Assignment no 08

In [ ]:
```
Aim:
1.Seaborn Library Basics
2. Know your Data
3. Finding patterns of data.
4. Checking how the price of the ticket (column name:
'fare') for each passenger is distributed by
plotting a histogram.
```

In [ ]:
```
Name:Sneha Navgire
Roll no:13246
Batch:B3
```

In [17]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [18]:
```
dataset = sns.load_dataset('titanic')
dataset
```

Out[18]:

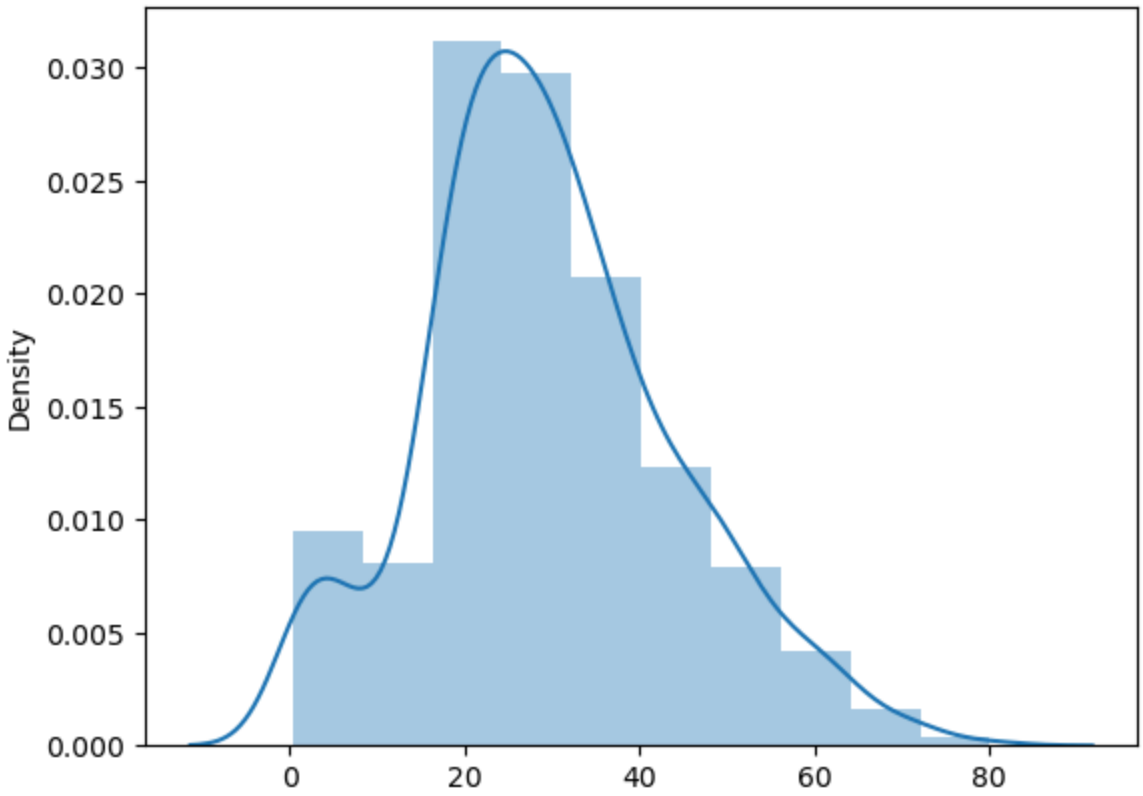
	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	ad
0	0	3	male	22.0	1	0	7.2500	S	Third	man	
1	1	1	female	38.0	1	0	71.2833	C	First	woman	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	
3	1	1	female	35.0	1	0	53.1000	S	First	woman	
4	0	3	male	35.0	0	0	8.0500	S	Third	man	
...	...	...	...	...	...	...	...	...	...	...	
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 [19]:
```
sns.distplot(x = dataset['age'], bins = 10)
```

```
C:\Users\Monuu\AppData\Local\Temp\ipykernel_6848\3970345142.py:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(x = dataset['age'], bins = 10)
```

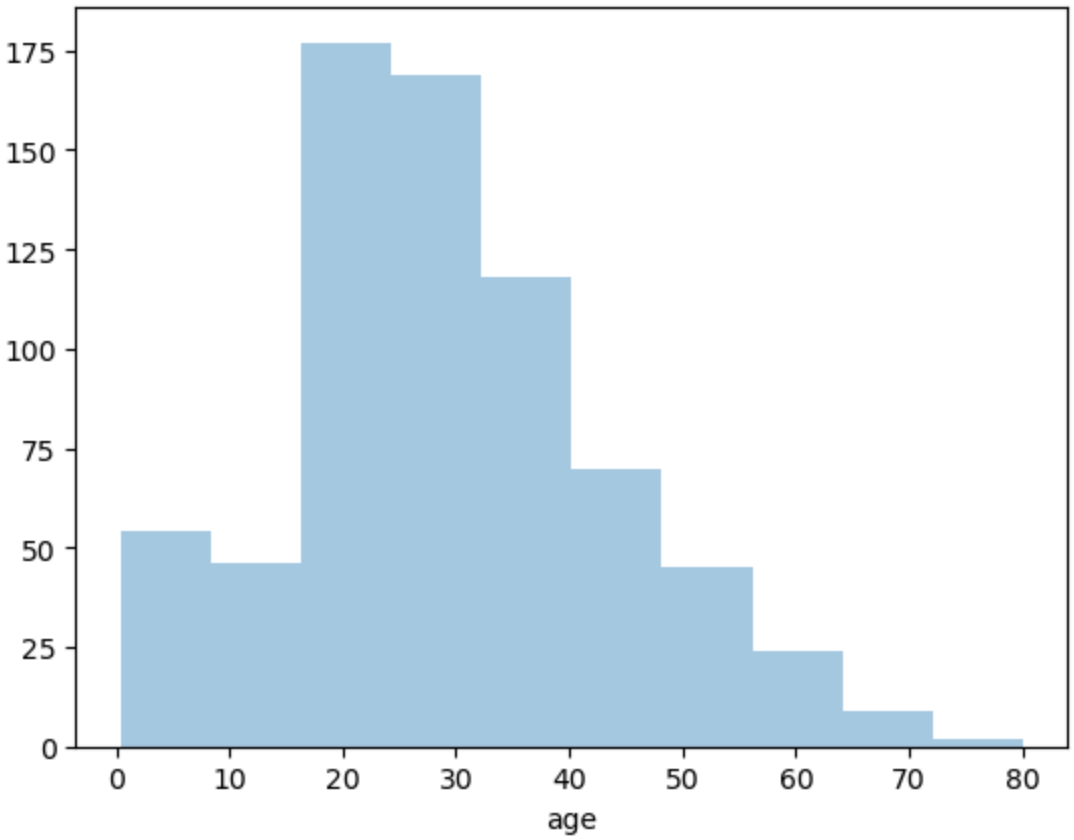
Out[19]: <Axes: ylabel='Density'>



In [20]:
```
sns.distplot(dataset['age'], bins = 10,kde=False)
```

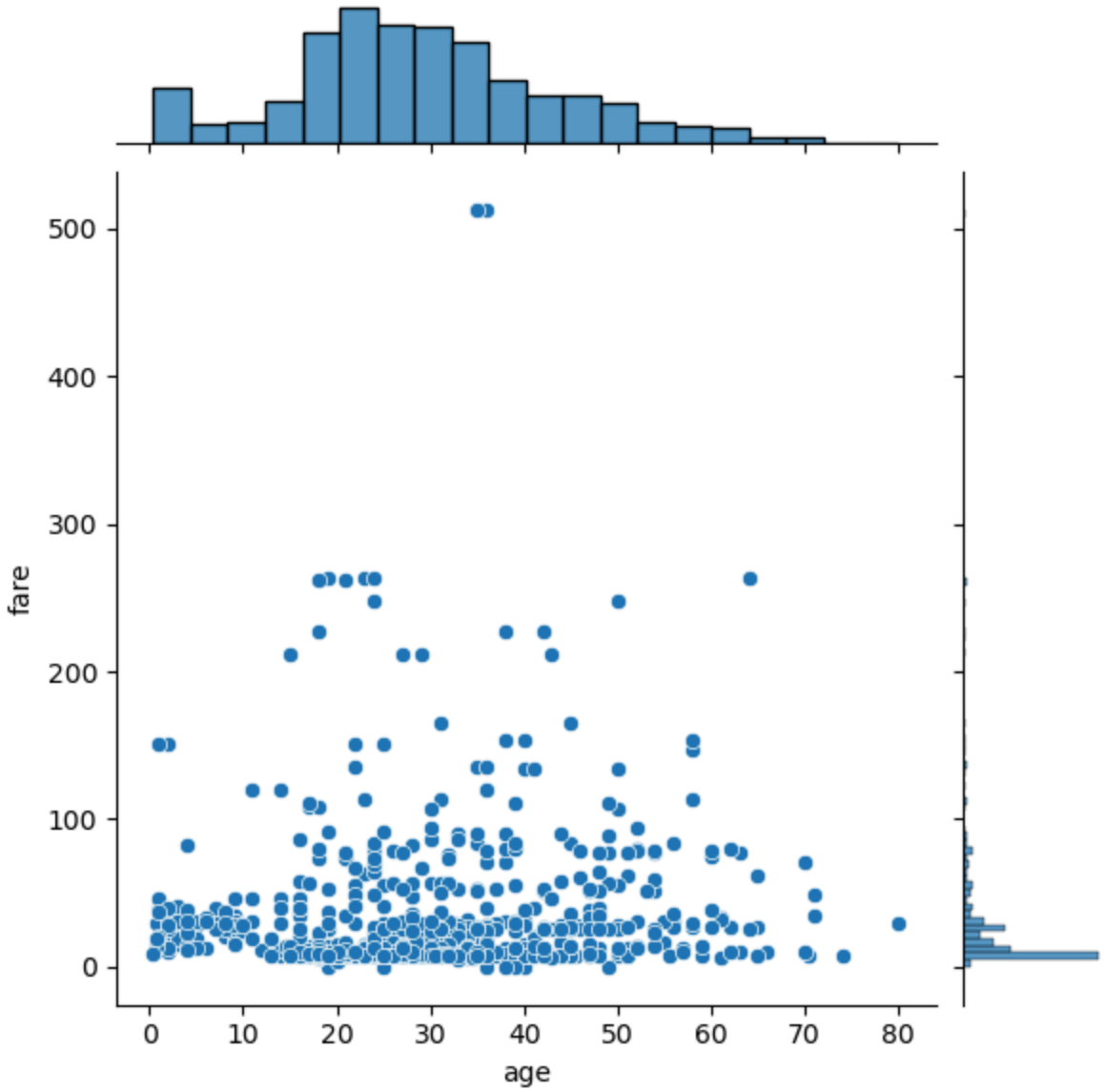
```
C:\Users\Monuu\AppData\Local\Temp\ipykernel_6848\3517108427.py:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dataset['age'], bins = 10,kde=False)
```

Out[20]: <Axes: xlabel='age'>



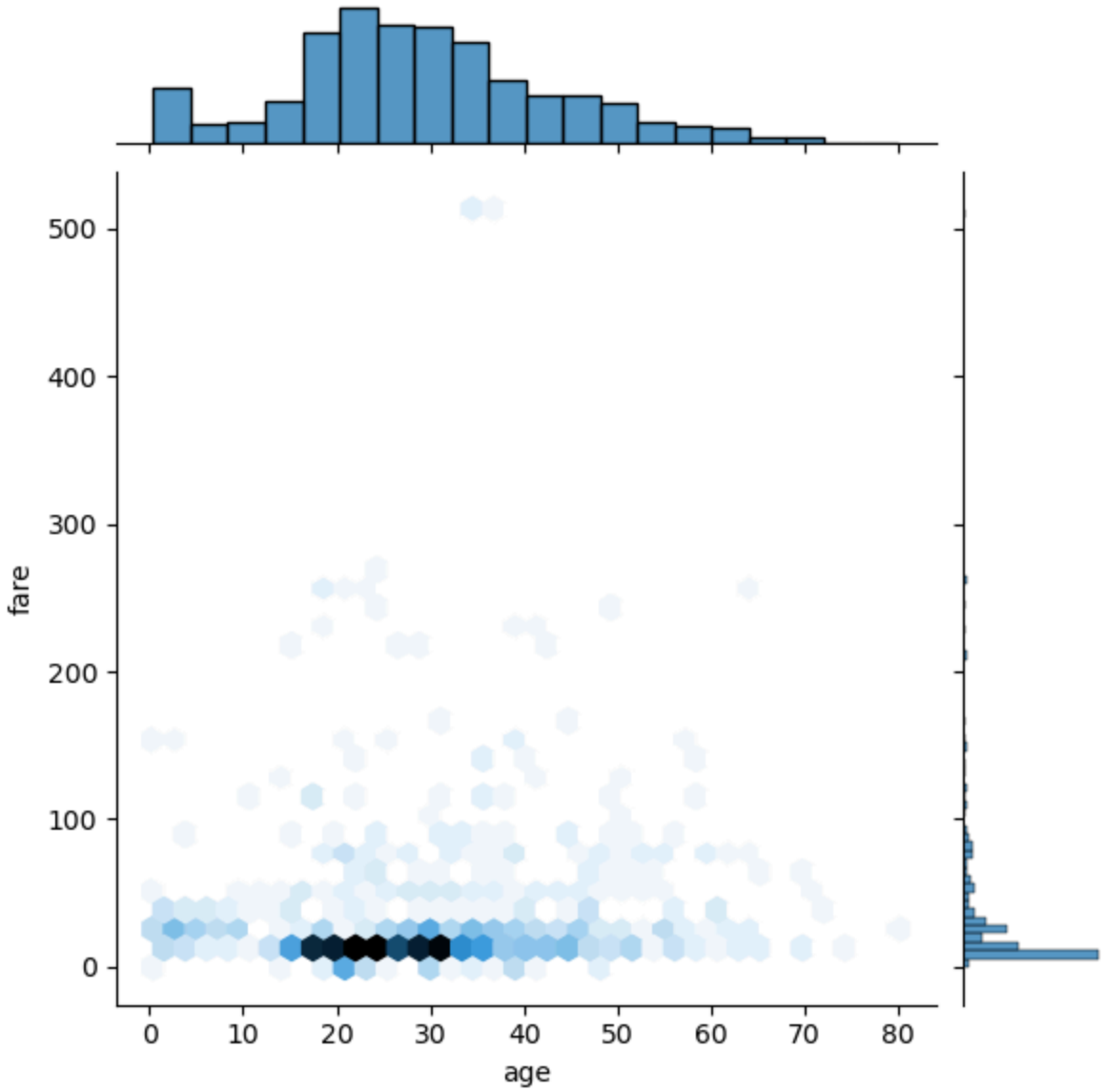
In [41]:
```
sns.jointplot(x = dataset['age'], y = dataset['fare'], kind =
'scatter')
```

