Pandas

- Pandas is a builtin library used for data analysis.we will be using pandas heavily foir data manipulation,visualization,building machine learning models
- there are two main data structures in pandas : Series and Dataframe

In [1]:

```
# import pandas library

import pandas as pd
```

In [2]:

```
pd.__version__
```

Out[2]:

```
'1.4.4'
```

Pandas series

- A series is one dimensional array holding data of anytype (numeric,character,datetime)
- data can be represented as a sequential format like a column in a table
- syntax: pandas.Series(item/element,label=index)
- here an index is provided by user that changes the default index

In [3]:

```
# pandas series using list

l=[1,2,34,4,56,78]
p=pd.Series(l)
p
```

Out[3]:

```
0     1
1     2
2    34
3     4
4    56
5    78
dtype: int64
```

In [6]:

```
# creating series by using tuple

t=(11,22,44)
p1=pd.Series(t)
p1
```

Out[6]:

```
0    11
1    22
2    44
dtype: int64
```

In [8]:

```
p1=pd.Series(t,index=['a','b','c'])
p1
```

Out[8]:

```
a    11
b    22
c    44
dtype: int64
```

In [9]:

```
# creating series by using dictionary

d={'stu1':89,'stu2':90,'stu3':78}
s=pd.Series(d)
s
```

Out[9]:

```
stu1    89
stu2    90
stu3    78
dtype: int64
```

In [10]:

```
s.index=["stu4","stu2","stu3"]  # reindexing
s
```

Out[10]:

```
stu4    89
stu2    90
stu3    78
dtype: int64
```

In [11]:

```
s['stu3']  # accessing values using index
```

Out[11]:

```
78
```

In [12]:

```
s['stu4']
```

Out[12]:

```
89
```

In [13]:

```
p1
```

Out[13]:

```
a    11
b    22
c    44
dtype: int64
```

In [14]:

```
p1['b']
```

Out[14]:

```
22
```

In [16]:

```
p1[1]
```

Out[16]:

```
22
```

In [17]:

```
p1[1:]  # start,stop
```

Out[17]:

```
b    22
c    44
dtype: int64
```

In [18]:

```
p1[::2]   # start,stop,step
```

Out[18]:

```
a    11
c    44
dtype: int64
```

Convertion of existing data sets into series

In [19]:

```
# using string

st=input()
pd.Series(st)
```

```
hello world
```

Out[19]:

```
0    hello world
dtype: object
```

In [21]:

```
s={7,8,9,7,60,9}
pd.Series(s)  # cannot be converted since it is unordered
```

```
---------------------------------------------------------------------------
-
TypeError                                 Traceback (most recent call las
t)
~\AppData\Local\Temp\ipykernel_13464\4045882660.py in <module>
      1 s={7,8,9,7,60,9}
----> 2 pd.Series(s)

~\anaconda3\lib\site-packages\pandas\core\series.py in __init__(self, dat
a, index, dtype, name, copy, fastpath)
    449                     data = data.copy()
    450             else:
--> 451                 data = sanitize_array(data, index, dtype, copy)
    452
    453                 manager = get_option("mode.data_manager")

~\anaconda3\lib\site-packages\pandas\core\construction.py in sanitize_arra
y(data, index, dtype, copy, raise_cast_failure, allow_2d)
    582         if isinstance(data, (set, frozenset)):
    583             # Raise only for unordered sets, e.g., not for dict_ke
ys
--> 584             raise TypeError(f"'{type(data).__name__}' type is unor
dered")
    585
    586         # materialize e.g. generators, convert e.g. tuples, abc.Va
lueView

TypeError: 'set' type is unordered
```

In [25]:

```
import numpy as np
na=np.array(range(10,20,2))
p=pd.Series(na)
p
```

Out[25]:

```
0    10
1    12
2    14
3    16
4    18
dtype: int32
```

In [26]:

```
p.index=[0,1,2,3,'python']
```

In [27]:

```
p
```

Out[27]:

```
0         10
1         12
2         14
3         16
python    18
dtype: int32
```

In [30]:

```
a1=np.random.randint(100,500,5)
p=pd.Series(a1,index=[1,2,3,4,5])
p
```

Out[30]:

```
1    291
2    363
3    145
4    102
5    211
dtype: int32
```

In [ ]:

```

```