

In [421]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import confusion_matrix, accuracy_score, classification_report
from sklearn.preprocessing import scale
import math
train_data='C:/Users/ramya/Desktop/New folder/train_new.csv'
full_data=pd.read_csv(train_data)
full_data.describe()
test_data='C:/Users/ramya/Desktop/New folder/test_new.csv'
full_test=pd.read_csv(test_data)
full_test.describe()
```

Out[421]:

	PassengerId	Pclass	Age	SibSp	Parch	Fare
count	418.000000	418.000000	332.000000	418.000000	418.000000	417.000000
mean	1100.500000	2.265550	30.272590	0.447368	0.392344	35.627188
std	120.810458	0.841838	14.181209	0.896760	0.981429	55.907576
min	892.000000	1.000000	0.170000	0.000000	0.000000	0.000000
25%	996.250000	1.000000	21.000000	0.000000	0.000000	7.895800
50%	1100.500000	3.000000	27.000000	0.000000	0.000000	14.454200
75%	1204.750000	3.000000	39.000000	1.000000	0.000000	31.500000
max	1309.000000	3.000000	76.000000	8.000000	9.000000	512.329200

In [422]:

```
a=full_data[['Age', 'Survived']].groupby(['Age'], as_index=False).mean().sort_values(by='Survived',
, ascending=False)
a.head(10)
```

Out[422]:

	Age	Survived
0	0.42	1.0
9	5.00	1.0
79	63.00	1.0
68	53.00	1.0
1	0.67	1.0
17	13.00	1.0
16	12.00	1.0
87	80.00	1.0
2	0.75	1.0
4	0.92	1.0

In [423]:

```
full_data[['Sex', 'Survived']].groupby(['Sex'], as_index=False).mean().sort_values(by='Survived',
ascending=True)
```

Out[423]:

	Sex	Survived
1	male	0.188908

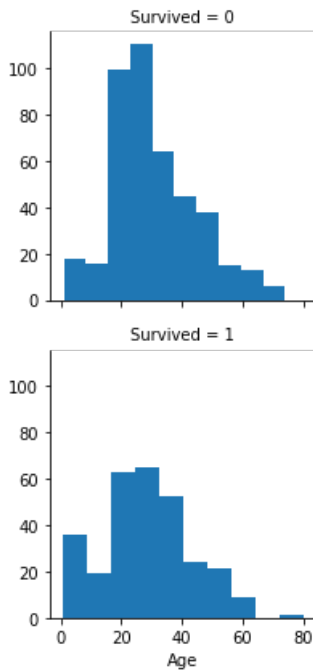
0 female 0.42038

In [424]:

```
g=sns.FacetGrid(full_data, 'Survived')
g.map(plt.hist, 'Age', bins=10)
```

Out[424]:

<seaborn.axisgrid.FacetGrid at 0x22eb83f16d8>



In [426]:

```
full_data[['Parch', 'Survived']].groupby(['Parch'],
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[426]:

	Parch	Survived
4	4	0.000000
6	6	0.000000
5	5	0.200000
0	0	0.343658
2	2	0.500000
1	1	0.550847
3	3	0.600000

In [427]:

```
full_data[['Pclass', 'Survived']].groupby(['Pclass'],
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[427]:

	Pclass	Survived
2	3	0.242363
1	2	0.472826
0	1	0.629630

In [428]:

```
full_data[['Cabin', 'Survived']].groupby(['Cabin'],
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[428]:

	Cabin	Survived
0	A10	0.0
45	B86	0.0
46	B94	0.0
52	C110	0.0
53	C111	0.0
54	C118	0.0
56	C124	0.0
59	C128	0.0
64	C30	0.0
67	C46	0.0
69	C49	0.0
79	C82	0.0
82	C86	0.0
83	C87	0.0
85	C91	0.0
44	B82 B84	0.0
88	C95	0.0
100	D30	0.0
106	D46	0.0
108	D48	0.0
110	D50	0.0
112	D6	0.0
122	E31	0.0
126	E38	0.0
129	E46	0.0
132	E58	0.0
133	E63	0.0
136	E77	0.0
139	F G63	0.0
140	F G73	0.0
...	...	...
68	C47	1.0
22	B3	1.0
66	C45	1.0
65	C32	1.0
57	C125	1.0
58	C126	1.0
48	C101	1.0
81	C85	1.0
19	B20	1.0
47	B96 B98	1.0
101	D33	1.0
30	B42	1.0

99	Cabin	Survived
15	B101	1.0
97	D21	1.0
96	D20	1.0
95	D19	1.0
94	D17	1.0
103	D36	1.0
93	D15	1.0
91	D10 D12	1.0
33	B50	1.0
89	C99	1.0
29	B41	1.0
87	C93	1.0
86	C92	1.0
17	B18	1.0
84	C90	1.0
92	D11	1.0
73	C62 C64	1.0

147 rows × 2 columns

In [376]:

```
full_data[['SibSp', 'Survived']].groupby(['SibSp'],
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[376]:

	SibSp	Survived
5	5	0.000000
6	8	0.000000
4	4	0.166667
3	3	0.250000
0	0	0.345395
2	2	0.464286
1	1	0.535885

In [429]:

```
full_data[['Embarked', 'Survived']].groupby(['Embarked'], as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[429]:

	Embarked	Survived
2	S	0.336957
1	Q	0.389610
0	C	0.553571

In [430]:

```
grid = sns.FacetGrid(full_data, row='Embarked', size=2.2, aspect=1.6)
grid.map(sns.pointplot, 'Cabin', 'Survived', 'Sex', palette='deep')
grid.add_legend()
```

C:\Users\ramya\Anaconda3\lib\site-packages\seaborn\axisgrid.py:230: UserWarning: The `size`

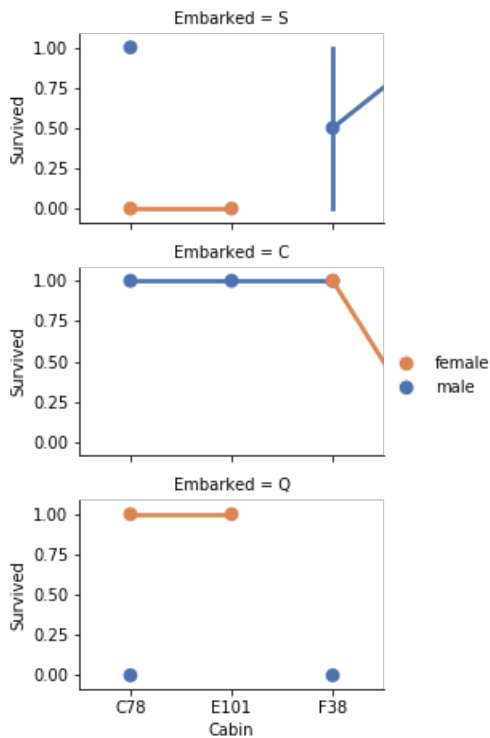
```

paramter has been renamed to `height`; please update your code.
warnings.warn(msg, UserWarning)
C:\Users\ramya\Anaconda3\lib\site-packages\seaborn\axisgrid.py:715: UserWarning: Using the
pointplot function without specifying `order` is likely to produce an incorrect plot.
warnings.warn(warning)
C:\Users\ramya\Anaconda3\lib\site-packages\seaborn\axisgrid.py:720: UserWarning: Using the
pointplot function without specifying `hue_order` is likely to produce an incorrect plot.
warnings.warn(warning)
C:\Users\ramya\Anaconda3\lib\site-packages\scipy\stats\stats.py:1713: FutureWarning: Using a non-t
uple sequence for multidimensional indexing is deprecated; use `arr[tuple(seq)]` instead of `arr[s
eq]`. In the future this will be interpreted as an array index, `arr[np.array(seq)]`, which will r
esult either in an error or a different result.
return np.add.reduce(sorted[indexer] * weights, axis=axis) / sumval

```

Out[430]:

<seaborn.axisgrid.FacetGrid at 0x22eb9b33a20>



In [432]:

```

full_data[['Age', 'Survived']].groupby(['Age'], as_index=False).mean().sort_values(by='Survived',
ascending=True)
#a.describe()

```

Out[432]:

	Age	Survived
77	61.00	0.000000
54	40.50	0.000000
30	23.50	0.000000
26	20.50	0.000000
37	28.50	0.000000
86	74.00	0.000000
40	30.50	0.000000
60	45.50	0.000000
19	14.50	0.000000
61	46.00	0.000000
32	24.50	0.000000
14	10.00	0.000000

80	Age	Survived
64.96	0.000000	0.000000
49	36.50	0.000000
85	71.00	0.000000
46	34.50	0.000000
81	65.00	0.000000
71	55.50	0.000000
73	57.00	0.000000
84	70.50	0.000000
75	59.00	0.000000
82	66.00	0.000000
83	70.00	0.000000
62	47.00	0.111111
50	37.00	0.166667
57	43.00	0.200000
25	20.00	0.200000
27	21.00	0.208333
13	9.00	0.250000
15	11.00	0.250000
...	...	...
78	62.00	0.500000
67	52.00	0.500000
12	8.00	0.500000
65	50.00	0.500000
18	14.00	0.500000
31	24.00	0.500000
76	60.00	0.500000
48	36.00	0.500000
43	32.50	0.500000
74	58.00	0.600000
35	27.00	0.611111
47	35.00	0.611111
63	48.00	0.666667
64	49.00	0.666667
10	6.00	0.666667
8	4.00	0.700000
5	1.00	0.714286
20	15.00	0.800000
7	3.00	0.833333
79	63.00	1.000000
0	0.42	1.000000
17	13.00	1.000000
16	12.00	1.000000
9	5.00	1.000000
4	0.92	1.000000
3	0.83	1.000000
2	0.75	1.000000
1	0.67	1.000000
68	53.00	1.000000
87	80.00	1.000000

88 rows × 2 columns

In [433]:

```
full_data.isnull().sum()
```

Out[433]:

```
PassengerId      0
Survived          0
Pclass           0
Name             0
Sex              0
Age             177
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin           687
Embarked         2
dtype: int64
```

In [434]:

```
full_data.drop("Cabin", axis=1, inplace=True)
```

In [435]:

```
full_data.isnull().sum()
```

Out[435]:

```
PassengerId      0
Survived          0
Pclass           0
Name             0
Sex              0
Age             177
SibSp            0
Parch            0
Ticket           0
Fare             0
Embarked         2
dtype: int64
```

In [436]:

```
full_data.Age.fillna(full_data.Age.mean(),axis = 0,inplace = True)
```

In [437]:

```
full_data.isnull().sum()
```

Out[437]:

```
PassengerId      0
Survived          0
Pclass           0
Name             0
Sex              0
Age              0
SibSp            0
Parch            0
Ticket           0
Fare             0
Embarked         2
dtype: int64
```

In [438]:

```
full_data.Embarked.fillna(method = 'ffill',axis=0,inplace = True)
```

In [439]:

```
full_data.isnull().sum()
```

Out[439]:

```
PassengerId    0
Survived        0
Pclass          0
Name            0
Sex             0
Age            0
SibSp           0
Parch           0
Ticket          0
Fare            0
Embarked        0
dtype: int64
```

In [440]:

```
sex=pd.get_dummies(full_data["Sex"],drop_first=True)
sex.head(5)
```

Out[440]:

male	
0	1
1	0
2	0
3	0
4	1

In [441]:

```
embark=pd.get_dummies(full_data["Embarked"],drop_first=True)
embark.head(5)
```

Out[441]:

Q S		
0	0	1
1	0	0
2	0	1
3	0	1
4	0	1

In [442]:

```
full_data=pd.concat([full_data,sex,embark],axis=1)
full_data.head(5)
```

Out[442]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked	male	Q	S
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	S	1	0	1
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C	0	0	0
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	S	0	0	1



3	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked	male	Q	S
4	5	0	3	Futrelle, Mrs. Jacob Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	S	0	0	1
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	S	1	0	1

In [443]:

```
full_data.drop(['Embarked', 'Sex', 'Ticket', 'Name'], axis=1, inplace=True)
full_data.head()
```

Out[443]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare	male	Q	S
0	1	0	3	22.0	1	0	7.2500	1	0	1
1	2	1	1	38.0	1	0	71.2833	0	0	0
2	3	1	3	26.0	0	0	7.9250	0	0	1
3	4	1	1	35.0	1	0	53.1000	0	0	1
4	5	0	3	35.0	0	0	8.0500	1	0	1

In [444]:

```
X = full_data.drop("Survived", axis = 1)
y = full_data["Survived"]
```

In [445]:

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import confusion_matrix, accuracy_score, classification_report
from sklearn.preprocessing import scale
```

In [446]:

```
X.isnull().sum()
```

Out[446]:

```
PassengerId    0
Pclass          0
Age            0
SibSp          0
Parch          0
Fare           0
male           0
Q              0
S              0
dtype: int64
```

In [447]:

```
xtrain, xtest, ytrain, ytest = train_test_split(X, y, test_size = 0.2, random_state = 10)
```

In [448]:

```
modell = LogisticRegression()
modell.fit(xtrain, ytrain)
```

C:\Users\ramya\Anaconda3\lib\site-packages\sklearn\linear\_model\logistic.py:433: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify a solver to silence this warning.  
FutureWarning)

Out[448]:

```
LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
intercept_scaling=1, max_iter=100, multi_class='warn',
n_jobs=None, penalty='l2', random_state=None, solver='warn',
```

```
tol=0.0001, verbose=0, warm_start=False)
```

```
In [ ]:
```

```
In [450]:
```

```
ypred = model1.predict(xtest)
accuracy_score(ytest,ypred)
```

```
Out[450]:
```

```
0.8212290502793296
```

```
In [451]:
```

```
full_test.isnull().sum()
```

```
Out[451]:
```

```
PassengerId      0
Pclass           0
Name             0
Sex              0
Age             86
SibSp            0
Parch            0
Ticket           0
Fare             1
Cabin           327
Embarked         0
dtype: int64
```

```
In [452]:
```

```
full_test.Age.fillna(full_test.Age.mean(),axis = 0,inplace = True)
```

```
In [453]:
```

```
full_test.isnull().sum()
```

```
Out[453]:
```

```
PassengerId      0
Pclass           0
Name             0
Sex              0
Age             0
SibSp            0
Parch            0
Ticket           0
Fare             1
Cabin           327
Embarked         0
dtype: int64
```

```
In [454]:
```

```
full_test.Fare.fillna(full_test.Age.mean(),axis = 0,inplace = True)
```

```
In [455]:
```

```
full_test.isnull().sum()
```

```
Out[455]:
```

```
PassengerId      0
Pclass           0
```

Name 0  
Sex 0  
Age 0  
SibSp 0  
Parch 0  
Ticket 0  
Fare 0  
Cabin 327  
Embarked 0  
dtype: int64

In [456]:

```
full_test.drop("Cabin", axis=1, inplace=True)
```

In [457]:

```
sex=pd.get_dummies(full_test["Sex"],drop_first=True)  
sex.head(5)
```

Out[457]:

male	
0	1
1	0
2	1
3	1
4	0

In [458]:

```
embark=pd.get_dummies(full_test["Embarked"],drop_first=True)  
embark.head(5)
```

Out[458]:

Q S		
0	1	0
1	0	1
2	1	0
3	0	1
4	0	1

In [459]:

```
full_test=pd.concat([full_test,sex,embark],axis=1)  
full_test
```

Out[459]:

PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked	male	Q	S	
0	892	3	Kelly, Mr. James	male	34.50000	0	0	330911	7.8292	Q	1	1	0
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.00000	1	0	363272	7.0000	S	0	0	1
2	894	2	Myles, Mr. Thomas Francis	male	62.00000	0	0	240276	9.6875	Q	1	1	0
3	895	3	Wirz, Mr. Albert	male	27.00000	0	0	315154	8.6625	S	1	0	1
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.00000	1	1	3101298	12.2875	S	0	0	1
5	897	3	Svensson, Mr. Johan Cervin	male	14.00000	0	0	7538	9.2250	S	1	0	1
6	898	3	Green, Mrs. Mary	female	38.00000	0	0	363272	7.0000	S	0	1	0

6	898	3	Connolly, Miss. Kate	female	30.00000	0	0	330972	7.6292	Q	0	1	0
PassengerId	Pclass		Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked	male	Q	S
7	899	2	Caldwell, Mr. Albert Francis	male	26.00000	1	1	248738	29.0000	S	1	0	1
8	900	3	Abraham, Mrs. Joseph (Sophie Halaut Easu)	female	18.00000	0	0	2657	7.2292	C	0	0	0
9	901	3	Davies, Mr. John Samuel	male	21.00000	2	0	A/4 48871	24.1500	S	1	0	1
10	902	3	Ilieff, Mr. Ylio	male	30.27259	0	0	349220	7.8958	S	1	0	1
11	903	1	Jones, Mr. Charles Cresson	male	46.00000	0	0	694	26.0000	S	1	0	1
12	904	1	Snyder, Mrs. John Pillsbury (Nelle Stevenson)	female	23.00000	1	0	21228	82.2667	S	0	0	1
13	905	2	Howard, Mr. Benjamin	male	63.00000	1	0	24065	26.0000	S	1	0	1
14	906	1	Chaffee, Mrs. Herbert Fuller (Carrie Constance...	female	47.00000	1	0	W.E.P. 5734	61.1750	S	0	0	1
15	907	2	del Carlo, Mrs. Sebastiano (Argenia Genovesi)	female	24.00000	1	0	SC/PARIS 2167	27.7208	C	0	0	0
16	908	2	Keane, Mr. Daniel	male	35.00000	0	0	233734	12.3500	Q	1	1	0
17	909	3	Assaf, Mr. Gerios	male	21.00000	0	0	2692	7.2250	C	1	0	0
18	910	3	Ilmakangas, Miss. Ida Livija	female	27.00000	1	0	STON/O2. 3101270	7.9250	S	0	0	1
19	911	3	Assaf Khalil, Mrs. Mariana (Miriam")"	female	45.00000	0	0	2696	7.2250	C	0	0	0
20	912	1	Rothschild, Mr. Martin	male	55.00000	1	0	PC 17603	59.4000	C	1	0	0
21	913	3	Olsen, Master. Artur Karl	male	9.00000	0	1	C 17368	3.1708	S	1	0	1
22	914	1	Flegenheim, Mrs. Alfred (Antoinette)	female	30.27259	0	0	PC 17598	31.6833	S	0	0	1
23	915	1	Williams, Mr. Richard Norris II	male	21.00000	0	1	PC 17597	61.3792	C	1	0	0
24	916	1	Ryerson, Mrs. Arthur Larned (Emily Maria Borie)	female	48.00000	1	3	PC 17608	262.3750	C	0	0	0
25	917	3	Robins, Mr. Alexander A	male	50.00000	1	0	A/5. 3337	14.5000	S	1	0	1
26	918	1	Ostby, Miss. Helene Ragnhild	female	22.00000	0	1	113509	61.9792	C	0	0	0
27	919	3	Daher, Mr. Shedid	male	22.50000	0	0	2698	7.2250	C	1	0	0
28	920	1	Brady, Mr. John Bertram	male	41.00000	0	0	113054	30.5000	S	1	0	1
29	921	3	Samaan, Mr. Elias	male	30.27259	2	0	2662	21.6792	C	1	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
388	1280	3	Canavan, Mr. Patrick	male	21.00000	0	0	364858	7.7500	Q	1	1	0
389	1281	3	Palsson, Master. Paul Folke	male	6.00000	3	1	349909	21.0750	S	1	0	1
390	1282	1	Payne, Mr. Vivian Ponsonby	male	23.00000	0	0	12749	93.5000	S	1	0	1
391	1283	1	Lines, Mrs. Ernest H (Elizabeth Lindsey James)	female	51.00000	0	1	PC 17592	39.4000	S	0	0	1
392	1284	3	Abbott, Master. Eugene Joseph	male	13.00000	0	2	C.A. 2673	20.2500	S	1	0	1
393	1285	2	Gilbert, Mr. William	male	47.00000	0	0	C.A. 30769	10.5000	S	1	0	1
394	1286	3	Kink-Heilmann, Mr. Anton	male	29.00000	3	1	315153	22.0250	S	1	0	1
395	1287	1	Smith, Mrs. Lucien Philip (Mary Eloise Hughes)	female	18.00000	1	0	13695	60.0000	S	0	0	1
396	1288	3	Colbert, Mr. Patrick	male	24.00000	0	0	371109	7.2500	Q	1	1	0
397	1289	1	Frolicher-Stehli, Mrs. Maxmillian (Margaretha ...	female	48.00000	1	1	13567	79.2000	C	0	0	0
398	1290	3	Larsson-Rondberg, Mr. Edvard A	male	22.00000	0	0	347065	7.7750	S	1	0	1
399	1291	3	Conlon, Mr. Thomas Henry	male	31.00000	0	0	21332	7.7333	Q	1	1	0
400	1292	1	Bonnell, Miss. Caroline	female	30.00000	0	0	36928	164.8667	S	0	0	1
401	1293	2	Gale, Mr. Harry	male	38.00000	1	0	28664	21.0000	S	1	0	1
402	1294	1	Gibson, Miss. Dorothy Winifred	female	22.00000	0	1	112378	59.4000	C	0	0	0

403	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked	male	Q	S
404	1296	1	Carrau, Mr. Joseph	male	17.00000	0	0	17765	27.7208	C	1	0	0
405	1297	2	Frauenthal, Mr. Isaac Gerald	male	43.00000	1	0	SC/PARIS 2166	13.8625	C	1	0	0
406	1298	2	Nourney, Mr. Alfred (Baron von Drachstedt)"	male	20.00000	0	0	28666	10.5000	S	1	0	1
407	1299	1	Ware, Mr. William Jeffery	male	23.00000	1	0	113503	211.5000	C	1	0	0
408	1300	3	Widener, Mr. George Dunton	male	50.00000	1	1	334915	7.7208	Q	0	1	0
409	1301	3	Riordan, Miss. Johanna Hannah"	female	30.27259	0	0	SOTON/O.Q. 3101315	13.7750	S	0	0	1
410	1302	3	Peacock, Miss. Treasteall	female	3.00000	1	1	365237	7.7500	Q	0	1	0
411	1303	3	Naughton, Miss. Hannah	female	30.27259	0	0	19928	90.0000	Q	0	1	0
412	1304	1	Minahan, Mrs. William Edward (Lillian E Thorpe)	female	37.00000	1	0	347086	7.7750	S	0	0	1
413	1305	3	Henriksson, Miss. Jenny Lovisa	female	28.00000	0	0	A.5. 3236	8.0500	S	1	0	1
414	1306	3	Spector, Mr. Woolf	male	30.27259	0	0	PC 17758	108.9000	C	0	0	0
415	1307	1	Oliva y Ocana, Dona. Fermina	female	39.00000	0	0	SOTON/O.Q. 3101262	7.2500	S	1	0	1
416	1308	3	Saether, Mr. Simon Sivertsen	male	38.50000	0	0	359309	8.0500	S	1	0	1
417	1309	3	Ware, Mr. Frederick	male	30.27259	0	0	2668	22.3583	C	1	0	0
417	1309	3	Peter, Master. Michael J	male	30.27259	1	1						

418 rows × 13 columns

In [460]:

```
full_test.drop(["Ticket", "Name", "Sex", "Embarked"], axis=1, inplace=True)
full_test.head()
```

Out[460]:

	PassengerId	Pclass	Age	SibSp	Parch	Fare	male	Q	S
0	892	3	34.5	0	0	7.8292	1	1	0
1	893	3	47.0	1	0	7.0000	0	0	1
2	894	2	62.0	0	0	9.6875	1	1	0
3	895	3	27.0	0	0	8.6625	1	0	1
4	896	3	22.0	1	1	12.2875	0	0	1

In [461]:

```
full_data.isnull().sum()
```

Out[461]:

```
PassengerId    0
Survived        0
Pclass         0
Age            0
SibSp          0
Parch          0
Fare           0
male           0
Q              0
S              0
dtype: int64
```

In [462]:

```
ydata_predict = model1.predict(full_test)
ydata_predict
```

Out[462]:

```
array([0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0,
       1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1,
       1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1,
       1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1,
       1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0,
       0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0,
       1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1,
       0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1,
       1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1,
       0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1,
       1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1,
       0, 0, 0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1,
       1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1,
       0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0,
       0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1,
       0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0,
       0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0,
       1, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0,
       0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0,
       0, 0],
      dtype=int64)
```

In [463]:

```
output = full_test.iloc[:,0]
```

In [464]:

```
ydata_predict1 = pd.DataFrame({"PassengerId":full_test["PassengerId"], "Survived":ydata_predict})
ydata_predict1
```

Out[464]:

	PassengerId	Survived
0	892	0
1	893	0
2	894	0
3	895	0
4	896	1
5	897	0
6	898	1
7	899	0
8	900	1
9	901	0
10	902	0
11	903	0
12	904	1
13	905	0
14	906	1
15	907	1
16	908	0
17	909	0
18	910	1
19	911	1
20	912	0
21	913	0
22	914	1
23	915	1
24	916	1

PassengerId	Survived
25	0
26	1
27	0
28	0
29	0
...	...
388	0
389	0
390	1
391	1
392	0
393	0
394	0
395	1
396	0
397	1
398	0
399	0
400	1
401	0
402	1
403	1
404	0
405	0
406	0
407	1
408	1
409	1
410	1
411	1
412	1
413	0
414	1
415	0
416	0
417	0

418 rows × 2 columns

In [ ]: