Python for data scientists SP1 2021/2022 Assignment 8 (Individual Assignment!)

There are several tasks in this assignment.

You can either submit:
- One Python file (solutionA8.py)
- A PDF file for the explanation part (if any).

Or a Jupyter Notebook with clearly marked answers for which tasks and any explanation.

During the pandemic, Zoom and other online tools have been used a lot as a medium for delivering education. In some educational activities, the teachers ask questions via Zoom polls. An example result of such a Zoom poll can be seen in one of the given files (Poll_Report.csv).

Considering that many teachers at school have no experience of analysing such poll results, the teachers have asked the data science students at GU to help.

The teacher supplies the following files:

- "Poll_Report.csv": the file consists of poll results of a particular test at school time.
- "Poll_Info.csv": the file consists of information about the correct answer for each of the questions asked at the test and the time when each question was launched by the teacher.

Your tasks:

1) Analyse how many questions are correctly answered by each student. Show the result with appropriate visualisation.
2) Analyse how many students answered correctly for each question. Show the result with appropriate visualisation.
3) To get some ideas where students need reinforcement, the teacher wants to know:
    a. For each student, which question(s) took the longest time for that student?
    b. Which question(s) took the longest time to answer among all student (that is, no student took longer to give a correct answer to this question or to any other question)?
    c. Which question(s) had the longest average response time?

    Do analysis to help the teacher get insights regarding this. Visualise the analysis results.

4) Investigate whether students who took a longer time to answer perform well or not.
5) Find all pairs of students who gave the same wrong answers for at least two questions.
6) Are there more interesting information one can derive from studying how long each student response to each question that might be of interest to the teacher? Convey your thoughts by doing analysis and showing the results.

Note: If you add any details or make any assumptions, please clearly describe those in your submission.