R Data Analysis Project: Using dplyr (select(), filter(), arrange())

Project Title: Student Performance Data Analysis Using dplyr

You are required to create and analyze your own dataset in **R Studio** using the **dplyr** package. The project will focus on manipulating data using the functions **select()**, **filter()**, and **arrange()**.

R Data Analysis Project Dataset: Student Performance Data

Use the following data to populate your data frame. Create a data frame with **20 records and 10 columns** exactly as shown below.

Column Names (Use Exactly These)

1. Student_ID
2. Name
3. Gender
4. Age
5. Course
6. Year
7. Maths_Marks
8. Programming
9. Statistics
10. Attendance

Data to Enter

Student_ID	Name	Gender	Age	Course	Year	Maths_Marks	Programming	Statistics	Attendance
101	David	Male	20	IT	1	78	85	72	90
102	Alex	Male	21	CS	2	88	79	81	85
103	Mary	Female	19	DS	1	92	87	90	95
104	John	Male	22	IT	3	65	70	68	75
105	Sarah	Female	20	CS	2	74	82	77	88
106	Brian	Male	23	IT	4	59	65	60	70
107	Grace	Female	21	DS	3	83	91	85	93
108	Peter	Male	20	CS	1	71	69	73	80

109	Lucy	Female	22	IT	4	90	88	92	96
110	Kevin	Male	19	DS	1	76	72	70	78
111	Faith	Female	21	CS	2	85	89	84	91
112	James	Male	23	IT	4	67	74	69	82
113	Irene	Female	20	DS	2	93	95	91	97
114	Mark	Male	22	CS	3	72	68	75	76
115	Diana	Female	19	IT	1	81	86	79	89
116	Samuel	Male	24	DS	4	58	62	64	68
117	Ruth	Female	21	CS	3	87	90	88	94
118	Daniel	Male	20	IT	2	69	73	71	83
119	Esther	Female	22	DS	3	91	93	89	96
120	Joseph	Male	23	CS	4	64	66	62	72

Part 2: Load dplyr

Install (if necessary) and load the dplyr package.

Part 3: Data Analysis Tasks

Use **dplyr functions only** to answer the following questions.

Section A: Using `select()`

1. Display only the columns: **Student_ID, Name, Course, and Programming**.
2. Select all columns related to marks only.
3. Display the dataset excluding the **Gender** and **Age** columns.
4. Reorder the dataset to show:
     o Name first
     o Course second

o   Attendance last

Section B: Using `filter()`

5. Display students who scored **above 70 in Maths**.
6. Show all students whose **Attendance is below 75%**.
7. Display students enrolled in the **IT course** only.
8. Show female students who scored **above 60 in Programming**.
9. Display students in **Year 3 or Year 4**.
10. Show students whose **Statistics marks are between 50 and 80**.

Section C: Using `arrange()`

11. Arrange the dataset by **Name in alphabetical order**.
12. Sort students by **Maths_Marks from highest to lowest**.
13. Arrange students by **Attendance in ascending order**.
14. Sort the dataset first by **Course**, then by **Programming marks (descending)**.

Section D: Combined Operations

15. Display only the **Name, Course, and Maths_Marks** of students who scored above 75 in Maths, sorted from highest to lowest.
16. Show students with attendance below 70%, displaying only their Name and Attendance, arranged from lowest attendance to highest.
17. Display Programming marks for students in the CS course, sorted alphabetically by name.