Out[41]: <seaborn.axisgrid.JointGrid at 0x1d6415092e0>



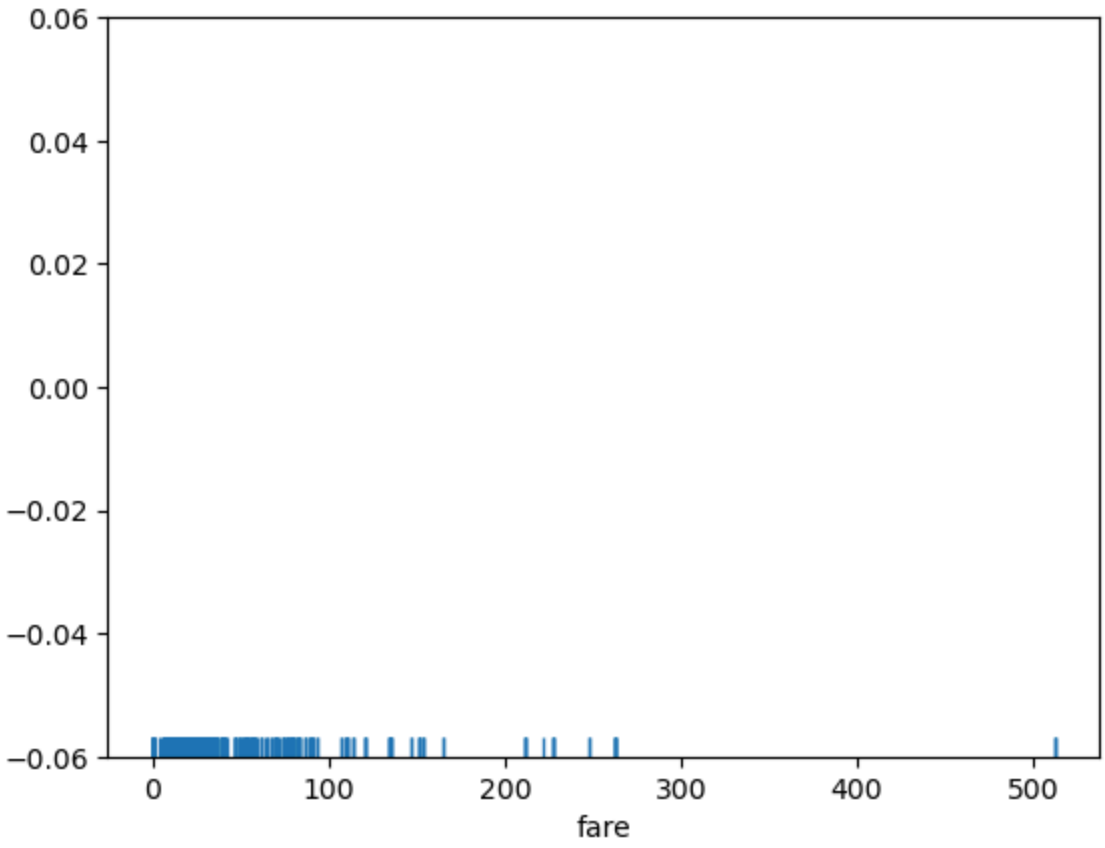
In [43]:
```
sns.jointplot(x = dataset['age'], y = dataset['fare'], kind = 'hex')
```

Out[43]: <seaborn.axisgrid.JointGrid at 0x1d646e82c60>



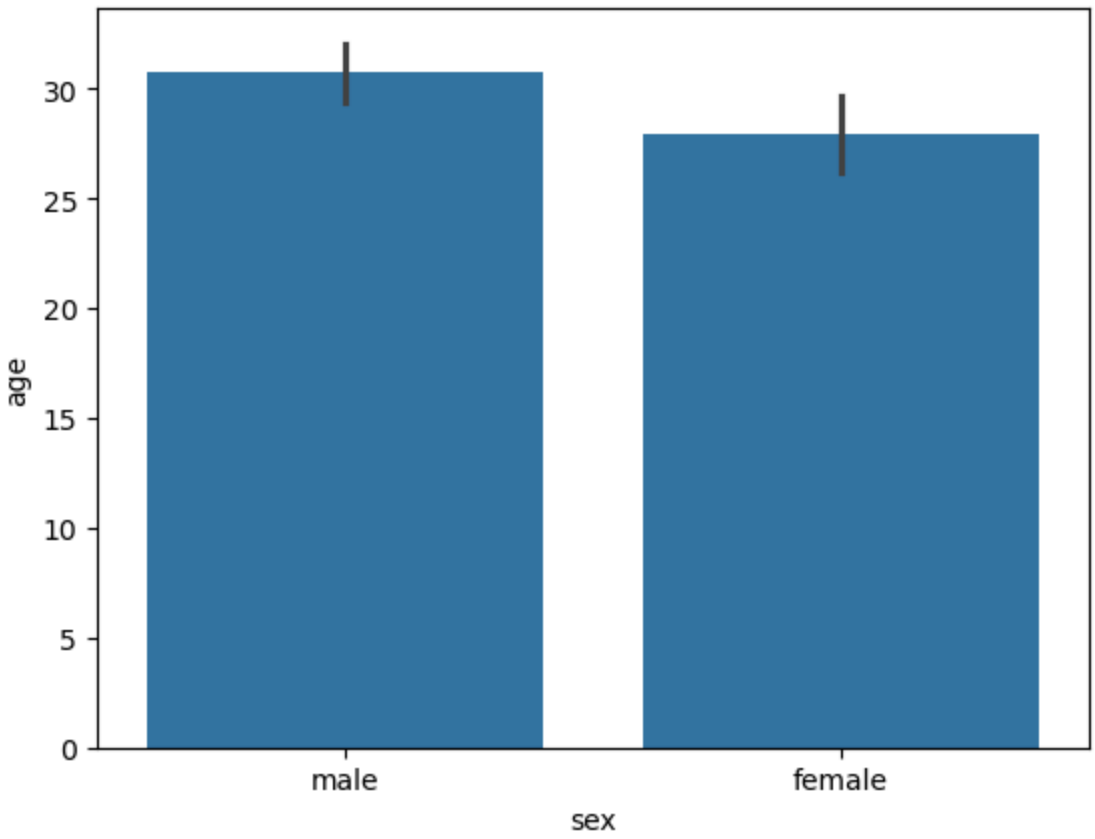
In [44]:
```
sns.rugplot(dataset['fare'])
```

Out[44]: <Axes: xlabel='fare'>



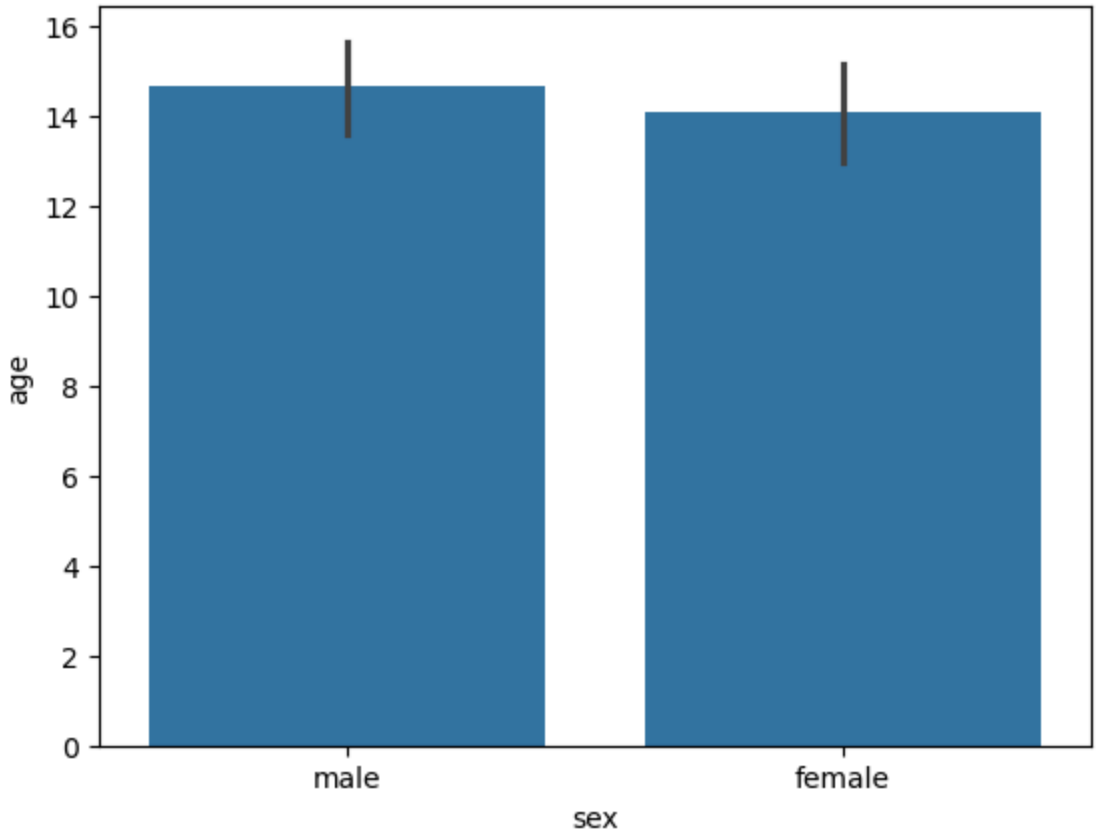
In [46]:
```
sns.barplot(x='sex', y='age', data=dataset)
```

Out[46]: <Axes: xlabel='sex', ylabel='age'>



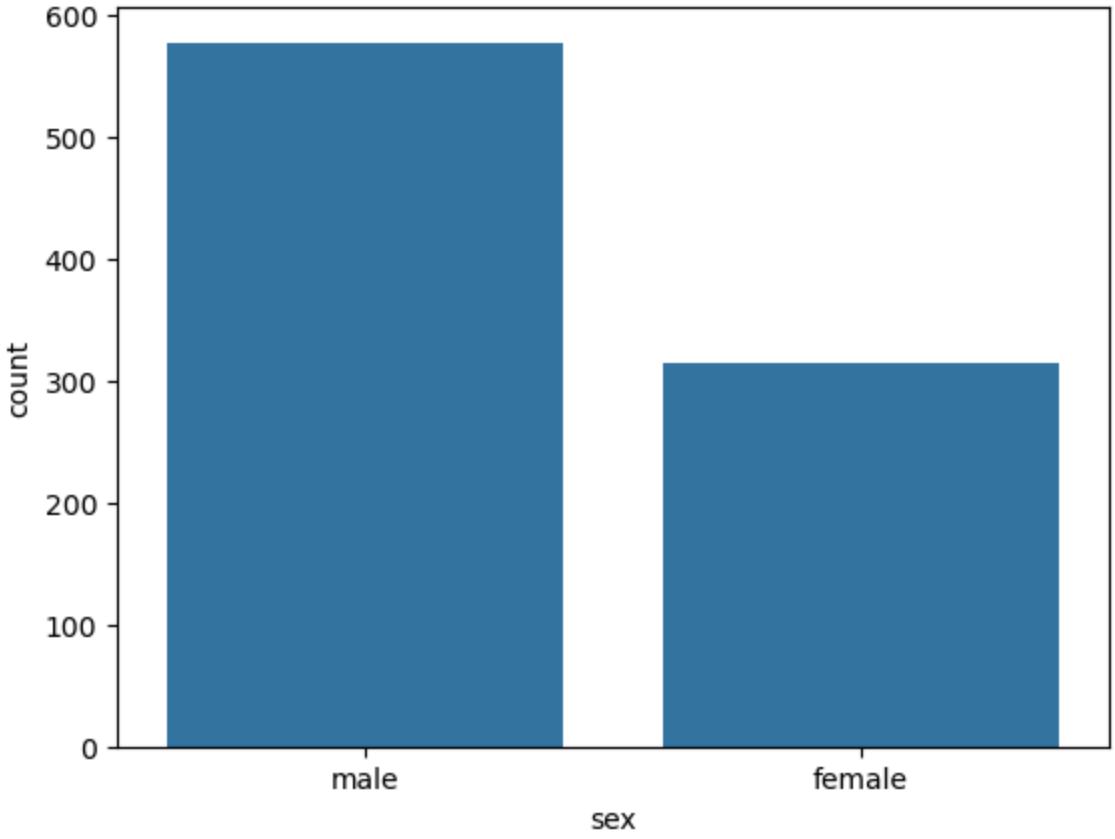
In [49]:
```
sns.barplot(x='sex', y='age', data=dataset, estimator=np.std)
```

Out[49]: <Axes: xlabel='sex', ylabel='age'>



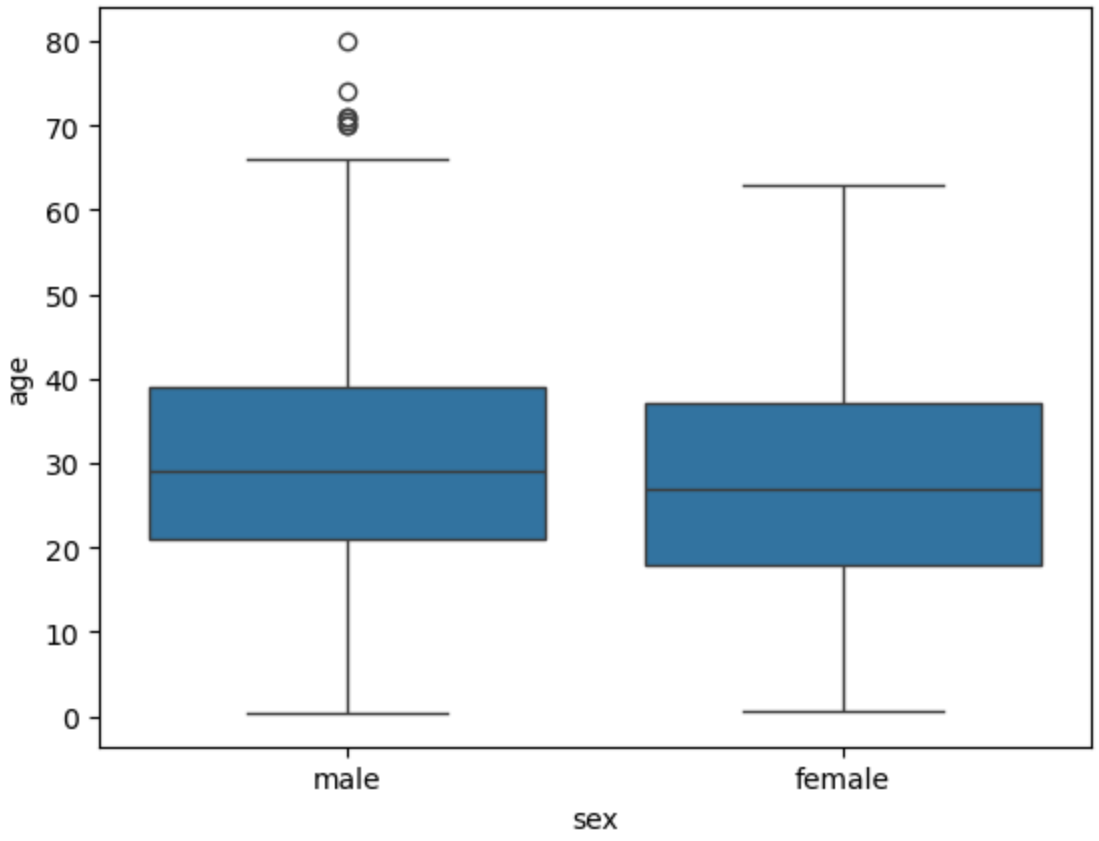
In [51]:
```
sns.countplot(x='sex', data=dataset)
```

Out[51]: <Axes: xlabel='sex', ylabel='count'>



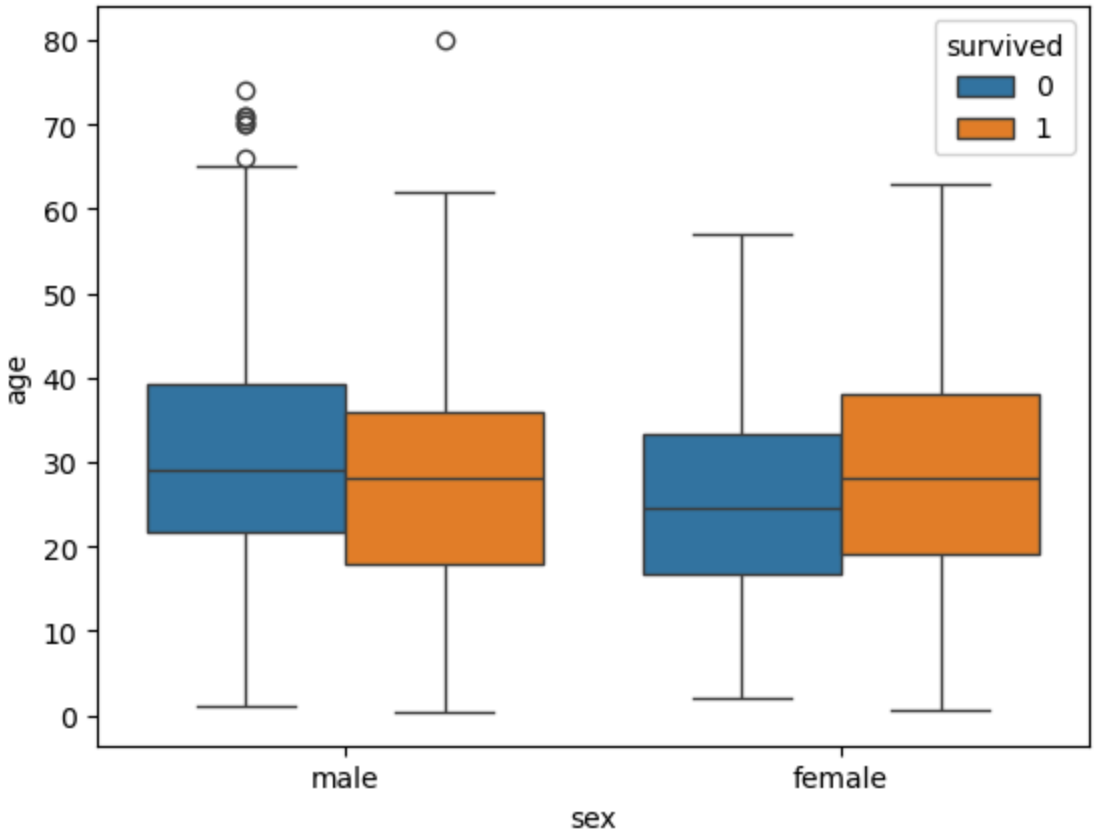
In [53]:
```
sns.boxplot(x='sex', y='age', data=dataset)
```

Out[53]: <Axes: xlabel='sex', ylabel='age'>



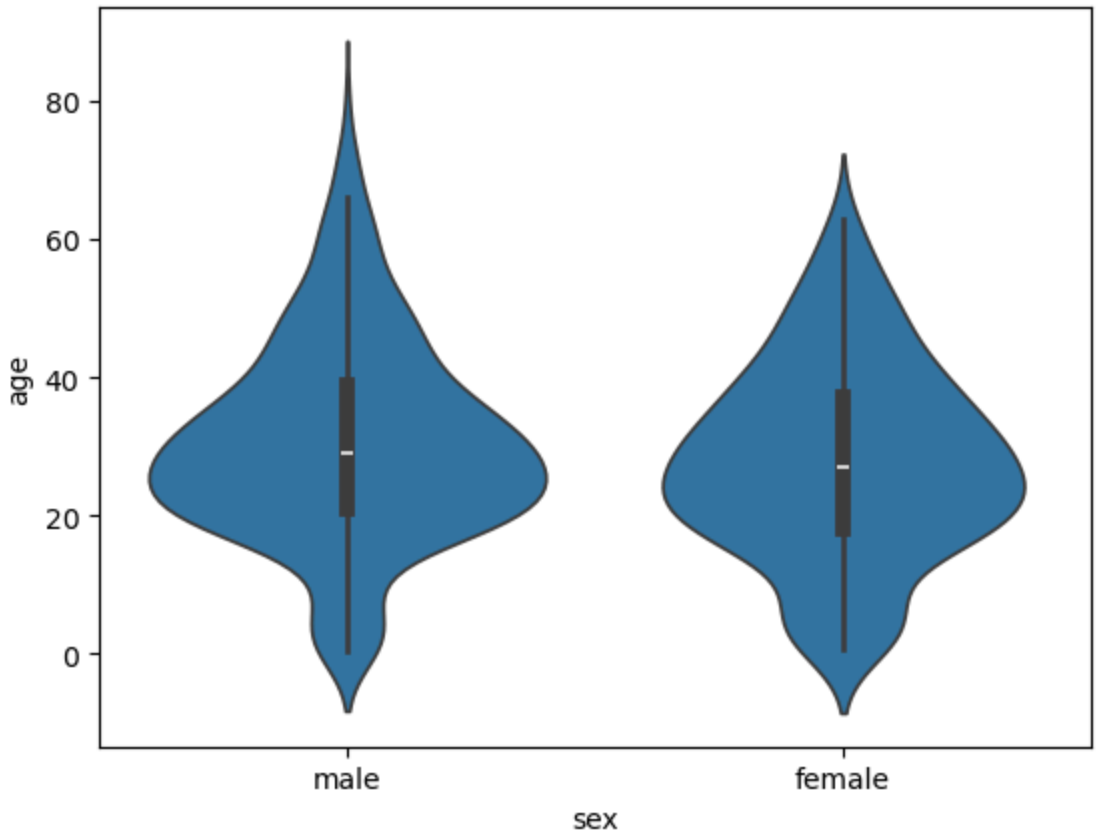
In [55]:
```
sns.boxplot(x='sex', y='age', data=dataset, hue="survived")
```

Out[55]:
```
<Axes: xlabel='sex', ylabel='age'>
```



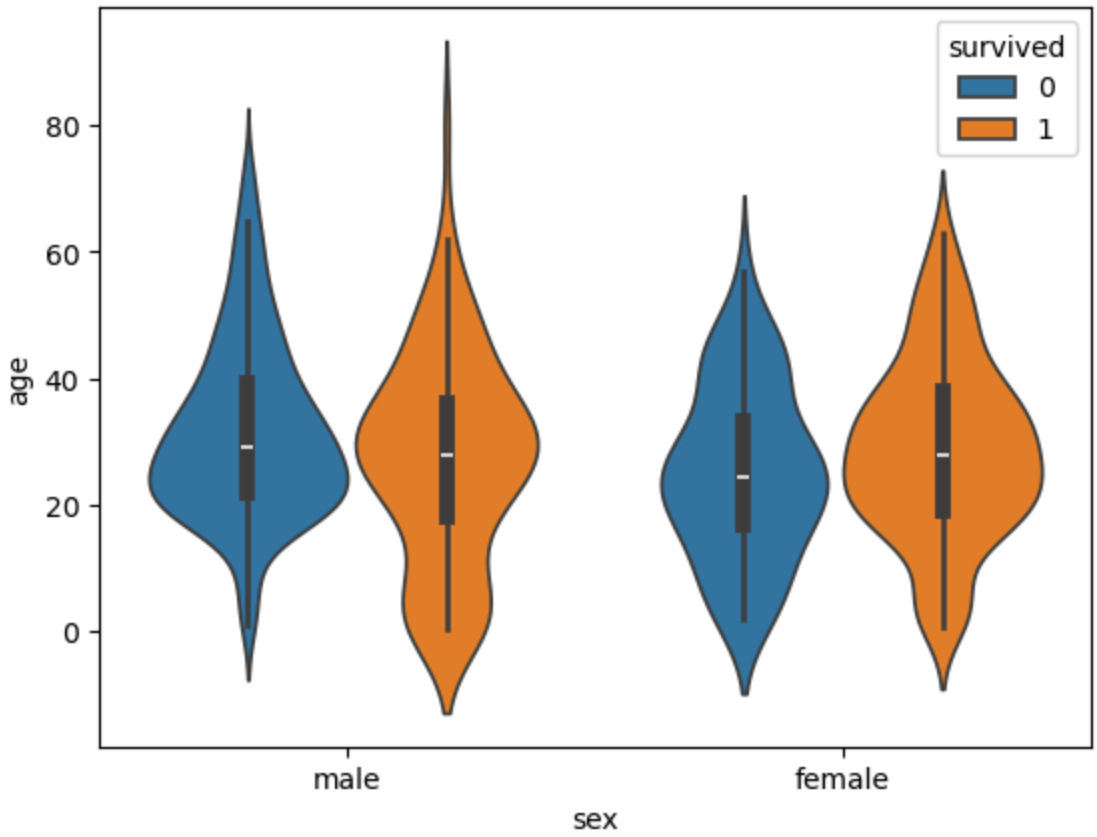
In [57]:
```
sns.violinplot(x='sex', y='age', data=dataset)
```

Out[57]:
```
<Axes: xlabel='sex', ylabel='age'>
```



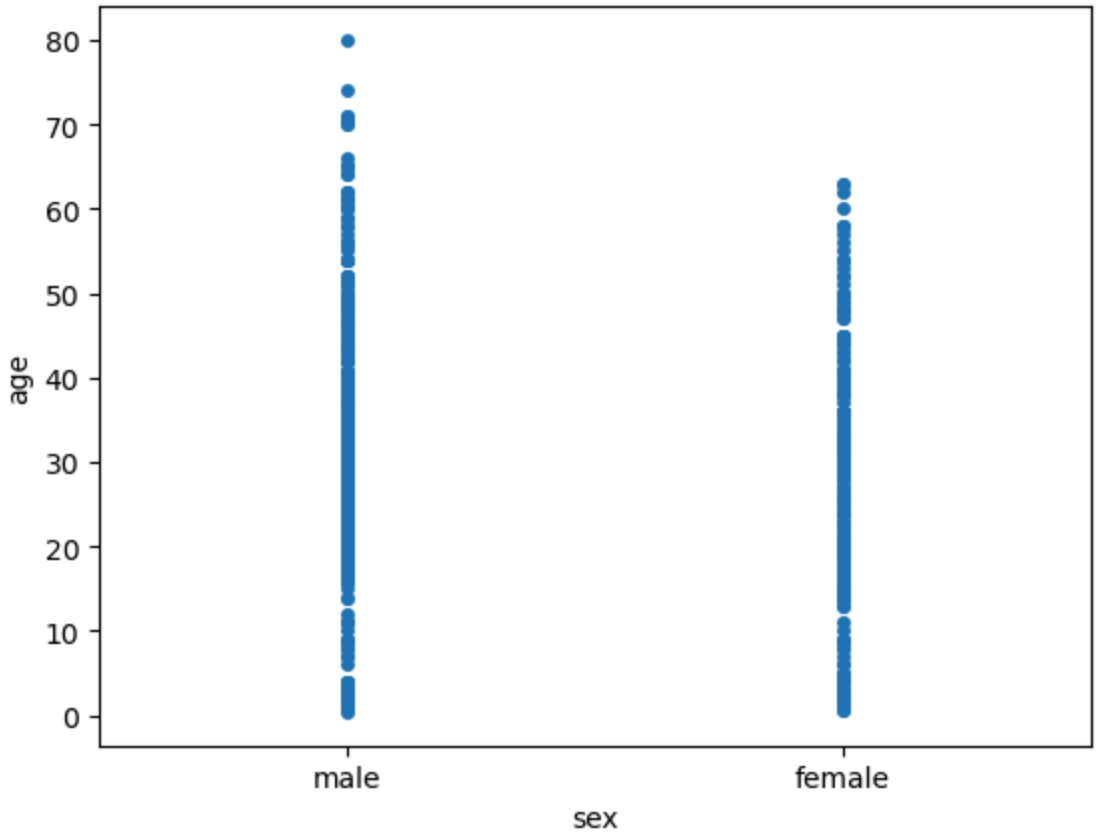
In [59]:
```
sns.violinplot(x='sex', y='age', data=dataset, hue='survived')
```

Out[59]: <Axes: xlabel='sex', ylabel='age'>



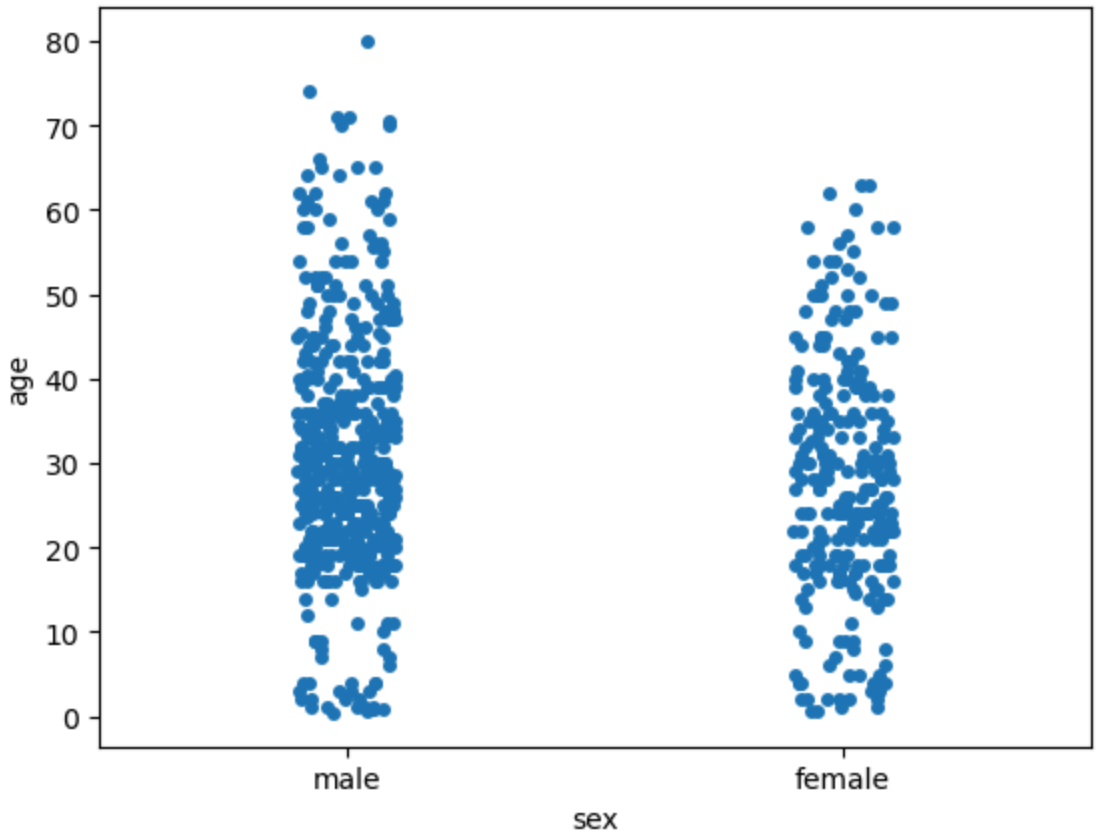
In [61]:
```
sns.stripplot(x='sex', y='age', data=dataset, jitter=False)
```

Out[61]: <Axes: xlabel='sex', ylabel='age'>



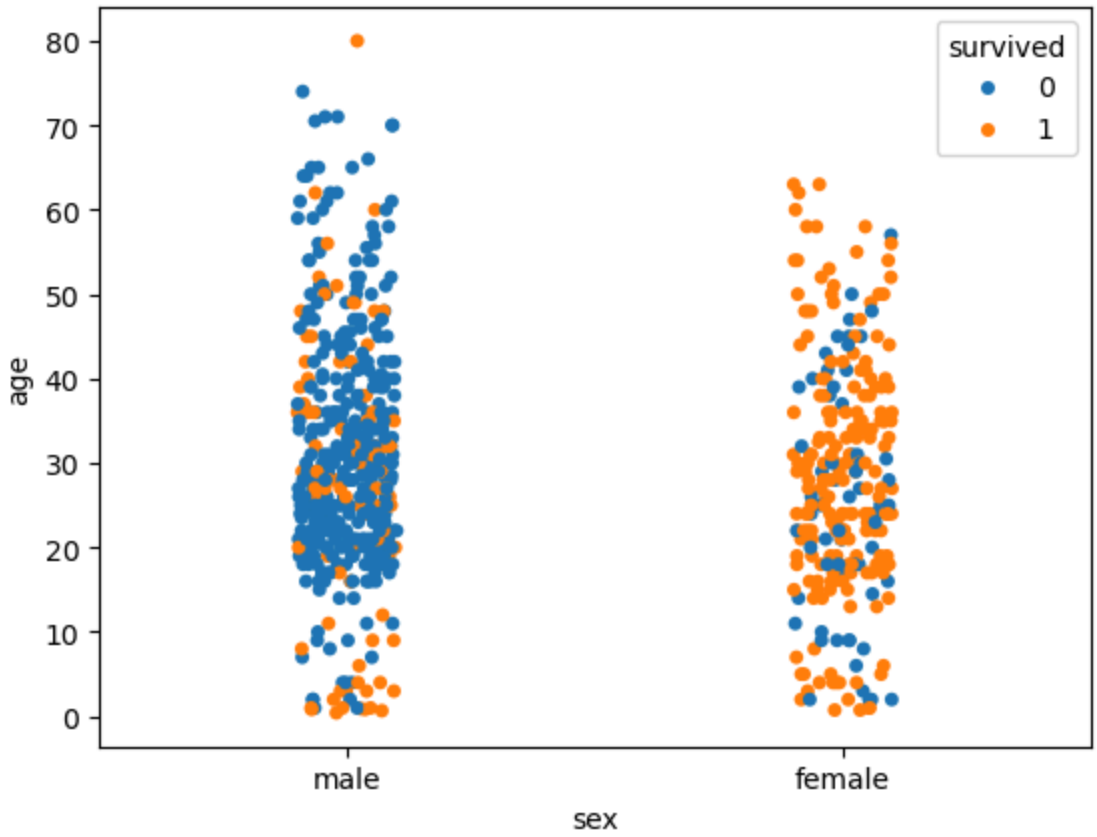
In [63]:
```
sns.stripplot(x='sex', y='age', data=dataset, jitter=True)
```

Out[63]: <Axes: xlabel='sex', ylabel='age'>



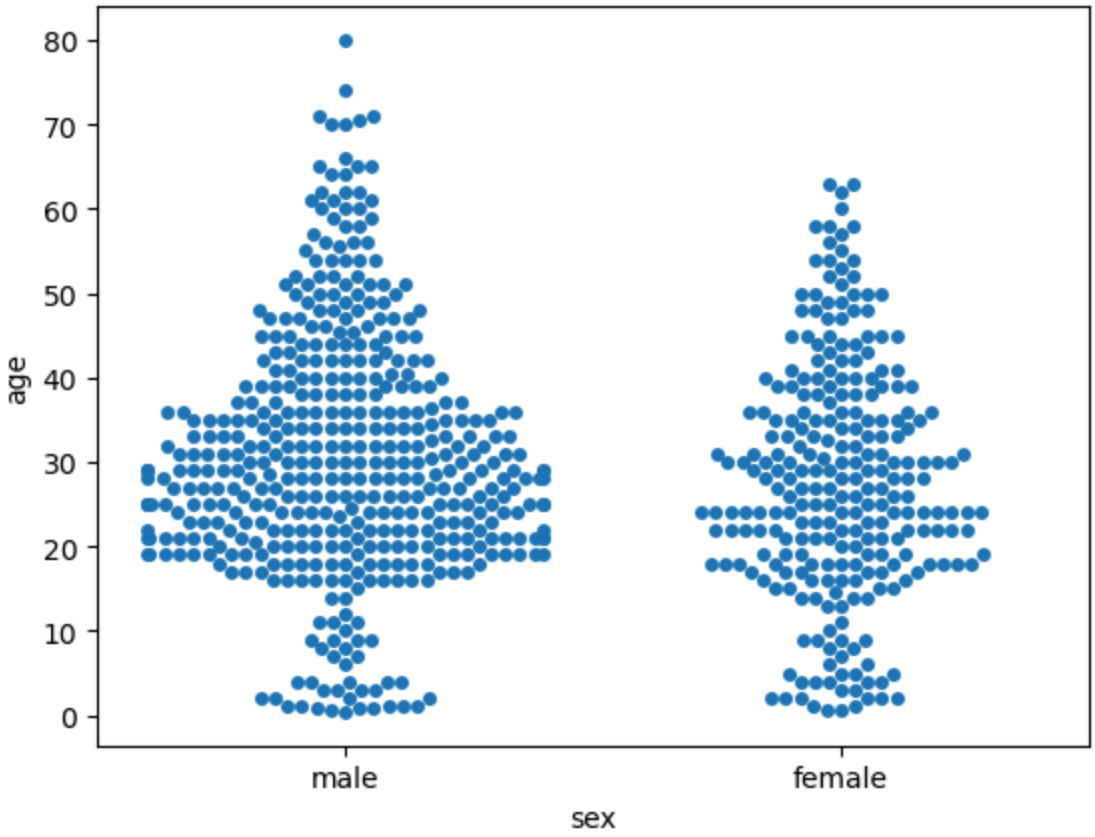
In [65]:
```
sns.stripplot(x='sex', y='age', data=dataset, jitter=True, hue='survived')
```

Out[65]: <Axes: xlabel='sex', ylabel='age'>



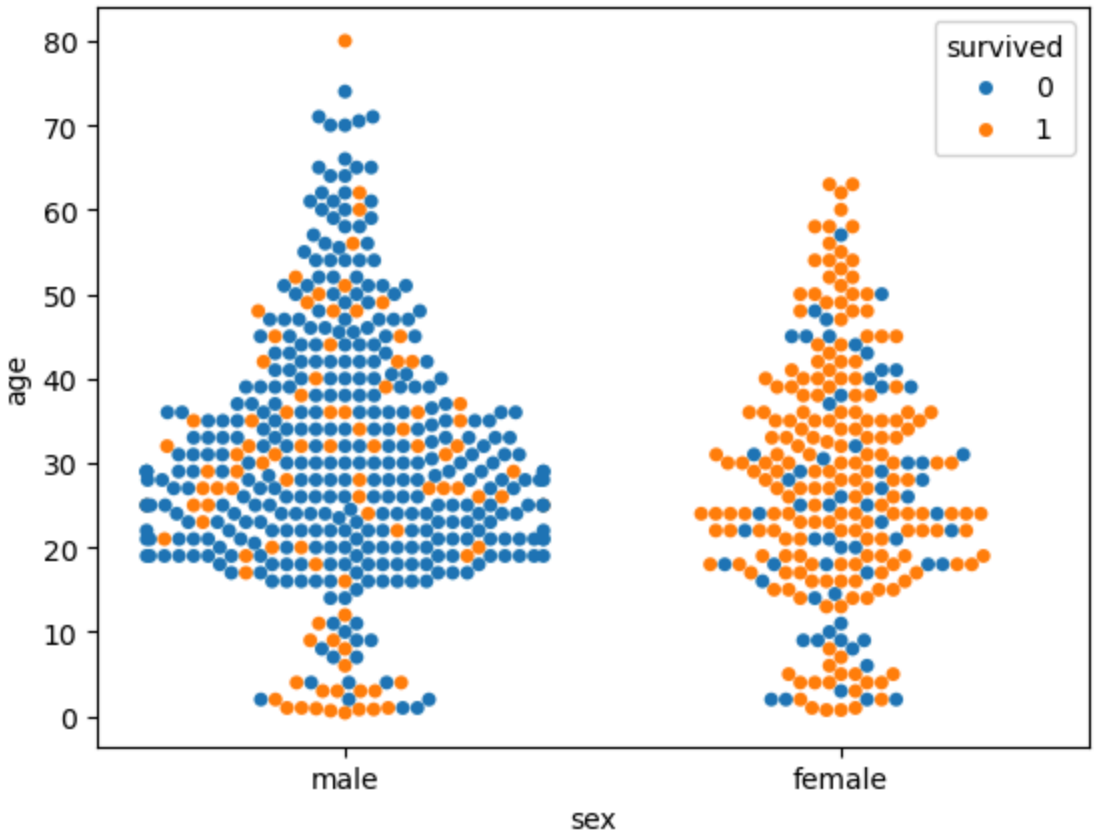
In [67]:
```
sns.swarmplot(x='sex', y='age', data=dataset)
```

Out[67]: <Axes: xlabel='sex', ylabel='age'>



In [68]:
```
sns.swarmplot(x='sex', y='age', data=dataset, hue='survived')
```

Out[68]: <Axes: xlabel='sex', ylabel='age'>



In [70]:
```
dataset = sns.load_dataset('titanic')
dataset.head()
```

Out[70]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_m
0	0	3	male	22.0	1	0	7.2500	S	Third	man	T
1	1	1	female	38.0	1	0	71.2833	C	First	woman	Fa
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	Fa
3	1	1	female	35.0	1	0	53.1000	S	First	woman	Fa
4	0	3	male	35.0	0	0	8.0500	S	Third	man	T

In [73]:
```
# Select only numeric columns
numeric_dataset = dataset.select_dtypes(include=['number'])
```

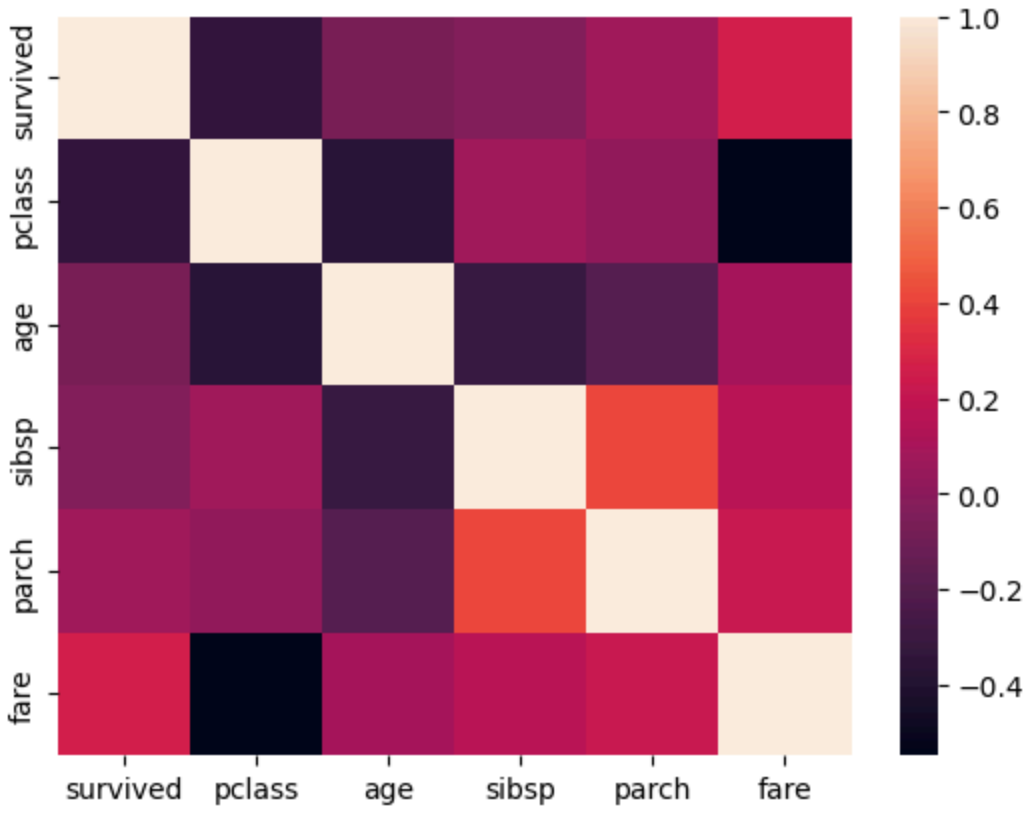
In [75]:
```
numeric_dataset.corr()
```

Out[75]:

	survived	pclass	age	sibsp	parch	fare
survived	1.000000	-0.338481	-0.077221	-0.035322	0.081629	0.257307
pclass	-0.338481	1.000000	-0.369226	0.083081	0.018443	-0.549500
age	-0.077221	-0.369226	1.000000	-0.308247	-0.189119	0.096067
sibsp	-0.035322	0.083081	-0.308247	1.000000	0.414838	0.159651
parch	0.081629	0.018443	-0.189119	0.414838	1.000000	0.216225
fare	0.257307	-0.549500	0.096067	0.159651	0.216225	1.000000

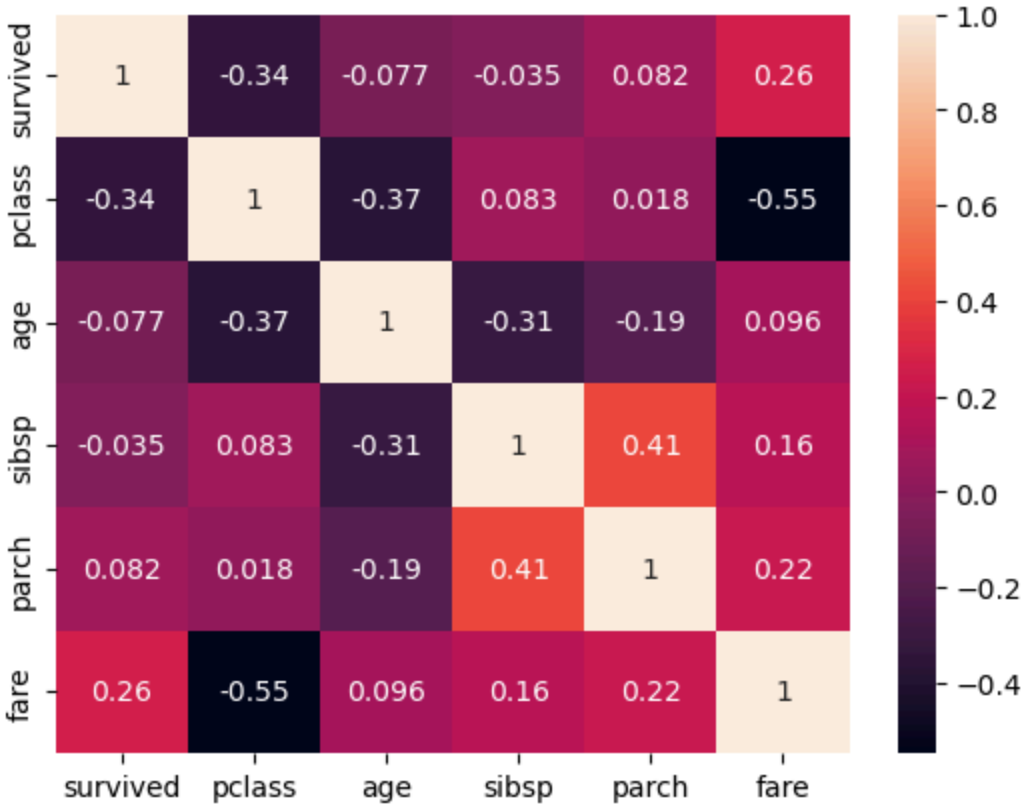
In [77]:
```
corr =numeric_dataset.corr()
sns.heatmap(corr)
```

Out[77]: <Axes: >



In [79]:
```
corr =numeric_dataset.corr()
sns.heatmap(corr, annot=True)
```

Out[79]: <Axes: >



In [85]:
```
sns.histplot(dataset['fare'], kde=False, bins=10)
```

Out[85]: <Axes: xlabel='fare', ylabel='Count'>