Amit Divekar | Assignment 1: Set C

GroupBy Operations and Aggregation - Set C

In [1]:
```
import pandas as pd
import numpy as np
```

Q. Create a following DataFrame named as "data" and split the dataframe into groups based on school code.

In [2]:
```
data = pd.DataFrame({
    'school': ['S1', 'S2', 'S3', 'S4', 'S5', 'S6'],
    'class': ['S001', 'S002', 'S003', 'S004', 'S005', 'S006'],
    'name': ['Asha', 'Fransis', 'Charlie', 'David', 'Shyam', 'Eashel'],
    'date_of_birth': ['15/05/2002', '17/05/2002', '16/02/1999', '25/09/1998'
    'age': [12, 12, 13, 14, 12, 16],
    'height': [173, 192, 186, 167, 151, 159],
    'weight': [35, 32, 33, 30, 31, 32],
    'address': ['Street1', 'Street2', 'Street3', 'Street4', 'Street5', 'Stre
})

print("Original DataFrame")
print(data)
```

```
Original DataFrame
  school class     name date_of_birth  age  height  weight  address
0     S1  S001     Asha    15/05/2002   12     173      35  Street1
1     S2  S002  Fransis    17/05/2002   12     192      32  Street2
2     S3  S003  Charlie    16/02/1999   13     186      33  Street3
3     S4  S004    David    25/09/1998   14     167      30  Street4
4     S5  S005    Shyam    11/05/2002   12     151      31  Street5
5     S6  S006   Eashel    15/09/1997   16     159      32  Street6
```

In [3]:
```
grouped = data.groupby('school')

print("\nGrouped object created")
```

```
Grouped object created
```

1. Check the type of GroupBy object.

In [4]:
```
print("\n1. Type of GroupBy object:")
print(type(grouped))
```

```
1. Type of GroupBy object:
<class 'pandas.core.groupby.generic.DataFrameGroupBy'>
```

2. Calculate mean, min, and max value of age for each school.

In [5]:
```
print("\n2. Mean, Min and Max age for each school:")
print(grouped['age'].agg(['mean', 'min', 'max']))
```

```
2. Mean, Min and Max age for each school:
        mean  min  max
school
S1      12.0   12   12
S2      12.0   12   12
S3      13.0   13   13
S4      14.0   14   14
S5      12.0   12   12
S6      16.0   16   16
```