Ex.No-4 **Data Loading and Storing**

LOADING

Aim:

To read excel/csv/text files and extract the relevant information

Description:

1.Read and display the excel file data

2. Through DataFrame get the details of column headings

3. Through DataFrame get the details of the shape of Excel table

4. Through DataFrame get the particular column values

5. Through DataFrame extract/slice the Excel table values

6. Through DataFrame get the particular row values

7. Through DataFrame make an average of particular column values

Program:

```
import pandas as pd d=pd.
read_csv("LAS.csv") #Get
the table data print("Get
the table data:\n") print(d)
#print(d.to_string()
)
df=pd.DataFrame(
d) #print(df)
#Get the column heading
```

```
print("\nGet the column heading\n",df.columns)

#Get the shape (no.of raws.no, of columns)

print("\nGet the shape (no.of rows,no.of

columns)\n",df.shape) #Get particular column values

print("\nGet particular column values\n",df['roll.no'])
```

```
#Extract/slice the table values (including this row, excluding this row

print("\nExtract/slice the table values-[including this row, excluding this row]\n",df[2:

5]) #Get the particular row values through row number identification

print("\nGet the particular row values-through row number

identification\n",df.loc[7]) #Get the particular row values-through 'Roll number'

identification

print("\nGet the particular row values-through 'Roll number'

identification\n",d.loc[d['roll.no']==5]) #Make an average of total mark

df=d['total']/5

print("\n Make an average of total marks:\n",df)
```

Output:

Get the table data:

	roll.no	name	maths	science	social	total
0	1	deepa	50	67	50	284
1	2	dinesh	56	89	56	346
2	3	kaviya	80	80	80	400
3	4	racheal	89	87	89	441
4	5	rajan	90	98	90	466
5	6	ramya	67	76	67	353
6	7	rohan	56	67	57	301
7	8	sandhya	58	56	58	286
8	9	saranya	49	45	49	237

Get the column heading

Index(['roll.no', 'name', 'maths', 'science', 'social', 'total'],

dtype='object') Get the shape (no.of rows,no.of columns)

(9, 6)

Get the column heading

```
Index(['roll.no', 'name', 'maths', 'science', 'social', 'total'], dtype='object')
```

Get the shape (no.of rows,no.of columns) (9, 6)

Get particular column values

```
0    1
1    2
2    3
3    4
4    5
5    6
6    7
7    8
8    9
Name: roll.no, dtype: int64
```

Extract/slice the table values-[including this row, excluding this row]

	roll.no	name	maths	science	social	total
2	3	kaviya	80	80	80	400
3	4	racheal	89	87	89	441
4	5	rajan	90	98	90	466

Get the particular row values-through row number identification

```
roll.no        8
name      sandhya
maths         58
science       56
social        58
total        286
Name: 7, dtype: object
```

Get the particular row values-through 'Roll number'

```
   identification roll.no  name  maths  science  social
   total
4      5  rajan   90     98     90  466
```

Make an average of total marks:

```
0    56.8
1    69.2
2    80.0
3    88.2
4    93.2
5    70.6
6    60.2
7    57.2
8    47.4
Name: total, dtype: float64
```

STORING

Aim:

To store and manipulate input data from DataFrame to Excel/CSV through Pandas.

Description:

1. Create a DataFrame and store the data into specified Excel file

2. To read two Excel file data and merge through append function and store the merged data in to the new Excel file.

3. Using sort function, to sort and store the resultant data into a new Excel file

4. Read and display the CSV file

5.List the column headings and get the length of the table data.

Program :

```
import pandas as pd

d=pd.read_csv("LAS.csv")

df=pd.DataFrame(d)

print("Original DataFrame:\n",df)

#Second Dataframe input to another Excel file

d=pd.DataFrame([[20,'divya',95,85,76,256],  [14,'lakshmi',90,80,58,228],
[32,'ganesh',70,47,88,205]],

columns=['roll.no', 'name', 'maths', 'science', 'social', 'total'])

d.to_csv('pandas_to_csv.csv')

#Merging two Excel files input into third file x=pd.

read_csv("LAS.csv")

y=pd.read_csv('pandas_to_csv.csv')

y.drop(['Unnamed: 0'],axis = 1,inplace=True) z=pd.

concat([x,y],ignore_index=True)

z.to_csv('pandas_to_csv3.csv')

#Sorting the column vaules

df=z.sort_values(["roll.no"])

print("\nSorted Values:\n",df)
```

```
df.to_csv('pandas_to_csv4.csv'

) df=pd.read_csv('LAS.csv')

print(list(df))

print(format(len(df)))
```

Output:

Original DataFrame:

	roll.no	name	maths	science	social	total
0	1	deepa	50	67	50	284
1	2	dinesh	56	89	56	346
2	3	kaviya	80	80	80	400
3	4	racheal	89	87	89	441
4	5	rajan	90	98	90	466
5	6	ramya	67	76	67	353
6	7	rohan	56	67	57	301
7	8	sandhya	58	56	58	286
8	9	saranya	49	45	49	237

Sorted Values:

	roll.no	name	maths	science	social	total
0	1	deepa	50	67	50	284
1	2	dinesh	56	89	56	346
2	3	kaviya	80	80	80	400
3	4	racheal	89	87	89	441
4	5	rajan	90	98	90	466
5	6	ramya	67	76	67	353
6	7	rohan	56	67	57	301
7	8	sandhya	58	56	58	286
8	9	saranya	49	45	49	237
10	14	lakshmi	90	80	58	228
9	20	divya	95	85	76	256

11 32 ganesh 70 47 88 205 ['roll.

no', 'name', 'maths', 'science', 'social',

'total'] 9

Result:

The programs were run successfully