Education SQL Project from Kaggle.com

Sourced dataset from
“Students Performance in Exams”
From Kaggle.com

Upload CSV into Postgres SQL database using Terminal commands.

```
postgres=# Create Database student_db;
CREATE DATABASE
postgres=# \c student_db
You are now connected to database "student_db" as user "susantan".
```

```
student_db=# DROP TABLE if EXISTS students_performance;
DROP TABLE
student_db=# Create table students_performance (
student_db(# "gender" TEXT,
student_db(# "race/ethnicity" TEXT,
student_db(# "parental level of education" TEXT,
student_db(# "lunch" TEXT,
student_db(# "test preparation course" TEXT,
student_db(# "math score" INTEGER,
student_db(# "reading score" INTEGER,
student_db(# "writing score" INTEGER
student_db(# );
CREATE TABLE
```

```
student_db=# \copy students_performance FROM '/Users/susantan/Downloads/Students
Performance.csv' DELIMITER ',' CSV HEADER;
COPY 1000
```

Question 1: Display the first 10 students with their genders and scores

```
student_db=# select "gender", "math score", "reading score", "writing score"
student_db-# from students_performance
student_db-# limit 10;
 gender | math score | reading score | writing score
--------+------------+---------------+---------------
 female |         72 |            72 |            74
 female |         69 |            90 |            88
 female |         90 |            95 |            93
 male   |         47 |            57 |            44
 male   |         76 |            78 |            75
 female |         71 |            83 |            78
 female |         88 |            95 |            92
 male   |         40 |            43 |            39
 male   |         64 |            64 |            67
 female |         38 |            60 |            50
(10 rows)
```

Question 2: Does gender affect math, reading, and writing scores?

```
student_db=# select "gender", avg("math score") as avg_math,
avg("reading score") as avg_reading,
avg("writing score") as avg_writing
from students_performance
group by "gender";
 gender |      avg_math       |     avg_reading     |     avg_writing
--------+---------------------+---------------------+---------------------
 male   | 68.7282157676348548 | 65.4730290456431535 | 63.3112033195020747
 female | 63.6332046332046332 | 72.6081081081081081 | 72.4671814671814672
(2 rows)
```

Males score higher than females in math scores on average.
Females score higher than males in reading and writing scores on average.

Question 3: Does parental level of education affect test scores?

```
student_db=# SELECT
student_db-#     "parental level of education",
student_db-#     ROUND(AVG("math score"), 2) AS avg_math,
student_db-#     ROUND(AVG("reading score"), 2) AS avg_reading,
student_db-#     ROUND(AVG("writing score"), 2) AS avg_writing
student_db-# FROM students_performance
student_db-# GROUP BY "parental level of education"
student_db-# ORDER BY CASE "parental level of education"
student_db-#     WHEN 'some high school' THEN 1
student_db-#     WHEN 'high school' THEN 2
student_db-#     WHEN 'some college' THEN 3
student_db-#     WHEN 'associate''s degree' THEN 4
student_db-#     WHEN 'bachelor''s degree' THEN 5
student_db-#     WHEN 'master''s degree' THEN 6
student_db-#     ELSE 7
student_db-# END;
 parental level of education | avg_math | avg_reading | avg_writing
-----------------------------+----------+-------------+-------------
 some high school            |    63.50 |       66.94 |       64.89
 high school                 |    62.14 |       64.70 |       62.45
 some college                |    67.13 |       69.46 |       68.84
 associate's degree          |    67.88 |       70.93 |       69.90
 bachelor's degree           |    69.39 |       73.00 |       73.38
 master's degree             |    69.75 |       75.37 |       75.68
(6 rows)
```

Students whose parents have higher levels of education tend to achieve higher average scores across math, reading, and writing.
The gap between high school and college degree parental education is noticeable, about a 7–8 point increase in math, reading, and writing scores.

Question 4: What is the total number of students who completed the test preparation course? Broken down by gender?

```
student_db=# select "test preparation course", count(*) as num_students
from students_performance
group by "test preparation course"
order by "test preparation course" ASC;
 test preparation course | num_students
-------------------------+--------------
 completed               |          358
 none                    |          642
(2 rows)
```

```
student_db=# select "test preparation course","gender", count(*) as num_students
from students_performance
group by "test preparation course", "gender"
;
 test preparation course | gender | num_students
-------------------------+--------+--------------
 completed               | male   |          174
 completed               | female |          184
 none                    | male   |          308
 none                    | female |          334
(4 rows)
```

There are 358 students who completed the test preparation course and 642 students who did not do the test preparation course.
Of the 358 students who took the test preparation course, 174 were male and 184 were female.
Of the 642 students who did not take the test preparation course, 308 were male and 334 were female.

Question 5: Does the test preparation course affect test scores? Broken down by gender?

```
student_db=# select "test preparation course", round(avg("math score"),2) as avg_math, round(avg("reading score"),2) as avg_reading, round(avg("writing score"),2) as avg_writing
from students_performance
group by "test preparation course"
order by "test preparation course" ASC;
 test preparation course | avg_math | avg_reading | avg_writing
-------------------------+----------+-------------+-------------
 completed               |    69.70 |       73.89 |       74.42
 none                    |    64.08 |       66.53 |       64.50
(2 rows)
```

```
student_db=# select "test preparation course","gender", round(avg("math score"),2) as avg_math, round(avg("reading score"),2) as avg_reading, round(avg("writing score"),2) as avg_writing
from students_performance
group by "test preparation course", "gender"
order by "test preparation course" ASC;
 test preparation course | gender | avg_math | avg_reading | avg_writing
-------------------------+--------+----------+-------------+-------------
 completed               | male   |    72.34 |       70.21 |       69.79
 completed               | female |    67.20 |       77.38 |       78.79
 none                    | male   |    66.69 |       62.80 |       59.65
 none                    | female |    61.67 |       69.98 |       68.98
(4 rows)
```

Those who took the test preparation course scored higher in math, reading, and writing.
Out of those who took the test preparation course, males scored higher in math and females scored higher in reading and writing. The same trend appeared in the those who did not take the test preparation course.
Males who took the test preparation course scored higher in all 3 subjects than their male counterparts who did not take the course, and likewise for females.