In [1]:
```
import pandas as pd
```

In [2]:
```
import numpy as np
```

In [3]:
```
import seaborn as sns
```

In [4]:
```
df=pd.read_csv("C:\\Users\\admn\\Desktop\\titanic_dataset.csv")
```

In [5]:
```
df
```

Out[5]:

	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
...	...	...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

891 rows × 12 columns

In [7]:
```
df.isnull().sum()
```

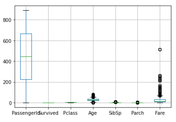
Out[7]:
```
PassengerId      0
Survived         0
Pclass           0
Name             0
Sex              0
Age            177
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin          687
Embarked         2
dtype: int64
```

In [8]:
```
df.drop(columns=['Cabin'],inplace=True)
```

In [10]:
```
df['Age']=df['Age'].fillna(df['Age'].median())
```

In [11]:
```
df.boxplot()
```

Out[11]:
```
<AxesSubplot:>
```



In [12]:
```
df.isnull().sum()
```

Out[12]:
```
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 [13]:
```
df['Embarked']=df['Embarked'].fillna(df['Embarked'].mode()[0])
```

In [14]:
```
df['Embarked'].value_counts()
```

Out[14]:
```
S    646
C    168
Q     77
Name: Embarked, dtype: int64
```

In [15]:
```
df['Pclass'].value_counts()
```

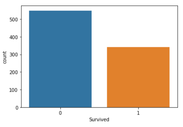
Out[15]:
```
3    491
1    216
2    184
Name: Pclass, dtype: int64
```

In [16]:
```
df['Survived'].value_counts()
```

Out[16]:
```
0    549
1    342
Name: Survived, dtype: int64
```

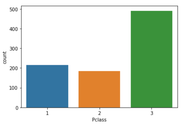
In [17]:
```
sns.countplot(x='Survived',data=df)
```

Out[17]:
```
<AxesSubplot:xlabel='Survived', ylabel='count'>
```



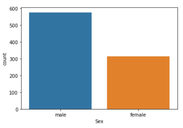
In [18]:
```
sns.countplot(x='Pclass',data=df)
```

Out[18]:
```
<AxesSubplot:xlabel='Pclass', ylabel='count'>
```



In [20]:
```
sns.countplot(x='Sex',data=df)
```

Out[20]:
```
<AxesSubplot:xlabel='Sex', ylabel='count'>
```

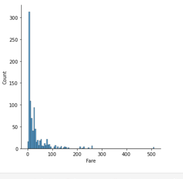


In [21]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 11 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          891 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  Embarked     891 non-null    object 
dtypes: float64(2), int64(5), object(4)
memory usage: 76.7+ KB
```

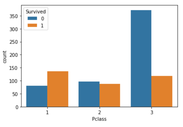
In [22]:
```
sns.displot(df['Fare'])
```

Out[22]:
```
<seaborn.axisgrid.FacetGrid at 0x1fa8ef241f0>
```



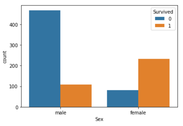
In [23]:
```
sns.countplot(x='Pclass',hue='Survived',data=df)
```

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



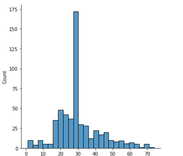
In [24]:
```
sns.countplot(x='Sex',hue='Survived',data=df)
```

Out[24]:
```
<AxesSubplot:xlabel='Sex', ylabel='count'>
```



In [25]:
```
sns.displot(df[df['Survived']==0]['Age'])
```

Out[25]:
```
<seaborn.axisgrid.FacetGrid at 0x1fa8f12fe20>
```



In [ ]:
```

```