In [13]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

In [2]:
```
df = sns.load_dataset('titanic')
```

In [3]:
```
df
```

Out[3]:

	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
886	0	2	male	27.0	0	0	13.0000	S	Second	man	True	NaN	Southampton	no	True
887	1	1	female	19.0	0	0	30.0000	S	First	woman	False	B	Southampton	yes	True
888	0	3	female	NaN	1	2	23.4500	S	Third	woman	False	NaN	Southampton	no	False
889	1	1	male	26.0	0	0	30.0000	C	First	man	True	C	Cherbourg	yes	True
890	0	3	male	32.0	0	0	7.7500	Q	Third	man	True	NaN	Queenstown	no	True

891 rows × 15 columns

In [4]:
```
df.describe()
```

Out[4]:

	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 [5]:
```
df.dtypes
```

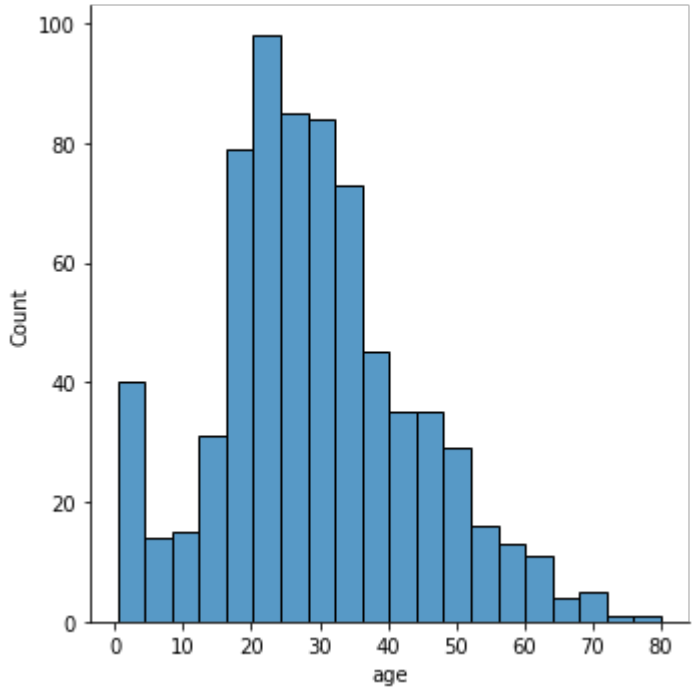
Out[5]:
```
survived          int64
pclass            int64
sex              object
age             float64
sibsp             int64
parch             int64
fare            float64
embarked         object
class          category
who              object
adult_male         bool
deck           category
embark_town      object
alive            object
alone              bool
dtype: object
```

In [6]:
```
df.info
```

Out[6]:
```
<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   Third
1           1       1  female  38.0      1      0  71.2833        C   First
2           1       3  female  26.0      0      0   7.9250        S   Third
3           1       1  female  35.0      1      0  53.1000        S   First
4           0       3    male  35.0      0      0   8.0500        S   Third
..        ...     ...     ...   ...    ...    ...      ...      ...     ...
886         0       2    male  27.0      0      0  13.0000        S  Second
887         1       1  female  19.0      0      0  30.0000        S   First
888         0       3  female   NaN      1      2  23.4500        S   Third
889         1       1    male  26.0      0      0  30.0000        C   First
890         0       3    male  32.0      0      0   7.7500        Q   Third

       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 [7]:
```
sns.displot(df['age'])
```

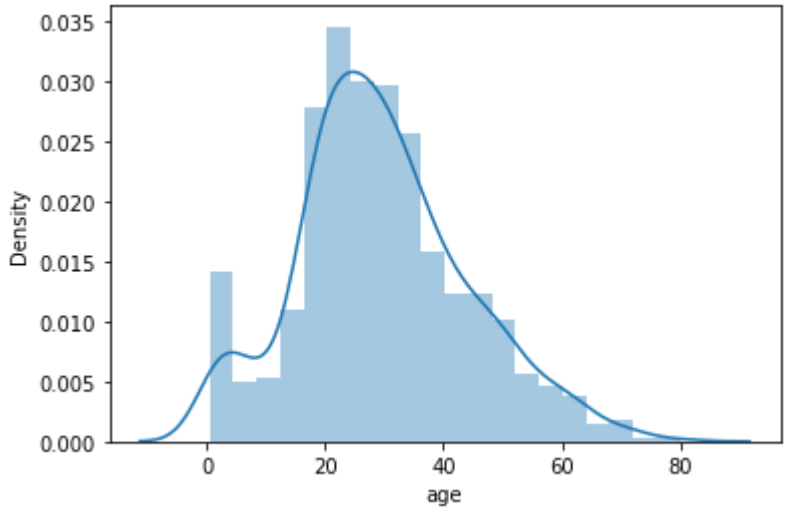
Out[7]:
```
<seaborn.axisgrid.FacetGrid at 0x7f527df13c10>
```



In [8]:
```
sns.distplot(df['age'])
```

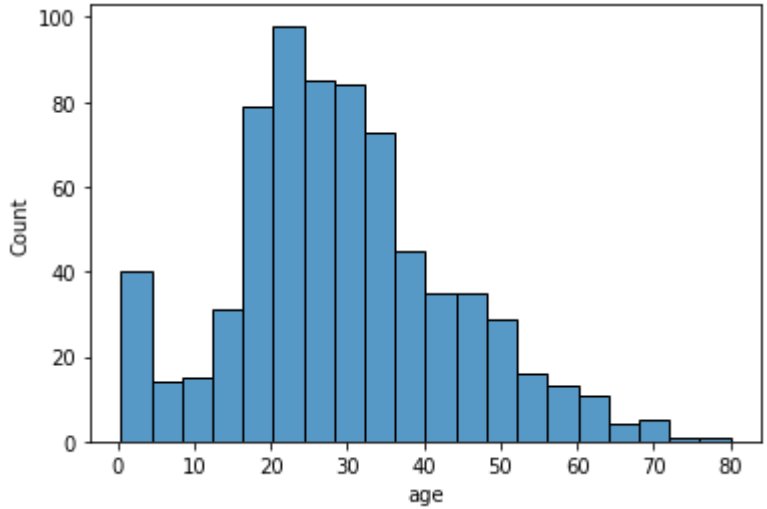
```
/tmp/ipykernel_6159/3234920688.py:1: UserWarning:

`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['age'])
```

Out[8]:
```
<AxesSubplot: xlabel='age', ylabel='Density'>
```



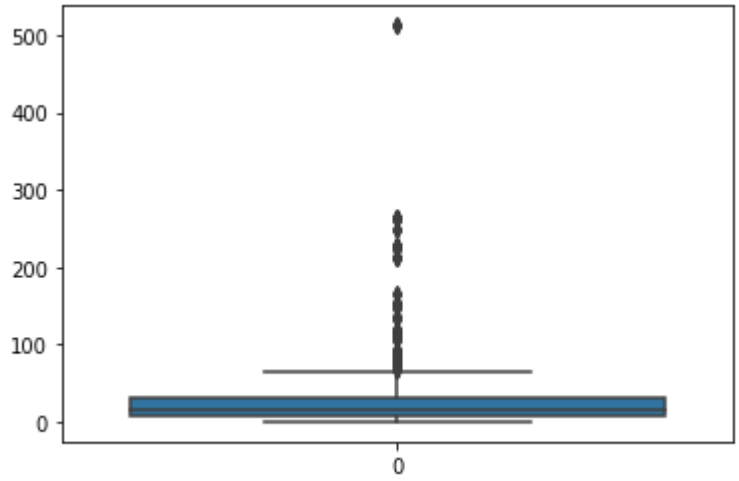
In [9]:
```
sns.histplot(df['age'])
```

Out[9]:
```
<AxesSubplot: xlabel='age', ylabel='Count'>
```



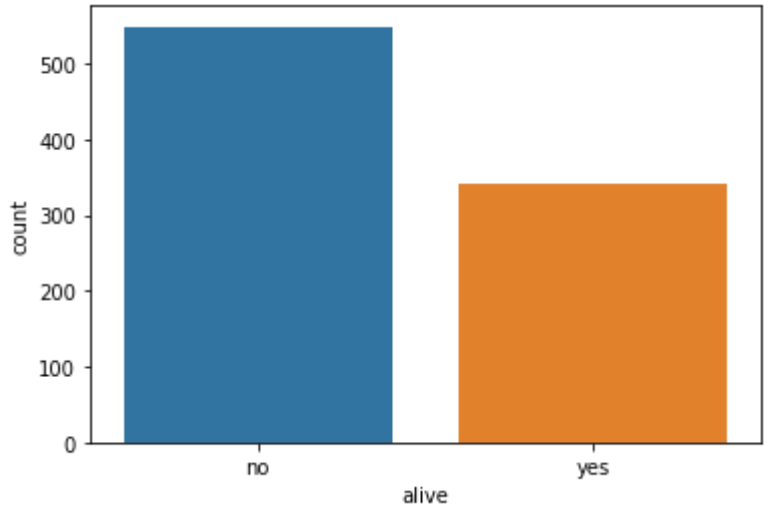
In [12]:
```
sns.boxplot(df['fare'])
```

Out[12]:
```
<AxesSubplot: >
```



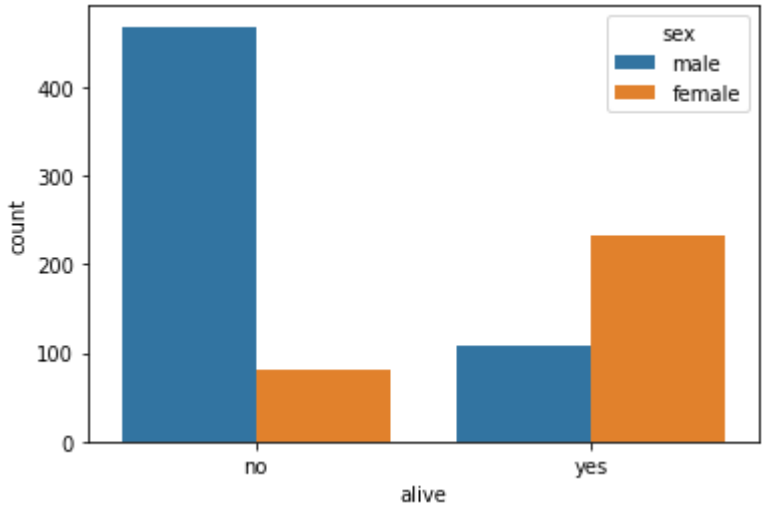
In [18]:
```
# Bar graph for classification of people who survived/not
sns.countplot(x='alive', data=df)
```

Out[18]:
```
<AxesSubplot: xlabel='alive', ylabel='count'>
```



In [17]:
```
# Bar graph for classification of people who survived/not wrt gender
sns.countplot(x='alive', data=df, hue='sex')
```

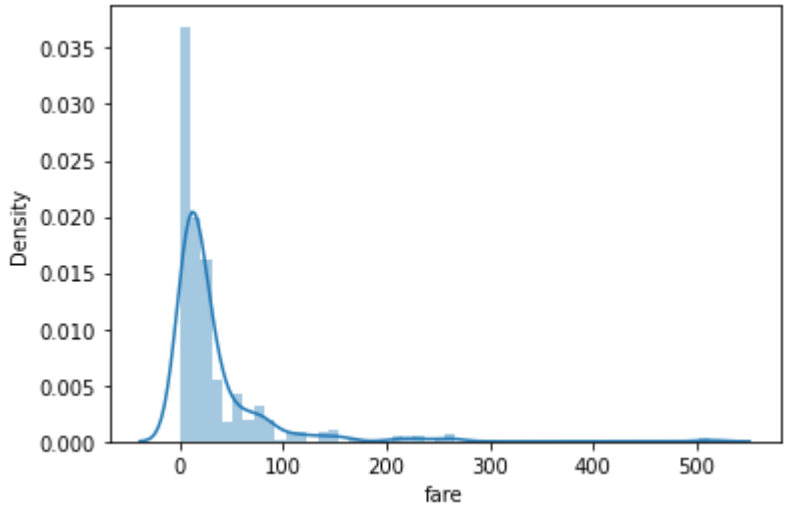
Out[17]:
```
<AxesSubplot: xlabel='alive', ylabel='count'>
```



In [20]:
```
# Distribution Plots

# 1. Distplot
sns.distplot(df['fare'])
```

```
/tmp/ipykernel_6159/3135303385.py:4: 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'])
```

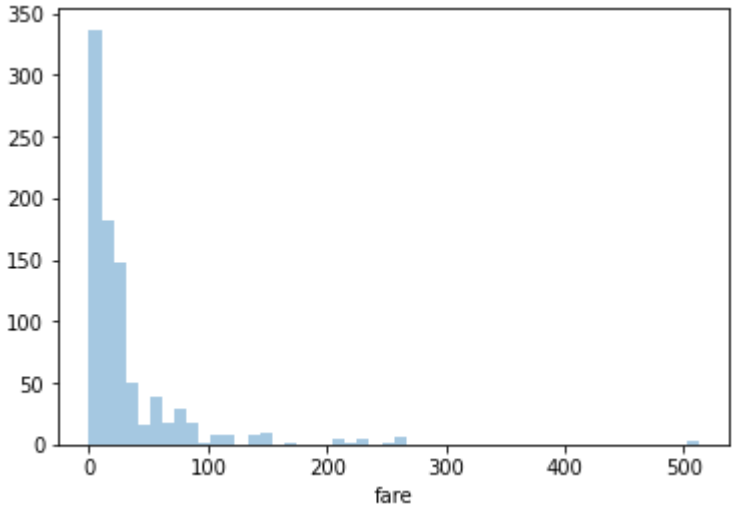
Out[20]:
```
<AxesSubplot: xlabel='fare', ylabel='Density'>
```



In [22]:
```
# removing the kde(kernel distribution estimation) line
sns.distplot(df['fare'], kde=False)
```

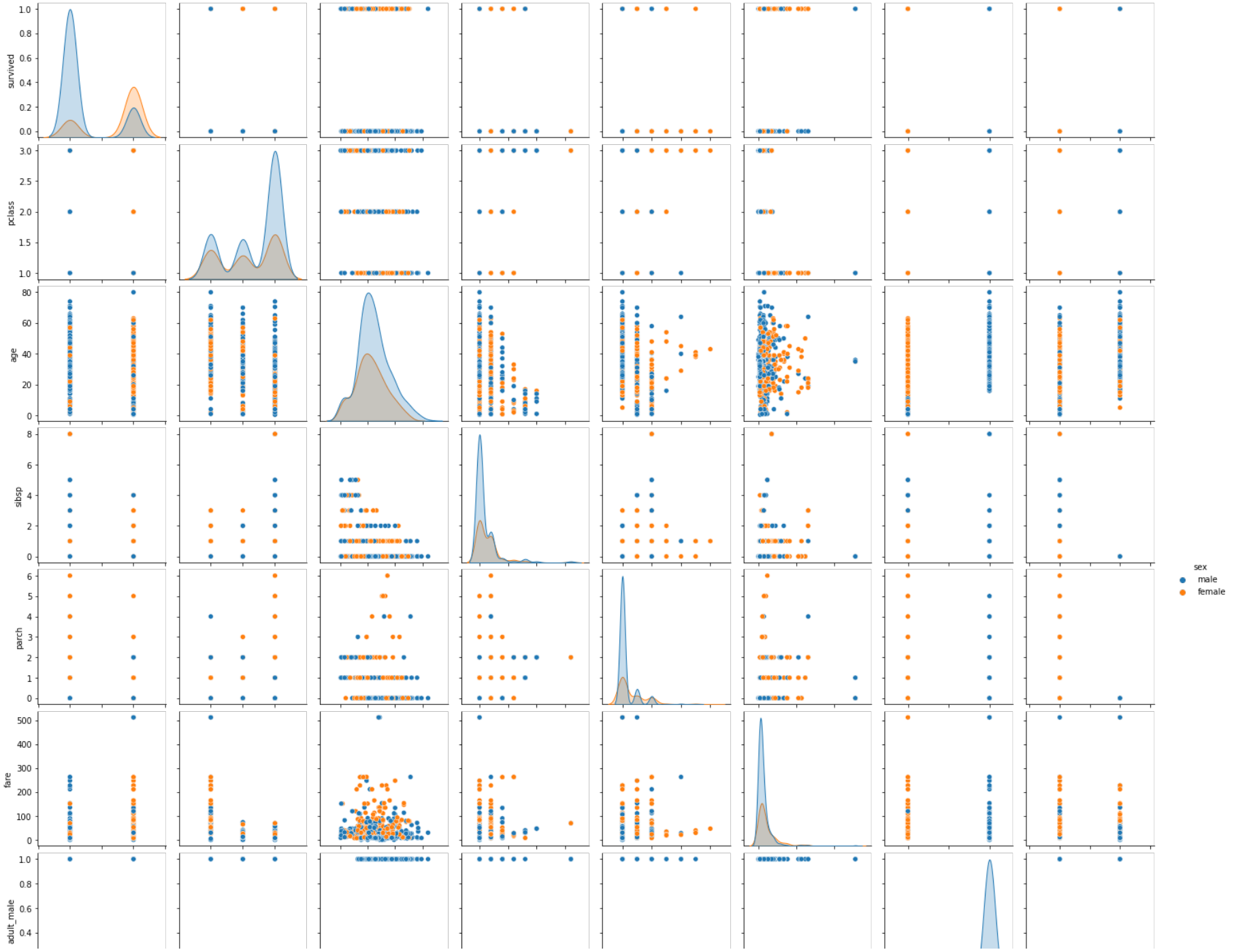
```
/tmp/ipykernel_6159/2623230642.py:2: UserWarning:

