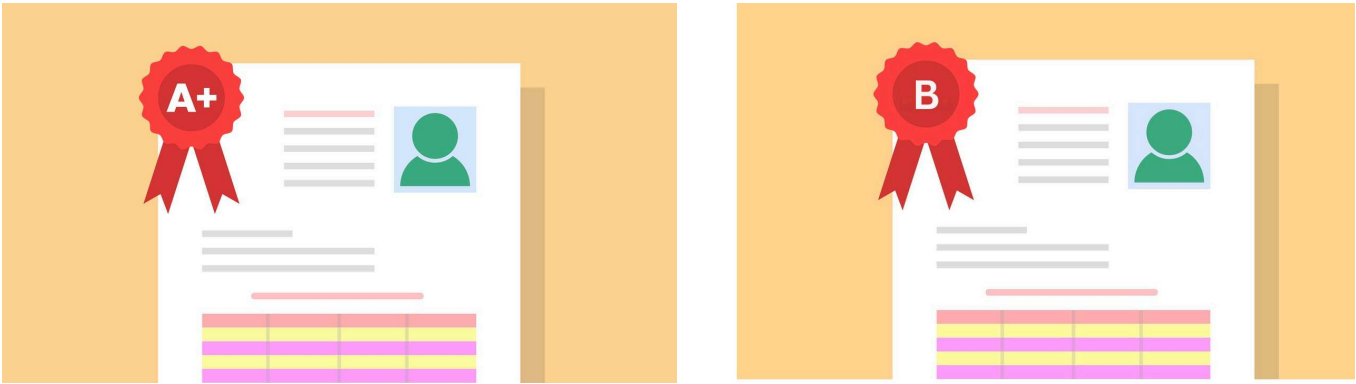


PROJECT: FACTORS THAT FUEL STUDENT PERFORMANCE



In today's fast-paced and competitive educational environment, understanding the factors that influence student success is more important than ever. Just like the transport system in a bustling city like London must adapt to serve its residents, schools and educators must adapt to meet the needs of students. In this project, we will take a deep dive into a dataset containing rich details about various aspects of student life, such as hours studied, sleep patterns, attendance, and more, to uncover what truly impacts exam performance.

The dataset we'll be working with includes a wide range of factors influencing student performance. By analyzing this data, we'll be able to identify key drivers of success and provide insights that could help students, teachers, and policymakers make informed decisions. The table we'll use for this project is called `student_performance` and includes the following data:

Column	Definition	Data type
attendance	Percentage of classes attended	float
extracurricular_activities	Participation in extracurricular activities	varchar (Yes, No)
sleep_hours	Average number of hours of sleep per night	float
tutoring_sessions	Number of tutoring sessions attended per month	integer
teacher_quality	Quality of the teachers	varchar (Low, Medium, High)
exam_score	Final exam score	float

You will execute SQL queries to answer three questions, as listed in the instructions.

Projects Data DataFrame as avg_exam_score_by_study_and_exti

-- avg_exam_score_by_study_and_extracurricular
-- Edit the query below as needed
SELECT

```
    hours_studied,  
    AVG(exam_score) AS avg_exam_score  
FROM student_performance  
WHERE extracurricular_activities= 'Yes'  
    AND hours_studied>10  
GROUP BY hours_studied  
ORDER BY hours_studied DESC;
```

index	...	↑↓	hours_studied	...	↑↓	avg_exam_score
			0			43
			1			39
			2			38
			3			37
			4			36
			5			35
			6			34
			7			33
			8			32
			9			31
			10			30
			11			29
			12			28
			13			27
			14			26
			15			25

Rows: 30

Expand

Projects Data DataFrame as avg_exam_score_by_hours_studied.

-- avg_exam_score_by_hours_studied_range
-- Add solution code below
SELECT
 CASE WHEN hours_studied>=1 AND hours_studied<=5 THEN '1-5 hours'
 WHEN hours_studied>=6 AND hours_studied<=10 THEN '6-10 hours'
 WHEN hours_studied>=11 AND hours_studied<=15 THEN '11-15 hours'
 WHEN hours_studied>=15 THEN '16+ hours'
 ELSE NULL END AS hours_studied_range,
 AVG(exam_score) AS avg_exam_score
FROM student_performance
GROUP BY hours_studied_range
ORDER BY avg_exam_score DESC;

index	...	↑↓	hours_studied_range	...	↑↓	avg_exam_score
			0 16+ hours			61
			1 11-15 hours			61
			2 6-10 hours			61
			3 1-5 hours			61

Rows: 4

Expand

 Projects Data DataFrame as s

-- student_exam_ranking
-- Add solution code below

```
SELECT
    attendance,
    hours_studied,
    sleep_hours,
    tutoring_sessions,
    DENSE_RANK() OVER(ORDER BY exam_score DESC) AS exam_rank
FROM student_performance
ORDER BY exam_rank
LIMIT 30;
```

...	↑↓ a...	...	↑↓	hours_...	...	↑↓	sle...	...	↑↓	tutoring_sessi...	...	↑↓	e...	...	↑↓
	0		↑↓	98		↑↓	27		↑↓	6		↑↓	5		↑↓
	1			89			18			4			3		2
	2			90			14			8			4		3
	3			83			23			4			1		3
	4			96			28			4			1		4
	5			90			28			9			0		4
	6			83			16			8			2		4
	7			83			15			7			2		5
	8			74			21			6			1		5
	9			99			25			7			0		5
	10			93			18			7			2		6
	11			90			18			6			1		7
	12			67			21			6			1		7
	13			96			24			6			2		8
	14			98			25			7			1		8
	15			76			29			8			2		8

Rows: 30

 Expand