Assignment no - 8

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
df1 = sns.load_dataset('titanic')
df1
```

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

891 rows × 15 columns

```
df = pd.DataFrame(df1)
df
```

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

891 rows × 15 columns

```
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

```
df.info()
```

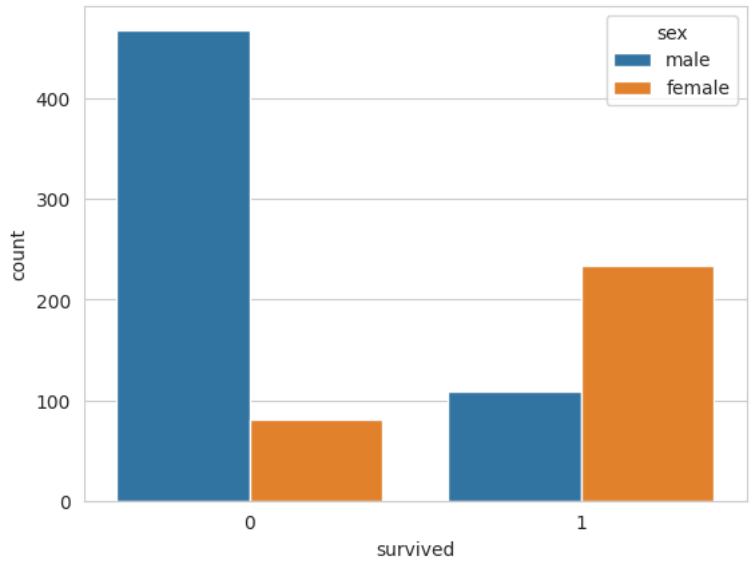
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 15 columns):
 #   Column       Non-Null Count  Dtype   
---  ------       --------------  -----   
 0   survived     891 non-null    int64   
 1   pclass       891 non-null    int64   
 2   sex          891 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
```

```
df.columns
```

```
Index(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
       'alive', 'alone'],
      dtype='object')
```

```
sns.set_style('whitegrid')
sns.countplot(x='survived', data=df, hue='sex')
```

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

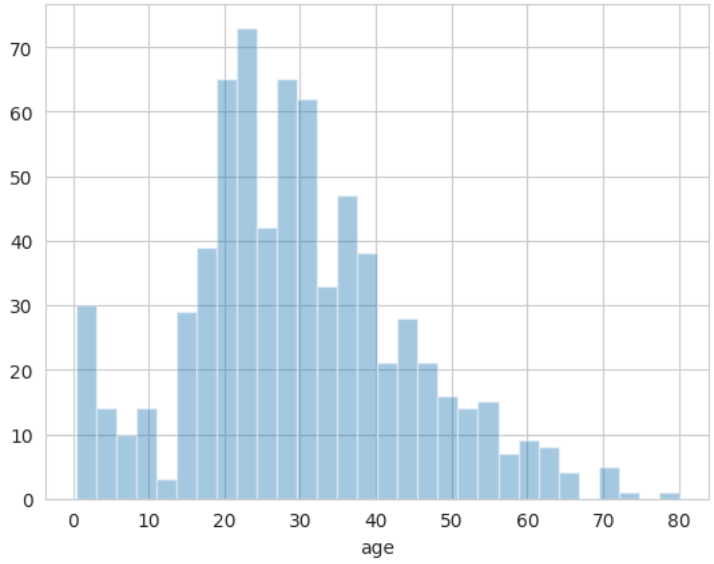



```
sns.countplot(x='survived', data=df, hue='class')
```

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

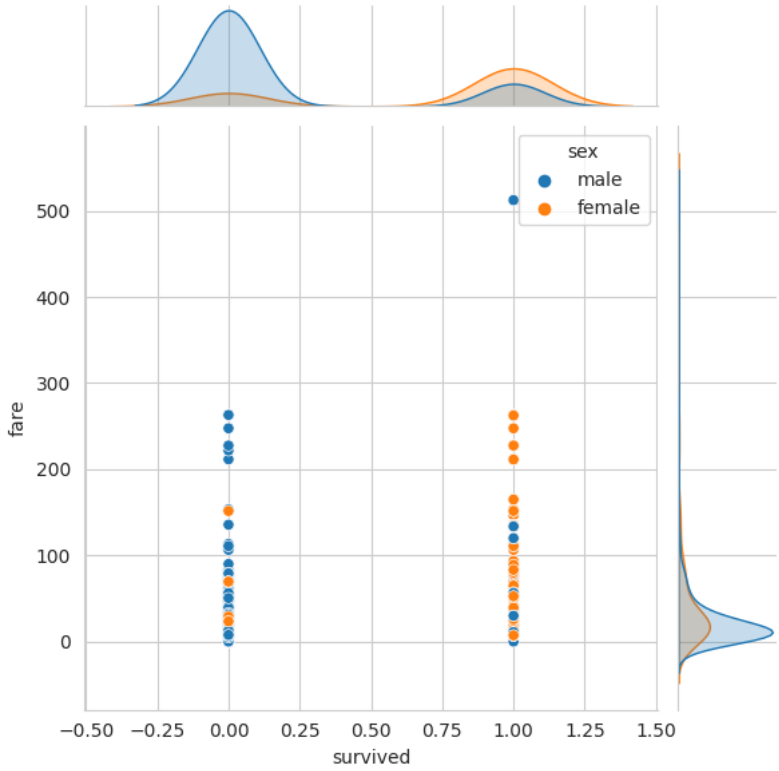

```
sns.distplot(df['age'].dropna() , kde = False , bins=30)
```

```
<ipython-input-26-49f290fce869>: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'].dropna() , kde = False , bins=30)
<Axes: xlabel='age'>
```



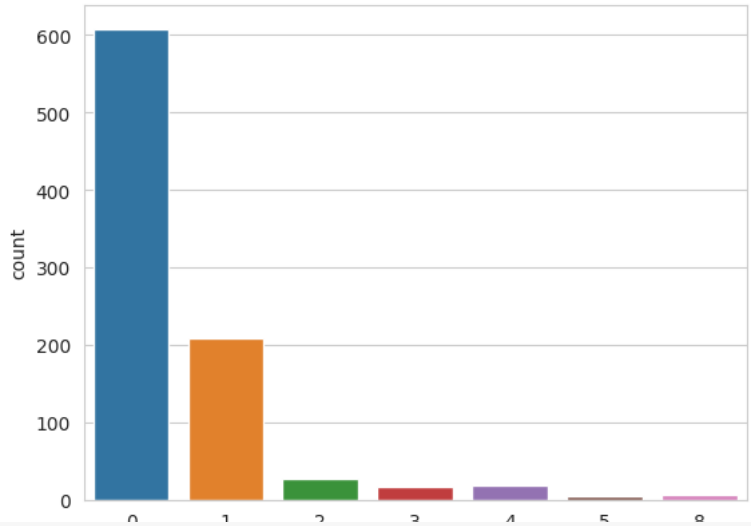

```
sns.jointplot(x = 'survived' , y = 'fare' , data = df , hue = 'sex')
```

```
<seaborn.axisgrid.JointGrid at 0x7f5720172460>
```



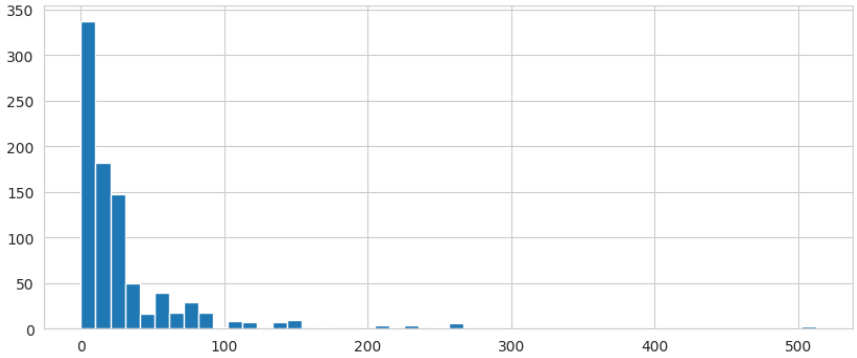

```
sns.countplot(x='sibsp' , data= df)
```

```
<Axes: xlabel='sibsp', ylabel='count'>
```



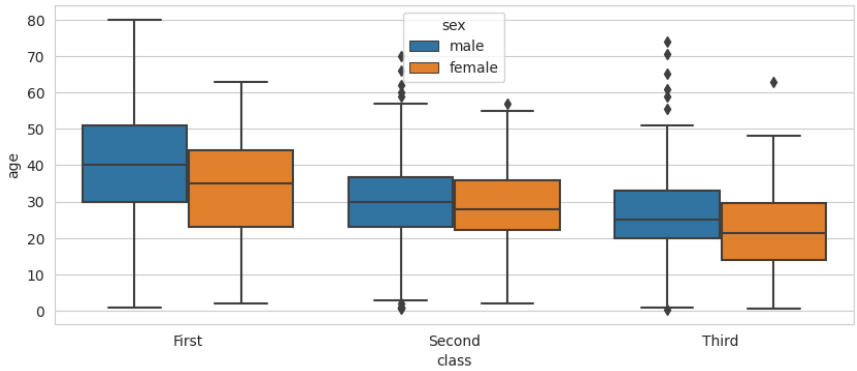

```
df['fare'].hist(bins = 50 , figsize = (10 , 4))
```

```
<Axes: >
```



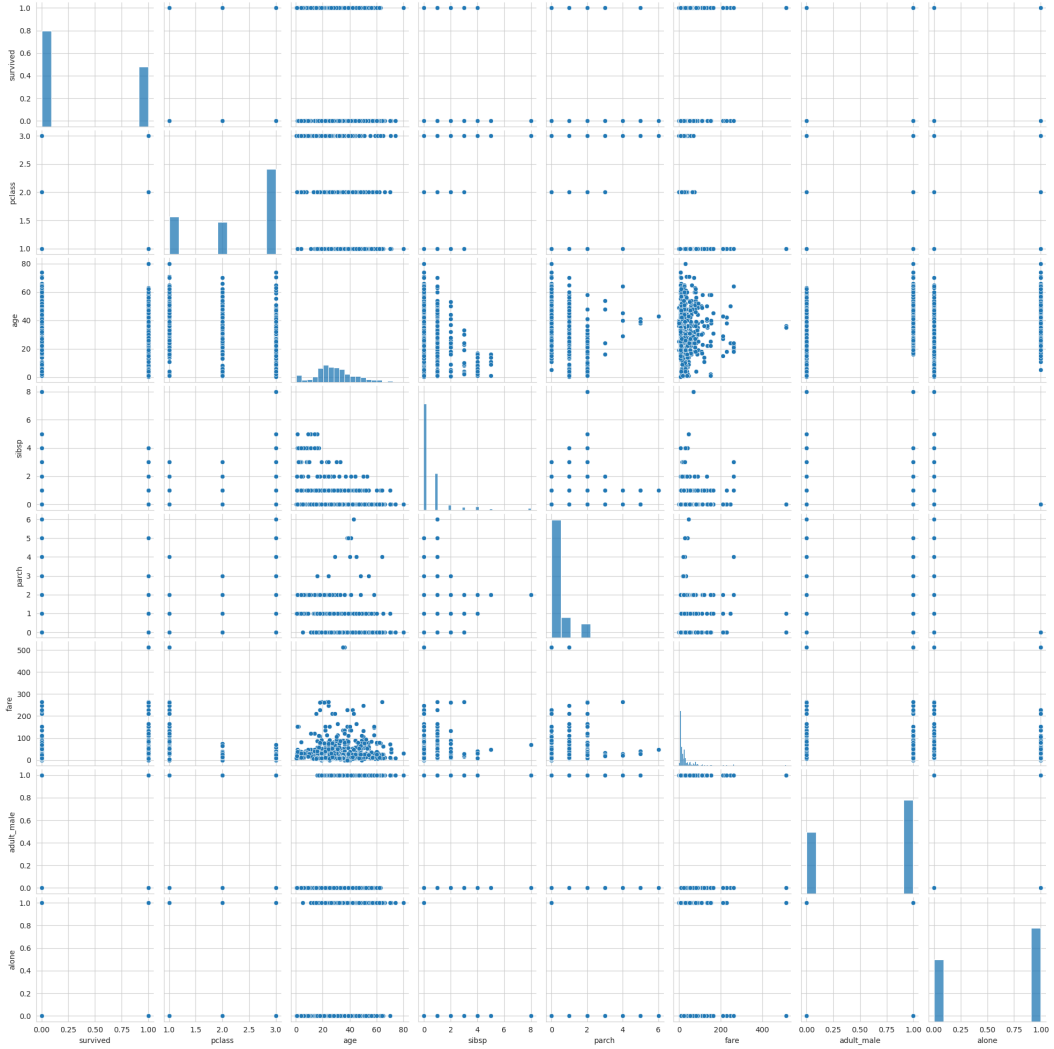

```
plt.figure(figsize= (10 , 4))
sns.boxplot(x = 'class' , y = 'age' , data = df , hue = 'sex')
```

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




```
sns.pairplot(df)
```

```
<__array_function__ internals>:180: RuntimeWarning: Converting input from bool to <class 'numpy.uint8'>
<__array_function__ internals>:180: RuntimeWarning: Converting input from bool to <class 'numpy.uint8'>
<seaborn.axisgrid.PairGrid at 0x7f571d9af430>
```




```
df['fare'].max()
```

```
512.3292
```

```
df['fare'].min()
```

```
0.0
```