8) Data Visualization I

1. Use the inbuilt dataset 'titanic'. The dataset contains 891 rows and contains information about the passengers who boarded the unfortunate Titanic ship. Use the Seaborn library to see if we can find any patterns in the data.

2. Write a code to check how the price of the ticket (column name: 'fare') for each passenger is distributed by plotting a histogram

In [63]:
```python
import pandas as pd
```

In [64]:
```python
df=pd.read_csv("Titanic.csv")
```

In [65]:
```python
df
```

Out[65]:

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

891 rows × 10 columns

In [66]:
```python
import matplotlib.pyplot as plt
import seaborn as sns
```
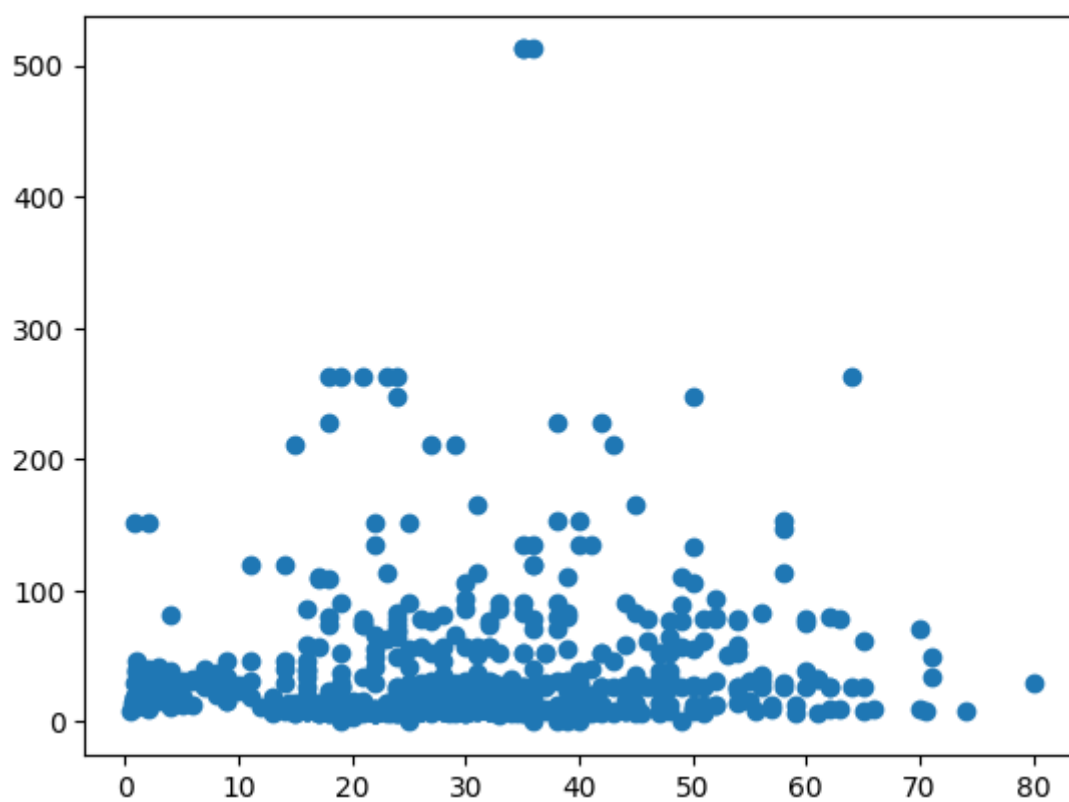
In [67]:
```python
plt.hist(df['fare'])
```

Out[67]:
```
(array([732., 106.,  31.,   2.,  11.,   6.,   0.,   0.,   0.,   3.]),
 array([  0.     ,  51.23292, 102.46584, 153.69876, 204.93168, 256.1646 ,
        307.39752, 358.63044, 409.86336, 461.09628, 512.3292 ]),
 <BarContainer object of 10 artists>)
```
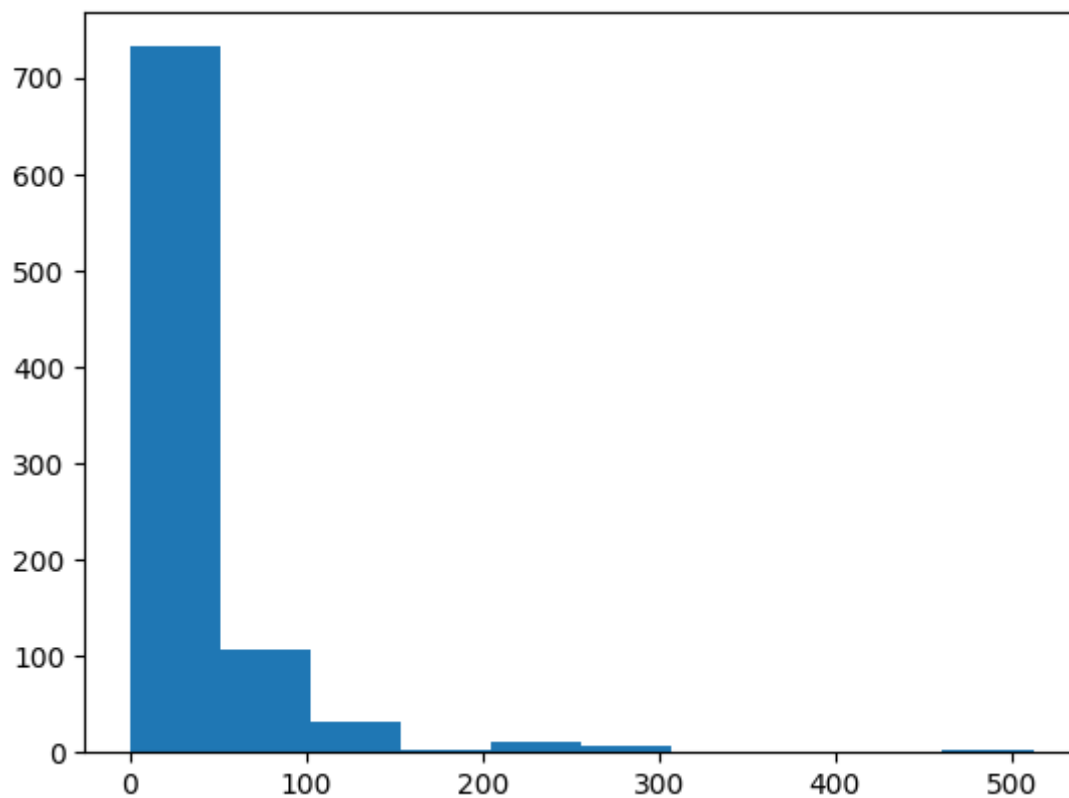
In [68]:
```python
plt.scatter(df['age'],df['fare'])
```

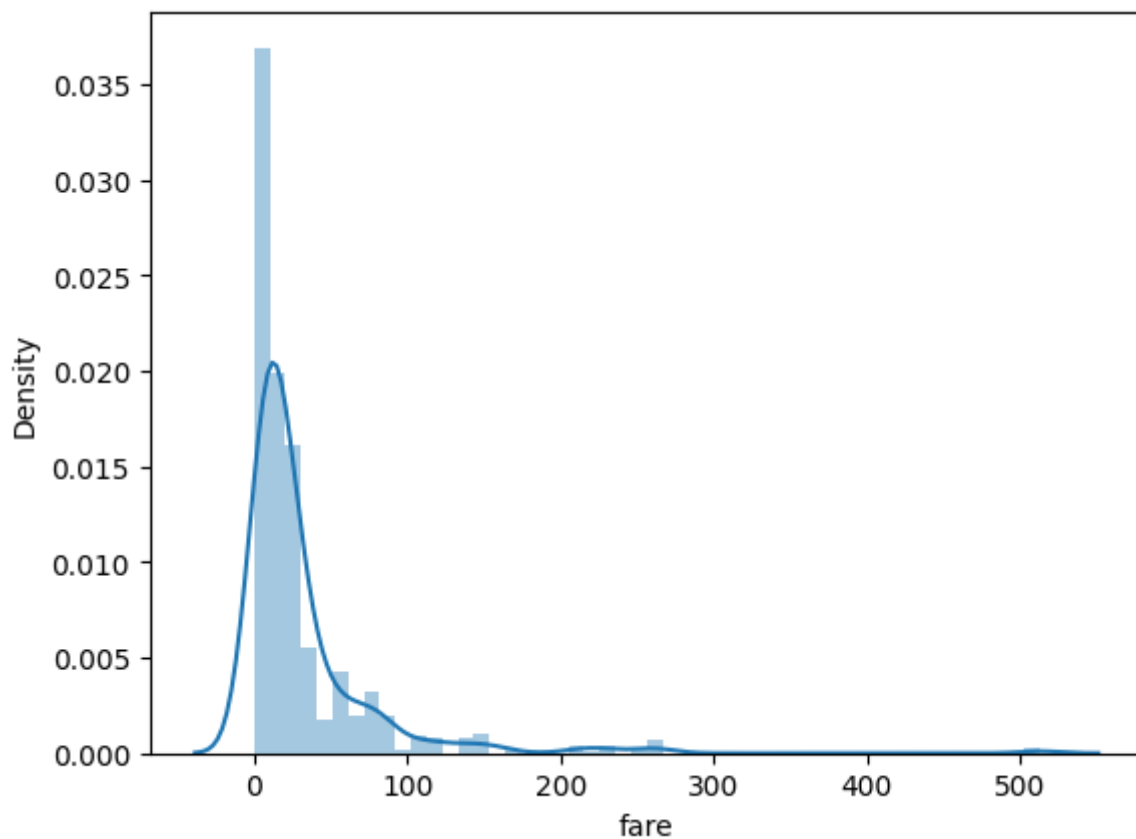Out[68]: <matplotlib.collections.PathCollection at 0x16061019b40>

In [69]:
```python
sns.distplot(df['fare'])
```

C:\Users\me\AppData\Local\Temp\ipykernel_12108\1195996103.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
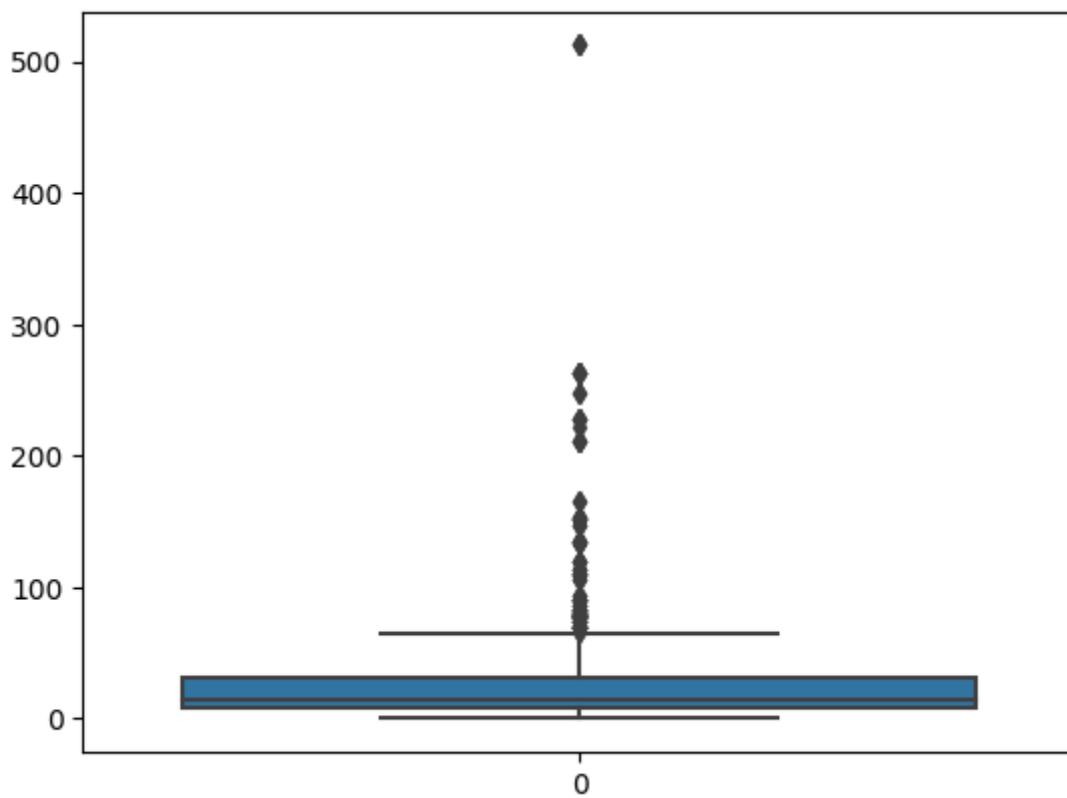similar flexibility) or `histplot` (an axes-level function for histograms).

