In [ ]:
```
NAME : ARYAN SIRDESAI
ROLL No.: TACO20175
Lab Assignment 9 : Data Visualization II

Problem Statement :
1. Use the inbuilt dataset 'titanic' as used in the above problem. Plot a box plot
respect to each gender along with the information about whether they survived or n
and 'age')
2. Write observations on the inference from the above statistics.
```

In [1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
```

In [2]:
```
df = pd.read_csv("titanic.csv")
```

In [3]:
```
df.columns
```

Out[3]:
```
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
       'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
      dtype='object')
```

In [4]:
```
df.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]:
```
df.isnull().sum()
```

Out[5]:
```
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 [6]:
```
Q1=df['Age'].quantile(0.25)
Q3=df['Age'].quantile(0.75)
IQR=Q3-Q1
```



```
")- Q1(", Q1, ")")
```

```
IQR( 17.875 ) = Q3( 38.0 )- Q1( 20.125 )
```

In [7]:
```
lower_limit=Q1-IQR
upper_limit=Q3+IQR
lower_limit,upper_limit
```

Out[7]:
```
(2.25, 55.875)
```

In [8]:
```
df_without_outliers=df[(df['Age']>lower_limit)&(df['Age']<upper_limit)]
df_without_outliers
```

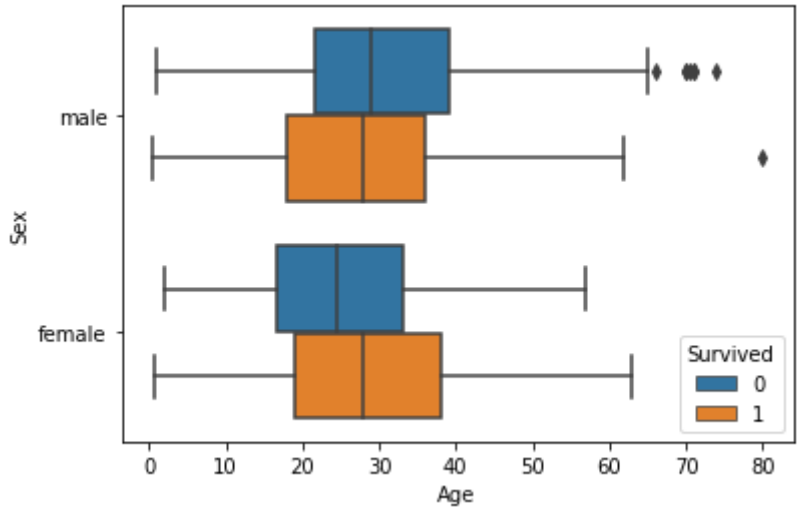
Out[8]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cal
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	C1
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	N
...	...	...	...	...	...	...	...	...	...	...	
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.1250	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	E
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C1
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	N

651 rows × 12 columns

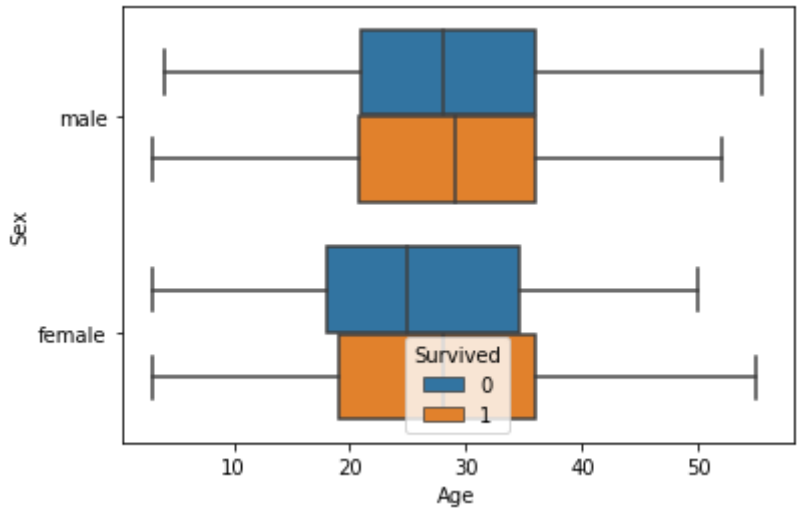
In [9]:
```
sns.boxplot(x='Age' , y='Sex', hue='Survived' ,data = df)
```

Out[9]:
```
<AxesSubplot:xlabel='Age', ylabel='Sex'>
```



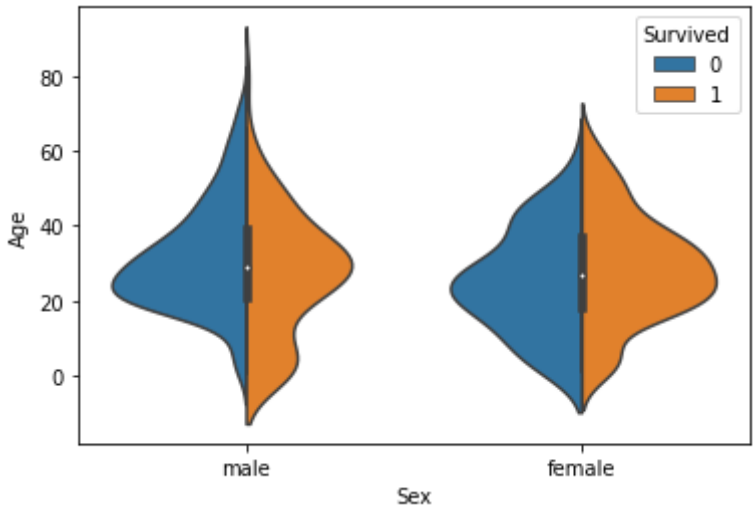
In [13]:
```
sns.boxplot(x='Age' , y='Sex', hue='Survived' ,data = df_without_outliers)
```

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



In [10]:
```
sns.violinplot(x='Sex', y='Age', data=df, hue="Survived",split=True)
```

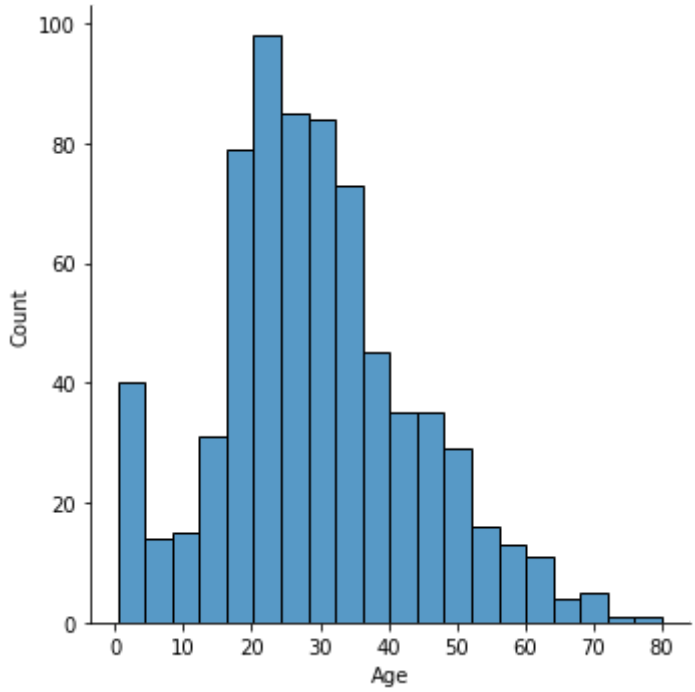
Out[10]:
```
<AxesSubplot:xlabel='Sex', ylabel='Age'>
```



Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

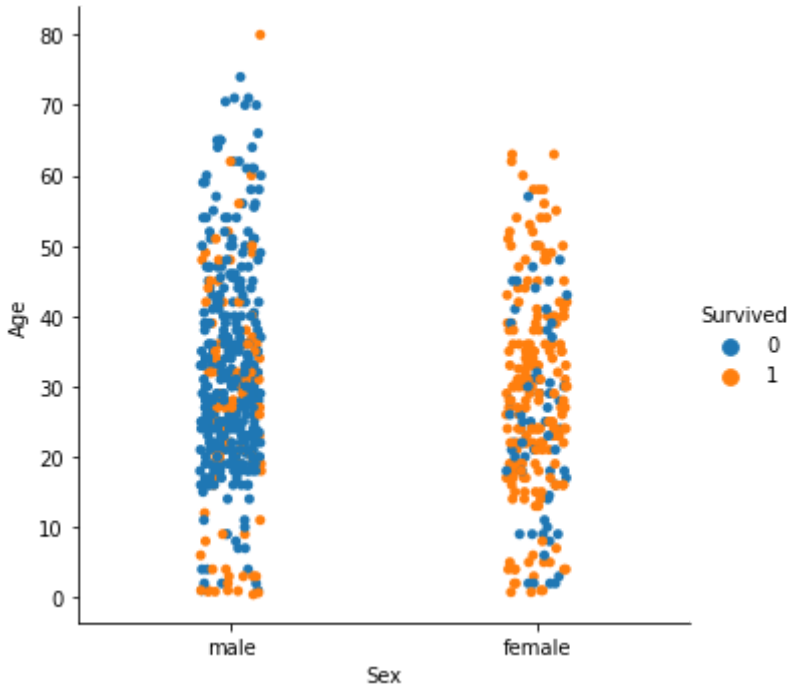
In [11]:
```
sns.displot(df['Age'])
```

Out[11]:
```
<seaborn.axisgrid.FacetGrid at 0x7fc97b81bdf0>
```



In [12]:
```
sns.catplot(x='Sex',y='Age' , data=df,hue='Survived')
```

Out[12]:
```
<seaborn.axisgrid.FacetGrid at 0x7fc97d9108e0>
```

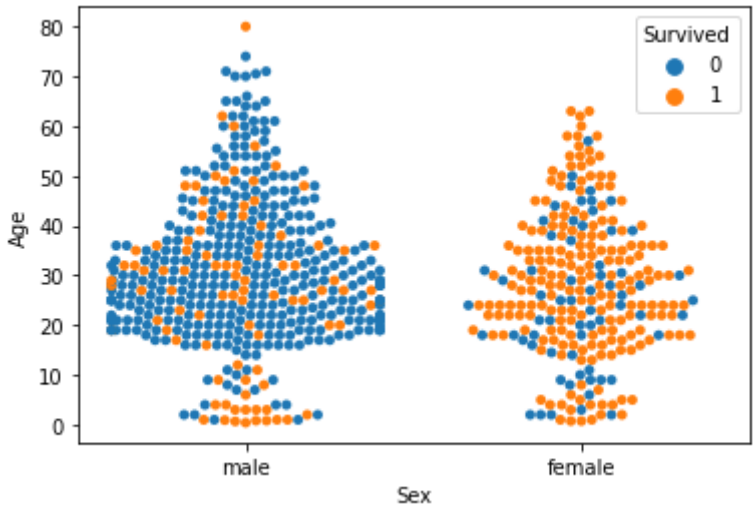


In [15]:
```
sns.swarmplot(x='Sex', y='Age', data=df, hue='Survived')
```

```
/home/pict/.local/lib/python3.8/site-packages/seaborn/categorical.py:1296: UserWar
ning: 5.9% of the points cannot be placed; you may want to decrease the size of th
e markers or use stripplot.
  warnings.warn(msg, UserWarning)
```

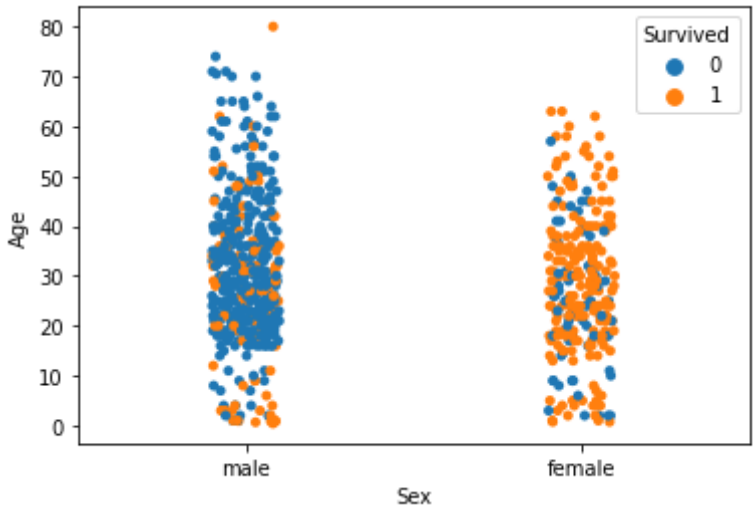
Out[15]:
```
<AxesSubplot:xlabel='Sex', ylabel='Age'>
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js



In [16]:
```
sns.stripplot(x='Sex', y='Age', data=df, hue='Survived')
```

Out[16]:
```
<AxesSubplot:xlabel='Sex', ylabel='Age'>
```



In [ ]:
```

```