University of Toronto

Exploratory Analysis of the Course GGR278 Quercus Data

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

Written by: Alan da Silveira Fleck

Course: GGR278H5S "Geographic Information Systems"

Instructor: Prof. Tingting Zhu

Toronto, September 6th, 2022

Background and Objectives

The course GGR278H5S "Geographic Information Systems" introduces students to models of representation and management of geographical data for scientific analysis, basic quantitative methods, and techniques for geographic data analysis, including collection, manipulation, description and interpretation. The course contains practical exercises using GIS and statistical software packages with examples drawn from both physical and human geography. As one of the methods of evaluation, students answered five biweekly quizzes distributed via Quercus that tested their understanding of the lecture content in multiple-choice, short-answer, and long-answer formats. In addition, students were encouraged to engage in discussion board activities in preparation for these quizzes.

The objective of this project was to investigate whether students used the resources provided in the discussion boards and if engagement with this material (captured by the number of discussion post views and number of participations in the discussion board) impacted the performance in the bi-weekly quizzes. Furthermore, we also aimed to evaluate whether email intervention effectively changes student engagement behaviour.

Data

Information on the number of views and timing of access to the discussion posts was collected from the New Analytics' Course Activity Report on the course's Quercus page. The data were filtered by the relevant course resources and the dates relative to the bi-weekly quizzes. The scores of the biweekly quizzes were extracted from the Quiz Report on the course's Quercus page.

Main Findings

The first set of analyses explored the level of student participation and views on the discussion boards related to each quiz and the median score for these quizzes. Figure 1 shows **higher levels of participation on the discussion board related to Quiz 1 compared to the discussion board of other Quizzes (P < 0.001).** The median participation on the discussion boards related to Quizzes 2, 3, 4 and 5 was zero.
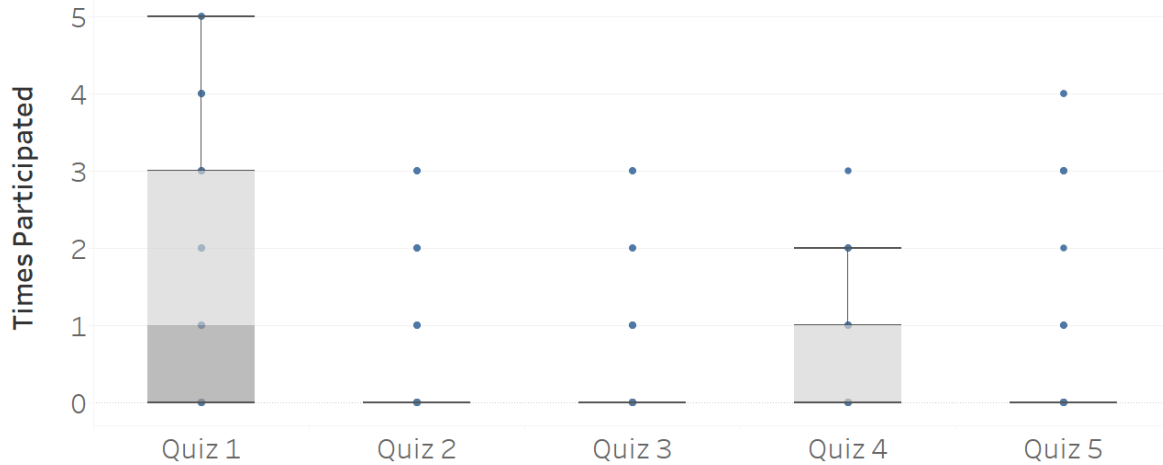
Figure 1: Box Plots of the Participation on the Discussion Boards Related to Each Quiz.

The discussion board related to Quiz 1 also had the highest median views: 2 (Interquartile Range (IQR: 0 – 4) (Figure 2). The number of views of the discussion board posts decreases with each quiz, with the higher median views of the discussion board found in Quiz 1 compared to the other quizzes (P = 0.02) (Figure 2).
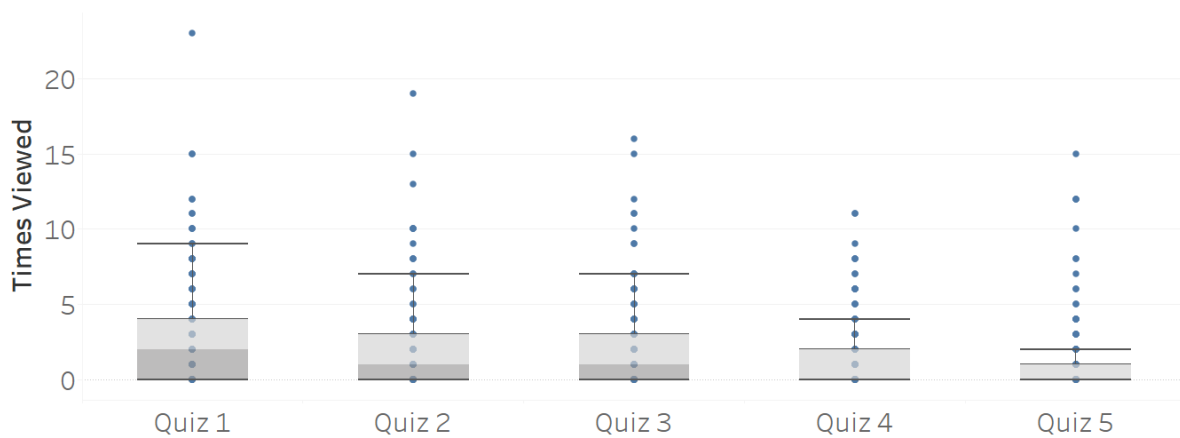
Figure 2: Box Plots of the Views on the Discussion Boards Related to Each Quiz.

Finally, the median (IQR) score for Quiz 4 was 18.5 (16.5 – 19.5). This value was statistically higher than for all other quizzes (P= 0.002) (Figure 3). Quiz 3, on the other hand, had the lowest median score: 12 (11 – 15).

Figure 3: Box Plots of Quiz Scores.
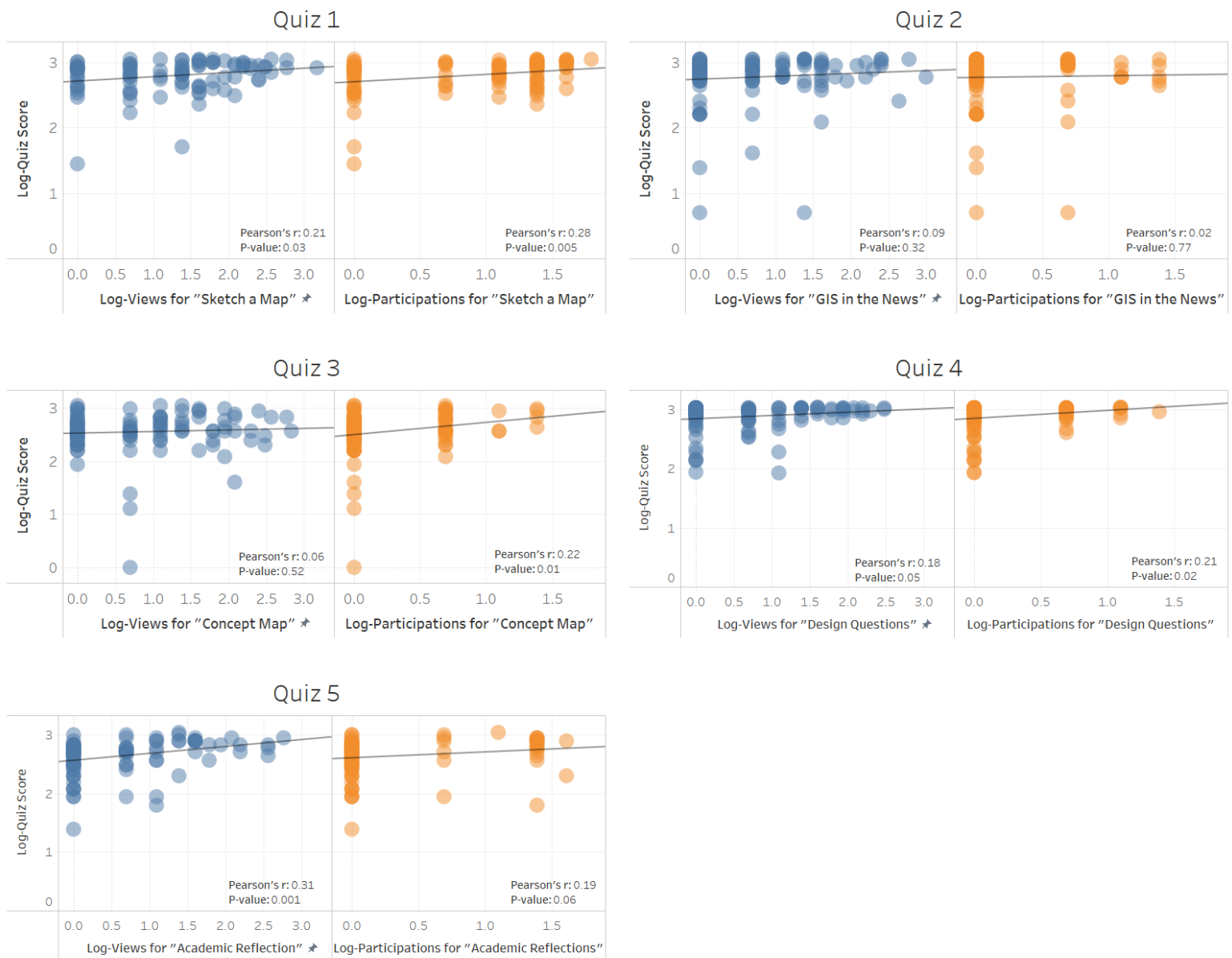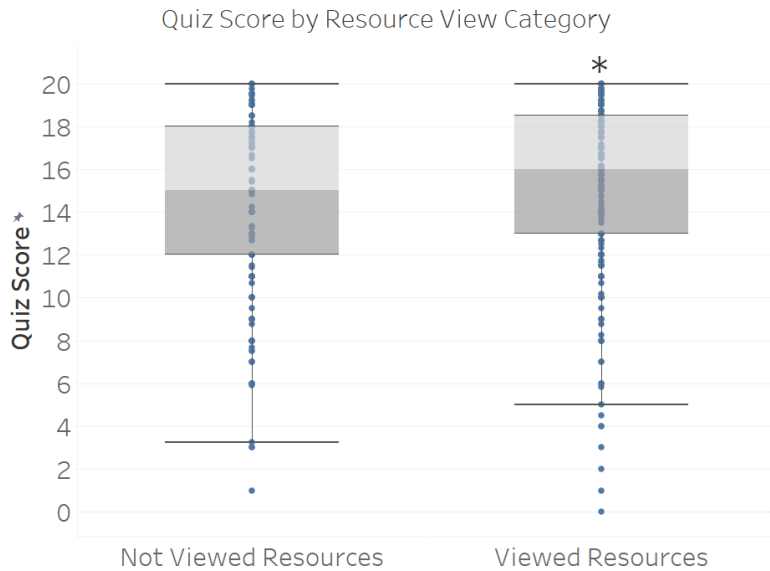
Next, the relationship between the views and participation in the discussion boards with the quiz scores was assessed for each quiz separately. Scatterplots were produced and associations were measured by Pearson correlation on the log-transformed data. There were **positive and significant correlations between views (in blue) and quiz score for Quiz 1 (Pearson's r = 0.21; P-value = 0.03), Quiz 4 (Pearson's r = 0.18; P-value = 0.05) and Quiz 5 (Pearson's r = 0.31; P-value = 0.001)** (Figure 4). In addition, **we found positive and significant correlations between participation (in orange) and quiz scores for Quiz 1 (Pearson's r = 0.28; P-value = 0.005), Quiz 3 (Pearson's r = 0.22; P-value = 0.01) and Quiz 4 (Pearson's r = 0.21; P-value = 0.02). In addition, Quiz 5 showed a tendency to a significant association (Pearson's r = 0.19; P-value = 0.06)** (Figure 4).

Figure 4: Scatterplots and Pearson correlations between the views (in blue) and participation (in orange) with the quiz scores.

A more detailed investigation of the relationship between engagement in the discussion board and quiz scores was conducted by separating views and participation into categories based on the distribution of these variables and by assessing the relationship between these variables and quiz scores.

Figure 5 shows that **students who accessed the discussion boards before each quiz had a statistically higher quiz score (Median (IQR) = 16 (13-18.5)) compared to the ones that did not access this resource (Median (IQR) = 15 (12 − 18)) (P = 0.003).** In addition, Figure 6 shows that, when dividing viewing status according to the number of discussion board accesses, **multiple views of the resources (i.e., at least 3 times) were associated with higher median quiz scores compared to fewer total views (i.e., 0, 1 or 2) (P < 0.05).**

Quiz Score by Resource View Category

Figure 5: Box Plots of Quiz Score by Resource View Category (Viewed vs Not Viewed; All Quizzes Combined).
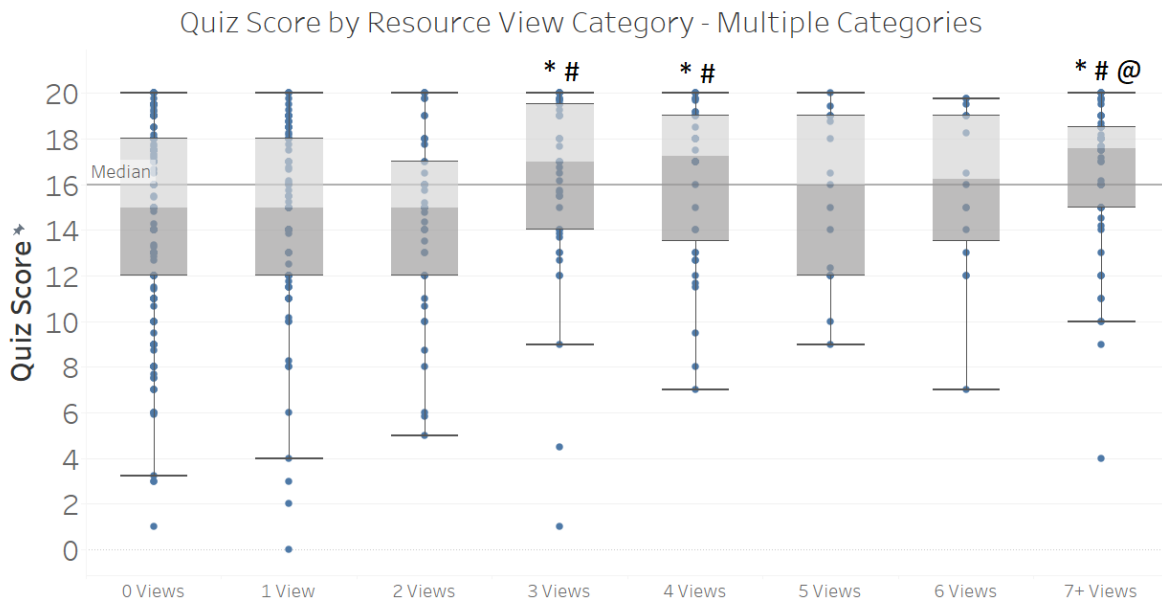
Quiz Score by Resource View Category - Multiple Categories

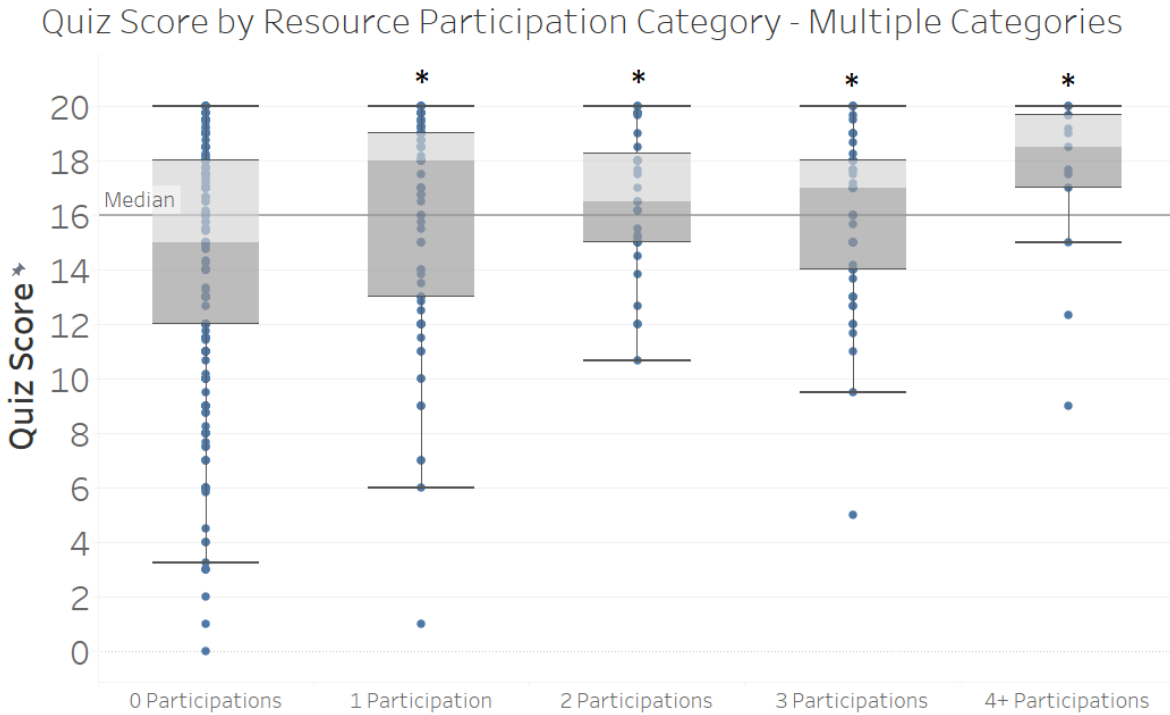**Figure 6:** Box Plots of Quiz Score by Resource View Category (Multiple Categories; All Quizzes Combined).

Similar analyses were also conducted for the participation variable. Figure 7 shows that **students who participated in the discussion boards before each quiz had a statistically higher quiz score (Median (IQR) = 17.2 (12 – 19)) compared to the ones that did not post in this resource. Median (IQR) = 15 (12-**

18) (P < 0.001) (Figure 7). In addition, **participation at all levels (i.e., 1, 2, 3 or 4+ participations) are all associated with a better quiz score compared to 0 participation (P < 0.05)** (Figure 8). However, the increase in the number of participants does not seem to result in increased scores since there was no statistical difference in the scores of the students that participated 1, 2, 3 or 4+ times. This suggests that participation at any level (> 0) should be stimulated.

Quiz Score by Resource Participation Category
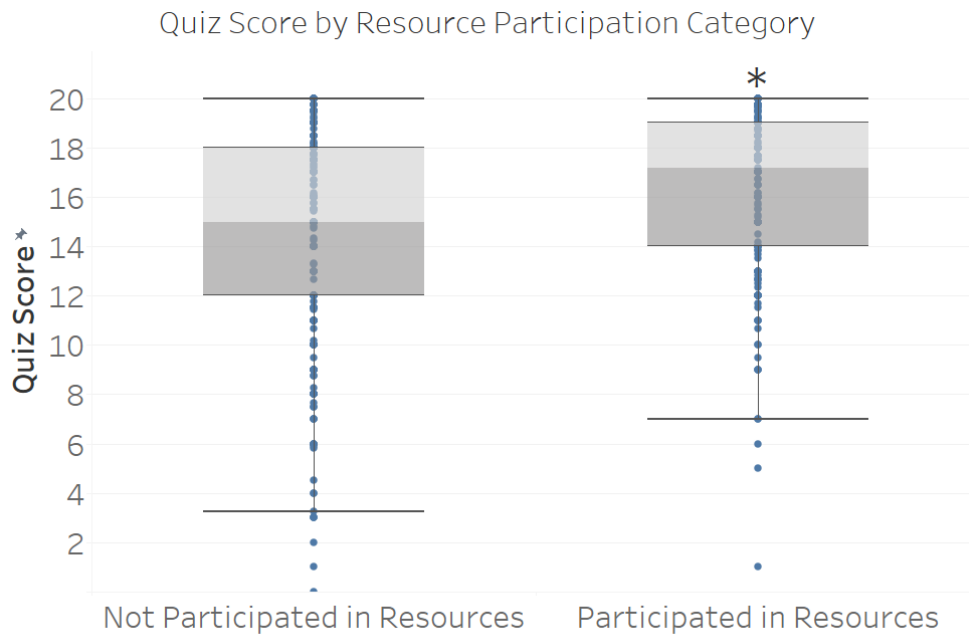
* P < 0.001. Mann Whitney's Test

Figure 7: Box Plots of Quiz Score by Participation Category (Participated versus not participated; All Quizzes Combined).

Figure 8: Box Plots of Quiz Score by Participation Category (Multiple Categories; All Quizzes Combined).

After finding a positive relationship between page views and student performance, the next step included the investigation of how to encourage students to access the resources provided on Quercus. For these analyses, an approach with customized emails was chosen.

Figure 9 shows the date of first access to Quercus. More than 30% of the students accessed Quercus for the first time on the first day the course page was available in the system. **In addition, on January 10, 2022, a message was sent to the students instructing the ones who had not visited the page yet to do so. This message coincides with the second peak of first access to Quercus around January 10 and 11, as seen in the figure.**
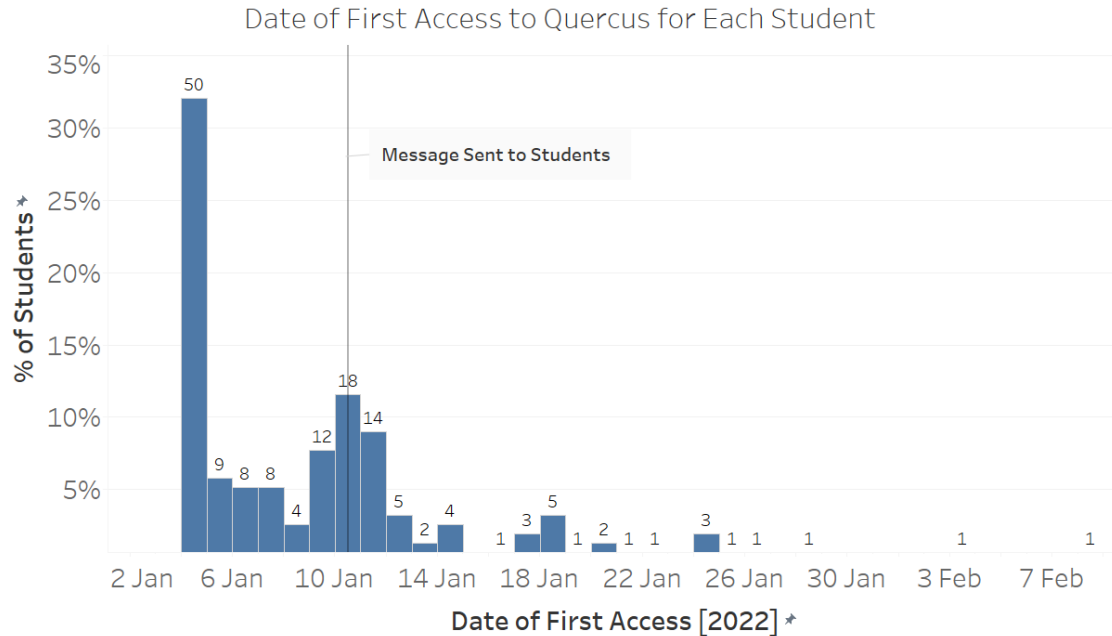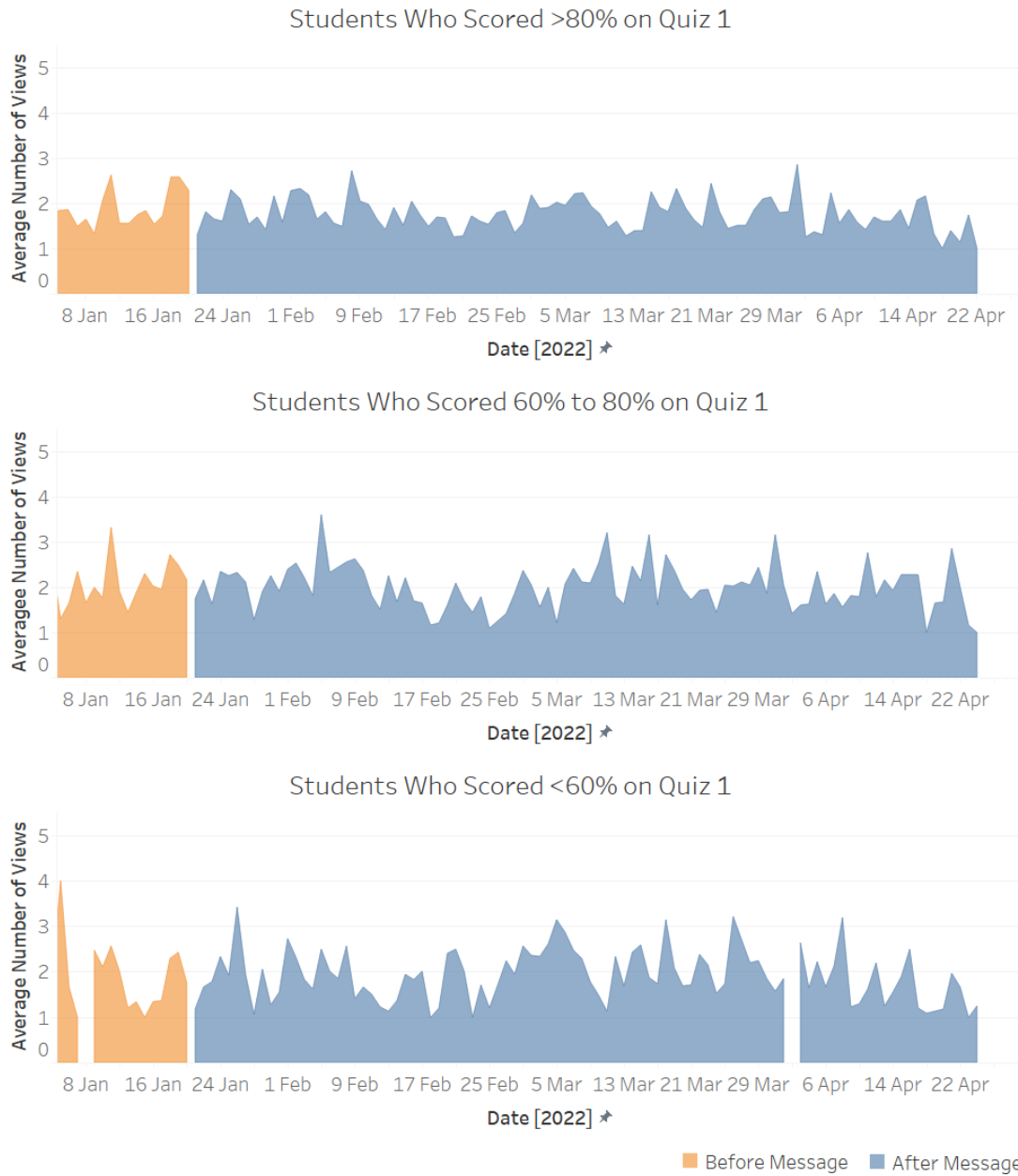
Figure 9: Date of First Access to Quercus and Effectiveness of Message to Students.

The second type of intervention was conducted on January 20th, 2022. On this date, a customized email was sent to students according to their scores on Quiz 1. The average number of resource views per day was monitored after this message was sent. For students with scores < 60%, the message included the following learning suggestions: spacing out learning, note-taking resources, encouragement of active participation, an invitation for office hours, posting questions on the discussion board, and suggestions for improved time management. For students with scores > 80%, the message included a congratulation for their performance on the quiz, and an encouragement to keep up the good work. As shown in Figure 10, **there was no significant change in the average number of resource views per day after the message regarding students' scores in Quiz 1 was sent.**
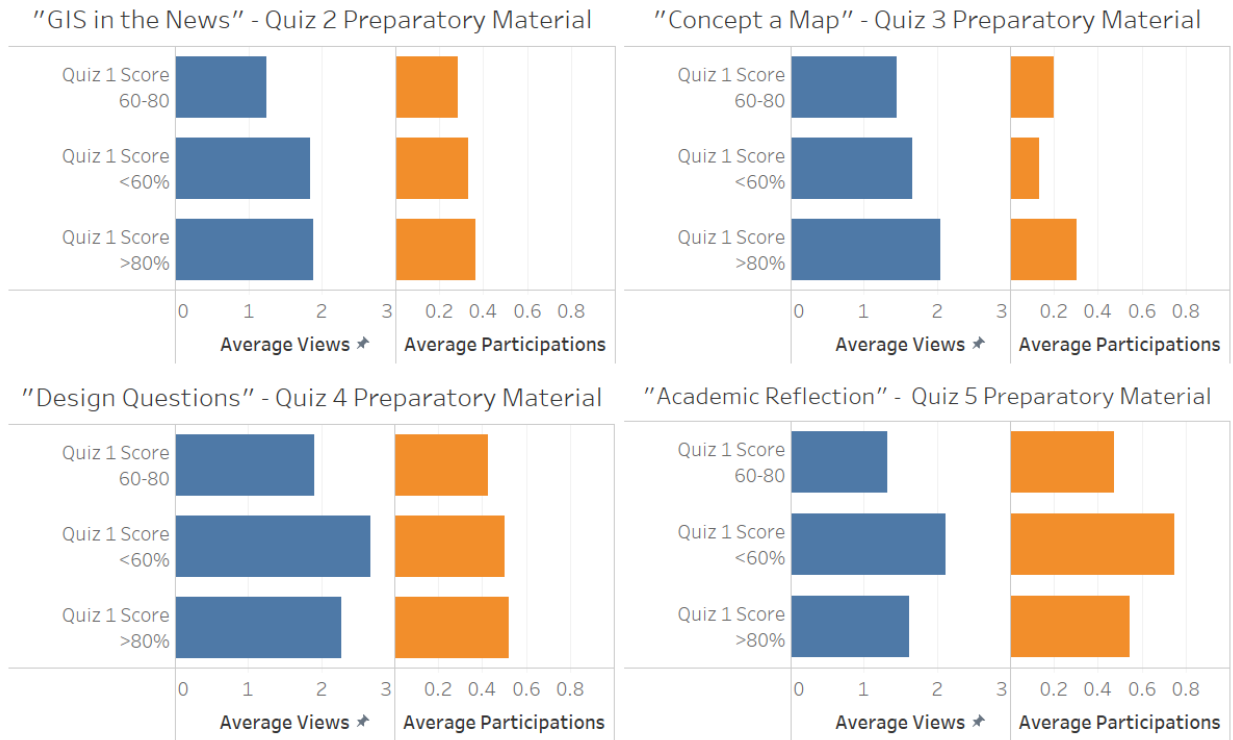
However, Figure 11 illustrates that **students who had the highest scores in Quiz 1 also had the highest average participation in the materials related to Quiz 2, 3 and 4**. On the other hand, **students with the worst performance in Quiz 1 showed the highest average views and participation in the preparatory material of Quiz 5, the last quiz of the course.** Finally, **students who did not receive a customized email (i.e., Quiz 1 score between 60% and 80%) had their views and participation remained low** (Figure 11).

Figure 10: Average Daily Number of Resource Views Before and After Message to Students Regarding Quiz 1 Score.

Figure 11: Average Views and Participations in the Quiz 2 to 5 Preparatory Material by Score-Category in Quiz 1.

Conclusions and Next Steps

In conclusion, this project demonstrated that multiple views of the relevant resources (i.e., at least 3 times) are associated with higher median quiz scores compared to fewer total views (i.e., 0, 1 or 2). Participations at all levels (i.e., 1, 2, 3 or 4+ participations) are all associated with a better quiz score compared to 0 participations. In addition, for three out of five quizzes, there were positive and significant correlations between views and quiz scores, as well as positive and significant correlations between participation and quiz scores. However, the correlations were weak, suggesting that other factors that may also contribute to students' performance were not investigated. Finally, customized emails did not change students' resource views. However, students who did not receive the customized emails remained low in levels of participation and views for the rest of the modules.

Future work could include the creation of a prediction model based on students' interaction with the discussion board and the exploration of additional engagement metrics. In addition, different intervention strategies to change student behaviours could be explored.