In [1]:
```
import pandas as pd
import seaborn as sns
```

In [2]:
```
datasetnames=sns.get_dataset_names()
```

In [3]:
```
datasetnames
```

Out[3]:
```
['anagrams',
 'anscombe',
 'attention',
 'brain_networks',
 'car_crashes',
 'diamonds',
 'dots',
 'dowjones',
 'exercise',
 'flights',
 'fmri',
 'geyser',
 'glue',
 'healthexp',
 'iris',
 'mpg',
 'penguins',
 'planets',
 'seaice',
 'taxis',
 'tips',
 'titanic']
```

In [4]:
```
print(datasetnames)
```

```
['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 [40]:
```
df1=sns.load_dataset("titanic")
df1
```

Out[40]:

	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 [41]:
```
df1.head()
```

Out[41]:

	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 [42]:
```
df1.tail(n=4)
```

Out[42]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	d
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	N
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	N

In [43]:
```
df1.index
```

Out[43]:
```
RangeIndex(start=0, stop=891, step=1)
```

In [44]:
```
df1.columns
```

Out[44]:
```
Index(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
       'alive', 'alone'],
      dtype='object')
```

In [45]:
```
df1.shape
```

Out[45]:
```
(891, 15)
```

In [46]:
```
df1.dtypes
```

Out[46]:
```
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 [47]:
```
df1.columns.values
```

Out[47]:
```
array(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
       'alive', 'alone'], dtype=object)
```

In [48]:
```
df1.describe(include='all')
```

Out[48]:

	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 [49]:
```
df1['pclass']
```

Out[49]:
```
0      3
1      1
2      3
3      1
4      3
      ..
886    2
887    1
888    3
889    1
890    3
Name: pclass, Length: 891, dtype: int64
```

In [50]:
```
df1.sort_index(axis=1,ascending=False)
```

Out[50]:

	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 [51]:
```
df1.sort_values(by="pclass")
```

Out[51]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male
445	1	1	male	4.0	0	2	81.8583	S	First	child	False
310	1	1	female	24.0	0	0	83.1583	C	First	woman	False
309	1	1	female	30.0	0	0	56.9292	C	First	woman	False
307	1	1	female	17.0	1	0	108.9000	C	First	woman	False
306	1	1	female	NaN	0	0	110.8833	C	First	woman	False
...	...	...	...	...	...	...	...	...	...	...	...
379	0	3	male	19.0	0	0	7.7750	S	Third	man	True
381	1	3	female	1.0	0	2	15.7417	C	Third	child	False
382	0	3	male	32.0	0	0	7.9250	S	Third	man	True
371	0	3	male	18.0	1	0	6.4958	S	Third	man	True
890	0	3	male	32.0	0	0	7.7500	Q	Third	man	True

891 rows × 15 columns

In [52]:
```
df1.iloc[5]
```

Out[52]:
```
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 [53]:
```
df1[0:3]
```

Out[53]:

	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 [54]:
```
df1.loc[:,["survived","pclass"]]
```

Out[54]:

	survived	pclass
0	0	3
1	1	1
2	1	3
3	1	1
4	0	3
...	...	...
886	0	2
887	1	1
888	0	3
889	1	1
890	0	3

891 rows × 2 columns

In [55]:
```
df1.iloc[:, :3]
```

Out[55]:

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

891 rows × 3 columns

In [56]:
```
df1.iloc[:3,:5]
```

Out[56]:

	survived	pclass	sex	age	sibsp
0	0	3	male	22.0	1
1	1	1	female	38.0	1
2	1	3	female	26.0	0

In [57]:
```
df1.isnull()
```

Out[57]:

	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 [58]:
```
df1.isnull().any()
```

Out[58]:
```
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 [59]:
```
df1.isnull().sum().sum()
```

Out[59]:
```
869
```

In [60]:
```
df1.isnull().sum(axis=1)
```

Out[60]:
```
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 [61]:
```
df1.isnull().sum()
```

Out[61]:
```
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 [62]:
```
df1.pclass.isnull().sum()
```

Out[62]:
```
0
```

In [63]:
```
df1.groupby(['pclass'])['age'].apply(lambda x:x.isnull().sum())
```

Out[63]:
```
pclass
1     30
2     11
3    136
Name: age, dtype: int64
```

In [64]:
```
df1.iloc[3:5,0:2]
```

Out[64]:

	survived	pclass
3	1	1
4	0	3

In [65]:
```
df1.iloc[[1,2,4],[0,2]]
```

Out[65]:

	survived	sex
1	1	female
2	1	female
4	0	male

In [66]:
```
df1.iloc[1:3,:]
```

Out[66]:

	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 [67]:
```
df1.iloc[:,1:3]
```

Out[67]:

	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 [68]:
```
df1.iloc[1,1]
```

Out[68]: 1

In [69]:
```
df1['pclass'].iloc[5]
```

Out[69]: 3

In [71]:
```
df[df.columns[2:4]].iloc[5:10]
```

Out[71]:

	petal_length	petal_width
5	1.7	0.4
6	1.4	0.3
7	1.5	0.2
8	1.4	0.2
9	1.5	0.1

In [73]:
```
df2=sns.load_dataset('iris')
df
```

Out[73]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	virginica
146	6.3	2.5	5.0	1.9	virginica
147	6.5	3.0	5.2	2.0	virginica
148	6.2	3.4	5.4	2.3	virginica
149	5.9	3.0	5.1	1.8	virginica

150 rows × 5 columns

```
# NAME:TANMAY DIXIT
# ROLL NO:13143
```