

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

```
[ ]: df=pd.read_csv('titanic-train.csv')
```

```
[10]: cols=df.columns
cols
```

```
[10]: Index(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
            'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
            'alive', 'alone'],
            dtype='object')
```

```
[11]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 15 columns):
survived      891 non-null int64
pclass        891 non-null int64
sex           891 non-null object
age           714 non-null float64
sibsp         891 non-null int64
parch         891 non-null int64
fare          891 non-null float64
embarked      889 non-null object
class         891 non-null category
who           891 non-null object
adult_male    891 non-null bool
deck          203 non-null category
embark_town   889 non-null object
alive         891 non-null object
alone         891 non-null bool
dtypes: bool(2), category(2), float64(2), int64(4), object(5)
memory usage: 80.6+ KB
```

```
[12]: df.describe()
```

```
[12]:
```

	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

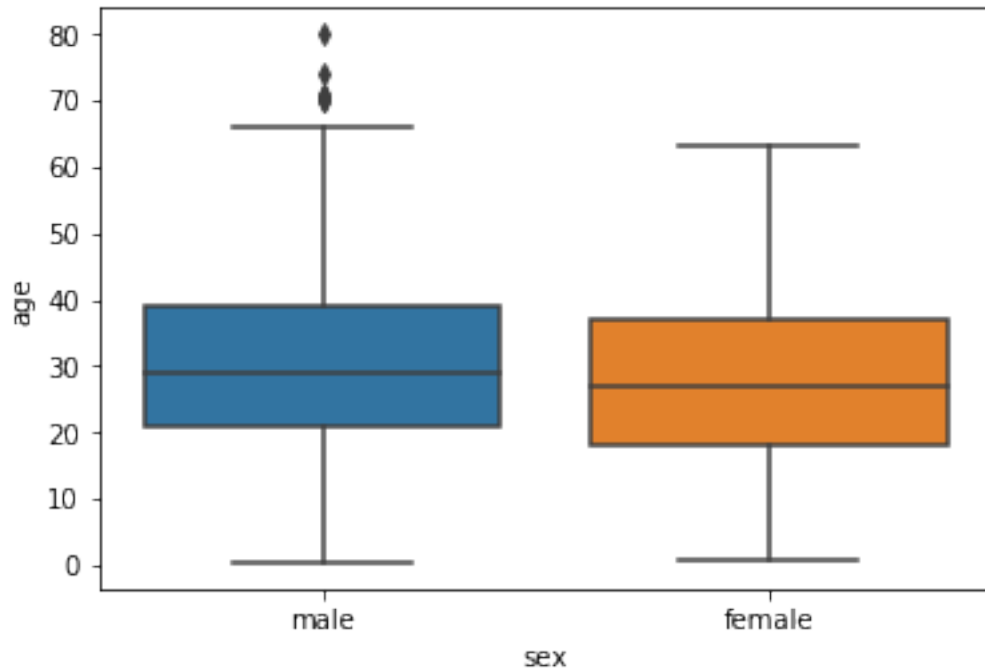
```
[14]: df.isnull().sum()
```

```
[14]: survived      0
pclass            0
sex              0
age             177
sibsp            0
parch            0
fare            0
embarked         2
class            0
who              0
adult_male       0
deck            688
embark_town      2
alive            0
alone            0
dtype: int64
```

```
[15]: sns.boxplot(df['sex'],df['age'])
```

```
C:\Users\dell\Anaconda3\lib\site-packages\seaborn\_decorators.py:43:
FutureWarning: Pass the following variables as keyword args: x, y. From version
0.12, the only valid positional argument will be `data`, and passing other
arguments without an explicit keyword will result in an error or
misinterpretation.
FutureWarning
```

```
[15]: <matplotlib.axes._subplots.AxesSubplot at 0xb655f60>
```



```
[16]: sns.boxplot(df['sex'],df['age'],df['survived'])
```

```
C:\Users\dell\Anaconda3\lib\site-packages\seaborn\_decorators.py:43:
FutureWarning: Pass the following variables as keyword args: x, y, hue. From
version 0.12, the only valid positional argument will be `data`, and passing
other arguments without an explicit keyword will result in an error or
misinterpretation.
FutureWarning
```

```
[16]: <matplotlib.axes._subplots.AxesSubplot at 0xb717e48>
```

