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

In [7]:
```
data_set_name=sb.get_dataset_names()
print(data_set_name)
```
```
['anagrams', 'anscombe', 'attention', 'brain_networks', 'car_crashes', 'diamonds', 'dots', 'dowjones', 'exercise', 'flights', 'fmri', 'gey
ser', 'glue', 'healthexp', 'iris', 'mpg', 'penguins', 'planets', 'seaice', 'taxis', 'tips', 'titanic']
```

In [8]:
```
df=sb.load_dataset("titanic")
df
```
Out[8]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	Southampton	yes	False
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton	no	True
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
886	0	2	male	27.0	0	0	13.0000	S	Second	man	True	NaN	Southampton	no	True
887	1	1	female	19.0	0	0	30.0000	S	First	woman	False	B	Southampton	yes	True
888	0	3	female	NaN	1	2	23.4500	S	Third	woman	False	NaN	Southampton	no	False
889	1	1	male	26.0	0	0	30.0000	C	First	man	True	C	Cherbourg	yes	True
890	0	3	male	32.0	0	0	7.7500	Q	Third	man	True	NaN	Queenstown	no	True

891 rows × 15 columns

In [9]:
```
df.head(n=5)
```
Out[9]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	Southampton	yes	False
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton	no	True

In [10]:
```
df.tail(n=5)
```
Out[10]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
886	0	2	male	27.0	0	0	13.00	S	Second	man	True	NaN	Southampton	no	True
887	1	1	female	19.0	0	0	30.00	S	First	woman	False	B	Southampton	yes	True
888	0	3	female	NaN	1	2	23.45	S	Third	woman	False	NaN	Southampton	no	False
889	1	1	male	26.0	0	0	30.00	C	First	man	True	C	Cherbourg	yes	True
890	0	3	male	32.0	0	0	7.75	Q	Third	man	True	NaN	Queenstown	no	True

In [11]:
```
df.index
```
Out[11]:
```
RangeIndex(start=0, stop=891, step=1)
```

In [12]:
```
df.columns
```
Out[12]:
```
Index(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
       'alive', 'alone'],
      dtype='object')
```

In [13]:
```
df.shape
```
Out[13]:
```
(891, 15)
```

In [14]:
```
df.dtypes
```
Out[14]:
```
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 [15]:
```
df.columns.values
```
Out[15]:
```
array(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
       'alive', 'alone'], dtype=object)
```

In [16]:
```
df.describe(include='all')
```
Out[16]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
count	891.000000	891.000000	891	714.000000	891.000000	891.000000	891.000000	889	891	891	891	203	889	891	891
unique	NaN	NaN	2	NaN	NaN	NaN	NaN	3	3	3	2	7	3	2	2
top	NaN	NaN	male	NaN	NaN	NaN	NaN	S	Third	man	True	C	Southampton	no	True
freq	NaN	NaN	577	NaN	NaN	NaN	NaN	644	491	537	537	59	644	549	537
mean	0.383838	2.308642	NaN	29.699118	0.523008	0.381594	32.204208	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
std	0.486592	0.836071	NaN	14.526497	1.102743	0.806057	49.693429	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
min	0.000000	1.000000	NaN	0.420000	0.000000	0.000000	0.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
25%	0.000000	2.000000	NaN	20.125000	0.000000	0.000000	7.910400	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
50%	0.000000	3.000000	NaN	28.000000	0.000000	0.000000	14.454200	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75%	1.000000	3.000000	NaN	38.000000	1.000000	0.000000	31.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
max	1.000000	3.000000	NaN	80.000000	8.000000	6.000000	512.329200	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [17]:
```
df['survived']
```
Out[17]:
```
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 [18]:
```
df.sort_index(axis=1,
ascending=False)
```
Out[18]:

	who	survived	sibsp	sex	pclass	parch	fare	embarked	embark_town	deck	class	alone	alive	age	adult_male
0	man	0	1	male	3	0	7.2500	S	Southampton	NaN	Third	False	no	22.0	True
1	woman	1	1	female	1	0	71.2833	C	Cherbourg	C	First	False	yes	38.0	False
2	woman	1	0	female	3	0	7.9250	S	Southampton	NaN	Third	True	yes	26.0	False
3	woman	1	1	female	1	0	53.1000	S	Southampton	C	First	False	yes	35.0	False
4	man	0	0	male	3	0	8.0500	S	Southampton	NaN	Third	True	no	35.0	True
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
886	man	0	0	male	2	0	13.0000	S	Southampton	NaN	Second	True	no	27.0	True
887	woman	1	0	female	1	0	30.0000	S	Southampton	B	First	True	yes	19.0	False
888	woman	0	1	female	3	2	23.4500	S	Southampton	NaN	Third	False	no	NaN	False
889	man	1	0	male	1	0	30.0000	C	Cherbourg	C	First	True	yes	26.0	True
890	man	0	0	male	3	0	7.7500	Q	Queenstown	NaN	Third	True	no	32.0	True

891 rows × 15 columns

In [19]:
```
df.sort_values(by="survived")
```
Out[19]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
519	0	3	male	32.0	0	0	7.8958	S	Third	man	True	NaN	Southampton	no	True
521	0	3	male	22.0	0	0	7.8958	S	Third	man	True	NaN	Southampton	no	True
522	0	3	male	NaN	0	0	7.2250	C	Third	man	True	NaN	Cherbourg	no	True
524	0	3	male	NaN	0	0	7.2292	C	Third	man	True	NaN	Cherbourg	no	True
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
546	1	2	female	19.0	1	0	26.0000	S	Second	woman	False	NaN	Southampton	yes	False
215	1	1	female	31.0	1	0	113.2750	C	First	woman	False	D	Cherbourg	yes	False
216	1	3	female	27.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True
218	1	1	female	32.0	0	0	76.2917	C	First	woman	False	D	Cherbourg	yes	True
445	1	1	male	4.0	0	2	81.8583	S	First	child	False	A	Southampton	yes	False

891 rows × 15 columns

In [20]:
```
df.iloc[5]
```
Out[20]:
```
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 [21]:
```
df[0:3]
```
Out[21]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True

In [22]:
```
df.loc[:, ["survived","embark_town"]]
```
Out[22]:

	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 [23]:
```
df.iloc[:5, :]
```
Out[23]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	Southampton	yes	False
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton	no	True

In [24]:
```
df.iloc[:, :4]
```
Out[24]:

	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 [25]:
```
df.iloc[:3, :7]
```
Out[25]:

	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 [26]:
```
df.iloc[3:5, 0:2]
```
Out[26]:

	survived	pclass
3	1	1
4	0	3

In [27]:
```
df.iloc[[1, 2,4], [0, 2]]
```
Out[27]:

	survived	sex
1	1	female
2	1	female
4	0	male

In [28]:
```
df.iloc[1:3, :]
```
Out[28]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True

In [29]:
```
df.iloc[:, 1:3]
```
Out[29]:

	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 [30]:
```
df.iloc[1, 1]
```
Out[30]:
```
1
```

In [31]:
```
df['sex'].iloc[5]
```
Out[31]:
```
'male'
```

In [32]:
```
cols_2_4=df.columns[2:4]
df[cols_2_4]
```
Out[32]:

	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 [33]:
```
df[df.columns[2:4]].iloc[5:10]
```
Out[33]:

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

In [34]:
```
df.isnull()
```
Out[34]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	False	False	False	False	False	False	False	False	False	False	False	True	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False	True	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False	True	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
886	False	False	False	False	False	False	False	False	False	False	False	True	False	False	False
887	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
888	False	False	False	True	False	False	False	False	False	False	False	True	False	False	False
889	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
890	False	False	False	False	False	False	False	False	False	False	False	True	False	False	False

891 rows × 15 columns

In [35]:
```
df.isnull().any()
```
Out[35]:
```
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 [36]:
```
df.isnull().sum().sum()
```
Out[36]:
```
869
```

In [37]:
```
df.isnull().sum(axis = 1)
```
Out[37]:
```
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 [38]:
```
df.isnull().sum()
```
Out[38]:
```
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 [39]:
```
df.isna().sum()
```
Out[39]:
```
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 [40]:
```
df.sex.isnull().sum()
```
Out[40]:
```
0
```

In [41]:
```
df.groupby(['sex'])['age'].apply(lambda x:x.isnull().sum())
```
Out[41]:
```
sex
female     53
male      124
Name: age, dtype: int64
```

In [ ]:

Name: Swayambhu Balasaheb Bhapkar

Roll no: 13121