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'). Write observations on the inference from the above statistics.

In [1]:

```
import seaborn as sns
titanic = sns.load_dataset("titanic")
```

In [2]:

```
titanic
```

Out[2]:

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

891 rows × 15 columns

In [3]:

```
titanic.head(10)
```

Out[3]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_n
0	0	3	male	22.0	1	0	7.2500	S	Third	man	-
1	1	1	female	38.0	1	0	71.2833	C	First	woman	F
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	F
3	1	1	female	35.0	1	0	53.1000	S	First	woman	F
4	0	3	male	35.0	0	0	8.0500	S	Third	man	-
5	0	3	male	NaN	0	0	8.4583	Q	Third	man	-
6	0	1	male	54.0	0	0	51.8625	S	First	man	-
7	0	3	male	2.0	3	1	21.0750	S	Third	child	F
8	1	3	female	27.0	0	2	11.1333	S	Third	woman	F
9	1	2	female	14.0	1	0	30.0708	C	Second	child	F

In [4]:

```
titanic.info
```

Out[4]:

```
<bound method DataFrame.info of      survived  pclass     sex   age  sibsp
parch     fare embarked   class  \
0           0       3    male  22.0      1      0   7.2500        S   Thir
d
1           1       1  female  38.0      1      0  71.2833        C   Firs
t
2           1       3  female  26.0      0      0   7.9250        S   Thir
d
3           1       1  female  35.0      1      0  53.1000        S   Firs
t
4           0       3    male  35.0      0      0   8.0500        S   Thir
d
..        ...     ...     ...   ...    ...    ...      ...      ...
...
886         0       2    male  27.0      0      0  13.0000        S  Secon
d
887         1       1  female  19.0      0      0  30.0000        S   Firs
t
888         0       3  female   NaN      1      2  23.4500        S   Thir
d
889         1       1    male  26.0      0      0  30.0000        C   Firs
t
890         0       3    male  32.0      0      0   7.7500        Q   Thir
d

       who  adult_male deck  embark_town alive  alone
0      man        True  NaN  Southampton    no  False
1    woman       False    C    Cherbourg   yes  False
2    woman       False  NaN  Southampton   yes   True
3    woman       False    C  Southampton   yes  False
4      man        True  NaN  Southampton    no   True
..     ...         ...  ...          ...   ...    ...
886    man        True  NaN  Southampton    no   True
887  woman       False    B  Southampton   yes   True
888  woman       False  NaN  Southampton    no  False
889    man        True    C    Cherbourg   yes   True
890    man        True  NaN   Queenstown    no   True

[891 rows x 15 columns]>
```

In [5]:

```
titanic.describe()
```

Out[5]:

	survived	pclass	age	sibsp	parch	fare
count	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

In [6]:

```
#Custom Columns with all rows
titanic.loc[:,["survived","alive"]]
```

Out[6]:

	survived	alive
0	0	no
1	1	yes
2	1	yes
3	1	yes
4	0	no
...	...	...
886	0	no
887	1	yes
888	0	no
889	1	yes
890	0	no

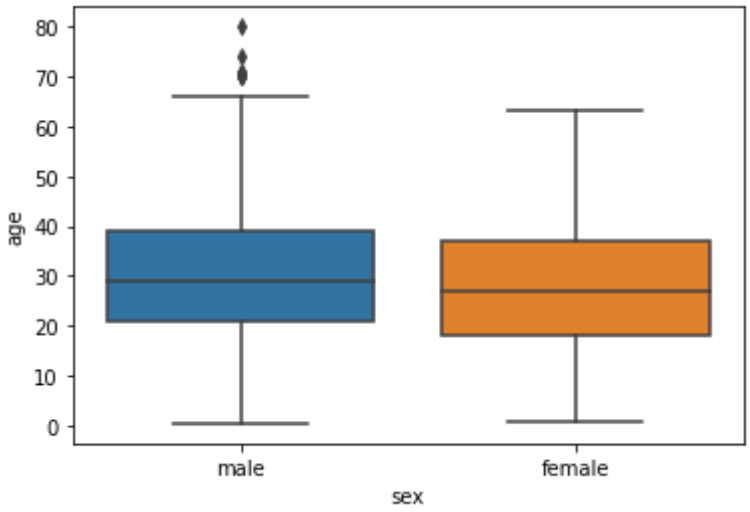
891 rows × 2 columns

In [7]:

```
#Now Plot boxplot
sns.boxplot(x="sex",y="age",data=titanic)
```

Out[7]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f9679321b50>
```



In [8]:

```
sns.boxplot(x="sex",y="age",data=titanic,hue="survived")
```

Out[8]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f96748b3390>
```