

Learning and Academic Analytics in the Realize^{it} System

Colm Howlin
CCKF Limited
Dublin, Ireland
colm.howlin@cckf-it.com

Danny Lynch
CCKF Limited
Dublin, Ireland
danny.lynch@cckf-it.com

Abstract: Analytics allows the discovery and communication of patterns and relationships hidden within data. In the context of education, capturing deeply granular data offers educators the potential to gain previously unobtainable insights into the learning process and allows them to make data-informed decisions. This paper discusses the analytics made available through the Realize^{it} adaptive learning platform. These can largely be split into two groups; *Learning Analytics* which is concerned with the ongoing learning process and *Academic Analytics* which focuses on institutional level insights. In particular the latter will be emphasized by presenting a series of modular outputs from the analytics being run on a real English Composition course delivered through Realize^{it}. These modular components will highlight some of the potential insights that are available from educational data.

Introduction

Over the last decade, we have witnessed an unprecedented growth in internet technology services and devices that provide ubiquitous connections to them. Coupled with the ever diminishing cost of data storage, it is now feasible for the billions of data points we generate to be stored and analyzed (Buckingham Shum 2012). With the large push for increased efficiency and reform in higher education, it has been natural to pivot our learning to new digital formats. This has certainly led to an explosion in the amount of educational data we now capture; data such as activity streams, interactions of learners with their peers & mentors, administration records, and many other rich data trails (Long & Siemens 2011). Although these data sources prove useful in themselves, it is the analytics that they inform which has educators excited. Analytics provides a new model for university leaders to improve teaching, learning, organizational efficiency and decision making (MacNiell et al. 2014).

The fundamental approach behind the Realize^{it} adaptive learning platform is the separation of the curriculum from the content. However placing a learner in a situation where they must both navigate a curriculum and select content adds enormous cognitive load. An instructor avoids this by making decisions and suggestions to the learners as they learn. Realize^{it} seeks to emulate this process and utilize data-driven decisions to bridge the gap between the curriculum, content, learner and instructor (Howlin 2013). The system supports all elements of the learning process including curriculum and content creation, delivery of learning and data capture. This paper will focus on the contributions to analytics that has been implemented in the Realize^{it} system.

Data is the critical element. Realize^{it} continuously learns about each student; it tracks progress and attainment, identifies knowledge strengths and weaknesses and in turn provides a personalized and adaptive learning experience to each individual. The platform has been designed from inception to generate key data points which go beyond the traditional summative metrics. Rather than requiring data for retrospective reports and analysis, the platform has been built so that in real-time it integrates the use of that data, and the associated analytics, into the learning process to support both the learner and the instructor. This type of analysis is known as Learning Analytics (Ferguson 2012) will be discussed briefly in the section *Learning Analytics in Realize^{it}*.

Academic analytics is the application of business intelligence methodologies and strategies to guide decision-making practices in educational institutions (van Barneveld 2012). In Realize^{it} the deeply granular learner data is available for any analysis that an institution might require. This drives insights and discovery of new patterns, relationships and features at any level, from the individual learner or instructor to course level metrics and institutional level analysis. The main focus of this paper we will be on the course level analytics and will be detailed, along with associated visualizations representing the insights, in the section *Academic Analytics in Realize^{it}*. We will examine the academic analytics generated from the data of an English Composition course run through the Realize^{it} system at a North American higher education institution. The data was collected over 13 successive instances of the course and includes 14,832 enrollments of which 12,067 are unique individuals.

Learning Analytics in Realize^{it}

The core of the Realize^{it} system is the Adaptive Intelligence Engine. This engine is built from many mathematical and statistical models, and algorithms which are tasked with learning, understanding and adapting to specific parts of the learning environment. The models are highly connected and rely intimately upon each other. From interaction with the system, a stream of highly granular data is generated by each student. The analysis of this data creates a highly granular picture of each individual's abilities, strengths and weakness, learning needs, preferences and overall experience.

The information generated by this analysis is provided to both learners and instructors in real-time. Learners are placed in control of their own learning and provided with data derived guidance. Instructors are supplied with historically unachievable insights into their student's current knowledge state and progress. They are delivered with the tools and functionality needed intervene in the learning process, provide additional support and direction, and track the outcome of that intervention.

Analytics for the Student

From learner interactions the system captures real-time measures of their progress, ability, knowledge growth, gaps, strengths, weaknesses and learning preferences. These provide the means for new and detailed insights into how each individual learns and enables the prediction of future progress.

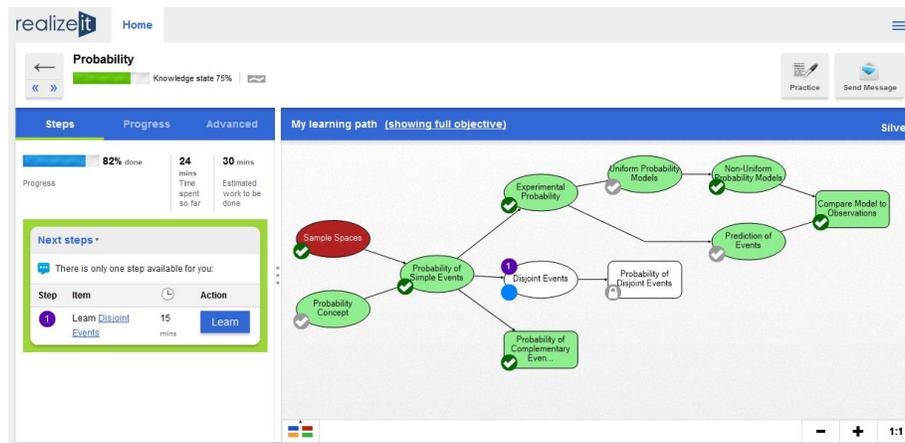


Figure 1: The main student learning dashboard

The analytics provided to each individual learner is integrated into their dashboard in the form of visualizations, feedback and guidance. Each student is provided with a summary of their overall progress across all courses and objectives. On each individual objective they can see a map of their learning path and its associated metrics; an example is shown in Figure 1. This provides them with a visual indicator of their own learning journey including their individual strengths and weaknesses. Along with data on their personal learning preferences and style, this data is used by Realize^{it} to provide assistance and guidance on the next steps in their learning path.

Analytics for the Instructor

Instructors are provided with their own set of analytics and associated dashboards to support academic advisement. The instructor has access to data at a deeply granular level so they can see precisely where people are in their learning path. This is generally unachievable in a traditional setting. Perhaps more interestingly the instructor is provided with real-time analytics and insights that are distilled from this data. Realize^{it} performs the analysis automatically for the instructor; the instructor is not forced to wade through the data themselves.

The analytics available on each individual learner, along with the class as a whole, allow instructors to measure and track engagement as well its effectiveness, to gain a better understanding of the needs of their students and tailor and target their techniques to better meet the needs of each individual. On the main instructor dashboard, of which an example is shown in Figure 2, they can visualize each learner's performance, engagement, ability, strengths and areas in need of improvement. Important insights such as problem areas in an objective or students who are struggling are highlighted and the opportunity to interact or intervene with a student is provided. The detailed data on each individual is also available for the instructor to view and analyze.

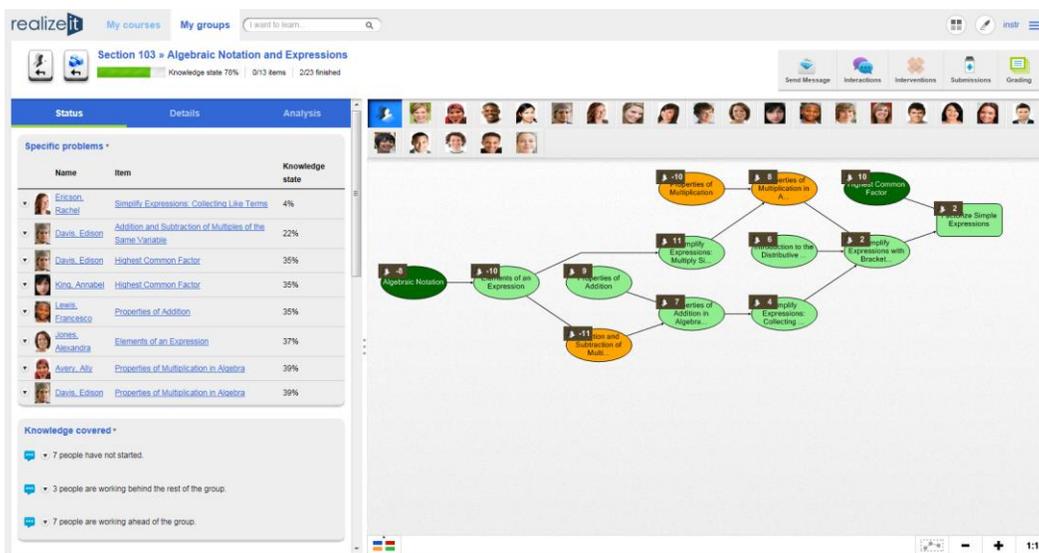


Figure 2: The main instructor dashboard

Academic Analytics in Realize^{it}

The Realize^{it} platform contains a number of tools that support enterprise level reporting and data analytics. Known as Analytical Reviews, these tools enable both the manual and automated generation and distribution of reports and analytics. They allow the institutions to gain an understanding of the performance and effectiveness of a curriculum, a course, their faculty, a class and individual learners from pre-enrolment to graduation. These insights can assist them in providing both academic and non-academics services and support. Data on faculty and departments can ensure effective and targeted faculty resource planning and to evolve this based on the response, feedback and behavior of learners and instructors. Content and curriculum usage patterns, together with student learning patterns can be correlated with successful learning outcomes and enable early visibility of student academic readiness and specific focus areas for improvement.

The Analytical Reviews automatically brings together all data affecting the learning process including curriculum, content, students, instructors and cross-course connections. It distinguishes between Inputs and Influencers and identifies relationships, trends and correlations between these factors and how they influence Outcomes. This functions as an early warning system for an institution. It can identify characteristics of best practice and make data and information actionable.

For the remainder of this paper, we will describe the output from one of the Analytical Reviews found in Realize^{it}. We will highlight some of the analytical modules through the use of real data that has been completely

anonymized. This data is from an English course (EN-01-01 Foundations of English Composition) delivered through Realize^{it} in a North American higher education institute. This particular course is made up of five weekly objectives. However we note that a learner may start an objective early or continue working on it after the due date. This particular dataset covers 13 iterations of this course, known as sessions, which were run consecutively.

Course Analytical Review

The Course Analytical Review is a one-click process that can be run on any desired course. The review aggregates all data relating to all sessions of the course and in real-time performs an automated and comprehensive analysis. The output of this is a series of modular components, each containing results and analytics related to the focus of that module. As shown in Figure 3, these modules can be preselected along with the chosen course before running the review. Furthermore the default of comparing the latest session to the previous one can also be changed. The following subsections will discuss each of the output modules individually.

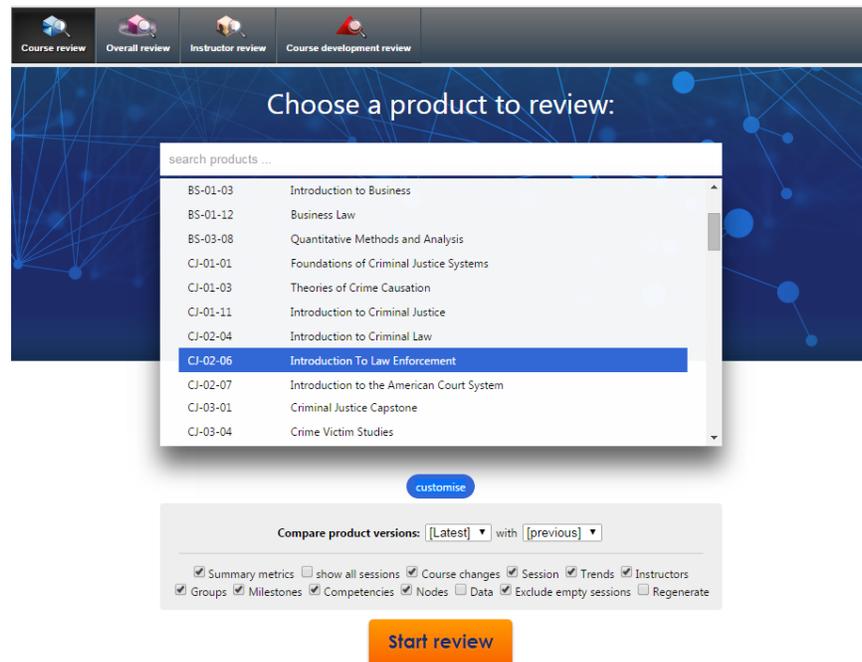


Figure 3: The course analytical review setup page

Course Structure

The course analytical review begins by identifying if any changes have been made to the structure of the course. The design of the current session is compared directly to that of the previous session, and these results are relayed to the user. If no changes were found, then the system checks back through all previous sessions to identify when changes were last made. They types of changes include:

- Course changes that relate to grading or scoring
- New, removed or changed objectives
- New or removed knowledge items
- New, removed or changed connections between knowledge items
- New, removed or changed content
- Changes in question or element metrics

Student Numbers

This module analyses the change in student numbers over all sessions of the course; a sample output is shown in Figure 4. For this course a total of 14,832 learners enrolled over the 13 sessions. This includes 12,067 unique learners and 2,765 repeat enrollments. The first 3 sessions of this course were trials with limited numbers of learners. The following 10 sessions were full scale implementations. We can clearly see that student enrollments in this course peaked at session Y13-10. Also visible is the increase in the percentage of enrollments that are repeat students.

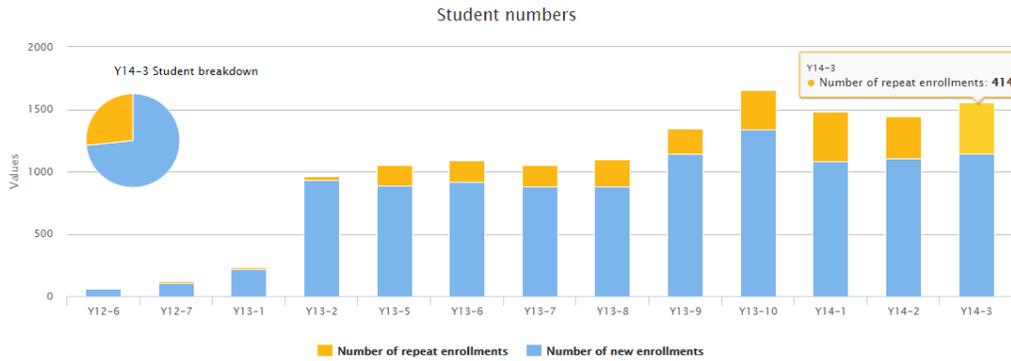


Figure 4: Student numbers

Student and Instructor Metrics

This module compares the current session to the previous one to identify and highlight changes in key student and instructor metrics. For the English composition course we can see an increase in student numbers but a drop in the number of instructors and groups. This has led to an increase in the number of students per group and per instructor. There has also been an increase in the number of repeat students. Within the system these factors are known as Inputs and are important factors to understand. They are outside of the instructors and students control yet they can have a significant impact on the learning process. These values along with the student numbers and course structure define the parameters under which the remaining analysis takes place.

	Y14.03	Y14.02	Difference	Percentage Difference
Number of students	1,555	1,446	109	8%
Number repeating	414	341	73	21%
Percentage repeating	26.62%	23.58%	3.04%	13%
Students per group	34.556	30.125	4.431	15%
Number of instructors	24	26	-2	-8%
Number of new instructors	0	1	-1	
Number of groups	45	48	-3	-6%
Students per instructor	64.792	55.615	9.176	16%
Groups per instructor	1.875	1.846	0.029	2%

