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", "Mummy", "Piyu", "Manu"])
print(students)
```

```
0    Dhara
1    Hetvi
2    Mummy
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([45,76,87,34,87], index=[101, 102, 103, 104, 105])
print(scores)
```

```
101    45
102    76
103    87
104    34
105    87
dtype: int64
```

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

In [7]:
```
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 [27]:
```
import pandas as pd
import numpy as np

np.random.seed(1)
roll_no = range(101, 111)
names = ["Dhara", "Riya", "Hetvi", "Nandi", "Bansi", "Piyu", "Manu", "Sujal",
pds = np.random.randint(70, 100, 10)
ca = np.random.randint(93, 100, 10)
cn = np.random.randint(23, 100, 10)
iat = np.random.randint(54, 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   Dhara   75  94  41   76
1      102    Riya   81  97  34   55
2      103   Hetvi   82  98  51   54
3      104   Nandi   78  97  52   71
4      105   Bansi   79  99  37   62
5      106    Piyu   81  94  73   78
6      107    Manu   75  95  91   67
7      108   Sujal   85  97  36   96
8      109   Rahul   70  99  32   62
9      110  Bhoomi   86  98  30   84
```

5. Display the first 3 rows of the DataFrame.

In [29]:
```
print(df.head(3))
```

```
   Roll No   Name  PDS  CA  CN  IAT
0      101  Dhara   75  94  41   76
1      102   Riya   81  97  34   55
2      103  Hetvi   82  98  51   54
```

6. Display the last 2 rows of the DataFrame.

In [30]:
```
print(df.tail(2))
```

```
   Roll No    Name  PDS  CA  CN  IAT
8      109   Rahul   70  99  32   62
9      110  Bhoomi   86  98  30   84
```

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

In [31]:
```
print(df.describe())
```

```
         Roll No        PDS         CA         CN       IAT
count   10.00000  10.000000  10.000000  10.000000  10.00000
mean   105.50000  79.200000  96.800000  47.700000  70.50000
std      3.02765   4.894441   1.873796  19.966917  13.28533
min    101.00000  70.000000  94.000000  30.000000  54.00000
25%    103.25000  75.750000  95.500000  34.500000  62.00000
50%    105.50000  80.000000  97.000000  39.000000  69.00000
75%    107.75000  81.750000  98.000000  51.750000  77.50000
max    110.00000  86.000000  99.000000  91.000000  96.00000
```

8. Select only the Name column.

In [32]:
```
print(df["Name"])
```

```
0     Dhara
1      Riya
2     Hetvi
3     Nandi
4     Bansi
5      Piyu
6      Manu
7     Sujal
8     Rahul
9    Bhoomi
Name: Name, dtype: object
```

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

In [33]:
```
print(df[["PDS", "CN", "IAT"]])
```

```
   PDS  CN  IAT
0   75  41   76
1   81  34   55
2   82  51   54
3   78  52   71
4   79  37   62
5   81  73   78
6   75  91   67
7   85  36   96
8   70  32   62
9   86  30   84
```

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

In [34]:
```
print(df.loc[df["Roll No"] == 105])
```

```
   Roll No   Name  PDS  CA  CN  IAT
4      105  Bansi   79  99  37   62
```

11. Select the 4th row using iloc.

In [35]:
```
print(df.iloc[3])
```

```
Roll No      104
Name       Nandi
PDS           78
CA            97
CN            52
IAT           71
Name: 3, dtype: object
```

12. Select students with marks in PDS > 80.

In [36]:
```
print(df[df["PDS"] > 80])
```

```
   Roll No    Name  PDS  CA  CN  IAT
1      102    Riya   81  97  34   55
2      103   Hetvi   82  98  51   54
5      106    Piyu   81  94  73   78
7      108   Sujal   85  97  36   96
9      110  Bhoomi   86  98  30   84
```

13. Select students with marks in CA < 70.

In [37]:
```
print(df[df["CA"] < 70])
```

```
Empty DataFrame
Columns: [Roll No, Name, PDS, CA, CN, IAT]
Index: []
```

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

In [38]:
```
print(df[(df["CN"] > 85) & (df["PDS"] > 80)])
```

```
Empty DataFrame
Columns: [Roll No, Name, PDS, CA, CN, IAT]
Index: []
```

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

In [39]:
```
df["Total Marks"] = df["PDS"] + df["CA"] + df["CN"] + df["IAT"]
print(df)
```

```
   Roll No    Name  PDS  CA  CN  IAT  Total Marks
0      101   Dhara   75  94  41   76          286
1      102    Riya   81  97  34   55          267
2      103   Hetvi   82  98  51   54          285
3      104   Nandi   78  97  52   71          298
4      105   Bansi   79  99  37   62          277
5      106    Piyu   81  94  73   78          326
6      107    Manu   75  95  91   67          328
7      108   Sujal   85  97  36   96          314
8      109   Rahul   70  99  32   62          263
9      110  Bhoomi   86  98  30   84          298
```

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

In [40]:
```
df_high = df[df["Total Marks"] > 320]
print(df_high)
```

```
   Roll No  Name  PDS  CA  CN  IAT  Total Marks
5      106  Piyu   81  94  73   78          326
6      107  Manu   75  95  91   67          328
```