Name: Pranav Mehendale

Roll No.: TCOD34

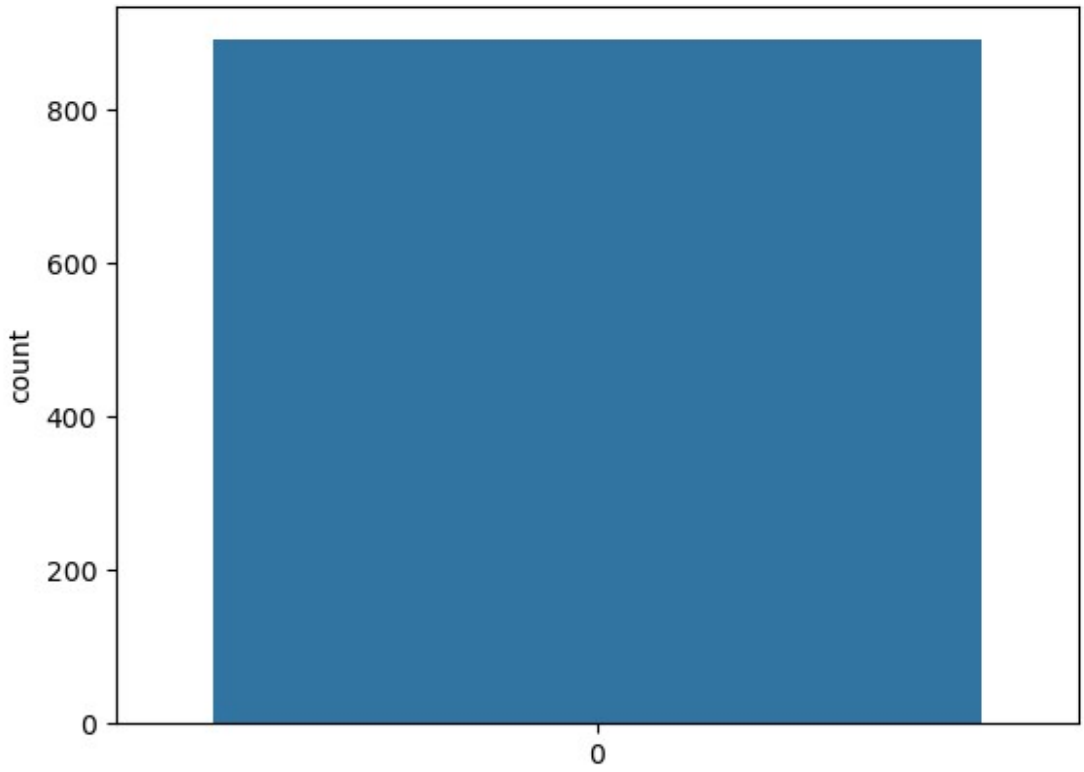
Batch: T11

Data Visualization II:

1. Use the inbuilt dataset 'titanic' as used in the above problem. Plot a box plot for distribution of age with respect to each gender along with the information about whether they survived or not. (Column names : 'sex' and 'age').
2. Write observations on the inference from the above statistics.

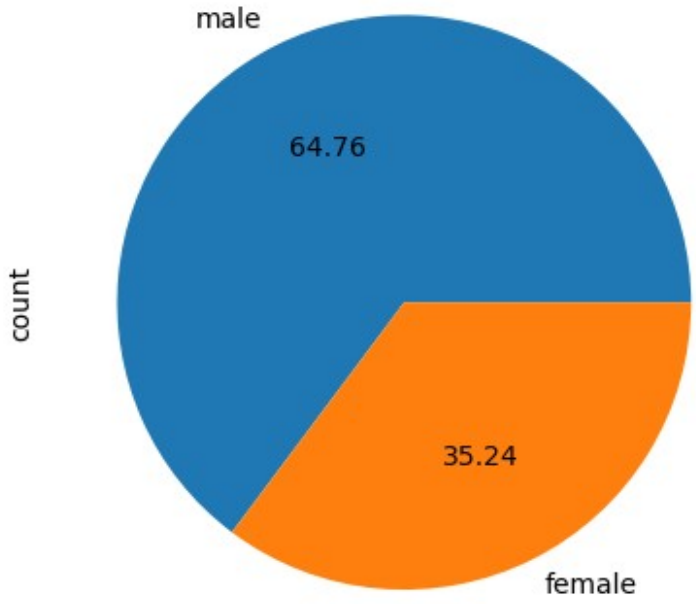
```
import pandas as pd
import matplotlib as plt
import seaborn as sns
import numpy as np
import warnings

warnings.filterwarnings('ignore')

%matplotlib inline

data = sns.load_dataset('titanic')
tips = sns.load_dataset('tips')

sns.countplot(data['parch'])
```

```
<Axes: ylabel='count'>
```



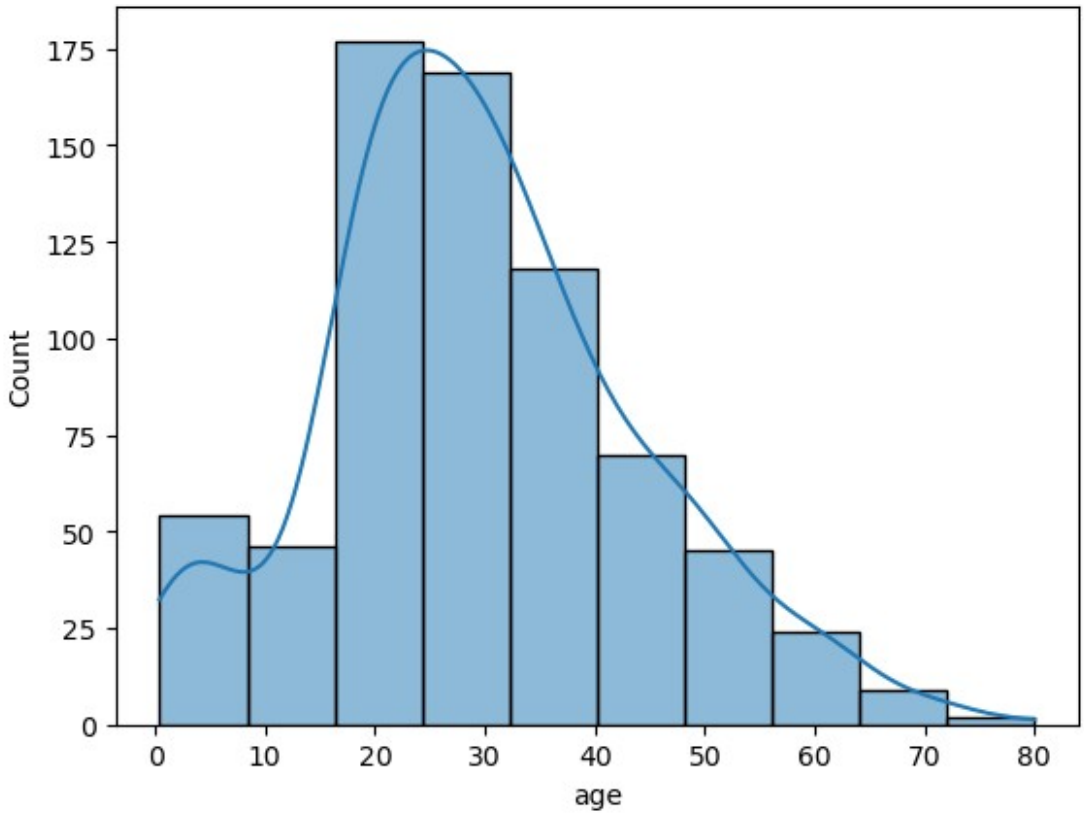

```
data['sex'].value_counts().plot(kind="pie", autopct="%.2f")
```

```
<Axes: ylabel='count'>
```



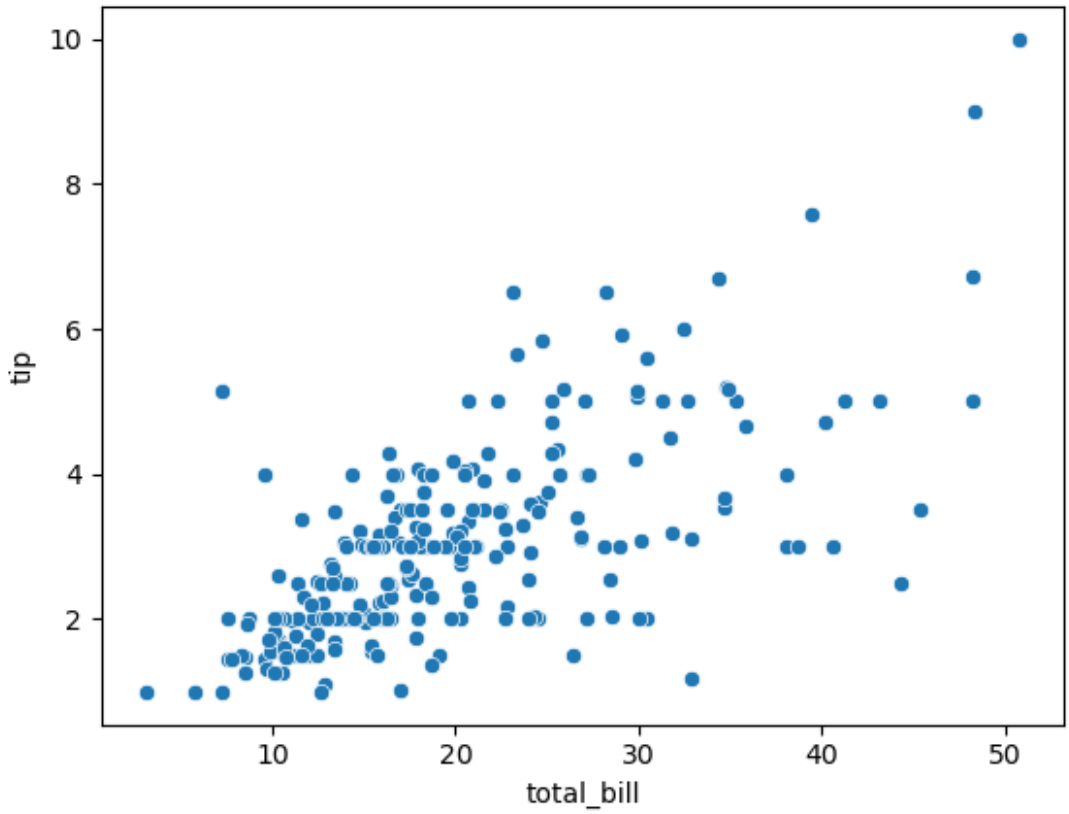

```
sns.histplot(data['age'], bins=10,kde=True)
```

```
<Axes: xlabel='age', ylabel='Count'>
```



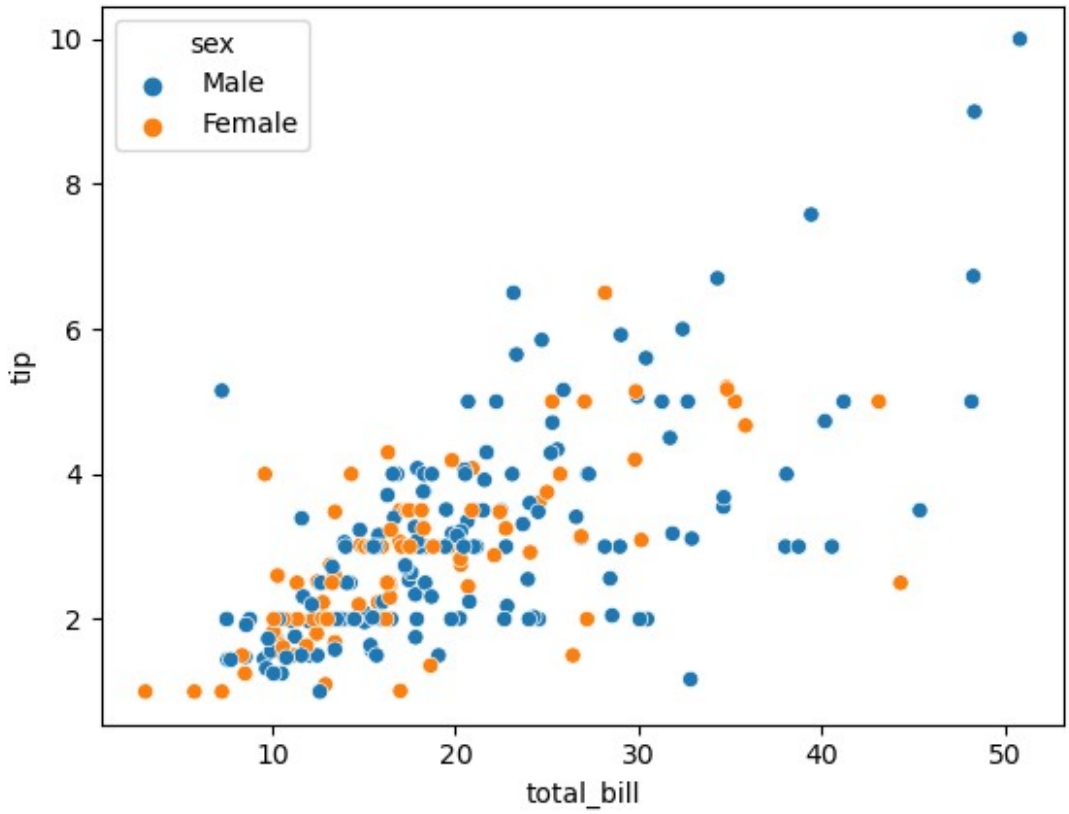

```
sns.scatterplot(x=tips["total_bill"], y=tips['tip'])
```

```
<Axes: xlabel='total_bill', ylabel='tip'>
```



