31444 Mehul

Assignment - 9

In [ ]:
```
import pandas as pd
import seaborn as sb
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

In [ ]:
```
data = pd.read_csv("titanic.csv")
```

In [ ]:
```
data.head()
```

Out[ ]:

	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	7.250(
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	7.925(
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.100(
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.050(

In [ ]:
```
data.drop(columns=["Cabin","PassengerId","Name","Ticket"],inplace=True)
data.isnull().sum()
```

Out[ ]:
```
Survived      0
Pclass        0
Sex           0
Age         177
SibSp         0
Parch         0
Fare          0
Embarked      2
dtype: int64
```

In [ ]:
```
data["Age"].fillna(value=data["Age"].median(), inplace=True)
data["Embarked"] = data["Embarked"].fillna('S')
data.isnull().sum()
```

Out[ ]:
```
Survived    0
Pclass      0
Sex         0
Age         0
SibSp       0
Parch       0
Fare        0
Embarked    0
dtype: int64
```

In [ ]:
```
data["Age"]= data["Age"].astype("int")
```

In [ ]:
```
data.dtypes
```

Out[ ]:
```
Survived      int64
Pclass        int64
Sex          object
Age           int64
SibSp         int64
Parch         int64
Fare        float64
Embarked     object
dtype: object
```

In [ ]:
```
sb.boxplot(x='Sex', y='Age',hue="Survived" ,data=data)
```

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