practice-assignment-9

May 4, 2025

[1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
```

[9]:
```
titanic=sns.load_dataset("titanic")
```

[13]:
```
print(titanic.head())
print(titanic.tail())
print(titanic.info())
print(titanic.describe())
print(titanic.isnull())
```

```
   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

     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
     survived  pclass     sex   age  sibsp  parch   fare embarked   class  \
886         0       2    male  27.0      0      0  13.00        S  Second
887         1       1  female  19.0      0      0  30.00        S   First
888         0       3  female   NaN      1      2  23.45        S   Third
889         1       1    male  26.0      0      0  30.00        C   First
890         0       3    male  32.0      0      0   7.75        Q   Third

       who  adult_male deck  embark_town alive  alone
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
<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
None
         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
     survived  pclass    sex    age  sibsp  parch   fare  embarked  class  \
0       False   False  False  False  False  False  False     False  False
1       False   False  False  False  False  False  False     False  False
2       False   False  False  False  False  False  False     False  False
3       False   False  False  False  False  False  False     False  False
4       False   False  False  False  False  False  False     False  False
..        …     …    …    …    …    …    …       …    …
886     False   False  False  False  False  False  False     False  False
887     False   False  False  False  False  False  False     False  False
888     False   False  False   True  False  False  False     False  False
889     False   False  False  False  False  False  False     False  False
890     False   False  False  False  False  False  False     False  False

       who  adult_male   deck  embark_town  alive  alone
0    False       False   True        False  False  False
```

```
1    False       False  False        False  False  False
2    False       False   True        False  False  False
3    False       False  False        False  False  False
4    False       False   True        False  False  False
..     …           …     …            …     …     …
886  False       False   True        False  False  False
887  False       False  False        False  False  False
888  False       False   True        False  False  False
889  False       False  False        False  False  False
890  False       False   True        False  False  False

[891 rows x 15 columns]
```

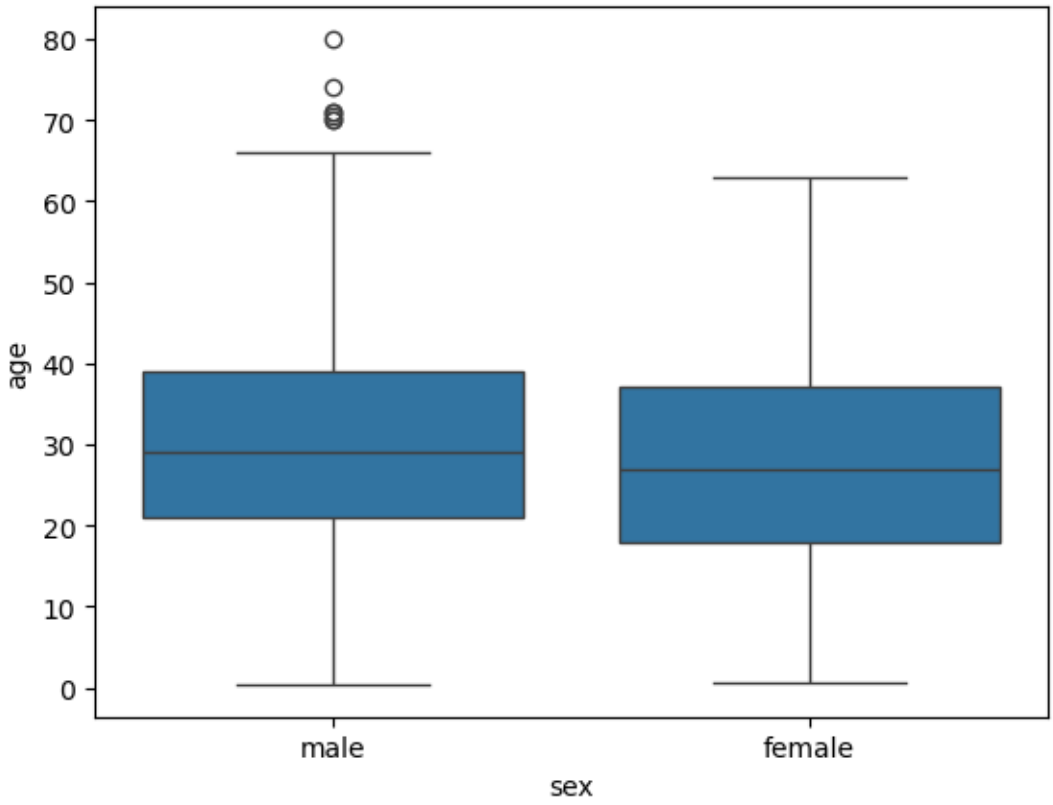
[53]:
```
print(titanic.isnull().any())
print("---------------------------")
print(titanic.isnull().sum())
```

