Darshan UNIVERSITY

योगः कर्मसु कौशलम्

Python for Data Science - 2305CS303

Lab - 9

Roll No. : 111

Name : Dhara Maru

1. Create a Pandas Series containing names of 5 students.

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

students = pd.Series(["Dhara", "Hetvi", "Udit", "Piyu", "Manu"])
print(students)
```

```
0    Dhara
1    Hetvi
2     Udit
3     Piyu
4     Manu
dtype: object
```

2. Create a Series with student roll numbers as index and their IAT scores as values..

In [2]:
```
import pandas as pd

scores = pd.Series([85, 78, 92, 74, 88], index=[101, 102, 103, 104, 105])
print(scores)
```

```
101    85
102    78
103    92
104    74
105    88
dtype: int64
```

3. Create a time series (daily) from 2025-08-01 to 2025-08-10 representing attendance tracking for a student.

In [3]:
```
import pandas as pd

dates = pd.date_range(start="2025-08-01", end="2025-08-10", freq="D")
attendance = pd.Series([1, 1, 0, 1, 1, 1, 0, 1, 1, 1], index=dates)
print(attendance)
```

```
2025-08-01    1
2025-08-02    1
2025-08-03    0
2025-08-04    1
2025-08-05    1
2025-08-06    1
2025-08-07    0
2025-08-08    1
2025-08-09    1
2025-08-10    1
Freq: D, dtype: int64
```

4. Create a DataFrame for 10 students with the following columns: Roll No, Name, PDS, CA, CN, IAT.

(Use NumPy random module to generate scores)

In [4]:
```
import pandas as pd
import numpy as np

np.random.seed(1)
roll_no = range(101, 111)
names = ["Amit", "Riya", "Karan", "Sneha", "Vikram", "Meena", "Arjun", "Pooja"
pds = np.random.randint(60, 100, 10)
ca = np.random.randint(60, 100, 10)
cn = np.random.randint(60, 100, 10)
iat = np.random.randint(60, 100, 10)

df = pd.DataFrame({
    "Roll No": roll_no,
    "Name": names,
    "PDS": pds,
    "CA": ca,
    "CN": cn,
    "IAT": iat
})
print(df)
```

```
   Roll No    Name  PDS  CA  CN  IAT
0      101    Amit   97  72  89   67
1      102    Riya   72  67  74   82
2      103   Karan   68  66  64   61
3      104   Sneha   69  85  83   60
4      105  Vikram   71  80  83   77
5      106   Meena   65  97  90   68
6      107   Arjun   75  78  92   84
7      108   Pooja   60  80  82   73
8      109   Rahul   76  71  73   68
9      110    Neha   61  88  69   90
```

5. Display the first 3 rows of the DataFrame.

In [5]:
```
print(df.head(3))
```

```
   Roll No   Name  PDS  CA  CN  IAT
0      101   Amit   97  72  89   67
1      102   Riya   72  67  74   82
2      103  Karan   68  66  64   61
```

6. Display the last 2 rows of the DataFrame.

In [6]:
```
print(df.tail(2))
```

```
   Roll No   Name  PDS  CA  CN  IAT
8      109  Rahul   76  71  73   68
9      110   Neha   61  88  69   90
```

7. Use .describe() to summarize the numeric data.

In [7]:
```
print(df.describe())
```

```
         Roll No        PDS         CA         CN        IAT
count   10.00000  10.000000  10.000000  10.000000  10.000000
mean   105.50000  71.400000  78.400000  79.900000  73.000000
std      3.02765  10.469002   9.811558   9.480389  10.033278
min    101.00000  60.000000  66.000000  64.000000  60.000000
25%    103.25000  65.750000  71.250000  73.250000  67.250000
50%    105.50000  70.000000  79.000000  82.500000  70.500000
75%    107.75000  74.250000  83.750000  87.500000  80.750000
max    110.00000  97.000000  97.000000  92.000000  90.000000
```

8. Select only the Name column.

In [8]:
```
print(df["Name"])
```

```
0      Amit
1      Riya
2     Karan
3     Sneha
4    Vikram
5     Meena
6     Arjun
7     Pooja
8     Rahul
9      Neha
Name: Name, dtype: object
```

9. Select the columns PDS, CN, and IAT.

In [9]:
```
print(df[["PDS", "CN", "IAT"]])
```

```
   PDS  CN  IAT
0   97  89   67
1   72  74   82
2   68  64   61
3   69  83   60
4   71  83   77
5   65  90   68
6   75  92   84
7   60  82   73
8   76  73   68
9   61  69   90
```

10. Select the row with Roll No = 105 using loc.

In [10]:
```
print(df.loc[df["Roll No"] == 105])
```

```
   Roll No    Name  PDS  CA  CN  IAT
4      105  Vikram   71  80  83   77
```

11. Select the 4th row using iloc.

In [11]:
```
print(df.iloc[3])
```

```
Roll No      104
Name       Sneha
PDS           69
CA            85
CN            83
IAT           60
Name: 3, dtype: object
```

12. Select students with marks in PDS > 80.

In [12]:
```
print(df[df["PDS"] > 80])
```

```
   Roll No  Name  PDS  CA  CN  IAT
0      101  Amit   97  72  89   67
```

13. Select students with marks in CA < 70.

In [13]:
```
print(df[df["CA"] < 70])
```

```
   Roll No   Name  PDS  CA  CN  IAT
1      102   Riya   72  67  74   82
2      103  Karan   68  66  64   61
```

14. Select students with marks in CN > 85 and PDS > 80

In [14]:
```
print(df[(df["CN"] > 85) & (df["PDS"] > 80)])
```

```
   Roll No  Name  PDS  CA  CN  IAT
0      101  Amit   97  72  89   67
```

15. Add a new column Total Marks = PDS + CA + CN + IAT.

In [15]:
```
df["Total Marks"] = df["PDS"] + df["CA"] + df["CN"] + df["IAT"]
print(df)
```

```
   Roll No    Name  PDS  CA  CN  IAT  Total Marks
0      101    Amit   97  72  89   67          325
1      102    Riya   72  67  74   82          295
2      103   Karan   68  66  64   61          259
3      104   Sneha   69  85  83   60          297
4      105  Vikram   71  80  83   77          311
5      106   Meena   65  97  90   68          320
6      107   Arjun   75  78  92   84          329
7      108   Pooja   60  80  82   73          295
8      109   Rahul   76  71  73   68          288
9      110    Neha   61  88  69   90          308
```

16. Create a new DataFrame of students with Total Marks > 320.

In [16]:
```
df_high = df[df["Total Marks"] > 320]
print(df_high)
```

```
   Roll No   Name  PDS  CA  CN  IAT  Total Marks
0      101   Amit   97  72  89   67          325
6      107  Arjun   75  78  92   84          329
```