`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'], kde=False)
```

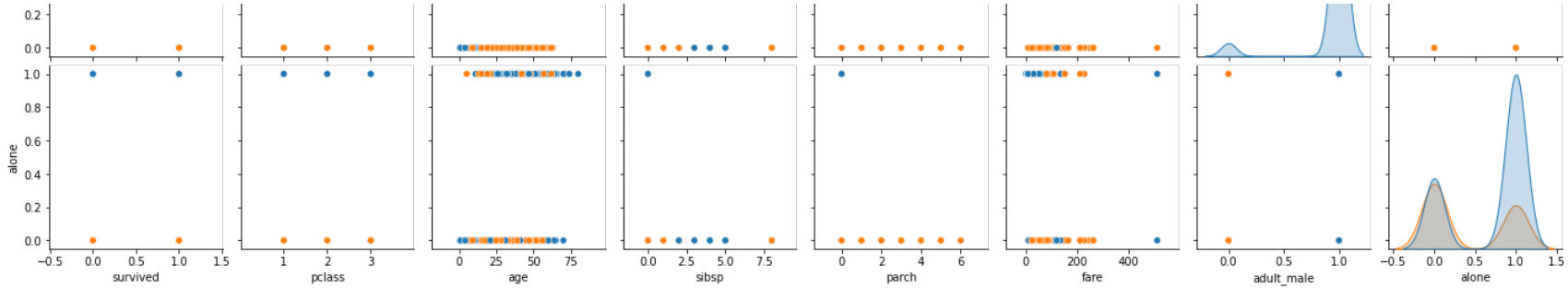
Out[22]:
```
<AxesSubplot: xlabel='fare'>
```



In [24]:
```
sns.pairplot(df, hue='sex')
```

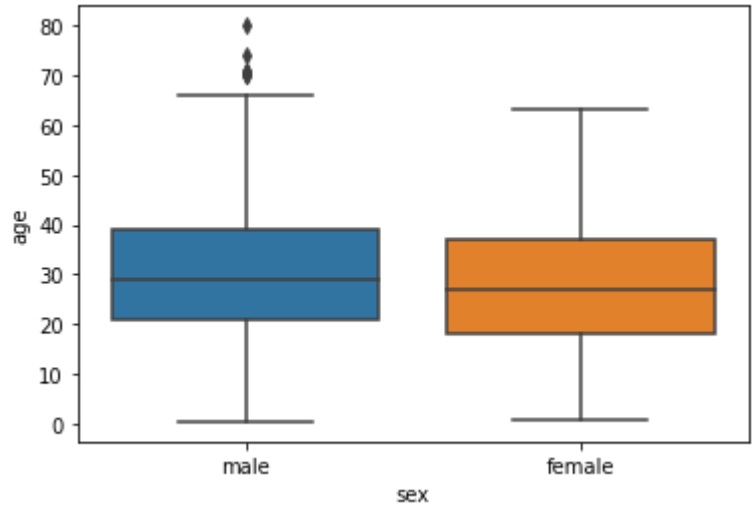
Out[24]:
```
<seaborn.axisgrid.PairGrid at 0x7f5270487a60>
```


survived
pclass
age
sibsp
parch
fare
adult_male
sex
male
female



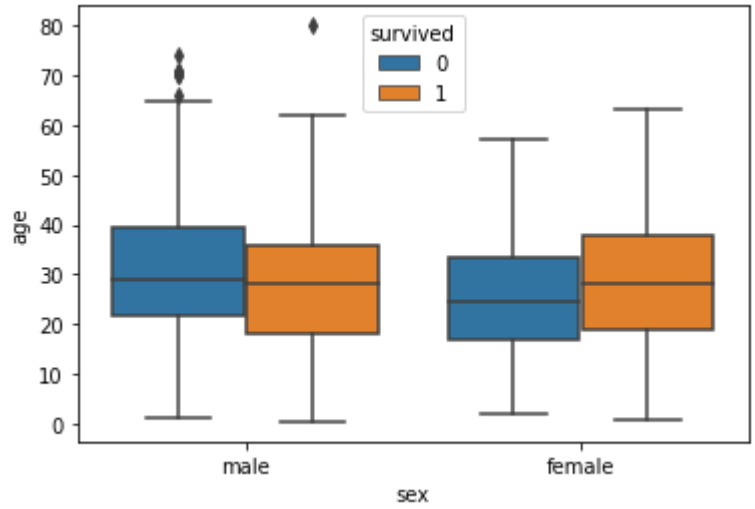
In [25]:
```
sns.boxplot(x='sex', y='age', data=df)
```

Out[25]:
```
<AxesSubplot: xlabel='sex', ylabel='age'>
```



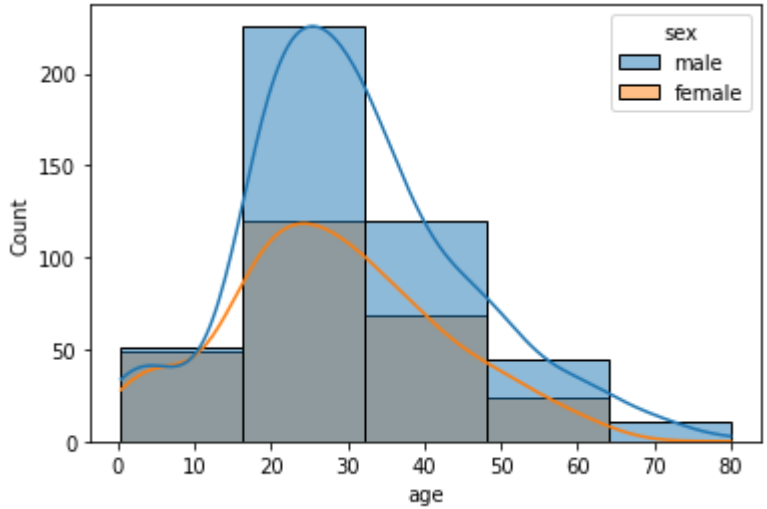
In [27]:
```
sns.boxplot(x='sex', y='age', data=df, hue="survived")
```

Out[27]:
```
<AxesSubplot: xlabel='sex', ylabel='age'>
```



In [28]:
```
sns.histplot(data=df, x="age", kde=True, bins = 5, hue = 'sex')
```

Out[28]:
```
<AxesSubplot: xlabel='age', ylabel='Count'>
```



In [ ]: