



A.Y. 2022-2023

Subject: Data Mining and Warehousing

SAP ID: 60004220253 – Devansh Mehta

Experiment 07

Aim: To perform OLAP operations such as slice, dice, roll up, drill down and pivot

Code:

```
import pandas as pd
```

```
data = {'Doctor': ['Dr. Smith', 'Dr. Johnson', 'Dr. Smith', 'Dr. Johnson'],  
        'Patient': ['Patient A', 'Patient B', 'Patient A', 'Patient B'],  
        'Time': ['2022-01-01', '2022-01-01', '2022-01-02', '2022-01-02'],  
        'Count': [1, 1, 1, 1],  
        'Charge': [100, 150, 120, 180]}
```

```
df = pd.DataFrame(data)
```

```
print("Original Data:")  
print(df)
```

```
sliced_data = df[df['Doctor'] == 'Dr. Smith']  
print("\n(i) Slice - Data for Doctor 'Dr. Smith':")  
print(sliced_data)
```

```
diced_data = df[(df['Doctor'] == 'Dr. Smith') & (df['Patient'] == 'Patient A')]  
print("\n(ii) Dice - Data for Doctor 'Dr. Smith' and Patient 'Patient A':")  
print(diced_data)
```

```
roll_up_data = df.groupby('Doctor').sum().reset_index()  
print("\n(iii) Roll Up - Total Count and Charge for each Doctor:")  
print(roll_up_data)
```

```
drill_down_data = df.groupby(['Doctor', 'Patient']).sum().reset_index()  
print("\n(iv) Drill Down - Count and Charge for each Doctor and Patient:")  
print(drill_down_data)
```

```
pivot_data = df.pivot_table(index='Doctor', columns='Patient', values='Charge',  
                              aggfunc='sum')  
print("\n(v) Pivot - Total Charge for each Doctor and Patient:")  
print(pivot_data)
```



A.Y. 2022-2023

Subject: Data Mining and Warehousing

SAP ID: 60004220253 – Devansh Mehta

Output:

```
Original Data:
  Doctor   Patient      Time  Count  Charge
0  Dr. Smith Patient A  2022-01-01    1    100
1  Dr. Johnson Patient B  2022-01-01    1    150
2  Dr. Smith Patient A  2022-01-02    1    120
3  Dr. Johnson Patient B  2022-01-02    1    180

(i) Slice - Data for Doctor 'Dr. Smith':
  Doctor   Patient      Time  Count  Charge
0  Dr. Smith Patient A  2022-01-01    1    100
2  Dr. Smith Patient A  2022-01-02    1    120

(ii) Dice - Data for Doctor 'Dr. Smith' and Patient 'Patient A':
  Doctor   Patient      Time  Count  Charge
0  Dr. Smith Patient A  2022-01-01    1    100
2  Dr. Smith Patient A  2022-01-02    1    120

(iii) Roll Up - Total Count and Charge for each Doctor:
  Doctor   Count  Charge
0  Dr. Johnson    2    330
1  Dr. Smith      2    220

(iv) Drill Down - Count and Charge for each Doctor and Patient:
  Doctor   Patient  Count  Charge
0  Dr. Johnson Patient B     2    330
1  Dr. Smith Patient A     2    220

(v) Pivot - Total Charge for each Doctor and Patient:
Patient   Patient A  Patient B
Doctor
Dr. Johnson    NaN    330.0
Dr. Smith    220.0    NaN
```