pandas2

September 16, 2020

```
In [74]: import numpy as np

In [75]: import pandas as pd

In [3]: #Example - 1

In [76]: # creating a empty dataframe
         import pandas as pd
         df1 = pd.DataFrame()
         print(df1)
```

```
Empty DataFrame
Columns: []
Index: []
```

```
In [6]: # Example  - 2

In [77]: import pandas as pd
         # creation of list
         data2 = [16,26,36,46,56,66,76,86,96]
         # create a dataframe from list
         df2 = pd.DataFrame(data2)
         print("Dataframe:")
         print(df2)
```

```
Dataframe:
    0
0  16
1  26
2  36
3  46
4  56
5  66
6  76
7  86
8  96
```

```
In [8]: # List of Lists
```

```
In [78]: # Example related creation of dataframe from list of list
         # creation of list of lists
         data3 = [['sayanthan',20],['Yathvi',4],['Harshika',10], ['Poojitha', 13], ['Goverdhini
         # creation of dataframe from the list of list
         df3 = pd.DataFrame(data3,columns=['Name','Age'])
         # printing the dataframe
         print(df3)
```

```
         Name  Age
0   sayanthan   20
1      Yathvi    4
2    Harshika   10
3    Poojitha   13
4  Goverdhini   15
```

```
In [79]: # small change in the Dataframe
         data4 = [['sayanthan',20000],['Yathvi',40000],['Harshika',10000], ['Poojitha', 13000]
         df4 = pd.DataFrame(data4,columns=['Name','salary'], dtype = float)
         print(df4)
```

```
         Name   salary
0   sayanthan  20000.0
1      Yathvi  40000.0
2    Harshika  10000.0
3    Poojitha  13000.0
4  Goverdhini  15000.0
```

```
In [11]: # Example - 3
```

```
In [80]: import pandas as pd
         data5 = {'Name':['harinisri', 'sushmitha', 'lakshmi', 'venkatesh'],'Age':[18,44,31,56]
         df5 = pd.DataFrame(data5)
         print(df5)
         # what is your observation
```

```
   Age       Name
0   18  harinisri
1   44  sushmitha
2   31    lakshmi
3   56  venkatesh
```

```
In [81]: import pandas as pd
         data6 = {'Name':['harinisri', 'sushmitha', 'lakshmi', 'venkatesh'],'Age':[18,44,31,56]
         df6 = pd.DataFrame(data6, index=['row1','row2','row3','row4'])
         print(df6)
         # what is your observation
```

```
      Age       Name
row1   18  harinisri
row2   44  sushmitha
row3   31    lakshmi
row4   56  venkatesh
```

```
In [14]: # Example - 4
```

```
In [82]: import pandas as pd
         data7 = [{'abc': 1, 'bcd': 2},{'abc': 5, 'bcd': 10, 'cdf': 20}]
         df7 = pd.DataFrame(data7)
         print(df7)
         # what is your observation
```

```
   abc  bcd   cdf
0    1    2   NaN
1    5   10  20.0
```

```
In [83]: import pandas as pd
         data8 = [{'abc': 1, 'bcd': 2},{'abc': 5, 'bcd': 10, 'cdf': 20}]
         df8 = pd.DataFrame(data8,index=['row1', 'row2'])
         print(df8)
         # what is your observation
```

```
      abc  bcd   cdf
row1    1    2   NaN
row2    5   10  20.0
```

```
In [84]: data9 = [{'abc': 1, 'bcd': 2},{'abc': 5, 'bcd': 10, 'cdf': 20}]

         #With two column indices, values same as dictionary keys
         df9 = pd.DataFrame(data9, index=['row1', 'row2'], columns=['abc', 'bcd'])
         print(df9)
         print("Another")
         #With two column indices with one index with other name
         df10 = pd.DataFrame(data9, columns=['abc', 'xyz'])
         print(df10)
```

```
      abc  bcd
row1    1    2
row2    5   10
Another
      abc  xyz
row1    1  NaN
row2    5  NaN
```

```
In [23]: # Example - 5

In [85]: import pandas as pd

         data10 = {'marks' : pd.Series([45, 48, 35], index=['abhi', 'bhargav', 'charan']),
                   'attedance' : pd.Series([79, 89, 100, 71], index=['abhi', 'bhargav', 'charan

         df10 = pd.DataFrame(data10)
         print(df10)
         # observation

         attedance  marks
abhi            79   45.0
bhargav         89   48.0
charan         100   35.0
dinesh          71    NaN


In [25]: #Example - 6

In [86]: import pandas as pd

         data10 = {'marks' : pd.Series([45, 48, 35], index=['abhi', 'bhargav', 'charan']),
                   'attendance' : pd.Series([79, 89, 100, 71], index=['abhi', 'bhargav', 'chara

         df10 = pd.DataFrame(data10)
         print("Attendance details of all students:")
         print(df10['attendance'])

Attendance details of all students:
abhi        79
bhargav     89
charan     100
dinesh      71
Name: attendance, dtype: int64


In [28]: #Example - 7

In [88]: import pandas as pd

         data11 = {'marks-1' : pd.Series([45, 48, 35], index=['abhi', 'bhargav', 'charan']),
                   'attendance' : pd.Series([79, 89, 100, 71], index=['abhi', 'bhargav', 'chara

         df11 = pd.DataFrame(data11)
         print(df11)
         # adding new column marks-2
         print ("Adding a new column by passing as Series:")
         df11['marks-2']=pd.Series([37,39,28],index=['abhi', 'bhargav', 'charan'])
         print(df11)
```

```
         attendance  marks-1
abhi             79     45.0
bhargav          89     48.0
charan          100     35.0
dinesh           71      NaN
Adding a new column by passing as Series:
         attendance  marks-1  marks-2
abhi             79     45.0     37.0
bhargav          89     48.0     39.0
charan          100     35.0     28.0
dinesh           71      NaN      NaN
```

```
In [89]: print ("Adding a new column using the existing columns in DataFrame:")
         df11['total']=df11['marks-1']+df11['marks-2']
         print(df11)
```

```
Adding a new column using the existing columns in DataFrame:
         attendance  marks-1  marks-2  total
abhi             79     45.0     37.0   82.0
bhargav          89     48.0     39.0   87.0
charan          100     35.0     28.0   63.0
dinesh           71      NaN      NaN    NaN
```

```
In [35]: # Example  -8
```

```
In [90]: import pandas as pd

         data12 = {'marks-1' : pd.Series([45, 48, 35], index=['abhi', 'bhargav', 'charan']),
                   'marks-2': pd.Series([37,39,28],index=['abhi', 'bhargav', 'charan']),
                   'marks-3' : pd.Series([41, 32, 39], index=['abhi', 'bhargav', 'charan']),
                  'attendance' : pd.Series([79, 89, 100, 71], index=['abhi', 'bhargav', 'chara

         df12 = pd.DataFrame(data12)
         print(df12)
         print ("finding total marks of each student:")
         df12['total']=df12['marks-1']+df12['marks-2']
         print(df12)

         # using del function
         print ("Deleting the total column using DEL function:")
         del(df12['total'])
         print(df12)

         # using pop function
         print ("Deleting attendance column using POP function:")
         df12.pop('attendance')
         print(df12)
```

```
         attendance  marks-1  marks-2  marks-3
abhi             79     45.0     37.0     41.0
bhargav          89     48.0     39.0     32.0
charan          100     35.0     28.0     39.0
dinesh           71      NaN      NaN      NaN
finding total marks of each student:
         attendance  marks-1  marks-2  marks-3  total
abhi             79     45.0     37.0     41.0   82.0
bhargav          89     48.0     39.0     32.0   87.0
charan          100     35.0     28.0     39.0   63.0
dinesh           71      NaN      NaN      NaN    NaN
Deleting the total column using DEL function:
         attendance  marks-1  marks-2  marks-3
abhi             79     45.0     37.0     41.0
bhargav          89     48.0     39.0     32.0
charan          100     35.0     28.0     39.0
dinesh           71      NaN      NaN      NaN
Deleting attendance column using POP function:
         marks-1  marks-2  marks-3
abhi        45.0     37.0     41.0
bhargav     48.0     39.0     32.0
charan      35.0     28.0     39.0
dinesh       NaN      NaN      NaN
```

In [40]:
```
# Example - 9
```

In [92]:
```
import pandas as pd

data13 = {'marks-1' : pd.Series([45, 48, 35], index=['abhi', 'bhargav', 'charan']),
          'marks-2': pd.Series([37,39,28],index=['abhi', 'bhargav', 'charan']),
          'marks-3' : pd.Series([41, 32, 39], index=['abhi', 'bhargav', 'charan']),
         'attendance' : pd.Series([79, 89, 100, 71], index=['abhi', 'bhargav', 'chara

df13 = pd.DataFrame(data13)
print(df13)
print("Retriving details of charan:")
print(df13.loc['charan'])

print("Retriving details of bhargav:")
print(df13.loc['bhargav'])

print("Retriving details of dinesh:")
print(df13.loc['dinesh'])
#The result is a series with labels as column names of the DataFrame.
# And, the Name of the series is the label with which it is retrieved.
```

```
         attendance  marks-1  marks-2  marks-3
abhi             79     45.0     37.0     41.0
```

```
bhargav          89     48.0     39.0     32.0
charan          100     35.0     28.0     39.0
dinesh           71      NaN      NaN      NaN
Retriving details of charan:
attendance    100.0
marks-1        35.0
marks-2        28.0
marks-3        39.0
Name: charan, dtype: float64
Retriving details of bhargav:
attendance    89.0
marks-1       48.0
marks-2       39.0
marks-3       32.0
Name: bhargav, dtype: float64
Retriving details of dinesh:
attendance    71.0
marks-1        NaN
marks-2        NaN
marks-3        NaN
Name: dinesh, dtype: float64
```

```
In [49]: # Example - 10
```

```
In [94]: import pandas as pd

         data14 = {'marks-1' : pd.Series([45, 48, 35], index=['abhi', 'bhargav', 'charan']),
                   'marks-2': pd.Series([37,39,28],index=['abhi', 'bhargav', 'charan']),
                   'marks-3' : pd.Series([41, 32, 39], index=['abhi', 'bhargav', 'charan']),
                  'attendance' : pd.Series([79, 89, 100, 71], index=['abhi', 'bhargav', 'chara

         df14 = pd.DataFrame(data14)
         print(df14)
         print("details of charan")
         print(df14.iloc[2])
```

```
         attendance  marks-1  marks-2  marks-3
abhi             79     45.0     37.0     41.0
bhargav          89     48.0     39.0     32.0
charan          100     35.0     28.0     39.0
dinesh           71      NaN      NaN      NaN
details of charan
attendance    100.0
marks-1        35.0
marks-2        28.0
marks-3        39.0
Name: charan, dtype: float64
```

```
In [55]: # Example - 11

In [95]: import pandas as pd

         data15 = {'marks-1' : pd.Series([45, 48, 35], index=['abhi', 'bhargav', 'charan']),
                   'marks-2': pd.Series([37,39,28],index=['abhi', 'bhargav', 'charan']),
                   'marks-3' : pd.Series([41, 32, 39], index=['abhi', 'bhargav', 'charan']),
                  'attendance' : pd.Series([79, 89, 100, 71], index=['abhi', 'bhargav', 'chara

         df15 = pd.DataFrame(data15)
         print("deatils from abhi and bhargav")
         print(df15[0:2])

deatils from abhi and bhargav
         attendance  marks-1  marks-2  marks-3
abhi             79     45.0     37.0     41.0
bhargav          89     48.0     39.0     32.0


In [58]: # Example - 12

In [96]: import pandas as pd

         df16 = pd.DataFrame([[1, 2], [3, 4]], columns = ['a','b'])
         print(df16)
         df17 = pd.DataFrame([[5, 6], [7, 8]], columns = ['a','b'])
         print(df17)

         df16 = df16.append(df17)
         print(df16)

   a  b
0  1  2
1  3  4
   a  b
0  5  6
1  7  8
   a  b
0  1  2
1  3  4
0  5  6
1  7  8


In [97]: import pandas as pd

         df16 = pd.DataFrame([[1, 2], [3, 4]], columns = ['a','b'])
         df17 = pd.DataFrame([[5, 6], [7, 8]], columns = ['a','b'])
```

```
df16 = df16.append(df17)
print(df16)
print("After deleting")
df16 = df16.drop(0)
print(df16)
```

```
   a  b
0  1  2
1  3  4
0  5  6
1  7  8
After deleting
   a  b
1  3  4
1  7  8
```