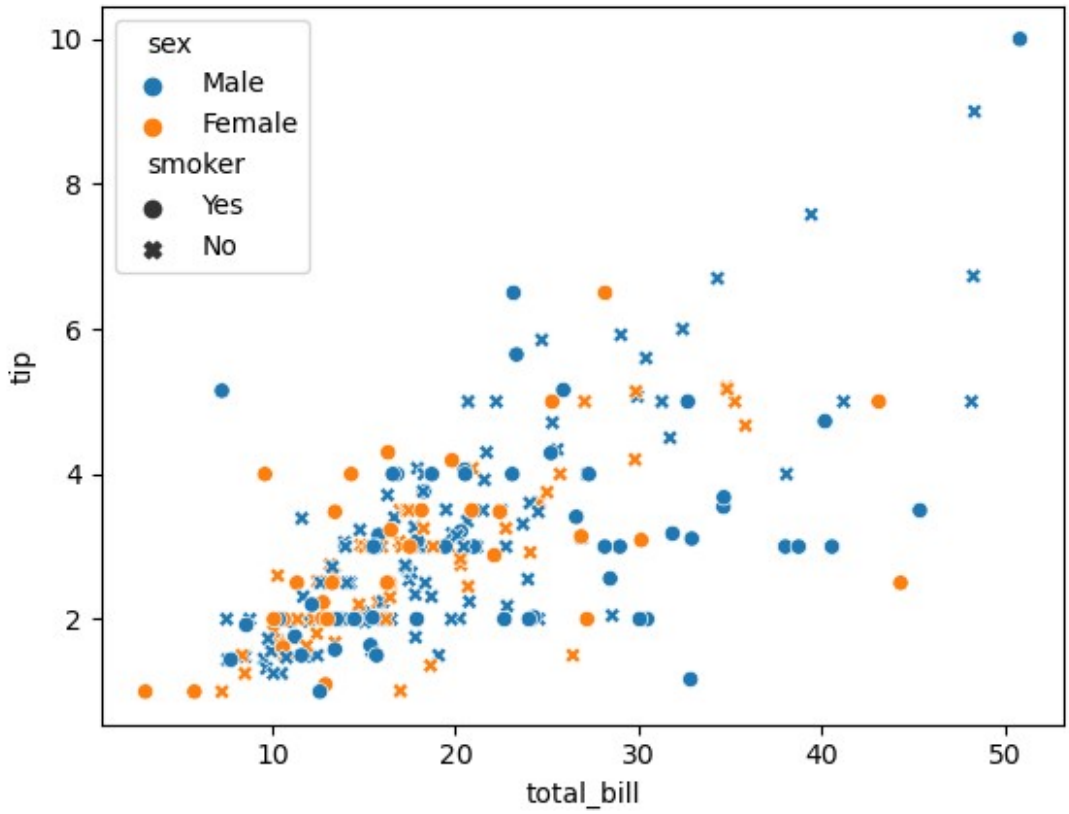

```
sns.scatterplot(x=tips["total_bill"], y=tips['tip'], hue=tips['sex'])
```

```
<Axes: xlabel='total_bill', ylabel='tip'>
```



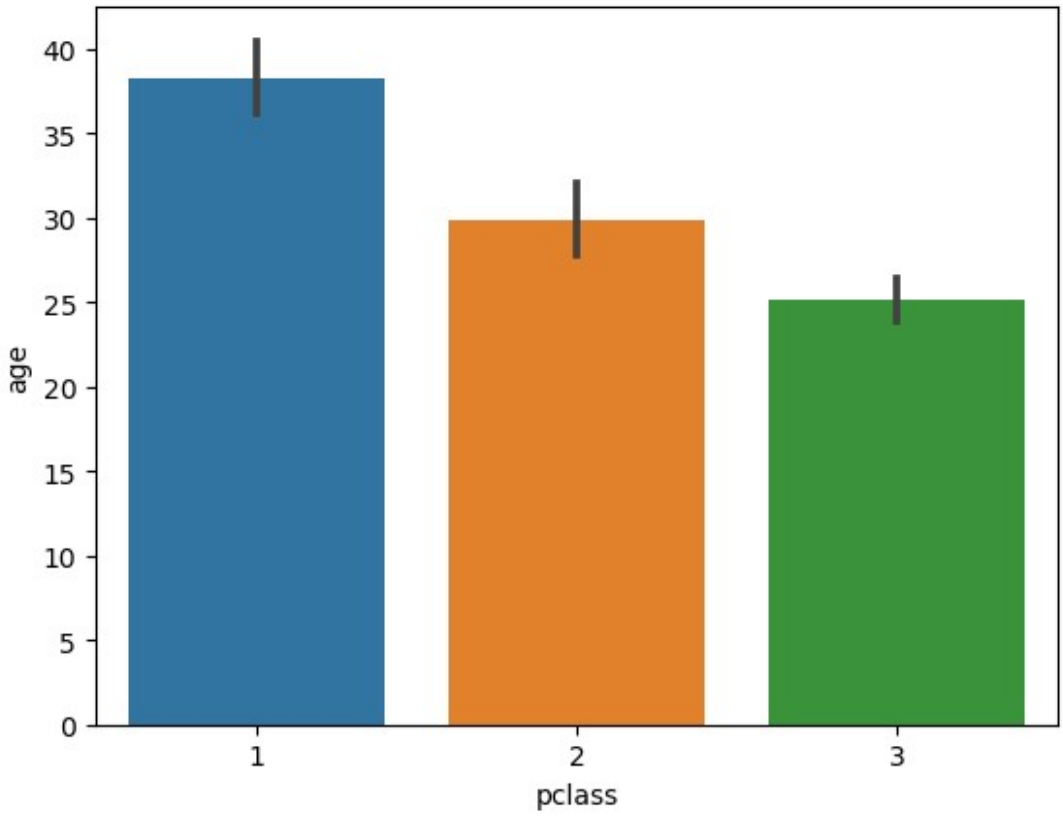

```
sns.scatterplot(x=tips["total_bill"], y=tips['tip'], hue=tips['sex'],
style=tips['smoker'])
```

```
<Axes: xlabel='total_bill', ylabel='tip'>
```



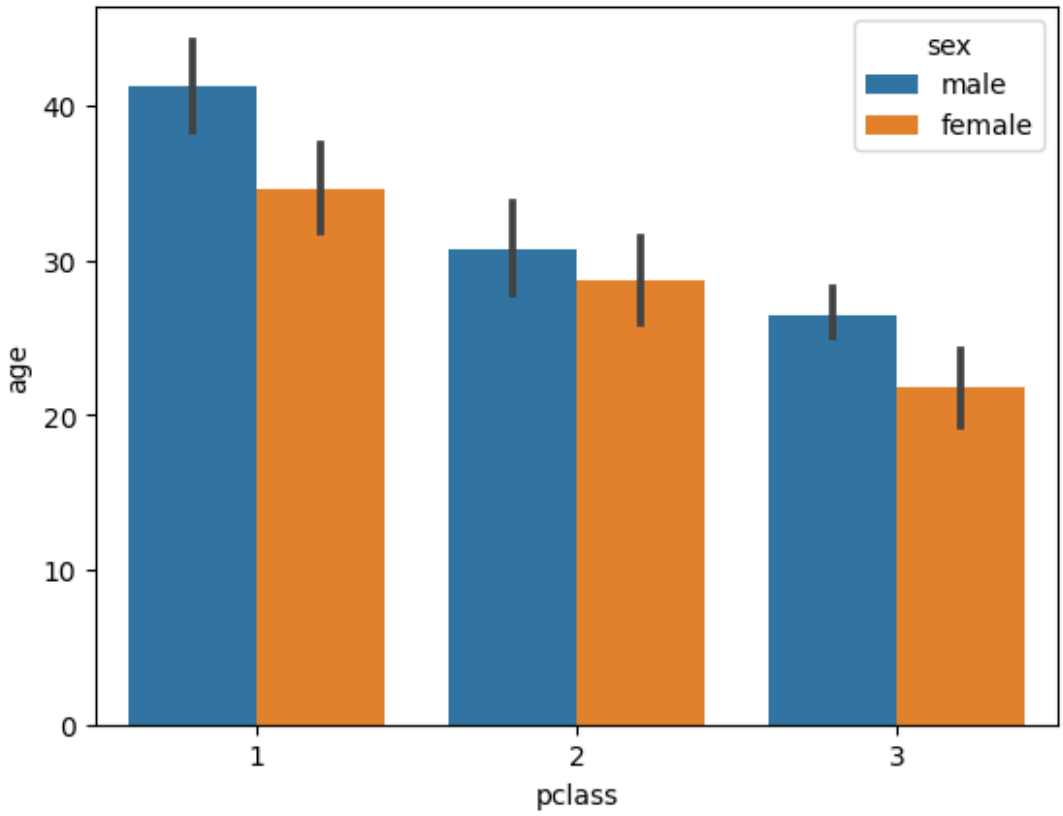

```
sns.barplot(x=data['pclass'], y=data['age'])
```

```
<Axes: xlabel='pclass', ylabel='age'>
```



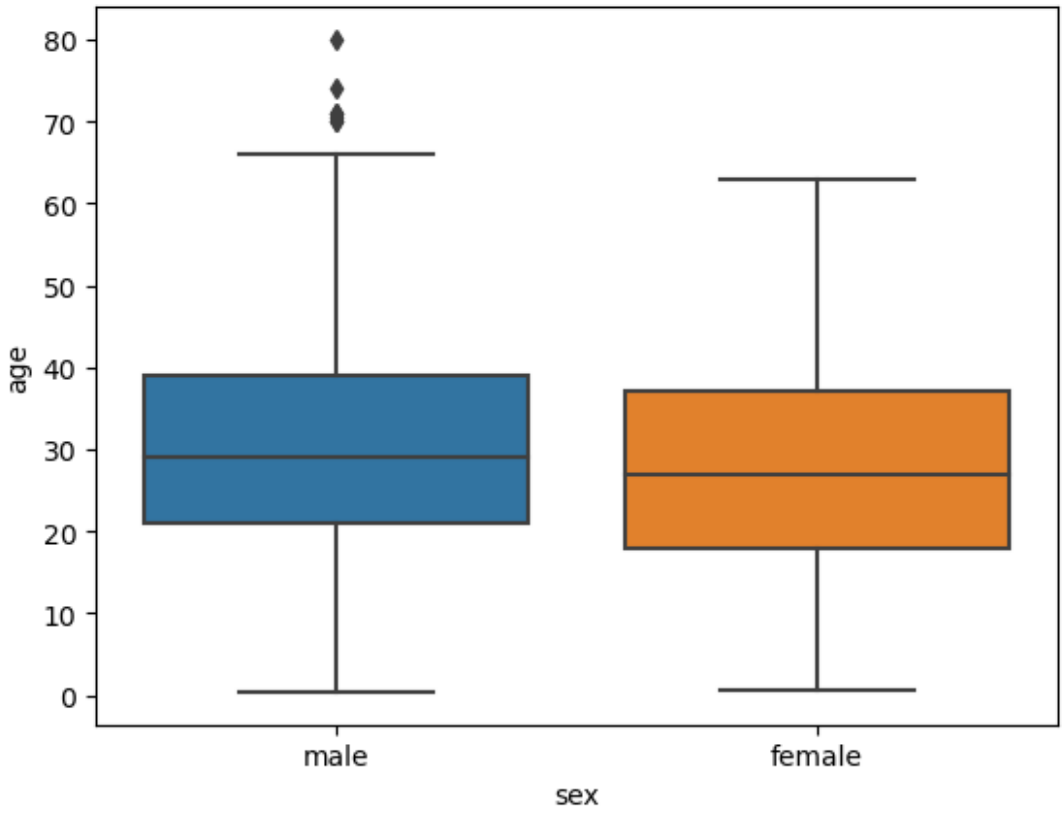

```
sns.barplot(x=data['pclass'], y=data['age'], hue=data['sex'])
```

```
<Axes: xlabel='pclass', ylabel='age'>
```



