

In [1]:

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

In [2]:

```
df = pd.read_csv("C:/Users/ameya/OneDrive/Desktop/DSBDAL/Titanic-Dataset.csv")
```

In [3]:

```
df.head()
```

Out[3]:

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



In [4]:

```
df.isnull().sum()
```

Out[4]:

```
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 [5]:

```
df['Age'].fillna(value = df['Age'].mean(),inplace=True)
```

In [6]:

```
df.isnull().sum()
```

Out[6]:

```
PassengerId      0
Survived          0
Pclass           0
Name             0
Sex              0
Age             0
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin           687
Embarked         2
dtype: int64
```

In [7]:

```
df2 = df.drop("Cabin",axis=1)
```

In [8]:

```
df2.isnull().sum()
```

Out[8]:

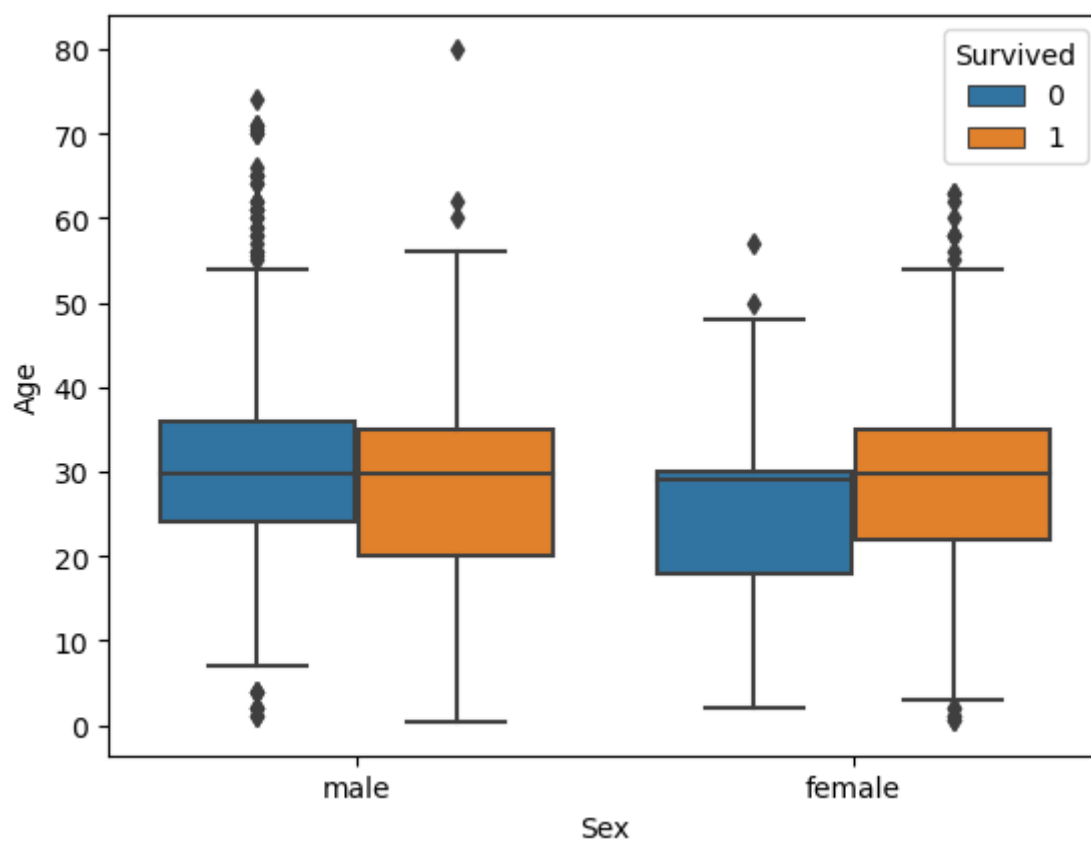
```
PassengerId    0
Survived        0
Pclass         0
Name           0
Sex            0
Age           0
SibSp          0
Parch          0
Ticket         0
Fare           0
Embarked       2
dtype: int64
```

In [9]:

```
sns.boxplot(data = df2,x='Sex',y='Age',hue='Survived')
```

Out[9]:

```
<AxesSubplot:xlabel='Sex', ylabel='Age'>
```



In [10]:

```
df2['Survived'].value_counts()
```

Out[10]:

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

In [11]:

```
df2['Sex'].value_counts()
```

Out[11]:

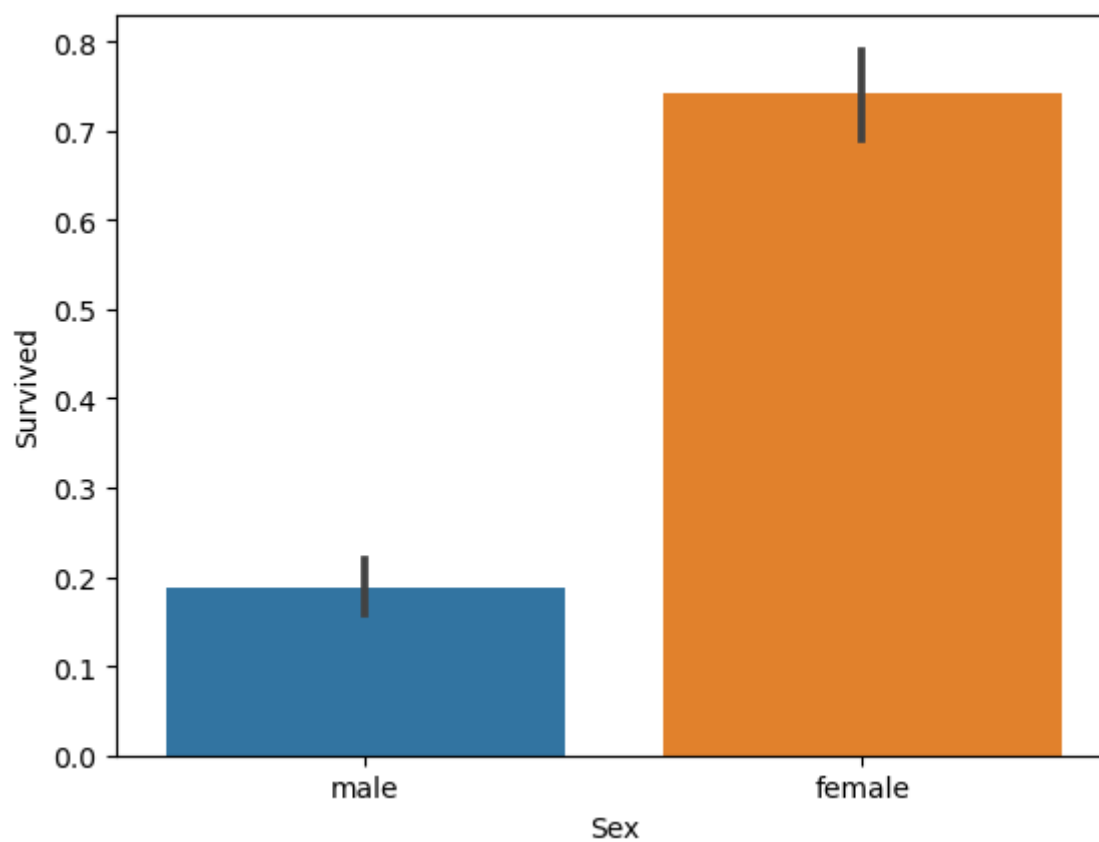
```
male    577
female  314
Name: Sex, dtype: int64
```

In [12]:

```
sns.barplot(data=df, x='Sex', y='Survived')
```

Out[12]:

<AxesSubplot:xlabel='Sex', ylabel='Survived'>

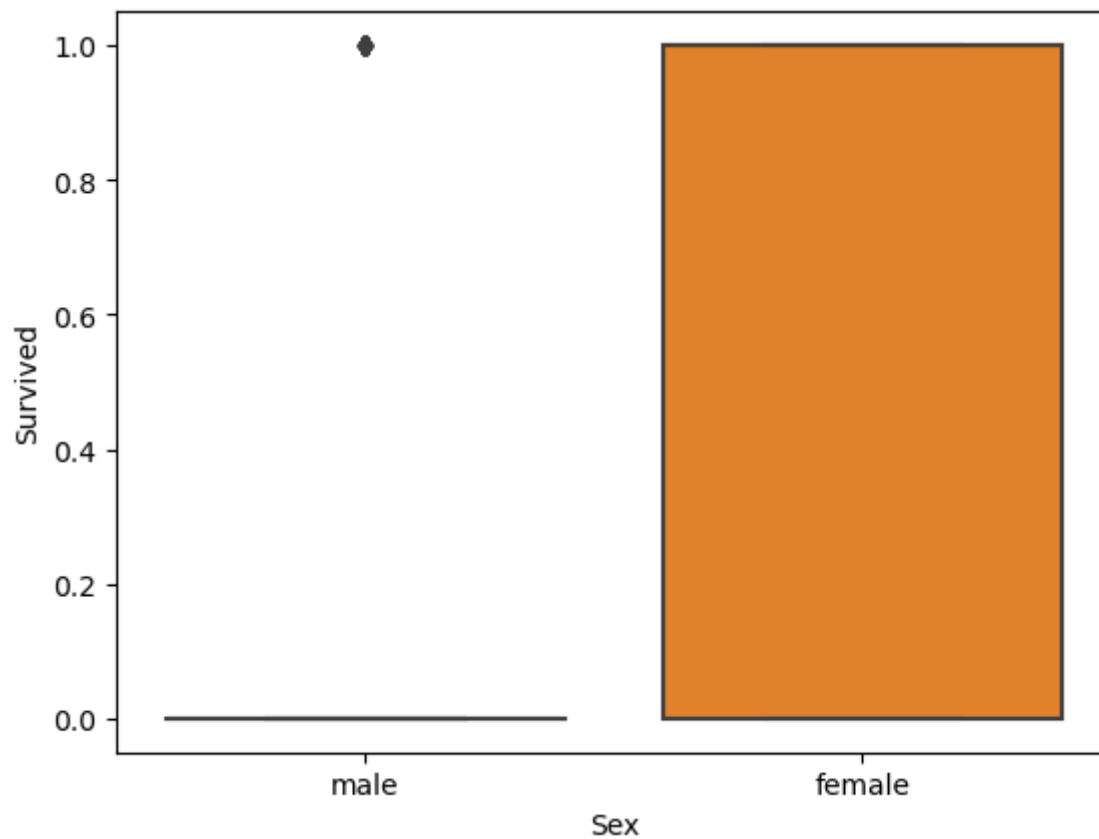


In [13]:

```
sns.boxplot(data=df,x='Sex',y='Survived')
```

Out[13]:

<AxesSubplot:xlabel='Sex', ylabel='Survived'>



In []: