In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [9]:

```
Titanic = pd.read_csv("E:/Data Sets/Titanic-Dataset.csv")
print(Titanic)
```

```
     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 [10]:
```
Titanic.head()
```

Out[10]:

	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 [11]:
```
Titanic.shape
```

Out[11]:
```
(891, 12)
```

In [12]:
```
Titanic.tail()
```

Out[12]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.00	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.00	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.45	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.00	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.75	NaN	Q

In [13]:

```
Titanic['Age']
```

Out[13]:

```
0      22.0
1      38.0
2      26.0
3      35.0
4      35.0
       ... 
886    27.0
887    19.0
888     NaN
889    26.0
890    32.0
Name: Age, Length: 891, dtype: float64
```

In [14]:

```
Titanic["Survived"].value_counts()
```

Out[14]:

```
Survived
0    549
1    342
Name: count, dtype: int64
```

In [15]:

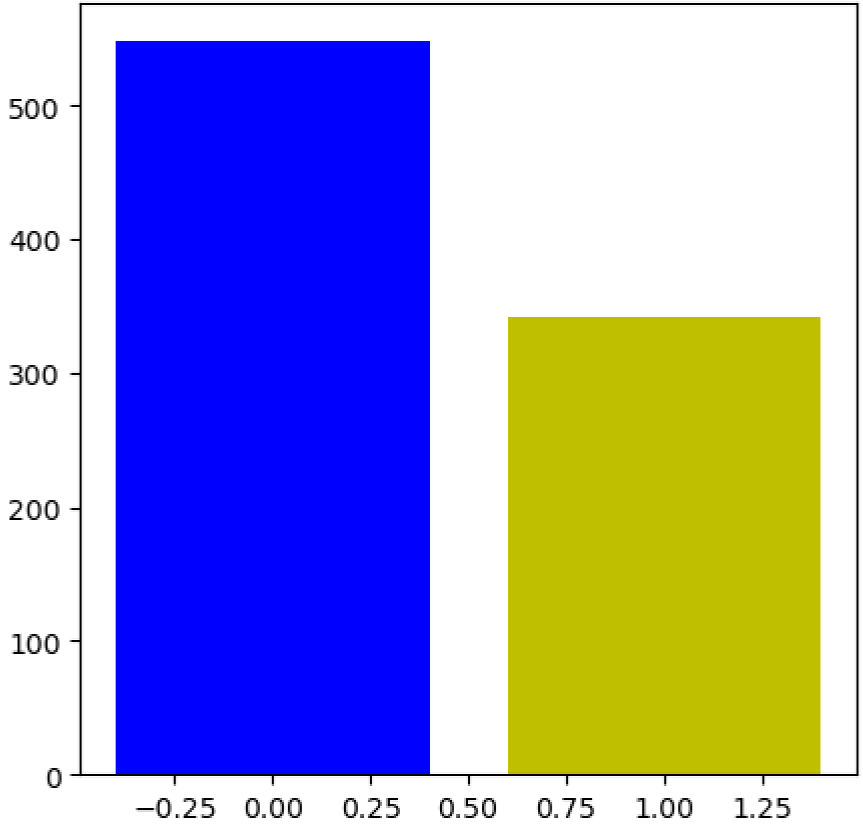
```
Titanic["Survived"].value_counts().keys()
```

Out[15]:

```
Index([0, 1], dtype='int64', name='Survived')
```

In [16]:

```
plt.figure(figsize=(5,5))
plt.bar(list(Titanic["Survived"].value_counts().keys()), list(Titanic["Survived"].value_counts()), color = ["b","y
plt.show()
```



In [17]:

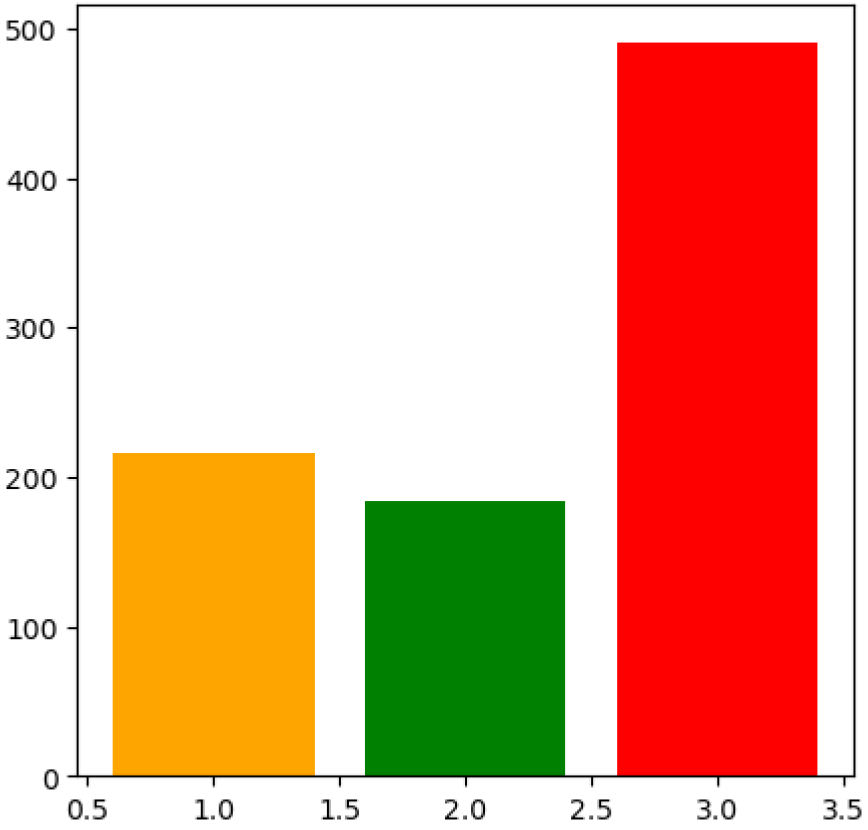
```
Titanic['Pclass'].value_counts()
```

Out[17]:

```
Pclass
3    491
1    216
2    184
Name: count, dtype: int64
```

In [18]:

```
plt.figure(figsize=(5,5))
plt.bar(list(Titanic["Pclass"].value_counts().keys()), list(Titanic["Pclass"].value_counts()), color = ["red","ora
plt.show()
```



In [19]:

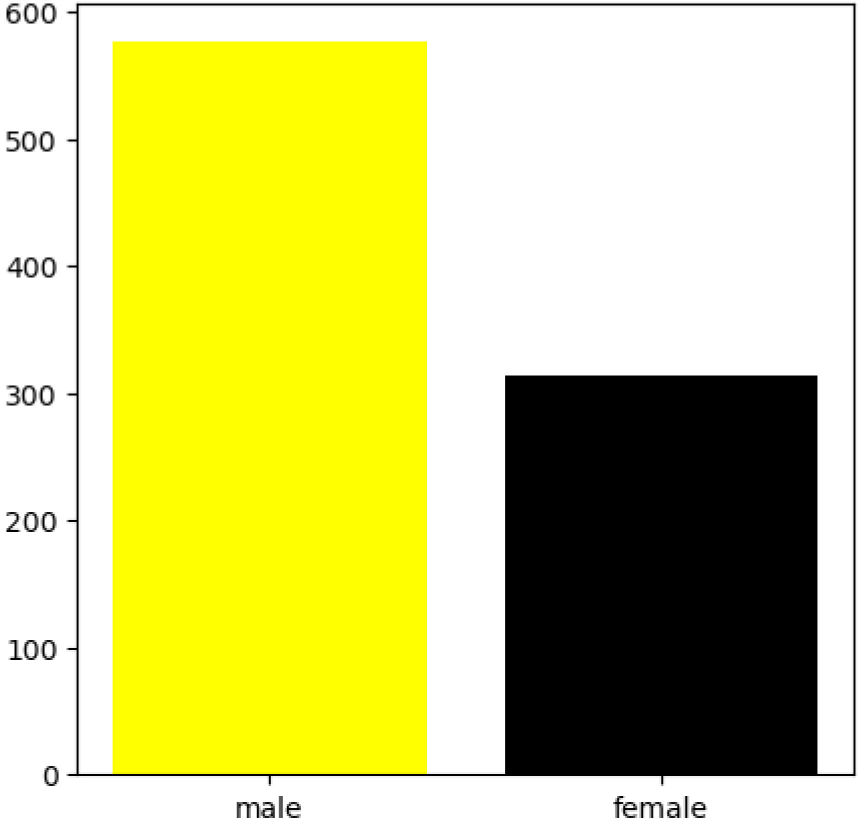
```
Titanic["Sex"].value_counts()
```

Out[19]:

```
Sex
male      577
female    314
Name: count, dtype: int64
```

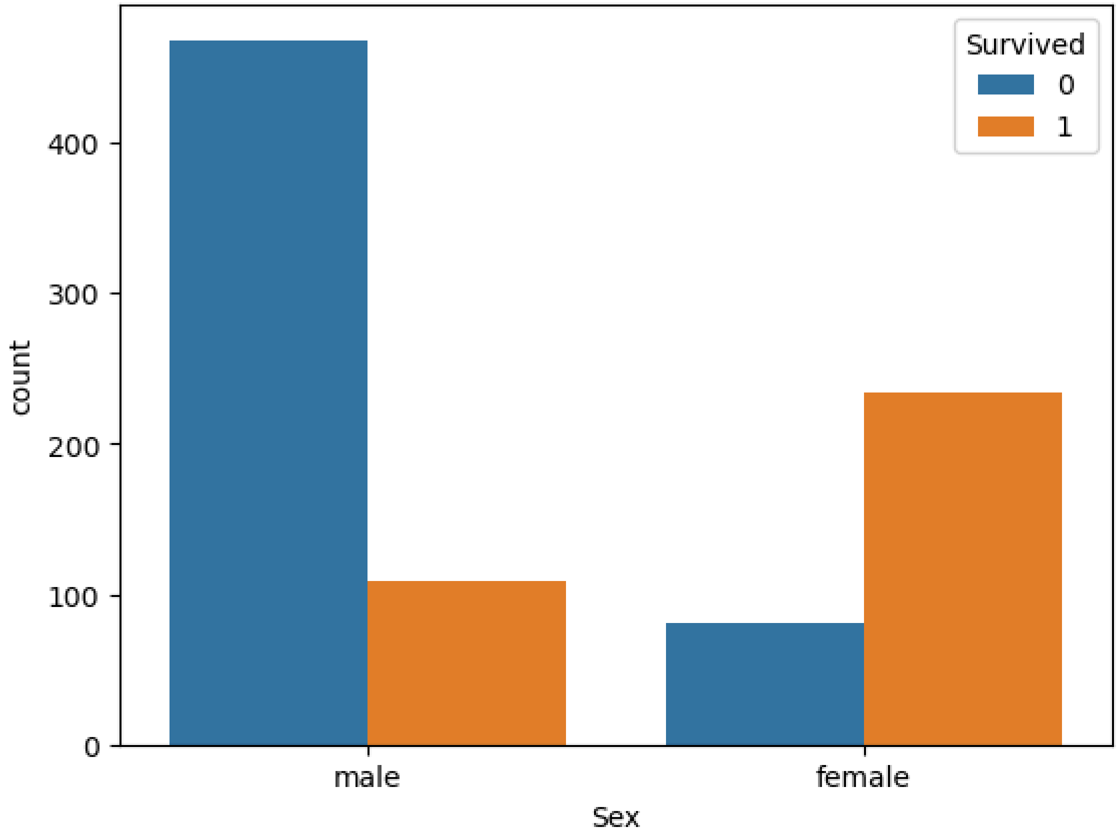
In [20]:

```
plt.figure(figsize=(5,5))
plt.bar(list(Titanic["Sex"].value_counts().keys()), list(Titanic["Sex"].value_counts()), color = ["yellow","black"
plt.show()
```



In [21]:
```
sns.countplot(x = Titanic["Sex"], hue = Titanic["Survived"])
```

Out[21]:
```
<Axes: xlabel='Sex', ylabel='count'>
```

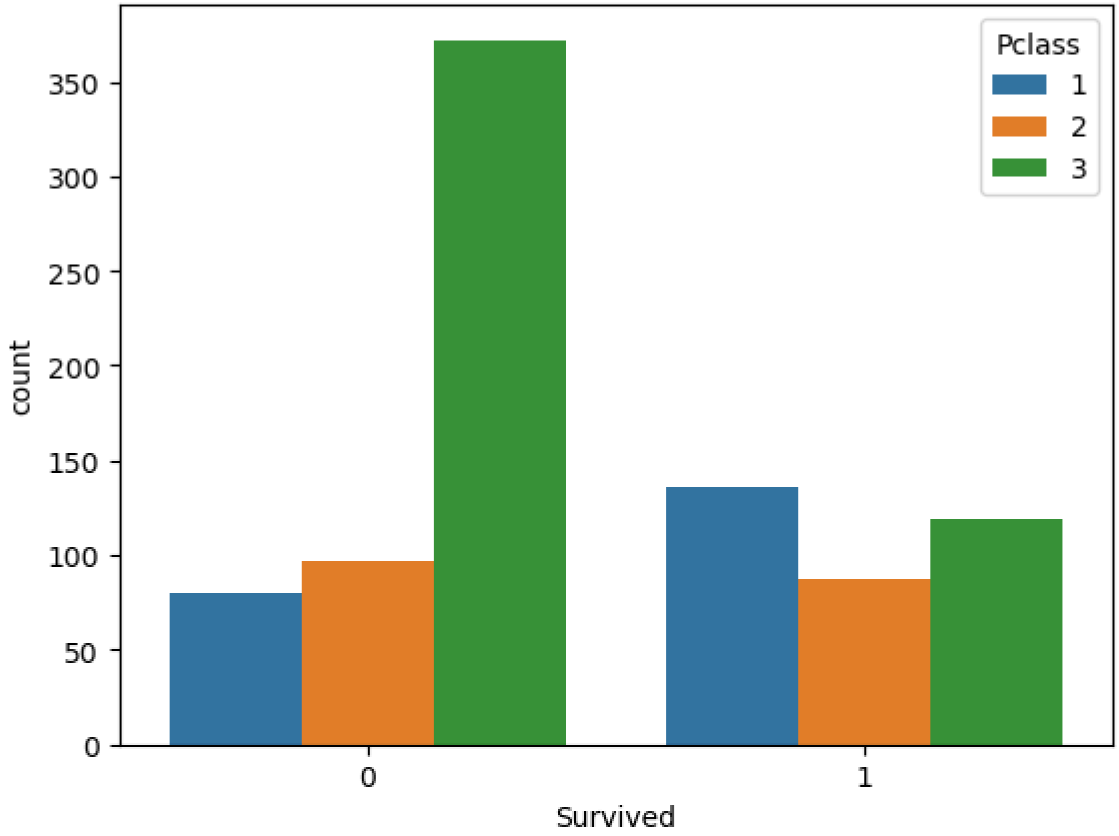


In [22]:

```
sns.countplot(x = Titanic["Survived"], hue = Titanic["Pclass"])
```

Out[22]:

```
<Axes: xlabel='Survived', ylabel='count'>
```



In [23]:

```
Titanic.isna().sum()
```

Out[23]:

```
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 [24]:

```
Titanic.replace({"Sex":{'male': 0, 'female': 1} }, inplace = True)
Titanic.head()
```

Out[24]:

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

In [25]:

```
x = Titanic[["Pclass", "Sex"]]
y = Titanic["Survived"]
```

In [26]:
```
x, y
```

Out[26]:
```
(     Pclass  Sex
 0         3    0
 1         1    1
 2         3    1
 3         1    1
 4         3    0
 ..      ...  ...
 886       2    0
 887       1    1
 888       3    1
 889       1    0
 890       3    0

 [891 rows x 2 columns],
 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 [27]:
```
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)
```

In [28]:
```
print (x.shape, x_train.shape, x_test.shape)
```

```
(891, 2) (712, 2) (179, 2)
```

In [29]:
```
print (y.shape, y_train.shape, y_test.shape)
```

```
(891,) (712,) (179,)
```

In [30]:
```
from sklearn.linear_model import LogisticRegression
ML = LogisticRegression()
ML.fit(x_train, y_train)
```

Out[30]:
```
▾ LogisticRegression
LogisticRegression()
```

In [31]:
```
x_test_prediction = ML.predict(x_test)
print(x_test_prediction)
```

```
[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 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 [32]:
```
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 [33]:

```
print(x_test)
```

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

[179 rows x 2 columns]
```

In [36]:

```
print(y_train)
```

```
140    0
439    0
817    0
378    0
491    0
      ..
835    1
192    1
629    0
559    1
684    0
Name: Survived, Length: 712, dtype: int64
```

In [37]:

```
print(x_train)
```

```
     Pclass  Sex
140       3    1
439       2    0
817       2    0
378       3    0
491       3    0
..      ...  ...
835       1    1
192       3    1
629       3    0
559       3    1
684       2    0

[712 rows x 2 columns]
```

In [41]:

```
print('x_train_prediction')
```

```
x_train_prediction
```

In [ ]:

```
from sklearn.metrics import accuracy_score
train_accuracy = accuracy_score(y_train_prediction, x_train)
train_accuracy = accuracy_score(y_test_prediction, x_test)
print("Accuracy scores of training and test data are", train_accuracy, "and", test_accuracy, "respectively")
```

In [ ]:

```
Accuracy scores of training and test data are 0.7865168539325843 and 0.7877094972067039  respectively
```