

University of Toronto

Exploratory Analysis of the Course CSC258 Quercus Data

A report derived from the Data-Driven Design: Quercus Analytics (D3:QA) Project

Written by: Alan da Silveira Fleck

Course: CSC258H5 - Computer Organization

Instructor: Prof. Pooja Vashisth

Toronto, September 6th, 2022

Background and Objectives

CSC258 is an introduction to computer organization and architecture. In this course, the students write assembly and examine the design of a simple RISC architecture (RISC-V). Core topics include data representations and computer arithmetic, processor organization, the memory hierarchy and caching, instruction set and addressing modes, and quantitative performance evaluation of computing systems. Each week, recommended readings and a quiz are posted. The quizzes are online, on Quercus, and are intended to guide the reading for the term tests and final exam. Students are allowed multiple attempts until they are satisfied with their final scores.

This analysis aimed to assist with spotting concepts which students were struggling with mid-course or earlier, giving instructors a proactive chance to provide supplemental resources to improve their chances of success. Specifically, analyses included when students were accessing the reading material and attempting the quizzes, identifying specific assignments and questions that the students were struggling with, the patterns of interaction with the quizzes, and the relationship between quizzes and exam performances.

Data

Information on the number of views and timing of access to the reading materials was collected from the New Analytics' Course Activity Report on the course's Quercus page. The data were filtered by the resources and dates relative to the weekly quizzes. The scores, number of attempts, and date and time of submission of the weekly quizzes were extracted from the Quiz Reports on the course's Quercus page.

Main Findings

Figure 1 shows that **every week the number of views of the reading material started to increase on Monday and has a peak on Wednesday, with an occasional extension of this peak through Thursday (the class date and deadline for the quizzes)**. The average number of students accessing the reading material per day was 41. The three larger peaks on February 15th, March 22nd, and April 22nd corresponded to the dates of the first and second term tests, and the final exam, respectively.

Number of Students Accessing the Reading Material by Date

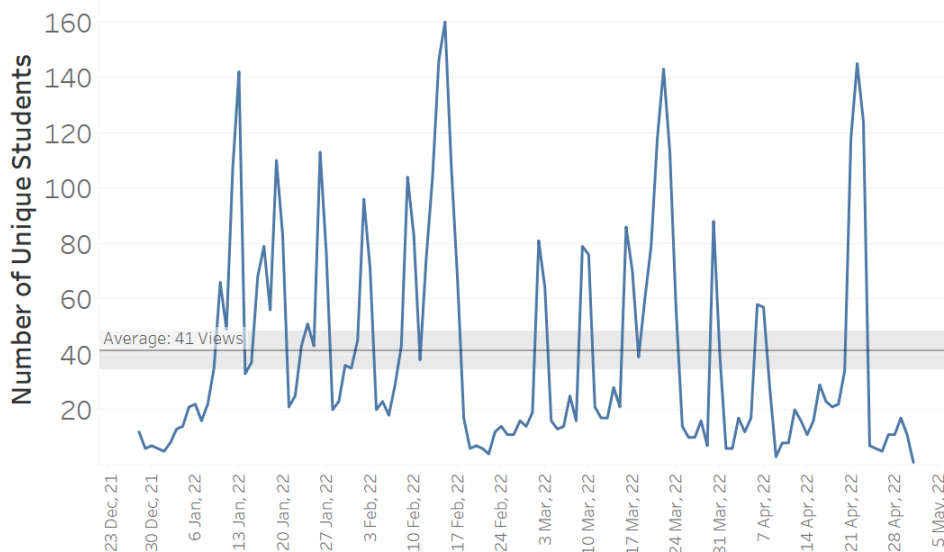


Figure 1: Number of Students Accessing the Reading Material by Date.

After identifying when the resources were being accessed, the next analysis investigated the percentage of students that accessed the reading material before each quiz deadline. In this regard, Figure 2 shows **an increase in the number of students doing the quizzes without reading the material** when comparing the percentage of students accessing the reading material at the beginning of the course (**Quiz 1: 13.5% not viewing the material**) versus the end of the course (**Quiz 12: 36.2% not viewing the material**).

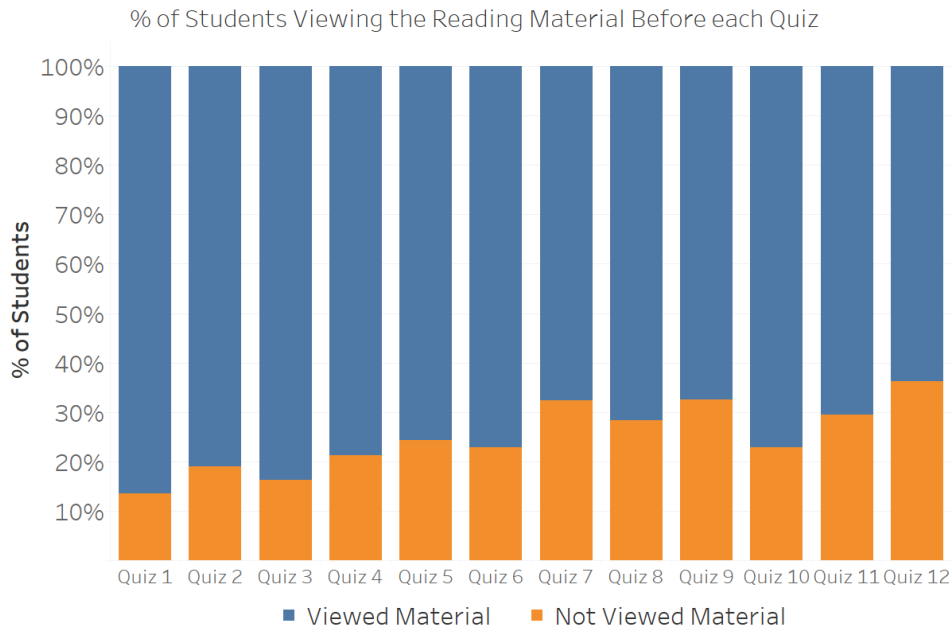


Figure 2: % of Students Viewing the Reading Material Before Each Quiz.

After finding this decrease in engagement with the quiz reading materials, we investigated the relationship between the average number of attempts per quiz and their average final score. The rationale for this analysis was that by searching for relationships between the number of attempts and the quiz score, we would identify quizzes, questions and concepts that students were struggling with the most.

Figure 3 shows that **the average number of attempts** (i.e., dark blue bars) **ranged from 4.3 in Quiz 1 to 16.8 in Quiz 3. The average number of attempts for all quizzes combined was 10.9.** In addition, in this figure, we can also see that the **Quizzes with a high average number of attempts seemed to have a lower average percentage score** (i.e., light blue lines). This last point was further investigated in the scatterplot between the average attempts and scores by each quiz. In this regard, Figure 4 shows that **there was a negative and statistically significant relationship between the Average Number of Attempts and the Quiz Percentage Score (p-value = 0.008).** In other words, quizzes with a higher number of attempts tended to have a lower average percentage score. Examples are **Quiz 3, Quiz 4 and Quiz 7 for a higher number of attempts and lower average scores**, and **Quiz 1, Quiz 2 and Quiz 11 for a lower number of attempts and higher average scores.** Inside each quiz, the questions with the lowest average scores were identified and flagged to the instructors (data not shown).

Average Number of Attempts and Percentage Score per Quiz

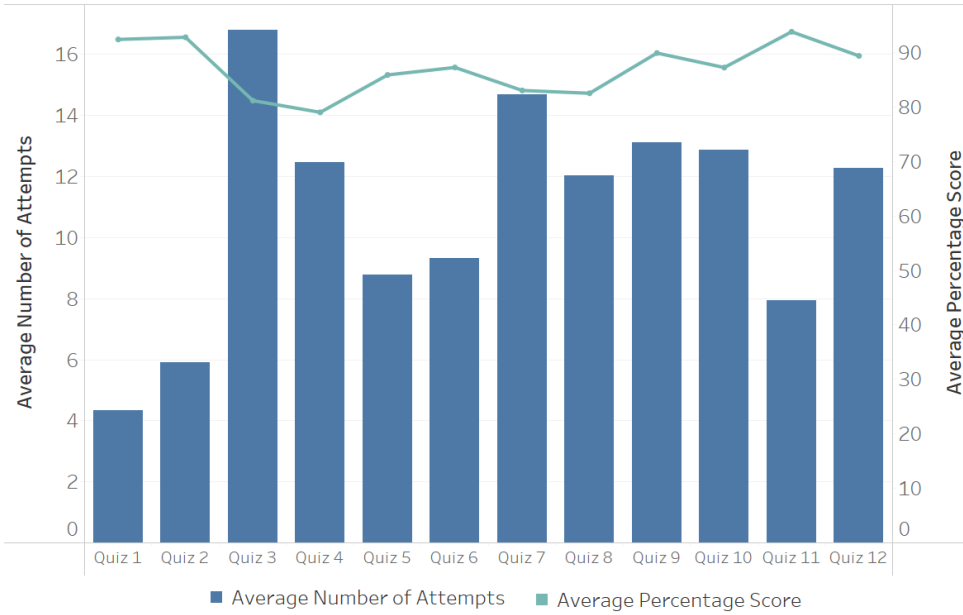


Figure 3: Average Number of Attempts and Percentage Score per Quiz

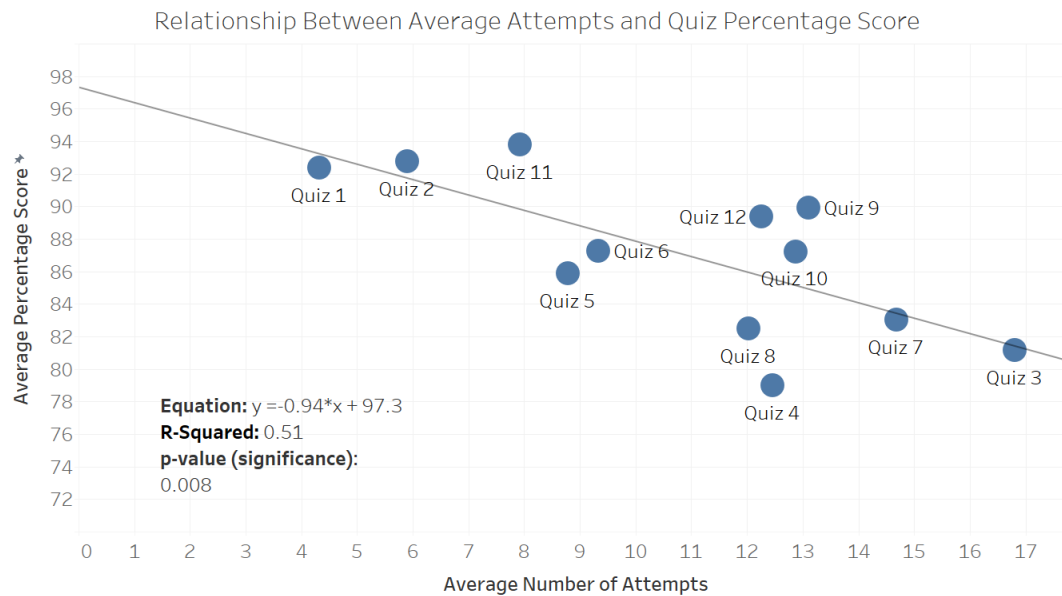


Figure 4: Relationship Between Average Attempts and Quiz Percentage Score

After finding this relationship between the average number of attempts and quiz scores, the next analysis comprehended the investigation of the patterns of interaction with the quiz attempts. In this regard, **for every Quiz, more than 50% of the attempts occurred on Thursdays, the same day as the deadline.** The other attempts showed a left-skewed distribution, with attempts becoming more frequent as the deadline approaches. Figure 5 illustrates these patterns for Quiz 8. The distribution shown in this quiz is representative of the pattern observed for the other quizzes.

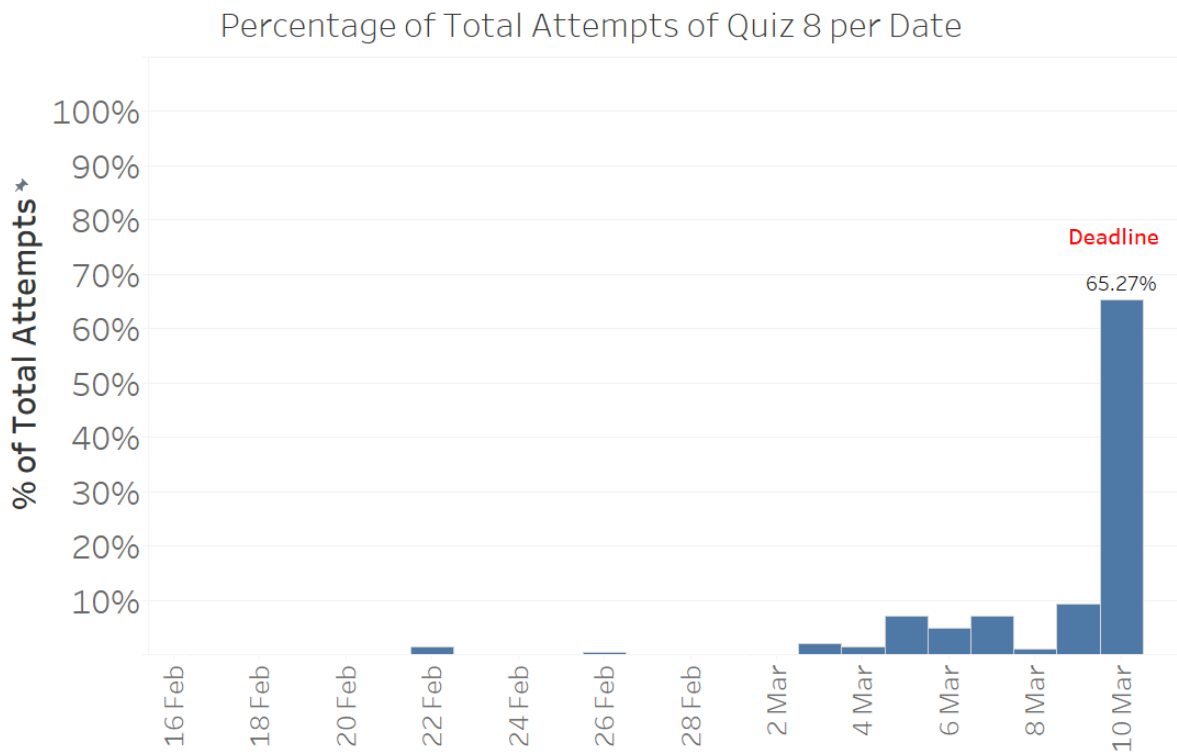


Figure 5: Percentage of Total Attempts of Quiz 8 per Date.

A follow-up investigation of this high number of attempts on the deadline date shows that **around 50% of all attempts had a duration of 1 minute or less (Figure 6)**. In addition, **after Quiz 3, most of the attempts included answers to only one question or two (Figure 7), and they were focused on specific question types (Figure 8)**. These question types allowed for multiple possible answers, which resulted in several repeated attempts until the students “brute forced” the correct one. Together, these findings suggest that students are mostly not reviewing the reading material after each attempt, and they might be using this quiz format to iterate through each question individually until the correct answer is found.

Histogram of Time Between Quiz Attempts

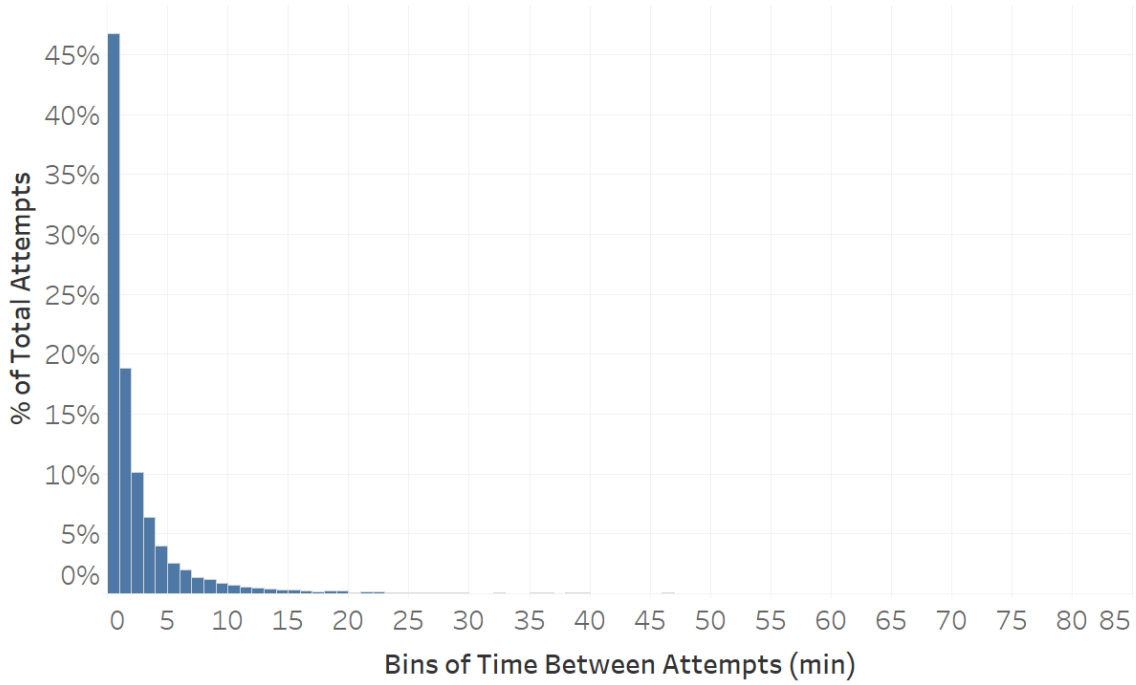


Figure 6: Time Between Each Quiz Attempt.

Total Questions Answered by Quiz (All Attempts)

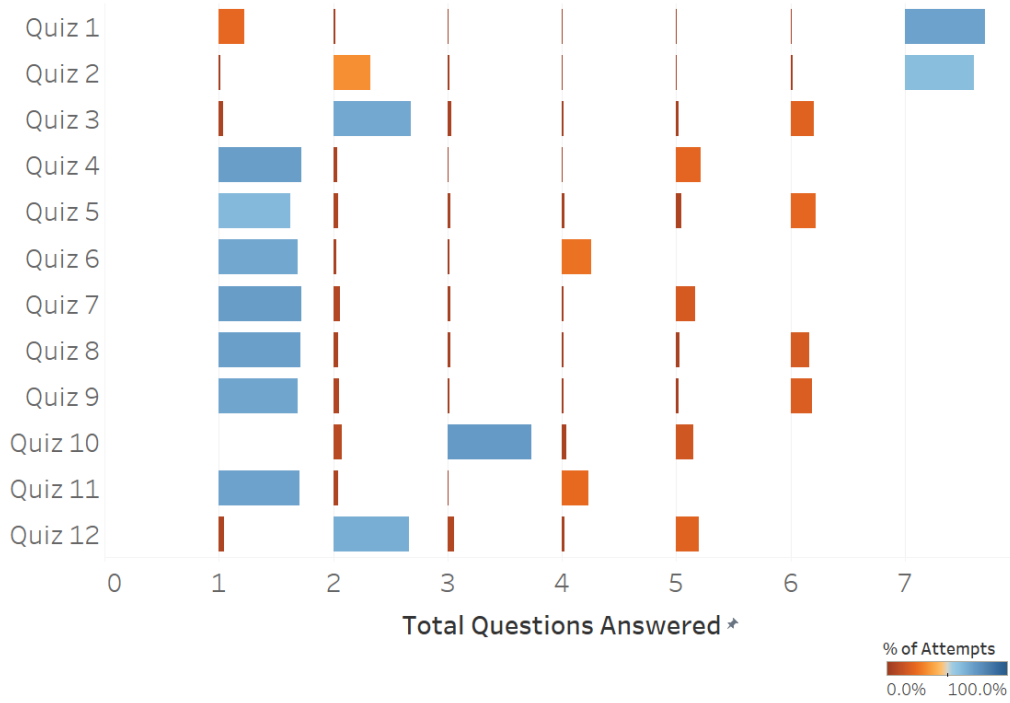


Figure 7: % of Total Number of Questions Answered by Quiz.

Frequency of Answering a Question by Quiz

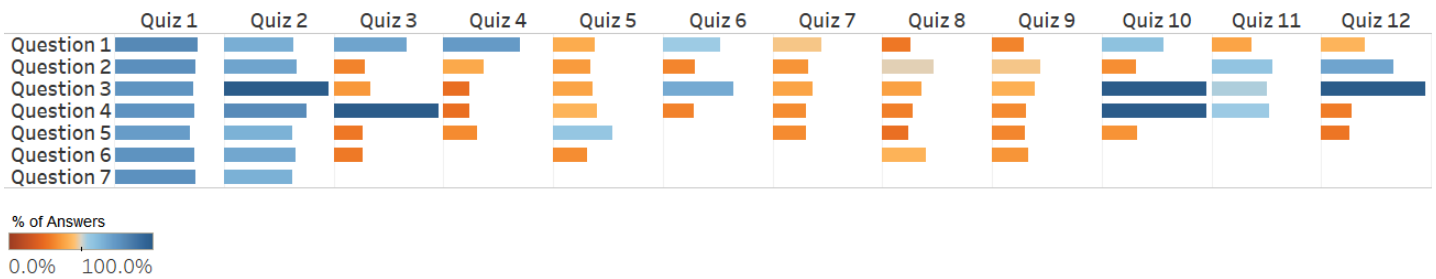


Figure 8: Frequency of Answering a Question by Quiz.

Finally, we analyzed the relationship between quizzes and exam performances. **Figure 9** shows that there were **positive and statistically significant** relationships between the **percentage score in quizzes 1 to 4 with the Test 1 score** (Linear regression: $R^2 = 0.12$; $P < 0.0001$), and the **normalized scores in quizzes 1 to 4 with Test 1 score** (Linear regression: $R^2 = 0.04$; $P = 0.003$). Furthermore, **Figure 10** shows that there was a **positive and statistically significant relationship between the percentage score in quizzes 5 to 8 and the Test 2 performance** (Linear regression: $R^2 = 0.13$; $P < 0.0001$). However, **no association was found between the normalized quiz scores and Test 2 score** (Linear regression: $R^2 = 0.01$; $P = 0.12$). Finally, **Figure 11** indicates **positive and statistically significant** relationships between the **percentage score in all quizzes with the Final Exam performance** (Linear regression: $R^2 = 0.17$; $P < 0.0001$), as well as between the **normalized scores in all quizzes with the Final Exam score** (Linear regression: $R^2 = 0.02$; $P = 0.02$).

Relationship Quizzes 1 to 4 Score and Test 1 Score

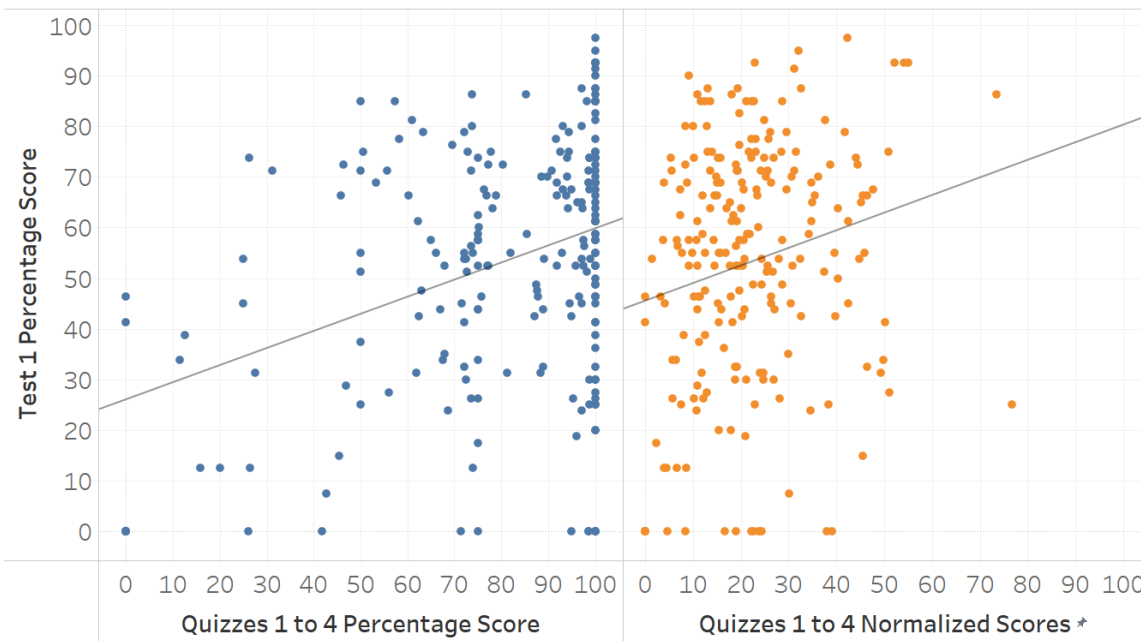


Figure 9: Relationship Between Quizzes 1 to 4 Scores and Test 1 Score. Note: Normalized score means percentage score divided by the number of attempts.

Relationship Quizzes 5 to 8 Score and Test 2 Score

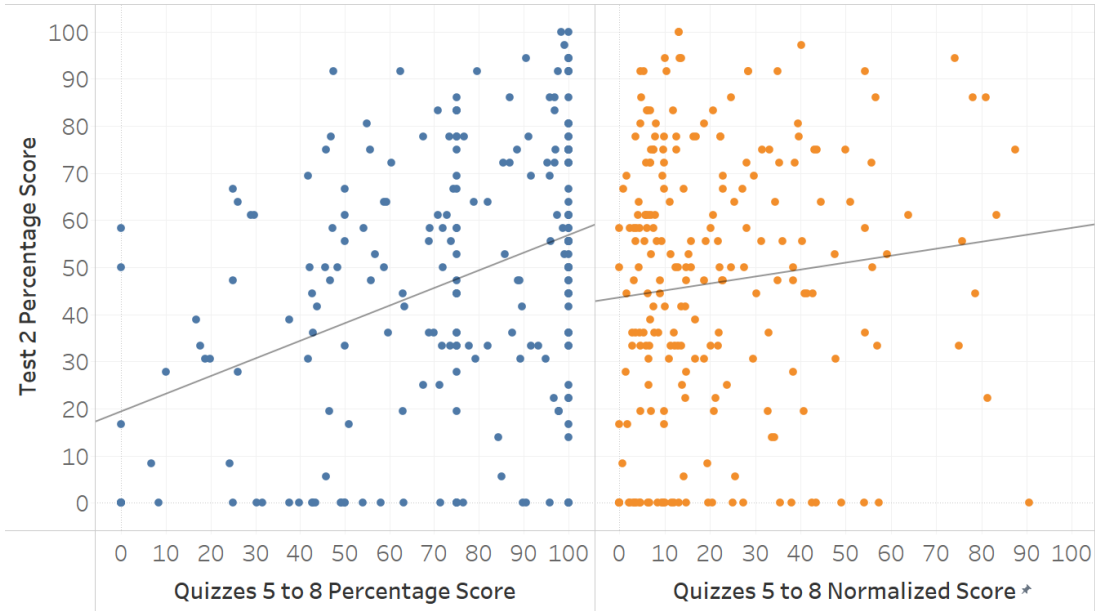


Figure 10: Relationship Between Quizzes 5 to 8 Scores and Test 2 Score. Note: Normalized score means percentage score divided by the number of attempts.

Relationship All Quizzes Average Score and Final Exam Score

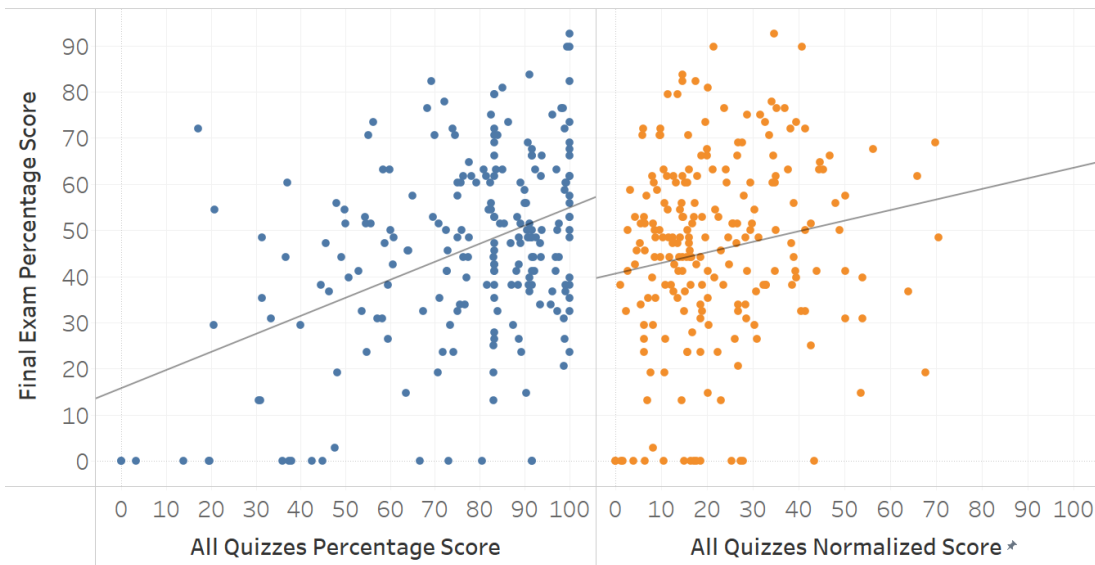


Figure 11: Relationship Between all Quizzes Score and the Final Exam Score. Note: Normalized score means percentage score divided by the number of attempts.

Recommendations

With the identification of the quizzes and questions that the students may be struggling with the most, the instructor can identify key concepts related to these quizzes and review them before the term tests and final

exam. This would help to identify the areas where students need help with the reading material and could be used as an indicator for the instructors to provide more support to the students by improving readings.

However, potential problems may not be always linked to a concept not well understood. It was identified by analyzing the students' attempt behaviour that the format of some quiz questions was being exploited for correct answers by applying the brute force approach. This was true as reading quizzes allowed unlimited attempts with feedback. In this regard, future questions should be developed in a format that minimizes the effect of random guessing.

Finally, the overall positive and significant relationships between quizzes and exam performances suggest that the weekly quizzes could be a positive preparatory material for the exams even after considering the issue of unlimited attempts. These results could be shared with future classes to stimulate engagement with the weekly quizzes and the related preparatory readings.