Volume 21 • Issue 2 (2018)

Empirical Results using Learning Analytics in the Classroom

Terry James

Abstract

The purpose is to improve insights and educational results by applying analytic methods. The focus is on the mathematics applied to learn from the kind of data available to most classes such as final examination marks or homework grades. The sample is 249 students learning introductory college statistics. The result is a predictive model for student success. final grade = 36.8 + 2.26(complete). The effect ratio is 0.43 at a 95% level of significance. Difficulty ratios were calculated from homework practice questions so professors can see, prior to the first test, what topics and levels of difficulty within

Terry James

Terry James has 16 years experience as a faculty member at Seneca College. He has a doctorate degree in business, master's in science, and bachelor's in philosophy. His doctorate dissertation examined small business innovation and commercialization of research at colleges. Terry lead information

a topic need review. The type of over-practice needed to deliver the most improvement is identified. Dashboards to drill down on findings are discussed.

Keywords: learning analytics, predictive model, difficulty ratio, student success, dashboard, data mining, analytics.

technology projects in financial services for 20 years. Terry is a certified project management professional. He has presented regularly at the OCMA (Ontario Colleges Mathematics Association).

Empirical Results using Learning Analytics in the Classroom

Analytics is a method using supporting technologies to find hidden patterns in data. Business leaders increasingly utilise analytics for improved decisions. Almost every client purchase, web click, phone call, or social media post creates data. Any device connected to the web provides data about clients. More data provides businesses with more opportunities to engage with customers. As more data becomes available from learning activities, educational stakeholders can experiment with data mining and analytic technologies to improve educational decisions. As educational organisations move resources online, systems capture increasing amounts of learning data. Most Learning Management Systems (LMS), for example, Blackboard, can capture which students are doing homework questions and summarise the results. The LMS can also log how often a posted document such as lecture video is accessed by students.

Learning analytics puts a focus on predicting and improving student performance by looking at patterns in the learning data. Instead of calculating basic metrics such as the proportion of students who passed the course or earned a good grade, we can do more sophisticated calculations. Who is at risk, in what chapter, in what topic, and at what level of topic difficulty? Which student needs help but is too shy to ask? Does a professor need to re-teach a topic for the whole class or just four students? How should a professor intervene in the most efficient way before the first test? Can we personalise feedback? Can we adapt to individual needs more effectively? Are we too late if we wait for the first test? If we plan to improve our teaching materials, what topics need the most attention? Learning analytics, using different mathematical calculations, can show sometimes surprising and counter-intuitive findings about hidden patterns that may answer these types of questions.

As background to this study, a statistics course was chosen as the candidate course for this research because of the high student failure rate in statistics. The course initially used traditional classroom lectures with a student success rate (passing grade) of about 50%. The course now posts all teaching materials online. Sixty video lectures are posted online so students can learn anywhere or any time using a computer or smartphone. The textbook is online. All practice questions are online. Students move through the topics at their own pace. Practice questions are personalised to student ability. The system adjusts the difficulty level of questions dynamically from one to seven levels as the student works. Feedback on questions is immediate. With more materials posted online, there is more data on homework, grades, and access to materials.

With learning material online, the teaching method evolved. A traditional classroom lecture delivery method evolved into a hybrid, flipped classroom. Students watch online lectures at home and do practice problems in class. The idea is to watch easier or introductory examples at home and do the difficult problems in class. The professor moves from desk to desk whenever a student needs help. As a self-paced course, every student has a challenging problem for their level. The professor lectures spontaneously for a few minutes based on the types of classroom questions. Dashboard reports give the professor real-time information about the level of class proficiency in a topic. The chart below shows the improvement using the personalised, self-paced, adaptive, flipped classroom approach. Most LMS systems can collect data on access to learning materials such as homework problems or final examination grades.

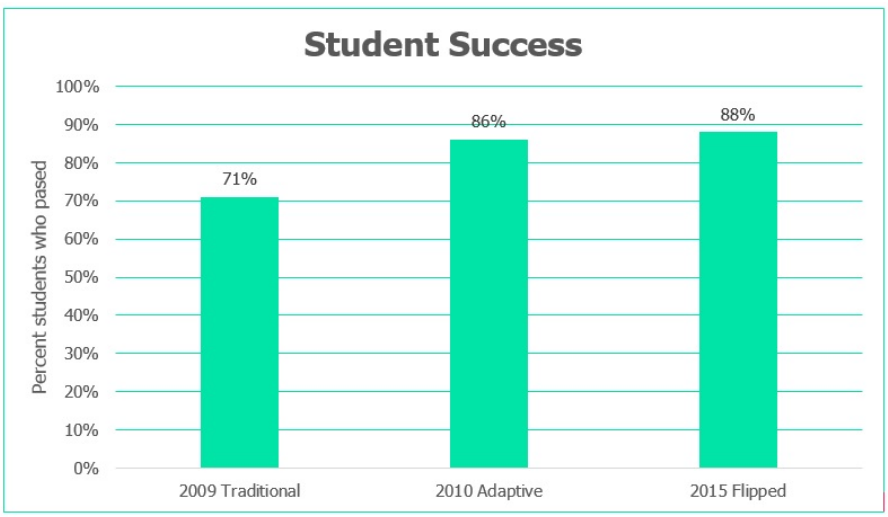


Figure 1: Background on teaching methods as the starting point for the analytics study.

Literature Review

Learning analytics is the measurement, collection, analysis, and reporting of patterns from the data to understand and optimise learning (Moissa, Gasparini & Kemczinski, 2015; ElAtia, Ipperciel & Hammad, 2011). The data is the evidence (Ubell, 2016). Researchers look for patterns in data, build understandings of the patterns, predict what follows, and take needed actions (Bienkowski, Feng & Means, 2012). Using analytics, professors can continuously improve and develop models and algorithms to provide insights that allow institutions to enhance learning.

Purposes listed in the literature for learning analytic studies are: prediction, models, personalisation, monitoring, assessment, adaptation, and reflection (Moissa, et al., 2015). Document goals for learning analytics are employed such as prediction of future learning, improving the subject domain model, optimising instructional methods, studying the impact on pedagogical approaches, and the advancement of the science of learning (AlShammari, Aldhafiri, and Al-Shammari, n.d.).

Data captured for analytic studies may include demographics such as age, gender, social economic status, and location (Ramaswami & Bhaskaran. 2010). Other important variables are the frequency of work, number of logons, time spent, and content visited such as web pages, videos watched, and problems attempted. Assessment details and interventions through the system such as performance alerts are logged. The decision to include data capture should be based on the actionable potential to improve learning.

Some of the techniques for learning analytics include multi-regression and random forests (Sweeney, Rangwala, Lester & Johri, 2016). Other techniques are CHAID, decision tree, classification, clustering, Bayesian networks, Naïve Bayes, and Support Vector Machine (Agarwal, Pandey & Tiwari, 2012). Papamitsiou and Economides (2014) include neural networks and association rule mining techniques. There is no standard of practice in learning analytics on the types of techniques to use or the format of data. There is no best technology. The plethora of analytic techniques may indicate the field of learning analytics is new so standards of practice by educators are in flux. Most learning analytics work appears to be published by computer scientists rather than educational scientists (ElAtia, et al., 2011). The analysis of data by most teachers is limited to basic statistics such as the average mark, standard deviation of grades, or proportion of students who pass a course (Greller & Dashsler, 2012).

Many courses include learning analytics to personalise lessons, so each student sees different learning material according to need. The order and pace of learning vary for each student. The adaptive system responds based on student ability and need. If a student knows a topic, they can skip it or move much faster through the material (Bienkowski, et al., 2012). Hints are provided where and if needed or requested. The data captured allows a feedback loop to provide for adaptive learning and personalization.

The knowledge gaps across classes or for individual students are identified immediately. What topic is the biggest struggle for a class or student to learn (Young, Blumenstyk, Hill & Feldstein, 2016)? Dashboards can report trends and teachers can drill-down in the report to different levels of detail as needed. The need for intervention for at-risk students can be reported dynamically to teachers and students as needed. Learning analytics allows foresight instead of hindsight (Educause, 2016). Rather than wait for the first test, the system flags at-risk students in the first few weeks.

Feedback loops can be fully automated, semi-automated, or manual (Moissa, et al., 2015). Students and professors can reflect and increase awareness of performance levels and expectations as reports and patterns emerge.

Prediction is a key goal of learning analytics. The system can predict student success rates and provide recommendations. We can provide models of learning that act as an early warning system to students and provide insight to possible pedagogical changes. Misconceptions can be identified, and educational system changes can be validated as data trends emerge (Campo-Avila, Conejo, Triguero & Moales-Bueno, 2015). The prediction model combined with data visualisations can help identify patterns, exceptions, and insights into teaching. All of this information is available before the final examination. For the types of questions learning analytics can address, Bienkowski, Feng, and Means (2012, pp.12) provide some suggestions: “what will predict student success”, “what student actions indicate more engagement and learning progress”, and “what features of an online learning environment lead to better learning?”

Method

The study had 249 students across eight different classes who took an introductory statistics course. The same professor taught all participants. Only students who completed at least one test and accessed online learning materials at least once, were included in the sample. Each class was 14 weeks of classroom and online instruction.

The learning materials were online. Each chapter consists of many topics. There were no surveys and no collection of data such as demographics (for example age, gender, postal codes). The study used only basic homework performance, access to online materials, and final grades. Data capture was at the topic level of granularity. Any access to the textbook, video lectures, or practice questions for a topic was logged using standard LMS features. There were 11 chapters and 34 topics.

Practice questions have from one to seven levels of difficulty depending on the complexity of the topic. For every homework question, for every student, a log of the topic, level of difficulty, correct or incorrect answer was summarised. A simple count of correct or wrong answers in a row and total number of questions at the level was calculated. The difficulty ratio was calculated as the total number of correct answers divided by the total questions attempted.

The LMS (growingknowing.com) marked a topic as complete if the student answered three practice questions correctly in a row at the most difficult level for the topic. The concept of completing a topic was important as a method for the LMS to measure what topic should be provided to students. The LMS provided dashboard reports that showed professors the number of completed topics summarised by class. The professor drilled down from class data to show topics completed by individual students. Once the course was complete, all individual identifiers were removed, and only summary data was kept. This research used only aggregated data with no individual identifiers available.

All summaried data from the database was ported into Microsoft Excel. Excel charts and pivot tables were used to gain different perspectives of aggregated data. Excel was used to generate descriptive statistics. Multiple regression was used to create a predictive model.

Results

High-level measures

As a high-level measure of engagement for the 249 students in the sample, the total for textbook views was 12,821, video lecture views 10,444, and practice problems attempted was 85,999.

As a measure of student success, the mean final grade was 67% (C+), the median grade was 68%. Eight-eight percent of students passed the course. The highest grade was 98%, and lowest was zero. Looking at Table 1, the students with higher grades are doing more practice problems, completing more topics, and reading the textbook more. The completion of a topic is defined as a student correctly answering three questions in a row at the highest level of difficulty for a topic. Top students did not access the video lectures as often as C students. The pattern below is a high-level view. If B and D grade students are included, some overlap on video and textbook access makes the pattern more difficult to see.

If we take a view of the data across 31 topics by student, the average for textbook access per student was 47 and median was 35. With 31 topics measured, a student needs to access some topics multiple times to obtain an average higher than the 31 topics available. A student may want to review the textbook before a test or when working on a practice problem. The maximum count for textbook access was 206, and the minimum was zero. The mean access to video lectures was 41, the median was 29, the maximum was 185, and the minimum was zero. Some students accessed the same video multiple times.

Table 1. Grades and Average Topic Access of Learning Material per Student

Grade	Book	Lecture	Practice	Complete	Average Grade
A	53	24	356	33	85%
C	49	44	358	32	65%
F	21	19	173	20	38%

The average number of practice problems per student was 319, median 293, the standard deviation was 144, maximum was 800, and the minimum was two. Every question given to a student is unique by varying data or text. With billions of possible questions, the probability of the same question twice to the same student is remote. The system is adaptive by varying the level of question difficulty based on the student ability.

Ease Ratio

We calculate an ease ratio by counting the number of correct answers divided by the total number of questions attempted (Frey, 2006). The higher the ease ratio score, the easier the question. The ease ratio is calculated per student for every topic and level of difficulty. The average of ease ratios calculated by topic for 249 students is shown in Table 2 with the most difficult topics listed first. Students found the topics of hypothesis testing and percentiles to be the most difficult topics in the course. Hypothesis testing and percentiles were equally difficult for students. A test for independence in probability calculations was the next most difficult.

Table2. Average Ease Ratio by Topic

Topic	Average ease ratio by topic
Hypothesis	.62
Percentile	.62
Independence	.66
Discrete Variance	.69
Binomial distribution	.73

Results on ease ratios for specific students can be shown to see exactly who and where an intervention is needed. The ease ratio can be run before the first graded test. This paper cannot publish individual results for privacy and ethical reasons.

Table 3 shows a summary level breakdown of average ease ratios by level of difficulty and topic. The percentile level 1 difficulty uses less data and a simpler calculation than level 2. Using the percentile example, for level 1 percentile, the student must find which data item in a provided list is in the requested percentile position. For level 2 difficulty, more data is given, the calculation includes taking an average, and the answer may not be a supplied data item. The easier technical question is found to be more difficult by students given the ease ratio. The ease ratio in table 3 shows binomial and hypothesis questions have a similar finding where easier technical calculations are found to be more difficult by students.

Table3. Ease Ratio by Problem Levels of Difficulty

Topic	Level 1	Level 2	Level 3
Percentile	.52	.73	
Binomial	.76	.67	.73
Hypothesis Proportion	.53	.69	.66

Predictive analytics

Multiple regression was used to create a predictive model. Independent variables included total-textbook reading, total-video lectures watched, total practice problems attempted, and a total number of topics completed by the student. The dependent variable was the final grade.

Correlation between the number of practice problems and final grade was statistically significant at the 95% confidence level. The coefficient of determination was 18%, and the p-value was 0.0000. The number of practice problems attempted by a student does determine the grade for the course.

Correlation between the number of topics completed and final grade was statistically significant at the 95% confidence level. The coefficient of determination was 43%, and the p-value was 0.0000. Results showed the number of completed topics is more important than the number of practice problems attempted in explaining the final grade. A student can complete a topic by doing every level of difficulty up to the highest level of difficulty for the topic and then answer three questions correctly in a row at the most difficult level. Results show the number of practice questions attempted is not as important as the levels of difficulty completed.

The predictive equation is grade = 36.8 + 2.26(complete). A simple equation with only the complete variable shows the same effect size of .43 when using both the total practice and the complete variable We effect size is .43.

A correlation between the total number of textbook readings and final grade was not statistically significant. A scatter diagram did not show a linear relationship between the two variables. The coefficient of determination was only three percent. The correlation with the total number of video lectures watched showed no effect size. The scatter diagram showed no linear or non-linear relationship. Further calculations using logistic regression with a pass-fail categorical grade, trimming 10% of the outliers, creating a book plus video index variable, and regression within grade clusters (A, B, C, D, and F) showed no effect size.

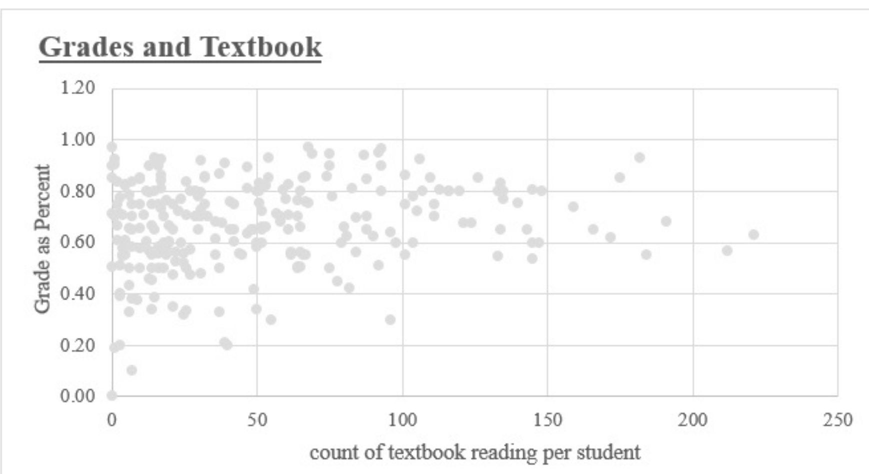


Figure 2: No linear relationship between grades and textbook reading. r2 = 0.03

Dashboard reports are available in the system, and each professor can run reports in real time. The reports show how many students have completed each topic. The LMS (growingknowing.com) permits the professor to drill down in real time to see the number of completed topics for individual students. Another report can send an automated alert to any student behind more than seven topics. The professor can dynamically alter the number of topics behind, the time allowed to catch up, and the maximum number of alerts that can be sent to any student. There is also a High Score report. Each time a student completes a question, points are awarded based on the difficulty and topic. Students compete for a high score with classmates. Since the high score is published online for anyone in the class to see, students use nicknames if they want privacy instead of fame. Results show some students find the high score highly motivating. No grades are given for high score, but some students do as many as 350 questions on a single topic to earn a better high score. For privacy reasons, no individual name can be published. Only the professor who needs access to individual student results in order to provide grades can see individual students results.

Since the student success rate is high at 88%, the sample for the alert report is too small to provide more than anecdotal evidence that alerts improve student success. Anecdotal results show about 50% of students who are failing and sent an alert, make a written commitment to work harder and eventually pass the course.

Discussion

The measure of 12,821 textbook views, 12,821 lecture views, and 85,999 practice problems for 249 students is a high-level indicator that students found value and were engaged in the learning material. The C+ 67% final grade shows the statistics course is not easy for most students. The 88% student success rate is outstanding given most colleges and universities have a 50% pass rate for mathematics courses such as statistics.

The finding that students earning higher grades are reading more and doing more practice problems is not surprising, but the video lectures views do not increase linearly as grades improve. Students can access materials at their own pace and in any order. If a student prefers to learn by reading instead of watching movies, that is a student choice.

The total number of practice problems per student has an average of 319 and median of 293 per student. In a traditional lecture delivery with a class of 30 students and no system automation, a professor have 9,000 questions to grade (30 × 300) which is impossible for one professor using pen and paper. The unique question generation, instant feedback, and adaptive ability of the system to provide questions with a difficulty-level matching student ability likely drives a higher success rate. Predictive modelling did confirm that more practice questions improve student success.

Ease ratio

The average ease ratio by topic shows which topics are a challenge for students. In a course without learning analytics, a professor would see how a class did on a test for each topic. The professor could only guess how much work was needed to learn a topic. The ease ratio shows exactly how many questions were answered incorrectly and can be studied at the class, topic, or individual student level of granularity. The ease ratio by topic informs the professor about what topics should be given more focus in lectures and learning the material to make the most difference. The ease ratio also showed some surprising results when averaged by the level of difficulty per topic. Some levels of difficulty that were mechanically easier to answer were found to be more difficult by the student. Easier questions that are harder for students is counter-intuitive. A challenge in grasping the concept of a percent versus percentile calculation may be harder to grasp than a slightly more difficult calculation. Once the student understands the concept, the progress through levels of difficulty is relatively easier than getting the first question correct.

Predictive model

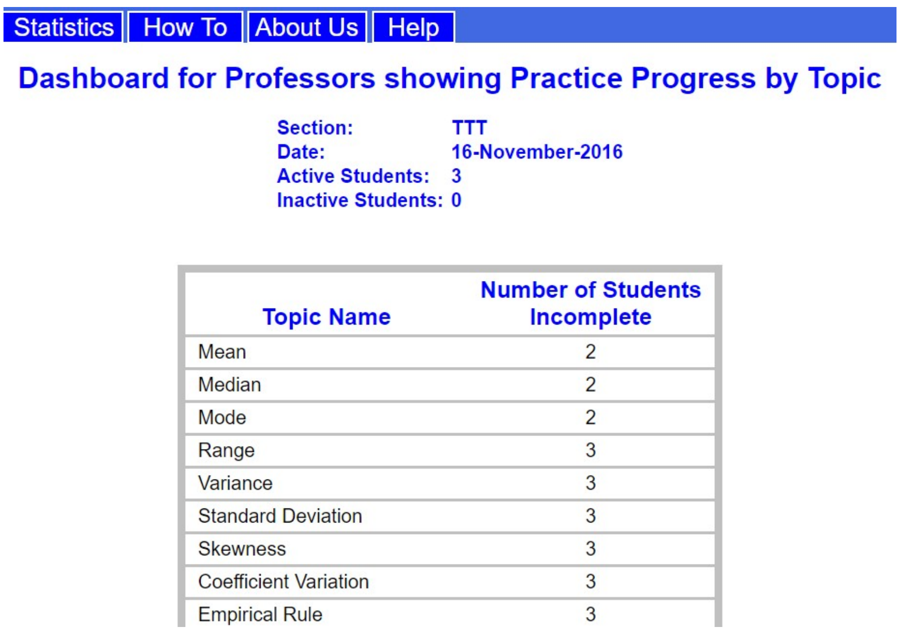
The multiple regression calculation for final grade as predicted by the number of practice questions shows some over-practice does improve student success in the course with an effect size of 0.18 (using the coefficient of determination for effect size). The finding is the quantity of over-practice is not as important as the quality of over practice. Some topics that are completed show the range of topics and difficulty levels of questions is more important than gross quantity of questions studied. The effect size for completing topics was .43, much higher than .18 for quantity of questions. Completing topics measures the number of topics studied and the level of difficulty for each topic. A simple predictive model was found: Grade = 36.8 + 2.26(complete). The same effect size is found whether we include total problems studied and topics completed or just topics completed.

There was no linear relationship between final grade and count of textbook access or video lecture access. Numerous analyses such as trimmed outliers, logistic regression, book and video index, and other methods showed little to no impact between reading, lectures, and final grades. The finding may surprise some professors who intuitively believe attendance at lectures and reading the textbook has some impact in determining a student grade. In a traditional course, assumptions are made that a high grade implies work was done reading the book and attending the lectures. In learning analytics, the researcher measures reading and lecture views and correlates forward to final grade. Learning analytics knows every page read, how often it was read, and when it was read.

The learning system used for this study does not measure entry level knowledge. A controlled experiment could perform a pre-assessment, students read a topic in the textbook, and then complete a post-assessment. A controlled experiment would provide more insight, but the system used in this study is personalised and adaptive to each student. A personalised, self-paced, and adaptive system makes for a difficult environment to perform a highly-controlled and standardised experiments. Also, the adaptive system does not measure outside resources. A student learning from other books, classmates, tutors, or conversations with the professor are not measured. The lack of correlation with reading and grades may be due to the influence of variables not measured in this study. This study uses regression. Analytics includes methods such as decision trees with random forests, Bayesian methodologies, and neural networks that may reveal relationships between the variables such as textbook reading and grades that are not available from regression studies.

Dashboard

Several dashboard reports allow the professor to see classroom performance much as we can look at the gauges on a car to understand critical aspects of car performance such as speed or gas reserves. The dashboard provides an overview of how many topics the class has completed, but the professor can drill down to individual student performance. A quick look at the dashboard tells the professor if the class is behind or ahead of schedule. The professor can see which student needs help, on what topic, and what difficulty level even if the student has not asked for help. Dashboard reports include a High Score report to gamify practice problem study.



Statistics | How To | About Us | Help

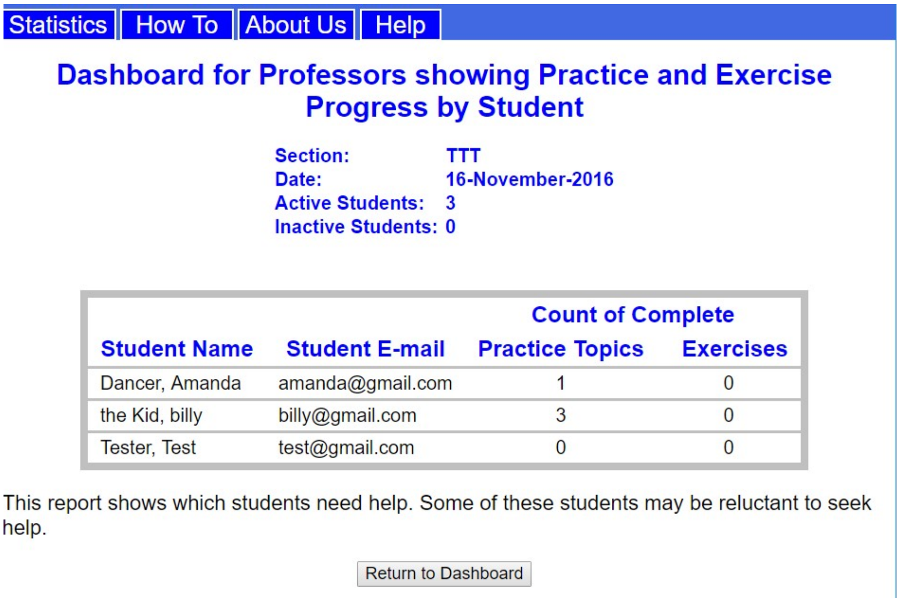
Dashboard for Professors showing Practice Progress by Topic

Section: TTT
Date: 16-November-2016
Active Students: 3
Inactive Students: 0

Topic Name	Number of Students Incomplete
Mean	2
Median	2
Mode	2
Range	3
Variance	3
Standard Deviation	3
Skewness	3
Coefficient Variation	3
Empirical Rule	3

Figure 3: Dashboard showing class performance. (Sample only – data is not real).

The High Score report is motivating to some students. There is an Alert Report that goes automatically to students at risk. The sample size for students getting alerts is too small to state performance with statistical significance. Anecdotally, about half the students who get an alert move from a likely failure to passing with a low grade. A sample of dashboard reports are shown below, but the data is a fictitious sample for privacy protection.



Statistics | How To | About Us | Help

Dashboard for Professors showing Practice and Exercise Progress by Student

Section: TTT
Date: 16-November-2016
Active Students: 3
Inactive Students: 0

		Count of Complete	
Student Name	Student E-mail	Practice Topics	Exercises
Dancer, Amanda	amanda@gmail.com	1	0
the Kid, billy	billy@gmail.com	3	0
Tester, Test	test@gmail.com	0	0

This report shows which students need help. Some of these students may be reluctant to seek help.

Return to Dashboard

Figure 4: Dashboard report by Student. (Sample only – data is not real)

Conclusion

Learning analytics is used in the classroom for a statistics course. The data provides information about class performance and individual performance prior to any tests. A predictive equation is provided. The data proves that doing many practice problems is key to success. The quantity of problem practice has an effect, but the quality of practice in topic scope and a range of question difficulty is more important for high grades. Ease ratios are used to direct teaching material improvements and provide insight into where students find difficulty. Some of the ease ratio results around question difficulty were counter-intuitive. Samples of dashboard reports allows the professor to have insights into student performance in real time prior to any testing. Student success in passing a subject that is notoriously difficult in most schools is 88% using the methods described in this paper.

References

Al Smammari, I., Aldhafiri, M., & Al-Shammari, Z. (n.d.). A meta-analysis of educational data mining on improvements in learning outcomes. *College Student Journal*, 326-332.

Agarwal., S., Pandey, G., & Tiwari, M. (2012). Data mining in education: Data classification and decision tree approach. *International Journal of e-Education, e-Business, and e-Learning*, *2*(2), 140-144.

Bienkowski, M., Feng, M., & Means, B. (2012). Enhanced teaching and learning through educational data mining and learning analytics: An issue brief. US Department of Education. Retrieved December 21st, 2016, from www.ed.gov/technology

Campo-Avila, J., Conejo, R., Triguero, F., & Morales-Bueno, R. (2015). Mining web-based educational systems to predict student learning achievements. *International Journal of Artificial Intelligence and Interactive Multimedia*, *3*(2), 49-53.

Educause. (2016). |Horizon report: 2016 higher education edition. Austin, Texas: New Media Consortium. Retrieved February 5, 2017 from www.nmc.org.

ElAtia, S., Ipperciel, D., & Hammad, A. (2011). Implications and challenges to using data mining in educational research in the Canadian context. *Canadian Journal of Education*, *35*(2), 101-119.

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational technology & Society*, *15*(3), 42-57.

Jones, S. (2012, Spring). Technology review: the possibilities of learning analytics to improve learner-centered decision making. *The Community College Enterprise*, 89-92.

Kumar, V., & Chadha, A. (2012). Mining association rules in student's assessment data. *International Journal of Computer Science*, *9*(5), 211-216.

Moissa, B., Gasparni, I., & Kemczinski, A. (2015). A systematic mapping on the learning analytics field and its analysis in the massive open online courses context. *International Journal of Distance Education*. *13*(3) 1-24.

Papamitsiou, Z. & Economides, A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society*, *17* (4), 49-64.

Ramaswami, M., & Bhaskaran, R. (2010). A CHAID based performance prediction model in educational data mining. *International Journal of Computer Science*, *7*, 10-18.

Sweeney, M., Rangwala, H., Lester, J., & Johri, A. (2016). *Next-term student performance prediction: A recommender system approach*. National Science Foundation.

Ubell, R. (2016, December 11). The good news about learning by the numbers. *The Chronicle of Higher Education*, 1-4.

Young., J., Blumenstyk, G., Hill, P., & Feldstein, M.

(2016). What you need to know about learning analytics.
The Chronicle of Higher Education, Focus, 1-24.

SenecaPress