Darshan UNIVERSITY

योगः कर्मसु कौशलम्

Python for Data Science - 2305CS303

Lab - 10

Roll No. : 111

Name : Dhara Maru

Student Score (.csv file)

1. Load the file student_scores.csv.

In [3]:
```
import pandas as pd
df = pd.read_csv("students_score.csv")
```

2. Show the first 5 rows.

In [4]:
```
df.head()
```

Out[4]:

	RollNo	Name	Math	Science	English
0	101	Aman	78	85	90
1	102	Riya	65	82	75
2	103	Kiran	90	88	92
3	104	Ravi	70	79	85
4	105	Meera	88	92	91

3. Display the index and column names.

In [5]:
```
df.index, df.columns
```

Out[5]:
```
(RangeIndex(start=0, stop=10, step=1),
 Index(['RollNo', 'Name', 'Math', 'Science', 'English'], dtype='object'))
```

4. Get descriptive statistics using .describe().

In [6]:
```
df.describe()
```

Out[6]:

	RollNo	Math	Science	English
count	10.00000	10.00000	10.000000	10.00000
mean	105.50000	77.40000	84.200000	85.60000
std	3.02765	8.40899	6.033241	6.60303
min	101.00000	65.00000	73.000000	75.00000
25%	103.25000	70.50000	79.750000	81.25000
50%	105.50000	77.50000	86.000000	87.00000
75%	107.75000	83.25000	88.000000	90.75000
max	110.00000	90.00000	92.000000	93.00000

5. Select the Name and Math columns.

In [7]:
```
df[["Name","Math"]]
```

Out[7]:

	Name	Math
0	Aman	78
1	Riya	65
2	Kiran	90
3	Ravi	70
4	Meera	88
5	John	81
6	Sara	77
7	Tom	69
8	Alice	84
9	Neha	72

6. Find all students who scored more than 80 in Science.

In [8]:
```
df[df["Science"]>80]
```

Out[8]:

	RollNo	Name	Math	Science	English
0	101	Aman	78	85	90
1	102	Riya	65	82	75
2	103	Kiran	90	88	92
4	105	Meera	88	92	91
5	106	John	81	87	93
6	107	Sara	77	90	89
8	109	Alice	84	88	85

7. Find all students with English < 75.

In [9]:
```
df[df["English"]<75]
```

Out[9]:

	RollNo	Name	Math	Science	English

8. Extract the last 3 rows.

In [10]:
```
df.tail(3)
```

Out[10]:

	RollNo	Name	Math	Science	English
7	108	Tom	69	73	80
8	109	Alice	84	88	85
9	110	Neha	72	78	76

9. Sort the DataFrame by Math in descending order.

(Hint : use df.sort_values(by = "column_name", ascending = True/ False))

In [11]:
```
df.sort_values(by="Math",ascending=False)
```

Out[11]:

	RollNo	Name	Math	Science	English
2	103	Kiran	90	88	92
4	105	Meera	88	92	91
8	109	Alice	84	88	85
5	106	John	81	87	93
0	101	Aman	78	85	90
6	107	Sara	77	90	89
9	110	Neha	72	78	76
3	104	Ravi	70	79	85
7	108	Tom	69	73	80
1	102	Riya	65	82	75

10. Set RollNo as the index and rename it "Student ID".

In [12]:
```
df2 = df.set_index("RollNo")
df2.index.name = "Student ID"
df2
```

Out[12]:

Student ID	Name	Math	Science	English
101	Aman	78	85	90
102	Riya	65	82	75
103	Kiran	90	88	92
104	Ravi	70	79	85
105	Meera	88	92	91
106	John	81	87	93
107	Sara	77	90	89
108	Tom	69	73	80
109	Alice	84	88	85
110	Neha	72	78	76

11. Reset the index back.

In [13]:
```
df2.reset_index()
```

Out[13]:

	Student ID	Name	Math	Science	English
0	101	Aman	78	85	90
1	102	Riya	65	82	75
2	103	Kiran	90	88	92
3	104	Ravi	70	79	85
4	105	Meera	88	92	91
5	106	John	81	87	93
6	107	Sara	77	90	89
7	108	Tom	69	73	80
8	109	Alice	84	88	85
9	110	Neha	72	78	76

12. Add a new column Total = Math + Science + English.

In [14]:
```
df["Total"] = df["Math"] + df["Science"] + df["English"]
df
```

Out[14]:

	RollNo	Name	Math	Science	English	Total
0	101	Aman	78	85	90	253
1	102	Riya	65	82	75	222
2	103	Kiran	90	88	92	270
3	104	Ravi	70	79	85	234
4	105	Meera	88	92	91	271
5	106	John	81	87	93	261
6	107	Sara	77	90	89	256
7	108	Tom	69	73	80	222
8	109	Alice	84	88	85	257
9	110	Neha	72	78	76	226

13. Find the student with the highest Total score.

In [15]:
```
df[df["Total"]==df["Total"].max()]
```

Out[15]:

	RollNo	Name	Math	Science	English	Total
4	105	Meera	88	92	91	271

14. Get the Top 3 students with the highest total score.

In [16]:
```
df.nlargest(3,"Total")
```

Out[16]:

	RollNo	Name	Math	Science	English	Total
4	105	Meera	88	92	91	271
2	103	Kiran	90	88	92	270
5	106	John	81	87	93	261

15. Get the average marks in each subject.

In [17]:
```
df[["Math","Science","English"]].mean()
```

Out[17]:
```
Math       77.4
Science    84.2
English    85.6
dtype: float64
```