```
survived       False
pclass         False
sex            False
age             True
sibsp          False
parch          False
fare           False
embarked        True
class          False
who            False
adult_male     False
deck            True
embark_town     True
alive          False
alone          False
dtype: bool
---------------------------
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
```

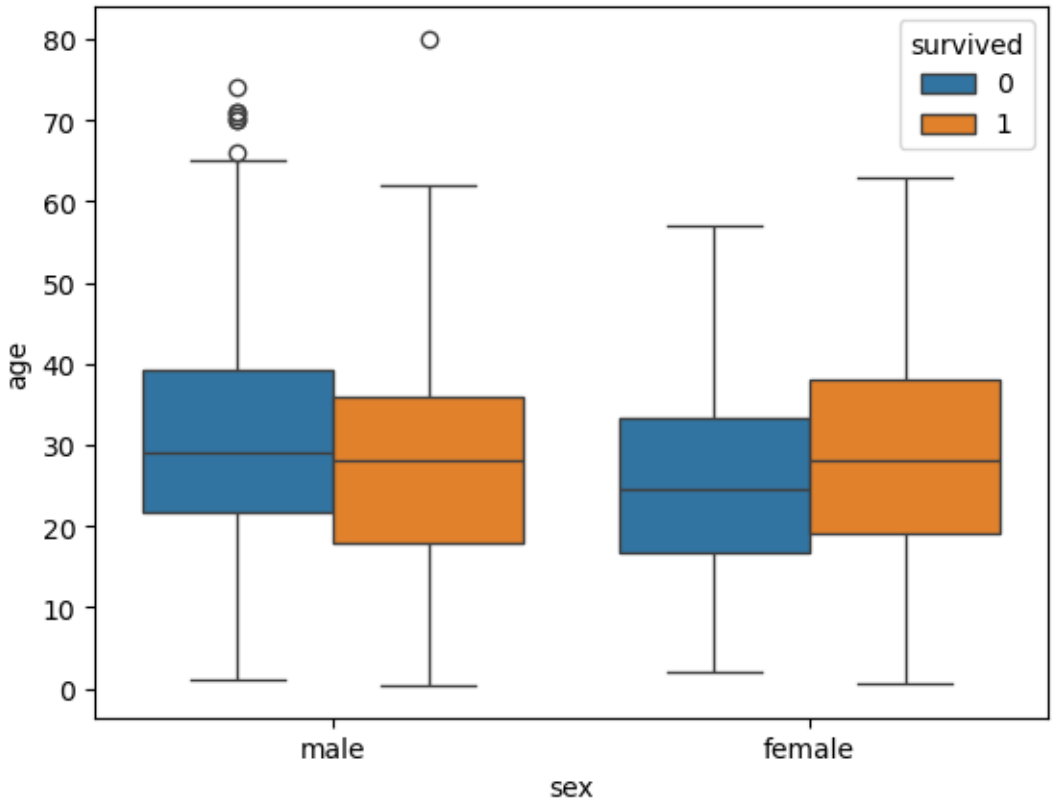
[19]:
```
sns.boxplot(x="sex",y="age",data=titanic)
```

[19]:
```
<Axes: xlabel='sex', ylabel='age'>
```



[21]:
```
sns.boxplot(x="sex",y="age",hue="survived",data=titanic)
```

[21]:
```
<Axes: xlabel='sex', ylabel='age'>
```

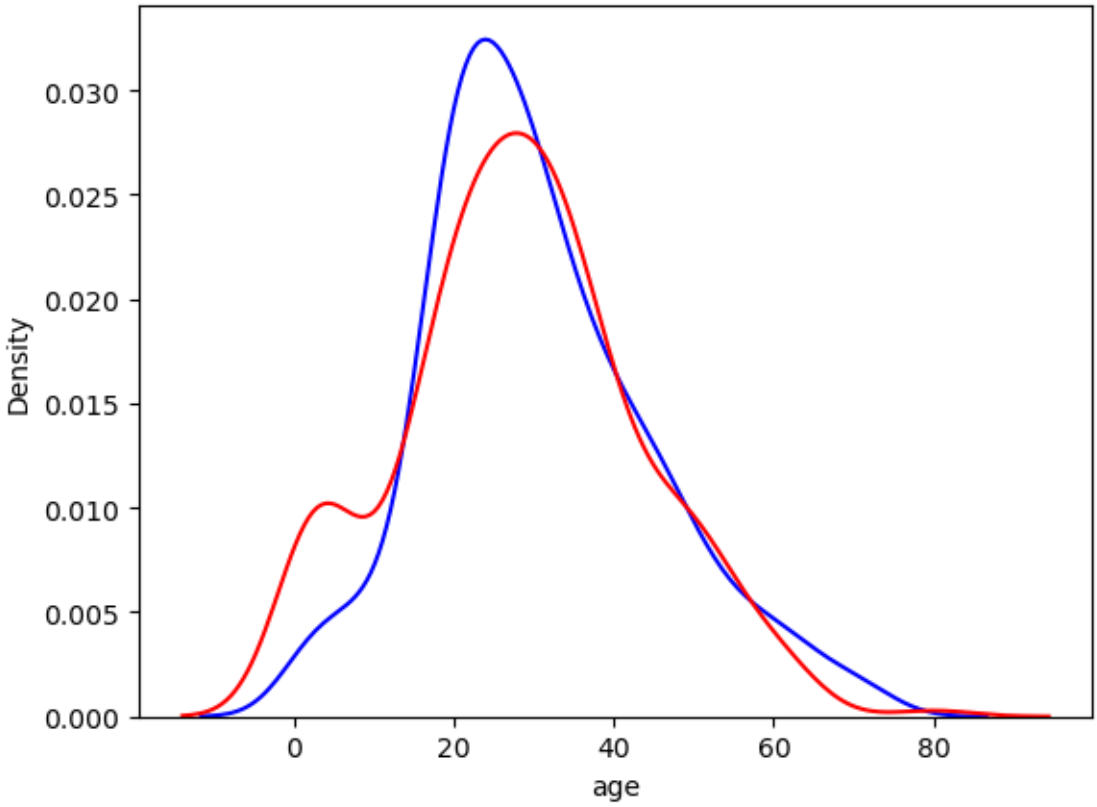


[31]:
```
sns.distplot(titanic[titanic["survived"]==0]["age"],hist=False,color='blue')
sns.distplot(titanic[titanic["survived"]==1]["age"],hist=False,color='red')
```

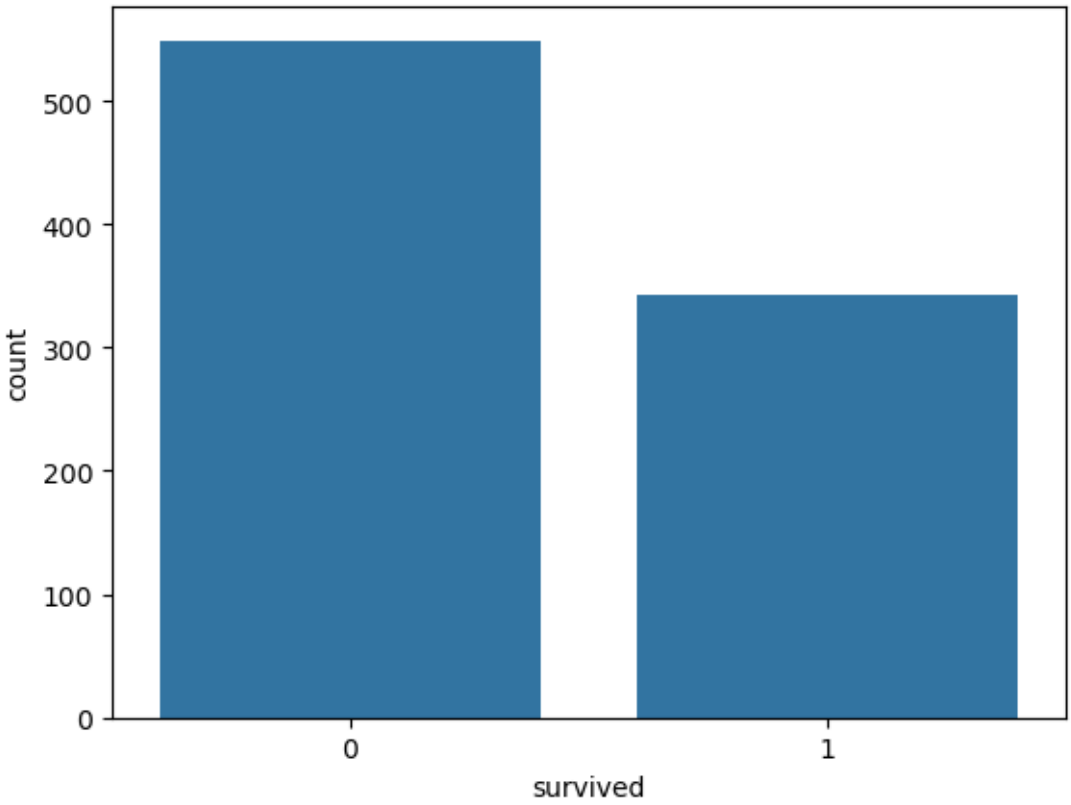
```
C:\Users\Varad\AppData\Local\Temp\ipykernel_16824\1306390126.py:1: UserWarning:

`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 `kdeplot` (an axes-level function for kernel density
plots).

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

  sns.distplot(titanic[titanic["survived"]==0]["age"],hist=False,color='blue')
C:\Users\Varad\AppData\Local\Temp\ipykernel_16824\1306390126.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 `kdeplot` (an axes-level function for kernel density
```

```
plots).

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

  sns.distplot(titanic[titanic["survived"]==1]["age"],hist=False,color='red')
```

[31]: <Axes: xlabel='age', ylabel='Density'>

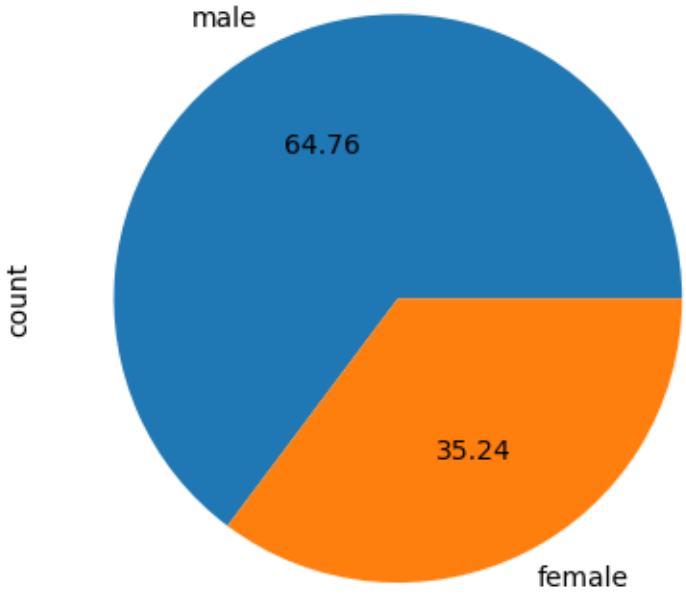


[35]:
```
sns.countplot(x="survived",data=titanic)
plt.show()
```



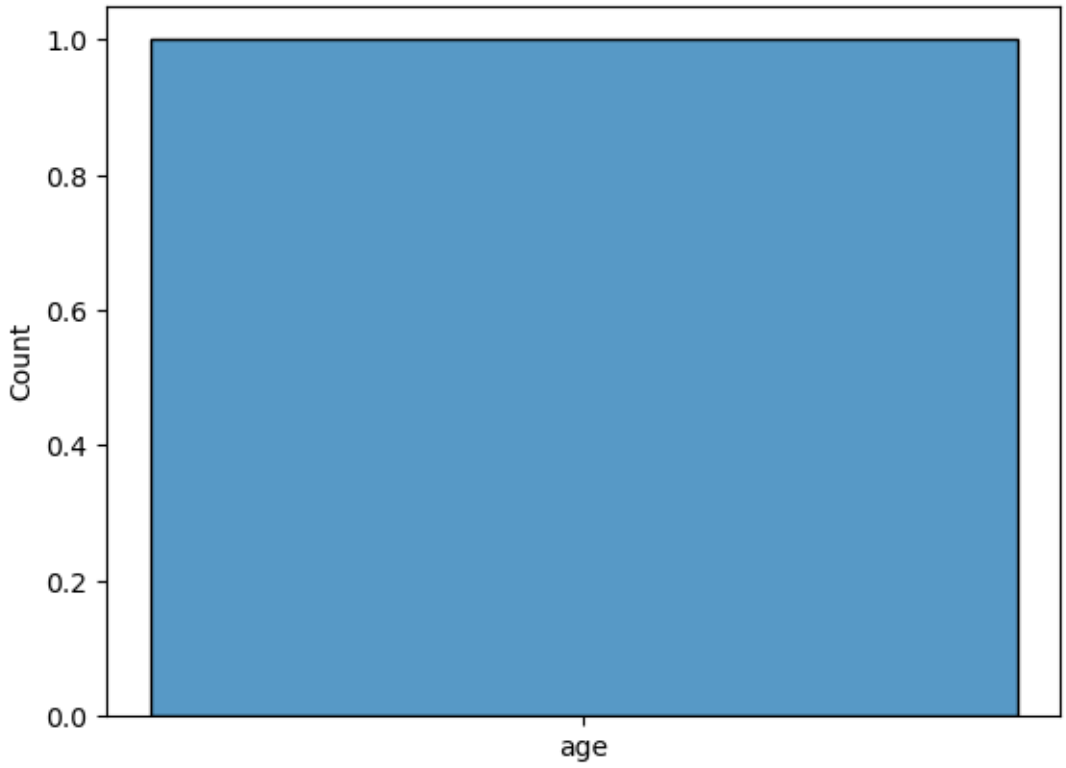
[37]:
```
titanic['sex'].value_counts().plot(kind="pie",autopct="%.2f")
```

[37]: <Axes: ylabel='count'>



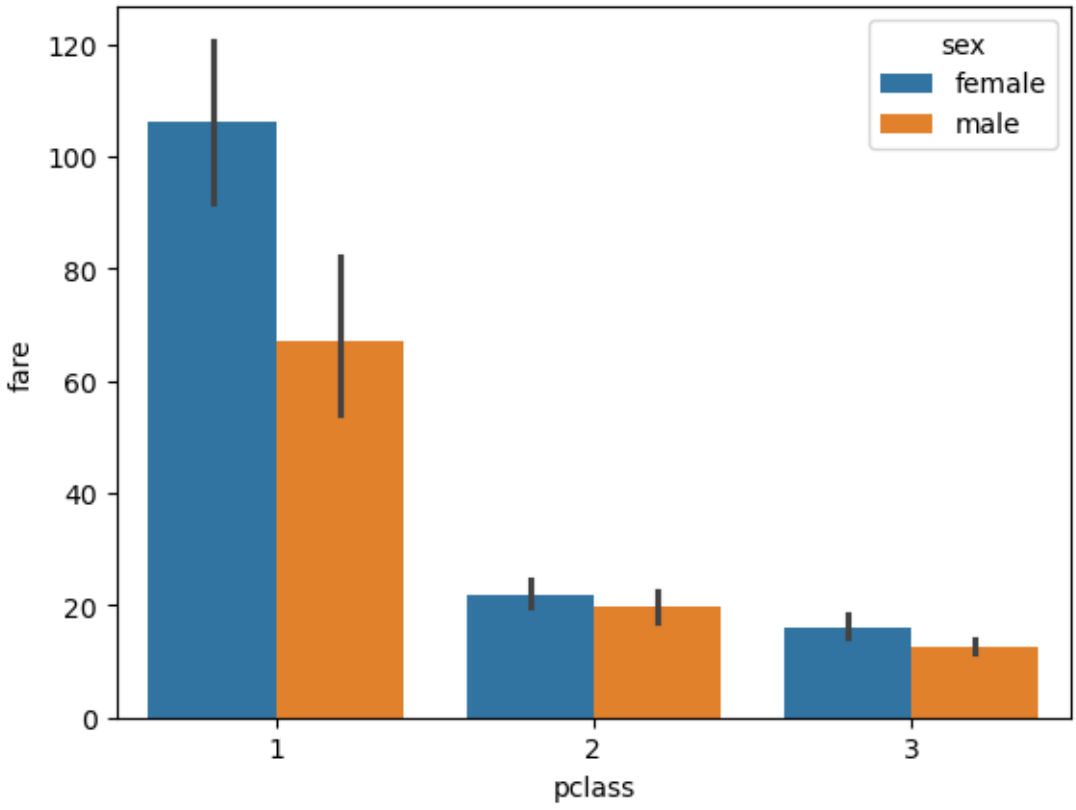
[39]:
```
sns.histplot(['age'],bins=5)
```

[39]: <Axes: ylabel='Count'>



[43]:
```
sns.barplot(x='pclass',y='fare',hue='sex',data=titanic)
```

[43]:
```
<Axes: xlabel='pclass', ylabel='fare'>
```

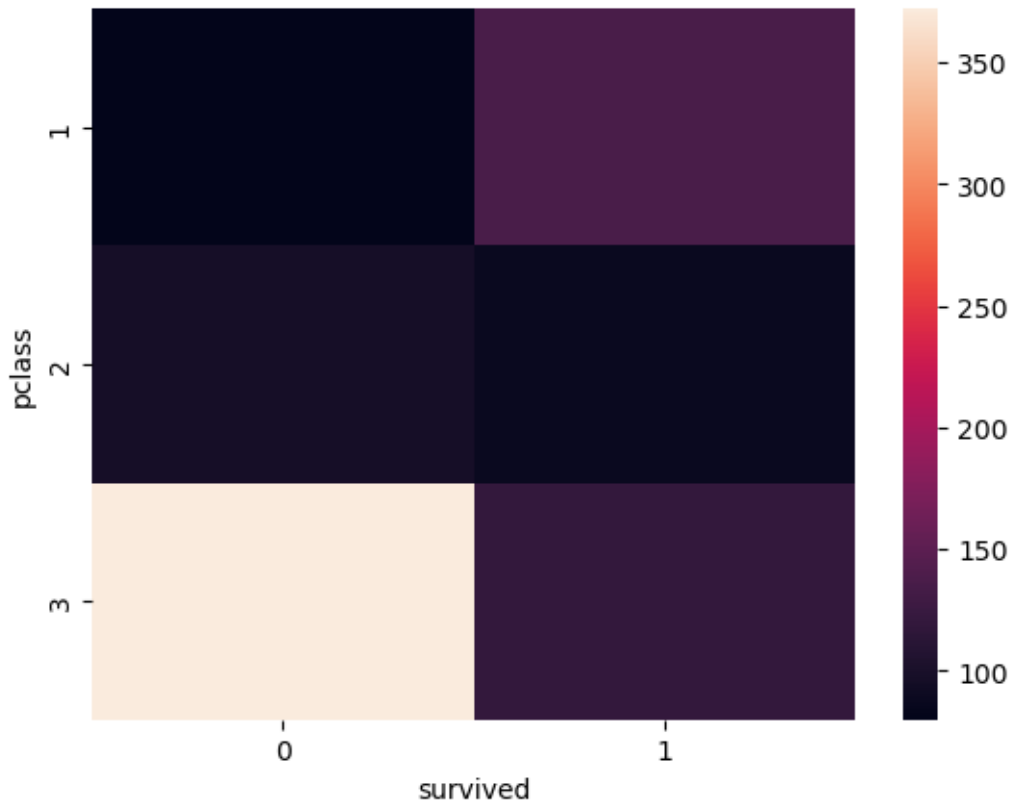


[45]:
```
pd.crosstab(titanic['pclass'],titanic['survived'])
```

[45]:
```
survived    0    1
pclass
