In [1]:
```python
import pandas as pd
import numpy as np
```

In [2]:
```python
df=pd.DataFrame({
    'name':['anu','mahe','suresh'],
    'age':[22,19,45],
    'sex':['female','male','male']
})
df
```

Out[2]:

	name	age	sex
0	anu	22	female
1	mahe	19	male
2	suresh	45	male

In [3]:
```python
df['age']
```

Out[3]:
```
0    22
1    19
2    45
Name: age, dtype: int64
```

In [5]:
```python
ages=pd.Series([22,35,58],name='age')
ages
```

Out[5]:
```
0    22
1    35
2    58
Name: age, dtype: int64
```

In [6]:
```python
df['age'].max()
```

Out[6]: 45

In [7]:
```python
ages.max()
```

Out[7]: 58

In [8]:
```python
df.describe()
```

Out[8]:

	age
count	3.000000
mean	28.666667
std	14.224392
min	19.000000
25%	20.500000
50%	22.000000
75%	33.500000
max	45.000000

In [10]:
```python
titanic=pd.read_csv('C:/Users/raviy/Downloads/titanic.csv')
titanic.head()
```

Out[10]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	

In [11]:
```python
titanic.dtypes
```

Out[11]:
```
PassengerId      int64
Survived         int64
Pclass           int64
Name            object
Sex             object
Age            float64
SibSp            int64
Parch            int64
Ticket          object
Fare           float64
Cabin           object
Embarked        object
dtype: object
```

In [12]:
```python
titanic.to_excel('titanic.xlsx',sheet_name='titanic',index=False)
```

```
In [14]:  titanic.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  891 non-null    int64
 1   Survived     891 non-null    int64
 2   Pclass       891 non-null    int64
 3   Name         891 non-null    object
 4   Sex          891 non-null    object
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64
 7   Parch        891 non-null    int64
 8   Ticket       891 non-null    object
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object
 11  Embarked     889 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

```
In [15]:  ages=titanic['Age']
          ages.head()
```

```
Out[15]:  0    22.0
          1    38.0
          2    26.0
          3    35.0
          4    35.0
          Name: Age, dtype: float64
```

```
In [16]:  type(titanic['Age'])
```

```
Out[16]:  pandas.core.series.Series
```

```
In [17]:  titanic['Age'].shape
```

```
Out[17]:  (891,)
```

```
In [18]:  age_sex=titanic[['Age','Sex']]
          age_sex.head()
```

Out[18]:

	Age	Sex
0	22.0	male
1	38.0	female
2	26.0	female
3	35.0	female
4	35.0	male

```
In [19]:  titanic[['Age','Sex']].shape
```

```
Out[19]:  (891, 2)
```

In [20]:
```python
above35=titanic[titanic['Age']>35]
above35.head()
```

Out[20]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	C
13	14	0	3	Andersson, Mr. Anders Johan	male	39.0	1	5	347082	31.2750	N
15	16	1	2	Hewlett, Mrs. (Mary D Kingcome)	female	55.0	0	0	248706	16.0000	N

In [21]:
```python
class23=titanic[titanic['Pclass'].isin([2,3])]
class23.head()
```

Out[21]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583	
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	

In [22]:
```python
class23=titanic[(titanic['Pclass']==2)| (titanic['Pclass']==3)]
class23.head()
```

Out[22]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583	
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	

In [23]:
```python
age_no_na=titanic[titanic['Age'].notna()]
age_no_na.head()
```

Out[23]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	I
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	I
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	I

In [24]:
```python
adult_names=titanic.loc[titanic['Age']>35]
adult_names.head()
```

Out[24]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	C
13	14	0	3	Andersson, Mr. Anders Johan	male	39.0	1	5	347082	31.2750	N
15	16	1	2	Hewlett, Mrs. (Mary D Kingcome)	female	55.0	0	0	248706	16.0000	N

In [25]:
```python
adult_names=titanic.loc[titanic['Age']>35,'Name']
adult_names.head()
```

Out[25]:
```
1      Cumings, Mrs. John Bradley (Florence Briggs Th...
6                           McCarthy, Mr. Timothy J
11                        Bonnell, Miss. Elizabeth
13                    Andersson, Mr. Anders Johan
15                  Hewlett, Mrs. (Mary D Kingcome)
Name: Name, dtype: object
```

In [26]: `titanic.iloc[9:25,2:5]`

Out[26]:

	Pclass	Name	Sex
9	2	Nasser, Mrs. Nicholas (Adele Achem)	female
10	3	Sandstrom, Miss. Marguerite Rut	female
11	1	Bonnell, Miss. Elizabeth	female
12	3	Saundercock, Mr. William Henry	male
13	3	Andersson, Mr. Anders Johan	male
14	3	Vestrom, Miss. Hulda Amanda Adolfina	female
15	2	Hewlett, Mrs. (Mary D Kingcome)	female
16	3	Rice, Master. Eugene	male
17	2	Williams, Mr. Charles Eugene	male
18	3	Vander Planke, Mrs. Julius (Emelia Maria Vande...	female
19	3	Masselmani, Mrs. Fatima	female
20	2	Fynney, Mr. Joseph J	male
21	2	Beesley, Mr. Lawrence	male
22	3	McGowan, Miss. Anna "Annie"	female
23	1	Sloper, Mr. William Thompson	male
24	3	Palsson, Miss. Torborg Danira	female

In [28]:
```
anon=titanic.iloc[0:3,3]='anonymous'
anon
```

Out[28]: `'anonymous'`

In [29]: `titanic.head()`

Out[29]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	
0	1	0	3	anonymous	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	anonymous	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	anonymous	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	

In [30]: `titanic['Age'].mean()`

Out[30]: 29.69911764705882

In [31]: `titanic[['Age','Fare']].median()`

Out[31]:
```
Age      28.0000
Fare     14.4542
dtype: float64
```

In [32]: `titanic[['Age','Fare']].describe()`

Out[32]:

	Age	Fare
count	714.000000	891.000000
mean	29.699118	32.204208
std	14.526497	49.693429
min	0.420000	0.000000
25%	20.125000	7.910400
50%	28.000000	14.454200
75%	38.000000	31.000000
max	80.000000	512.329200

In [33]:
```
titanic.agg({
    'Age':['min','max','median','skew'],
    'Fare':['min','max','median','mean']
})
```

Out[33]:

	Age	Fare
min	0.420000	0.000000
max	80.000000	512.329200
median	28.000000	14.454200
skew	0.389108	NaN
mean	NaN	32.204208

In [34]: `titanic[['Sex','Age']].groupby('Sex').mean()`

Out[34]:

	Age
Sex	
female	27.915709
male	30.726645

In [35]: `titanic[['Sex','Age']].groupby('Sex').max()`

Out[35]:

	Age
Sex	
female	63.0
male	80.0

In [36]: `titanic[['Sex','Age']].groupby('Sex').first()`

Out[36]:

	Age
Sex	
female	38.0
male	22.0

In [37]: `titanic.head(2)`

Out[37]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabi
0	1	0	3	anonymous	male	22.0	1	0	A/5 21171	7.2500	Nal
1	2	1	1	anonymous	female	38.0	1	0	PC 17599	71.2833	C8

◀ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ▶

In [38]: `titanic.groupby('Sex')['Age'].mean()`

Out[38]:
```
Sex
female    27.915709
male      30.726645
Name: Age, dtype: float64
```

In [39]: `titanic.groupby(['Sex','Pclass'])['Fare'].mean()`

Out[39]:
```
Sex     Pclass
female  1         106.125798
        2          21.970121
        3          16.118810
male    1          67.226127
        2          19.741782
        3          12.661633
Name: Fare, dtype: float64
```

In [40]: `titanic['Pclass'].value_counts()`

Out[40]:
```
Pclass
3    491
1    216
2    184
Name: count, dtype: int64
```

In [41]: `titanic.groupby('Pclass')['Pclass'].count()`

Out[41]:
```
Pclass
1    216
2    184
3    491
Name: Pclass, dtype: int64
```

In [42]: `titanic.sort_values(by='Age',ascending=False).head()`

Out[42]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	C
630	631	1	1	Barkworth, Mr. Algernon Henry Wilson	male	80.0	0	0	27042	30.0000	
851	852	0	3	Svensson, Mr. Johan	male	74.0	0	0	347060	7.7750	
493	494	0	1	Artagaveytia, Mr. Ramon	male	71.0	0	0	PC 17609	49.5042	
96	97	0	1	Goldschmidt, Mr. George B	male	71.0	0	0	PC 17754	34.6542	
116	117	0	3	Connors, Mr. Patrick	male	70.5	0	0	370369	7.7500	

In [43]: `titanic.sort_values(by=['Pclass','Age'],ascending=False).head()`

Out[43]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cab
851	852	0	3	Svensson, Mr. Johan	male	74.0	0	0	347060	7.7750	Na
116	117	0	3	Connors, Mr. Patrick	male	70.5	0	0	370369	7.7500	Na
280	281	0	3	Duane, Mr. Frank	male	65.0	0	0	336439	7.7500	Na
483	484	1	3	Turkula, Mrs. (Hedwig)	female	63.0	0	0	4134	9.5875	Na
326	327	0	3	Nysveen, Mr. Johan Hansen	male	61.0	0	0	345364	6.2375	Na

In [45]: `titanic.dtypes`

Out[45]:
```
PassengerId      int64
Survived         int64
Pclass           int64
Name            object
Sex             object
Age            float64
SibSp            int64
Parch            int64
Ticket          object
Fare           float64
Cabin           object
Embarked        object
dtype: object
```

In [46]: `titanic['Name'].str.lower()`

Out[46]:
```
0                                anonymous
1                                anonymous
2                                anonymous
3       futrelle, mrs. jacques heath (lily may peel)
4                      allen, mr. william henry
                         ...
886                    montvila, rev. juozas
887                  graham, miss. margaret edith
888        johnston, miss. catherine helen "carrie"
889                      behr, mr. karl howell
890                        dooley, mr. patrick
Name: Name, Length: 891, dtype: object
```

In [47]: `titanic['Name'].str.split(',')`

Out[47]:
```
0                                [anonymous]
1                                [anonymous]
2                                [anonymous]
3       [Futrelle,  Mrs. Jacques Heath (Lily May Peel)]
4                      [Allen,  Mr. William Henry]
                         ...
886                    [Montvila,  Rev. Juozas]
887                  [Graham,  Miss. Margaret Edith]
888        [Johnston,  Miss. Catherine Helen "Carrie"]
889                      [Behr,  Mr. Karl Howell]
890                        [Dooley,  Mr. Patrick]
Name: Name, Length: 891, dtype: object
```

In [49]: `titanic['Surname']=titanic['Name'].str.split(',').str.get(0)`
`titanic['Surname']`

Out[49]:
```
0         anonymous
1         anonymous
2         anonymous
3          Futrelle
4             Allen
            ...
886        Montvila
887          Graham
888        Johnston
889            Behr
890          Dooley
Name: Surname, Length: 891, dtype: object
```

In [50]:
```python
titanic['Name_main']=titanic['Name'].str.split(',').str.get(1)
titanic['Name_main']
```

Out[50]:
```
0                                         NaN
1                                         NaN
2                                         NaN
3            Mrs. Jacques Heath (Lily May Peel)
4                           Mr. William Henry
                        ...
886                               Rev. Juozas
887                       Miss. Margaret Edith
888              Miss. Catherine Helen "Carrie"
889                           Mr. Karl Howell
890                               Mr. Patrick
Name: Name_main, Length: 891, dtype: object
```

In [51]:
```python
titanic['Name'].str.split(',')
```

Out[51]:
```
0                                         [anonymous]
1                                         [anonymous]
2                                         [anonymous]
3          [Futrelle,  Mrs. Jacques Heath (Lily May Peel)]
4                           [Allen,  Mr. William Henry]
                        ...
886                              [Montvila,  Rev. Juozas]
887                       [Graham,  Miss. Margaret Edith]
888            [Johnston,  Miss. Catherine Helen "Carrie"]
889                             [Behr,  Mr. Karl Howell]
890                               [Dooley,  Mr. Patrick]
Name: Name, Length: 891, dtype: object
```

In [52]:
```python
titanic['Real_Name']=titanic['Name'].str.split(',').str.get(0)
titanic.head()
```

Out[52]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	C
0	1	0	3	anonymous	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	anonymous	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	anonymous	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	

In [53]:
```python
titanic['Surname']=titanic['Name'].str.split(',').str.get(1)
titanic.head()
```

Out[53]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	C
0	1	0	3	anonymous	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	anonymous	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	anonymous	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	

In [54]:
```python
titanic['Salutation']=titanic['Surname'].str.split(',').str.get(0)
titanic.head()
```

Out[54]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	C
0	1	0	3	anonymous	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	anonymous	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	anonymous	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	

In [55]:
```python
titanic['Name'].str.contains('Mr')
```

Out[55]:
```
0       False
1       False
2       False
3        True
4        True
        ...
886     False
887     False
888     False
889      True
890      True
Name: Name, Length: 891, dtype: bool
```

In [56]: `titanic[titanic['Name'].str.contains('Countess')]`

Out[56]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
759	760	1	1	Rothes, the Countess. of (Lucy Noel Martha Dye...	female	33.0	0	0	110152	86.5	B77

In [57]: `titanic['Name'].str.len()`

Out[57]:
```
0        9
1        9
2        9
3       44
4       24
        ..
886     21
887     28
888     40
889     21
890     19
Name: Name, Length: 891, dtype: int64
```

In [59]: `titanic['Name'].str.len().idxmax()`

Out[59]: 307

In [ ]: