```
In [ ]:  TATANIC SURVIVAL PREDICTION TASK(1)
```

```
In [19]:  import pandas as pd
          import numpy as np
          import seaborn as sns
          import matplotlib.pyplot as plt
          from sklearn.model_selection import train_test_split
```

```
In [20]:  titanic_data=pd.read_csv("Titanic-Dataset.csv")
```

```
In [21]:  titanic_data.head()
```

Out[21]:

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

```
In [22]:  titanic_data.shape
```

```
Out[22]:  (891, 12)
```

```
In [23]:  titanic_data.describe
```

```
Out[23]:   <bound method NDFrame.describe of      PassengerId  Survived  Pclass  \
           0              1         0       3
           1              2         1       1
           2              3         1       3
           3              4         1       1
           4              5         0       3
           ..           ...       ...     ...
           886          887         0       2
           887          888         1       1
           888          889         0       3
           889          890         1       1
           890          891         0       3

                                                        Name     Sex   Age  SibSp  \
           0                              Braund, Mr. Owen Harris    male  22.0      1
           1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
           2                               Heikkinen, Miss. Laina  female  26.0      0
           3         Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
           4                             Allen, Mr. William Henry    male  35.0      0
           ..                                                 ...     ...   ...    ...
           886                              Montvila, Rev. Juozas    male  27.0      0
           887                       Graham, Miss. Margaret Edith  female  19.0      0
           888           Johnston, Miss. Catherine Helen "Carrie"  female   NaN      1
           889                              Behr, Mr. Karl Howell    male  26.0      0
           890                                Dooley, Mr. Patrick    male  32.0      0

                Parch           Ticket     Fare Cabin Embarked
           0        0        A/5 21171   7.2500   NaN        S
           1        0         PC 17599  71.2833   C85        C
           2        0  STON/O2. 3101282   7.9250   NaN        S
           3        0           113803  53.1000  C123        S
           4        0           373450   8.0500   NaN        S
           ..     ...              ...      ...   ...      ...
           886      0           211536  13.0000   NaN        S
           887      0           112053  30.0000   B42        S
           888      2       W./C. 6607  23.4500   NaN        S
           889      0           111369  30.0000  C148        C
           890      0           370376   7.7500   NaN        Q

           [891 rows x 12 columns]>
```
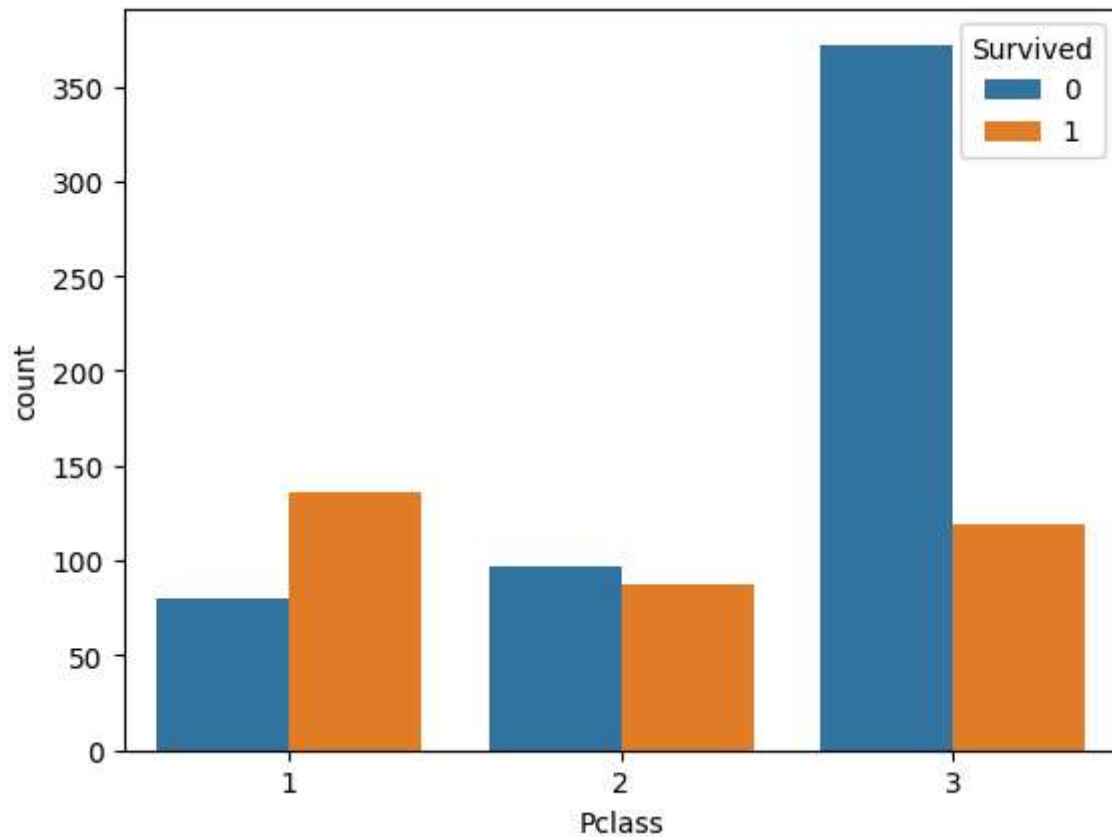
In [24]: `titanic_data['Survived'].value_counts()`

```
Out[24]:  Survived
          0    549
          1    342
          Name: count, dtype: int64
```
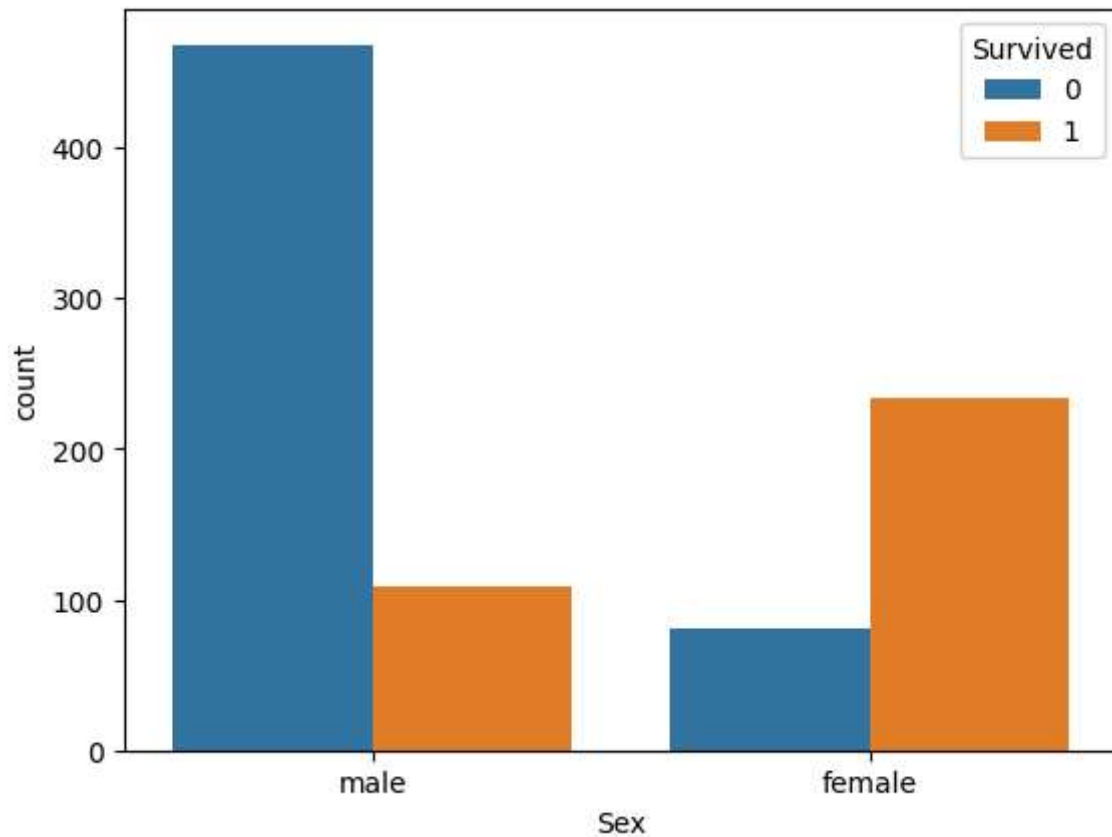
In [26]:
```
sns.countplot(x='Pclass', hue='Survived', data=titanic_data)
plt.show()
```

In [28]: `titanic_data['Sex']`

Out[28]:
```
0         male
1       female
2       female
3       female
4         male
         ...  
886       male
887     female
888     female
889       male
890       male
Name: Sex, Length: 891, dtype: object
```

In [29]:
```python
sns.countplot(x='Sex', hue='Survived', data=titanic_data)
plt.show()
```

```
In [30]: titanic_data.groupby('Sex')[['Survived']].mean()
```

Out[30]:

Sex	Survived
female	0.742038
male	0.188908

```
In [31]: titanic_data['Sex'].unique()
```

Out[31]: array(['male', 'female'], dtype=object)

```
In [32]: from sklearn.preprocessing import LabelEncoder
         LabelEncoder=LabelEncoder()
         titanic_data['Sex'] = LabelEncoder.fit_transform(titanic_data['Sex'])
         titanic_data.head()
```

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

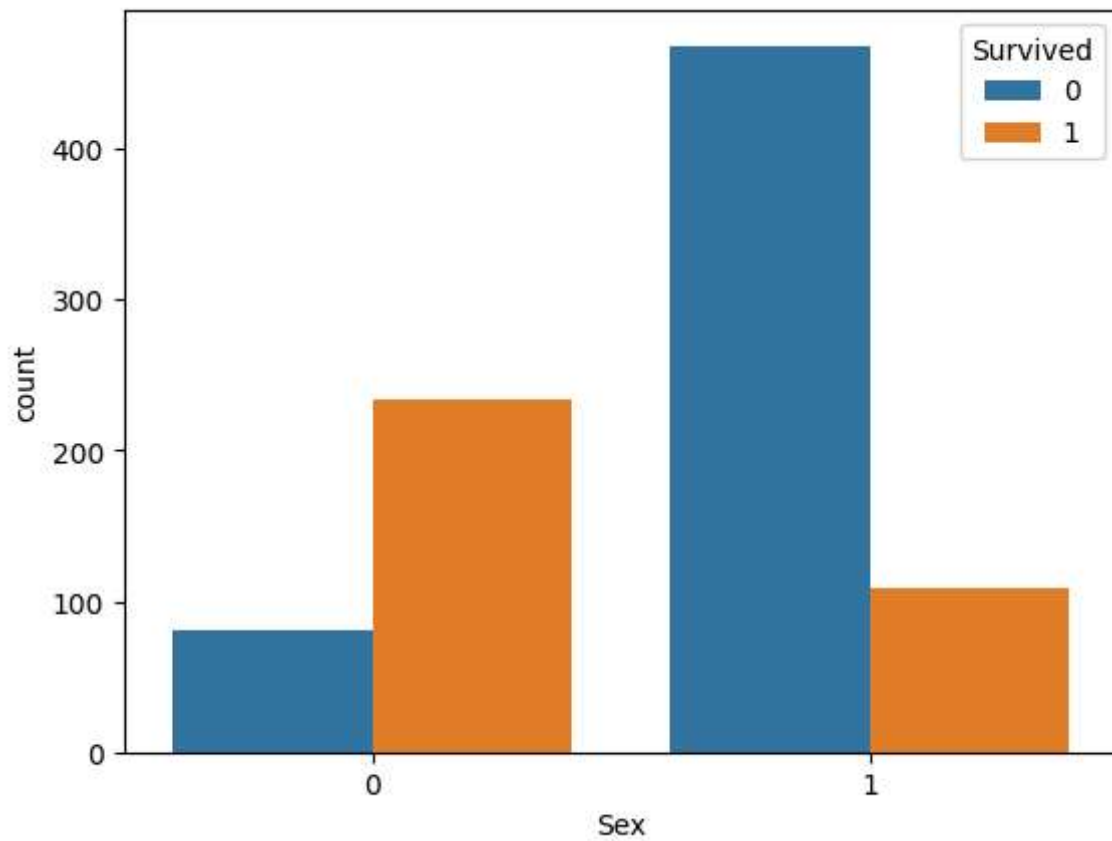In [33]: `titanic_data['Sex'],titanic_data['Survived']`

Out[33]:
```
(0      1
 1      0
 2      0
 3      0
 4      1
        ..
 886    1
 887    0
 888    0
 889    1
 890    1
 Name: Sex, Length: 891, dtype: int32,
 0      0
 1      1
 2      1
 3      1
 4      0
        ..
 886    0
 887    1
 888    0
 889    1
 890    0
 Name: Survived, Length: 891, dtype: int64)
```

```
In [34]:  sns.countplot(x=titanic_data['Sex'], hue=titanic_data['Survived'], data=titanic_dat
          plt.show()
```



```
In [35]:  titanic_data=titanic_data.drop(['Age'],axis=1)
```

```
In [36]:  titanic_data
```

Out[36]:

	PassengerId	Survived	Pclass	Name	Sex	SibSp	Parch	Ticket	Fare	Cabi
0	1	0	3	Braund, Mr. Owen Harris	1	1	0	A/5 21171	7.2500	Na
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	0	1	0	PC 17599	71.2833	C8
2	3	1	3	Heikkinen, Miss. Laina	0	0	0	STON/O2. 3101282	7.9250	Na
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	0	1	0	113803	53.1000	C12
4	5	0	3	Allen, Mr. William Henry	1	0	0	373450	8.0500	Na
...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	1	0	0	211536	13.0000	Na
887	888	1	1	Graham, Miss. Margaret Edith	0	0	0	112053	30.0000	B4
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	0	1	2	W./C. 6607	23.4500	Na
889	890	1	1	Behr, Mr. Karl Howell	1	0	0	111369	30.0000	C14
890	891	0	3	Dooley, Mr. Patrick	1	0	0	370376	7.7500	Na

891 rows × 11 columns

```
In [37]:  X=titanic_data[['Pclass','Sex']]
          Y=titanic_data[['Survived']]
```

```
In [38]:  from sklearn.model_selection import train_test_split
          X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_sta
```

```
In [39]:  from sklearn.linear_model import LogisticRegression
          log = LogisticRegression(random_state=0)
          log.fit(X_train,Y_train)
```

C:\Users\rushi\AppData\Local\Programs\Python\Python312\Lib\site-packages\sklearn\uti
ls\validation.py:1339: DataConversionWarning: A column-vector y was passed when a 1d
array was expected. Please change the shape of y to (n_samples,), for example using
ravel().
  y = column_or_1d(y, warn=True)

Out[39]:   ▼   LogisticRegression   ⓘ ⍰

          LogisticRegression(random_state=0)

```
In [40]:  pred=print(log.predict(X_test))
```

[0 0 0 1 1 0 1 1 0 1 0 1 0 1 1 1 0 0 0 0 0 1 0 0 1 1 0 1 1 1 0 1 0 0 0 0 0
 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 1 1 0 1 0 1 0 1 1 1 0 0 0
 0 1 0 0 0 0 0 1 0 0 1 1 1 1 0 0 0 0 1 1 0 1 0 0 0 0 0 0 1 1 1 1 0 1 0
 1 0 1 0 1 1 1 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 1 0 1 1 1 0 1
 1 0 0 1 1 0 1 0 1 0 1 1 0 0 1 1 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0]

```
In [41]:  print(Y_test)
```

         Survived
    495   0
    648   0
    278   0
    31   1
    255   1
    ..   ...
    780   1
    837   0
    215   1
    833   0
    372   0

[179 rows x 1 columns]

```
In [42]:  data = {
              'PassengerId': [1, 2, 3, 4, 5],
              'Survived': [1, 0, 0, 1, 1]
          }
          df = pd.DataFrame(data)

          def check_survival_status(survived):
              if survived == 1:
                  return 'Survived'
```

```
        else:
            return 'Not Survived'
```

In [44]: `titanic_data['Survival_Status'] = titanic_data['Survived'].apply(check_survival_sta`

In [45]: `titanic_data`

	PassengerId	Survived	Pclass	Name	Sex	SibSp	Parch	Ticket	Fare	Cabi
0	1	0	3	Braund, Mr. Owen Harris	1	1	0	A/5 21171	7.2500	Na
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	0	1	0	PC 17599	71.2833	C8
2	3	1	3	Heikkinen, Miss. Laina	0	0	0	STON/O2. 3101282	7.9250	Na
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	0	1	0	113803	53.1000	C12
4	5	0	3	Allen, Mr. William Henry	1	0	0	373450	8.0500	Na
...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	1	0	0	211536	13.0000	Na
887	888	1	1	Graham, Miss. Margaret Edith	0	0	0	112053	30.0000	B4
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	0	1	2	W./C. 6607	23.4500	Na
889	890	1	1	Behr, Mr. Karl Howell	1	0	0	111369	30.0000	C14
890	891	0	3	Dooley, Mr. Patrick	1	0	0	370376	7.7500	Na

891 rows × 12 columns

```
In [46]:  titanic_data.isnull().sum()
```

```
Out[46]:  PassengerId          0
          Survived             0
          Pclass               0
          Name                 0
          Sex                  0
          SibSp                0
          Parch                0
          Ticket               0
          Fare                 0
          Cabin              687
          Embarked             2
          Survival_Status      0
          dtype: int64
```

```
In [ ]:
```