

```
In [56]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
In [57]: titanic=pd.read_csv('E://TASK 1/train.csv')
titanic.head()
```

Out[57]:

	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 [12]: titanic.shape
```

Out[12]: (891, 12)

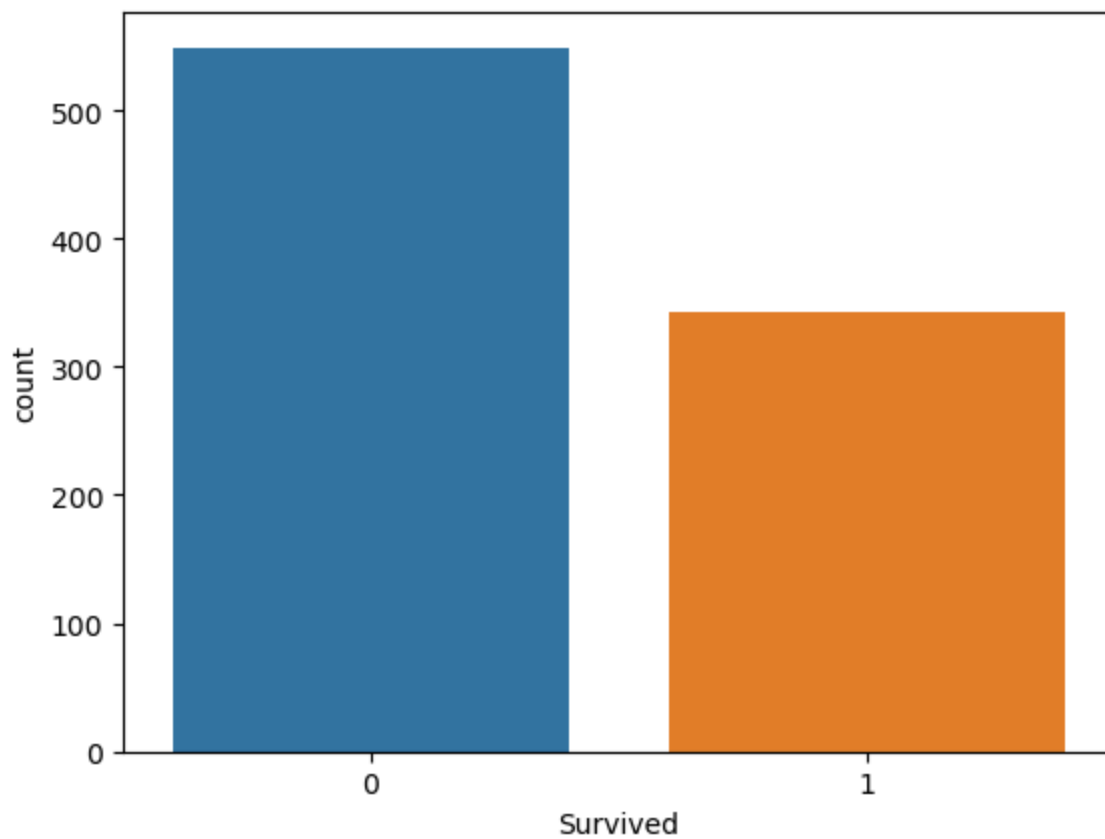
```
In [27]: titanic.describe()
```

Out[27]:

	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 [15]: import seaborn as sns
sns.countplot(x='Survived',data=titanic)
```

Out[15]: <Axes: xlabel='Survived', ylabel='count'>

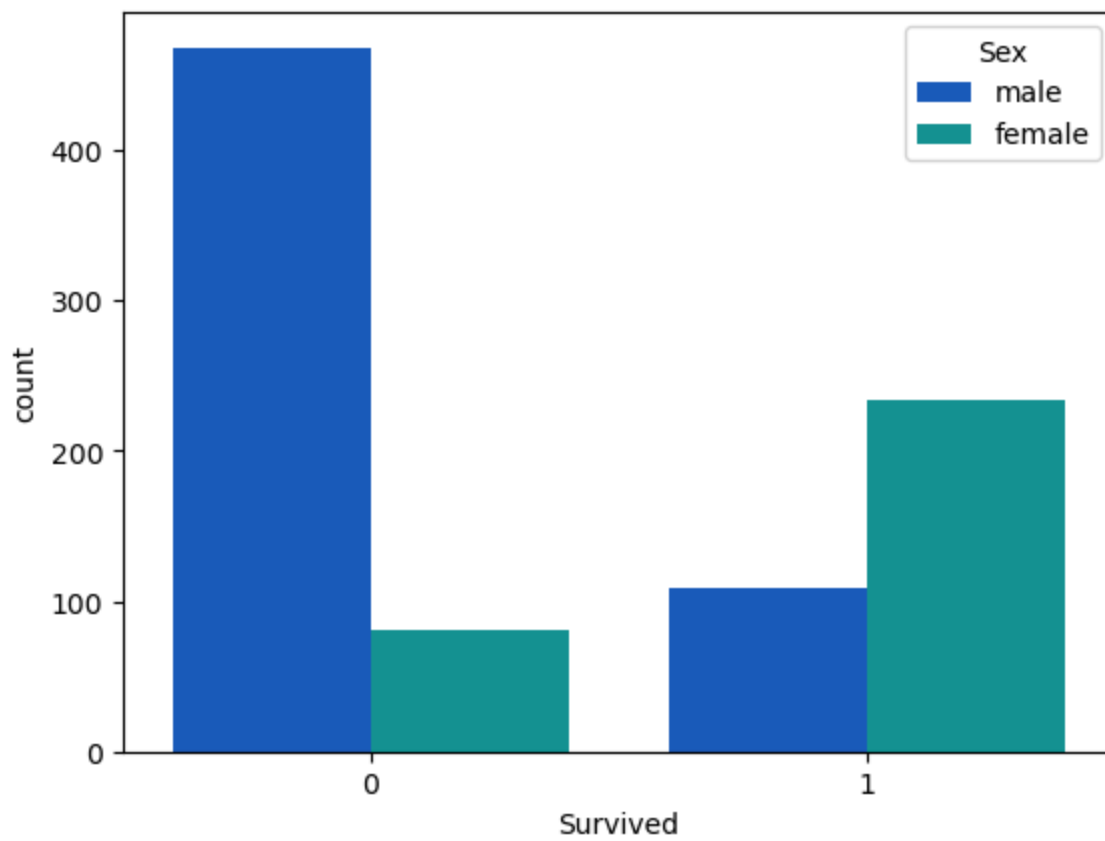


```
In [21]: titanic['Survived'].value_counts()
```

```
Out[21]: Survived  
0      549  
1      342  
Name: count, dtype: int64
```

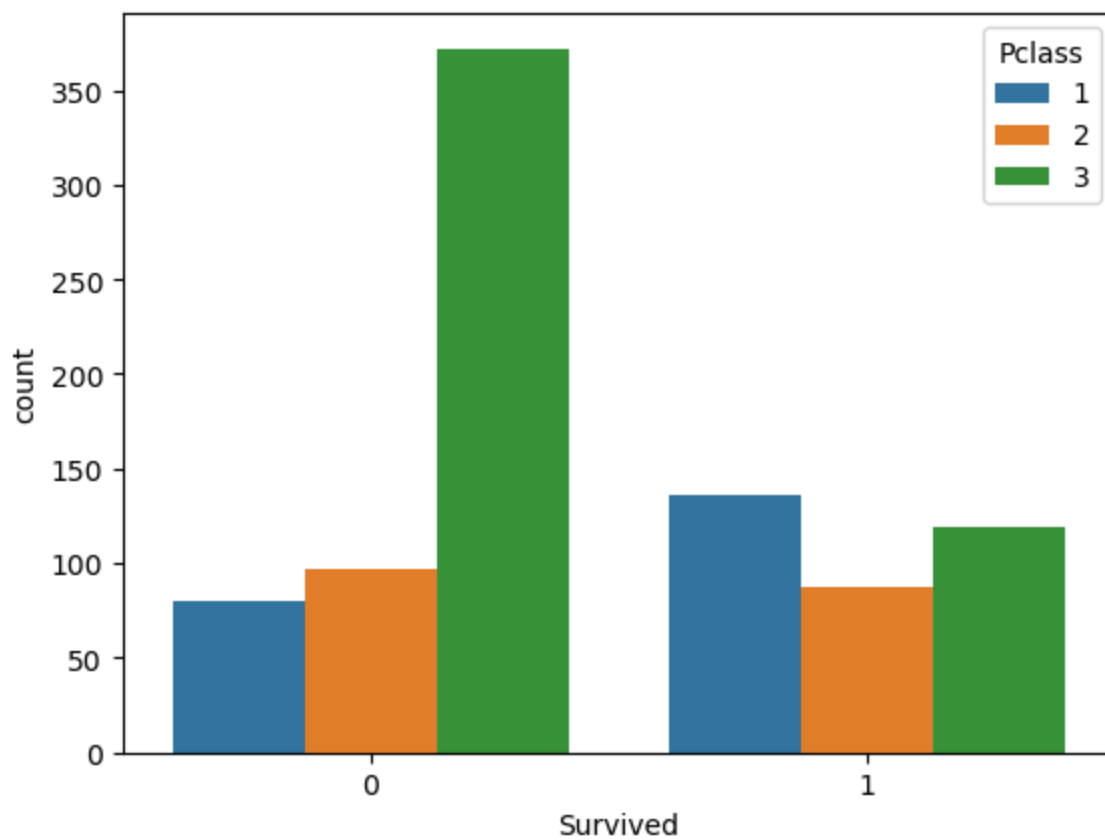
```
In [19]: sns.countplot(x='Survived', hue='Sex', data=titanic, palette='winter')
```

```
Out[19]: <Axes: xlabel='Survived', ylabel='count'>
```



```
In [22]: sns.countplot(x=titanic['Survived'],hue=titanic['Pclass'])
```

```
Out[22]: <Axes: xlabel='Survived', ylabel='count'>
```

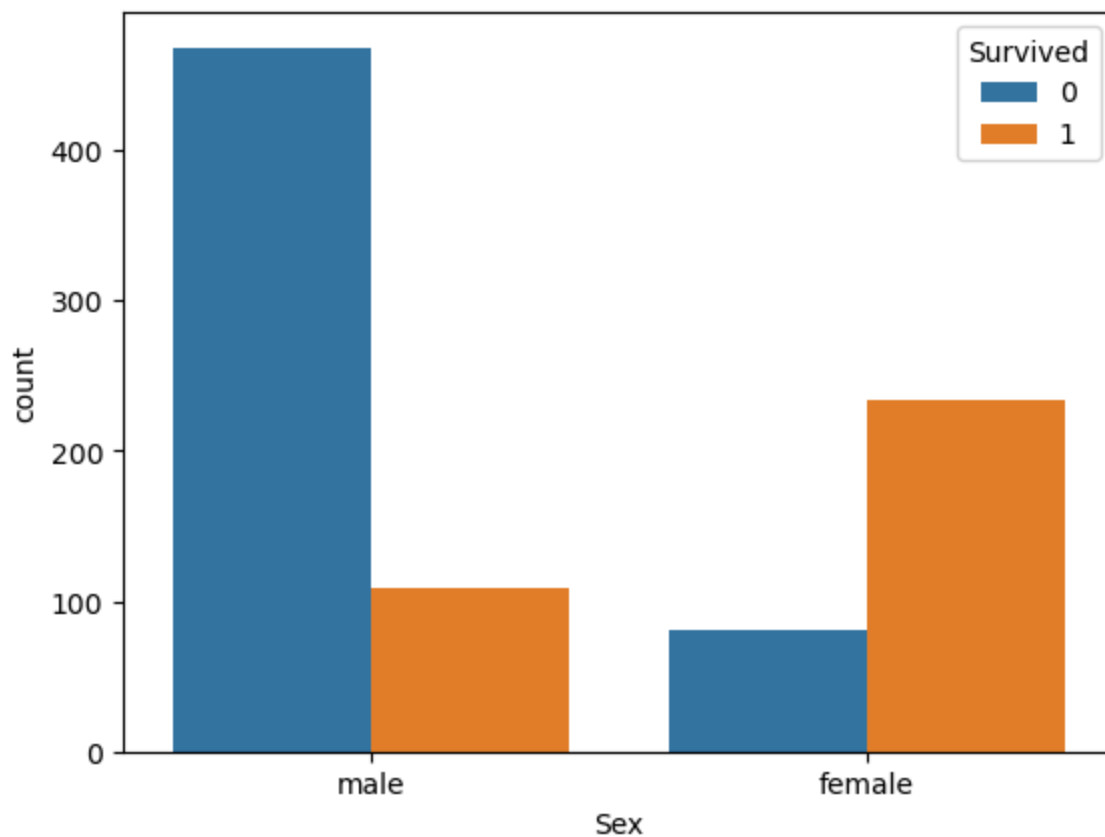


```
In [23]: titanic['Sex']
```

```
Out[23]: 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 [24]: sns.countplot(x=titanic['Sex'], hue=titanic['Survived'])
```

```
Out[24]: <Axes: xlabel='Sex', ylabel='count'>
```



```
In [28]: titanic.groupby('Sex')[['Survived']].mean()
```

```
Out[28]:      Survived
Sex
female  0.742038
male    0.188908
```

```
In [31]: from sklearn.preprocessing import LabelEncoder
          labelencoder = LabelEncoder()

          titanic['Sex'] = labelencoder.fit_transform(titanic['Sex'])
          titanic.head(10)
```

Out [31]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	1	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	0	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	0	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	0	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	1	35.0	0	0	373450	8.0500	NaN	S
5	6	0	3	Moran, Mr. James	1	NaN	0	0	330877	8.4583	NaN	Q
6	7	0	1	McCarthy, Mr. Timothy J	1	54.0	0	0	17463	51.8625	E46	S
7	8	0	3	Palsson, Master. Gosta Leonard	1	2.0	3	1	349909	21.0750	NaN	S
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	0	27.0	0	2	347742	11.1333	NaN	S
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	0	14.0	1	0	237736	30.0708	NaN	C

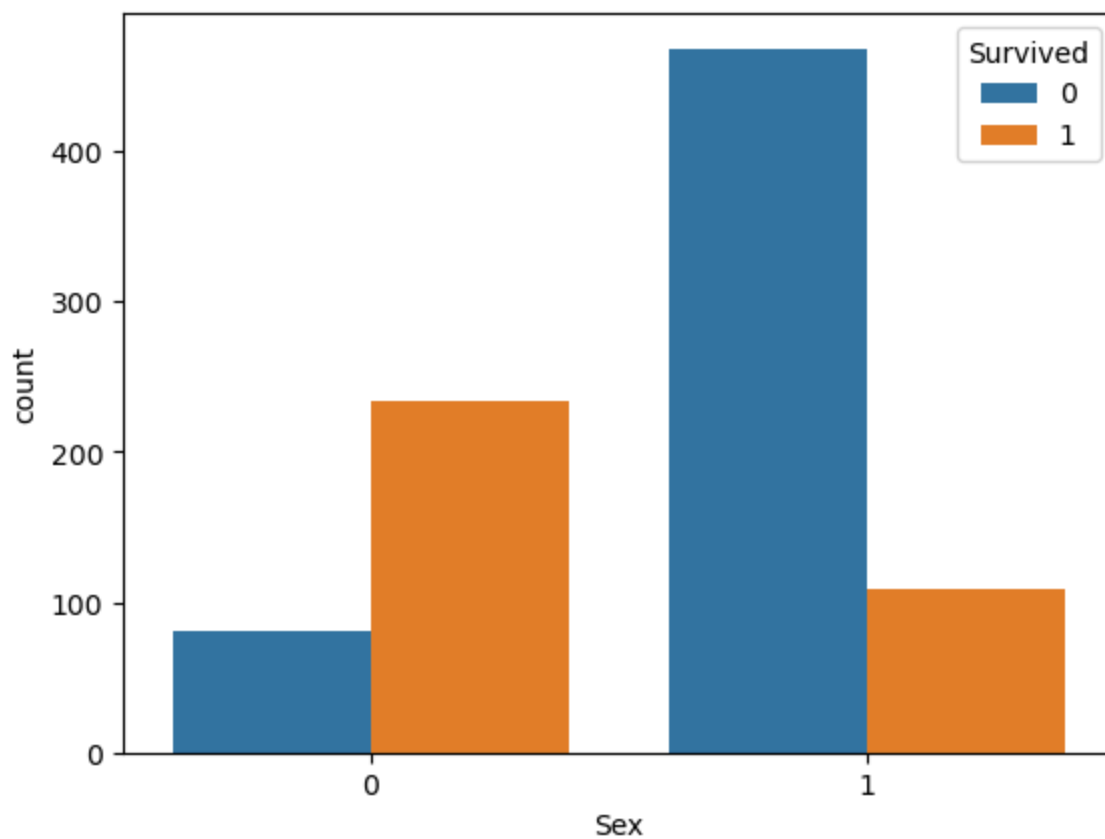
In [35]:

titanic['Sex'],titanic['Survived']

```
Out[35]: (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: int64,
          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 [36]: sns.countplot(x=titanic['Sex'],hue=titanic['Survived'])
```

```
Out[36]: <Axes: xlabel='Sex', ylabel='count'>
```



```
In [38]: titanic.isna().sum()
```

```
Out[38]: 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 [40]: titanic=titanic.drop(['Age'],axis=1)
titanic_final =titanic
titanic_final.head(10)
```

```
Out[40]:
```

	PassengerId	Survived	Pclass	Name	Sex	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	1	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	1	0	0	373450	8.0500	NaN	S
5	6	0	3	Moran, Mr. James	1	0	0	330877	8.4583	NaN	Q
6	7	0	1	McCarthy, Mr. Timothy J	1	0	0	17463	51.8625	E46	S
7	8	0	3	Palsson, Master. Gosta Leonard	1	3	1	349909	21.0750	NaN	S
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	0	0	2	347742	11.1333	NaN	S
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	0	1	0	237736	30.0708	NaN	C

```
In [48]: x=titanic[['Pclass','Sex']]
y=titanic['Survived']
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_state=0)
from sklearn.linear_model import LogisticRegression

log= LogisticRegression(random_state=0)
log.fit(x_train,y_train)
```

```
Out[48]: ▼ LogisticRegression
LogisticRegression(random_state=0)
```

```
In [49]: 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 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 0 1 0 0 1 1 1 1 0 0 0 0 1 1 0 1 0 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 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 1 0 0 1 1 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0]
```

```
In [50]: print(x_test)

      Pclass  Sex
495         3    1
648         3    1
278         3    1
31          1    0
255         3    0
..         ...  ...
780         3    0
837         3    1
215         1    0
833         3    1
372         3    1

[179 rows x 2 columns]
```

```
In [51]: print(y_test)

495    0
648    0
278    0
31     1
255    1
..
780    1
837    0
215    1
833    0
372    0
Name: Survived, Length: 179, dtype: int64
```

```
In [54]: import warnings
warnings.filterwarnings("ignore")
res=log.predict([[2,0]])
if(res==0):
    print("Not Survived!")
else:
    print("Survived")

Survived
```

```
In [55]: import warnings
warnings.filterwarnings("ignore")
res=log.predict([[2,1]])
if(res==0):
    print("Not Survived!")
else:
    print("Survived")

Not Survived!
```