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

```
sns.distplot(df['fare'])
```

Out[69]:   `<AxesSubplot: xlabel='fare', ylabel='Density'>`

In [70]:
```
sns.boxplot(df['fare'])
```

Out[70]:   `<AxesSubplot: >`

In [71]:
```python
sns.jointplot(x='age', y='fare', data=df)
```

Out[71]:   <seaborn.axisgrid.JointGrid at 0x1606e658eb0>
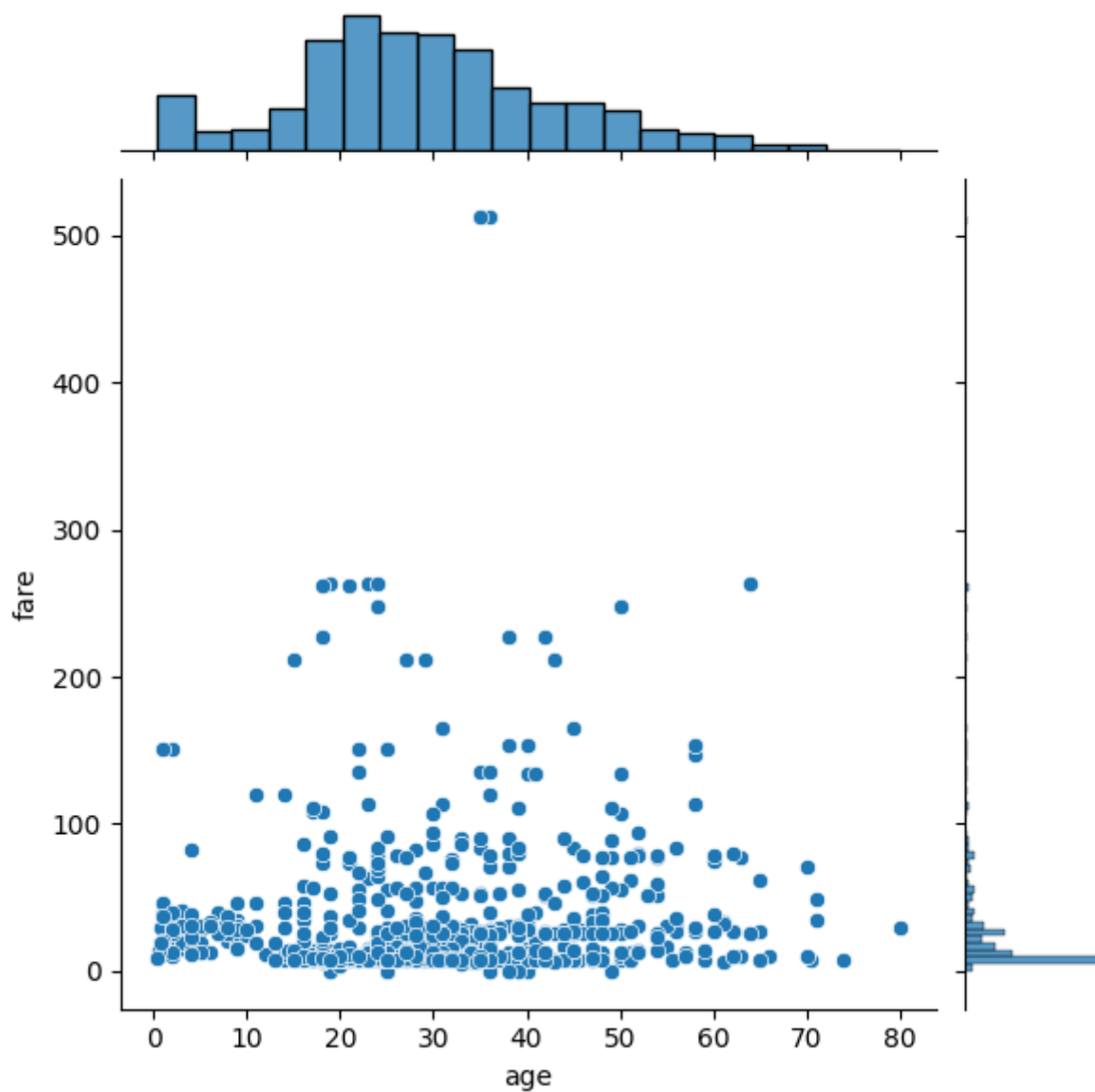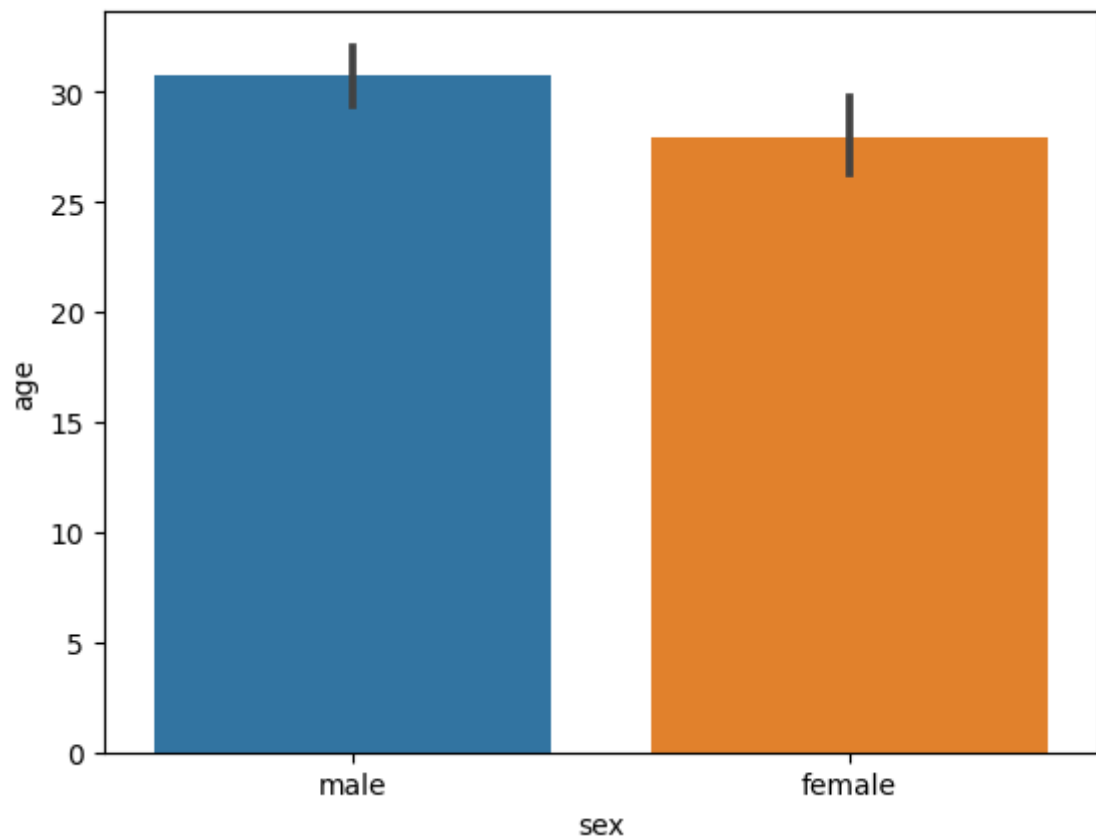
```
In [72]:  sns.pairplot(df, hue='sex')
```

```
Out[72]:  <seaborn.axisgrid.PairGrid at 0x1606dd25f60>
```
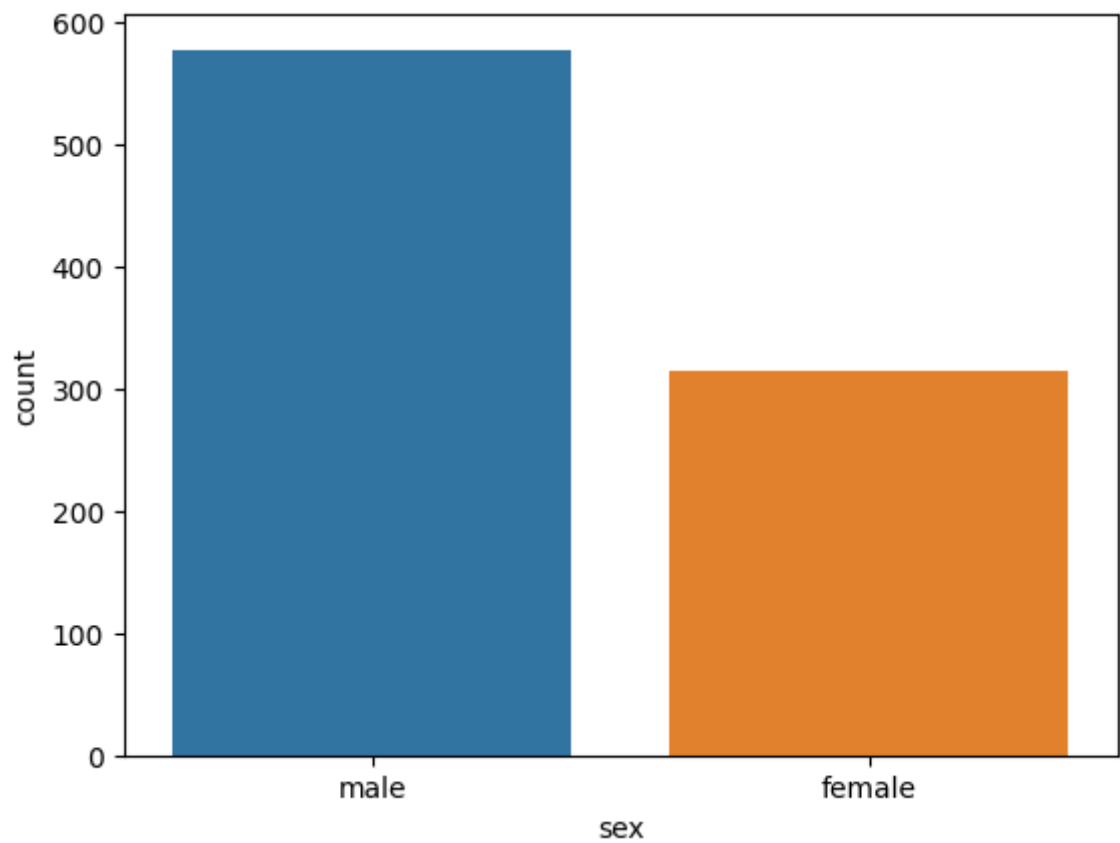
In [73]:
```python
sns.rugplot(df['fare'])
```

Out[73]:   <AxesSubplot: xlabel='fare'>
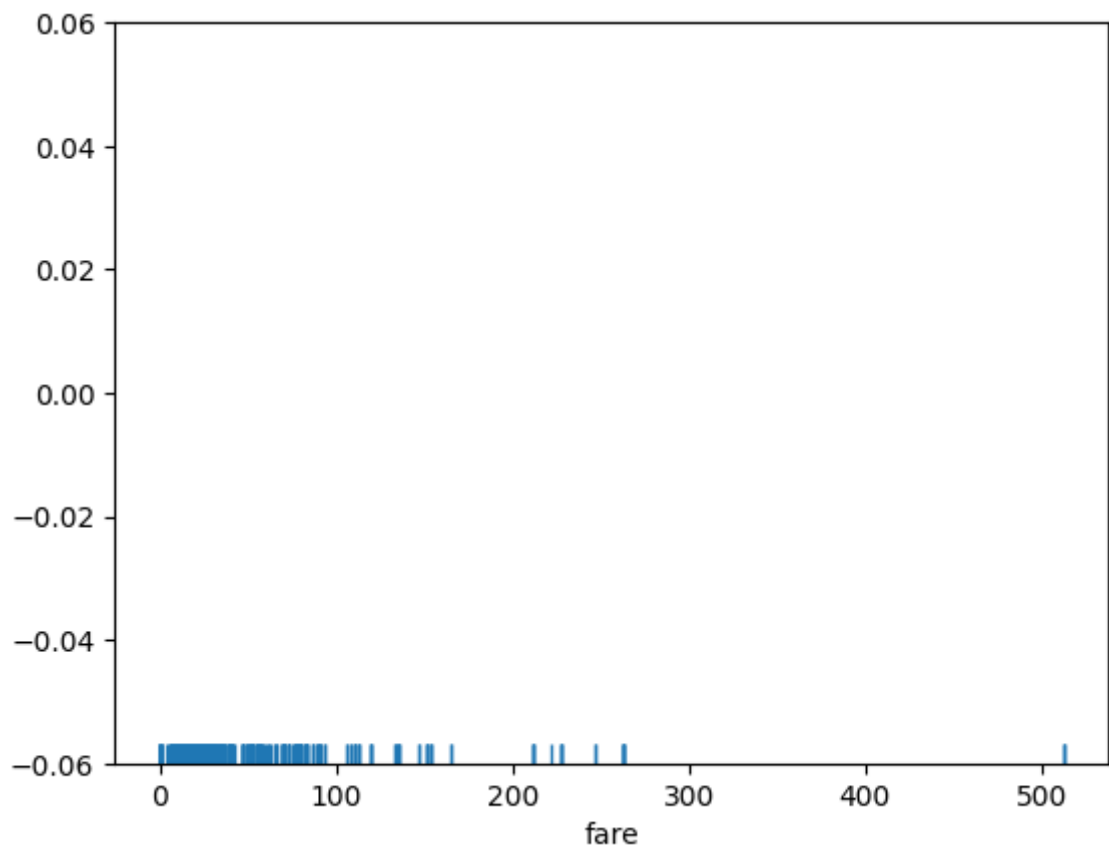
In [74]:
```python
sns.barplot(x='sex', y='age', data=df)
```
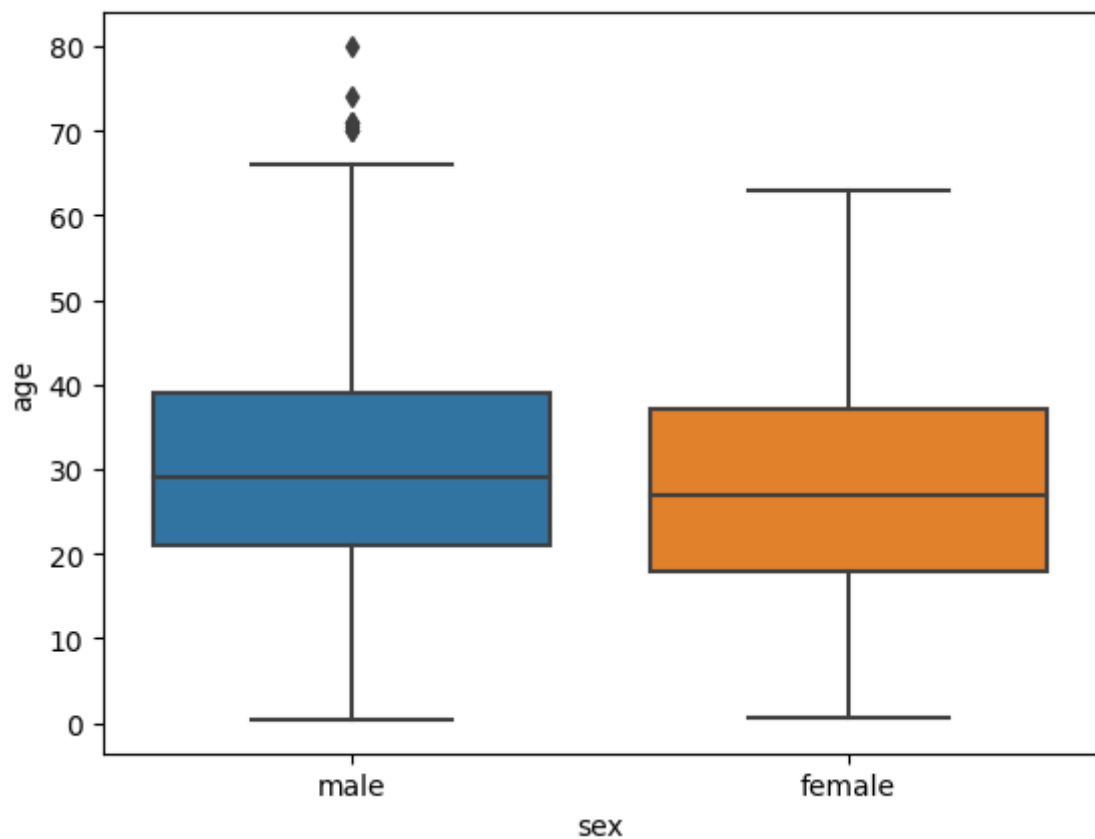
Out[74]: <AxesSubplot: xlabel='sex', ylabel='age'>

In [75]:
```python
sns.countplot(x='sex', data=df)
```

Out[75]: `<AxesSubplot: xlabel='sex', ylabel='count'>`

In [76]:
```python
sns.boxplot(x='sex', y='age', data=df)
```
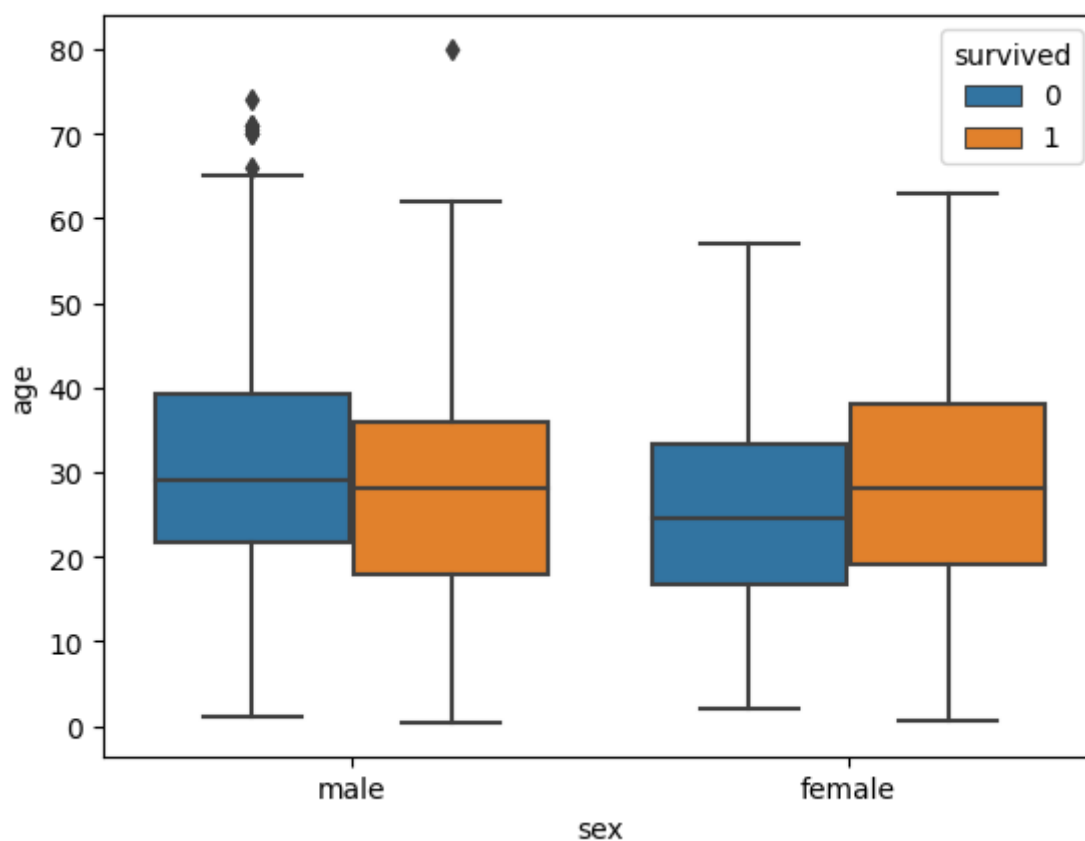
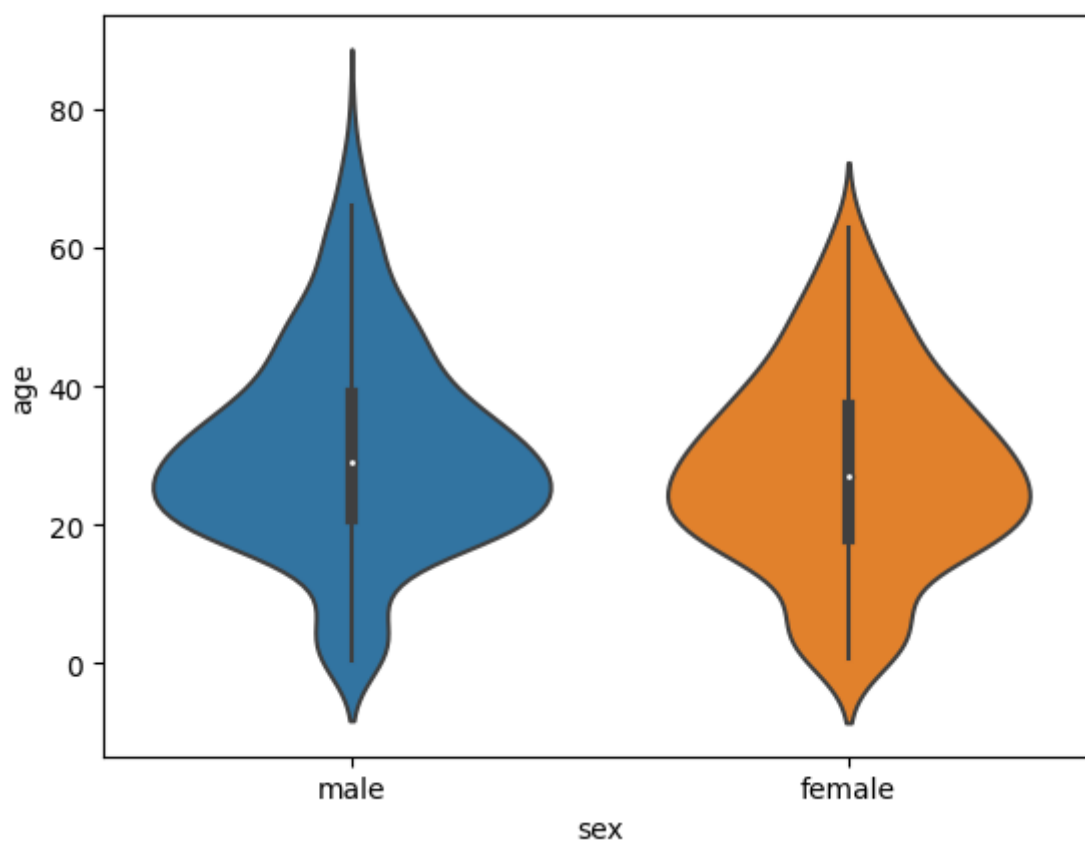Out[76]: `<AxesSubplot: xlabel='sex', ylabel='age'>`

In [77]:
```python
sns.boxplot(x='sex', y='age', data=df, hue="survived")
```
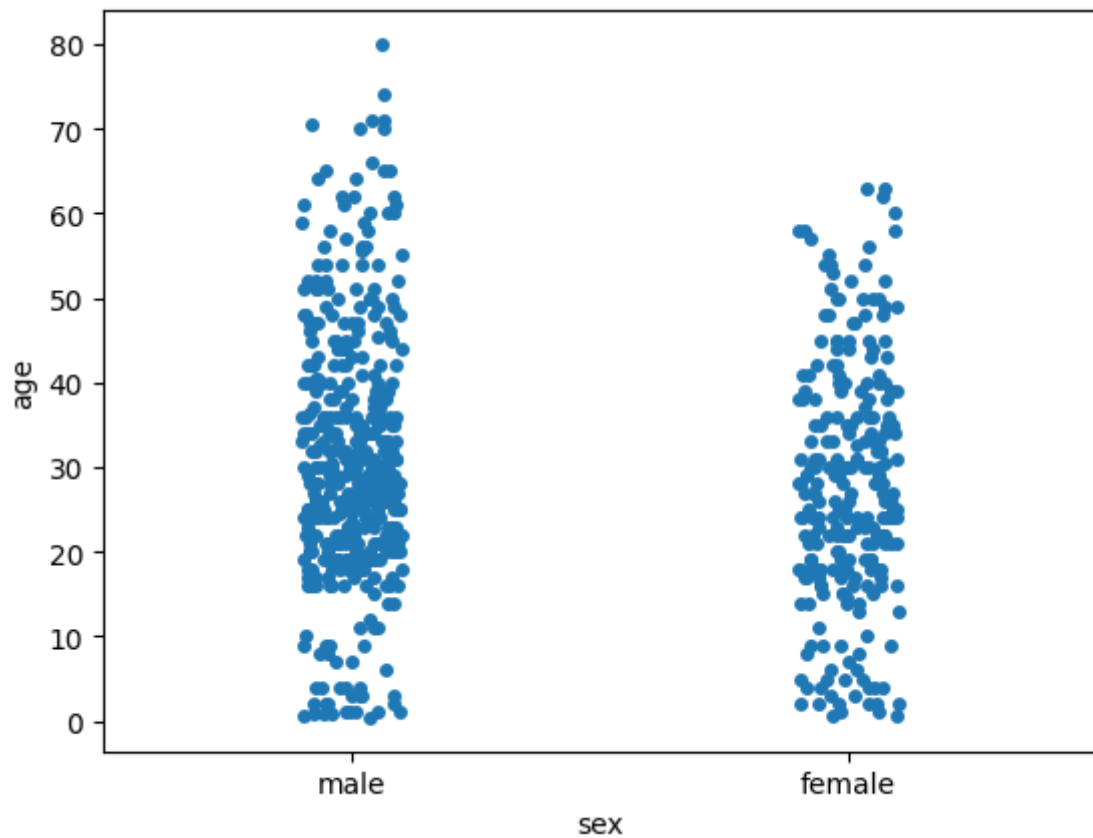
Out[77]:   `<AxesSubplot: xlabel='sex', ylabel='age'>`

In [78]:
```python
sns.violinplot(x='sex', y='age', data=df)
```

Out[78]:   `<AxesSubplot: xlabel='sex', ylabel='age'>`

In [79]:
```python
sns.stripplot(x='sex', y='age', data=df)
```

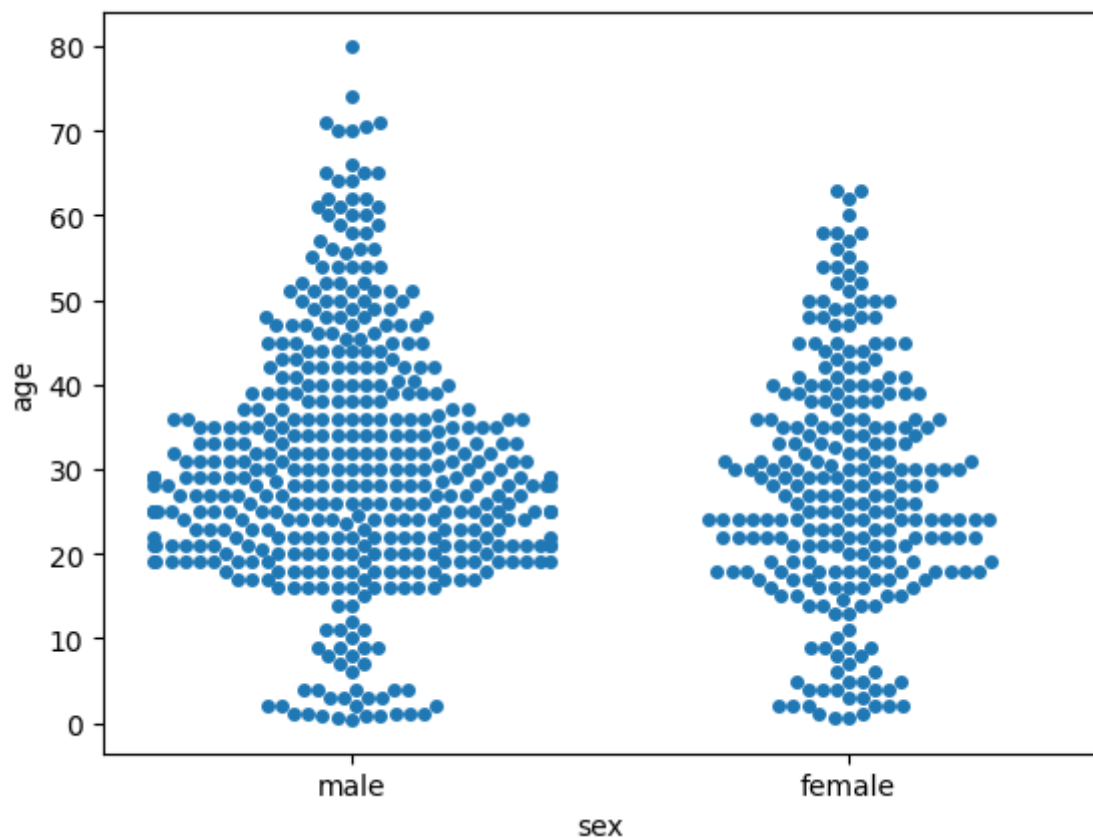Out[79]:   `<AxesSubplot: xlabel='sex', ylabel='age'>`

In [80]:
```python
sns.swarmplot(x='sex', y='age', data=df)
```
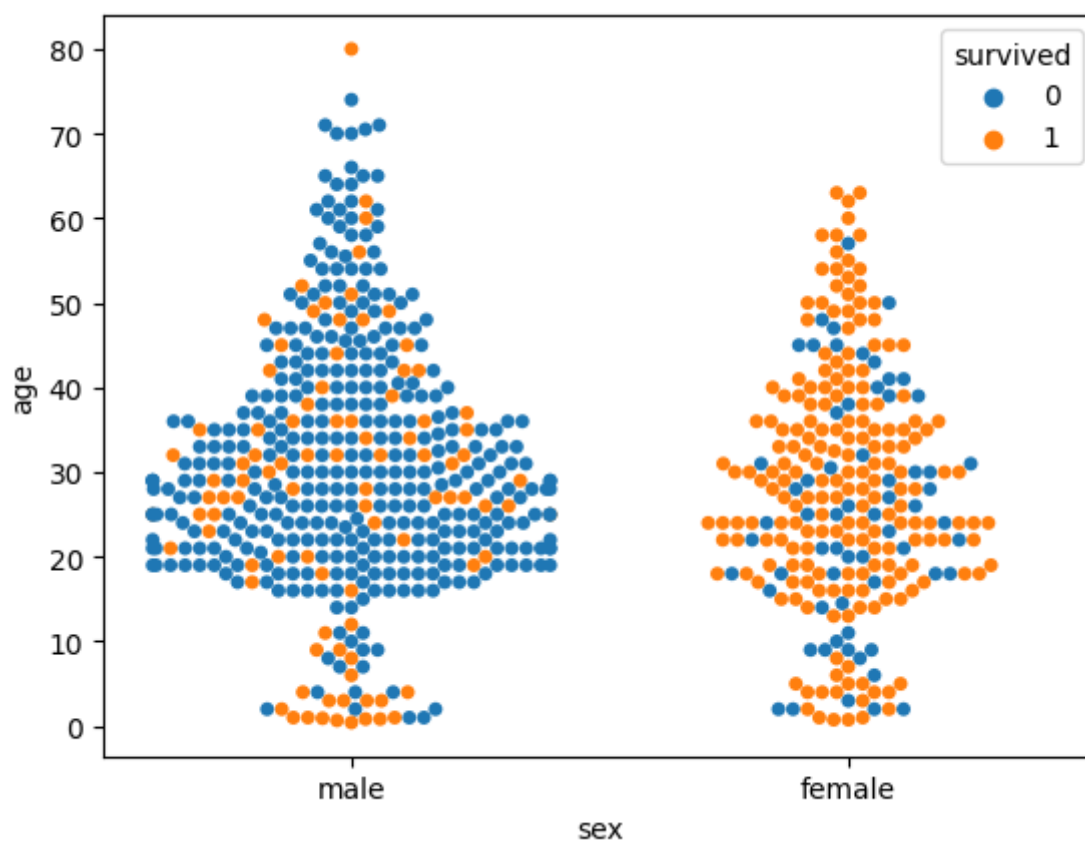
Out[80]:   `<AxesSubplot: xlabel='sex', ylabel='age'>`

In [81]:
```python
sns.swarmplot(x='sex', y='age', data=df, hue='survived')
```
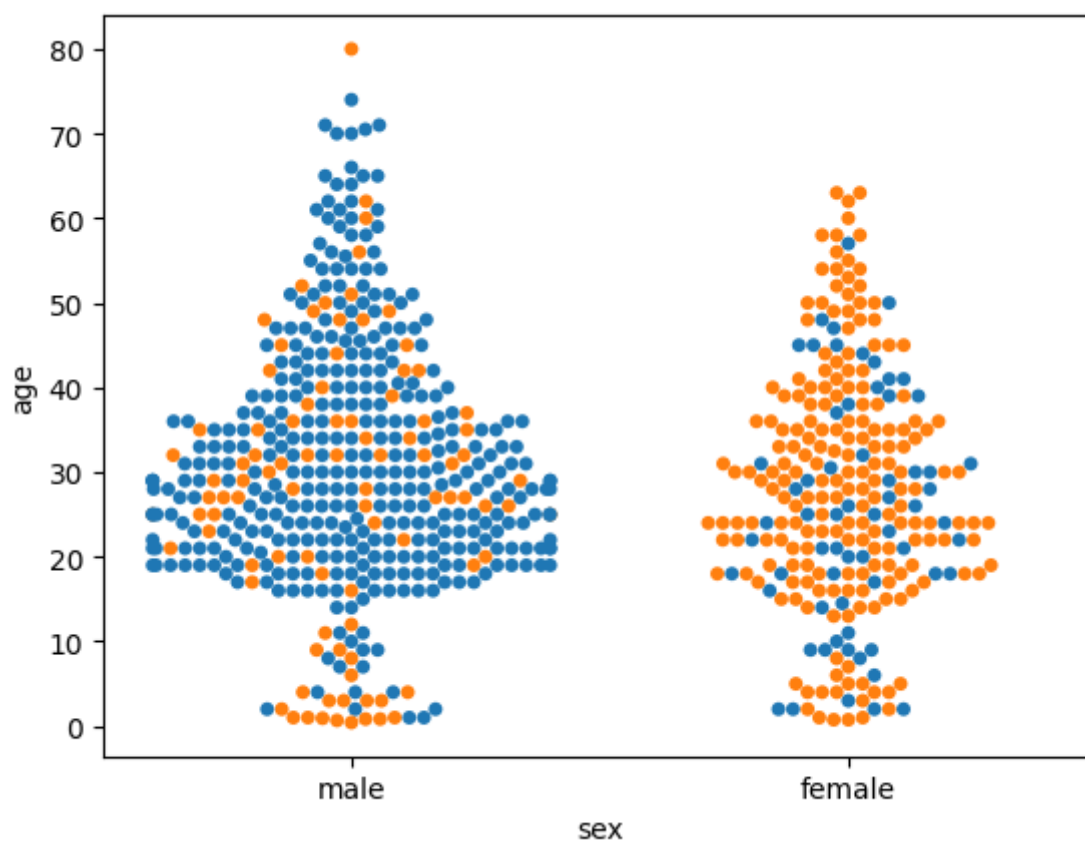
Out[81]:   `<AxesSubplot: xlabel='sex', ylabel='age'>`

In [82]:
```python
sns.swarmplot(x='sex', y='age', data=df, hue='survived', legend=False)
```

Out[82]:   `<AxesSubplot: xlabel='sex', ylabel='age'>`

In [ ]: