Assignment8

March 25, 2022

[ ]:
```
import pandas as pd
import seaborn as sb
```

[ ]:
```
df = sb.load_dataset('titanic')
```

[ ]:
```
df.head()
```

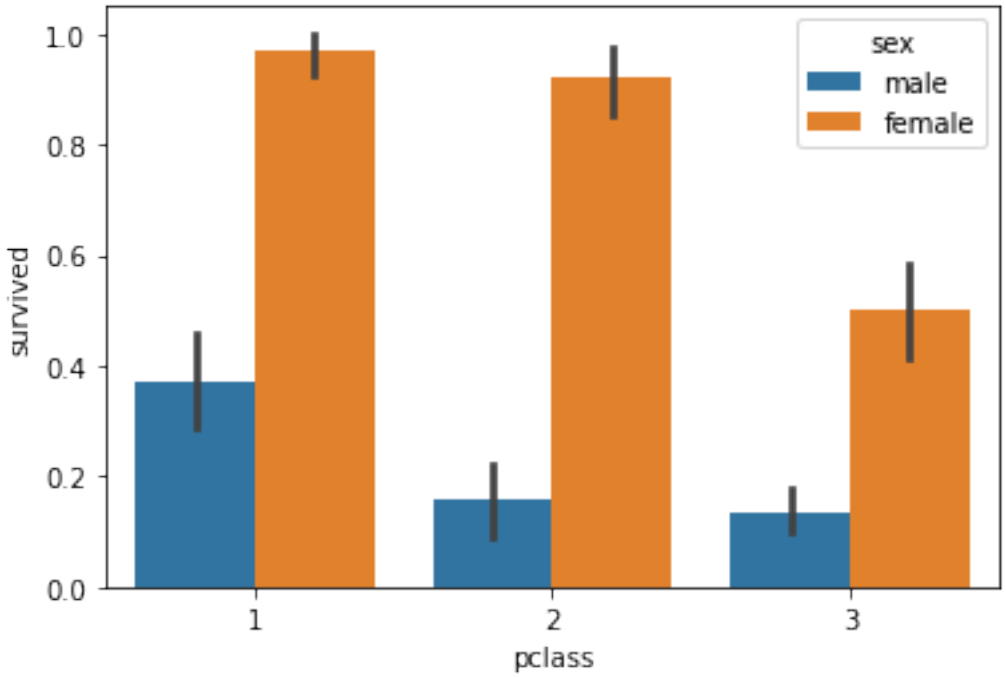
[ ]:
```
   survived  pclass     sex   age  sibsp  parch     fare embarked  class  \
0         0       3    male  22.0      1      0   7.2500        S  Third
1         1       1  female  38.0      1      0  71.2833        C  First
2         1       3  female  26.0      0      0   7.9250        S  Third
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
```

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

[ ]:
```
         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
```

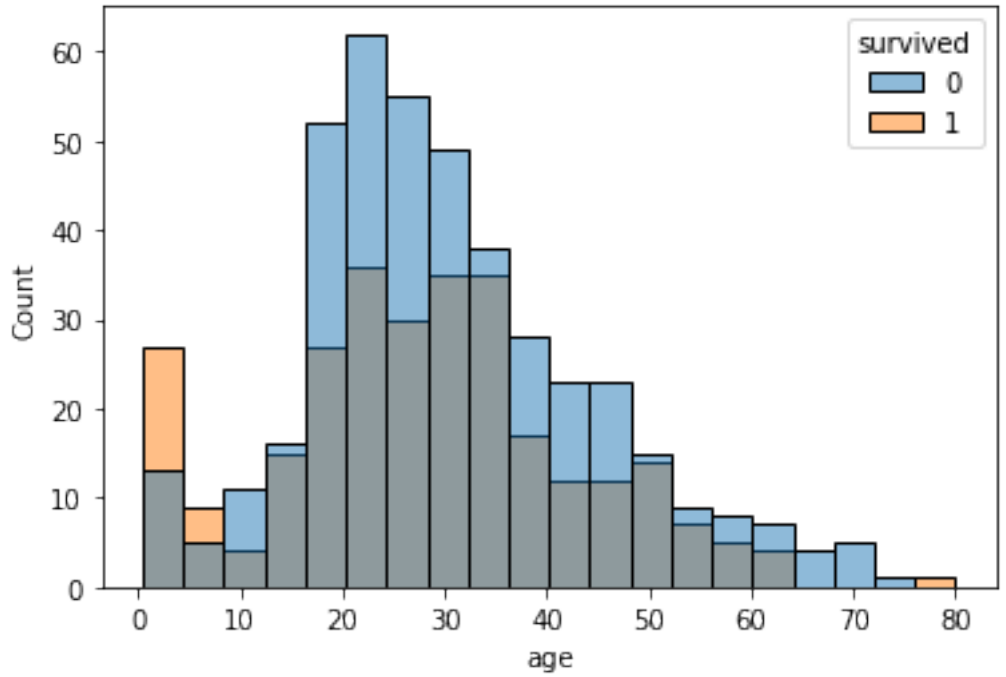
[ ]:
```
sb.barplot(data=df, hue='sex', x='pclass', y='survived')
```

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



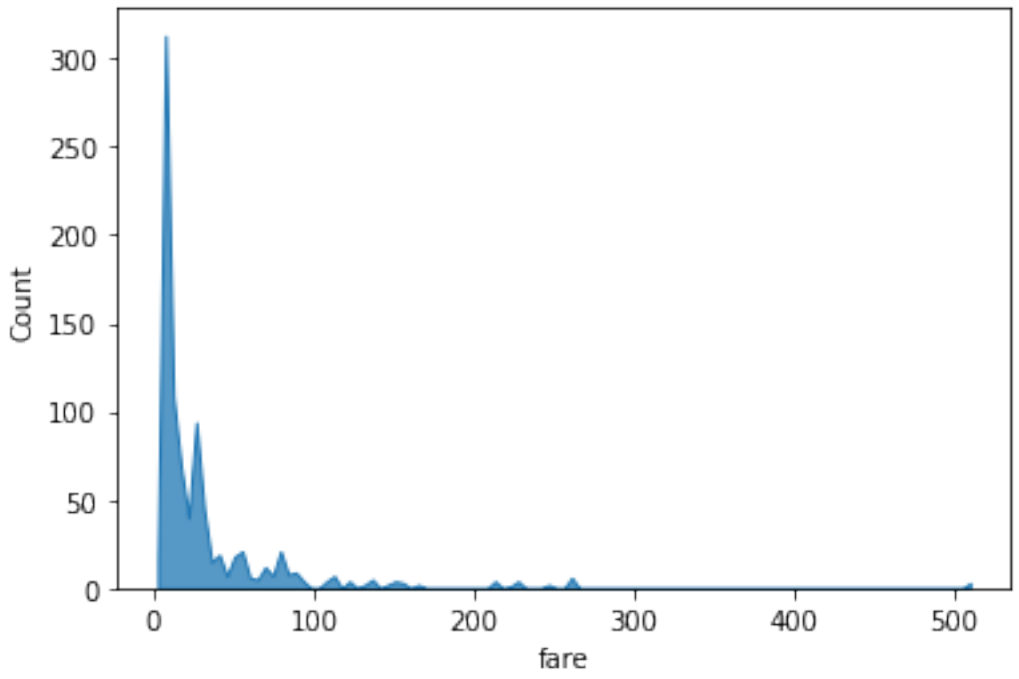
[ ]:
```
sb.histplot(data=df, x='age', hue='survived')
```

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



[ ]:
```
sb.histplot(x=df['fare'],element='poly')
```

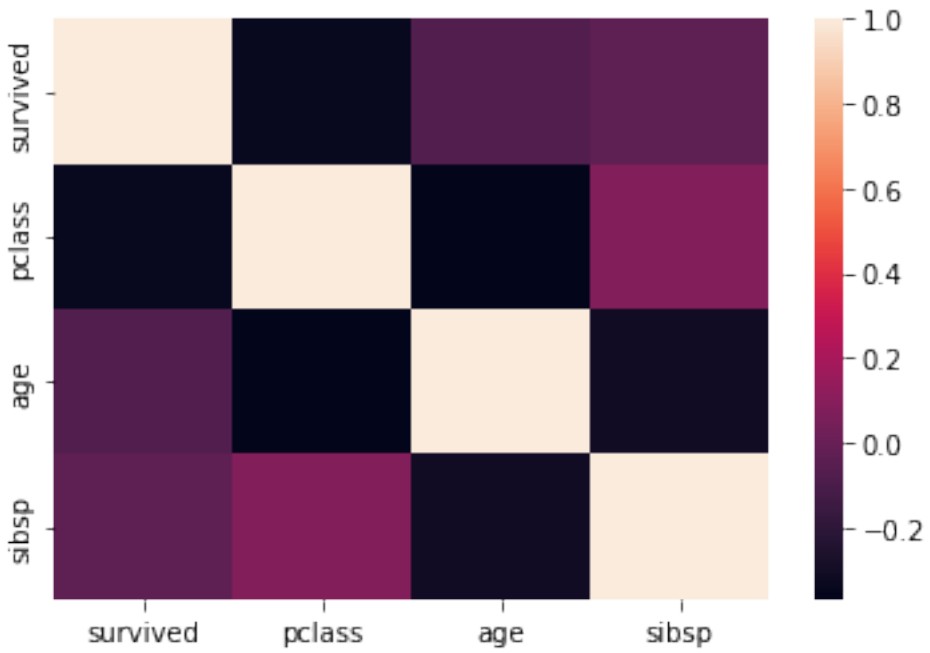
[ ]:
```
<AxesSubplot:xlabel='fare', ylabel='Count'>
```



[ ]:
```
data2 = df[['survived', 'pclass', 'sex', 'age', 'sibsp']]
```

[ ]:
```
sb.heatmap(data2.corr())
```

[ ]:
```
<AxesSubplot:>
```

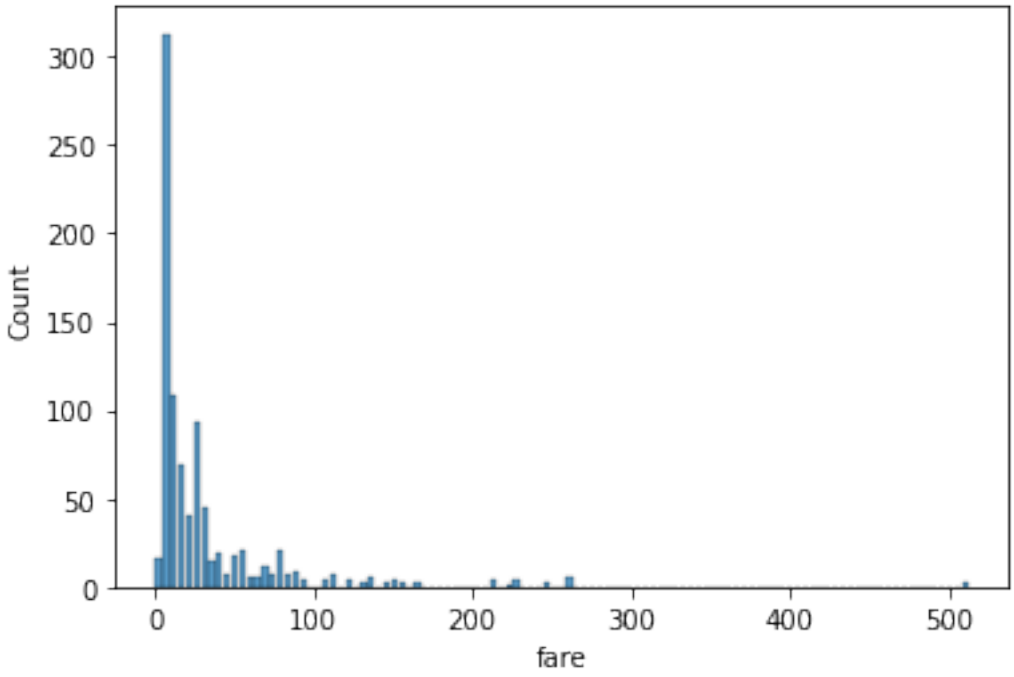


[ ]:
```
data2.corr()
```

[ ]:
```
           survived    pclass       age     sibsp
survived   1.000000 -0.338481 -0.077221 -0.035322
pclass    -0.338481  1.000000 -0.369226  0.083081
age       -0.077221 -0.369226  1.000000 -0.308247
sibsp     -0.035322  0.083081 -0.308247  1.000000
```

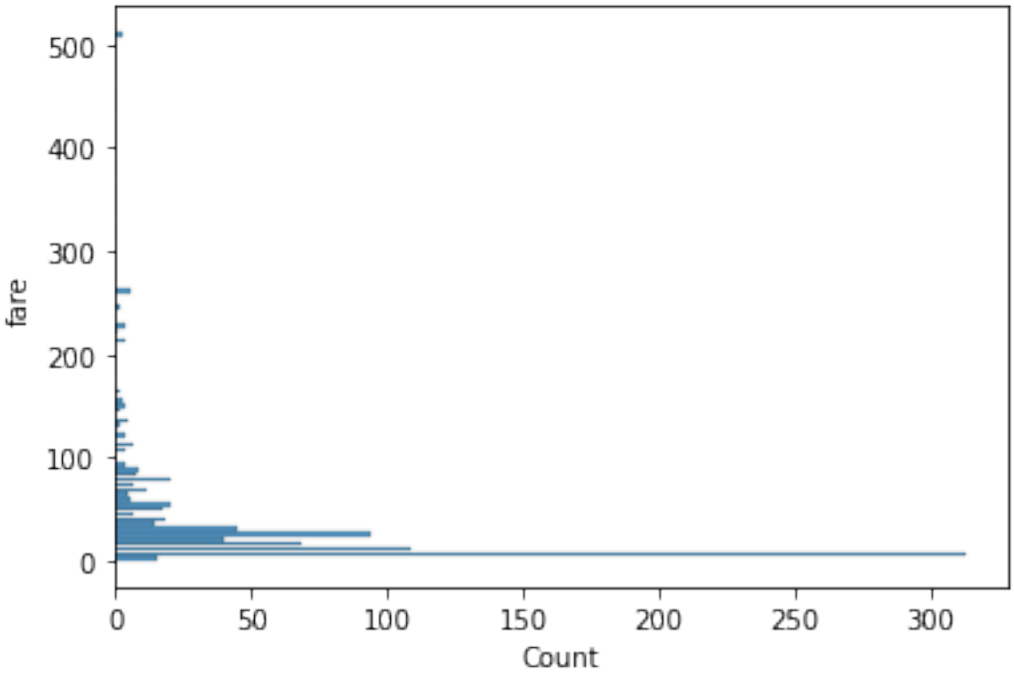
[ ]:
```
sb.histplot(data=df,x='fare',shrink=0.8)
```

[ ]:
```
<AxesSubplot:xlabel='fare', ylabel='Count'>
```



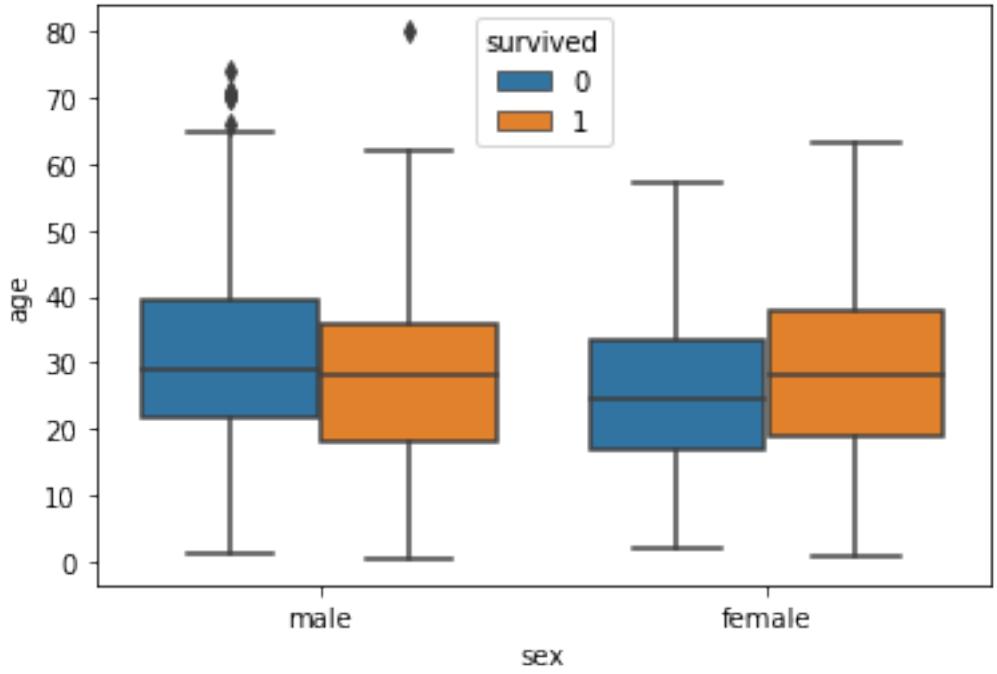
[ ]:
```
sb.histplot(data=df,y='fare',shrink=0.8)
```

[ ]: <AxesSubplot:xlabel='Count', ylabel='fare'>



[ ]:
```
sb.boxplot(data=df,x='sex', y='age', hue='survived')
```

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



[ ]:
```
df.corr()
```

[ ]:
```
             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
adult_male  -0.557080  0.094035  0.280328 -0.253586 -0.349943 -0.182024
alone       -0.203367  0.135207  0.198270 -0.584471 -0.583398 -0.271832

             adult_male     alone
survived      -0.557080 -0.203367
pclass         0.094035  0.135207
age            0.280328  0.198270
sibsp         -0.253586 -0.584471
parch         -0.349943 -0.583398
fare          -0.182024 -0.271832
```

```
adult_male    1.000000  0.404744
alone         0.404744  1.000000
```

[ ]:
```
sb.heatmap(df.corr())
```

[ ]:
```
<AxesSubplot:>
```