Assignment No. 9

Aim : Data Visualization II

1. Use the inbuilt dataset 'titanic' as used in the above problem. Plot a box plot for distribution of age with respect to each gender along with the information about whether they survived or not. (Column names : 'sex' and 'age')
2. Write observations on the inference from the above statistics.

Code :

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df1 = sns.load_dataset('titanic')
df1
```

Out[1]:

| | survived | pclass | sex | age | sibsp | parch | fare | embarked | class | who | adult_mal |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 3 | male | 22.0 | 1 | 0 | 7.2500 | S | Third | man | Tru |
| **1** | 1 | 1 | female | 38.0 | 1 | 0 | 71.2833 | C | First | woman | Fals |
| **2** | 1 | 3 | female | 26.0 | 0 | 0 | 7.9250 | S | Third | woman | Fals |
| **3** | 1 | 1 | female | 35.0 | 1 | 0 | 53.1000 | S | First | woman | Fals |
| **4** | 0 | 3 | male | 35.0 | 0 | 0 | 8.0500 | S | Third | man | Tru |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | . |
| **886** | 0 | 2 | male | 27.0 | 0 | 0 | 13.0000 | S | Second | man | Tru |
| **887** | 1 | 1 | female | 19.0 | 0 | 0 | 30.0000 | S | First | woman | Fals |
| **888** | 0 | 3 | female | NaN | 1 | 2 | 23.4500 | S | Third | woman | Fals |
| **889** | 1 | 1 | male | 26.0 | 0 | 0 | 30.0000 | C | First | man | Tru |
| **890** | 0 | 3 | male | 32.0 | 0 | 0 | 7.7500 | Q | Third | man | Tru |

891 rows × 15 columns

In [2]:
```
df1.head()
```

Out[2]:

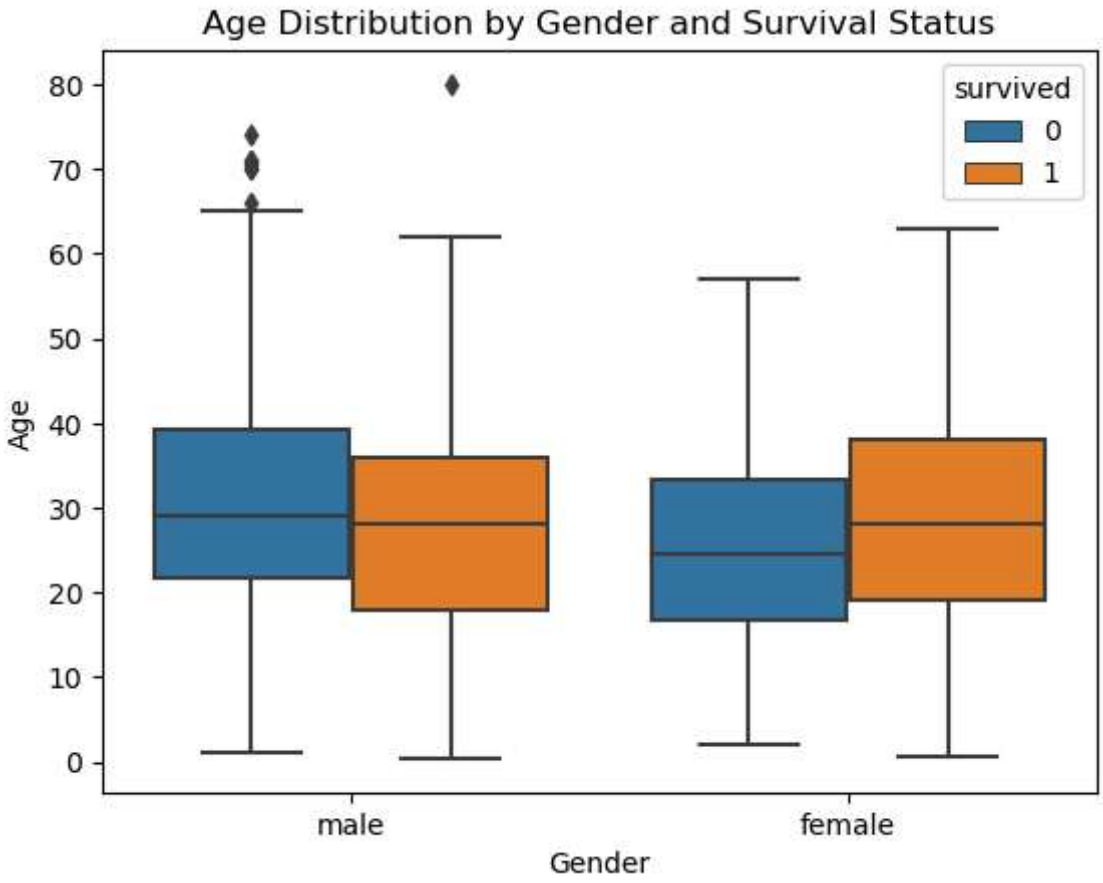
| | survived | pclass | sex | age | sibsp | parch | fare | embarked | class | who | adult_male | d |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 3 | male | 22.0 | 1 | 0 | 7.2500 | S | Third | man | True | N |
| **1** | 1 | 1 | female | 38.0 | 1 | 0 | 71.2833 | C | First | woman | False | |
| **2** | 1 | 3 | female | 26.0 | 0 | 0 | 7.9250 | S | Third | woman | False | N |
| **3** | 1 | 1 | female | 35.0 | 1 | 0 | 53.1000 | S | First | woman | False | |
| **4** | 0 | 3 | male | 35.0 | 0 | 0 | 8.0500 | S | Third | man | True | N |

In [3]:
```
plt.figure(figsize=(10, 6))
```

Out[3]:
```
<Figure size 1000x600 with 0 Axes>
```


```
<Figure size 1000x600 with 0 Axes>
```


In [10]:
```
sns.boxplot(x='sex', y='age', hue='survived', data=df1)
plt.title('Age Distribution by Gender and Survival Status')
plt.xlabel('Gender')
plt.ylabel('Age')
plt.show()
```



In [ ]:

Name : Mohan Kadamabnde

Roll No. : 13212 (TECO-b1)