In [8]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
import math
titanic=pd.read_csv("Titanic")
titanic.head(10)
```

Out[8]:

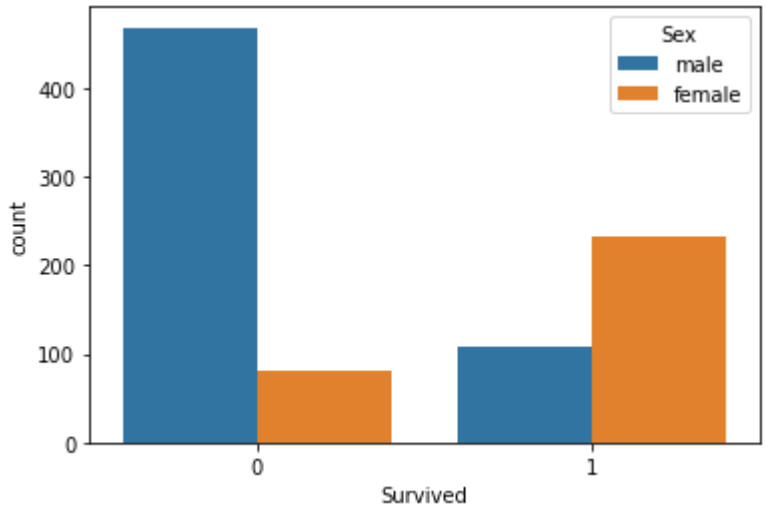
	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	N
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	N
3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

In [11]:
```
print("number of passsagers:"+str(len(titanic.index)))
```

```
number of passsagers:891
```

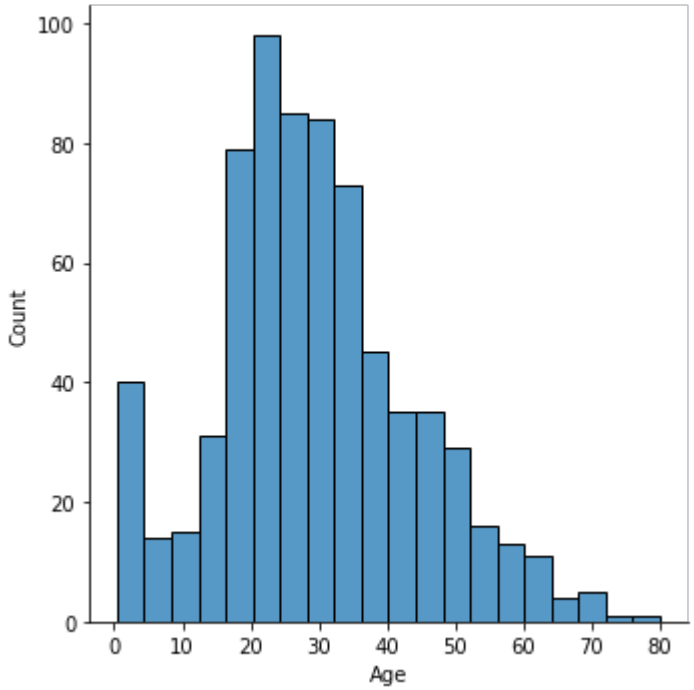
In [14]:
```
sns.countplot(x="Survived",hue="Sex",data=titanic)
```

Out[14]:
```
<AxesSubplot:xlabel='Survived', ylabel='count'>
```



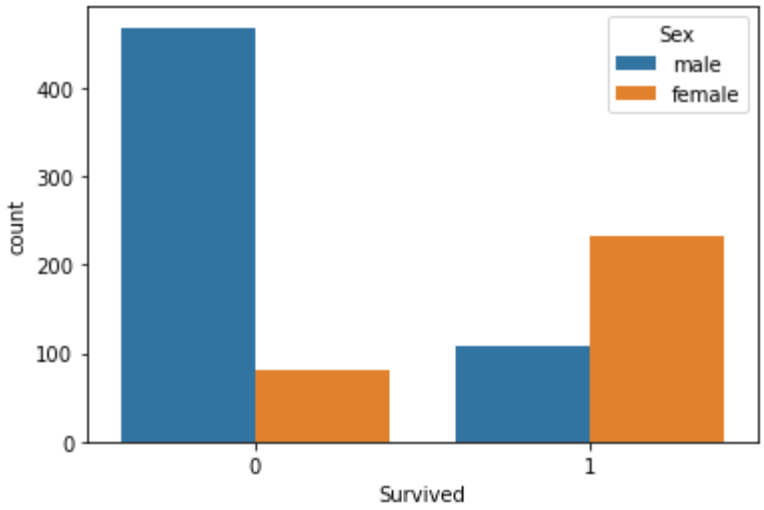
In [26]:
```
sns.displot(x="Age",data=titanic)
```

Out[26]:
```
<seaborn.axisgrid.FacetGrid at 0x1d1bf62bf70>
```



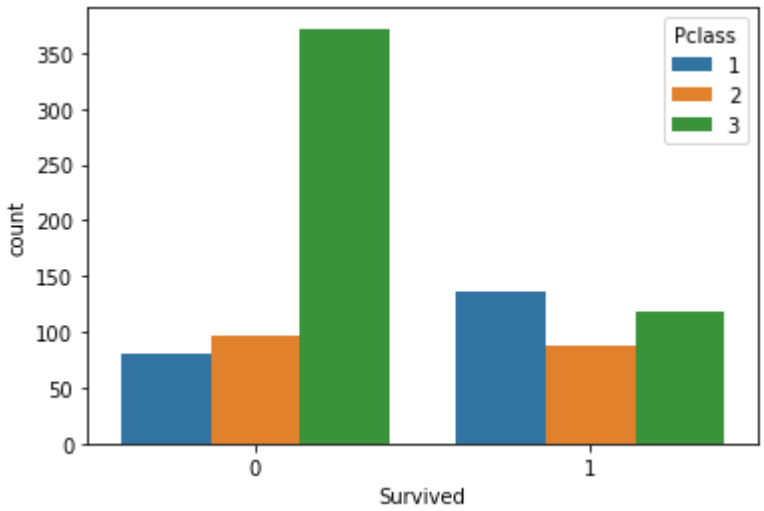
In [27]:
```
sns.countplot(x="Survived",hue="Sex",data=titanic)
```

Out[27]:
```
<AxesSubplot:xlabel='Survived', ylabel='count'>
```



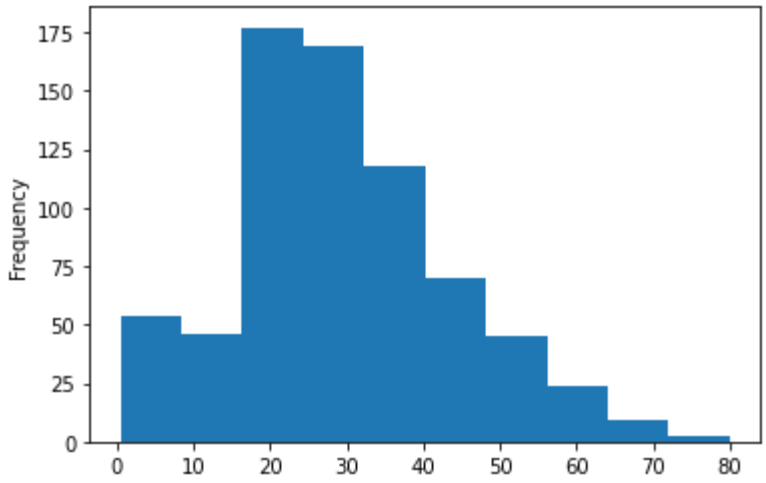
In [31]:
```
sns.countplot(x="Survived",hue="Pclass",data=titanic)
```

Out[31]:
```
<AxesSubplot:xlabel='Survived', ylabel='count'>
```



In [35]:
```
titanic["Age"].plot.hist()
```

Out[35]:
```
<AxesSubplot:ylabel='Frequency'>
```



In [36]:
```
titanic.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   PassengerId  891 non-null    int64  
 1   Survived     891 non-null    int64  
 2   Pclass       891 non-null    int64  
 3   Name         891 non-null    object 
 4   Sex          891 non-null    object 
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64  
 7   Parch        891 non-null    int64  
 8   Ticket       891 non-null    object 
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object 
 11  Embarked     889 non-null    object 
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

In [41]:
```
titanic.isnull()
```

Out[41]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	En
0	False	False	False	False	False	False	False	False	False	False	True	
1	False	False	False	False	False	False	False	False	False	False	False	
2	False	False	False	False	False	False	False	False	False	False	True	
3	False	False	False	False	False	False	False	False	False	False	False	
4	False	False	False	False	False	False	False	False	False	False	True	
...	...	...	...	...	...	...	...	...	...	...	...	
886	False	False	False	False	False	False	False	False	False	False	True	
887	False	False	False	False	False	False	False	False	False	False	False	
888	False	False	False	False	False	True	False	False	False	False	True	
889	False	False	False	False	False	False	False	False	False	False	False	
890	False	False	False	False	False	False	False	False	False	False	True	

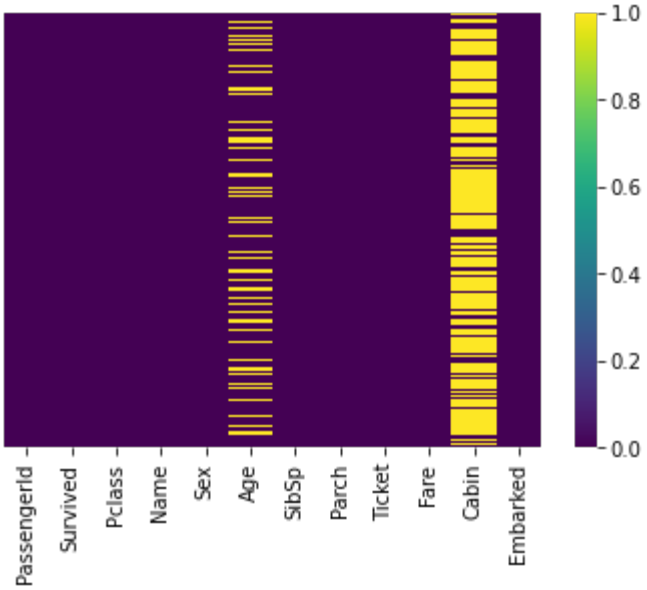
891 rows × 12 columns

In [42]:
```
titanic.isnull().sum()
```

Out[42]:
```
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
```

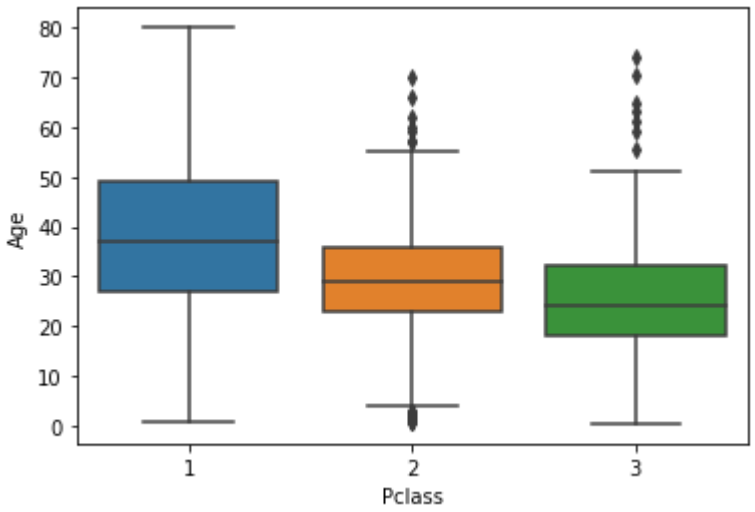
In [48]:
```
sns.heatmap(titanic.isnull(),yticklabels=False,cmap="viridis")
```

Out[48]: <AxesSubplot:>



In [53]:
```
sns.boxplot(x="Pclass",y="Age",data=titanic)
```

Out[53]: <AxesSubplot:xlabel='Pclass', ylabel='Age'>



In [55]:
```
titanic.head(5)
```

Out[55]:

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

In [56]:
```
titanic.drop("Cabin",axis=1,inplace=True)
```

In [57]:
```
titanic.head(5)
```

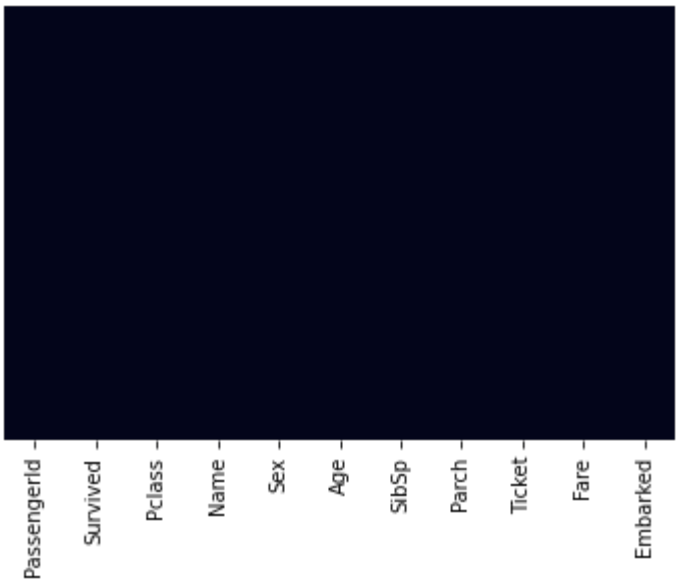
Out[57]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	En
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 [58]:
```
titanic.dropna(inplace=True)
```

In [61]:
```
sns.heatmap(titanic.isnull(),yticklabels=False,cbar=False)
#sns.heatmap(titanic.isnull(),yticklabels=False,cmap="viridis")
```

Out[61]: <AxesSubplot:>



In [62]:
```
print("number of passsagers:"+str(len(titanic.index)))
```

```
number of passsagers:712
```

In [67]:
```
sex=pd.get_dummies(titanic["Sex"],drop_first=True)
sex.head(5)
```

Out[67]:

	male
0	1
1	0
2	0
3	0
4	1

In [71]:
```
embarked=pd.get_dummies(titanic["Embarked"],drop_first=True)
embarked.head(5)
```

Out[71]:

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

In [72]:
```
pclass=pd.get_dummies(titanic["Pclass"],drop_first=True)
pclass.head(5)
```

Out[72]:

	2	3
0	0	1
1	0	0
2	0	1
3	0	0
4	0	1

In [73]:
```
titanic=pd.concat([titanic,sex,embarked,pclass],axis=1)
```

In [74]:
```
titanic.head(10)
```

Out[74]:

	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
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.7000

In [75]:
```
titanic.drop(["Sex","Embarked","PassengerId","Name","Ticket"],axis=1,inplace=T
```

In [78]:
```
titanic.head()
```

Out[78]:

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

In [ ]:
```

```

In [77]:
```
titanic.drop(["Pclass"],axis=1,inplace=True)
```

In [79]:
```
x=titanic.drop("Survived",axis=1)
y=titanic["Survived"]
```

In [81]:
```
from sklearn.model_selection import train_test_split
```

In [83]:
```
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.25, rand
```

In [85]:
```
from sklearn.linear_model import LogisticRegression
```

In [97]:
```
log_model = LogisticRegression(solver='lbfgs', max_iter=1000)
```

In [98]:
```
log_model.fit(x_train.values,y_train.values)
```

Out[98]:
```
LogisticRegression(max_iter=1000)
```

In [102]:
```
y_pred=log_model.predict(x_test.values)
```

In [103]:
```
from sklearn.metrics import classification_report
```

In [104]:
```
classification_report(y_test,y_pred)
```

Out[104]:
```
'              precision    recall  f1-score   support\n\n           0
0.81      0.83      0.82       102\n           1       0.77      0.74      0.
75        76\n\n    accuracy                           0.79       178\n   mac
ro avg       0.79      0.79      0.79       178\nweighted avg       0.79
0.79      0.79       178\n'
```

In [105]:
```
from sklearn.metrics import confusion_matrix
```

In [106]:
```
confusion_matrix(y_test,y_pred)
```

Out[106]:
```
array([[85, 17],
       [20, 56]], dtype=int64)
```

In [107]:
```
from sklearn .metrics import accuracy_score
```

In [109]:
```
accuracy_score(y_test,y_pred)*100
```

Out[109]:
```
79.21348314606742
```

In [ ]:
```

```