The Impact of COVID-19 on Students’ Marks: A Bayesian Hierarchical Modeling Approach

Jabed Tomal · Saeed Rahmati · Shirin Boroushaki · Lingling Jin · Ehsan Ahmed

Received: date / Accepted: date

Abstract Due to COVID-19, universities across Canada were forced to undergo a transition from classroom-based face-to-face learning and invigilated assessments to online-based learning and non-invigilated assessments. This study attempts to empirically measure the impacts of COVID-19 on students’ marks from eleven science, technology, engineering, and mathematics (STEM) courses using a Bayesian linear mixed effects model fitted to longitudinal data. The Bayesian linear mixed effects model is designed for this application which allows student-specific error variances to vary. The novel Bayesian missing value imputation method is flexible which seamlessly generates missing values given complete data. We observed an increase in overall average marks for the courses requiring lower-level cognitive
skills according to Bloom’s Taxonomy and a decrease in marks for the courses requiring higher-level cognitive skills, where larger changes in marks were observed for the underachieving students. About half of the disengaged students who did not participate in any course assessments after the transition to online delivery were in special support.

**Keywords** COVID-19, Science Education, Hierarchical Linear Modeling, Missing value estimation

### 1 Introduction

The spread of the novel coronavirus disease 2019 (COVID-19) started in December 2019 in Wuhan, China [25]. Due to the rising health concerns, many universities around the world transitioned from face-to-face to online course delivery and assessments [18]. In the province of British Columbia (BC), Canada, all on-campus classes and activities in the post-secondary schools were cancelled starting March 15, 2020. At Thompson Rivers University (TRU), the face-to-face classes shifted to alternative modes of delivery mostly conducted through the online learning management systems. In order to adapt to the changes, most, if not all, of remaining course assessments were switched to open-book non-invigilated exams with adjusted weights for the evaluation components. In this study, we proposed a Bayesian statistical model to evaluate the effects of the sudden change from classroom-based to online delivery and assessments in response to COVID-19 pandemic on students’ academic performance. It is important to understand comprehensively how students’ performance has changed after March 15, 2020 due to the COVID-19 effect and to investigate factors that potentially contributed to the changes.

Due to the unique situation of COVID-19 and its social and educational ramifications, there have been a few published articles discussing how COVID-19 and the shutdown of universities have impacted students’ performance in post-secondary education. For instance, Sintema [20] investigated the possible impacts that the closure of secondary schools due to COVID-19 in Zambia would have on the general performance of students in specific subject areas. Having interviewed STEM educators at a public secondary school in Zambia, the study concluded that learners’ performance would be negatively affected in STEM subjects in the upcoming national examination if the COVID-19 epidemic is not being lessened in the shortest possible time. This is because of not only the lack of contact and meaningful interactions between learners and teachers, but also insufficient e-learning tools to facilitate such interactions. Basilaia and Kvavadze [3] studied the capacities of the country Georgia and its population to continue the education process at the schools in the online form of distance learning. The study was conducted in a private school during the COVID-19 pandemic which reviewed different available platforms that were used with the support of the government for online education and live communications. Their findings showed that the transition from traditional to online education system was successful in terms of adaptability and

---

1 STEM is a curriculum based on the idea of educating students in four specific disciplines — Science, Technology, Engineering and Mathematics — in an interdisciplinary and applied approach.
gained skills by students, teachers and administrative staff. Nevertheless, none of these studies compared students’ performances empirically.

During the Severe Acute Respiratory Syndrome (SARS) epidemic in Hong Kong, when all schools and universities were ordered closed and governments invoked quarantine laws to isolate those who might be the carriers, Wong [23] conducted a study which briefly described the impact of e-commerce on the local community with an emphasis on the use of e-learning technology as a contingency measure in tertiary institutions. Their study showed that, given limited time available for the course design and delivery, the examination result of the e-learning class was slightly better than the traditional class. However, their study lacks the rigorous empirical comparisons of students’ marks using a sound statistical model. They suggested further rigorous study.

There are some studies evaluating how students perform in online versus face-to-face delivery in post-secondary mathematics education. Jones and Vena [10] focused on equity in learning as reflected in the final grades of online and on-site students from the same post-secondary mathematics course taught repeatedly over 10 semesters. On-site students attended regular class sessions, while online students only attended an orientation session and a final exam. For both groups, the evaluations were invigilated. Their findings revealed significant statistical differences in online and on-site students’ final grades, in favor of on-site student achievement. Rey [17] examined the associations between taking basic skills mathematics courses online versus face-to-face and student success and persistence. The study stressed on the difficulties associated with effective communication of mathematics topics and ideas via the Internet and found no noticeable associations with learning outcomes or persistence. Their study also pointed out that the quality of education gained from online basic skills mathematics courses is relatively equivalent to face-to-face courses. Weems [21] compared two sections of beginning algebra course where one was taught online and the other was taught on site. The study reported no significant difference for exam average between the two formats, but highlighted a decrease in performance by the online students across the exams. Ashby et al [2] compared student success in a Developmental Math course offered in three different learning environments: online, blended, and face-to-face. Using a one way analysis of variance (ANOVA), the authors showed that there were significant differences between learning environments with the students in the blended courses having the least success. With regards to the efficiency of learning outcomes in the online pedagogy, Arias et al [1] studied the effectiveness of online delivery relative to face-to-face delivery by randomly assigning students into two categories: online and face-to-face. These two sections were taught by the same instructor and the course objectives and exams were the same for both sections as well. The authors concluded that “both course objectives and the mechanism used to assess the relative effectiveness of the two modes of education may play an important role in determining the relative effectiveness of alternative delivery methods.” One other area of interest in this respect is the notion of exams and their competence in the digital space. In a study by Williams and Wong [22], the authors examined the efficacy of closed-book invigilated final exams versus open-book and open-web (OBOW) final exams in a completely online university. They analyzed students’ experience who had completed both formats of the exam by surveying them on the merits of each exam format. Their findings showed that 100% of students found OBOW preferable to traditional closed-book exams. On
the issue of academic integrity, their results indicated that in both exam formats “there has been equal opportunity of plagiarism in the view of students.” On the other hand, there are some studies indicating that online exams and access to technology provide more opportunities for dishonest behaviors [11]. Even though there are some similarities and differences with the aforementioned studies, the research environment of our study is qualitatively different. In our study, a Bayesian statistical model is proposed to compare students’ performances reflected by their empirical marks received in a sudden unexpected scenario where both classes and evaluations were virtually forced to be conducted online.

On the other hand, the closure of universities and transitioning to online teaching may have prompted mental, psychological and educational challenges for students and faculty members. A recent study by Statistics Canada based on crowdsourcing data completed by over 100,000 post-secondary students from April 19 to May 1, 2020 provides insight into how students’ academic life was impacted by the COVID-19 pandemic [8]. Admittedly, the adoption of alternate modes of delivery and distance learning under these circumstances are not ideal. Students accustomed to on-campus learning have expressed concerns over the loss of social and interactive side of education. In addition, the rapid transition to online instruction poses challenges to students who were not equipped to adjust to this mode of learning either because they learn better in-person, lack appropriate tools, suffer financial hardship, or do not have a home environment suitable for learning online [8]. The transition of classes to online from in-person lectures due to COVID-19 had an even more severe impact on students with mental health disabilities who needed extra care and frequent face-to-face interactions with the instructors in order to keep maintaining motivation, interest and persistence to succeed in the course [7]. In a recent study, Sahu [18] highlighted the impacts of COVID-19 on education and mental health of students and academic staff and indicated the challenges posed by the closure of universities and other restrictive measures such as the shift in the delivery mode of the course, the change in the format of assessments, the travel restrictions on students, etc. In another study, Cao et al [5] investigated students’ anxiety level during COVID-19 in a Chinese medical college located in Hubei Province. Their findings showed that about a quarter of students in their sample had experienced mild, moderate or severe levels of anxiety during COVID-19 outbreak. In this paper, we have studied some aspects of stress among students with disabilities and special needs that contributed to their disengagement in any of the remaining evaluation components after the emergence of COVID-19.

We developed a Bayesian hierarchical linear mixed effects model to measure the effects of COVID-19 on students’ marks. The literature of random effects model goes back to Laird and Ware [13]. The classical maximum likelihood estimation and inference of the linear mixed effects model using expectation maximization (EM) algorithm were shown by Laird et al [12] and Lindstrom et al [15]. These articles approached the problem of model building and inference from a frequentist viewpoint. On the other hand, the general design of Bayesian method for the linear mixed effects model was described by Zhao et al [24]. The missing value imputation for linear mixed effects model, using a frequentist viewpoint, was proposed by Schafer et al [19]. In this paper, we redesigned the linear mixed effects model for the change of students’ raw marks before and after COVID-19 effect in Winter 2020 semester at TRU. We described the complete Bayesian methodology for the
proposed model using conjugate and semi-conjugate prior distributions. We then derived the full conditional posterior distributions of the parameters and proposed the Gibbs sampling [14] to generate Markov Chain Monte Carlo (MCMC) samples [9] from the posterior distributions. In order to impute missing values of marks, we assumed the mechanism of missingness completely at random [4]. Our novel contributions are the design and implementation of the fully Bayesian missing value imputation in a linear mixed effects modeling setup. When most of the classical statistical missing value imputation methods are performed only once given complete data before the model building [6], our method of Bayesian missing value imputation is flexible which seamlessly generates missing values before every generation of the Markov chain from the posterior distributions of the missing values given observed data. In our model, we allow student-specific error variances to vary which is a considerable extension of the methodology in a Bayesian linear mixed effects hierarchical modeling setup. We wrote our codes for the proposed fully Bayesian hierarchical model using R [16]. The R codes and data are available on GitHub (https://github.com/jhtomal/COVID19_STUDENT_PERF.git). We are in the process of wrapping up the codes in an R package to publish in the Comprehensive R Archive Network (CRAN) so that the broad scientific community can apply our method in their applications.

We hypothesize that the COVID-19 has negative effects on students’ performances, reflected on their marks, and on their stress level especially on students needing special supports. As such, this article aims to explore the following questions:

(i) How are the raw marks of all students in a course compared before and after the university switched to online delivery due to COVID-19?
(ii) How are the raw marks of individual students within a course compared before and after the transition to online delivery?
(iii) Are the stronger or weaker students getting higher or lower average raw marks due to online delivery relative to in-person delivery?
(iv) Are there disengaged students who did not participate in any evaluation components in a course after the university shifted to online delivery? Are these students on special support?
(v) Are the trends of marks consistent across all courses or departments in this study? If there are different trends, what are the potential factors that make the difference?

In order to cope with the unexpected transition to online delivery and open-book non-invigilated evaluations, the instructors from different departments in the Faculty of Science at TRU came up with new sets of weighted schemes which were strikingly different from what they had at the beginning of the semester. We thus analyze students’ raw marks, as opposed to weighted aggregated marks, because the raw numbers reflect the unbiased scenario. Subsequently in the Discussion section, we discussed the results according to the level of cognitive skills and hands-on experiences required in the courses we investigated. In our analysis, we classified the cognitive skills with reference to the Bloom’s Taxonomy of Knowledge 2.

2 Bloom’s Taxonomy of Knowledge is a set of three hierarchical models used to classify educational learning objectives into levels of complexity and specificity: https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/
2 Design of the study and description of data

The analysis of empirical data starts with some summary statistics followed by a fully Bayesian linear mixed effects model. The posterior statistical inferences of the model parameters proceed following posterior model inspections.

The study population consists of all of the students in the Faculty of Science at Thompson Rivers University. We collect longitudinal data by sampling 11 courses across three departments from the Faculty of Science, including seven courses in Mathematics (MATH) and Statistics (STAT), two courses in Computing Science (COMP) and two courses in Architectural & Engineering Technology (ARET). Each course has been taught, marked\(^3\), and graded by the same instructor from the start to the end of the Winter 2020 semester (before and after COVID-19 effects) which ensures instructor’s effects being adjusted while leaving only students’ effects to compare. Moreover, the alpha test of reliability analysis has been performed to each course to make sure that the course data are reliable. After the selection of the courses, student-specific raw marks are recorded over time using a series of evaluations such as multiple assignments, quizzes, tests, exams, and projects. Specific to each course, we have analyzed marks from all sorts of evaluations and ensured covering all possible grading practices. The raw marks specific to each evaluation components are then converted into percents (from 0 to 100) where larger numbers indicate better marks. We also record whether the marks are observed or missing. Given that all the classes went online starting March 15, 2020, the index variable is simply defined as a vector of boolean values indicating whether the effects of COVID-19 have occurred (before March 15 vs after March 15). As there is no randomness in the index variable, we consider it deterministic. A Bayesian hierarchical model with missing value imputation technique is then applied to determine whether students’ raw marks were increased or decreased before and after March 15. After model fitting, statistical analysis and inference are performed to student specific raw marks in each course. Our findings generalize to the Faculty of Science at TRU, and may eventually to all faculties of science across universities in Canada. Figure 1 shows the overall design of the study.

3 Methods

3.1 The Bayesian linear mixed effects model

Let \( Y_{it} \) be the raw marks in percent for the \( i \)th student evaluated at time \( t \) in a particular course. The subscript \( t \) represents a series of evaluations conducted over time from multiple assignments, quizzes, tests, exams, and projects. Consider that \( x_{it} \) be another variable measured at time \( t \) for the \( i \)th student which can explain the variation in \( Y_{it} \). We consider a linear mixed effects model as following

\[
Y_{it} = \beta_0 + \beta_1 x_{it} + \epsilon_{it},
\]

where

\[
x_{it} = \begin{cases} 
1 & \text{evaluation taken after March 15, 2020} \\
0 & \text{otherwise},
\end{cases}
\]

\(^3\) There is no teaching assistant for marking.
and \( \epsilon_{it} \sim \text{Normal}(0, \sigma^2_t) \) with student-specific error variance \( \sigma^2_t \). The sampling distribution of \( Y_{it} \) is

\[
Y_{it} \sim \text{Normal} \left( \beta_{0i} + \beta_{1i} x_{it}, \sigma^2_t \right).
\]  

(2)

In this model, \( \beta_{0i} \) and \( \beta_{0i} + \beta_{1i} \) are the average marks of the \( i \)th student before and after March 15, 2020, respectively. Here, \( \beta_{1i} \) is the difference of average marks before and after March 15, 2020 for the \( i \)th student. Let \( \beta_i = (\beta_{0i}, \beta_{1i})^T \) represents the vector of regression coefficients for the \( i \)th student. We assume that the \( i \)th student is a randomly and independently selected student from the pool of students in a specific course in the faculty of science at Thompson Rivers University (TRU). This leads us to consider the sampling distribution for \( \beta_i \) as

\[
\beta_i | \Sigma \sim \text{Multivariate Normal} \left( \theta, \Sigma \right).
\]  

(3)

where \( \theta = (\theta_0, \theta_1)^T \) and

\[
\Sigma = \begin{bmatrix}
\Sigma_{00} & \Sigma_{01} \\
\Sigma_{10} & \Sigma_{11}
\end{bmatrix},
\]

where \( \text{Var-Cov}(\beta_i, \beta_j) = \Sigma_{ij} \) for \( i, j = 0, 1 \). This part of the model explains that \( \theta_0 \) and \( \theta_0 + \theta_1 \) are the average marks of all students in a course before and after March 15, 2020, respectively.

The sampling distribution for the student-specific error variance is

\[
\sigma^2_t \sim \text{Inverse-Gamma} \left( \frac{\nu_0}{2}, \frac{\nu_0 \sigma^2_0}{2} \right).
\]  

(4)

where the course specific error variance is \( \sigma^2_0 \) with strength \( \nu_0 (>0) \).

The overall picture of the hierarchical model is shown above in Figure 2.
3.2 The prior distributions

The prior distribution represents our prior belief through hyper-parameters. We note that our prior belief should be unbiased which should not pull the posterior belief towards prior belief unless we have strong evidence towards our belief. In situations where we have weak or no prior belief, we prefer to be objective and follow the lead by the data.

We restrict \( \nu_0 \) to be a whole number and choose the prior on \( \nu_0 \) to be a discrete analogue of exponential distribution on \( \{1, 2, 3, \cdots\} \) as following:

\[
p(\nu_0) \propto (1 - e^{-\alpha_3})e^{-\alpha_3 \nu_0}.
\]

(5)

The prior distribution for the course-specific error variance \( \sigma_0^2 \) is

\[
\sigma_0^2 \sim \text{Gamma}(\alpha_1, \alpha_2),
\]

(6)

where the prior belief regarding \( \sigma_0^2 \) is represented by the hyper-parameters \( \alpha_1 \) and \( \alpha_2 \). Here, small and large numbers of \( \alpha_1 \) represent weak and strong belief regarding the prior course-specific error variance.

The mean parameter vector \( \theta \) for the course has the following prior distribution

\[
\theta \sim \text{Multivariate-Normal}(\theta_0, \Sigma_0).
\]

(7)

The prior belief about the course specific mean vector is represented by \( \theta_0 \) and the strength of prior belief is represented by the variance-covariance matrix \( \Sigma_0 \). Here, \( \Sigma_0 \) is a positive-definite matrix where the diagonal elements contain the variances and the off-diagonal elements contain the covariances.

The prior distribution corresponding to the variance-covariance matrix \( \Sigma \) of the course-specific mean vector \( \theta \) is

\[
\Sigma \sim \text{Inverse-Wishart}(\eta_0, S_0^{-1}),
\]

(8)

where the prior belief about \( \Sigma \) is represented by \( S_0 \) and the strength of prior belief is represented by \( \eta_0 \). As large value of \( \eta_0 \) represents strong prior belief, we may wish to make the distribution diffuse or flat by assigning small values for \( \eta_0 \).
3.3 The posterior distributions

After collecting data, we combine the information from the data with our prior belief and obtain posterior belief. In other words, the posterior belief is our updated belief after observing the data.

The posterior distribution for the strength parameter $\nu_0$ for the overall error variance $\sigma_0^2$ is obtained using equations 4 and 5 as following:

$$
\nu_0|\sigma_0^2, \sigma_1^2, \cdots, \sigma_k^2 \propto \left(\frac{\nu_0 + \sum \sigma_i^2}{\frac{\nu_0}{\sigma_0^2}}\right)^{\frac{k}{2}} \times \left(\prod \frac{1}{\sigma_i^2}\right)^{\frac{k}{2} - 1} e^{-\nu_0 \left(\frac{2\nu_0 + \sum \sigma_i^2}{\sigma_0^2}\right)}
$$

The posterior distribution for the overall error variance $\sigma_0^2$ is Gamma which is obtained using equations 4 and 6 as following:

$$
\sigma_0^2|\sigma_1^2, \cdots, \sigma_k^2, \nu_0 \sim \text{Gamma} \left(\alpha_1 + \frac{k\nu_0}{2}, \alpha_2 + \frac{\nu_0}{2} \sum \frac{1}{\sigma_i^2}\right).
$$

The posterior distribution for the student-specific error variance $\sigma_i^2$ is independent Inverse Gamma which is obtained by using equations 2 and 4 as following:

$$
\sigma_i^2|y_i, x_i, \beta, \nu_0, \sigma_0^2 \sim \text{Inverse-Gamma} \left(\nu_0 + d, \frac{\nu_0\sigma_0^2 + \text{SSE}(\beta_i)}{2}\right),
$$

where $\text{SSE}(\beta_i) = (y_i - x_i\beta_i)^T (y_i - x_i\beta_i)$ is the student-specific sum of squares of errors.

The posterior distribution for each $\beta_i$ is independent Multivariate Normal which is obtained using equations 2 and 3 as following:

$$
\beta_i|y_i, x_i, \theta, \sigma_i, \Sigma \sim \text{Multivariate-Normal} \left(E(\beta_i)_p, \text{Var-Cov}(\beta_i)_p\right),
$$

with posterior mean

$$
E(\beta_i)_p = \left(\frac{x_i^T x_i}{\sigma_i^2} + \Sigma^{-1}\right)^{-1} \left(\frac{x_i^T y_i}{\sigma_i^2} + \Sigma^{-1}\theta\right)
$$

and posterior variance-covariance matrix

$$
\text{Var-Cov}(\beta_i)_p = \left(\frac{x_i^T x_i}{\sigma_i^2} + \Sigma^{-1}\right)^{-1}.
$$

The posterior distribution for the overall course specific mean vector $\theta$ is Multivariate Normal which is obtained using equations 3 and 7

$$
\theta|\beta_1, \cdots, \beta_k, \Sigma \sim \text{Multivariate-Normal} \left(E(\theta)_p, \text{Var-Cov}(\theta)_p\right),
$$

with posterior mean

$$
E(\theta)_p = \left(\Sigma_0^{-1} + K \Sigma^{-1}\right)^{-1} \left(\Sigma_0^{-1}\theta_0 + \Sigma^{-1} \sum \beta_i\right).
$$
and posterior variance-covariance matrix

\[
\text{Var-Cov}(\theta)_p = \left( \Sigma_0^{-1} + K \Sigma_0^{-1} \right)^{-1}.
\]

The posterior distribution of \( \Sigma \) is Inverse Wishart which is obtained via equations 3 and 8 as following

\[
\Sigma|\theta, \beta_1, \ldots, \beta_k \sim \text{Inverse-Wishart} \left( \nu_0 + k, (S_{\theta} + S_0)^{-1} \right),
\]

where \( S_{\theta} = \sum_{i=1}^{k} (\beta_i - \theta)(\beta_i - \theta)^T \).

### 3.4 Distribution of missing values

Let \( Y_{i,o} \) and \( Y_{i,m} \) be the marks that are observed and missing, respectively, for the \( i \)th student in a course. Given the distribution of the marks

\[
Y_{i}|x_i, \beta_i, \Sigma, \sigma_i^2 \sim \text{MN} \left( x_i \beta_i, x_i \Sigma x_i^T + \sigma_i^2 I_{d \times d} \right),
\]

the missing marks for the \( i \)th student are imputed independently by generating data from the following distribution

\[
p \left( Y_{i,m}|y_{i,o}, \beta_i, \Sigma, \sigma_i^2 \right),
\]

using the properties of conditional multivariate normal distribution. Note that, in generating the missing values, the \( \beta_i, \Sigma \) and \( \sigma_i^2 \) are generated first from their respective posterior distributions.

### 3.5 The Gibbs sampling algorithm

The approximation of posterior distribution via Gibbs sampling is briefly presented below. For a given state of the parameters

\[
\left\{ \nu_0, \sigma_0^2, \beta_1^{(s)}, \ldots, \beta_k^{(s)}, \sigma_1^2, \ldots, \sigma_k^2, \theta^{(s)}, \Sigma^{(s)}, y_{i,m}^{(s)} \right\},
\]

a new state is generated as follows:

1. Sample \( \nu_0 \) using

   \[
   \nu_0^{(s+1)} \sim P \left( \nu_0 | \sigma_0^{2(s)}, \sigma_1^{2(s)}, \ldots, \sigma_k^{2(s)} \right).
   \]

2. Sample \( \sigma_0^2 \) using

   \[
   \sigma_0^{2(s+1)} \sim P \left( \sigma_0^2 | \sigma_1^{2(s)}, \ldots, \sigma_k^{2(s)}, \nu_0^{(s+1)} \right).
   \]

3. For each \( i \in \{1, 2, \ldots, k\} \), independently sample \( \sigma_i^2 \) using

   \[
   \sigma_i^{2(s+1)} \sim P \left( \sigma_i^2 | y_{i,o}, y_{i,m}^{(s)}, x_i, \beta_i^{(s)}, \nu_0^{(s+1)}, \sigma_0^{2(s+1)} \right).
   \]
4. For each $i \in \{1, 2, \ldots, k\}$, independently sample $\beta_i$ using

$$
\beta_i^{(s+1)} \sim P(\beta_i | y_{i,o}, y_{i,m}^{(s)},x_i, \theta, \sigma_i^{2(s+1)}, \Sigma^{(s)}).
$$

5. Sample $\theta$ using

$$
\theta^{(s+1)} \sim P(\theta | \beta_1^{(s+1)}, \ldots, \beta_k^{(s+1)}, \Sigma^{(s)}).
$$

6. Sample $\Sigma$ using

$$
\Sigma^{(s+1)} \sim P(\Sigma | \theta^{(s+1)}, \beta_1^{(s+1)}, \ldots, \beta_k^{(s+1)}).
$$

7. For each $i \in \{1, 2, \ldots, k\}$, sample the missing marks using

$$
y_{i,m}^{(s+1)} \sim P(y_{i,m} | y_{i,o}, y_{i,m}^{(s)}, x_i, \beta_i^{(s+1)}, \Sigma^{(s+1)}, \sigma_i^{2(s+1)}).
$$

The order in which the new parameters are generated does not matter. What does matter is that each parameter is updated conditional upon the most current value of the parameters and imputed missing values. The Gibbs sampling algorithm is implemented using the R [16] language.

3.6 Prior specification

The hyper-parameters for the prior distributions are chosen as described below. The hyper-parameters for the prior distribution of $\sigma_0^2$ (equation 6) are chosen as $\alpha_1 = 1$ and $\alpha_2 = 1/100$. This specification implies $\sigma_0^2$ to be 0.01. In other words, our prior specification considers very small variability between $\sigma_1^2$, $\sigma_2^2$, $\ldots$, $\sigma_k^2$. That is, all the student-specific error variances are approximately equal. At the same time, we consider $\alpha_3$ (the hyper-parameter for the prior of $\nu_0$) to be 1 which dictates $\nu_0$ to be a small number. A small value for $\nu_0$ makes our prior belief for $\sigma_0^2$ weak or diffuse. Given that we have no prior information about student specific error variances, a weak prior imposes less subjectivity and lets the data objectively determine the parameter values.

To choose the hyper-parameters for the prior distribution of $\theta$ (equation 7), we fitted ordinary least squares (OLS) regression to student-specific marks and saved the OLS coefficients. The OLS coefficients are averaged to specify $\theta_0$ (the prior mean vector for $\theta$). The prior variance-covariance matrix $\Sigma_0$ is considered to be the sample covariance of the OLS coefficients. Such a prior distribution represents the information of someone with relatively unbiased prior information.

Similarly, we consider the prior sum of squares matrix $S_0$ (equation 6) to be equal to the sample covariance of the ordinary least squares estimates of the coefficients. This ensures relatively unbiasedness as the expectation of the prior distribution of $\Sigma$ equal to the ordinary least squares estimates of the coefficients. But at the same time, we consider $\eta_0 = d + 2 = 4$. This makes the prior distribution flat or diffuse to make the prior information weak. Such a prior specification ensures less subjectivity or more objectivity.

In order to jump-start the Gibbs sampling algorithm, the initial parameter values are obtained from the OLS estimates of the coefficients. The initial missing values are obtained by simple average of the student-specific observed marks. We
used the first 1000 scans of the Gibbs sampler as burn-in and threw the values of the realizations out. This eliminates the effect of the initial values we have chosen to jump-start the Gibbs sampler. We then ran the Gibbs sampler for another 10,000 scans and saved every 10th scan to produce a sequence of 1,000 values for each parameter. We checked if the Markov chain for each parameter obtained stationarity or not. The autocorrelation values and plots for the sequence of saved scans for each parameter are also checked. After confirming convergence in terms of stationarity and minimal autocorrelation, we proceeded next to perform Bayesian posterior inference.

Figure 3 shows the MCMC trace-plots against thinned scans of the chains for $\theta_0$ (top-left) and $\theta_1$ (top-right). These plots show the generated values of $\theta_0$ and $\theta_1$ from their respective posterior distributions saved at every 10th scan of the chain after throwing out the burned-in scans. It is visible that the chains achieved stationarity. The bottom portion of Figure 3 shows the auto-correlation functions (ACFs) for $\theta_0$ (bottom-left) and $\theta_1$ (bottom-right). It is also clear that the generated thinned scans of the chains are nearly uncorrelated. In fact, their effective sample sizes are straight 1000 in each. Similarly, we have confirmed stationarity for every other course selected in our study.

4 Results

We summarized the number of students, the percentage of disengaged students, and the percentage of missing marks within each course in Table 1. The percentage of disengaged students is defined as the percentage of students within a course who did not participate in any of the assessment/evaluation components after classes and assessments/evaluations transitioned to online. The percentage of disengaged students varies from 0% to 5.56% across different courses. Even though these numbers are relatively small, the implications are significant. We investigated the reasons as to why students disengaged after March 15, 2020. According to students’ records/responses, at least 43% of these students were in need of special supports and accommodations due to mental illness caused by concussion, severe disability in coping with stress, and being slow in processing information. Most of such students had difficulties following the course content in an online delivery mode without face-to-face interactions with the instructors. The other 57% of disengaged students were those who struggled the most with the course contents. Since there were no marks available for the evaluation components for the disengaged students after March 15, 2020, we excluded them from further analysis.

The percentage of missing marks within a course is defined as the proportion of assessment/evaluation components that students did not participate relative to all evaluation components multiplied by the number of the remaining students, regardless of whether the evaluation components were taken before or after March 15. These numbers vary from 3.73% to 14.07%. As shown in Table 2, the percentage of missing marks are higher for the courses in which more direct hands-on supports are required in the form of labs, coding, programming, and seminars than the courses evaluated mainly using assignments, quizzes, tests, and exams. It is worth mentioning that one of the features of our model is that it can objectively impute the missing values seamlessly from their conditional posterior distributions given observed data.
Figures 4–12 show the results of each individual course in our study. In each figure, there are four panels. Panel (a) shows the distributions of marks before and after COVID-19 took effect on March 15, 2020 for all the students in a course. Panels (b) to (d) show three interesting student-specific cases selected from that course. Within each panel, there are two distributions: the grey (for either $\theta_0$ or $\beta_0$) and black (for either $\theta_0 + \theta_1$ or $\beta_0 + \beta_1$) curves represent the distributions of marks before and after March 15, respectively. A horizontal shift of the black curve towards the right (left) of the grey curve indicates an increase (decrease) of marks after March 15. The wider and narrower distributions (curves) indicate larger and smaller variability (spread) of marks, respectively. The increase or decrease of marks is statistically significant when the two distributions do not overlap. As we have adjusted the instructor’s effects in the form of a statistical design known as matching, the observed effects are attributed to the changes of students efforts due to COVID-19.

Tables 4–12 report the summary of the corresponding figures in the form of numbers. The first row of each table contains information for the entire class, and the next three rows show three student-specific cases. In each row, there is a lower limit, a median and an upper limit which represent the 2.5th, 50th and 97.5th percentiles of the distribution of marks before and after COVID-19, respectively. The median shows the central value (the middle most value) of the distribution of marks; whereas the lower and upper limits bound the 95% confidence interval, also known as the “credible interval” in Bayesian statistics. The increase or decrease of marks is statistically significant at 5% level when the two credible intervals do not overlap.

4.1 Course-specific analysis

Figure 4 and Table 4 show the comparisons of overall performances and three student-specific performances before and after March 15 for the course MATH 1070: Mathematics for Business and Economics. The overall marks for all the students in this course went up from a median of 66.683% (95% credible interval of 61.203% to 71.763%) to 74.991% (with 95% credible interval of 69.353% to 80.647%). Even though there is a small overlap in the 95% credible intervals for the overall marks, student-specific marks for some students increased by a large margin. The first case (Student 5) made a significant shift of marks from a failing letter grade to a passing letter grade of C (TRU grading scales are summarized in Table 3). This student was most likely failing as his/her average marks had a 95% credible interval ranging from 47.384% to 50.470%. The second case (Student 14) made an even larger shift from a failing letter grade to a strong letter grade of A-. The third case (Student 27) also showed a large shift from a barely passing letter grade of D to a passing letter grade of A. None of the distributions of marks before and after March 15 had any overlaps, making the shifts highly statistically significant.

Figures 5–7 and Tables 5–7 show the comparisons of overall performance and student-specific cases in each of the courses MATH 1240: Calculus 2, MATH 1250: Calculus for Biological Sciences 2, and MATH 1640: Technical Mathematics 1. After the transition to online delivery, the overall medians in these courses were increased by about 20%, 8% and 7%, respectively. The overall increase of marks
was statistically significant for the course MATH 1240. On the other hand, the increase of student-specific marks within the course was not uniform. For example, the performance of Student 5 in MATH 1240 improved by only 1.67%, while Students 6 and 29 had an increase of marks by 33.66% and 29.18%, respectively. The last two students shifted their marks from a failing grade to passing letter grades of $A^-$ and $B$, respectively. A more or less similar trend is observed for Students 6 and 24 in MATH 1250, with a small variability in their grades. In other words, a few weaker students managed to improve their marks consistently to a significantly higher range of grades after March 15. However, some good-standing students, such as Student 26 in MATH 1250 and Student 5 in MATH 1640, experienced a decline in their marks. Note that the decline of marks for these good students were statistically insignificant. While the decline of marks for Student 5 in MATH 1640 was practically insignificant, the decline for Student 26 in MATH 1250 was practically significant as the change of the letter grade was from $A^+$ to $A$.

Figures 8–9 and Tables 8–9 show the comparisons of overall marks and student-specific cases for the second-year courses MATH 2200: Introduction to Analysis, and MATH 2240: Differential Equations 1, respectively. MATH 2200 is a proof course, required for Math Majors and a gateway to heavy-proof math courses. After the transition to online delivery, the overall median marks were increased by about 10%. For most of the students the improvement was not notable. However, for some students, such as Students 5 and 9, the improvements in marks (41.96% and 41.28%, respectively) were practically and statistically significant. Both of the students’ marks moved from a failing letter grade to passing letter grades of $B^+$ and $C^-$, respectively. In MATH 2240, the overall marks increased by about 11% with almost zero overlap before and after COVID-19. A similar trend of growth in the marks, especially for the struggling students, has been observed in this course. For instance, Students 14 and 21 have shown a significant jump in their marks after March 15. On the other hand, the $A^+$ level Student 29 experienced a very minimum change in his/her grade. This change is insignificant both practically and statistically.

Figure 10 and Table 10 display the comparisons of marks before and after March 15 for the course STAT 2000: Probability and Statistics. The overall marks went up slightly from a median of 65.653% to 67.675%. As there is an overlap in the two distributions, the increase is not statistically significant. Regarding the student-specific cases, Students 5 and 22 who were barley passing the course had an increase by about 15% and 16%, respectively. None of the two distributions overlap each other before and after March 15, hence, the shifts are highly statistically significant. On the other hand, a few good-standing students experienced a decline in their marks from higher percentages to lower percentages. For example, Student 12 was in the $A^+$ grade range before March 15, while his/her marks went down significantly by around 11% to the letter grade of $B+$ after March 15.

Figures 11–12 and Tables 11–12 compare the overall performance and student-specific cases for the courses ARET 1400: Civil Technology 1, and ARET 2600: Statics and Strength of Materials. The overall students’ performances in these two courses declined after the transition to online delivery. In ARET 1400, the median decreased significantly by 12.14%, whereas in ARET 2600, the median decreased by 4.93%. The decrease in ARET 2600 is not statistically significant. The amount of decrease or increase of marks varies from one student to another.
For example, the decrease in marks for Student 29 in ARET 1400 and Student 15 in ARET 2600 were 30.75% and 18.74%, respectively. On the other hand, Student 6 in ARET 1400 and Student 4 in ARET 2600 did not experience a large drop in their marks. For Student 4 in ARET 2600, the decrease is insignificant both statistically and practically. There are a few exceptions to this trend as well. For example, Student 12 in ARET 1400 and Student 12 in ARET 2600 did experience some increase in their marks after transitioning to online delivery mode. Again, the increase of 2.36% for Student 12 in ARET 1400 was insignificant both statistically and practically.

Figures 13–14 and Tables 13–14 show overall performances and student-specific cases for the courses COMP 2680: Web Development, and COMP 4980: Bioinformatics. Students’ performances in these Computing Science courses were negatively affected by the transition to online delivery mode. The median of overall marks went down by 3% for COMP 2680 and by 19% for COMP 4980. The decrease is statistically insignificant for COMP 2680, while significant for COMP 4980. The marks decreased for a large number of students after March 15. For example, the performance of Students 18 and 22 in COMP 2680 decreased by 14.86% and 33.82% and the performance of Students 8 and 11 in COMP 4980 decreased by 1.95% and 40.56%, respectively. One the other hand, some good-standing students were able to maintain their good performance after March 15, such as Student 16 in COMP 2680 and Student 19 in COMP 4980. However, these increases were insignificant both practically and statistically.

5 Discussion of results

Among the 11 courses in this study, both increasing and decreasing trends in students’ marks were observed. Specifically, a general increase in the marks are shown in theory-based courses requiring lower-level cognitive skills according to Bloom’s Taxonomy of Knowledge, whereas a general decrease in the marks are shown in the courses requiring either interactive hands-on support or higher-level cognitive skills.

5.1 Rising trend in courses requiring lower-level cognitive skills

University-level math courses are normally delivered in a traditional lecture format with the instructor teaching core concepts and theories accompanied by related examples and applications varying from direct applications to more conceptual and intricate ones. Student assessment is then composed of several in-class quizzes, written homeworks, one or two midterms and a final exam with generally more weight on summative assessments than formative ones. Knowing that a standard first- or second-year math course is often taken by a large group of students enrolled in various university programs with a wide range of backgrounds in math, assessments in these courses are mainly focused on questions with a medium-difficulty level in order to reasonably evaluate students’ learning. In other words, according to Bloom’s Taxonomy of Knowledge, in-person math exams are normally testing low- to medium-level skills and abilities with limited allocation of questions to
higher level skills such as analysis and synthesis. However, transitioning to online and open-book exams did change all equations.

The unprecedented closure of universities due to COVID-19 imposed an unexpected shock to academia in particular to the traditional culture of course delivery and assessment in mathematics. Given the very limited time for instructors to prepare for switching to online modes of delivery, creating online open-book tests and restructuring in-person exams to be suitable for the online format was rather infeasible. Moreover, although some instructors did attempt to design tests relatively different than in-person exams in order to target deeper levels of understanding, they were faced with students’ complaints and resistance. This is completely understandable though on account of the lack of training throughout the semester for such exams. Students in MATH 1250, for example, even in normal circumstances, are mostly categorized in the struggling group with levels of math anxiety; hence it is unrealistic to expect them to perform well in an exam format to which they are not used to. Therefore, most MATH and STAT courses maintained a format similar to in-person exams with the online assessments being open-book.

Nevertheless, with the availability of resources in an open-book exam the low- and medium-level question types that target memory and comprehension skills such as recalling, defining, describing or explaining concepts were no longer truly examining students’ learning as these can be easily found in textbooks, class notes, etc. It was observed that students had a better performance in these question types in the online version of exams compared to similar face-to-face exams prior to university closure. For instance, students in MATH 1240 are normally in the category of moderately strong Science students and their performance improved after the transition to online delivery. The improvement took place mostly because the assessment components and their structures did not change, except the exams being open-book, and students could adapt to the new delivery method with relative ease. Similarly, in MATH 2240 which is a second-year course, the overall performance has grown significantly which can be partly attributed to open-book exams maintaining a similar structure to face-to-face ones.

Furthermore, one can consider the role of technology and online resources available to students during an online math exam. Tools were no longer limited to a basic scientific calculator. Advanced online calculators and math programs, along with many online forums were at hand during an open-book exam. For example, at the beginning of the Winter 2020 semester, all students in MATH 2200 were struggling with the course content as writing a rigorous proof is a skill never taught in the first-year courses. After about a month, many students in this class could improve their proof writing skills and consequently improved their grades to some extent. When the course transitioned to online delivery, the assessment components remained unchanged, most of the students continued their upward trend and their overall performance improved. But this transition might have provided some weaker students the opportunity to seek other resources for answers in non-invigilated tests. This pattern can be seen for students 5 and 9 (Figure 8). This is also observed in Student 24 in MATH 1250 and Student 21 in MATH 2240, whose large grade improvements turned their grade from failing to the letter grades of C+ and B, respectively. The performance of Student 29 in MATH 2240 and Student 17 in MATH 2200, however, did not change significantly after March 15. These students were among the average- to high-performing students.
The statistics course (STAT 2000) we investigated in this study tests medium-level cognitive skills of students, as per Bloom’s Taxonomy, and falls slightly above the Math courses. This course requires students to understand the methods, organize the information in the data, apply the methodology to the data to gain insightful knowledge, and provide summary and explanations of the findings. Nurturing these skills requires some discussions between the instructor and the students. After the transition to online delivery, the support that students needed were provided to the best of the instructor’s ability, especially to the students who requested support via online meetings and discussions. As the structure of the course is very similar to Math courses with some added applications, the overall performance of students improved slightly by about 2%. This increase of overall performance is not as large as the Math courses which are aligned with lower-level cognitive skills. Moreover, when most of the weaker students in this course improved their grades by a large margin, the top students experienced a slight decrease in their marks despite their potential.

In addition, there is one other factor that may contribute to the growth in grades. The time that unexpectedly became available to students. No face-to-face classes meant that no more time needed to be spent on transportation to school and running from one class to the other. Moreover, a large portion of students working at part-time jobs were to stay at home due to the lock-down; as a result, some students could possibly turn this into an opportunity to invest more time and effort into their course work, and hence improve their performance in the course.

All of these factors may have contributed to the increase in students’ performance.

5.2 Decreasing trend in courses requiring higher-level cognitive skills and interactive hands-on support

In ARET courses, application and analysis of learned theories and concepts play a key role in the assessments. While assessments in math courses revolve around defining, calculating or reproducing facts pertaining to a topic, ARET courses demand students’ expertise in applying the knowledge learned to novel applications/situations. Assessments for the ARET courses investigated in this study entail developing problem solving skills with emphasis on analyzing and devising the concepts and principles in applied situations. These skills are categorized as medium- to high-level cognitive skills, as per Bloom’s Taxonomy. Students in these ARET courses experienced a decline in their grades after March 15 mainly because they needed some hands-on and face-to-face support to develop their analyzing and problem solving skills required for the final exam. Providing such hands-on support was not feasible after March 15 and, as a result, many students in ARET 1400 and ARET 2600 did not perform well in their final exam. A few of these students happened to be among the stronger cohort of students; for example, Student 29 in ARET 1400 and Student 15 in ARET 2600.

Computing Science is known as a ‘learning-by-doing’ subject, and most COMP courses require enormous interactive supports in hands-on programming and laboratories, without which it is hard for students to succeed in these courses. Table 2 shows that the computing science courses at TRU are more geared towards hands-on practices in lab components compared to MATH and STAT courses. For the
two COMP courses investigated in this study, students in COMP 2680 learn web development skills in lectures, then practice and apply these skills in hands-on laboratories. COMP 4980 is an interdisciplinary course where students learn how to apply computing science skills to analyze, synthesize and interpret the biological data. It requires not only hands-on laboratories to practice problem solving skills, but also supports from instructor’s domain-specific knowledge to help students connect the biological problems that they try to solve with computational models and interpret their results both biologically and mathematically. In other words, COMP courses investigated in this study test students’ medium- to high-level cognitive skills, as per Bloom’s Taxonomy.

Unsurprisingly, students’ performances in both courses were negatively affected by the rapid switch from face-to-face to online delivery mode. The marks dropped after March 15 for most students as shown in Figures 13 and 14, including students with good standings (e.g., Student 18 in COMP 2680 and Student 8 in COMP 4980) and relatively weak students (e.g., Student 22 in COMP 2680 and Student 11 in COMP 4980). There might be a few reasons that potentially caused the decrease of students’ marks after March 15 in these courses. For instance, most students did better in the first half of the semester, because the first few topics of the course were introductory topics which were easy to pick up, and more difficult topics were introduced in the second half of the semester (after March 15). During the face-to-face delivery, students could also get hands-on help from instructors or TAs in lectures, labs, or Computing Science Help Center, but the hands-on labs and the Help Center were no longer accessible after March 15. For COMP 4980, there was a group project in the last three weeks of the semester, but students could not get face-to-face interaction with each other and did not receive the same level of support from the instructor due to the online delivery mode, therefore, some students struggled in the term project.

However, there are some exceptions to this trend, such as Student 16 in COMP 2680, Student 19 in COMP 4980, Student 12 in ARET 1400 and Student 12 in ARET 2600. This group of students showed extra efforts, compared to others, to adapt to the change of delivery mode by always showing up in virtual lectures, asking questions during lecture hours, discussing problems with other students in assignment forums, requesting extra office hours with instructors through video conferences, etc.

6 Conclusion

We report results from a moderately large scale study from 11 courses, where the effects of COVID-19 on students’ performance were compared with empirical rigor. This study shows that a sudden change of delivery mode has an immense impact on students’ marks. After switching to online delivery mode and assessments due to COVID-19, students’ marks were increased in theory-based courses that required lower-level cognitive skills based on Bloom’s Taxonomy, whereas in courses with hands-on lab, coding and programming components, or courses that required higher-level cognitive skills, the marks were decreased. The larger increase (for MATH and STAT courses) or decrease (for COMP and ARET courses) of marks are mainly observed for weaker students as opposed to stronger students. The group of stronger students experienced a smaller decrease of marks, while some
very hard working students were able to maintain a good standing of marks towards their credentials. The impact has been much more significant on students with special needs who disengaged from the course after March 15. We also emphasize on the fact that the COVID-19 outbreak, lock-down and closure of schools have exposed students to an extraordinary stress level. Students faced the sudden shock of online transition with virtually no education and training on how to take ownership over their submitted work in the online space and be accountable for that.

The authors of this paper observed similar trends in results across Canada which was discussed in many educational workshops and meetings, such as SSC Webinar on Teaching Statistics Online and CMS COVID-19 Research and Education Meeting (CCREM). Hence, the results of the study can also be generalized across Canada. This is because most, if not all, of the universities across Canada follow the same educational system and experienced moving towards online teaching and assessment at around the same time.

Our novel contribution is analysing and comparing COVID-19 effects on students' marks. In addition to this novel application, we also designed and developed novel computational models. A Bayesian linear mixed effect model was designed to fully address the comparison of marks. The implementation of Bayesian missing value imputation is novel both in terms of statistics and application.

Acknowledgement

We would like to thank Shi Shi for sharing the course information of COMP 2680. We would also like to thank Dr. Musfiq Rahman for his helpful discussions in this study.

Declarations

Funding: No authors declare any funding.
Conflicts of interest/Competing interests: No authors declare any conflicts of interest.
Availability of data and material: Upon acceptance, data will be available at GitHub (https://github.com/jhtomal/COVID19_STUDENT_PERF.git).
Code availability: Upon acceptance, the R codes will be available on GitHub (https://github.com/jhtomal/COVID19_STUDENT_PERF.git).

---

4 https://ssc.ca/en/node/10244
5 https://ccrem20.cms.math.ca/
7 Figures

Fig. 3 Trace plots for MCMC samples (top) and autocorrelation functions (bottom) for $\theta_0$ and $\theta_1$ for MATH 1070

Fig. 4 Overall performances and interesting cases for MATH 1070: Mathematics for Business and Economics
Fig. 5 Overall performances and interesting cases for MATH 1240: Calculus 2

Fig. 6 Overall performances and interesting cases for MATH 1250: Calculus for Biological Sciences 2
Fig. 7 Overall performances and interesting cases for MATH 1640: Technical Mathematics 1

Fig. 8 Overall performances and interesting cases for MATH 2200: Introduction to Analysis
Fig. 9 Overall performances and interesting cases for MATH 2240: Differential Equations 1

Fig. 10 Overall performances and interesting cases for STAT 2000: Probability and Statistics
Fig. 11 Overall performances and interesting cases for ARET 1400: Civil Technology 1

Fig. 12 Overall performances and interesting cases for ARET 2600: Statics and Strength of Materials
Fig. 13 Overall performances and interesting cases for COMP 2680: Web Site Design and Development

Fig. 14 Overall performances and interesting cases for COMP 4980: Introduction to Bioinformatics
8 Tables

Table 1 Number of students, percentage of disengaged students, and percentage of missing marks

| Courses       | Number of students | Disengaged students (%) | Missing marks (%) |
|---------------|--------------------|--------------------------|-------------------|
| MATH1070      | 36                 | 5.56                     | 3.78              |
| MATH1240      | 29                 | 0                        | 3.73              |
| MATH1250      | 35                 | 0                        | 5.14              |
| MATH1640      | 29                 | 3.45                     | 3.86              |
| MATH2200      | 19                 | 0                        | 6.22              |
| MATH2240      | 38                 | 0                        | 4.67              |
| STAT2000      | 29                 | 3.45                     | 6.34              |
| ARET1400      | 34                 | 2.94                     | 7.67              |
| ARET2600      | 21                 | 0                        | 13.09             |
| COMP2680      | 28                 | 3.57                     | 14.07             |
| COMP4980      | 28                 | 0                        | 5.65              |

Table 2 Number of evaluation components and percentage of hands-on (lab, coding and programming) components

| Courses       | Number of evaluation components | Hands-on percentages |
|---------------|---------------------------------|----------------------|
| MATH1070      | 7                               | 0                    |
| MATH1240      | 12                              | 0                    |
| MATH1250      | 10                              | 0                    |
| MATH1640      | 12                              | 0                    |
| MATH2200      | 11                              | 0                    |
| MATH2240      | 9                               | 0                    |
| STAT2000      | 9                               | 0                    |
| ARET1400      | 15                              | 26.7                 |
| ARET2600      | 20                              | 60                   |
| COMP2680      | 15                              | 60                   |
| COMP4980      | 12                              | 60                   |
Table 3  Grading Scale for Undergraduate Academic Programs at Thompson Rivers University

| Letter Grade | Numerical Grade | Grade Points |
|--------------|-----------------|--------------|
| A+           | 90-100          | 4.33         |
| A            | 85-89           | 4.00         |
| A-           | 80-85           | 3.67         |
| B+           | 77-79           | 3.33         |
| B            | 73-76           | 3.00         |
| B-           | 70-72           | 2.67         |
| C+           | 65-69           | 2.33         |
| C            | 60-64           | 2.00         |
| C-           | 55-59           | 1.67         |
| D            | 50-54           | 1.00         |
| F            | 0-49            | 0.00         |

Table 4  Overall performances and interesting cases for MATH 1070: Mathematics for Business and Economics

| Index | Statistics | COVID-19 | Before | After |
|-------|------------|----------|--------|-------|
| Overall | Lower Limit | 64.253 | 69.353 |
|        | Median     | 66.683 | 74.991 |
|        | Upper Limit | 71.763 | 80.647 |
| Student 5 | Lower Limit | 47.384 | 61.043 |
|          | Median     | 48.827 | 63.618 |
|          | Upper Limit | 50.470 | 66.098 |
| Student 14 | Lower Limit | 41.539 | 74.345 |
|           | Median     | 45.215 | 81.242 |
|           | Upper Limit | 50.226 | 86.742 |
| Student 27 | Lower Limit | 52.770 | 85.118 |
|           | Median     | 54.246 | 87.676 |
|           | Upper Limit | 55.663 | 89.824 |

Table 5  Overall performances and interesting cases for MATH 1240: Calculus 2

| Index | Statistics | COVID-19 | Before | After |
|-------|------------|----------|--------|-------|
| Overall | Lower Limit | 50.752 | 70.954 |
|        | Median     | 57.443 | 77.655 |
|        | Upper Limit | 63.931 | 84.020 |
| Student 5 | Lower Limit | 49.610 | 49.650 |
|           | Median     | 54.974 | 56.645 |
|           | Upper Limit | 60.289 | 64.299 |
| Student 6 | Lower Limit | 45.104 | 77.593 |
|           | Median     | 47.317 | 80.984 |
|           | Upper Limit | 49.515 | 84.194 |
| Student 29 | Lower Limit | 42.372 | 70.388 |
|           | Median     | 44.162 | 73.348 |
|           | Upper Limit | 45.991 | 76.251 |
Table 6 Overall performances and interesting cases for MATH 1250: Calculus for Biological Sciences 2

| Index    | Statistics  | COVID-19 Before | COVID-19 After |
|----------|-------------|-----------------|----------------|
| Overall  | Lower Limit | 58.007          | 67.038         |
|          | Median      | 65.298          | 73.057         |
|          | Upper Limit | 72.111          | 79.085         |
| Student 6| Lower Limit | 51.292          | 77.433         |
|          | Median      | 53.754          | 80.809         |
|          | Upper Limit | 56.679          | 83.703         |
| Student 24| Lower Limit | 48.274          | 62.440         |
|          | Median      | 48.279          | 65.082         |
|          | Upper Limit | 50.673          | 67.650         |
| Student 26| Lower Limit | 87.516          | 84.943         |
|          | Median      | 90.833          | 88.016         |
|          | Upper Limit | 93.522          | 91.788         |

Table 7 Overall performances and interesting cases for MATH 1640: Technical Mathematics 1

| Index    | Statistics  | COVID-19 Before | COVID-19 After |
|----------|-------------|-----------------|----------------|
| Overall  | Lower Limit | 75.653          | 81.511         |
|          | Median      | 79.584          | 86.368         |
|          | Upper Limit | 83.197          | 91.297         |
| Student 5| Lower Limit | 95.986          | 94.657         |
|          | Median      | 96.652          | 95.672         |
|          | Upper Limit | 97.313          | 96.739         |
| Student 12| Lower Limit | 53.355          | 74.291         |
|          | Median      | 56.996          | 79.491         |
|          | Upper Limit | 60.992          | 84.669         |
| Student 25| Lower Limit | 57.587          | 73.124         |
|          | Median      | 63.029          | 83.916         |
|          | Upper Limit | 70.621          | 91.720         |

Table 8 Overall performances and interesting cases for MATH 2200: Introduction to Analysis

| Index    | Statistics  | COVID-19 Before | COVID-19 After |
|----------|-------------|-----------------|----------------|
| Overall  | Lower Limit | 50.782          | 64.223         |
|          | Median      | 61.399          | 71.676         |
|          | Upper Limit | 71.497          | 79.387         |
| Student 5| Lower Limit | 31.422          | 70.392         |
|          | Median      | 36.763          | 78.721         |
|          | Upper Limit | 43.312          | 85.561         |
| Student 9| Lower Limit | 14.921          | 55.265         |
|          | Median      | 16.164          | 57.443         |
|          | Upper Limit | 17.428          | 59.480         |
| Student 17| Lower Limit | 83.927          | 83.610         |
|          | Median      | 89.224          | 90.670         |
|          | Upper Limit | 94.369          | 97.751         |
### Table 9 Overall performances and interesting cases for MATH 2240: Differential Equations

| Index        | Statistics | COVID-19 | Before  | After  |
|--------------|------------|----------|---------|--------|
| Overall      | Lower Limit| 60.021   | 71.390  |
|              | Median     | 65.683   | 76.214  |
|              | Upper Limit| 71.397   | 81.067  |
| Student 14   | Lower Limit| 58.864   | 77.164  |
|              | Median     | 59.203   | 77.717  |
|              | Upper Limit| 59.773   | 78.204  |
| Student 21   | Lower Limit| 39.808   | 68.417  |
|              | Median     | 43.660   | 73.472  |
|              | Upper Limit| 49.035   | 77.973  |
| Student 29   | Lower Limit| 93.213   | 93.281  |
|              | Median     | 94.972   | 94.961  |
|              | Upper Limit| 96.569   | 96.903  |

### Table 10 Overall performances and interesting cases for STAT 2000: Probability and Statistics

| Index        | Statistics | COVID-19 | Before  | After  |
|--------------|------------|----------|---------|--------|
| Overall      | Lower Limit| 59.315   | 62.279  |
|              | Median     | 65.653   | 67.675  |
|              | Upper Limit| 72.045   | 73.220  |
| Student 5    | Lower Limit| 51.056   | 65.219  |
|              | Median     | 51.849   | 66.706  |
|              | Upper Limit| 52.668   | 68.160  |
| Student 12   | Lower Limit| 89.326   | 77.278  |
|              | Median     | 90.761   | 79.874  |
|              | Upper Limit| 91.950   | 82.403  |
| Student 22   | Lower Limit| 52.412   | 77.089  |
|              | Median     | 53.504   | 79.135  |
|              | Upper Limit| 54.587   | 80.874  |

### Table 11 Overall performances and interesting cases for ARET 1400: Civil Technology 1

| Index        | Statistics | COVID-19 | Before  | After  |
|--------------|------------|----------|---------|--------|
| Overall      | Lower Limit| 82.387   | 67.281  |
|              | Median     | 85.131   | 72.933  |
|              | Upper Limit| 87.899   | 79.046  |
| Student 6    | Lower Limit| 91.180   | 86.049  |
|              | Median     | 92.351   | 88.559  |
|              | Upper Limit| 93.529   | 90.766  |
| Student 12   | Lower Limit| 84.014   | 85.621  |
|              | Median     | 84.843   | 87.205  |
|              | Upper Limit| 85.634   | 88.729  |
| Student 29   | Lower Limit| 84.563   | 51.266  |
|              | Median     | 86.614   | 55.863  |
|              | Upper Limit| 88.573   | 60.593  |
Table 12 Overall performances and interesting cases for ARET 2600: Statics and Strength of Materials

| Index   | Statistics | Before | After |
|---------|------------|--------|-------|
| Overall | Lower Limit| 74.795 | 67.592|
|         | Median     | 78.120 | 73.199|
|         | Upper Limit| 80.904 | 78.190|
| Student 4 | Lower Limit| 78.975 | 76.232|
|         | Median     | 79.537 | 77.069|
|         | Upper Limit| 80.089 | 77.871|
| Student 12 | Lower Limit| 72.386 | 86.270|
|         | Median     | 73.467 | 88.037|
|         | Upper Limit| 74.548 | 89.737|
| Student 15 | Lower Limit| 73.177 | 92.305|
|         | Median     | 74.522 | 55.779|
|         | Upper Limit| 75.826 | 59.157|

Table 13 Overall performances and interesting cases for COMP 2680: Web Site Design and Development

| Index   | Statistics | Before | After |
|---------|------------|--------|-------|
| Overall | Lower Limit| 89.318 | 84.949|
|         | Median     | 91.591 | 88.953|
|         | Upper Limit| 93.654 | 92.900|
| Student 16 | Lower Limit| 93.964 | 94.873|
|         | Median     | 94.950 | 95.962|
|         | Upper Limit| 95.988 | 97.036|
| Student 18 | Lower Limit| 97.425 | 82.524|
|         | Median     | 97.869 | 83.006|
|         | Upper Limit| 98.249 | 83.491|
| Student 22 | Lower Limit| 90.997 | 56.612|
|         | Median     | 92.492 | 58.664|
|         | Upper Limit| 94.024 | 60.570|

Table 14 Overall performances and interesting cases for COMP 4980: Introduction to Bioinformatics

| Index   | Statistics | Before | After |
|---------|------------|--------|-------|
| Overall | Lower Limit| 80.916 | 58.211|
|         | Median     | 83.927 | 64.378|
|         | Upper Limit| 86.774 | 70.547|
| Student 8 | Lower Limit| 97.972 | 95.892|
|         | Median     | 98.401 | 96.444|
|         | Upper Limit| 98.815 | 97.033|
| Student 11 | Lower Limit| 77.197 | 35.682|
|         | Median     | 81.164 | 40.692|
|         | Upper Limit| 85.171 | 45.960|
| Student 19 | Lower Limit| 88.281 | 88.216|
|         | Median     | 90.469 | 91.361|
|         | Upper Limit| 92.449 | 94.331|
References

1. Arias, J.J., Swinton, J., Anderson, K.: Online vs. face-to-face: A comparison of student outcomes with random assignment. e-Journal of Business Education and Scholarship Teaching 12(2), 23 (2018). URL https://files.eric.ed.gov/fulltext/EJ1193426.pdf

2. Ashby, J., Sadera, W.A., McNary, S.W.: Comparing student success between developmental math courses offered online, blended, and face-to-face. Journal of Interactive Online Learning 10(3), 10 (2011). URL https://www.ncolr.org/jiol/issues/pdf/10.3.2.pdf

3. Basilaia, G., Kvavadze, D.: Transition to online education in schools during a SARS-CoV-2 Coronavirus (COVID-19) pandemic in Georgia. Pedagogical Research 5, 1–9 (2020). DOI 10.29333/pr/7937

4. Bhaskaran, K., Smeeth, L.: What is the difference between missing completely at random and missing at random? International Journal of Epidemiology 43(4), 1336–1339 (2014). DOI 10.1093/ije/dyu080

5. Cao, W., Fang, Z., Hou, G., Han, M., Xu, X., Dong, J., Zheng, J.: The psychological impact of the covid-19 epidemic on college students in china. Psychiatry Research 287(112934) (2020). DOI https://doi.org/10.1016/j.psychres.2020.112934

6. Celton, M., Malpertuy, A., Lelandais, G., Brevern, A.G.d.: Comparative analysis of missing value imputation methods to improve clustering and interpretation of microarray experiments. BMC Genomics 11(15) (2010). DOI https://doi.org/10.1186/1471-2164-11-15

7. Dennon, A.: Coronavirus impacts on students and online learning (2020). URL https://www.bestcolleges.com/blog/coronavirus-impacts-on-students/

8. Doreleyers, A., Knighton, T.: COVID-19 pandemic: Academic impacts on postsecondary students in canada. Statistics Canada (2020). URL https://www150.statcan.gc.ca/n1/pub/45-28-0001/2020001/article/00015-eng.htm

9. Geyer, C.J.: Practical markov chain monte carlo. Statistical Science 7(4), 473–483 (1992). URL https://www.jstor.org/stable/2246094

10. Jones, S.J., Vena, M.L.: Learning equity between online and on-site mathematics courses. MERLOT Journal of Online Learning and Teaching 9(1), 1–12 (2013). URL https://jolt.merlot.org/vol9no1/jones_0313.pdf

11. Kleiner, C., Lord, M.: The Cheating Game: ‘Everyone’s doing it,’ from grade school to graduate school. U.S. News and World Report 127, 54–63 (1999)

12. Laird, N., Lange, N., Stram, D.: Maximum likelihood computations with repeated measures: Application of the EM algorithm. Journal of the American Statistical Association 82(397), 97–105 (1987). DOI https://doi.org/10.2307/2289129. URL http://www.jstor.org/stable/2289129

13. Laird, N.M., Ware, J.H.: Random-effects models for longitudinal data. Biometrics 38(4), 963–974 (1982). DOI https://doi.org/10.2307/2529876. URL http://www.jstor.org/stable/2529876

14. Lawrence, C., Altschul, S., Boguski, M., Liu, J., Neuwald, A., Woottton, J.: Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. Science 262(5131), 208–214 (1993). DOI https://doi.org/10.1126/science.8211139. URL https://science.sciencemag.org/content/
15. Lindstrom, M.J., Bates, D.M.: Newton-raphson and EM algorithms for linear mixed-effects models for repeated-measures data. Journal of the American Statistical Association 83(404), 1014–1022 (1988). DOI https://doi.org/10.2307/2290128. URL http://www.jstor.org/stable/2290128
16. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2020). URL https://www.R-project.org
17. Rey, J.G.: The effects of online courses for student success in basic skills mathematics classes at california community colleges. Ph.D. thesis, University of Southern California (2010). URL https://search.proquest.com/openview/8872a71d22481e7919cde71eda5992/1?pq-origsite=gscholar&cbl=18750&diss=y
18. Sahu, P.: Closure of universities due to coronavirus disease 2019 (covid-19): Impact on education and mental health of students and academic staff. Cureus 12(4)(e7541) (2020). DOI 10.7759/cureus.7541. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7198094/
19. Schafer, J.L., Yucel, R.M.: Computational strategies for multivariate linear mixed-effects models with missing values. Journal of Computational and Graphical Statistics 11(2), 437–457 (2002). DOI https://doi.org/10.1198/106186002760180608. URL https://doi.org/10.1198/106186002760180608
20. Sintema, E.: Effect of covid-19 on the performance of grade 12 students: Implications for STEM education. EURASIA Journal of Mathematics, Science and Technology Education 16(7), em1851 (2020). DOI https://doi.org/10.29333/ejmste/7893
21. Weems, G.H.: Comparison of beginning algebra taught onsite versus online. Journal of Developmental Education 26(1), 10 (2002). URL https://www.learntechlib.org/p/95604/
22. Williams, J.B., Wong, A.: The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. British Journal of Educational Technology 40, 227 – 236 (2009). DOI https://doi.org/10.1111/j.1467-8535.2008.00929.x
23. Wong, E.T.: Impact of E-learning during SARS outbreak in Hong Kong. The social and cognitive impacts of e-commerce on modern organizations pp. 185–203 (2004). DOI 10.4018/978-1-59140-249-7.ch009
24. Zhao, Y., Staudenmayer, J., Coull, B.A., Wand, M.P.: General design bayesian generalized linear mixed models. Statistical Science 21(1), 35–51 (2006). DOI https://doi.org/10.2307/27645735. URL http://www.jstor.org/stable/27645735
25. Zu, Z.Y., Jiang, M.D., Xu, P.P., Chen, W., Ni, Q.Q., Lu, G.M., Zhang, L.J.: Coronavirus disease 2019 (covid-19): A perspective from china. Radiology 0(0), 200490 (2020). DOI 10.1148/radiol.2020200490