Practical 1: Analyse the Exam Data

This dataset contains math, reading and writing scores based on students' gender, race, level of parent education, whether they receive a free or reduced-cost lunch, and whether they completed a test preparation course.

For every question, create a markdown block for the question number (Question 1, Question 2 and so on).

Directly underneath the markdown for a specific question, create a Python code block and enter the code for your answer into the code block.

Answer all questions with appropriate code.

The dataset name is `exams.csv` and it is provided as a download on the learn site.

Questions

1. Create a folder named **PRT2030_Practical_1** in a location this is fully under your control.
2. Copy the dataset (**exams.csv**) into the folder.
3. Open the folder in your IDE of choice (either JupyterLabs or VS Code), and create a new notebook file inside the folder named **Practical_1.ipynb**.
4. Read the data from the CSV file into a DataFrame named **df_exams** and display the first five rows.
5. Display the basic information for the DataFrame and its columns using the **info()** method.
6. Display statistical information for the math score, reading score, and writing score columns using the **describe()** method.
7. Group the data by the race/ethnicity column and display the mean scores.
8. Display the **gender** column as a DataFrame with bracket notation.
9. Display the **gender** column as a Series with bracket notation.
10. Display the **gender** column as a Series with dot notation.
11. Display only rows for female students with a math score greater than or equal to 90.
12. Does taking a test preparation course improve average scores?
13. Which gender is better on average at math?
14. Which gender is better on average at all three subjects? Hint: Start by adding a column to the DataFrame with the total score.
15. Does the parents' level of education influence the average scores?

-END-