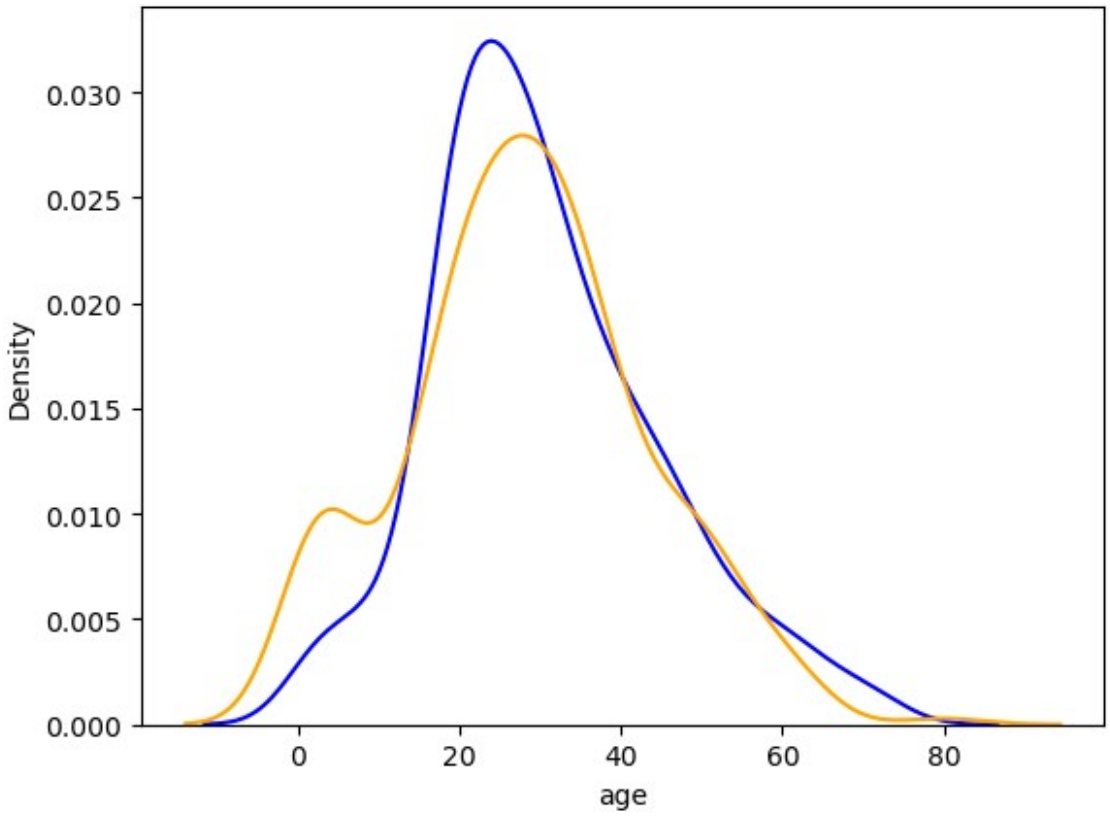

```
sns.boxplot(x=data['sex'], y=data["age"])
```

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




```
sns.distplot(data[data['survived'] == 0]['age'], hist=False,
color="blue")
sns.distplot(data[data['survived'] == 1]['age'], hist=False,
color="orange")
```

```
<Axes: xlabel='age', ylabel='Density'>
```

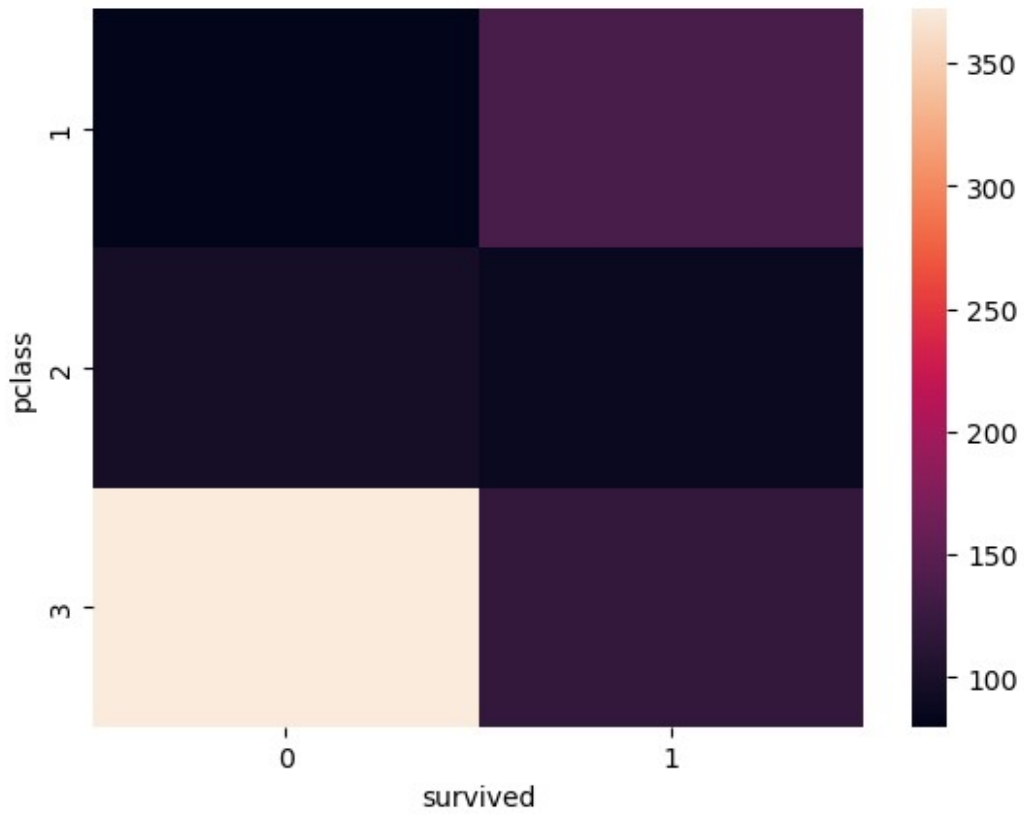



```
pd.crosstab(data['pclass'], data['survived'])
```

```
survived    0    1
pclass
1          80  136
2          97   87
3         372  119
```

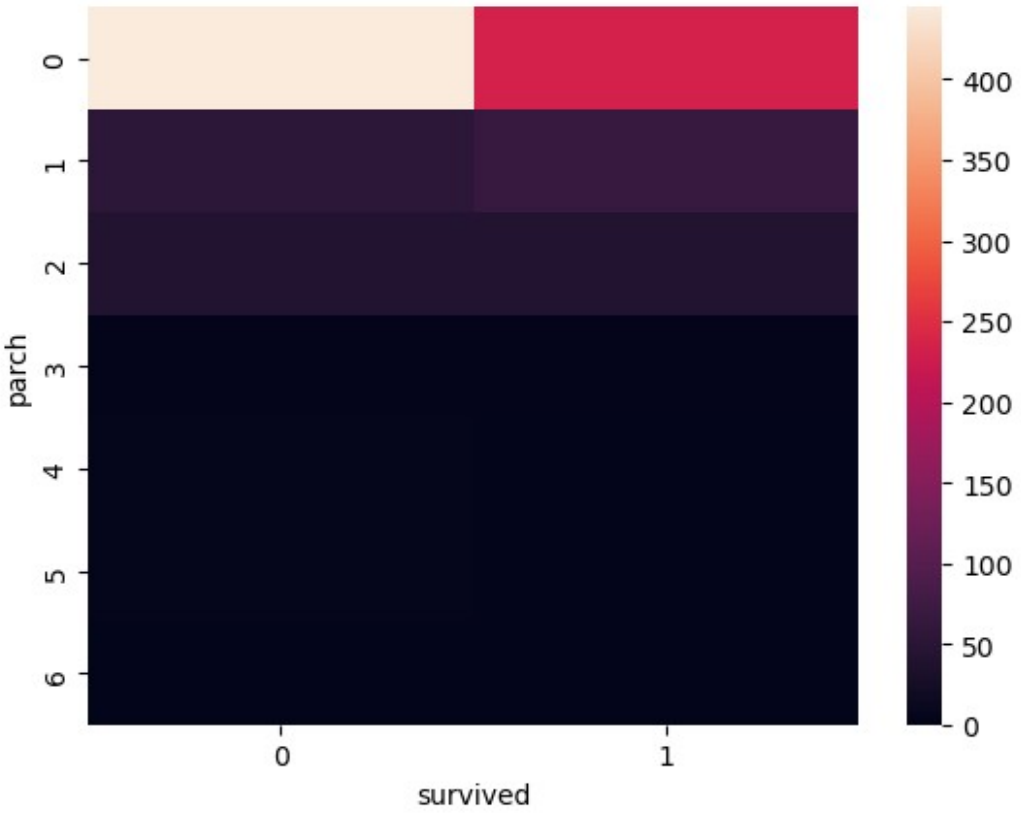
```
sns.heatmap(pd.crosstab(data['pclass'], data['survived']))
```

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




```
sns.heatmap(pd.crosstab(data['parch'], data['survived']))
```

```
<Axes: xlabel='survived', ylabel='parch'>
```


parch
0
1
2
3
4
5
6
0
1
survived
400
350
300
250
200
150
100
50
0