



Student Performance Analysis

Using SQL



JULY 11, 2025
MUHAMMAD SAAD NAEEM



Student Performance Analysis Using SQL

By: MUHAMMAD SAAD NAEEM

Date: 11TH July 2025

1. Introduction

This project analyzes student performance based on gender, parental education, lunch type, and test preparation status. The goal is to discover insights that explain what factors influence student grades in math, reading, and writing.

2. Dataset Description

- Source:** Kaggle – [Student Performance Dataset](#)
- Total Rows:** 1000 students
- Columns:**

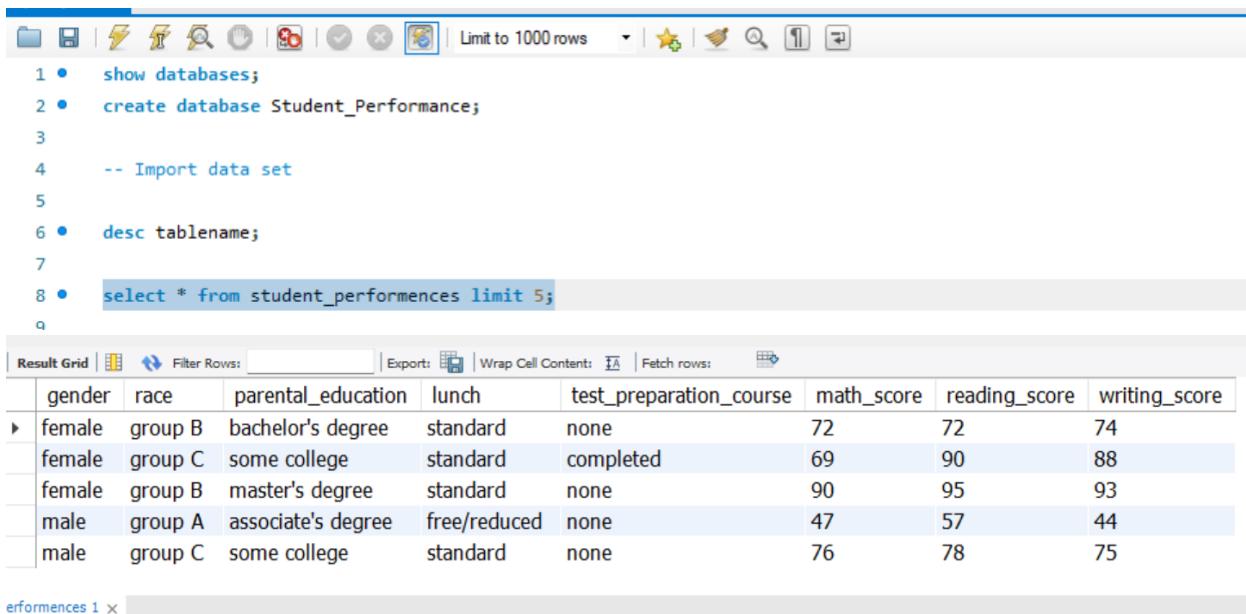
Column Name	Description
gender	Male or Female
race	Socioeconomic group (Group A to E)
parental_education	Highest education level of the parent
lunch	Standard or free/reduced
test_prep_course	None or completed
math_score	Score out of 100
reading_score	Score out of 100
writing_score	Score out of 100

❓ 3. Questions to Explore

1. What is the average score in math, reading, and writing?
2. What is the average score by gender?
3. How does parental education impact performance?
4. Does lunch type affect scores?
5. What is the effect of test preparation course?
6. Who are the top 10 highest scoring students?
7. How many students passed all three subjects?
8. How many students Fail all three subjects?

🧠 4. SQL Queries and Screenshots

◆ Starters



The screenshot shows a MySQL Workbench interface. At the top, there's a toolbar with various icons. Below it is a query editor window containing the following SQL code:

```

1 • show databases;
2 • create database Student_Performance;
3
4 -- Import data set
5
6 • desc tablename;
7
8 • select * from student_performences limit 5;
q

```

Below the query editor is a results grid titled "Result Grid". It displays the following data:

gender	race	parental_education	lunch	test_preparation_course	math_score	reading_score	writing_score
female	group B	bachelor's degree	standard	none	72	72	74
female	group C	some college	standard	completed	69	90	88
female	group B	master's degree	standard	none	90	95	93
male	group A	associate's degree	free/reduced	none	47	57	44
male	group C	some college	standard	none	76	78	75

```

10    -- checking total column
11
12 •  select count(*) from student_performences;
13
14    -- checking null values
15
16 •  select count(*) from

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
count(*)				

▶ 1000

Checking null Values

```

13
14    -- checking null values
15
16 •  select count(*) from
17    student_performences
18    where gender is NULL
19    or race is null
20    or parental_education is null
21    or lunch is null
22    or test_preparation_course is null
23    or math_score is null
24    or reading_score is null
25    or writing_score is null;

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
count(*)				

▶ 0

Finding Unique Values

```

28    -- finding unique values
29
30 •  select * from student_performences limit 5;
31 •  select distinct(gender) from student_performences;
32 •  select distinct(race) from student_performences;
33 •  select distinct(parental_education) from student_performences;
34 •  select distinct(lunch) from student_performences;
35 •  select distinct(test_preparation_course) from student_performences;
36
37

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
gender				

▶ female

male

◆ Q1: Average Score in Each Subject

Average score of math

```
38    -- Finding Average score  
39  
40 •  select avg(math_score) from student_performences; -- 66.08  
41 •  select avg(reading_score) from student_performences; -- 69.16  
42 •  select avg(writing_score) from student_performences; -- 68.05  
43
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	avg(math_score)			
▶	66.0890			

Average score of reading

```
38    -- Finding Average score  
39  
40 •  select avg(math_score) from student_performences; -- 66.08  
41 •  select avg(reading_score) from student_performences; -- 69.16  
42 •  select avg(writing_score) from student_performences; -- 68.05  
43
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	avg(reading_score)			
▶	69.1690			

Average score of writing

```
37  
38    -- Finding Average score  
39  
40 •  select avg(math_score) from student_performences; -- 66.08  
41 •  select avg(reading_score) from student_performences; -- 69.16  
42 •  select avg(writing_score) from student_performences; -- 68.05  
43
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	avg(writing_score)			
▶	68.0540			

Overall Average Score

```

43
44 •   select avg((math_score + writing_score + reading_score)/3) as average_combine from student_performences;
45
46 -- whole class average score lies between 67.77
47
48 -- Gender wise average score
49

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
average_combine				
67.77066667				

◆ Q2: Average Score by Gender

Math-female

```

48 -- Gender wise average score
49
50 •   select * from student_performences limit 5;
51
52 •   select avg(math_score) as female_math_score from student_performences where gender = 'female'; -- 63.66
53 •   select avg(math_score) as male_math_score from student_performences where gender = 'male'; -- 68.72
54

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
female_math_score				
63.6332				

Math-Male

```

48 -- Gender wise average score
49
50 •   select * from student_performences limit 5;
51
52 •   select avg(math_score) as female_math_score from student_performences where gender = 'female'; -- 63.66
53 •   select avg(math_score) as male_math_score from student_performences where gender = 'male'; -- 68.72
54

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
male_math_score				
68.7282				

Reading-Female

```

54
55 • select avg(reading_score) as female_reading_score from student_performences where gender = 'female'; -- 72.60
56 • select avg(reading_score) as male_reading_score from student_performences where gender = 'male'; -- 65.47
57

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
female_reading_score				

▶ 72.6081

Reading-Male

```

54
55 • select avg(reading_score) as female_reading_score from student_performences where gender = 'female'; -- 72.60
56 • select avg(reading_score) as male_reading_score from student_performences where gender = 'male'; -- 65.47
57

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
male_reading_score				

▶ 65.4730

Writing-Female

```

57
58 • select avg(writing_score) as female_writing_score from student_performences where gender = 'female'; -- 72.46
59 • select avg(writing_score) as male_writing_score from student_performences where gender = 'male'; -- 63.31
60

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
female_writing_score				

▶ 72.4672

Writing-Male

```

58 • select avg(writing_score) as female_writing_score from student_performences where gender = 'female'; -- 72.46
59 • select avg(writing_score) as male_writing_score from student_performences where gender = 'male'; -- 63.31
60

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
male_writing_score				

▶ 63.3112

Overall-Female

```

62
63 •   select (avg(math_score + reading_score + writing_score)/3)
64     as total_female_avg_score from student_performences
65     where gender = 'female'; -- 69.56
66

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	total_female_avg_score			
▶	69.56949807			

Overall-Male

```

66
67 •   select (avg(math_score + reading_score + writing_score)/3)
68     as total_male_avg_score from student_performences
69     where gender = 'male'; -- 65.83
70
71

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	total_male_avg_score			
▶	65.83748271			

◆ Q3: Average score based on race group

Group-A

```

72      -- Average score based on Race group
73
74 •   select distinct(race) from student_performences;
75
76      -- Group A
77
78 •   select (avg(math_score + reading_score + writing_score)/3)
79     as total_groupA_avg_score from student_performences
80     where race = 'group A'; -- 62.99
81

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	total_groupA_avg_score			
▶	62.99250936			

Group-B

```

81
82      -- Group B
83
84 •   select (avg(math_score + reading_score + writing_score)/3)
85      as total_groupB_avg_score from student_performences
86      where race = 'group B'; -- 65.46
87

```

Result Grid | Filter Rows: Export: Wrap Cell Content:

total_groupA_avg_score

62.99250936

Group-C

```

87
88      -- GROUP C
89
90 •   select (avg(math_score + reading_score + writing_score)/3)
91      as total_groupC_avg_score from student_performences
92      where race = 'group C'; -- 67.13
93

```

Result Grid | Filter Rows: Export: Wrap Cell Content:

total_groupC_avg_score

67.13166144

Group-D

```

94      -- Group D
95
96 •   select (avg(math_score + reading_score + writing_score)/3)
97      as total_groupB_avg_score from student_performences
98      where race = 'group D'; -- 69.17
99

```

Result Grid | Filter Rows: Export: Wrap Cell Content:

total_groupB_avg_score

69.17938931

Group-E

```

99
100    -- Group E
101
102 •   select (avg(math_score + reading_score + writing_score)/3)
103     as total_groupE_avg_score from student_performences
104     where race = 'group E'; -- 72.75
105

```

Result Grid	
	total_groupE_avg_score
▶	72.75238095

◆ Q4: Student score based on Parental Education

Associate Degree

```

106
107    -- parental education level affect student performance
108
109 •   select distinct(parental_education) from student_performences;
110
111    -- Associate degree
112
113 •   select (avg(math_score + reading_score + writing_score)/3)
114     as student_score_associate_degree from student_performences
115     where parental_education = "associate's degree"; -- 69.56
116

```

Result Grid	
	student_score_associate_degree
▶	69.56906907

Bachelor Degree

```

118
119 •   select (avg(math_score + reading_score + writing_score)/3)
120     as student_score_bachelors_degree from student_performences
121     where parental_education = "bachelor's degree"; -- 71.92
122

```

Result Grid	
	student_score_bachelors_degree
▶	71.92372881

High School

```

123      -- high school
124
125 •   select (avg(math_score + reading_score + writing_score)/3)
126     as student_score_high_school from student_performences
127     where parental_education = "high school"; -- 63.03
128

```

		Result Grid		Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	student_score_high_school								
▶	63.09693878								

Master Degree

```

129      -- master degree
130
131 •   select (avg(math_score + reading_score + writing_score)/3)
132     as student_score_Masters_degree from student_performences
133     where parental_education = "master's degree"; -- 73.59
134

```

		Result Grid		Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	student_score_Masters_degree								
▶	73.59887006								

Some College

```

135      -- some college
136
137 •   select (avg(math_score + reading_score + writing_score)/3)
138     as student_score_some_college from student_performences
139     where parental_education = "some college"; -- 68.47
140

```

		Result Grid		Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	student_score_some_college								
▶	68.47640118								

Some high school

```

140
141      -- some high school
142
143 •   select (avg(math_score + reading_score + writing_score)/3)
144      as student_score_some_high_school from student_performences
145      where parental_education = "some high school"; -- 65.11
146

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
student_score_some_high_school		65.10800745		

◆ Q5: does lunch type effect score?

Standard Lunch

```

151
152      -- standard lunch
153
154 •   select (avg(math_score + reading_score + writing_score)/3)
155      as standard_lunch_score from student_performences
156      where lunch = "standard"; -- 70.83
157

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
standard_lunch_score		70.83720930		

Free Lunch

```

157
158      -- free/reduced lunch
159
160 •   select (avg(math_score + reading_score + writing_score)/3)
161      as free_lunch_score from student_performences
162      where lunch = "free/reduced"; -- 62.199
163

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
free_lunch_score		62.19906103		

◆ Q6: Preparation Completeness effect score?

Completed-Score

```

164 •      select distinct(test_preparation_course) from student_performences;
165
166      -- who complete preparations
167
168 •      select (avg(math_score + reading_score + writing_score)/3)
169      as complete_prepartion_score from student_performences
170      where test_preparation_course = "completed"; -- 72.66

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	complete_prepartion_score			
▶	72.66945996			

None

```

172      -- didn't complete preparations
173
174 •      select (avg(math_score + reading_score + writing_score)/3)
175      as none_prepation_score from student_performences
176      where test_preparation_course = "none"; -- 65.03

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	none_prepation_score			
▶	65.03894081			

- ◆ Q7: Who are the top 10 highest scoring students based on total or average score?

Total top 10

```

182      -- top 10  total score
183
184 •      select gender,math_score,reading_score,writing_score ,
185      (math_score+reading_score+writing_score) as total_score from
186      student_performences order by total_score desc limit 10j
187
188

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:	Fetch rows:
	gender	math_score	reading_score	writing_score	total_score
▶	female	100	100	100	300
	male	100	100	100	300
	female	100	100	100	300
	female	99	100	100	299
	female	97	100	100	297
	female	98	100	99	297
	female	96	100	100	296
	male	100	97	99	296
	female	94	99	100	293
	male	100	100	93	293

Average top 10

```

188 -- top 10 average score
189
190 • select gender,math_score,reading_score,writing_score ,
191 ((math_score+reading_score+writing_score)/3) as average_score from
192 student_performences order by average_score desc limit 10;

```

Result Grid | Filter Rows: Export: Wrap Cell Content: Fetch rows:

gender	math_score	reading_score	writing_score	average_score
female	100	100	100	100.0000
male	100	100	100	100.0000
female	100	100	100	100.0000
female	99	100	100	99.6667
female	97	100	100	99.0000
female	98	100	99	99.0000
female	96	100	100	98.6667
male	100	97	99	98.6667
female	94	99	100	97.6667
male	100	100	93	97.6667

◆ Q8: Students Pass/Fail subject wise

Math

```

194 -- How many students passed all three subjects (assuming pass mark = 40)?
195
196 -- math score
197
198 • select count(math_score) from student_performences where math_score >= 40;
199
200 -- pass 960 in math
201 -- fail 40 in math
202

```

Result Grid | Filter Rows: Export: Wrap Cell Content: Fetch rows:

count(math_score)
960

Reading

```

---
203 -- reading score
204
205 • select count(reading_score) from student_performences where reading_score >= 40;
206
207 -- pass 974 in reading
208 -- fail 26 in reading

```

Result Grid | Filter Rows: Export: Wrap Cell Content: Fetch rows:

count(reading_score)
974

Writing

```

211
212 •   select count(writing_score) from student_performences where writing_score >= 40;
213
214      -- pass 968 in writing
215      -- fail 32 in writing
216
217

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
count(writing_score)				

▶ 968

◆ Q9: Students who pass/Fail in all subjects?

Pass in all

```

219
220 •   select count(*) as pass_studnets
221     from student_performences
222     where reading_score >= 40 and writing_score >= 40 and math_score >= 40 ;
223
224      -- pass studnets 949

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
pass_studnets				

▶ 949

Fail in all

```

226
227      -- students who fail in all subjects
228
229 •   select count(*) as fail_studnets
230     from student_performences
231     where reading_score < 40 and writing_score < 40 and math_score < 40 ;
232
233      -- failers 18

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
fail_studnets				

▶ 18

◆ Q10: Which subject has highest average?

```

235
236    -- what subject has the highest average
237
238 •   select avg(reading_score),avg(writing_score),avg(math_score) from student_performences;
239
240    -- reading avg score is highest
241
242

```

Result Grid			
	avg(reading_score)	avg(writing_score)	avg(math_score)
▶	69.1690	68.0540	66.0890

📌 5. Key Insights

- Students who completed the **test prep course** scored significantly higher.
- **Females scored better in reading and writing**, while **males slightly led in math**.
- Students with **parents holding master's or bachelor's degrees** performed best.
- **Free/reduced lunch students** had consistently lower scores.
- Students in **Group E** performed the best overall among all race/ethnic groups.
- On Average students performed higher in **reading** score

✓ 6. Conclusion

The analysis shows that **preparation, family background, and economic support** are key factors in student performance. Targeted support programs like test prep and parental awareness may improve outcomes.

⚠ 7. Limitations

- Dataset only includes scores for three subjects.
- No age, school name, location, or behavioral data.
- Limited to 1000 rows; may not generalize globally.

8. Future Exploration

- Add attendance, school performance, and teacher feedback.
- Compare student results across schools or cities.
- Use machine learning models for prediction based on input factors.

MUHAMMAD SAAD NAEEM