**PRACTICAL – 4**

**AIM**: Perform Reshaping of the hierarchical data and pivoting data frame data.

**THEORY**

In pandas, we can arrange data within the data frame from the existing data frame. For example, we have the same name with different features, instead of writing the name all time, we can write only once.

We can create hierarchical data from the existing data frame using pandas. Stacking a Data Frame means moving (also rotating or pivoting) the innermost column index to become the innermost row index. The inverse operation is called unstacking. It means moving the innermost row index to become the innermost column index.

The pivot function is used to create a new derived table out of a given one. Pivot takes 3 arguments with the following names: index, columns, and values. As a value for each of these parameters you need to specify a column name in the original table. Then the pivot function will create a new table, whose row and column indices are the unique values of the respective parameters

**SOURCE CODE AND OUTPUT**

**# Reshaping of the hierarchical data**

import numpy as np

import pandas as pd

df=pd.read\_csv("/content/sample\_data/nba.csv")

df.head()



**#Used to stack (or melt) the DataFrame**

df\_stacked=df.stack()

df\_stacked.head(26)

0 Name Avery Bradley

Team Boston Celtics

Number 0.0

Position PG

Age 25.0

Height 6-2

Weight 180.0

College Texas

Salary 7730337.0

1 Name Jae Crowder

Team Boston Celtics

Number 99.0

Position SF

Age 25.0

Height 6-6

Weight 235.0

College Marquette

Salary 6796117.0

2 Name John Holland

Team Boston Celtics

Number 30.0

Position SG

Age 27.0

Height 6-5

Weight 205.0

College Boston University

dtype: object

df\_unstacked=df\_stacked=df.unstack()

df\_unstacked.head(10)

Name 0 Avery Bradley

1 Jae Crowder

2 John Holland

3 R.J. Hunter

4 Jonas Jerebko

5 Amir Johnson

6 Jordan Mickey

7 Kelly Olynyk

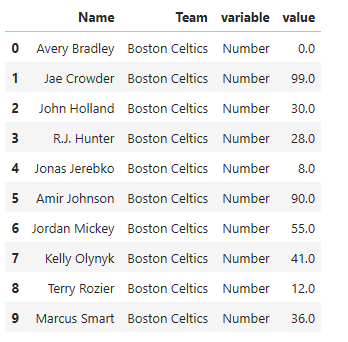
8 Terry Rozier

9 Marcus Smart

dtype: object

df\_melt = df.melt(id\_vars =['Name', 'Team'])

df\_melt.head(10)



**#Pivoting data frame data**

import pandas as pd

import numpy as np

import warnings

warnings.filterwarnings('ignore')

**#Create a dummy datasframe from dictionary**

dict={'Test':[1,2,3,1,2,3,1,2,3],

'Names':['Ahmed','Ahmed','Ahmed','Renu','Renu','Renu',

'Deepak','Deepak','Deepak'],

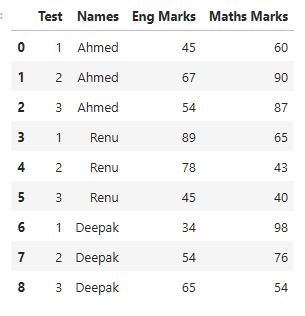
'Eng Marks':[45,67,54,89,78,45,34,54,65],

'Maths Marks':[60,90,87,65,43,40,98,76,54]}

**#Converts dictionary into dataframe**

df=pd.DataFrame(dict)

df

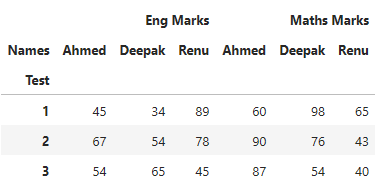


**#Here long dataset means no. of rows=9 and no. of columns=4**

**#9>4**

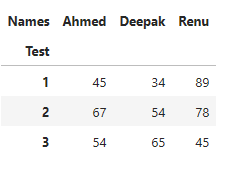
**#pivot() inbuilt method of Pandas , use for DataFrame**

df.pivot(index='Test',columns='Names')



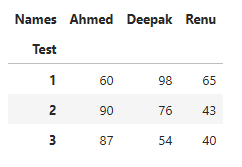
**#Only show marks of english of each students**

df.pivot(index='Test',columns='Names',values='Eng Marks')



**#Only show marks of maths of each students**

df.pivot(index='Test',columns='Names',values='Maths Marks')



**#Use pivot\_table()**

df1=df.pivot\_table(index='Names',columns='Test',values='Maths Marks',

aggfunc='mean')

#by default aggfunc='mean'

#aggragate function : sum() ,mean() ,max() ,min() ,count()

df1

