In [5]:
```
import pandas as pd
import seaborn as sns
```

In [9]:
```
data_set_name=sns.get_dataset_names()
print(data_set_name)
```

```
['anagrams', 'anscombe', 'attention', 'brain_networks', 'car_crashes', 'diamo
nds', 'dots', 'dowjones', 'exercise', 'flights', 'fmri', 'geyser', 'glue', 'h
ealthexp', 'iris', 'mpg', 'penguins', 'planets', 'seaice', 'taxis', 'tips',
'titanic']
```

In [21]:
```
df=sns.load_dataset("titanic")
df
```

Out[21]:

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

891 rows × 15 columns

In [24]:
```
df.head(n=5)
```

Out[24]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	d
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	N
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	N
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	N

In [25]:
```
df.tail(n=5)
```

Out[25]:

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

In [26]:
```
df.index
```

Out[26]:
```
RangeIndex(start=0, stop=891, step=1)
```

In [27]:
```
df.columns
```

Out[27]:
```
Index(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
       'alive', 'alone'],
      dtype='object')
```

In [29]:
```
df.shape
```

Out[29]:
```
(891, 15)
```

In [30]:
```
df.dtypes
```

Out[30]:
```
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 [31]:
```
df.columns.values
```

Out[31]:
```
array(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
       'alive', 'alone'], dtype=object)
```

In [32]:
```
df.describe(include='all')
```

Out[32]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked
count	891.000000	891.000000	891	714.000000	891.000000	891.000000	891.000000	889
unique	NaN	NaN	2	NaN	NaN	NaN	NaN	3
top	NaN	NaN	male	NaN	NaN	NaN	NaN	S
freq	NaN	NaN	577	NaN	NaN	NaN	NaN	644
mean	0.383838	2.308642	NaN	29.699118	0.523008	0.381594	32.204208	NaN
std	0.486592	0.836071	NaN	14.526497	1.102743	0.806057	49.693429	NaN
min	0.000000	1.000000	NaN	0.420000	0.000000	0.000000	0.000000	NaN
25%	0.000000	2.000000	NaN	20.125000	0.000000	0.000000	7.910400	NaN
50%	0.000000	3.000000	NaN	28.000000	0.000000	0.000000	14.454200	NaN
75%	1.000000	3.000000	NaN	38.000000	1.000000	0.000000	31.000000	NaN
max	1.000000	3.000000	NaN	80.000000	8.000000	6.000000	512.329200	NaN

In [57]:
```
df['survived']
```

Out[57]:
```
0      0
1      1
2      1
3      1
4      0
      ..
886    0
887    1
888    0
889    1
890    0
Name: survived, Length: 891, dtype: int64
```

In [37]:
```
df.sort_index(axis=1,
ascending=False)
```

Out[37]:

	who	survived	sibsp	sex	pclass	parch	fare	embarked	embark_town	deck	c
0	man	0	1	male	3	0	7.2500	S	Southampton	NaN	T
1	woman	1	1	female	1	0	71.2833	C	Cherbourg	C	
2	woman	1	0	female	3	0	7.9250	S	Southampton	NaN	T
3	woman	1	1	female	1	0	53.1000	S	Southampton	C	
4	man	0	0	male	3	0	8.0500	S	Southampton	NaN	T
...	...	...	...	...	...	...	...	...	...	...	
886	man	0	0	male	2	0	13.0000	S	Southampton	NaN	Sec
887	woman	1	0	female	1	0	30.0000	S	Southampton	B	
888	woman	0	1	female	3	2	23.4500	S	Southampton	NaN	T
889	man	1	0	male	1	0	30.0000	C	Cherbourg	C	
890	man	0	0	male	3	0	7.7500	Q	Queenstown	NaN	T

891 rows × 15 columns

In [56]:
```
df.sort_values(by="survived")
```

Out[56]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_ma
0	0	3	male	22.0	1	0	7.2500	S	Third	man	Tru
519	0	3	male	32.0	0	0	7.8958	S	Third	man	Tru
521	0	3	male	22.0	0	0	7.8958	S	Third	man	Tru
522	0	3	male	NaN	0	0	7.2250	C	Third	man	Tru
524	0	3	male	NaN	0	0	7.2292	C	Third	man	Tru
...	...	...	...	...	...	...	...	...	...	...	
546	1	2	female	19.0	1	0	26.0000	S	Second	woman	Fals
215	1	1	female	31.0	1	0	113.2750	C	First	woman	Fals
216	1	3	female	27.0	0	0	7.9250	S	Third	woman	Fals
218	1	1	female	32.0	0	0	76.2917	C	First	woman	Fals
445	1	1	male	4.0	0	2	81.8583	S	First	child	Fals

891 rows × 15 columns

In [41]:
```
df.iloc[5]
```

Out[41]:
```
survived               0
pclass                 3
sex                 male
age                  NaN
sibsp                  0
parch                  0
fare              8.4583
embarked               Q
class              Third
who                  man
adult_male          True
deck                 NaN
embark_town    Queenstown
alive                  no
alone                True
Name: 5, dtype: object
```

In [42]:
```
df[0:3]
```

Out[42]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	d
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	N
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	N

In [58]:
```
df.loc[:, ["survived","embark_town"]]
```

Out[58]:

	survived	embark_town
0	0	Southampton
1	1	Cherbourg
2	1	Southampton
3	1	Southampton
4	0	Southampton
...	...	...
886	0	Southampton
887	1	Southampton
888	0	Southampton
889	1	Cherbourg
890	0	Queenstown

891 rows × 2 columns

In [49]:
```
df.iloc[:5, :]
```

Out[49]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	d
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	N
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	N
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	N

In [59]:
```
df.iloc[:, :4]
```

Out[59]:

	survived	pclass	sex	age
0	0	3	male	22.0
1	1	1	female	38.0
2	1	3	female	26.0
3	1	1	female	35.0
4	0	3	male	35.0
...	...	...	...	...
886	0	2	male	27.0
887	1	1	female	19.0
888	0	3	female	NaN
889	1	1	male	26.0
890	0	3	male	32.0

891 rows × 4 columns

In [61]:
```
df.iloc[:3, :7]
```

Out[61]:

	survived	pclass	sex	age	sibsp	parch	fare
0	0	3	male	22.0	1	0	7.2500
1	1	1	female	38.0	1	0	71.2833
2	1	3	female	26.0	0	0	7.9250

In [ ]:
```

```

In [73]:
```
df.iloc[3:5, 0:2]
```

Out[73]:

	survived	pclass
3	1	1
4	0	3

In [74]:
```
df.iloc[[1, 2,4], [0, 2]]
```

Out[74]:

	survived	sex
1	1	female
2	1	female
4	0	male

In [75]:
```
df.iloc[1:3, :]
```

Out[75]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	d
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	N

In [76]:
```
df.iloc[:, 1:3]
```

Out[76]:

	pclass	sex
0	3	male
1	1	female
2	3	female
3	1	female
4	3	male
...	...	...
886	2	male
887	1	female
888	3	female
889	1	male
890	3	male

891 rows × 2 columns

In [77]:
```
df.iloc[1, 1]
```

Out[77]:
```
1
```

In [79]:
```
df['sex'].iloc[5]
```

Out[79]:
```
'male'
```

In [81]:
```
cols_2_4=df.columns[2:4]
df[cols_2_4]
```

Out[81]:

	sex	age
0	male	22.0
1	female	38.0
2	female	26.0
3	female	35.0
4	male	35.0
...	...	...
886	male	27.0
887	female	19.0
888	female	NaN
889	male	26.0
890	male	32.0

891 rows × 2 columns

In [83]:
```
df[df.columns[2:4]].iloc[5:10]
```

Out[83]:

	sex	age
5	male	NaN
6	male	54.0
7	male	2.0
8	female	27.0
9	female	14.0

In [ ]:
```

```

In [67]:
```
df.isnull()
```

Out[67]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	dec
0	False	False	False	False	False	False	False	False	False	False	False	Tru
1	False	False	False	False	False	False	False	False	False	False	False	Fals
2	False	False	False	False	False	False	False	False	False	False	False	Tru
3	False	False	False	False	False	False	False	False	False	False	False	Fals
4	False	False	False	False	False	False	False	False	False	False	False	Tru
...	...	...	...	...	...	...	...	...	...	...	...	
886	False	False	False	False	False	False	False	False	False	False	False	Tru
887	False	False	False	False	False	False	False	False	False	False	False	Fals
888	False	False	False	True	False	False	False	False	False	False	False	Tru
889	False	False	False	False	False	False	False	False	False	False	False	Fals
890	False	False	False	False	False	False	False	False	False	False	False	Tru

891 rows × 15 columns

In [68]:
```
df.isnull().any()
```

Out[68]:
```
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
```

In [69]:
```
df.isnull().sum().sum()
```

Out[69]:
```
869
```

In [71]:
```
df.isnull().sum(axis = 1)
```

Out[71]:
```
0      1
1      0
2      1
3      0
4      1
      ..
886    1
887    0
888    2
889    0
890    1
Length: 891, dtype: int64
```

In [72]:
```
df.isnull().sum()
```

Out[72]:
```
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
```

In [84]:
```
df.isna().sum()
```

Out[84]:
```
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
```

In [91]:
```
df.sex.isnull().sum()
```

Out[91]:
```
0
```

In [92]:
```
df.groupby(['sex'])['age'].apply(lambda x:x.isnull().sum())
```

Out[92]:
```
sex
female     53
male      124
Name: age, dtype: int64
```

In [ ]:

Name- Anurag Jadhav
Roll No-12371
Practical-1
Class- TE A