

```

import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
import numpy as np
import warnings
warnings.filterwarnings("ignore")

dataset=sns.load_dataset('titanic')
dataset.head()

```

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	Southampton	yes	False
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton	no	True

Next steps: [Generate code with dataset](#) [New interactive sheet](#)

```
dataset.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 15 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   survived    891 non-null    int64  
 1   pclass      891 non-null    int64  
 2   sex         891 non-null    object  
 3   age         714 non-null    float64 
 4   sibsp       891 non-null    int64  
 5   parch       891 non-null    int64  
 6   fare         891 non-null    float64 
 7   embarked    889 non-null    object  
 8   class        891 non-null    category 
 9   who          891 non-null    object  
 10  adult_male  891 non-null    bool   
 11  deck         203 non-null    category 
 12  embark_town 889 non-null    object  
 13  alive        891 non-null    object  
 14  alone        891 non-null    bool  
dtypes: bool(2), category(2), float64(2), int64(4), object(5)
memory usage: 80.7+ KB

```

```
dataset.describe()
```

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

```

print("Survived: ",dataset['survived'].value_counts()[1])
print("Dead: ",dataset['survived'].value_counts()[0])

```

```

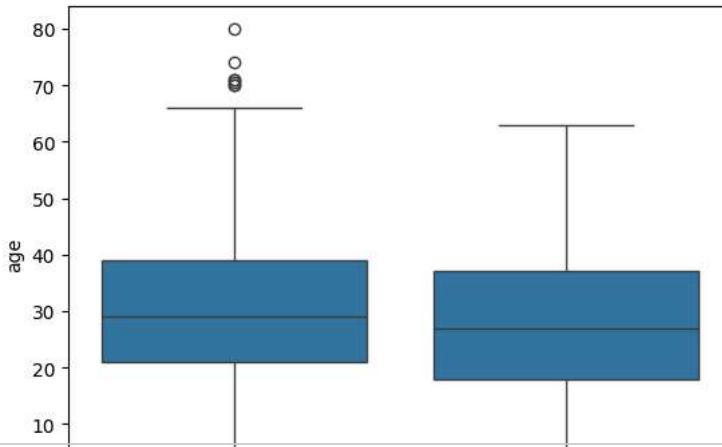
Survived: 342
Dead: 549

```

```

sns.boxplot(x='sex', y='age', data=dataset)
plt.show()

```



```
sns.boxplot(x='sex', y='age', hue='survived', data=dataset)  
plt.show()
```

