In [ ]:

```
1
```

In [ ]:

```
1
```

In [2]:

```
1  import numpy as np
2  a = np.arange(1,11).reshape(5,2)
3  a
```

Out[2]:

```
array([[ 1,  2],
       [ 3,  4],
       [ 5,  6],
       [ 7,  8],
       [ 9, 10]])
```

In [4]:

```
1  import numpy as np
2  a = np.arange(1,17).reshape(4,4)
3  a
```

Out[4]:

```
array([[ 1,  2,  3,  4],
       [ 5,  6,  7,  8],
       [ 9, 10, 11, 12],
       [13, 14, 15, 16]])
```

In [5]:

```
1  np.expand_dims(a,axis=0)
```

Out[5]:

```
array([[[ 1,  2,  3,  4],
        [ 5,  6,  7,  8],
        [ 9, 10, 11, 12],
        [13, 14, 15, 16]]])
```

In [6]:

```
1  np.expand_dims(a,axis=1)
```

Out[6]:

```
array([[[ 1,  2,  3,  4]],

       [[ 5,  6,  7,  8]],

       [[ 9, 10, 11, 12]],

       [[13, 14, 15, 16]]])
```

In [8]:

```
1  np.expand_dims(a,axis=1)
```

Out[8]:

```
array([[[ 1,  2,  3,  4]],

       [[ 5,  6,  7,  8]],

       [[ 9, 10, 11, 12]],

       [[13, 14, 15, 16]]])
```

In [9]:

```
1  np.expand_dims(a,axis=2)
```

Out[9]:

```
array([[[ 1],
        [ 2],
        [ 3],
        [ 4]],

       [[ 5],
        [ 6],
        [ 7],
        [ 8]],

       [[ 9],
        [10],
        [11],
        [12]],

       [[13],
        [14],
        [15],
        [16]]])
```

In [11]:

```python
np.expand_dims(a,axis=2)
```

Out[11]:

```
array([[[ 1],
        [ 2],
        [ 3],
        [ 4]],

       [[ 5],
        [ 6],
        [ 7],
        [ 8]],

       [[ 9],
        [10],
        [11],
        [12]],

       [[13],
        [14],
        [15],
        [16]]])
```

In [12]:

```python
b = np.expand_dims(a,axis=0)
b
```

Out[12]:

```
array([[[ 1,  2,  3,  4],
        [ 5,  6,  7,  8],
        [ 9, 10, 11, 12],
        [13, 14, 15, 16]]])
```

pandas

In [13]:

```python
import pandas as pd
import numpy as np

```

In [15]:

```
1  pd.Series([1,2,3,4,5])
```

Out[15]:

```
0    1
1    2
2    3
3    4
4    5
dtype: int64
```

In [17]:

```
1  s=pd.Series(["a","b","c","d"],index = ["I","II","III","IV"])
```

In [32]:

```
1  s
```

Out[32]:

```
['I', 'II']
```

In [19]:

```
1  s[0]
```

Out[19]:

```
'a'
```

In [20]:

```
1  s["I"]
```

Out[20]:

```
'a'
```

In [33]:

```
1  s[0:2]
```

Out[33]:

```
['I', 'II']
```

In [22]:

```python
1 s["I":"III"]
```

Out[22]:

```
I      a
II     b
III    c
dtype: object
```

In [25]:

```python
1 s=["I","II"]
```

In [26]:

```python
1 s
```

Out[26]:

```
['I', 'II']
```

In [36]:

```python
1 s=["I":"III"]
```

```
  File "<ipython-input-36-adcabcd0e616>", line 1
    s=["I":"III"]
          ^
SyntaxError: invalid syntax
```

In [35]:

```python
1 s[0:2]
```

Out[35]:

```
['I', 'II']
```

In [38]:

```python
1 marks = {"maths":87,"science":66,"english":82}
2 a=pd.Series(marks)
3 a
```

Out[38]:

```
maths      87
science    66
english    82
dtype: int64
```

In [39]:

```
1  a.index
```

Out[39]:

```
Index(['maths', 'science', 'english'], dtype='object')
```

In [40]:

```
1  pd.Series(np.arange(1,5))
```

Out[40]:

```
0    1
1    2
2    3
3    4
dtype: int32
```

Data Frame(table)

In [41]:

```
1  marks = {"naame":["avinash","sai","ram","mahi"],"maths":[87,98,75,68],"science":[78,67,
2  pd.DataFrame(marks)
```

Out[41]:

	naame	maths	science	english
0	avinash	87	78	89
1	sai	98	67	92
2	ram	75	87	76
3	mahi	68	90	55

In [45]:

```
1  a=np.array([["avinash",87,78,89],["sai",98,67,92]])
2  pd.DataFrame(a,columns=["name","maths","science","english"],index=["row1","row2"])
```

Out[45]:

	name	maths	science	english
row1	avinash	87	78	89
row2	sai	98	67	92

Reading a file

In [67]:

```python
df=pd.read_csv("marks.csv",header = None,names=["studentnames","maths","science","engli
```

In [68]:

```python
df
```

Out[68]:

	studentnames	maths	science	english
0	lalitha	45.0	67	89.0
1	venkatesh	89.0	76	65.0
2	phani	67.0	89	90.0
3	ranga	56.0	86	NaN
4	sathish	78.0	67	45.9
5	raja	56.7	98	NaN

In [54]:

```python
df["science"]
```

Out[54]:

```
0    67
1    76
2    89
3    86
4    67
5    98
Name: science, dtype: int64
```

In [55]:

```python
df.maths
```

Out[55]:

```
0    45.0
1    89.0
2    67.0
3    56.0
4    78.0
5    56.7
Name: maths, dtype: float64
```

In [69]:

```
1  df
```

Out[69]:

	studentnames	maths	science	english
0	lalitha	45.0	67	89.0
1	venkatesh	89.0	76	65.0
2	phani	67.0	89	90.0
3	ranga	56.0	86	NaN
4	sathish	78.0	67	45.9
5	raja	56.7	98	NaN

In [60]:

```
1  df.student name
```

```
  File "<ipython-input-60-97eb7dd23813>", line 1
    df.student name
               ^
SyntaxError: invalid syntax
```

In [72]:

```
1  df["studentnames"]
```

Out[72]:

```
0      lalitha
1    venkatesh
2        phani
3        ranga
4      sathish
5         raja
Name: studentnames, dtype: object
```

In [73]:

```
1  df[["studentnames","maths"]]
```

Out[73]:

	studentnames	maths
0	lalitha	45.0
1	venkatesh	89.0
2	phani	67.0
3	ranga	56.0
4	sathish	78.0
5	raja	56.7

In [74]:

```
1  df[["studentnames","science"]]
```

Out[74]:

	studentnames	science
0	lalitha	67
1	venkatesh	76
2	phani	89
3	ranga	86
4	sathish	67
5	raja	98

In [76]:

```
1  df[["studentnames","maths"]]
```

Out[76]:

	studentnames	maths
0	lalitha	45.0
1	venkatesh	89.0
2	phani	67.0
3	ranga	56.0
4	sathish	78.0
5	raja	56.7

In [75]:

```
1 df.values
```

Out[75]:

```
array([['lalitha', 45.0, 67, 89.0],
       ['venkatesh', 89.0, 76, 65.0],
       ['phani', 67.0, 89, 90.0],
       ['ranga', 56.0, 86, nan],
       ['sathish', 78.0, 67, 45.9],
       ['raja', 56.7, 98, nan]], dtype=object)
```

In [77]:

```
1 df.dtypes
```

Out[77]:

```
studentnames      object
maths            float64
science            int64
english          float64
dtype: object
```

In [78]:

```
1 df.shape
```

Out[78]:

```
(6, 4)
```

In [79]:

```
1 df.shape[1]
```

Out[79]:

```
4
```

indexing

*position based indexing(iloc)
*label based indexing(loc)

In [81]:

```
1  df["studentnames"]
```

Out[81]:

```
0       lalitha
1     venkatesh
2         phani
3         ranga
4       sathish
5          raja
Name: studentnames, dtype: object
```

In [82]:

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

Out[82]:

```
studentnames      sathish
maths                  78
Name: 4, dtype: object
```

In [83]:

```
1  df
```

Out[83]:

	studentnames	maths	science	english
0	lalitha	45.0	67	89.0
1	venkatesh	89.0	76	65.0
2	phani	67.0	89	90.0
3	ranga	56.0	86	NaN
4	sathish	78.0	67	45.9
5	raja	56.7	98	NaN

In [85]:

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

Out[85]:

```
studentnames      sathish
maths                  78
Name: 4, dtype: object
```

In [86]:

```
1  df.iloc[4,0:3]
```

Out[86]:

```
studentnames      sathish
maths                  78
science                67
Name: 4, dtype: object
```

In [88]:

```
1  df.iloc[0:5:2,:]
```

Out[88]:

	studentnames	maths	science	english
0	lalitha	45.0	67	89.0
2	phani	67.0	89	90.0
4	sathish	78.0	67	45.9

In [90]:

```
1  df.iloc[2,3]
```

Out[90]:

90.0

In [100]:

```
1  df.loc[1,"maths"]
```

Out[100]:

89.0

In [98]:

```
1  df.loc[2:5,"studentnames":"english"]
```

Out[98]:

	studentnames	maths	science	english
2	phani	67.0	89	90.0
3	ranga	56.0	86	NaN
4	sathish	78.0	67	45.9
5	raja	56.7	98	NaN

In [101]:

```
1  df.loc[2:5,["science","studentnames"]]
```

Out[101]:

	science	studentnames
2	89	phani
3	86	ranga
4	67	sathish
5	98	raja

In [102]:

```
1  d1 = df.set_index('studentnames')
2  d1
```

Out[102]:

	maths	science	english
studentnames			
lalitha	45.0	67	89.0
venkatesh	89.0	76	65.0
phani	67.0	89	90.0
ranga	56.0	86	NaN
sathish	78.0	67	45.9
raja	56.7	98	NaN

In [103]:

```
1  d1 = df.set_index('maths')
2  d1
```

Out[103]:

	studentnames	science	english
maths			
45.0	lalitha	67	89.0
89.0	venkatesh	76	65.0
67.0	phani	89	90.0
56.0	ranga	86	NaN
78.0	sathish	67	45.9
56.7	raja	98	NaN

In [104]:

```
1  df
```

Out[104]:

	studentnames	maths	science	english
0	lalitha	45.0	67	89.0
1	venkatesh	89.0	76	65.0
2	phani	67.0	89	90.0
3	ranga	56.0	86	NaN
4	sathish	78.0	67	45.9
5	raja	56.7	98	NaN

In [105]:

```
1  df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6 entries, 0 to 5
Data columns (total 4 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   studentnames  6 non-null      object
 1   maths         6 non-null      float64
 2   science       6 non-null      int64
 3   english       4 non-null      float64
dtypes: float64(2), int64(1), object(1)
memory usage: 320.0+ bytes
```

In [108]:

```
1  df["total"]=df["english"]+df["maths"]+df["science"]
```

In [110]:

```
1  df
```

Out[110]:

	studentnames	maths	science	english	total
0	lalitha	45.0	67	89.0	201.0
1	venkatesh	89.0	76	65.0	230.0
2	phani	67.0	89	90.0	246.0
3	ranga	56.0	86	NaN	NaN
4	sathish	78.0	67	45.9	190.9
5	raja	56.7	98	NaN	NaN

In [111]:

```
1  df.sum()
```

Out[111]:

```
studentnames    lalithavenkateshphanirangasathishraja
maths                                              391.7
science                                              483
english                                            289.9
total                                              867.9
dtype: object
```

In [112]:

```
1  df.sum(axis = 1)
```

Out[112]:

```
0    402.0
1    460.0
2    492.0
3    142.0
4    381.8
5    154.7
dtype: float64
```

In [114]:

```
1  df.describe()
```

Out[114]:

	maths	science	english	total
count	6.000000	6.000000	4.000000	4.000000
mean	65.283333	80.500000	72.475000	216.975000
std	16.114641	12.597619	21.152679	25.476836
min	45.000000	67.000000	45.900000	190.900000
25%	56.175000	69.250000	60.225000	198.475000
50%	61.850000	81.000000	77.000000	215.500000
75%	75.250000	88.250000	89.250000	234.000000
max	89.000000	98.000000	90.000000	246.000000

In [115]:

```python
1  df.describe(include = "all")
```

Out[115]:

	studentnames	maths	science	english	total
count	6	6.000000	6.000000	4.000000	4.000000
unique	6	NaN	NaN	NaN	NaN
top	venkatesh	NaN	NaN	NaN	NaN
freq	1	NaN	NaN	NaN	NaN
mean	NaN	65.283333	80.500000	72.475000	216.975000
std	NaN	16.114641	12.597619	21.152679	25.476836
min	NaN	45.000000	67.000000	45.900000	190.900000
25%	NaN	56.175000	69.250000	60.225000	198.475000
50%	NaN	61.850000	81.000000	77.000000	215.500000
75%	NaN	75.250000	88.250000	89.250000	234.000000
max	NaN	89.000000	98.000000	90.000000	246.000000

In [116]:

```python
1  df.describe(include = "object")
```

Out[116]:

	studentnames
count	6
unique	6
top	venkatesh
freq	1

In [117]:

```
1  df.head()
```

Out[117]:

	studentnames	maths	science	english	total
0	lalitha	45.0	67	89.0	201.0
1	venkatesh	89.0	76	65.0	230.0
2	phani	67.0	89	90.0	246.0
3	ranga	56.0	86	NaN	NaN
4	sathish	78.0	67	45.9	190.9

In [119]:

```
1  df.tail()
```

Out[119]:

	studentnames	maths	science	english	total
1	venkatesh	89.0	76	65.0	230.0
2	phani	67.0	89	90.0	246.0
3	ranga	56.0	86	NaN	NaN
4	sathish	78.0	67	45.9	190.9
5	raja	56.7	98	NaN	NaN

In [120]:

```
1  df.fillna(0)
```

Out[120]:

	studentnames	maths	science	english	total
0	lalitha	45.0	67	89.0	201.0
1	venkatesh	89.0	76	65.0	230.0
2	phani	67.0	89	90.0	246.0
3	ranga	56.0	86	0.0	0.0
4	sathish	78.0	67	45.9	190.9
5	raja	56.7	98	0.0	0.0

Merging

In [123]:

```python
df1 = pd.read_csv("stu1.csv")
df2 = pd.read_csv("col1.csv")
print(df1)
print(df2)
```

```
      name  science  maths  english
0    meena       90     80     67.0
1      sai       67     45     76.0
2  lalitha       78     67     56.0
3     rani       78     45     65.5
4  sathish       69     55      NaN
5     raja       88     61      NaN
      name   college    course
0    meena    Aditya    Python
1  lalitha  Pragathi        DS
2  lakshmi   Avanthi        Ml
3     nani      IIIT    Python
4 poogitha    Aditya        ML
5  mounika       RVR        DS
6     rani      IIIT   BigData
```

In [126]:

```python
pd.merge(df1,df2,on = "name")
```

Out[126]:

	name	science	maths	english	college	course
0	meena	90	80	67.0	Aditya	Python
1	lalitha	78	67	56.0	Pragathi	DS
2	rani	78	45	65.5	IIIT	BigData

In [122]:

```python
pd.merge(df1,df2,on = "name",how="inner") #==>intersection
```

Out[122]:

	name	science	maths	english	college	course
0	meena	90	80	67.0	Aditya	Python
1	lalitha	78	67	56.0	Pragathi	DS
2	rani	78	45	65.5	IIIT	BigData

In [124]:

```
1 pd.merge(df1,df2,on = "name",how="outer")   #==>union
```

Out[124]:

	name	science	maths	english	college	course
0	meena	90.0	80.0	67.0	Aditya	Python
1	sai	67.0	45.0	76.0	NaN	NaN
2	lalitha	78.0	67.0	56.0	Pragathi	DS
3	rani	78.0	45.0	65.5	IIIT	BigData
4	sathish	69.0	55.0	NaN	NaN	NaN
5	raja	88.0	61.0	NaN	NaN	NaN
6	lakshmi	NaN	NaN	NaN	Avanthi	Ml
7	nani	NaN	NaN	NaN	IIIT	Python
8	poogitha	NaN	NaN	NaN	Aditya	ML
9	mounika	NaN	NaN	NaN	RVR	DS

In [125]:

```
1 pd.merge(df1,df2,on = "name",how="left")
```

Out[125]:

	name	science	maths	english	college	course
0	meena	90	80	67.0	Aditya	Python
1	sai	67	45	76.0	NaN	NaN
2	lalitha	78	67	56.0	Pragathi	DS
3	rani	78	45	65.5	IIIT	BigData
4	sathish	69	55	NaN	NaN	NaN
5	raja	88	61	NaN	NaN	NaN

In [127]:

```
1  pd.merge(df1,df2,on = "name",how="right")
```

Out[127]:

	name	science	maths	english	college	course
0	meena	90.0	80.0	67.0	Aditya	Python
1	lalitha	78.0	67.0	56.0	Pragathi	DS
2	rani	78.0	45.0	65.5	IIIT	BigData
3	lakshmi	NaN	NaN	NaN	Avanthi	Ml
4	nani	NaN	NaN	NaN	IIIT	Python
5	poogitha	NaN	NaN	NaN	Aditya	ML
6	mounika	NaN	NaN	NaN	RVR	DS

In [ ]:

```
1
```