

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.

 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'
group by hours_studied
order by hours_studied desc
limit 30;
```

index	...	↑↓	hours_studied	...	↑↓	avg_exam_score	...	↑↓
		0			43			
		1			39			
		2			38			
		3			37			
		4			36	70.4285714		
		5			35	72.3		
		6			34	71.1		
		7			33	70.3333333		
		8			32	71.		
		9			31	70.5531914		
		10			30	71.4328358		
		11			29	70.256097		
		12			28	69.8256880		
		13			27	69.7768595		
		14			26	68.801369		
		15			25			

Rows: 30    ↓

 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 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'
  when hours_studied >= 16 then '16+ hours'
END 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			67.9233633	
		1	11-15 hours			65.2043859	
		2	6-10 hours			64.2254901	
		3	1-5 hours			62.6271186	

Rows: 4    ↓

 Projects Data    DataFrame as    student\_exam\_ranking

```
-- 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 asc
limit 30;
```

i...	...	↑↓	attenda...	...	↑↓	hours_studied	...	↑↓	sleep_hours	...	↑↓	tutoring_sessions	...	↑↓	exam...	...	↑
		0			98			27			6			5			▲
		1			89			18			4			3			
		2			90			14			8			4			
		3			83			23			4			1			
		4			96			28			4			1			
		5			90			28			9			0			
		6			83			16			8			2			
		7			83			15			7			2			
		8			74			21			6			1			
		9			99			25			7			0			
		10			93			18			7			2			
		11			90			18			6			1			
		12			67			21			6			1			
		13			96			24			6			2			
		14			98			25			7			1			
		15			76			29			8			2			▼

Rows: 30    ↓