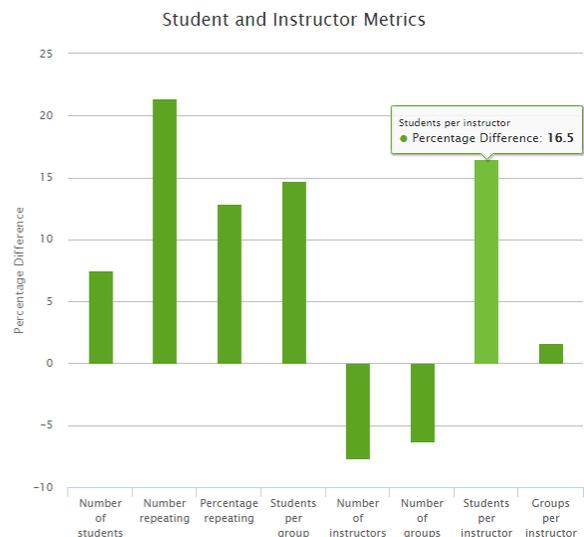


Figure 5: Student and Instructor metrics

Session Outcomes

The session outcomes module summaries changes to the five fundamental student metrics, known as Outcomes, across the comparison sessions. It examines average student knowledge state (measure of ability across the course) and the average knowledge covered (percentage of knowledge items completed). It also measures growth in each of these factors from participating in the course. This is the difference between the metrics at the start of the course, measured using the Realizeⁱⁱ functionality known as Determine Knowledge (Lynch & Howlin 2014), and the values at the end of the course. This module also captures the change in the average calculated score which is determined by an institution supplied grading formula. This formula is used to determine a student’s final grade and can be calculated from a variety of student metrics available in the system. The session outcomes for the English Composition course are shown in Figure 6. We can see that while there have been tiny decreases in the growth metrics, the other three metrics have increased slightly.

Outcome	Y14-03	Y14-02	Difference
Calculated score	78.700%	77.160%	+1.540%
Knowledge state	75.070%	74.460%	+0.610%
Knowledge covered	87.270%	84.870%	+2.400%
Knowledge state growth	2.390%	3.210%	-0.820%
Knowledge covered growth	24.090%	24.680%	-0.590%

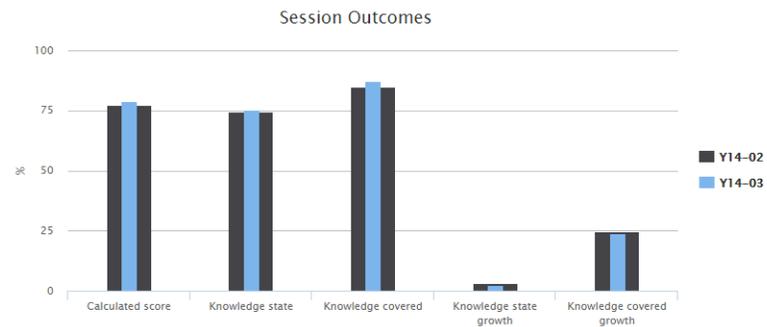


Figure 6: Session Outcomes

Grades

This module highlights changes in the distribution of grades between the two comparison sessions. For the course under review, Figure 7 shows us that there has been a slight increase in the percentage of students receiving an A or C grade combined with a decrease in the D, E and F grades. This demonstrates the large effect on grades from the small change in calculated score, discussed in the previous section, which is itself a result of the increase in knowledge state and knowledge covered.

Grade	Y14-03	Y14-02	Difference
A	44.7%	38.2%	7%
B	20.6%	20.7%	0%
C	14.0%	13.3%	1%
D	14.2%	15.0%	-1%
E	3.9%	7.8%	-4%
F	2.7%	5.0%	-2%
Percentage pass	93.44%	87.21%	6.23%

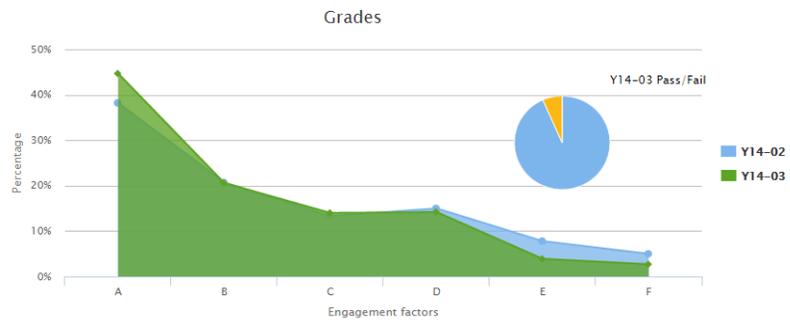


Figure 7: Comparison of grades

Milestones

A course can be split into separate milestone or objectives. For instance the English Composition course is split into four objectives. The milestones analysis section provides a breakdown of the session outcomes by objective. This allows the user to understand where the changes in session outcome originate from. The output of this module is similar to Figure 6 and so it has been omitted from the paper.

Engagement

The Realize^{it} platform records all interactions between students, instructors and the system. This includes the activity of the users in Realize^{it} as well as any interactions between the users themselves. Further messages between students and instructors are associated with any resulting outcomes. This allows the system to calculate and analyze 21 engagement metrics. In later analysis these are referred to as Influencers. It is worth noting that not all metrics associate an increase as a positive effect. For instance, the desired movement of 'number of messages with no response' is decreasing.

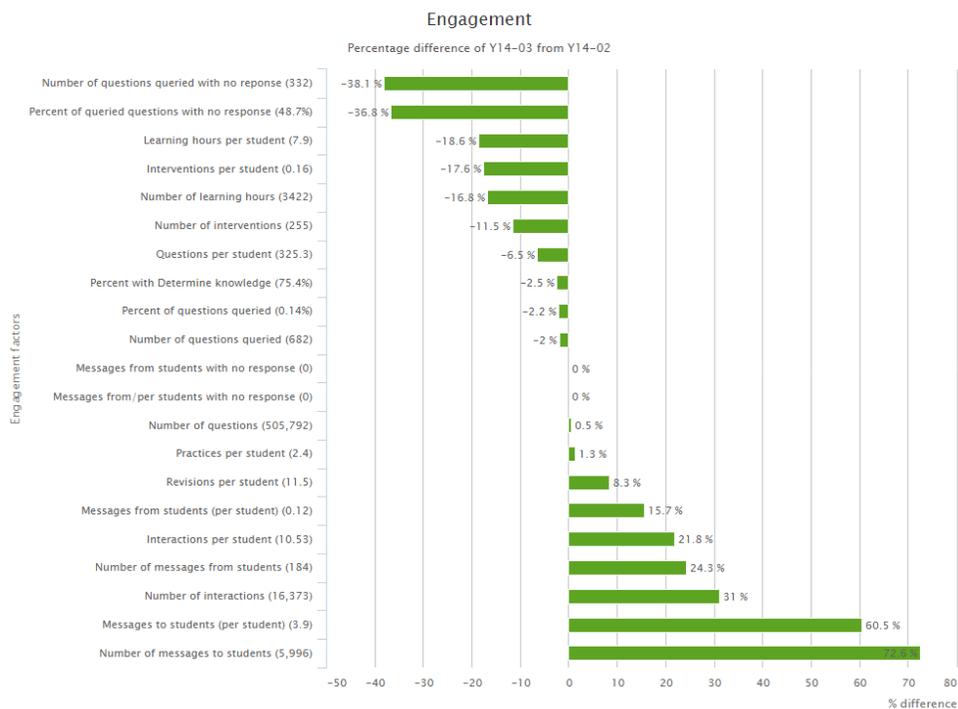


Figure 8: Engagement

Course Relationships

This module seeks to identify if other courses in the institution have had an effect on this course session. Each valid course is considered separately and used to partition the students of this current session into those who have previously completed the course versus those who have not. Hypothesis testing is then performed to determine if the students who have completed the course previously have a statistically significantly different outcome. From this we can infer which other courses have had positive or negative effects on this course.

Instructor Characteristics

The instructor level module aims to identify instructors that are performing beyond or behind their peers. Moreover it also aims to identify insights into the particular approaches adopted by different instructors. This begins with examining changes in the distribution of various instructor characteristics. These characteristics are measured relative to their peers and capture measures such as work load and engagement.

Instructor Outcomes

This module identifies the instructors who are either over-performing, underperforming or who have mixed performance relative to their peers. This is based across a range of student and engagement factors. The output of which can be useful to support instructors and to help develop a set of best practices.

Comparison of Groups for Instructor 7289
 Values have been scaled against the average across all groups (100)

Some of the groups belonging to Instructor 7289 perform below average compared to their peers, while some perform above average.

Below average grouping performance

- G893976-3

Above average grouping performance

- G894245-3



Figure 9: Instructor outcomes.

For this analysis, Instructor 7289 has been listed as having mixed performance. As shown in Figure 9, this instructor had one group that over-performed and another that underperformed (compared to the average across all groups). The values used in this analysis are provided along with the average scaled to 100. We can see that the main differences between the two groups are on the *Question queried* and *Knowledge state growth* metrics.

Instructor Correlations

The metrics derived from the previous two instructor modules are compared to determine if any correlations exists. For example, do instructors with more experience have students with higher knowledge states? Statistically significant relationships are highlighted and the correlations between all metrics are displayed in a heat map. The results for the most recent session of the English Composition course can be seen in Figure 10. We find that two correlations have been highlighted. The level of exposure of an instructor (number of groups) is significantly correlated with both the knowledge state and the growth in knowledge covered.

Two significant correlations have been found.

- There is a *positive* correlation between level of **Exposure and Av Group knowledge state**.
- There is a *positive* correlation between level of **Exposure and Av Group knowledge covered growth**.

The correlations between the other instructor metrics can be explored using the chart opposite.

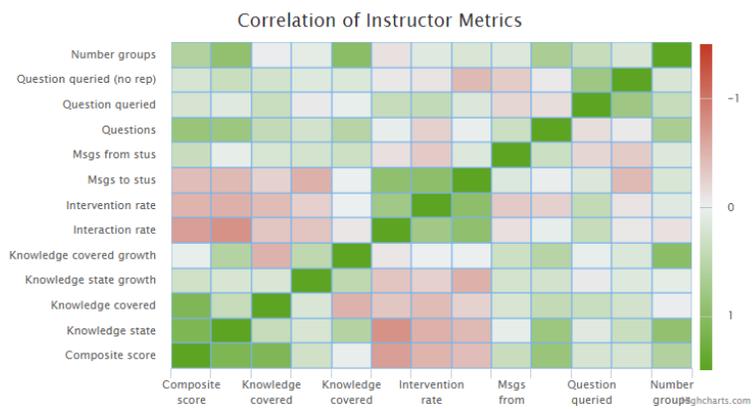


Figure 10: Instructor correlations

Trends

A major module of the automated analytical course review is concerned with determining linear trends in the data across *all* sessions of the course. It begins by examining the trends of over 30 individual factors that the Realize^{it} system captures during the learning process. The trends highlighted are those that have been identified as statistically significant, either positive or negative, or those where a significant flat trend is determined.

Global Versus Recent Trends

This builds on the trends module by comparing the global trend over all sessions to the trend over the most recent sessions. This allows early visibility on shifts and changes that are occurring within the course. It can act as an early warning system before these changes can significantly affect the global trend. A comparison will be highlighted if a significant trend is identified and if there is a difference between the global and the recent trend.

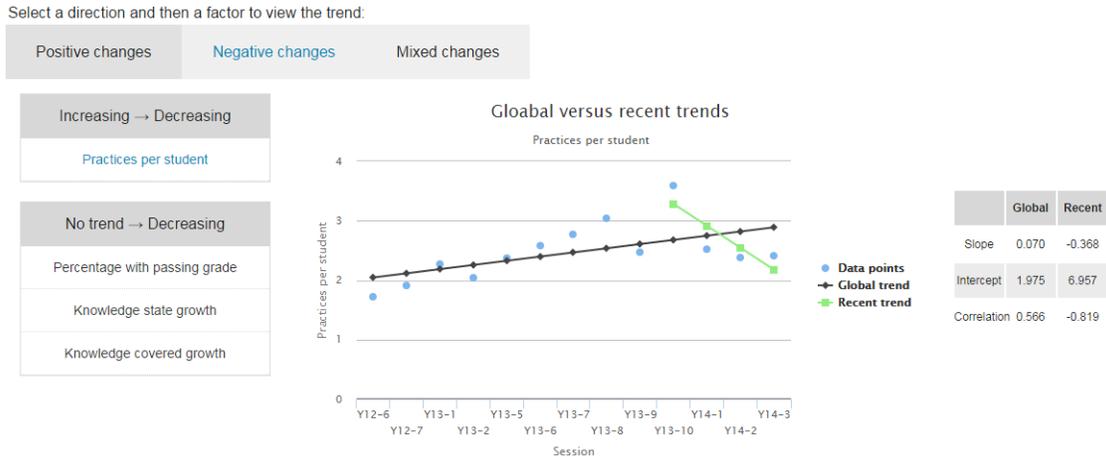


Figure 11: Global versus recent trends

For example in the English Composition course, as shown in Figure 11, there has been a negative change in the number of practices per student. The global trend suggests that the values have generally been increasing but it can be seen that in the most recent sessions it has generally been decreasing. This early warning can prompt instructors to investigate further and perhaps take action to encourage students to practice their knowledge more.

Correlations

In this module the system examines if there is a linear correlation between any of the factors across all previous sessions. Factors are grouped into inputs, influencers and outcomes.

- Inputs are the factors which describe how the course is structured. These factors are outside of the control of students and instructors but can still impact the learning process. Examples of these include the number of students or the number of groups per instructor.
- Influencers are factors that measure the performance and engagement of the students and the instructors and are directly influenced by their behavior. Examples of these include the number of learning hours or the number of interactions between students and instructors.
- Outcomes capture the results of the learning process. Examples of these include the knowledge state and knowledge covered.

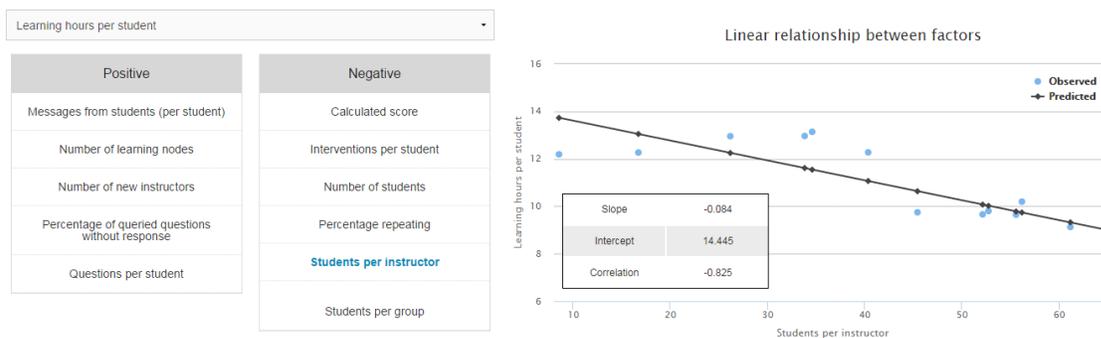


Figure 12: Correlations

These groups are then compared in order to determine correlations. The following comparisons take place: inputs

to influencers, influencers to influencers, influencers to outcomes, and inputs to outcomes. In Figure 12 we can see that the system has determined a negative relationship between the input, students per instructor, and the influencer, learning hours per student. This allows the institution to determine and quantify the negative impact increasing the student to instructor ratio has on learning.

Summary

The sections above have detailed and demonstrated just a small portion of the analytics available on a course. Yet this glimpse has generated valuable insights. Changes to the course structure between comparison sessions become immediately apparent. The review highlights important differences in the Inputs; aspects such as student instructor ratios or the number of new instructors. Outcome metrics are analyzed with changes emphasized to allow visibility into learner achievements and progress across sessions. This can be broken down further to enable users to trace these changes back to their source. Learner and instructor engagement metrics, or Influencers, are analyzed to determine the change in makeup and behavior of each of these groups. This includes a detailed analysis of the Instructors performance relative to their peers. Any course relationships or difference in student performance that can be determined are also reported in order to highlight their possible influence on the Outcomes. Finally the review seeks to determine any long term, recent or changes in trends, correlations or relationships that exist between the Input, Influencers and Outcomes across all sessions of the course.

Conclusion

There is a wealth of educational data being generated and captured by learning systems. Realize^{it} uses this as a powerful resource that supports both learners and instructors with integrated analytics during the learning process. Realize^{it} also makes this granular learner data available for institutional analysis which facilitates institutional level decision making. In particular the Analytical Reviews provide an institution with a one-click process that will aggregate all appropriate educational data and in real-time perform a series of automated and comprehensive analyses. The modular outputs of these reviews drive insights, allow the discovery of new patterns and relationships along with displaying the relevant data points. They allow the appropriate educational stakeholders to make data-informed decisions.

References

- Buckingham Shum, S. (2012). *Learning Analytics*. UNESCO Institute for Information Technologies in Education.
- Ferguson, R. (2012). *Learning Analytics: Drivers, Developments and Challenge*. International Journal of Technology Enhanced Learning
- Howlin, C. (2013). *The Realize^{it} System*. <http://realizeitlearning.com/papers/The%20RealizeIt%20System.pdf>
- Long P., & Siemens, G. (2011). *Penetrating the Fog: Analytics in Learning and Education*. Educause Review Online, 45, 5 31-40.
- Lynch D.J., & Howlin C.P. (2014). *Uncovering Latent Knowledge: A Comparison of Two Algorithms*. UMAP 2014, LNCS 8538, pp. 363-368. Springer International Publishing Switzerland
- MacNeill, S., Campbell, L. & Hawksey, M. (2014). *Analytics for Education*. Journal of Interactive Media in Education
- van Barneveld, A., Arnold, K.E. and Campbell, J.P. (2012). *Analytics in Higher Education: Establishing a Common Language* (ELI Paper 1), EDUCAUSE.