

Name: Shubham Shigwan
Email: shubhamshigwan2k@gmail.com

In []:

```
import pandas as pd
```

In []:

```
titanic = pd.read_csv('https://github.com/YBI-Foundation/Dataset/raw/main/Titanic.csv')
```

In []:

```
titanic.head()
```

Out[]:

	pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embark
0	1	1	Allen, Miss. Elisabeth Walton	female	29.00	0	0	24160	211.3375	B5	
1	1	1	Allison, Master. Hudson Trevor	male	0.92	1	2	113781	151.5500	C22 C26	
2	1	0	Allison, Miss. Helen Lorraine	female	2.00	1	2	113781	151.5500	C22 C26	
3	1	0	Allison, Mr. Hudson Joshua Creighton	male	30.00	1	2	113781	151.5500	C22 C26	
4	1	0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	25.00	1	2	113781	151.5500	C22 C26	

In []:

titanic.head(2)

Out[]:

	pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embark
0	1	1	Allen, Miss. Elisabeth Walton	female	29.00	0	0	24160	211.3375	B5	
1	1	1	Allison, Master. Hudson Trevor	male	0.92	1	2	113781	151.5500	C22 C26	

In []:

titanic.tail(2)

Out[]:

	pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarl
1307	3	0	Zakarian, Mr. Ortin	male	27.0	0	0	2670	7.225	NaN	
1308	3	0	Zimmerman, Mr. Leo	male	29.0	0	0	315082	7.875	NaN	

In []:

titanic.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1309 entries, 0 to 1308
Data columns (total 14 columns):
 #   Column      Non-Null Count  Dtype  
 --- 
 0   pclass      1309 non-null   object 
 1   survived    1309 non-null   int64  
 2   name        1309 non-null   object 
 3   sex         1309 non-null   object 
 4   age         1046 non-null   float64
 5   sibsp       1309 non-null   int64  
 6   parch       1309 non-null   int64  
 7   ticket      1309 non-null   object 
 8   fare         1308 non-null   float64
 9   cabin       295 non-null   object  
 10  embarked    1307 non-null   object 
 11  boat        486 non-null   object  
 12  body        121 non-null   float64
 13  home.dest   745 non-null   object  
dtypes: float64(3), int64(3), object(8)
memory usage: 143.3+ KB
```

In []:

titanic.describe()

Out[]:

	pclass	survived	age	sibsp	parch	fare	
count	1309.000000	1309.000000	1046.000000	1309.000000	1309.000000	1308.000000	121.0
mean	2.294882	0.381971	29.881138	0.498854	0.385027	33.295479	160.8
std	0.837836	0.486055	14.413493	1.041658	0.865560	51.758668	97.6
min	1.000000	0.000000	0.170000	0.000000	0.000000	0.000000	1.0
25%	2.000000	0.000000	21.000000	0.000000	0.000000	7.895800	72.0
50%	3.000000	0.000000	28.000000	0.000000	0.000000	14.454200	155.0
75%	3.000000	1.000000	39.000000	1.000000	0.000000	31.275000	256.0
max	3.000000	1.000000	80.000000	8.000000	9.000000	512.329200	328.0

In []:

titanic.describe(include='all')

Out[]:

	pclass	survived	name	sex	age	sibsp	parch	t
count	1309.000000	1309.000000	1309	1309	1046.000000	1309.000000	1309.000000	
unique	Nan	Nan	1307	2	Nan	Nan	Nan	
top	Nan	Nan	Connolly, Miss. Kate	male	Nan	Nan	Nan	
freq	Nan	Nan	2	843	Nan	Nan	Nan	
mean	2.294882	0.381971	Nan	Nan	29.881138	0.498854	0.385027	
std	0.837836	0.486055	Nan	Nan	14.413493	1.041658	0.865560	
min	1.000000	0.000000	Nan	Nan	0.170000	0.000000	0.000000	
25%	2.000000	0.000000	Nan	Nan	21.000000	0.000000	0.000000	
50%	3.000000	0.000000	Nan	Nan	28.000000	0.000000	0.000000	
75%	3.000000	1.000000	Nan	Nan	39.000000	1.000000	0.000000	
max	3.000000	1.000000	Nan	Nan	80.000000	8.000000	9.000000	

In []:

```
titanic.corr()
```

Out[]:

	pclass	survived	age	sibsp	parch	fare	body
pclass	1.000000	-0.312469	-0.408106	0.060832	0.018322	-0.558629	-0.034642
survived	-0.312469	1.000000	-0.055512	-0.027825	0.082660	0.244265	NaN
age	-0.408106	-0.055512	1.000000	-0.243699	-0.150917	0.178740	0.058809
sibsp	0.060832	-0.027825	-0.243699	1.000000	0.373587	0.160238	-0.099961
parch	0.018322	0.082660	-0.150917	0.373587	1.000000	0.221539	0.051099
fare	-0.558629	0.244265	0.178740	0.160238	0.221539	1.000000	-0.043110
body	-0.034642	NaN	0.058809	-0.099961	0.051099	-0.043110	1.000000

In []:

```
titanic['age']
```

Out[]:

```
0      29.00
1      0.92
2      2.00
3     30.00
4     25.00
...
1304    14.50
1305    NaN
1306    26.50
1307    27.00
1308    29.00
Name: age, Length: 1309, dtype: float64
```

In []:

```
titanic['age'].nlargest(3)
```

Out[]:

```
14      80.0
61      76.0
1235    74.0
Name: age, dtype: float64
```

In []:

```
titanic['age'].nsmallest(4)
```

Out[]:

```
763      0.17  
747      0.33  
1240     0.42  
427      0.67  
Name: age, dtype: float64
```

In []:

```
titanic['age'].max()
```

Out[]:

```
80.0
```

In []:

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

Out[]:

```
0.0
```

Change Data Type

In []:

```
titanic['pclass']= titanic['pclass'].astype('object')
```

In []:

```
titanic['pclass']= titanic['pclass'].astype('int64')
```

In []:

```
titanic['embarked'].nunique()
```

Out[]:

```
3
```

In []:

```
titanic['embarked'].unique()
```

Out[]:

```
array(['S', 'C', nan, 'Q'], dtype=object)
```

In []:

```
titanic['embarked'].value_counts()
```

Out[]:

```
S    914  
C    270  
Q    123  
Name: embarked, dtype: int64
```

In []:

```
titanic['sex'].value_counts
```

Out[]:

```
<bound method IndexOpsMixin.value_counts of 0      female  
1      male  
2      female  
3      male  
4      female  
      ...  
1304    female  
1305    female  
1306    male  
1307    male  
1308    male  
Name: sex, Length: 1309, dtype: object>
```

In []:

```
titanic['age'].value_counts
```

Out[]:

```
<bound method IndexOpsMixin.value_counts of 0      29.00  
1      0.92  
2      2.00  
3      30.00  
4      25.00  
      ...  
1304    14.50  
1305    NaN  
1306    26.50  
1307    27.00  
1308    29.00  
Name: age, Length: 1309, dtype: float64>
```

In []:

```
titanic.replace({'sex':{'male':0,'female':1}},inplace=True)
```

In []:

titanic.head()

Out[]:

	pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked
0	1	1	Allen, Miss. Elisabeth Walton	female	29.00	0	0	24160	211.3375	B5	S
1	1	1	Allison, Master. Hudson Trevor	male	0.92	1	2	113781	151.5500	C22 C26	S
2	1	0	Allison, Miss. Helen Loraine	female	2.00	1	2	113781	151.5500	C22 C26	S
3	1	0	Allison, Mr. Hudson Joshua Creighton	male	30.00	1	2	113781	151.5500	C22 C26	S
4	1	0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	25.00	1	2	113781	151.5500	C22 C26	S

Selection

In []:

titanic.loc[2:5, ['age', 'fare', 'name']]

Out[]:

	age	fare	name
2	2.0	151.55	Allison, Miss. Helen Loraine
3	30.0	151.55	Allison, Mr. Hudson Joshua Creighton
4	25.0	151.55	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)
5	48.0	26.55	Anderson, Mr. Harry

In []:

```
titanic.iloc[2:6,[4,8,2]]
```

Out[]:

	age	fare	name
2	2.0	151.55	Allison, Miss. Helen Loraine
3	30.0	151.55	Allison, Mr. Hudson Joshua Creighton
4	25.0	151.55	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)
5	48.0	26.55	Anderson, Mr. Harry

In []:

```
titanic[(titanic['age']>75) & (titanic['sex']==0)]
```

Out[]:

pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	bc
14	1	Barkworth, Mr. Algernon Henry Wilson	1	80.0	0	0	27042	30.0	A23	S	