```
# Pandas


# Data structure
# series
# dataframe


import pandas as pd


x = pd.Series([5,7,8,9,2])
print(x)
```

```
0    5
1    7
2    8
3    9
4    2
dtype: int64
```

```
x[2:4]
```

```
2    8
3    9
dtype: int64
```

```
#loc=> ley index kei kaam garchhaa
#iloc()=> implicit lock
#index deda yesle aafno user index pani store gareko hunchhaa iloc le



y=pd.Series([5,7,8,9],index=['a','b','c','d'])
print(y)
```

```
a    5
b    7
c    8
d    9
dtype: int64
```

```
y.iloc[1]
```

```
np.int64(7)
```

```
# .loc()
# .iloc()



x = pd.Series([5,7,8,9,2])
print(x)
```

```
0    5
1    7
2    8
3    9
4    2
dtype: int64
```

```
x.loc[1]
```

```
np.int64(7)
```

```
Start coding or generate with AI.


Start coding or generate with AI.


Start coding or generate with AI.
```

Dataframe

```
# tabular data handle garnaa ko lagi dataframe chahenchhaa.


#dictionary bataa dataframe banauna sakenchhaa


 #empty bracket deyo bhane dictionary banchhaa
dis = {
    'roll':4,  #dictionary=> key:values
    'name':'sita',
    'faculty':'DS'
}


# roll, name, faculty => column banchhaa
# 4,sita,DS=>values


# dicyionay ma yeutaa set ko multiple value hunchhaa vaney list banaunu parchhaa


dis = {
    'roll':[4,5,7,9,12],
    'name':['sita','c','d','e','f'],
    'faculty':['DS','CS','DS','DS','CS']
}


df = pd.DataFrame(dis)
print(df)
```

```
   roll  name faculty
0     4  sita      DS
1     5     c      CS
2     7     d      DS
3     9     e      DS
4    12     f      CS
```

```
type(df)
```

```
pandas.core.frame.DataFrame
```

```
df.shape #5 row 3 columns
```

```
(5, 3)
```

```
df.columns
```

```
Index(['roll', 'name', 'faculty'], dtype='object')
```

```
data = pd.read_csv('Salary Data.csv') #file lai load garnaa lako vayeraaa .read-csv
#salary data .csv vanne file save gareraa cha ra teslai access garnaa try gareko


print(data)
```

```
       Age  Gender Education Level                     Job Title  \
0    32.0    Male      Bachelor's              Software Engineer
1    28.0  Female        Master's                   Data Analyst
2    45.0    Male             PhD                 Senior Manager
3    36.0  Female      Bachelor's                Sales Associate
4    52.0    Male        Master's                       Director
..    ...     ...             ...                            ...
370  35.0  Female      Bachelor's       Senior Marketing Analyst
371  43.0    Male        Master's         Director of Operations
372  29.0  Female      Bachelor's         Junior Project Manager
373  34.0    Male      Bachelor's  Senior Operations Coordinator
374  44.0  Female             PhD        Senior Business Analyst

     Years of Experience    Salary
0                    5.0   90000.0
1                    3.0   65000.0
2                   15.0  150000.0
3                    7.0   60000.0
4                   20.0  200000.0
```

```
..                   ...       ...
370                  8.0   85000.0
371                 19.0  170000.0
372                  2.0   40000.0
373                  7.0   90000.0
374                 15.0  150000.0

[375 rows x 6 columns]
```

```
# ### attributes and dunctions of dataframe
# dot(.) ley garnaa garnee


#Attributes
1.shapes
2.columns
3.
4.


#functions
1> head(),
tail(),
sample()
describe()
info()
```

```
data.head()# top 5 rows denchhaa
```

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
0	32.0	Male	Bachelor's	Software Engineer	5.0	90000.0
1	28.0	Female	Master's	Data Analyst	3.0	65000.0
2	45.0	Male	PhD	Senior Manager	15.0	150000.0
3	36.0	Female	Bachelor's	Sales Associate	7.0	60000.0
4	52.0	Male	Master's	Director	20.0	200000.0

```
data.head(10)# top 10 rows denchhaa
```

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
0	32.0	Male	Bachelor's	Software Engineer	5.0	90000.0
1	28.0	Female	Master's	Data Analyst	3.0	65000.0
2	45.0	Male	PhD	Senior Manager	15.0	150000.0
3	36.0	Female	Bachelor's	Sales Associate	7.0	60000.0
4	52.0	Male	Master's	Director	20.0	200000.0
5	29.0	Male	Bachelor's	Marketing Analyst	2.0	55000.0
6	42.0	Female	Master's	Product Manager	12.0	120000.0
7	31.0	Male	Bachelor's	Sales Manager	4.0	80000.0
8	26.0	Female	Bachelor's	Marketing Coordinator	1.0	45000.0
9	38.0	Male	PhD	Senior Scientist	10.0	110000.0

```
data.tail() #last ko 5 data denchha
```

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
370	35.0	Female	Bachelor's	Senior Marketing Analyst	8.0	85000.0
371	43.0	Male	Master's	Director of Operations	19.0	170000.0
372	29.0	Female	Bachelor's	Junior Project Manager	2.0	40000.0
373	34.0	Male	Bachelor's	Senior Operations Coordinator	7.0	90000.0
374	44.0	Female	PhD	Senior Business Analyst	15.0	150000.0

```
data.sample()#randomly rows lenchhaa
```

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
190	39.0	Female	Bachelor's	Senior Account Executive	12.0	95000.0

```
data.sample(10)#randomly rows lenchhaa
```

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
5	29.0	Male	Bachelor's	Marketing Analyst	2.0	55000.0
287	35.0	Female	Bachelor's	Senior Marketing Analyst	8.0	85000.0
261	37.0	Female	Bachelor's	Senior Financial Manager	10.0	120000.0
284	35.0	Male	Bachelor's	Senior Financial Manager	9.0	100000.0
368	44.0	Female	PhD	Senior Data Engineer	16.0	160000.0
232	27.0	Female	Master's	Junior Research Scientist	1.5	50000.0
107	36.0	Male	Bachelor's	IT Support Specialist	7.0	60000.0
206	31.0	Male	Bachelor's	Junior HR Generalist	4.0	50000.0
117	48.0	Male	PhD	Principal Engineer	20.0	170000.0
32	29.0	Male	Master's	Data Scientist	3.0	75000.0

```
# data ma kaam garnaa aaghi teslai explore garenchhaa ie;
# Exploratory Data Analysis(EDA)


data.shape #no of rows and columns denchaa
```

```
(375, 6)
```

```
data.shape[0]#index 0 deye rows matraa denchaa
```

```
375
```

```
data.shape[1]# index 1 deye columns matraa denchha
```

```
6
```

```
data.columns #Attribute ho so () use hudenaa
```

```
Index(['Age', 'Gender', 'Education Level', 'Job Title', 'Years of Experience',
       'Salary'],
      dtype='object')
```

```
# select particular rows or column


#columns
#syntax: dataframe.column_name

data.Age
```

```
0      32.0
1      28.0
2      45.0
3      36.0
4      52.0
       ...
370    35.0
371    43.0
372    29.0
373    34.0
374    44.0
Name: Age, Length: 375, dtype: float64
```

```
data.Gender
```

```
0        Male
1      Female
2        Male
3      Female
4        Male
        ...
370    Female
371      Male
372    Female
373      Male
374    Female
Name: Gender, Length: 375, dtype: object
```

```
data['Gender']#another way
```

```
0        Male
1      Female
2        Male
3      Female
4        Male
        ...
370    Female
371      Male
372    Female
373      Male
374    Female
Name: Gender, Length: 375, dtype: object
```

```
# #data.gender yesari access garda problem k cha vandaaa space ma vako kuraa lai garnaa ssakedenaa
# suppose attribute ko naam 2 ottaaa chaa EDucation level then dot gardaa error aauchaha
```

```
data.Education Level
```

```
  Cell In[23], line 1
    data.Education Level
                       ^
SyntaxError: invalid syntax
```

```
data['Education Level']
```

```
0      Bachelor's
1        Master's
2             PhD
3      Bachelor's
4        Master's
         ...
370    Bachelor's
371      Master's
372    Bachelor's
373    Bachelor's
374           PhD
Name: Education Level, Length: 375, dtype: object
```

```
data[0]#data ko 0 index ma vako row chaiyo bahne yesari garnaa medenaa loc,iloc chahenchhaa
```

```
data.loc[0]#0 index ma vako row chaiye
```

```
Age                                32.0
Gender                             Male
Education Level               Bachelor's
Job Title             Software Engineer
Years of Experience                5.0
Salary                         90000.0
Name: 0, dtype: object
```

```
data.loc[0:5] #0 ra 5 both aaucha ie; inclusive
```

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
0	32.0	Male	Bachelor's	Software Engineer	5.0	90000.0
1	28.0	Female	Master's	Data Analyst	3.0	65000.0
2	45.0	Male	PhD	Senior Manager	15.0	150000.0
3	36.0	Female	Bachelor's	Sales Associate	7.0	60000.0
4	52.0	Male	Master's	Director	20.0	200000.0
5	29.0	Male	Bachelor's	Marketing Analyst	2.0	55000.0

iloc=> exclusicve

```
data.iloc[0:5] # 0 to 4 ko matraa aayo so exclusive
```

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
0	32.0	Male	Bachelor's	Software Engineer	5.0	90000.0
1	28.0	Female	Master's	Data Analyst	3.0	65000.0
2	45.0	Male	PhD	Senior Manager	15.0	150000.0
3	36.0	Female	Bachelor's	Sales Associate	7.0	60000.0
4	52.0	Male	Master's	Director	20.0	200000.0

```
# 2 ottaa or badi  column access garna paryo bhaen as dataframe nai excess garnu parchhaa
data[['Age','Gender']]
```

	Age	Gender
0	32.0	Male
1	28.0	Female
2	45.0	Male
3	36.0	Female
4	52.0	Male
...	...	...
370	35.0	Female
371	43.0	Male
372	29.0	Female
373	34.0	Male
374	44.0	Female

375 rows × 2 columns

```
# 2 ottaa or badi  column access garna paryo bhaen as dataframe nai excess garnu parchhaa
data[['Age','Gender','Job Title']] #2 otaaa bracket ley nai hunchhaa jati otaa column chaiye ni
```

	Age	Gender	Job Title
0	32.0	Male	Software Engineer
1	28.0	Female	Data Analyst
2	45.0	Male	Senior Manager
3	36.0	Female	Sales Associate
4	52.0	Male	Director
...	...	...	...
370	35.0	Female	Senior Marketing Analyst
371	43.0	Male	Director of Operations
372	29.0	Female	Junior Project Manager
373	34.0	Male	Senior Operations Coordinator
374	44.0	Female	Senior Business Analyst

375 rows × 3 columns

```
data[['Age']]#yesto garnaa ni melchhaa yessko output is not series, it is dataframe
```

	Age
0	32.0
1	28.0
2	45.0
3	36.0
4	52.0
...	...
370	35.0
371	43.0
372	29.0
373	34.0
374	44.0

375 rows × 1 columns

```
data.head()
```

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
0	32.0	Male	Bachelor's	Software Engineer	5.0	90000.0
1	28.0	Female	Master's	Data Analyst	3.0	65000.0
2	45.0	Male	PhD	Senior Manager	15.0	150000.0
3	36.0	Female	Bachelor's	Sales Associate	7.0	60000.0
4	52.0	Male	Master's	Director	20.0	200000.0

```
#column ko type chhutaunaa
Numerical => continuos datatype => (int, float)
Categorical => object (text haru)


data.info() #data ko dtype pani dechhaa
#object bahnnee betekei categorical
#float64 bhannee betekei numerical
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 375 entries, 0 to 374
Data columns (total 6 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Age                  373 non-null    float64
 1   Gender               373 non-null    object
```

```
 2   Education Level      373 non-null    object
 3   Job Title            373 non-null    object
 4   Years of Experience  373 non-null    float64
 5   Salary               373 non-null    float64
dtypes: float64(3), object(3)
memory usage: 17.7+ KB
```

```
data.describe()#data ko overview(summary)denchhaa
```

	Age	Years of Experience	Salary
count	373.000000	373.000000	373.000000
mean	37.431635	10.030831	100577.345845
std	7.069073	6.557007	48240.013482
min	23.000000	0.000000	350.000000
25%	31.000000	4.000000	55000.000000
50%	36.000000	9.000000	95000.000000
75%	44.000000	15.000000	140000.000000
max	53.000000	25.000000	250000.000000

```
#  age 23-53 samma raichhaa  , mean(37),median(36)
```