Ex. No: 7

Aim: *Perform following operations using pandas*

a. Filling NaN with string
b. Sorting based on column values
c. groupby().

7 (A):

Filling NaN with string: The fillna() method replaces the NULL values with a specified value.

Example:

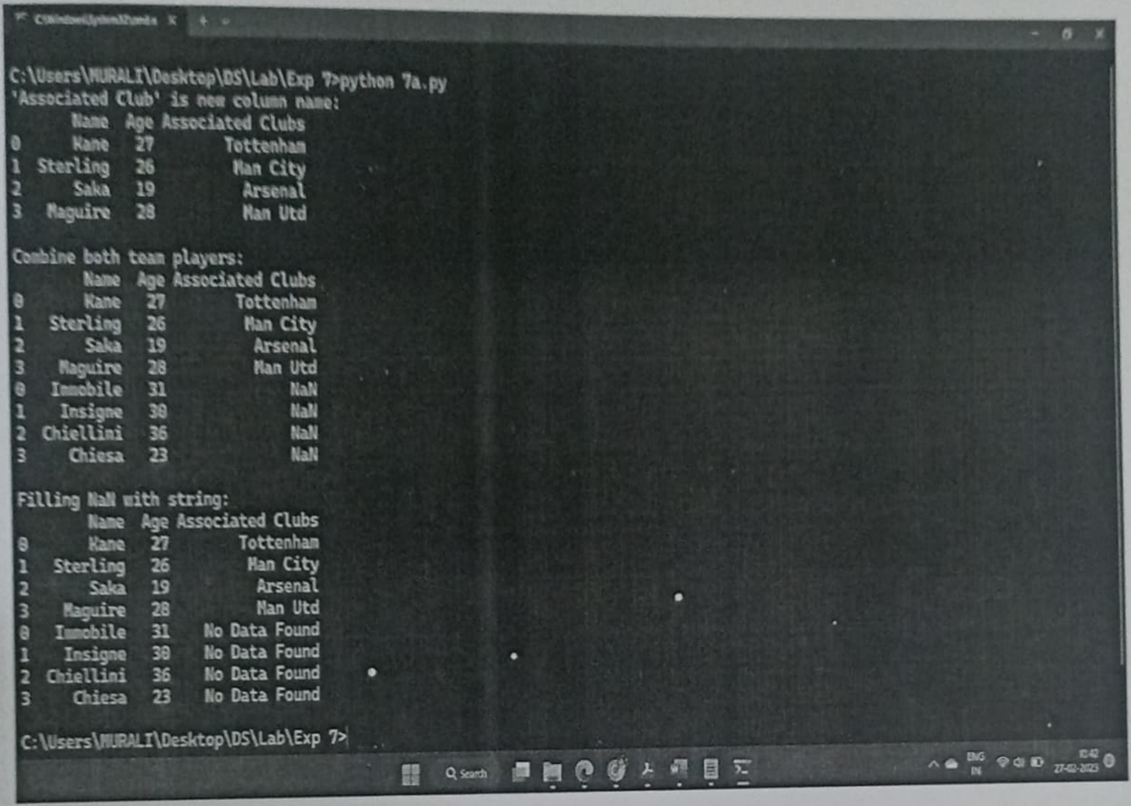
dataframe.py

```
import pandas as pd
def england():
   data_england = {'Name': ['Kane', 'Sterling', 'Saka', 'Maguire'], 'Age': [27, 26, 19, 28]}
   df_england = pd.DataFrame(data_england)
   return df_england
def italy():
   data_italy = {'Name': ['Immobile', 'Insigne', 'Chiellini', 'Chiesa'], 'Age': [31, 30, 36, 23]}
   df_italy = pd.DataFrame(data_italy)
   return df_italy
```

7a.py

```
import pandas as pd
from dataframe import *
df_england=england()
df_italy=italy()
club=['Tottenham', 'Man City', 'Arsenal', 'Man Utd']
print("'Associated Club' is new column name:")
df_england['Associated Clubs']=club
print(df_england)
print("\nCombine both team players:")
frames=[df_england, df_italy]
both_teams=pd.concat(frames)
print(both_teams)
print("\nFilling NaN with string:")
both_teams['Associated Clubs'].fillna('No Data Found', inplace=True)
print(both_teams)
```

Output:



```
C:\Users\MURALI\Desktop\DS\Lab\Exp 7>python 7a.py
'Associated Club' is new column name:
       Name  Age Associated Clubs
0      Kane   27        Tottenham
1  Sterling   26         Man City
2      Saka   19          Arsenal
3   Maguire   28          Man Utd

Combine both team players:
        Name  Age Associated Clubs
0       Kane   27        Tottenham
1   Sterling   26         Man City
2       Saka   19          Arsenal
3    Maguire   28          Man Utd
0   Immobile   31              NaN
1    Insigne   30              NaN
2  Chiellini   36              NaN
3     Chiesa   23              NaN

Filling NaN with string:
        Name  Age Associated Clubs
0       Kane   27        Tottenham
1   Sterling   26         Man City
2       Saka   19          Arsenal
3    Maguire   28          Man Utd
0   Immobile   31    No Data Found
1    Insigne   30    No Data Found
2  Chiellini   36    No Data Found
3     Chiesa   23    No Data Found

C:\Users\MURALI\Desktop\DS\Lab\Exp 7>
```

7 (B):

Sorting based on column values: The sort_values() method sorts the DataFrame by the specified label.

Example:

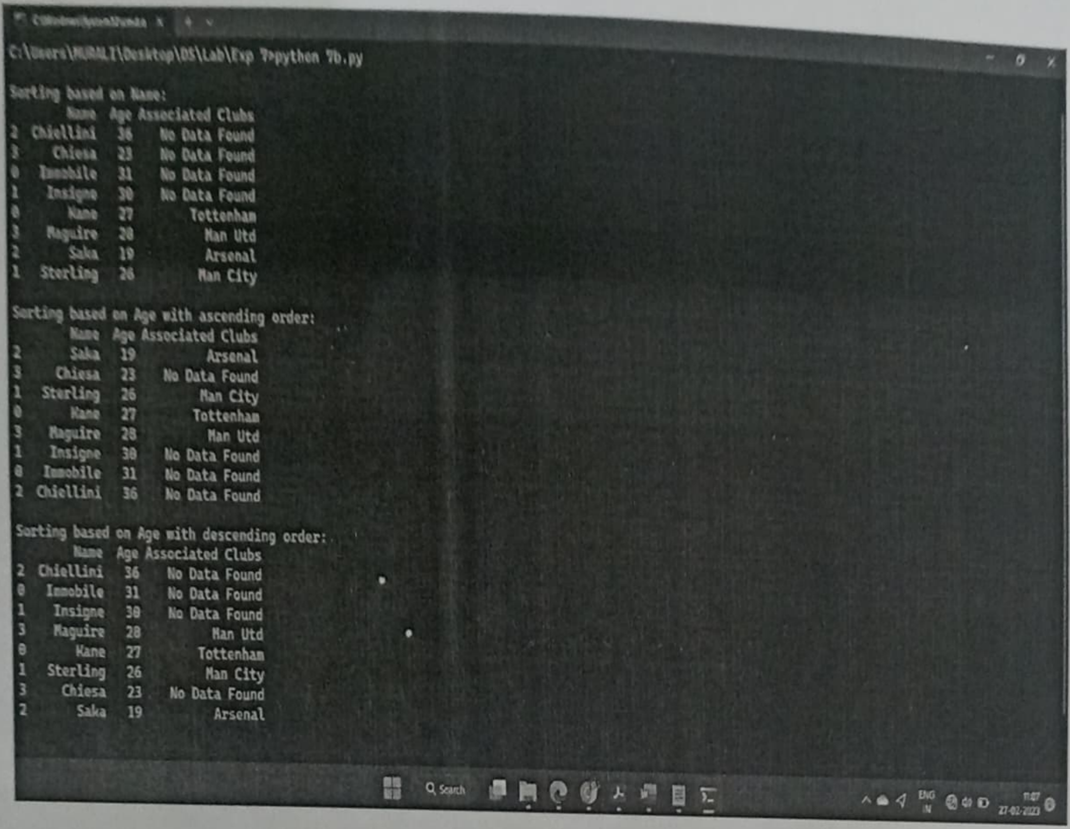
dataframe.py

```
import pandas as pd
def england():
   data_england = {'Name': ['Kane', 'Sterling', 'Saka', 'Maguire'], 'Age': [27, 26, 19, 28]}
   df_england = pd.DataFrame(data_england)
   return df_england
def italy():
   data_italy = {'Name': ['Immobile', 'Insigne', 'Chiellini', 'Chiesa'], 'Age': [31, 30, 36, 23]}
   df_italy = pd.DataFrame(data_italy)
   return df_italy
```

7b.py

```
import pandas as pd
from dataframe import *
df_england=england()
df_italy=italy()
club=['Tottenham', 'Man City', 'Arsenal', 'Man Utd']
#print("'Associated Club' is new column name:")
df_england['Associated Clubs']=club
#print(df_england)
#print("\nCombine both team players:")
frames=[df_england, df_italy]
both_teams=pd.concat(frames)
#print(both_teams)
#print("\nFilling NaN with string:")
both_teams['Associated Clubs'].fillna('No Data Found', inplace=True)
#print(both_teams)
print("\nSorting based on Name:")
print(both_teams.sort_values('Name'))
print("\nSorting based on Age with ascending order:")
print(both_teams.sort_values('Age'))
print("\nSorting based on Age with descending order:")
print(both_teams.sort_values('Age',ascending=False))
```

Output:



```
C:\Users\MURALI\Desktop\DS\Lab\Exp 7>python 7b.py

Sorting based on Name:
       Name  Age Associated Clubs
2 Chiellini   36    No Data Found
3    Chiesa   23    No Data Found
0  Immobile   31    No Data Found
1   Insigne   30    No Data Found
0      Kane   27        Tottenham
3   Maguire   28          Man Utd
2      Saka   19          Arsenal
1  Sterling   26         Man City

Sorting based on Age with ascending order:
       Name  Age Associated Clubs
2      Saka   19          Arsenal
3    Chiesa   23    No Data Found
1  Sterling   26         Man City
0      Kane   27        Tottenham
3   Maguire   28          Man Utd
1   Insigne   30    No Data Found
0  Immobile   31    No Data Found
2 Chiellini   36    No Data Found

Sorting based on Age with descending order:
       Name  Age Associated Clubs
2 Chiellini   36    No Data Found
0  Immobile   31    No Data Found
1   Insigne   30    No Data Found
3   Maguire   28          Man Utd
0      Kane   27        Tottenham
1  Sterling   26         Man City
3    Chiesa   23    No Data Found
2      Saka   19          Arsenal
```

7 (C):

groupby():

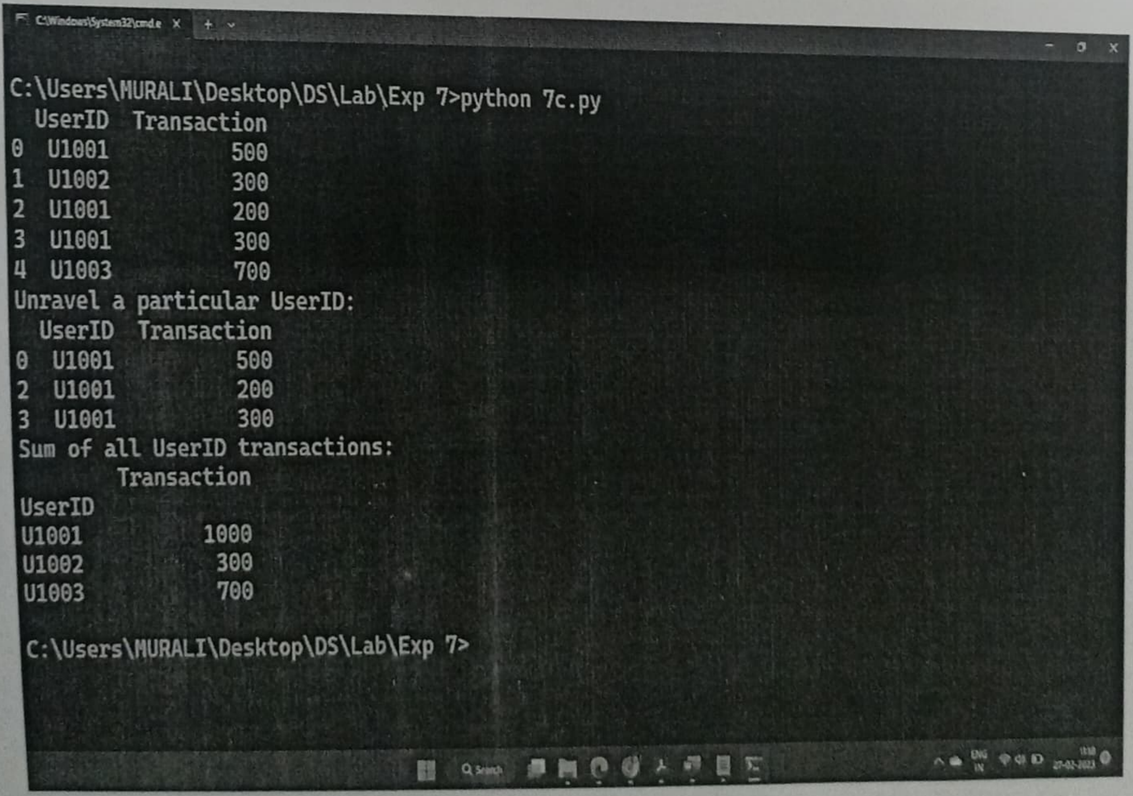
A groupby operation involves some combination of splitting the object, applying a function, and combining the results. This can be used to group large amounts of data and compute operations on these groups.

Example:

7c.py

```
import pandas as pd
a = {
    'UserID': ['U1001', 'U1002', 'U1001', 'U1001', 'U1003'],
    'Transaction': [500, 300, 200, 300, 700]
}
df_a = pd.DataFrame(a)
print(df_a)
print("Unravel a particular UserID:")
print(df_a.groupby('UserID').get_group('U1001'))
print("Sum of all UserID transactions:")
print(df_a.groupby('UserID').sum())
```

Output:



```
C:\Users\MURALI\Desktop\DS\Lab\Exp 7>python 7c.py
  UserID  Transaction
0  U1001          500
1  U1002          300
2  U1001          200
3  U1001          300
4  U1003          700
Unravel a particular UserID:
  UserID  Transaction
0  U1001          500
2  U1001          200
3  U1001          300
Sum of all UserID transactions:
        Transaction
UserID
U1001          1000
U1002           300
U1003           700

C:\Users\MURALI\Desktop\DS\Lab\Exp 7>
```