1          80  136
2          97   87
3         372  119
```

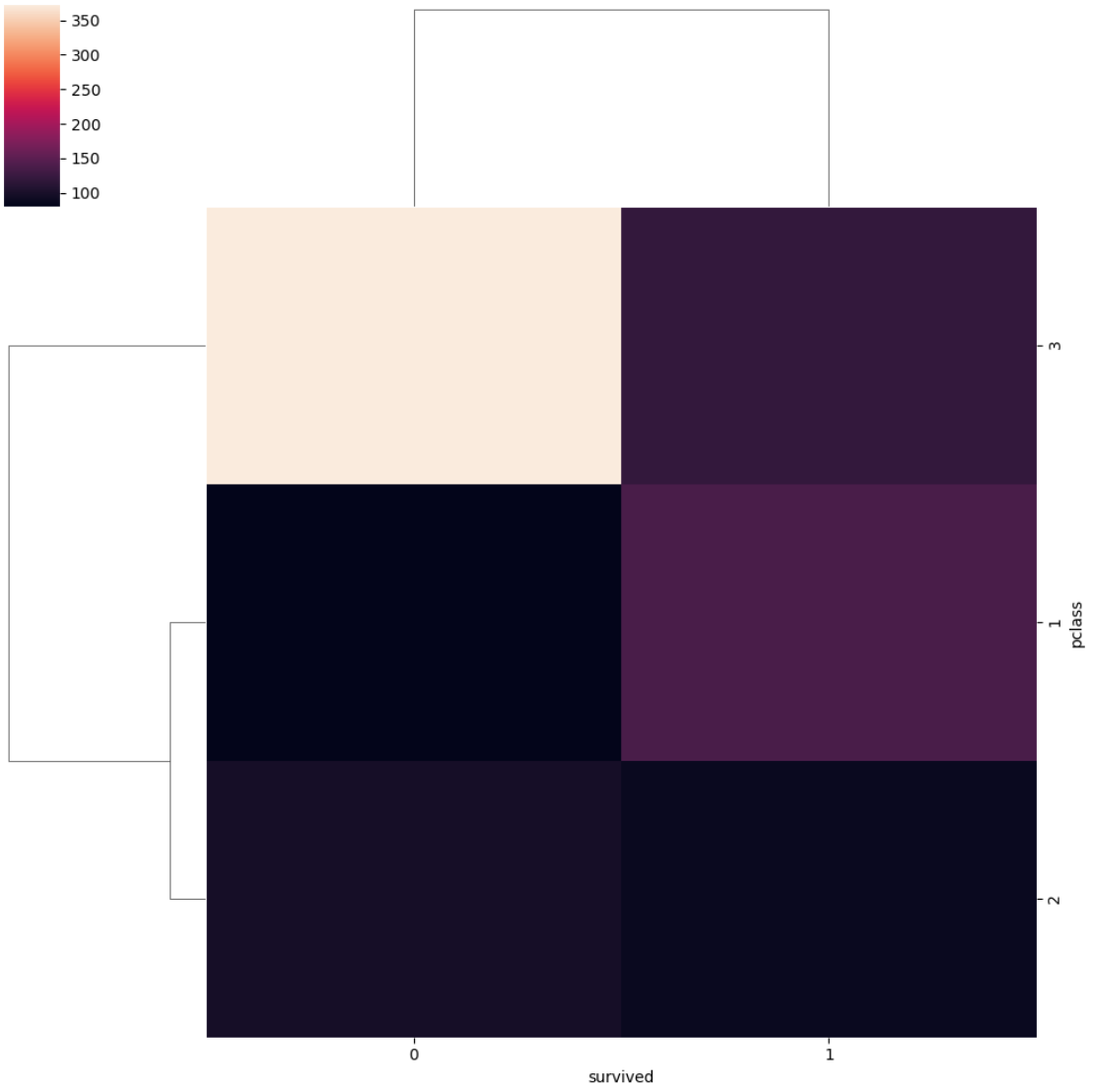
[47]:
```
sns.heatmap(pd.crosstab(titanic['pclass'],titanic['survived']))
```

[47]:
```
<Axes: xlabel='survived', ylabel='pclass'>
```



[49]:
```
sns.clustermap(pd.crosstab(titanic['pclass'],titanic['survived']))
```

[49]: <seaborn.matrix.ClusterGrid at 0x1fc74a03a70>


350
300
250
200
150
100
3
1
2
pclass
0
1
survived

[ ]: