In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
data = pd.read_csv('https://raw.githubusercontent.com/dphi-official/Datasets/master
data
```

Out[2]:

	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
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	N
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	N
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	N

891 rows × 12 columns

In [3]:
```
data.shape
```

Out[3]:
```
(891, 12)
```

In [4]:
```
data.describe()
```

Out[4]:

	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 [5]:
```
data.describe(include = 'object')
```

Out[5]:

	Name	Sex	Ticket	Cabin	Embarked
count	891	891	891	204	889
unique	891	2	681	147	3
top	Braund, Mr. Owen Harris	male	347082	B96 B98	S
freq	1	577	7	4	644

In [6]:
```
data.isnull().sum()
```

Out[6]:
```
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 [7]:
```
data['Age'] = data['Age'].fillna(np.mean(data['Age']))
```

In [8]:
```
data['Cabin'] = data['Cabin'].fillna(data['Cabin'].mode()[0])
```

In [9]:
```
data['Embarked'] = data['Embarked'].fillna(data['Embarked'].mode()[0])
```

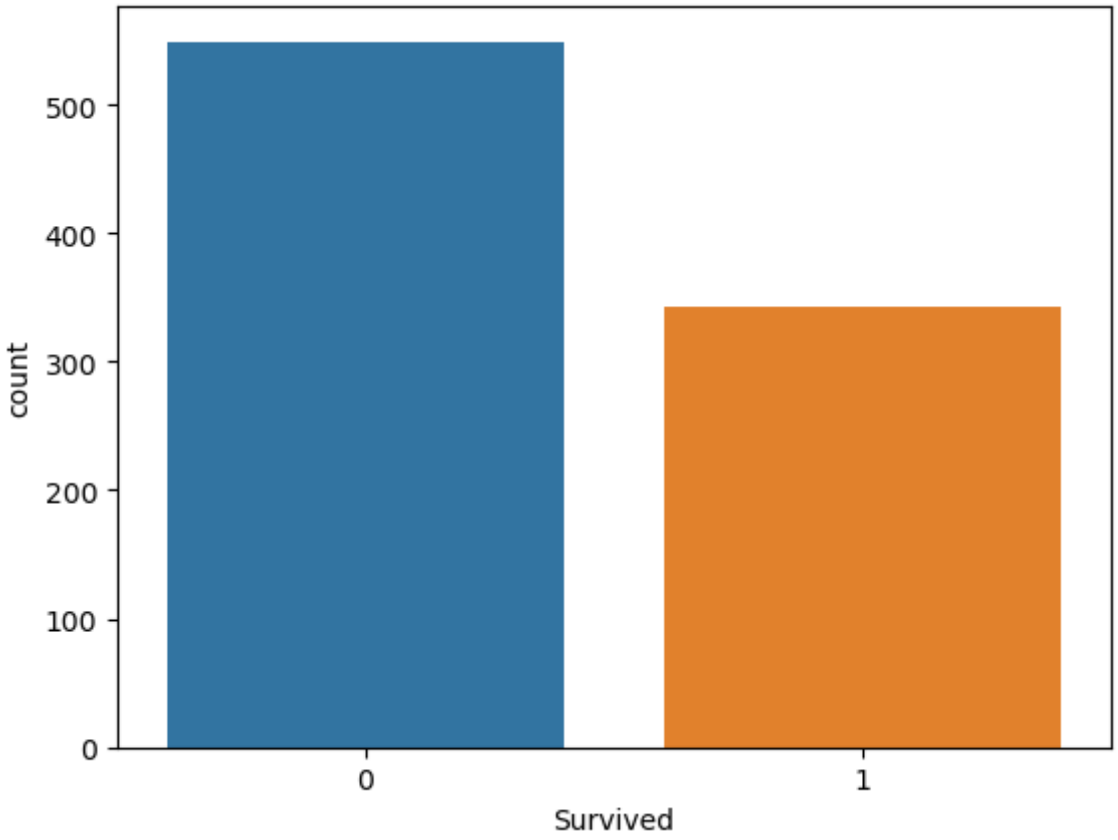
In [10]:
```
data.isnull().sum()
```

Out[10]:
```
PassengerId    0
Survived       0
Pclass         0
Name           0
Sex            0
Age            0
SibSp          0
Parch          0
Ticket         0
Fare           0
Cabin          0
Embarked       0
dtype: int64
```

In [11]:
```
sns.countplot(data['Survived'])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarnin
g: Pass the following variable as a keyword arg: x. From version 0.12, the only va
lid positional argument will be `data`, and passing other arguments without an exp
licit keyword will result in an error or misinterpretation.
  warnings.warn(
```

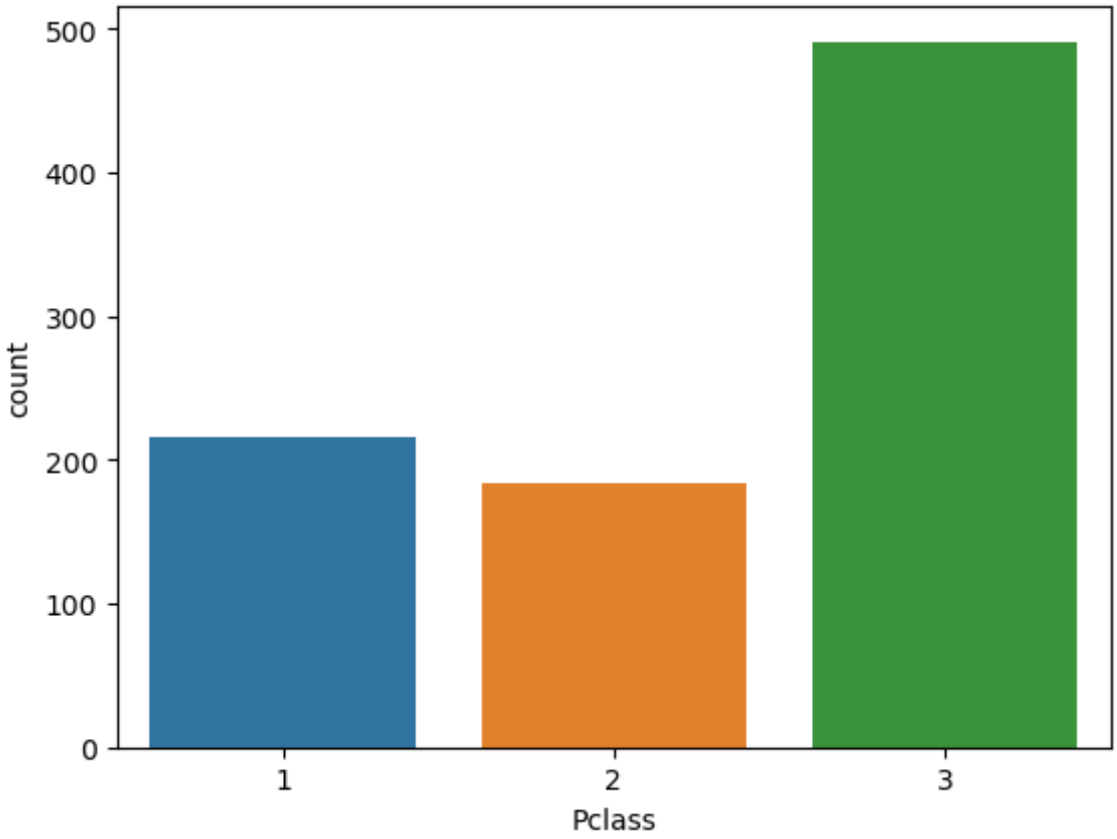
Out[11]:
```
<AxesSubplot:xlabel='Survived', ylabel='count'>
```



In [12]:
```
sns.countplot(data['Pclass'])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarnin
g: Pass the following variable as a keyword arg: x. From version 0.12, the only va
lid positional argument will be `data`, and passing other arguments without an exp
licit keyword will result in an error or misinterpretation.
  warnings.warn(
```

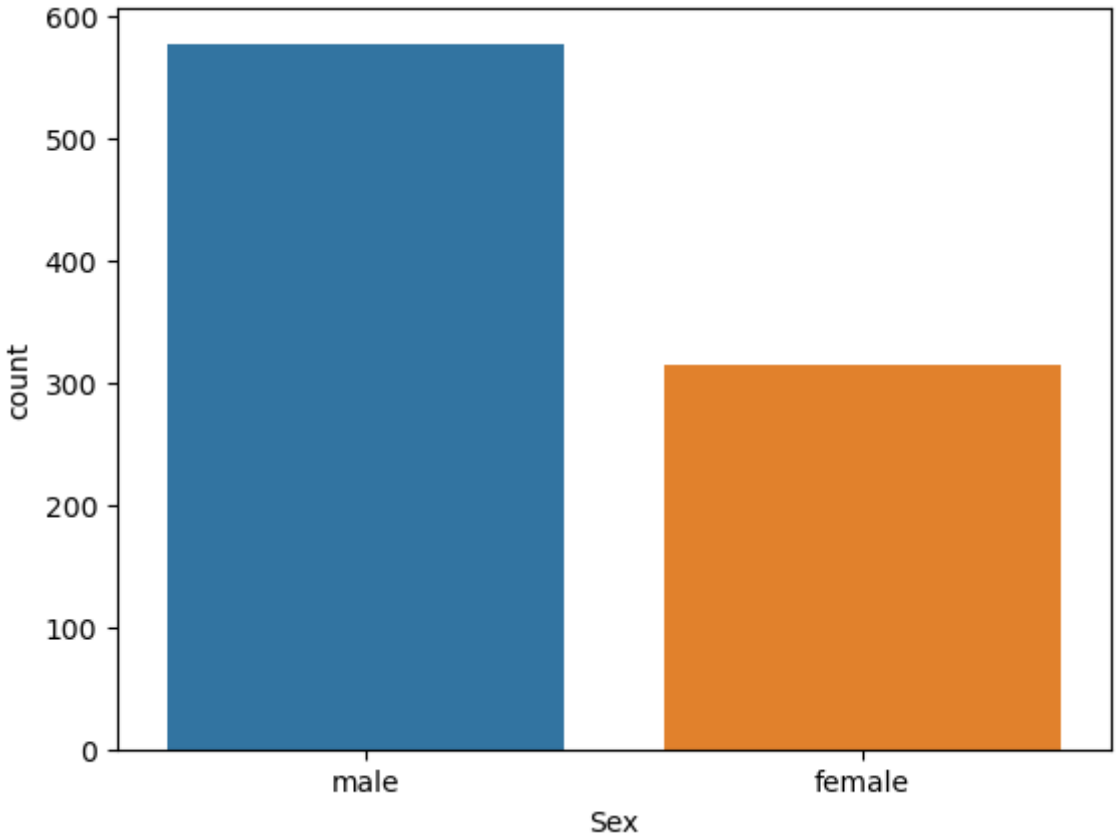
Out[12]:
```
<AxesSubplot:xlabel='Pclass', ylabel='count'>
```



In [13]:
```
sns.countplot(data['Sex'])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarnin
g: Pass the following variable as a keyword arg: x. From version 0.12, the only va
lid positional argument will be `data`, and passing other arguments without an exp
licit keyword will result in an error or misinterpretation.
  warnings.warn(
```

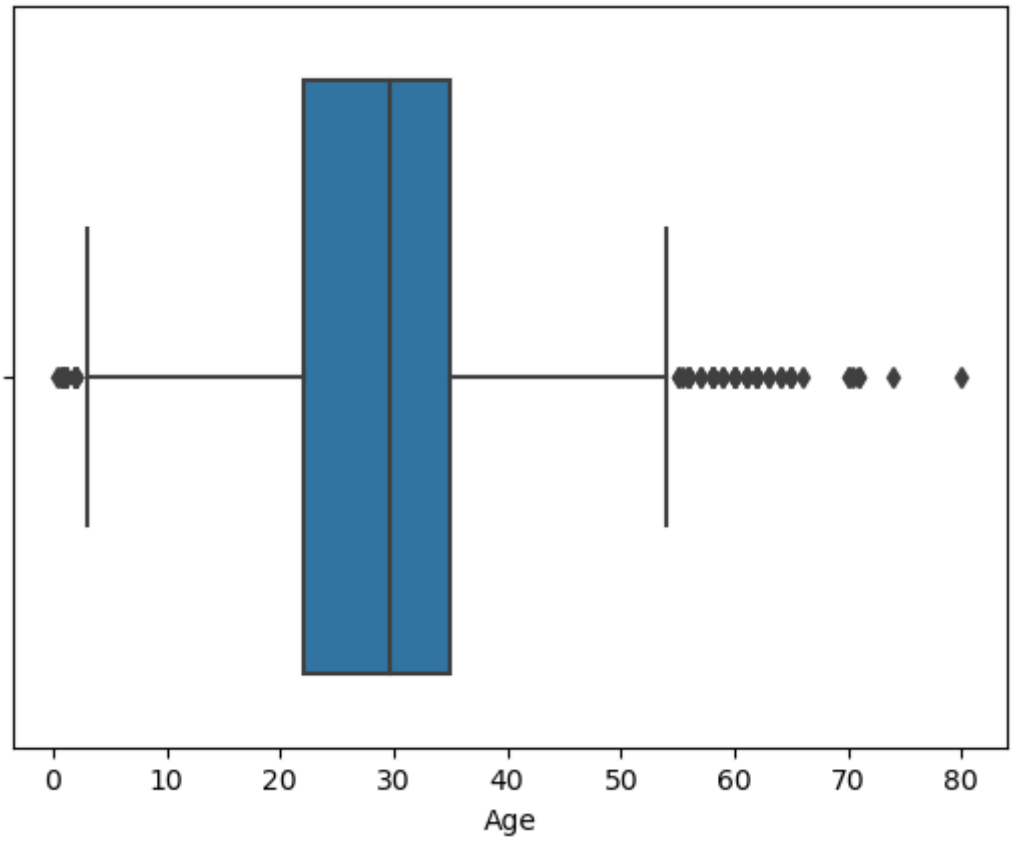
Out[13]:
```
<AxesSubplot:xlabel='Sex', ylabel='count'>
```



In [14]:
```
sns.boxplot(data['Age'])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarnin
g: Pass the following variable as a keyword arg: x. From version 0.12, the only va
lid positional argument will be `data`, and passing other arguments without an exp
licit keyword will result in an error or misinterpretation.
  warnings.warn(
```

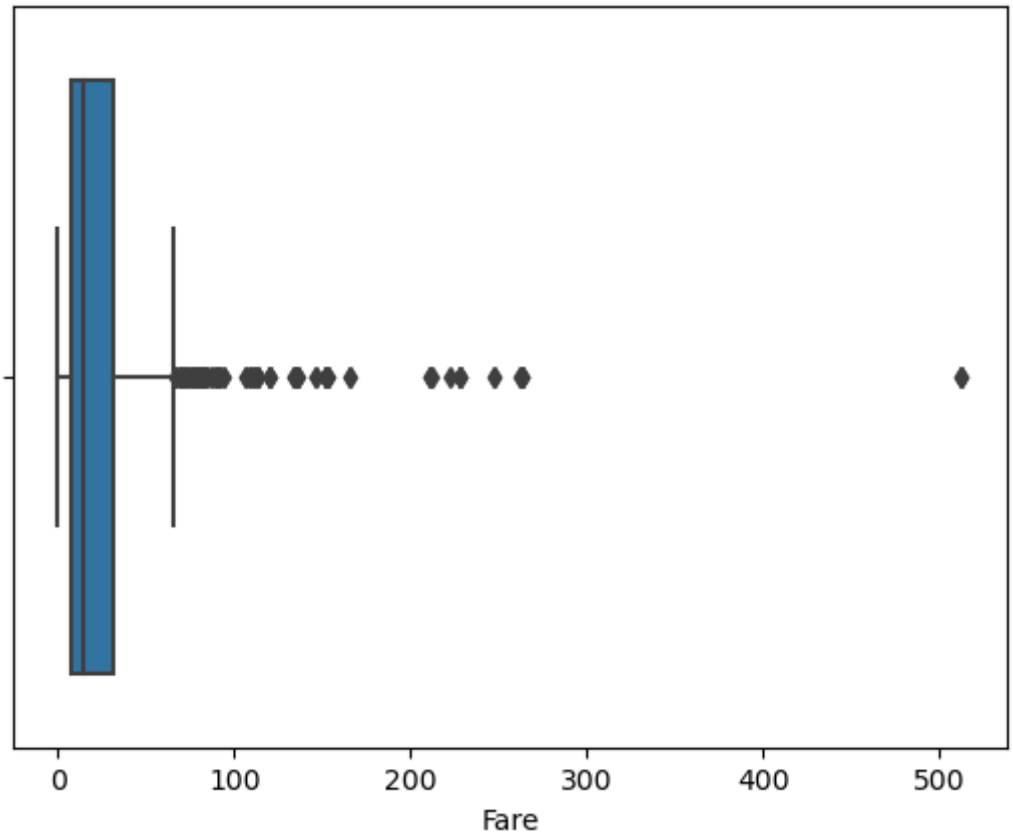
Out[14]:
```
<AxesSubplot:xlabel='Age'>
```



In [15]:
```
sns.boxplot(data['Fare'])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarnin
g: Pass the following variable as a keyword arg: x. From version 0.12, the only va
lid positional argument will be `data`, and passing other arguments without an exp
licit keyword will result in an error or misinterpretation.
  warnings.warn(
```

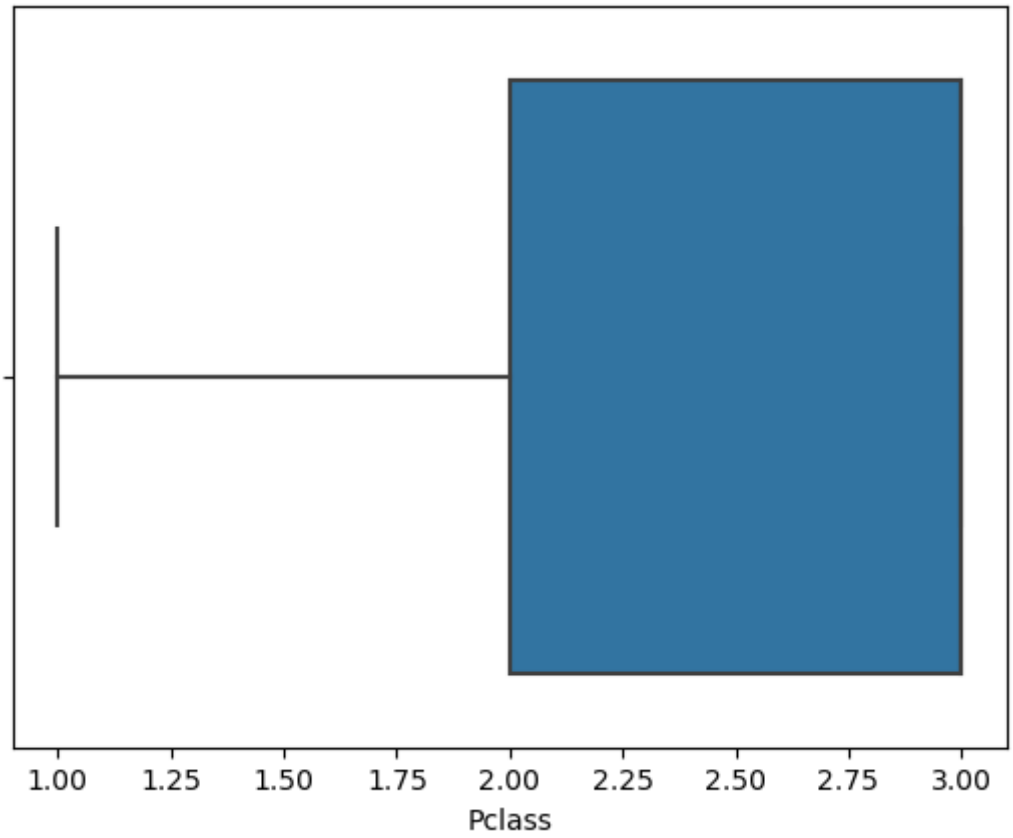
Out[15]:
```
<AxesSubplot:xlabel='Fare'>
```



In [16]:
```
sns.boxplot(data['Pclass'])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarnin
g: Pass the following variable as a keyword arg: x. From version 0.12, the only va
lid positional argument will be `data`, and passing other arguments without an exp
licit keyword will result in an error or misinterpretation.
  warnings.warn(
```

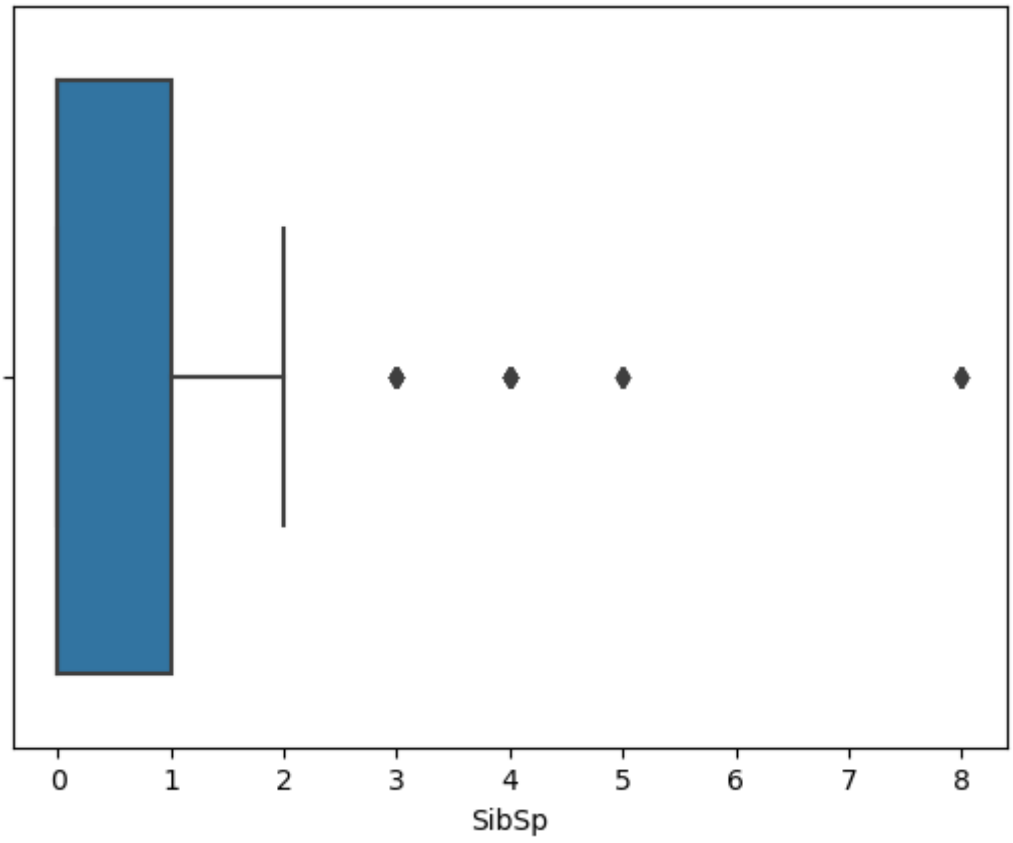
Out[16]:
```
<AxesSubplot:xlabel='Pclass'>
```



In [17]:
```
sns.boxplot(data['SibSp'])
```

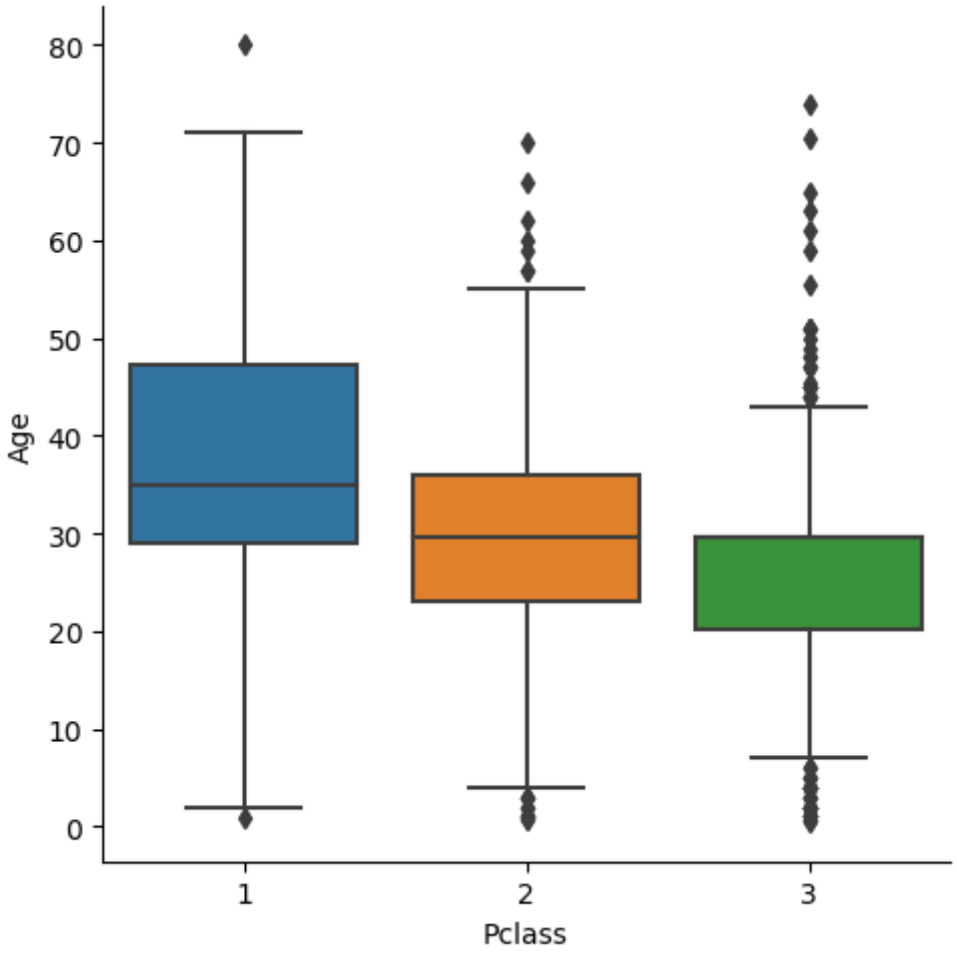
```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarnin
g: Pass the following variable as a keyword arg: x. From version 0.12, the only va
lid positional argument will be `data`, and passing other arguments without an exp
licit keyword will result in an error or misinterpretation.
  warnings.warn(
```

Out[17]:
```
<AxesSubplot:xlabel='SibSp'>
```



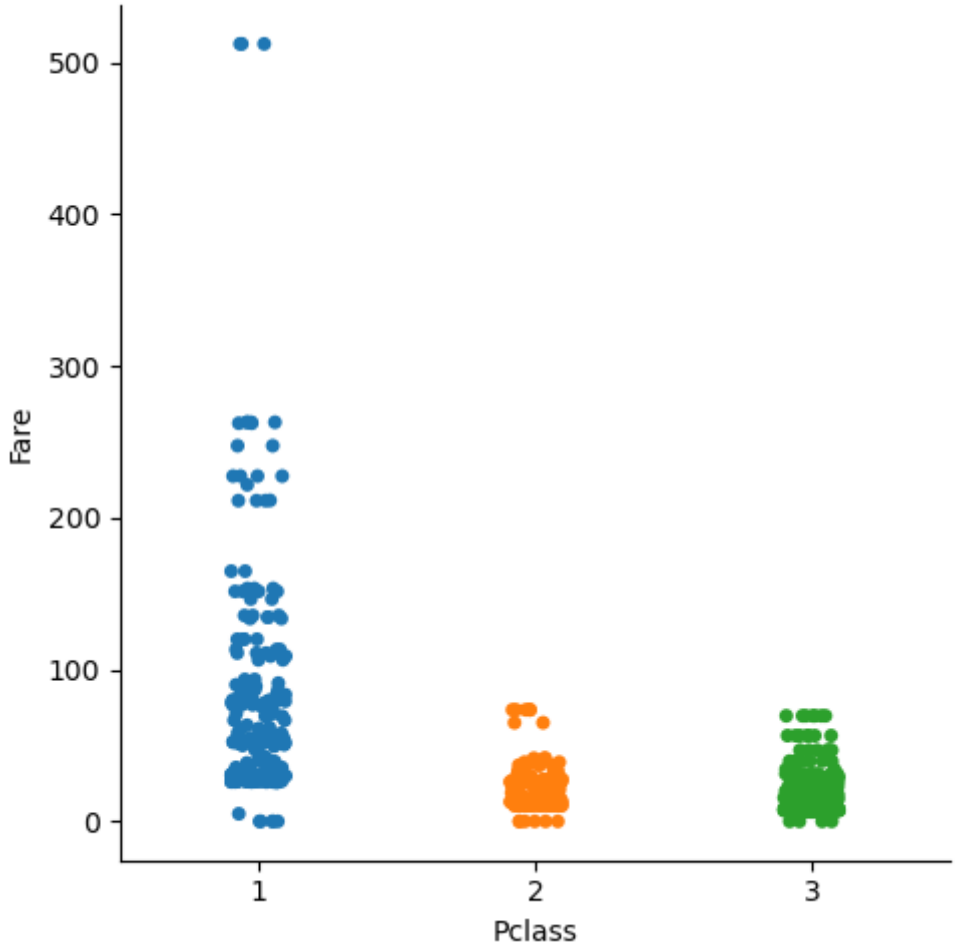
In [18]:
```
sns.catplot(x= 'Pclass', y = 'Age', data=data, kind = 'box')
```

Out[18]:
```
<seaborn.axisgrid.FacetGrid at 0x1f84a886c40>
```



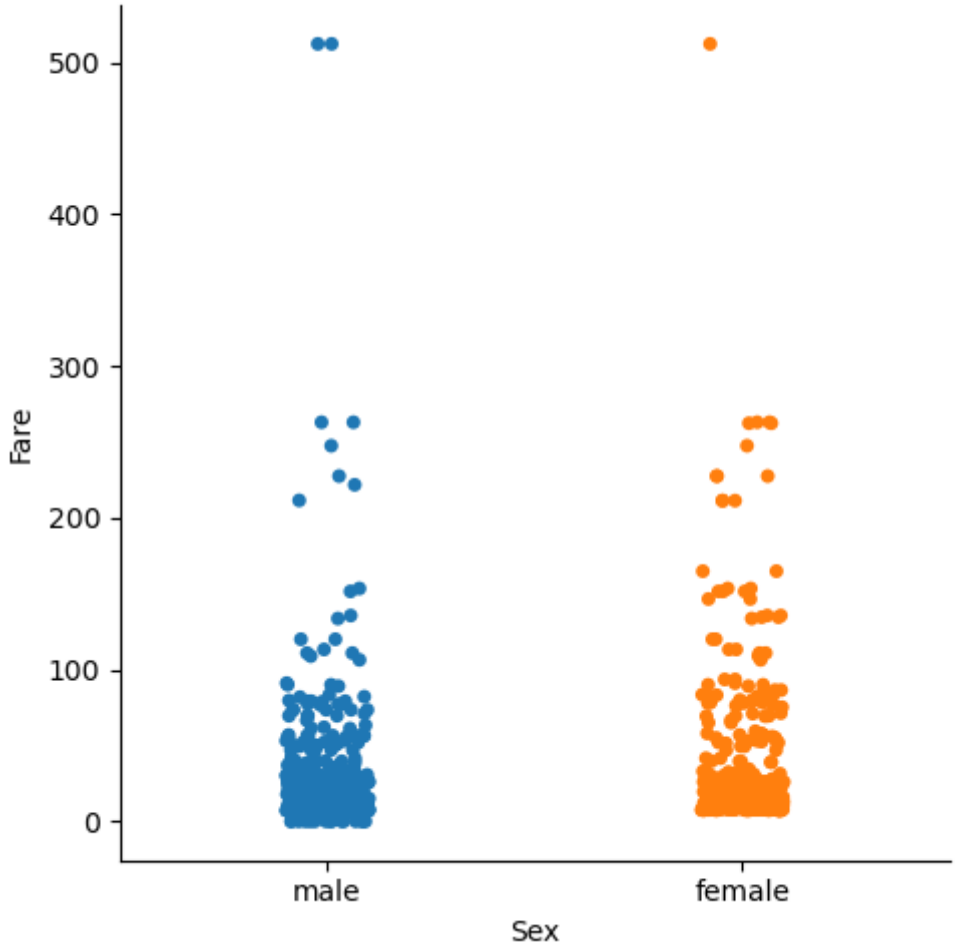
In [19]:
```
sns.catplot(x= 'Pclass', y = 'Fare', data=data, kind = 'strip')
```

Out[19]:
```
<seaborn.axisgrid.FacetGrid at 0x1f84a886df0>
```



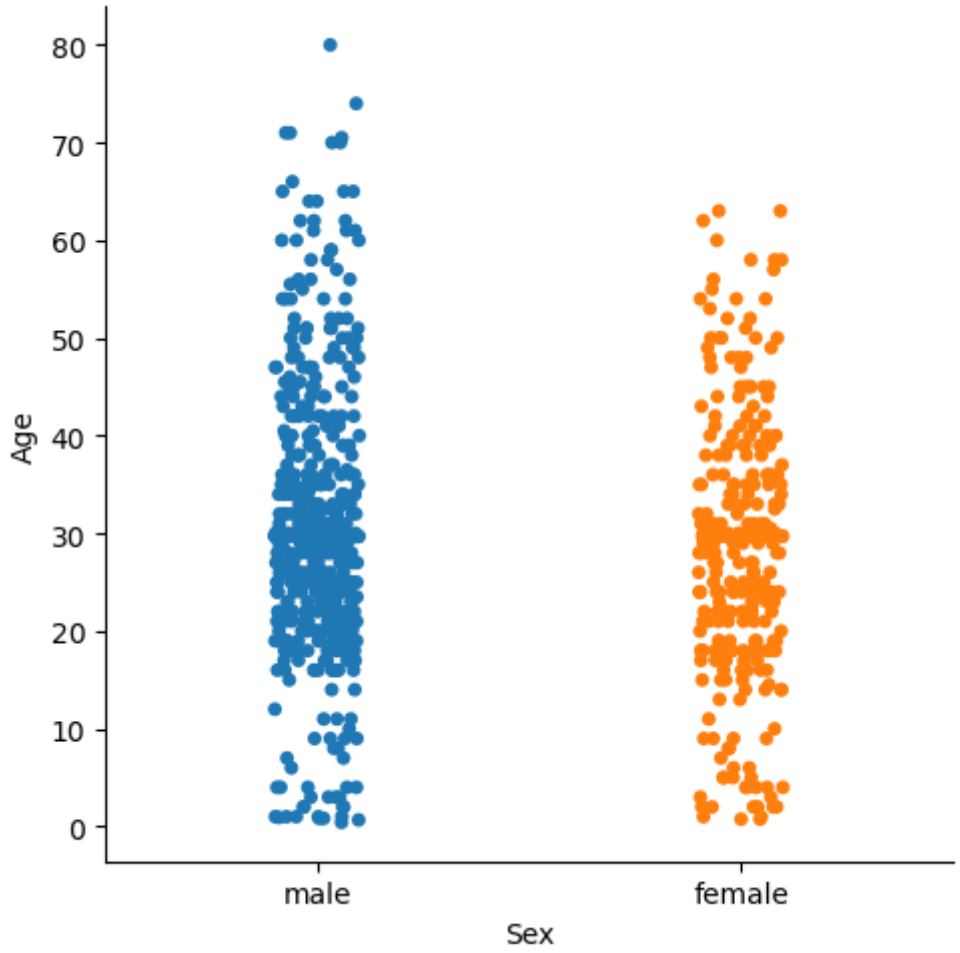
In [20]:
```
sns.catplot(x= 'Sex', y = 'Fare', data=data, kind = 'strip')
```

Out[20]:
```
<seaborn.axisgrid.FacetGrid at 0x1f84a6d1df0>
```



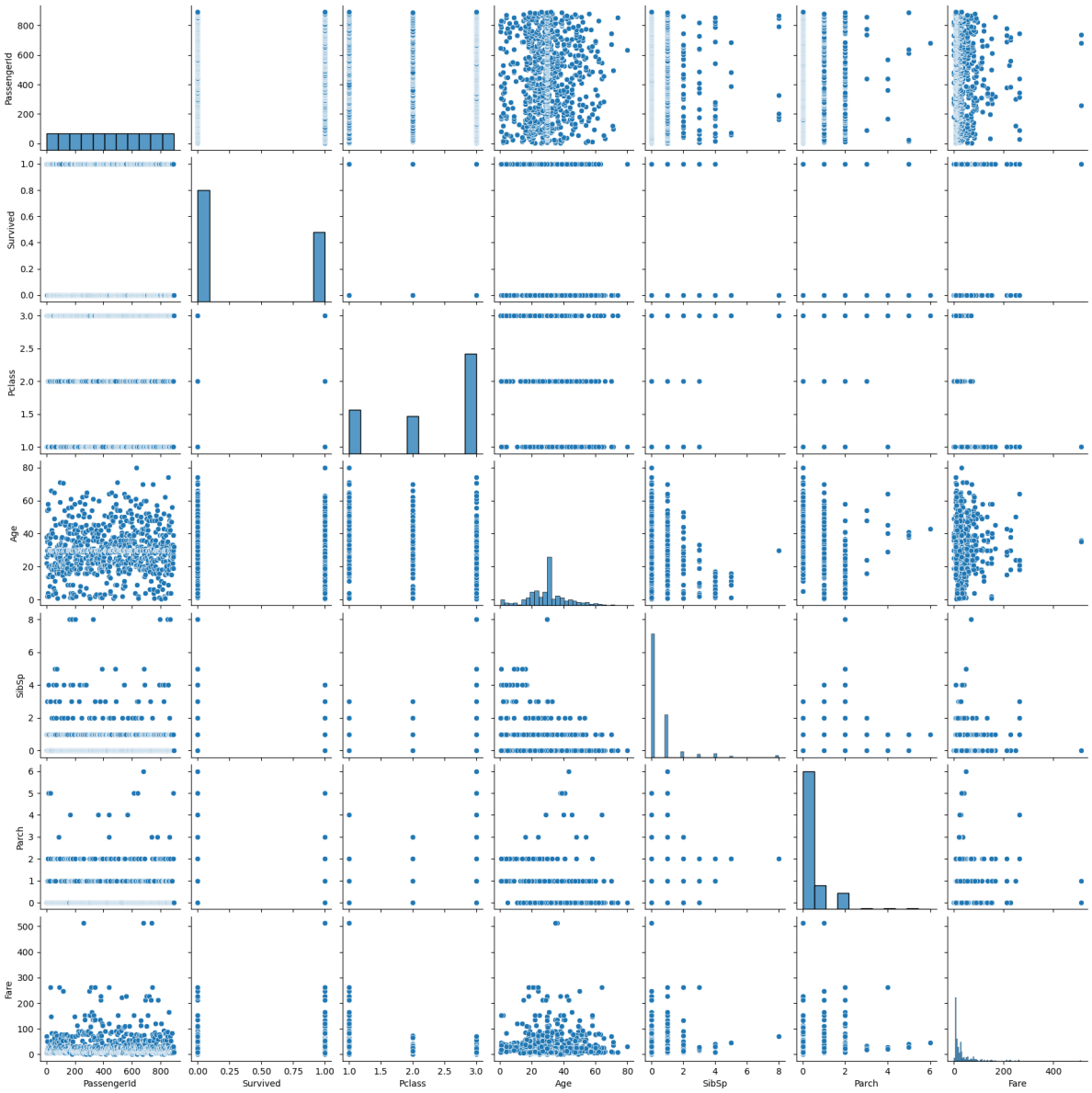
In [21]:
```
sns.catplot(x= 'Sex', y = 'Age', data=data, kind = 'strip')
```

Out[21]:
```
<seaborn.axisgrid.FacetGrid at 0x1f84aa4f910>
```



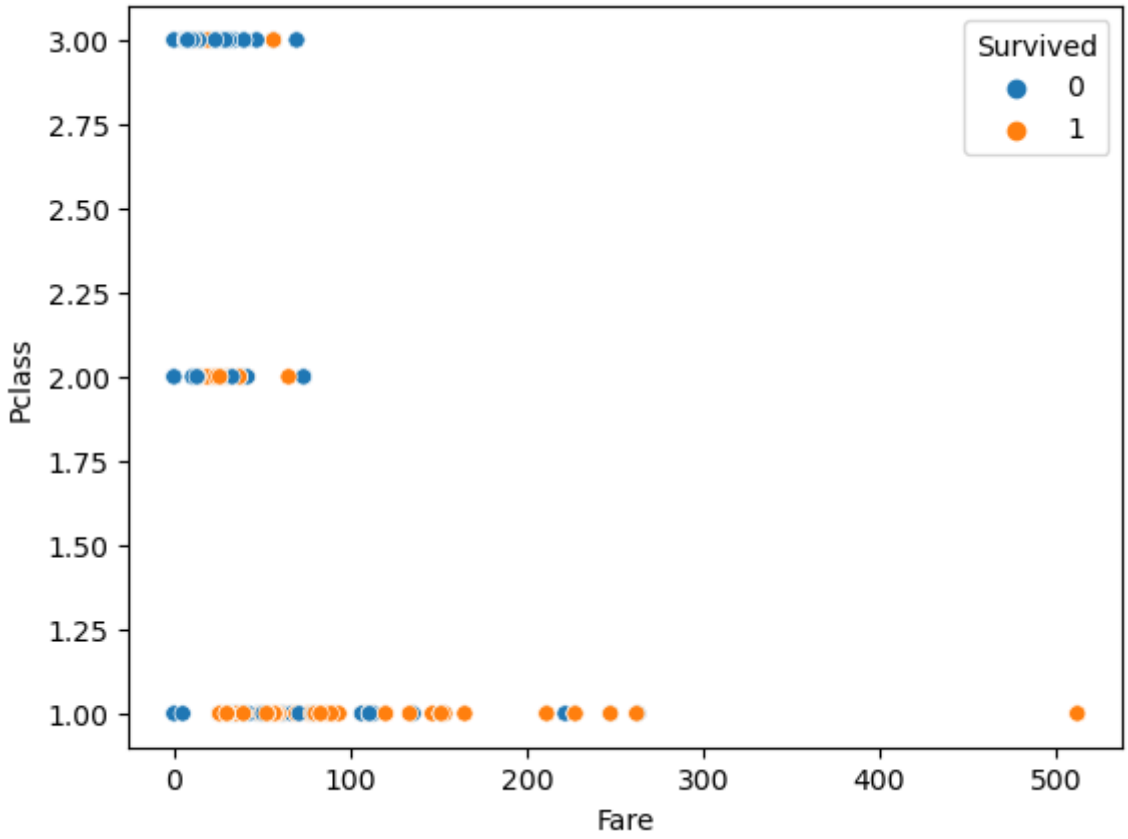
In [22]:
```
sns.pairplot(data)
```

Out[22]:
```
<seaborn.axisgrid.PairGrid at 0x1f84a922a90>
```



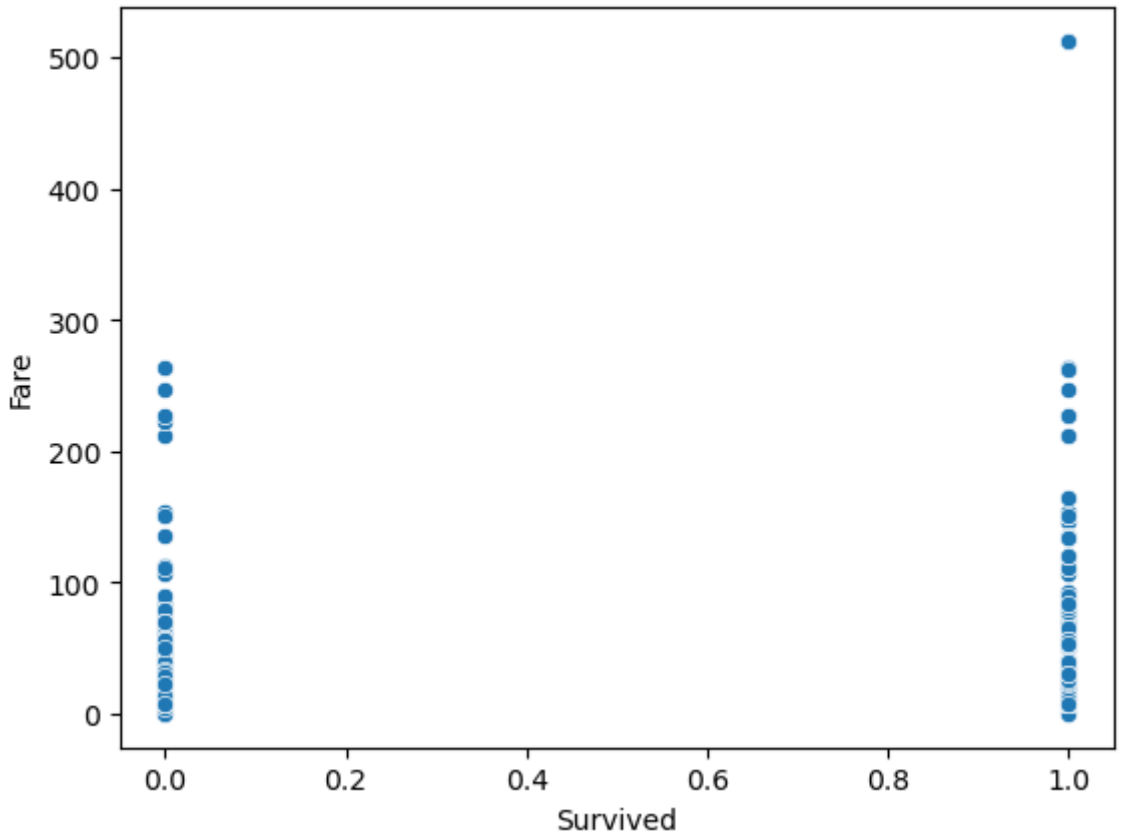
In [23]:
```
sns.scatterplot(x = 'Fare', y = 'Pclass', hue = 'Survived', data = data)
```

Out[23]:
```
<AxesSubplot:xlabel='Fare', ylabel='Pclass'>
```



In [24]:
```
sns.scatterplot(x = 'Survived', y = 'Fare', data = data)
```

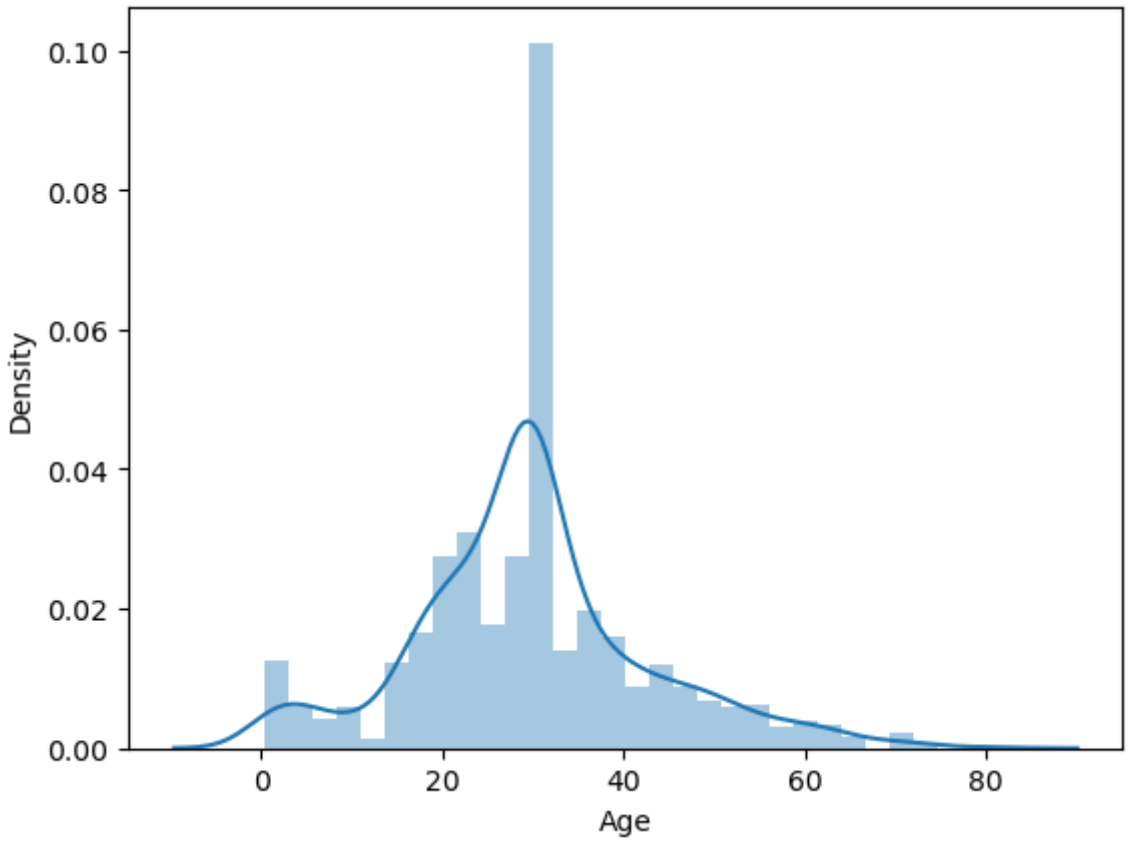
Out[24]:
```
<AxesSubplot:xlabel='Survived', ylabel='Fare'>
```



In [25]:
```
sns.distplot(data['Age'])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWa
rning: `distplot` is a deprecated function and will be removed in a future versio
n. Please adapt your code to use either `displot` (a figure-level function with si
milar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

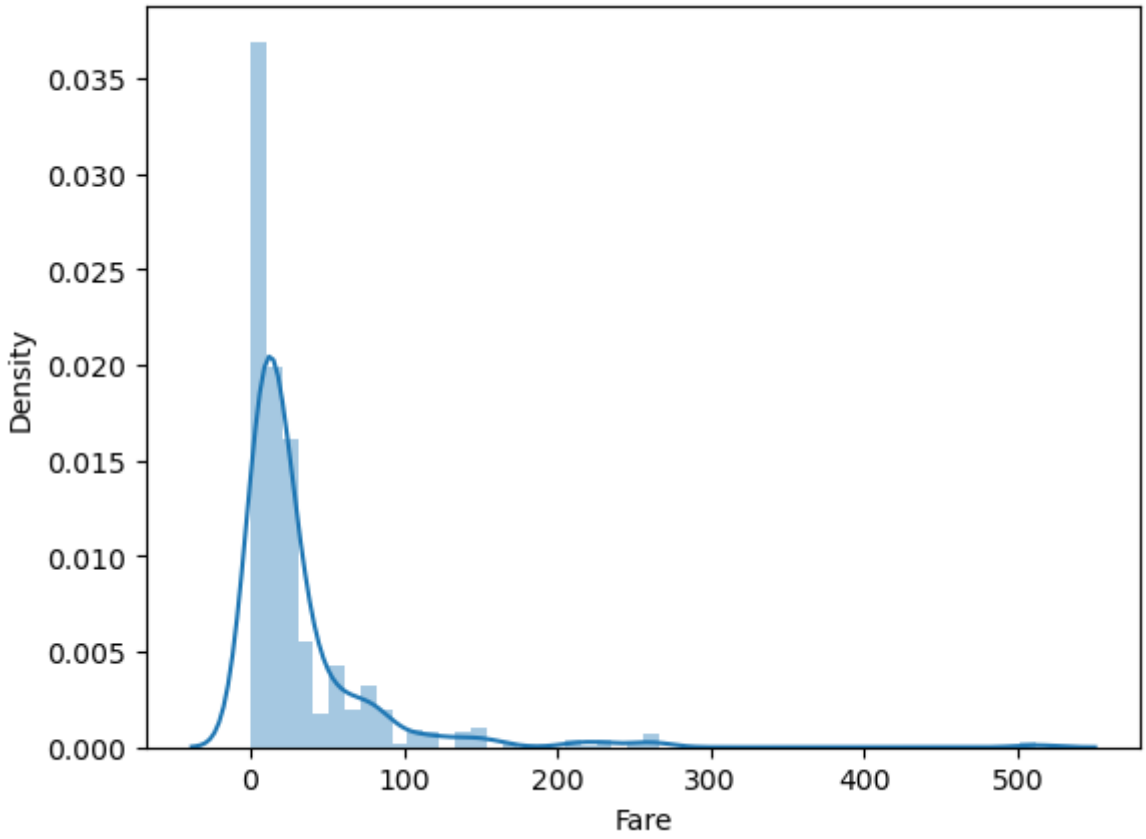
Out[25]: <AxesSubplot:xlabel='Age', ylabel='Density'>



In [26]:
```
sns.distplot(data['Fare'])
```

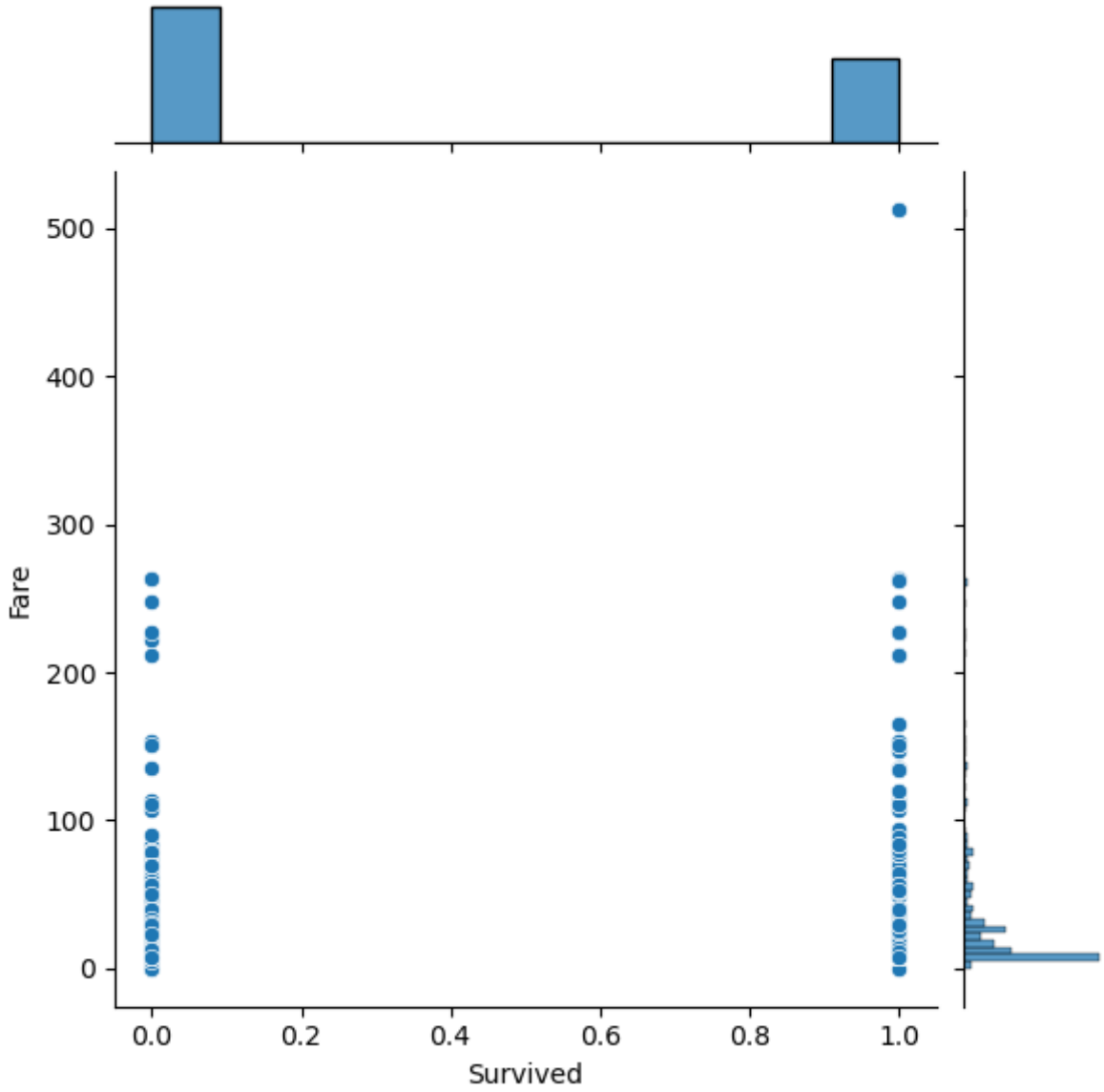
```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWa
rning: `distplot` is a deprecated function and will be removed in a future versio
n. Please adapt your code to use either `displot` (a figure-level function with si
milar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Out[26]: <AxesSubplot:xlabel='Fare', ylabel='Density'>



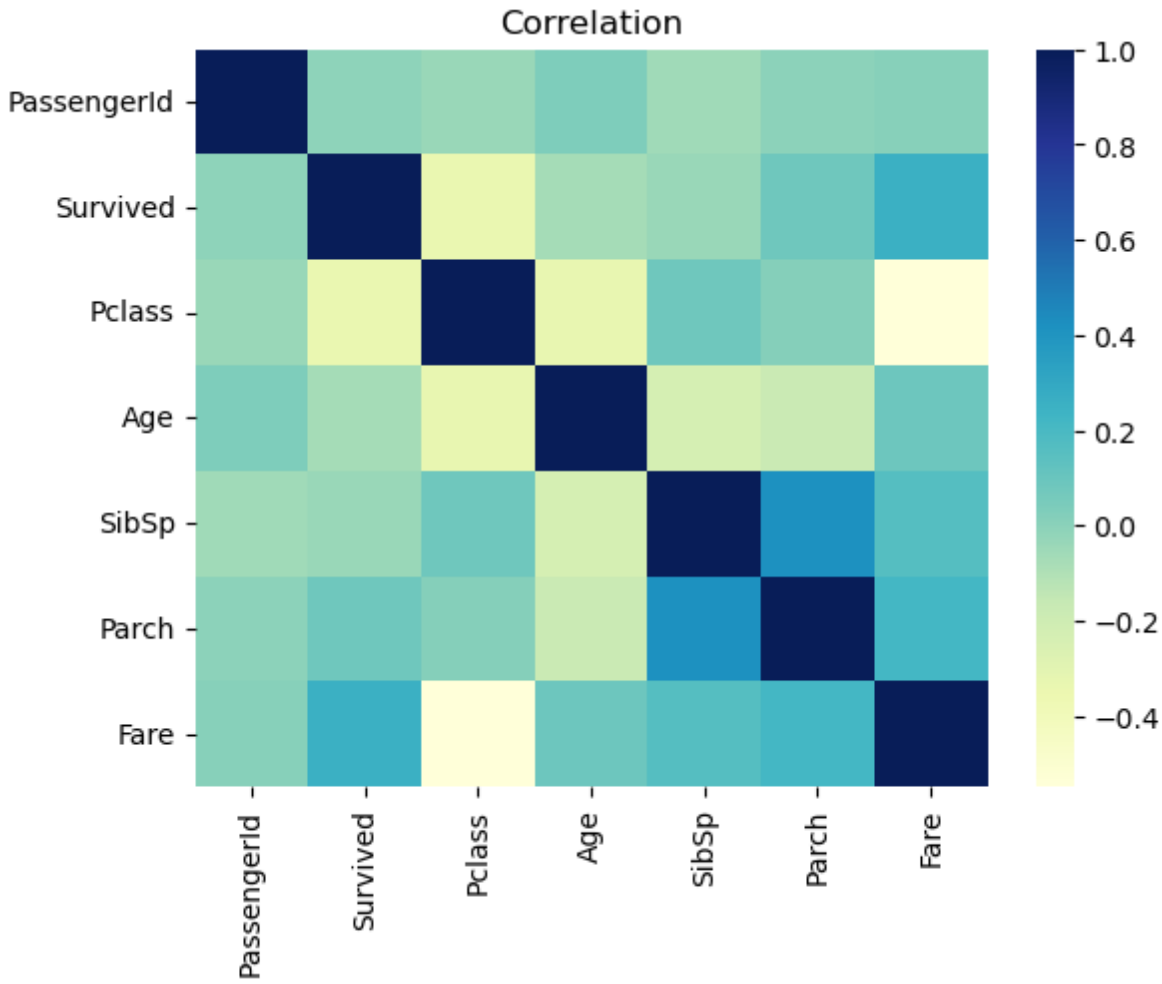
In [27]:
```
sns.jointplot(x = "Survived", y = "Fare", kind = "scatter", data = data)
```

Out[27]:
```
<seaborn.axisgrid.JointGrid at 0x1f84f8c5850>
```



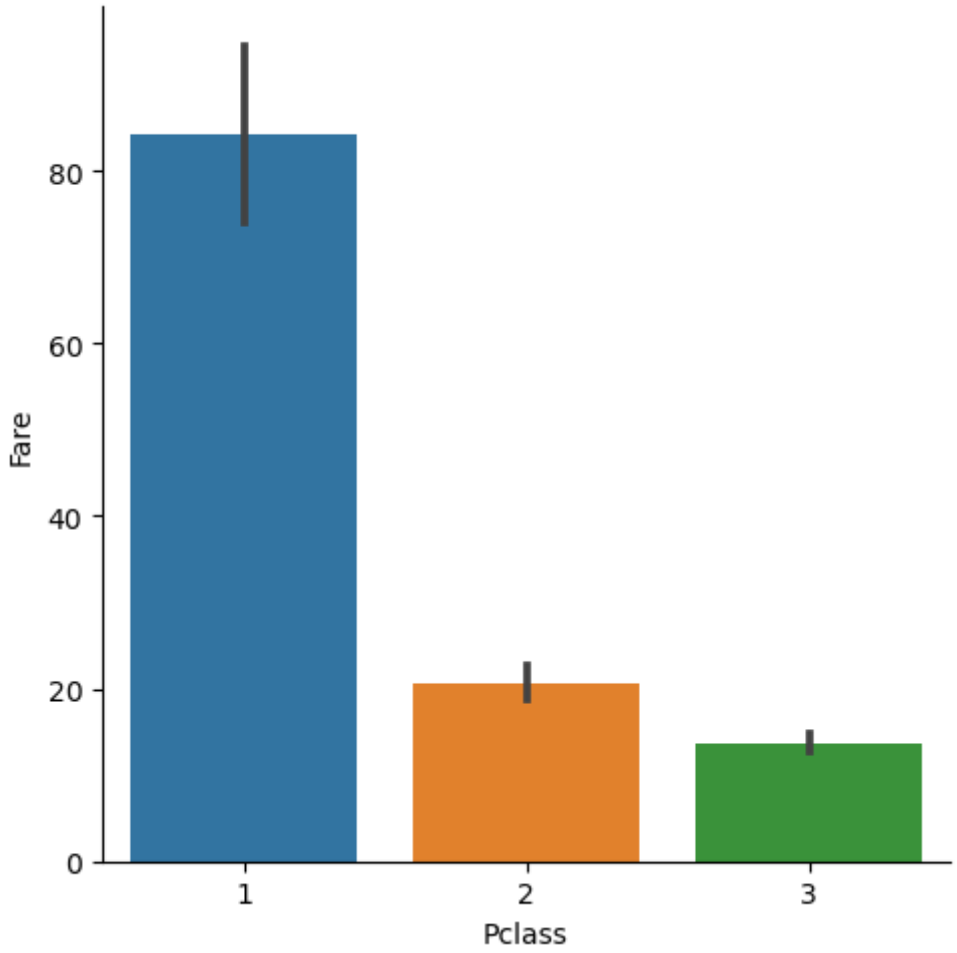
In [28]:
```
tc = data.corr()
sns.heatmap(tc, cmap="YlGnBu")
plt.title('Correlation')
```

Out[28]:
```
Text(0.5, 1.0, 'Correlation')
```



In [29]:
```
sns.catplot(x='Pclass', y='Fare', data=data, kind='bar')
```

Out[29]:
```
<seaborn.axisgrid.FacetGrid at 0x1f84f9b5610>
```



In [ ]: