



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

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

1) Do more study hours and extracurricular activities lead to better scores?

Projects Data DataFrame as avg_exam_score_by_study_and_exti

```
-- avg_exam_score_by_study_and_extracurricular
SELECT
    hours_studied,
    AVG(exam_score) AS avg_exam_score
FROM student_performance
WHERE hours_studied > 10 AND extracurricular_activities = 'Yes'
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

From our output, we can see that students who participate in extracurricular activities and study more hours tend to perform better on exams. As study hours increased, the average exam score increased.

2) Is there a sweet spot for study hours? Explore how different ranges of study hours impact exam performance by calculating the average exam score for each study range.

Projects Data DataFrame as avg_exam_score_by_hours_studied.

```
-- avg_exam_score_by_hours_studied_range
SELECT
    CASE
        WHEN hours_studied BETWEEN 1 AND 5 THEN '1-5 hours'
        WHEN hours_studied BETWEEN 6 AND 10 THEN '6-10 hours'
        WHEN hours_studied BETWEEN 11 AND 15 THEN '11-15 hours'
        ELSE '16+ hours'
    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_sco
		0	16+ hours			
		1	11-15 hours			
		2	6-10 hours			
		3	1-5 hours			

Rows: 4

Expand

From our output, we can see that the more hours a student studies, they perform better on the exams.

3) How can we display each student's relative academic standing in the class using exam scores while preserving score confidentiality?

Projects Data DataFrame as stuc

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

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

Rows: 30

Expand