.0	0.0	26.000000	0	0	635.0	7.9250	2.0
3	1	254.0	0.0	35.000000	1	0	41.0	53.1000	2.0
4	3	15.0	1.0	35.000000	0	0	447.0	8.0500	2.0
...	...	...	...	...	...	...	...	...	...
886	2	515.0	1.0	27.000000	0	0	87.0	13.0000	2.0
887	1	283.0	0.0	19.000000	0	0	12.0	30.0000	2.0
888	3	390.0	0.0	29.699118	1	2	641.0	23.4500	2.0
889	1	74.0	1.0	26.000000	0	0	6.0	30.0000	0.0
890	3	206.0	1.0	32.000000	0	0	441.0	7.7500	1.0

827 rows × 9 columns

In [29]:

```
from sklearn.preprocessing import StandardScaler
sc=StandardScaler()
features.iloc[:,:]=sc.fit_transform(features.iloc[:,:])
```

In [30]:

```
df.drop(['Name','Ticket','PassengerId'],axis=1,inplace=True)
```

In [31]:

```
df.head()
```

Out[31]:

	Survived	Pclass	Sex	Age	SibSp	Parch	Fare	Embarked
0	0	3	male	22.0	1	0	7.2500	S
1	1	1	female	38.0	1	0	71.2833	C
2	1	3	female	26.0	0	0	7.9250	S
3	1	1	female	35.0	1	0	53.1000	S
4	0	3	male	35.0	0	0	8.0500	S

In [32]:

```
x=df.iloc[:,1:]
y=df.iloc[:,0]
```

In [33]:

```
x
```

Out[33]:

	Pclass	Sex	Age	SibSp	Parch	Fare	Embarked
0	3	male	22.000000	1	0	7.2500	S
1	1	female	38.000000	1	0	71.2833	C
2	3	female	26.000000	0	0	7.9250	S
3	1	female	35.000000	1	0	53.1000	S
4	3	male	35.000000	0	0	8.0500	S
...	...	...	...	...	...	...	...
886	2	male	27.000000	0	0	13.0000	S
887	1	female	19.000000	0	0	30.0000	S
888	3	female	29.699118	1	2	23.4500	S
889	1	male	26.000000	0	0	30.0000	C
890	3	male	32.000000	0	0	7.7500	Q

814 rows × 7 columns

In [34]:

```
from sklearn.preprocessing import OrdinalEncoder
oe=OrdinalEncoder()

x[['Sex','Embarked']]=oe.fit_transform(x[['Sex','Embarked']])
```

In [35]:

```
from sklearn.model_selection import train_test_split

xtrain,xtest,ytrain,ytest=train_test_split(x,y,test_size=0.3,random_state=1)
```

In [36]:

```
from sklearn.linear_model import LogisticRegression

logreg=LogisticRegression()
logreg.fit(xtrain,ytrain)

ypred=logreg.predict(xtest)
```

In [37]:

```
from sklearn.metrics import accuracy_score,confusion_matrix,classification_report

ac= accuracy_score(ytest,ypred)
cm=confusion_matrix(ytest,ypred)
cr=classification_report(ytest,ypred)

print(f'Accuracy:- {ac}\n{cm}\n{cr}')
```

```
Accuracy:- 0.763265306122449
[[124  31]
 [ 27  63]]
              precision    recall  f1-score   support

           0       0.82      0.80      0.81       155
           1       0.67      0.70      0.68        90

    accuracy                           0.76       245
   macro avg       0.75      0.75      0.75       245
weighted avg       0.77      0.76      0.76       245
```

In [38]:

```
train=logreg.score(xtrain,ytrain)
test=logreg.score(xtest,ytest)

print(f'Training score:- {train}\nTesting score:- {test}')
```

```
Training score:- 0.8260105448154658
Testing score:- 0.763265306122449
```

In [39]:

```
oe
```

Out[39]:

```
OrdinalEncoder()
```

In [40]:

```
oe.categories
```

Out[40]:

```
'auto'
```

In [41]:

```
def predictsurvived():
    pclass=int(input('Enter passenger Class:- '))
    sex=input('Enter Gender of Passenger:- ')
    age=float(input('Enter the age of passenger:- '))
    sibsp=int(input('Enter no of SibSp of the passenger:- '))
    parch=int(input('Enter no of parch of the passenger:- '))
    fare=int(input('Enter Ticket price:- '))
    embarked=input('Enter Embarked/port of the passenger:- ')

    newob=[pclass,sex,age,sibsp,parch,fare,embarked]
    newob[1],newob[-1]=oe.transform([[newob[1],newob[-1]]])[0]

    v=logreg.predict([newob])[0]

    if v==1:
        print(f'\nWith the given feature passenger will survive')
        return v
    else:
        print(f'\nWith the given feature passenger will not survive')
        return v
```

In [45]:

```
predictsurvived()
```

```
Enter passenger Class:- 1
Enter Gender of Passenger:- male
Enter the age of passenger:- 20
Enter no of SibSp of the passenger:- 0
Enter no of parch of the passenger:- 0
Enter Ticket price:- 512
Enter Embarked/port of the passenger:- C

With the given feature passenger will survive
```

Out[45]:

```
1
```

In [ ]: