Very Good Morning to All

Numpy

Pandas

Matplot Lib

Seaborn

Scipy

Scikit-Learn

Pandas

It is the one of package.It is used for Data Manipulation , Vis,Data Cleaning

Data is stored in Two Ways or data Structure

- Series
- DataFrame

Series:

It is similar one-dim array of Numpy,It is having only one column

In [1]:

```
1 import numpy as np
```

In [2]:

```
1 np.array([100,200,300,400])
```

Out[2]:

```
array([100, 200, 300, 400])
```

In [3]:

```
1 import pandas
```

In [5]:

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

Out[5]:

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

In [6]:

```
import pandas as pd
```

In [7]:

```
pd.Series(["one","two","three"])
```

Out[7]:

```
0      one
1      two
2    three
dtype: object
```

In [8]:

```
p = pd.Series(["Ranganayakulu","Ayyappa","Anil","Ravi","Srinu"])
print(p)
```

```
0    Ranganayakulu
1          Ayyappa
2             Anil
3             Ravi
4            Srinu
dtype: object
```

In [9]:

```
p[0]
```

Out[9]:

```
'Ranganayakulu'
```

In [10]:

```
p[1:4]
```

Out[10]:

```
1    Ayyappa
2       Anil
3       Ravi
dtype: object
```

In [11]:

```
m = pd.Series(["Ranga","Ayyappa","Anil","Ravi","Srinu"],
              index=["Day1","Day2","Day3","Day4","Day5"])
print(m)
```

```
Day1      Ranga
Day2    Ayyappa
Day3       Anil
Day4       Ravi
Day5      Srinu
dtype: object
```

In [12]:

```
m[0]
```

Out[12]:

```
'Ranga'
```

In [13]:

```
m['Day1']
```

Out[13]:

```
'Ranga'
```

In [14]:

```
m[0][0]
```

Out[14]:

```
'R'
```

In [15]:

```
m.index[1]
```

Out[15]:

```
'Day2'
```

In [16]:

```
m
```

Out[16]:

```
Day1      Ranga
Day2    Ayyappa
Day3       Anil
Day4       Ravi
Day5      Srinu
dtype: object
```

In [17]:

```
m['Day1']
```

Out[17]:

```
'Ranga'
```

In [18]:

```
m['Day1'][0]
```

Out[18]:

```
'R'
```

In [20]:

```
m['Day1'][-1]
```

Out[20]:

```
'a'
```

In [21]:

```
h = pd.Series({"name":"lakshmi","sal":35000,"desig":"Dev"})
h
```

Out[21]:

```
name     lakshmi
sal        35000
desig        Dev
dtype: object
```

In [22]:

```
pd.date_range(start="2-10-2020",end="2-15-2020")
```

Out[22]:

```
DatetimeIndex(['2020-02-10', '2020-02-11', '2020-02-12', '2020-02-13',
               '2020-02-14', '2020-02-15'],
              dtype='datetime64[ns]', freq='D')
```

In [23]:

```
pd.date_range("1-10-2020","10-10-2020",freq="M")
```

Out[23]:

```
DatetimeIndex(['2020-01-31', '2020-02-29', '2020-03-31', '2020-04-30',
               '2020-05-31', '2020-06-30', '2020-07-31', '2020-08-31',
               '2020-09-30'],
              dtype='datetime64[ns]', freq='M')
```

In [24]:

```
pd.date_range("1-10-2020","1-11-2020",freq="30T")
```

...

In [25]:

```
pd.date_range("1-10-2020","1-15-2020",freq="120T")
```

...

In [38]:

```
pd.date_range("1-10-2020","1-11-2020",freq="0.1T")
```

Out[38]:

```
DatetimeIndex(['2020-01-10 00:00:00', '2020-01-10 00:00:06',
               '2020-01-10 00:00:12', '2020-01-10 00:00:18',
               '2020-01-10 00:00:24', '2020-01-10 00:00:30',
               '2020-01-10 00:00:36', '2020-01-10 00:00:42',
               '2020-01-10 00:00:48', '2020-01-10 00:00:54',
               ...
               '2020-01-10 23:59:06', '2020-01-10 23:59:12',
               '2020-01-10 23:59:18', '2020-01-10 23:59:24',
               '2020-01-10 23:59:30', '2020-01-10 23:59:36',
               '2020-01-10 23:59:42', '2020-01-10 23:59:48',
               '2020-01-10 23:59:54', '2020-01-11 00:00:00'],
              dtype='datetime64[ns]', length=14401, freq='6S')
```

In [39]:

```
g = pd.Series(pd.date_range("18-5-2020","29-5-2020"))
g
```

Out[39]:

```
0    2020-05-18
1    2020-05-19
2    2020-05-20
3    2020-05-21
4    2020-05-22
5    2020-05-23
6    2020-05-24
7    2020-05-25
8    2020-05-26
9    2020-05-27
10   2020-05-28
11   2020-05-29
dtype: datetime64[ns]
```

In [41]:

```
k = pd.Series(["Ranga","Srinu","Anil","Sastry","Ayyappa"],
              index=[1," ",3," "," "])
k
```

Out[41]:

```
1      Ranga
       Srinu
3       Anil
      Sastry
     Ayyappa
dtype: object
```

DataFrame

In [48]:

```
g=pd.Series({"names":["vijay","lakshmi","raja"],
             "Sal":[25000,35000,45000],
            "Desig":["Dev","Trainer","texter"]})
g
```

Out[48]:

```
names    [vijay, lakshmi, raja]
Sal       [25000, 35000, 45000]
Desig    [Dev, Trainer, texter]
dtype: object
```

In [49]:

```
g=pd.DataFrame({"names":["vijay","lakshmi","raja"],
             "Sal":[25000,35000,45000],
            "Desig":["Dev","Trainer","texter"]})
g
```

Out[49]:

	names	Sal	Desig
0	vijay	25000	Dev
1	lakshmi	35000	Trainer
2	raja	45000	texter

In [50]:

```
g.shape
```

Out[50]:

```
(3, 3)
```

In [51]:

```
g.values
```

Out[51]:

```
array([['vijay', 25000, 'Dev'],
       ['lakshmi', 35000, 'Trainer'],
       ['raja', 45000, 'texter']], dtype=object)
```

In [52]:

```
g.keys
```

Out[52]:

```
<bound method NDFrame.keys of      names    Sal    Desig
0    vijay  25000      Dev
1  lakshmi  35000  Trainer
2     raja  45000   texter>
```

In [53]:

```
1 g.columns
```

Out[53]:

```
Index(['names', 'Sal', 'Desig'], dtype='object')
```

In [54]:

```
1 g.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3 entries, 0 to 2
Data columns (total 3 columns):
names    3 non-null object
Sal      3 non-null int64
Desig    3 non-null object
dtypes: int64(1), object(2)
memory usage: 152.0+ bytes
```

In [55]:

```
1 g.describe()
```

Out[55]:

	Sal
count	3.0
mean	35000.0
std	10000.0
min	25000.0
25%	30000.0
50%	35000.0
75%	40000.0
max	45000.0

In [56]:

```
1 k=g[['names','Desig']]
2 k
```

Out[56]:

	names	Desig
0	vijay	Dev
1	lakshmi	Trainer
2	raja	texter

In [58]:

```
1 k.describe()
```

Out[58]:

	names	Desig
count	3	3
unique	3	3
top	raja	Dev
freq	1	1

In [59]:

```
1 k
```

Out[59]:

	names	Desig
0	vijay	Dev
1	lakshmi	Trainer
2	raja	texter

In [ ]:

```
1 ## iloc
2 ## loc
```

In [61]:

```
1 k.iloc[0]
```

Out[61]:

```
names    vijay
Desig      Dev
Name: 0, dtype: object
```

In [65]:

```
1 k['names'][0]
```

Out[65]:

```
'vijay'
```

In [66]:

```
1 k
```

Out[66]:

	names	Desig
0	vijay	Dev
1	lakshmi	Trainer
2	raja	texter

In [67]:

```
1 k['names']
```

Out[67]:

```
0      vijay
1    lakshmi
2       raja
Name: names, dtype: object
```

In [68]:

```
1 k['Desig']
```

Out[68]:

```
0        Dev
1    Trainer
2     texter
Name: Desig, dtype: object
```

In [69]:

```
1 k['Desig'][1]
```

Out[69]:

```
'Trainer'
```

In [71]:

```
1 k[['names','Desig']]
```

Out[71]:

	names	Desig
0	vijay	Dev
1	lakshmi	Trainer
2	raja	texter

In [72]:

```
k.loc[1,'names']
```

Out[72]:

```
'lakshmi'
```

In [73]:

```
k.loc[:,'Desig']
```

Out[73]:

```
0        Dev
1    Trainer
2     texter
Name: Desig, dtype: object
```

In [74]:

```
k.iloc[0][1]
```

Out[74]:

```
'Dev'
```

In [75]:

```
k.loc[0][1]
```

Out[75]:

```
'Dev'
```

In [76]:

```
k.loc[0]
```

Out[76]:

```
names    vijay
Desig      Dev
Name: 0, dtype: object
```

In [77]:

```
#read_csv
k = pd.read_csv("students.csv")
k
```

Out[77]:

	Roll_Num	Name	TMarks	Percentage
0	501	sai	502	92
1	502	vijay	359	65
2	512	sankhar	321	63
3	524	giri	521	94
4	505	hari	565	96
5	504	ran	456	83

In [79]:

```
k.set_index("Roll_Num")
```

Out[79]:

	Name	TMarks	Percentage
Roll_Num			
501	sai	502	92
502	vijay	359	65
512	sankhar	321	63
524	giri	521	94
505	hari	565	96
504	ran	456	83

In [80]:

```
k
```

Out[80]:

	Roll_Num	Name	TMarks	Percentage
0	501	sai	502	92
1	502	vijay	359	65
2	512	sankhar	321	63
3	524	giri	521	94
4	505	hari	565	96
5	504	ran	456	83

In [81]:

```
k.set_index('Roll_Num',inplace=True)
k
```

Out[81]:

Roll_Num	Name	TMarks	Percentage
501	sai	502	92
502	vijay	359	65
512	sankhar	321	63
524	giri	521	94
505	hari	565	96
504	ran	456	83

In [82]:

```
k.sort_index()
```

Out[82]:

Roll_Num	Name	TMarks	Percentage
501	sai	502	92
502	vijay	359	65
504	ran	456	83
505	hari	565	96
512	sankhar	321	63
524	giri	521	94

In [83]:

```
1 k.sort_values(by="Percentage")
```

Out[83]:

Roll_Num	Name	TMarks	Percentage
512	sankhar	321	63
502	vijay	359	65
504	ran	456	83
501	sai	502	92
524	giri	521	94
505	hari	565	96

In [84]:

```
1 k.sort_values(by="Percentage",ascending=True)
```

Out[84]:

Roll_Num	Name	TMarks	Percentage
512	sankhar	321	63
502	vijay	359	65
504	ran	456	83
501	sai	502	92
524	giri	521	94
505	hari	565	96

In [85]:

```
1 k.sort_values(by="Percentage",ascending=False)
```

Out[85]:

Roll_Num	Name	TMarks	Percentage
505	hari	565	96
524	giri	521	94
501	sai	502	92
504	ran	456	83
502	vijay	359	65
512	sankhar	321	63

In [86]:

```
1 k['Name']
```

Out[86]:

```
Roll_Num
501        sai
502      vijay
512    sankhar
524       giri
505       hari
504        ran
Name: Name, dtype: object
```

In [87]:

```
1 k[['Name','Percentage']]
```

Out[87]:

Roll_Num	Name	Percentage
501	sai	92
502	vijay	65
512	sankhar	63
524	giri	94
505	hari	96
504	ran	83

In [88]:

```
1 k
```

Out[88]:

Roll_Num	Name	TMarks	Percentage
501	sai	502	92
502	vijay	359	65
512	sankhar	321	63
524	giri	521	94
505	hari	565	96
504	ran	456	83

In [89]:

```
1 k.iloc[2][2]
```

Out[89]:

63

In [91]:

```
1 k.loc[512,"Percentage"]
```

Out[91]:

63

In [92]:

```
1 k.loc[k['TMarks']>500]
```

Out[92]:

Roll_Num	Name	TMarks	Percentage
501	sai	502	92
524	giri	521	94
505	hari	565	96

In [93]:

```
1 k.loc[505,'TMarks']
```

Out[93]:

565

In [94]:

```
1 k.loc[505,'TMarks']=555
```

In [95]:

```
1 k
```

Out[95]:

Roll_Num	Name	TMarks	Percentage
501	sai	502	92
502	vijay	359	65
512	sankhar	321	63
524	giri	521	94
505	hari	555	96
504	ran	456	83

Tomorrow is the last session

Please provide your valuable feedback and feedback videos

manasa.p@apssdc send mail to this mail if you didn't receive the mail

In [ ]:

```
1
```