In [1]:
```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot  as plt
import sklearn
from sklearn.preprocessing import LabelEncoder
```

In [2]:
```python
data = pd.read_csv('Titanic-Dataset.csv')
```

Out[2]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarl
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	
...	...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	

891 rows × 12 columns

In [3]:

```
data.describe()
```

Out[3]:

	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 [4]:

```
data.shape
```

Out[4]:

```
(891, 12)
```

In [5]:

```
data.isna().sum()
#finding null value
```

Out[5]:

```
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 [51]:

```
data1 = data.drop(["Age", "Cabin"], axis =1)
data1['Embarked'].fillna(data1['Embarked'].mode()[0],inplace=True)
data1
```

Out[51]:

	PassengerId	Survived	Pclass	Name	Sex	SibSp	Parch	Ticket	Fare	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	1	0	A/5 21171	7.2500	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	1	0	PC 17599	71.2833	C
2	3	1	3	Heikkinen, Miss. Laina	female	0	0	STON/O2. 3101282	7.9250	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	1	0	113803	53.1000	S
4	5	0	3	Allen, Mr. William Henry	male	0	0	373450	8.0500	S
...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	male	0	0	211536	13.0000	S
887	888	1	1	Graham, Miss. Margaret Edith	female	0	0	112053	30.0000	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	1	2	W./C. 6607	23.4500	S
889	890	1	1	Behr, Mr. Karl Howell	male	0	0	111369	30.0000	C
890	891	0	3	Dooley, Mr. Patrick	male	0	0	370376	7.7500	Q

891 rows × 10 columns

In [52]:

```
data1.isna().sum()
```

Out[52]:

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

In [53]:

```
data1.isna().sum()
```

Out[53]:

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

In [54]:
```
data1.shape
```

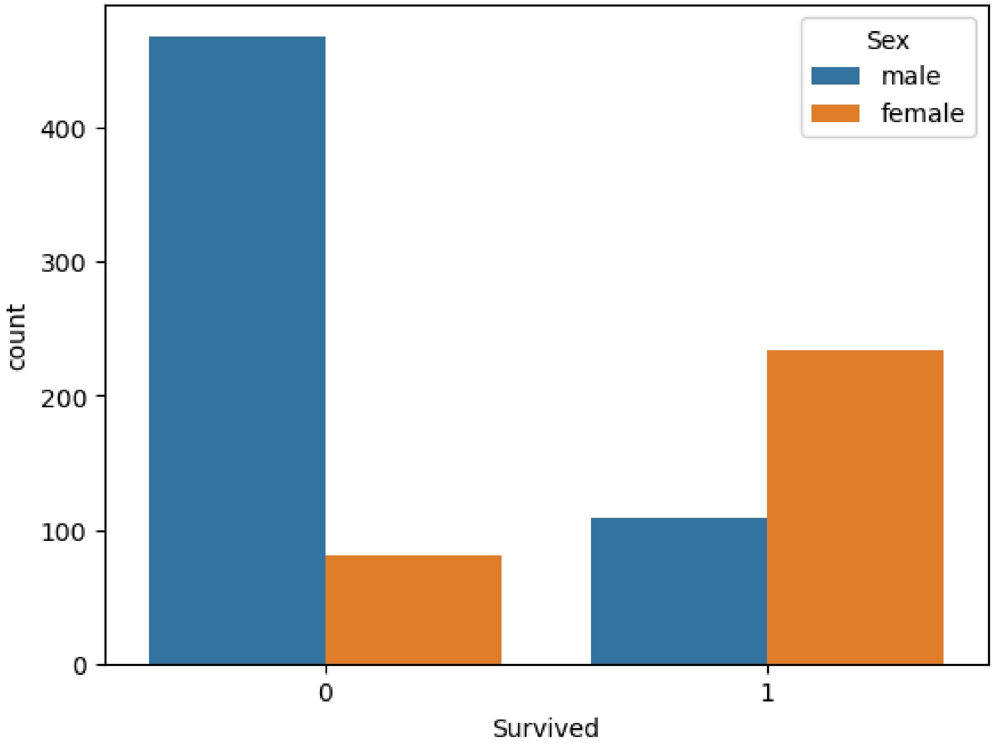
Out[54]:
```
(891, 10)
```

In [55]:
```
data1["Survived"].value_counts()
```

Out[55]:
```
Survived
0    549
1    342
Name: count, dtype: int64
```

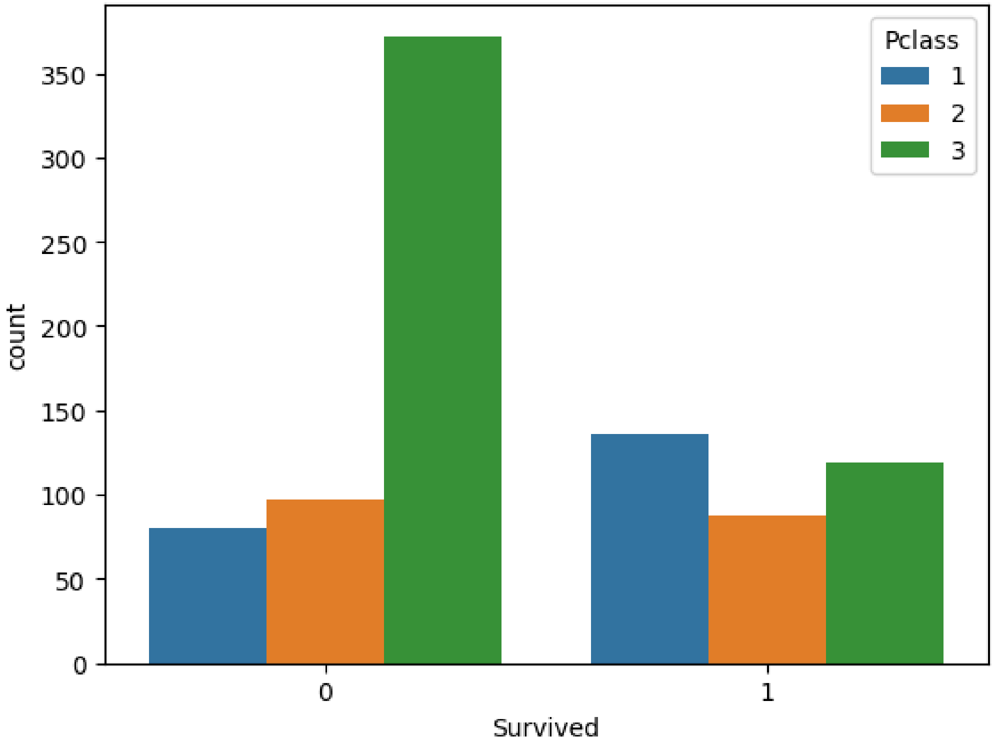
In [56]:
```
sns.countplot(x=data1["Survived"] , hue=data1["Sex"] )
```

Out[56]:
```
<Axes: xlabel='Survived', ylabel='count'>
```



In [57]:
```
sns.countplot(x=data1["Survived"] , hue=data1["Pclass"] )
```

Out[57]:
```
<Axes: xlabel='Survived', ylabel='count'>
```



In [58]:
```
data1.groupby('Sex')[('Survived')].mean()
```

Out[58]:
```
Sex
female    0.742038
male      0.188908
Name: Survived, dtype: float64
```

#Average female survival rate is more than an male.

In [59]:
```
Label = LabelEncoder()
data1['Sex'] = Label.fit_transform(data1['Sex'])
```

Out[59]:

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

In [60]:
```
X= data1[['Pclass', 'Sex']]
```

In [61]:
```
 from sklearn.model_selection import train_test_split
```

In [62]:
```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(X_train,Y_train)
```

Out[62]:
```
▾ LinearRegression

LinearRegression()
```

In [63]:
```
print(Y_test)
```

```
707    1
37     0
615    1
169    0
68     1
      ..
842    1
671    0
696    0
42     0
396    0
Name: Survived, Length: 268, dtype: int64
```

In [64]:
```
 model.predict(X_train)
```

```
       0.08051491, 0.08051491, 0.90997442, 0.08051491, 0.60842143,
       0.08051491, 0.2312914 , 0.08051491, 0.75919792, 0.60842143,
       0.60842143, 0.08051491, 0.08051491, 0.60842143, 0.75919792,
       0.38206789, 0.2312914 , 0.08051491, 0.38206789, 0.08051491,
       0.90997442, 0.08051491, 0.2312914 , 0.38206789, 0.38206789,
       0.38206789, 0.75919792, 0.60842143, 0.08051491, 0.90997442,
       0.60842143, 0.90997442, 0.08051491, 0.08051491, 0.08051491,
       0.90997442, 0.08051491, 0.08051491, 0.38206789, 0.08051491,
       0.75919792, 0.60842143, 0.38206789, 0.08051491, 0.60842143,
       0.08051491, 0.08051491, 0.90997442, 0.38206789, 0.60842143,
       0.08051491, 0.2312914 , 0.60842143, 0.08051491, 0.2312914 ,
       0.90997442, 0.08051491, 0.75919792, 0.38206789, 0.08051491,
       0.08051491, 0.90997442, 0.90997442, 0.90997442, 0.75919792,
       0.60842143, 0.2312914 , 0.38206789, 0.38206789, 0.60842143,
       0.08051491, 0.75919792, 0.38206789, 0.90997442, 0.38206789,
       0.08051491, 0.08051491, 0.38206789, 0.60842143, 0.90997442,
       0.08051491, 0.2312914 , 0.60842143, 0.08051491, 0.38206789,
       0.38206789, 0.08051491, 0.2312914 , 0.38206789, 0.08051491,
       0.08051491, 0.2312914 , 0.90997442, 0.60842143, 0.08051491,
       0.38206789, 0.38206789, 0.38206789])
```

In [32]:
```
import warnings
warnings.filterwarnings("ignore")

res = model.predict([[2, 1]])

if res== 0:
    print("Not Survived")
else:
```

```
Survived
```

In [ ]: