Chapter 8
Learning from Learning Analytics: How Much Do We Know About Patterns of Student Engagement?

Belinda A. Chiera, Małgorzata W. Korolkiewicz, and Lisa J. Schultz

Abstract Learning analytics is a rapidly growing field of research to inform the process of using data to improve learning and teaching. A common focus is to explore and develop strategies to address student learning and engagement issues at the institutional level. Typically, the data are captured by virtual learning environment systems and often complemented with student demographic information. At present, there is no standardised, holistic approach to data storage and collection across learning institutions, resulting in student learning data sets that are very large, incomplete and non-homogeneous.

In this chapter, we aim to harness these data in order to gain a better understanding of students' online engagement patterns specifically in a standard university setting that combines face-to-face learning with online resources and learning environments. Moreover, in the context of first-year University students, it is particularly vital to understand student patterns of online engagement. The data for the research conducted here has been sourced from a large Australian university. Patterns of student engagement are explored using a variety of data visualisations and statistical analysis techniques to develop an understanding of student engagement with course resources over the duration of a semester.

Keywords Learning analytics · University · Online engagement · Patterns

B. A. Chiera (✉) · M. W. Korolkiewicz · L. J. Schultz
University of South Australia, Adelaide, SA, Australia
e-mail: Belinda.Chiera@unisa.edu.au; malgorzata.korolkiewicz@unisa.edu.au;
Lisa.Schultz@unisa.edu.au

Introduction

Learning analytics is a rapidly growing field of research to inform the process of using data to improve learning and teaching. According to the definition proposed by the Society for Learning Analytics Research (SoLAR), learning analytics refers to the measurement, collection, analysis, and reporting of data about the progress of learners and the contexts in which learning takes place (Siemens & Gasevic, 2012). Typically, the data are captured by virtual learning environment systems and complemented with student demographic information. Although learning analytics in higher education is still at relatively early stages of development, it is expected to make significant contributions in a number of areas, including quality assurance, improvement of retention rates, identifying students at risk, assessment of outcomes among distinct student sub-groups, and adaptive learning (Sclater et al., 2016).

The widespread adoption by universities of learning management systems to create virtual learning environments has fundamentally changed how students engage with their studies. There is no single definition of student engagement, however students' effort, involvement in learning activities, motivation to learn, and consequently their academic achievement are typically considered (see for example Beer et al., 2010; Henrie et al., 2015). Virtual learning environments offer a unique opportunity to capture aspects of students' engagement in a rich data source providing information about changes in learner behaviour. In this chapter, we study different aspects of students' online activity to identify how and when they engage with online course resources.

To date, student online engagement has been explored with a predominant focus on Massive Open Online Courses (MOOCs), with latent engagement patterns examined in Ramesh et al. (2014), connections with performance explored in Coffrin et al. (2014) and Phan et al. (2015), and learner profiles studied in Kizilcec et al. (2013) and Ferguson and Clow (2015), just to name a few. There is however comparatively less emphasis on student patterns of engagement in the standard university setting that combines face-to-face learning with online resources and learning environments. Studies that do exist (see for example Beer et al., 2010 or Henrie et al., 2015) indicate a vital link between student engagement online and a successful learning experience overall. In the context of first-year University students, it therefore seems to be particularly vital to understand student patterns of online engagement as early experiences are likely to set the foundations for meaningful interaction with the University learning environment for the duration of their study.

The data considered here have been sourced from a large Australian university and encompass all student records across the 2016 academic year. The data combine student demographic information and weekly counts of visits to online course resources, used as a proxy for student online engagement. The weekly course click-count data provide a fine granularity of online engagement data and reflect the heterogeneous design and delivery of online resources in courses from distinct learning

areas. For the purposes of this chapter, we focus on a specific first-year course selected from the health area and examine click-count data for its virtual learning environment for semester one 2016. Patterns of student engagement are explored using data visualisations and statistical analysis techniques to develop an understanding of student behaviour, including how students engage with course resources over the duration of a semester. As in studies dedicated to student engagement patterns in MOOCs (see for example Hung & Zhang, 2008; Kizilcec et al., 2013; Hughes & Dobbins, 2015), we apply data mining techniques to analyse weekly virtual learning environment visits, in particular a hybrid hierarchical *k*-means clustering approach.

In light of existing studies and our own experience with virtual learning environments, we aim to gain insight into the following questions:

- How do levels of access to online resources vary over the course of a semester?
- How do student characteristics influence their engagement with online resources over the course of a semester?
- What role does assessment play in guiding the use of online resources?

The layout of this chapter is as follows. In section "Course Data", the student engagement data used for the analysis is described, and the results of the analysis are presented in section "Results". In particular, the visualisations of patterns in engagement behaviours are presented and discussed in section "Patterns of Student Engagement with Course Virtual Environment". This is then followed by a statistical analysis of the weekly engagement data using hybrid hierarchical *k*- means clustering across all students section "Clustering Weekly Engagement Patterns", with the analysis extended to clustering between student cohorts in section "Clustering Weekly Engagement Patterns by Cohort Classification". Conclusions are then drawn in section "Conclusions" on the patterns of student engagement identified in this chapter.

Course Data

The course selected for analysis is a first-year biology subject dedicated to the study of the structure and function of the human body. Our data consist of counts based on online learning platform logs for semester one in 2016 and captures the online activity of 745 individuals. There were both internal and external (online) offerings in the course, with the majority (75%) of students enrolled to study internally. The student cohort was predominantly female (85%) and 66% of students were under 21 years of age. To reflect their status as recent high school graduates, these students who were under 21 years of age are hereafter referred to as 'school leavers', with the remaining 34% of students who were over 21 years old referred to as 'mature age' students. It is worth noting that mature age students were the majority in the external offering of the course (66%), in contrast to the internal offering, where the

majority were school leavers (76%). The gender split in the internal and external offering was similar to that in the course cohort overall.

The course online learning environment was created within the university Moodle-based platform and included a range of resources and activities. The course syllabus was divided into weekly topics and organised into 'book' resources populated with an overview, readings, lecture notes, exercises, lecture recordings, 'test your knowledge' formative quizzes and additional resources. There was also a social forum created for each course offering. Course assessment included both formative and summative online quizzes as well as in-person tests and exams. Our analysis of student engagement with the virtual learning environment for this course is based on course site visits as well as student interactions with the 'book' resource and the 'forum' activity.

One of the limitations of our data is that for a selection of cases where final grades were not recorded, we are not able to distinguish between students who have unenrolled from the course and students who stopped engaging online however remained enrolled in the course. We are also not able to tell whether the recorded counts are simply multiple instances of the course website or a particular resource being opened in a student's browser or the number of student's clicks through the sections of the website. We are thus not able to meaningfully interpret the size of the counts or infer the exact nature of the student's engagement with the virtual learning environment. For those reasons, in what follows we mainly focus on general patterns only of student's online activity.

Results

Results presented in this section are based on weekly counts of clicks or 'visits' to the course website as well as specific Moodle resources and activities that exist within it. We focus on the main teaching weeks plus two 'mid-break' weeks, the end-of-semester examination period and the dedicated revision week separating the teaching weeks and the examination period, known as the 'swotvac' week. Based on our own teaching experience, we expected assessment to be an important factor in determining when and how students engage with online learning resources. In the course we have selected for analysis, the major pieces of assessment were a mid-test scheduled for teaching Week 5, a series of six online quizzes in Weeks 3–4 and 7–10, a practical examination in Week 12, as well as a final examination. In addition to summative assessment, there were also a number of formative online quizzes and so one can reasonably expect the timing of assessment to be strongly reflected in student engagement with the course website.

Patterns of Student Engagement with Course Virtual Environment

The overall pattern of student engagement is depicted in Fig. 8.1. Engagement with the course website during a week was defined as visiting the course website at least once. In general, the proportion of students visiting the course website was quite high at 70–80% for most weeks. It is however immediately apparent that students engaged with course online resources predominantly in the teaching weeks. The proportion of students visiting the course website during the mid-break and the second exam week was approximately 30% compared to over 90% in Week 1. Interestingly, the proportion of students visiting the course website during the first exam week was very high at 80%. Although we are not able to verify this, we expect the likely reason for such a high level of interaction with the online learning environment to be that the final examination was scheduled early in the second exam week. We also note a tendency in the student cohort to engage progressively less as the semester unfolded, with fewer visits to the course website in weeks without summative assessment, for example in Week 6 following the mid-test or in Weeks 11–13, despite the fact that summative practical examination was scheduled in Week 12.

In order to better understand the weekly patterns of visits to the course website, we also examine the proportions of students who engaged with the course website by study mode (internal versus external), gender (female versus male) and age (school leavers versus mature age students). From Figs. 8.2 and 8.3 there appears to be a very similar pattern of engagement across weeks for internal female and male students, with one noteworthy difference, namely much higher proportion of male
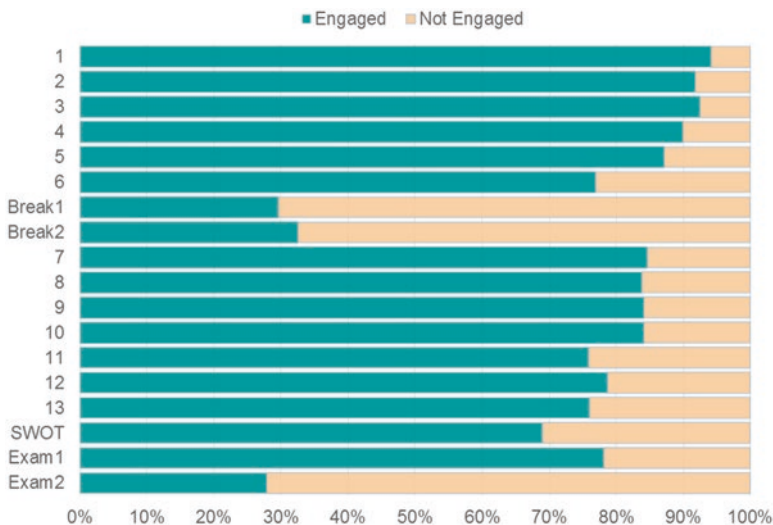

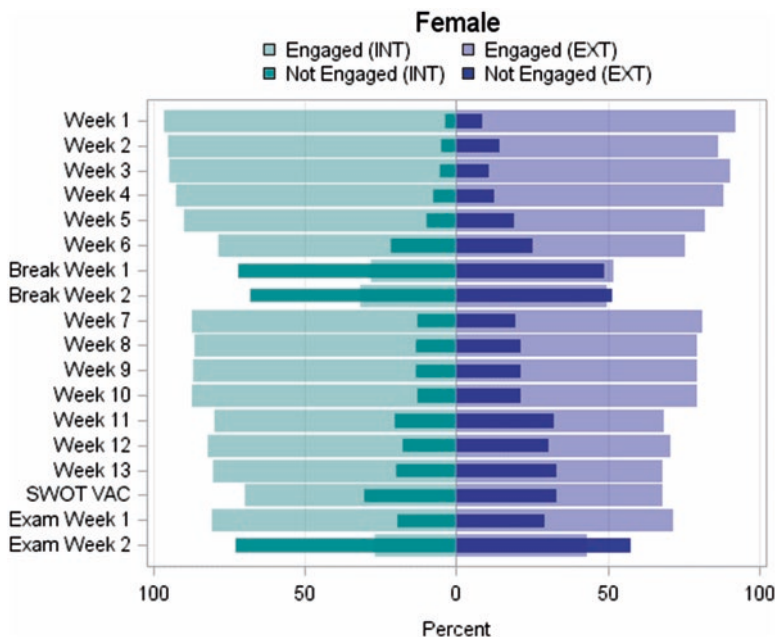**Fig. 8.1** Student engagement with course virtual environment

Fig. 8.2 Student engagement with course virtual environment by study mode among female students
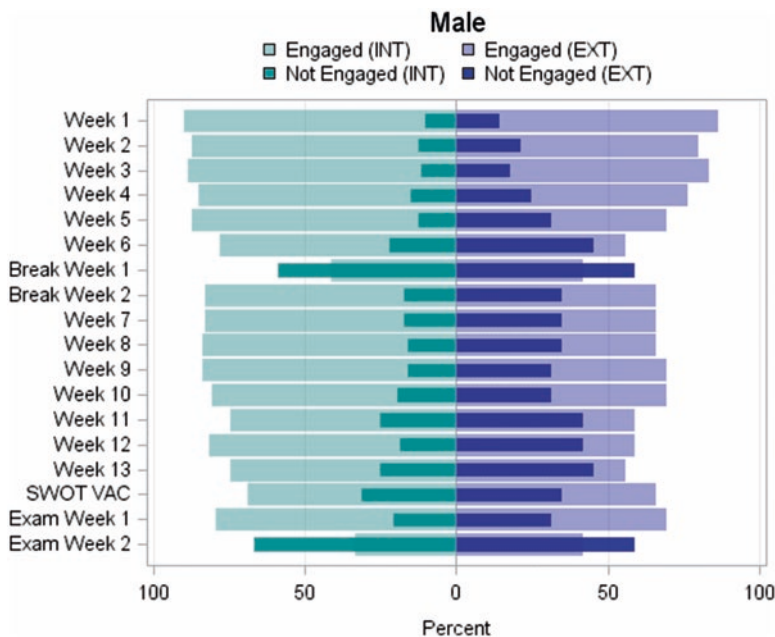

Fig. 8.3 Student engagement with course virtual environment by study mode among male students

students visiting the course website during the second week of the mid-break. In contrast, the proportion of female students visiting the course website was higher in the first few weeks of the semester. Comparing weekly proportions for internal and external female students we see a very similar pattern, however the proportion of female external students visiting the website is generally slightly lower.

This is particularly the case after Week 10 when there are no more summative online quizzes. The only exceptions are the mid-break and the second exam week when a much higher proportion of female external students visited the website. Finally, male external students appear to stop vising the course website much sooner in the semester. In short external students, in particular male external students, tended to disengage from the course sooner than their internal or female counterparts. Based on Figs. 8.4 and 8.5 we can also say that generally, a greater proportion of mature age students continued to visit the course website throughout the semester. There was little difference in engagement pattern of mature age students between the internal and external offering, with the exception of the mid-break. Approximately 60% of mature age external students continued to visit the course website during the break, compared with 30% of mature age internal students. For school leavers we note a higher proportion of students progressively disengaging from the course website, and likely from the course in general, particularly in the external offering.

The frequency with which students engaged with the course website each week was then explored, with the results shown in Fig. 8.6. Note that this analysis focussed only on students who accessed the course site at least once in a given week. From
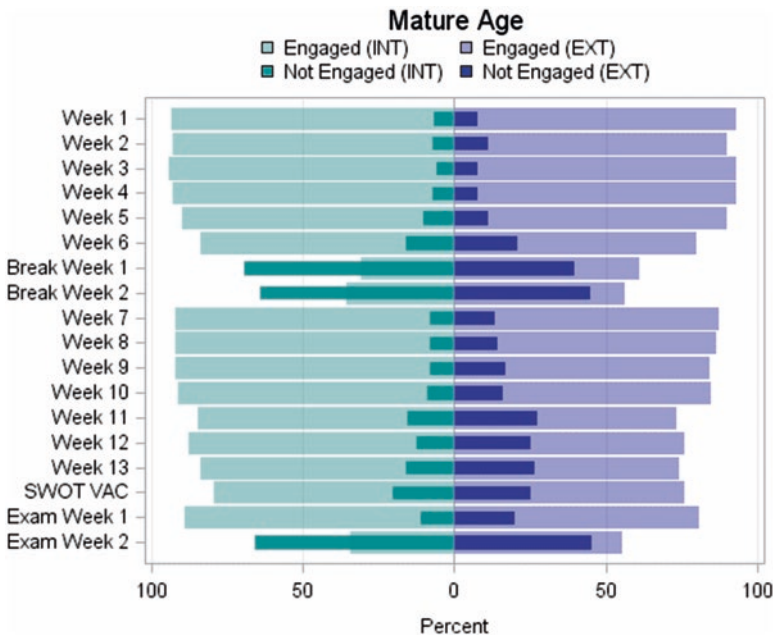

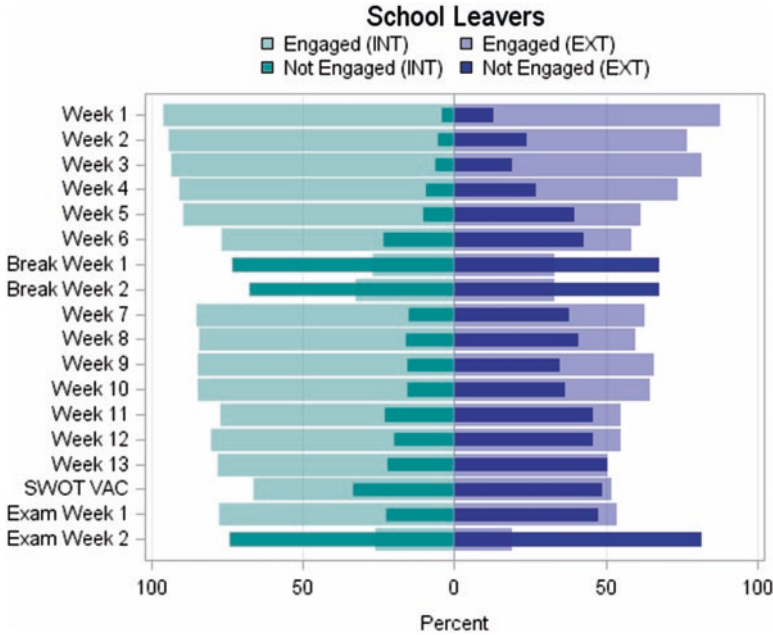**Fig. 8.4** Student engagement with course virtual environment by study mode and age category

Fig. 8.5 Student engagement with course virtual environment by study mode and age category
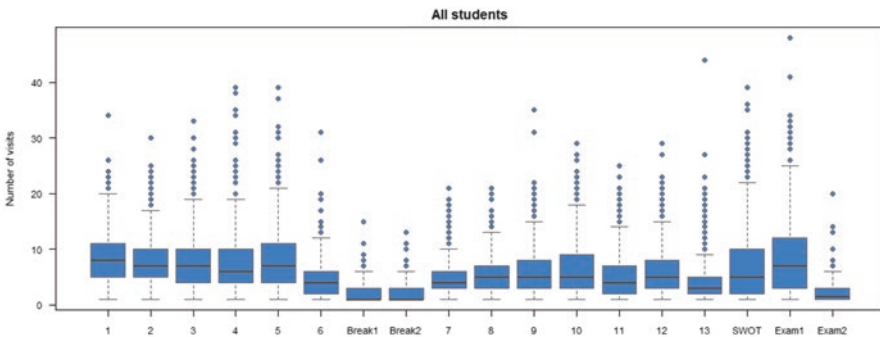

Fig. 8.6 Weekly course site visits for all students

Fig. 8.6 it can be seen that there are a selection of students with unusually high counts of course site visits across all teaching weeks, with some students accessing the course website more than 40 times in the space of a week. For the early teaching weeks, namely Weeks 1–5, the distribution of engagement counts is relatively consistent in terms of the typical number of course site visits and the associated variability in engagement between students. Overall engagement counts appear to decrease notably in Week 6 however and remain low across each of the two mid-semester break weeks. Following the mid-semester teaching break, course site visits become relatively high in variability.

Figure 8.6 also shows engagement counts for visits to the course website are higher during the first week of the exam period. This is suspected to be a reflection of an increased level of revision by the students in the lead-up to their final exam for the course.

Among students who engaged with the course site visit, the weekly frequency with which engagement occurred was then explored by the students' study mode, gender and age. Figure 8.7 shows the distribution of weekly engagement counts by gender, from which it can be seen that patterns of engagement frequency do not appear to differ substantially between male and female students.

Figure 8.8 shows the weekly engagement counts for internal and external students. The overall temporal pattern in weekly engagement appears to be similar
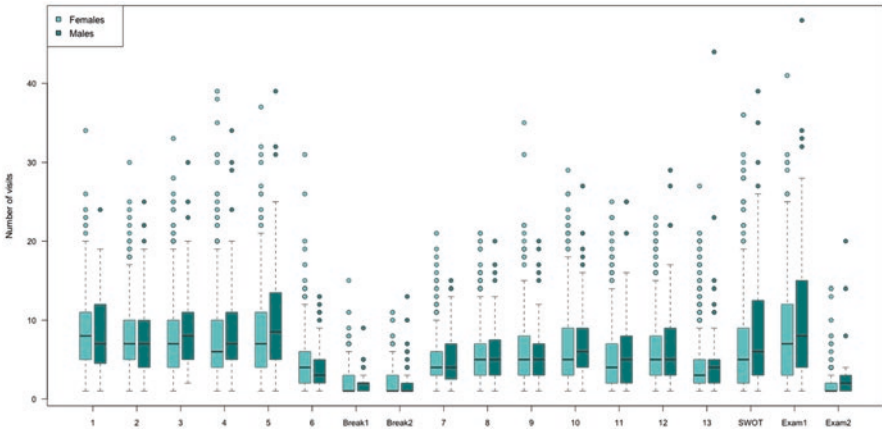

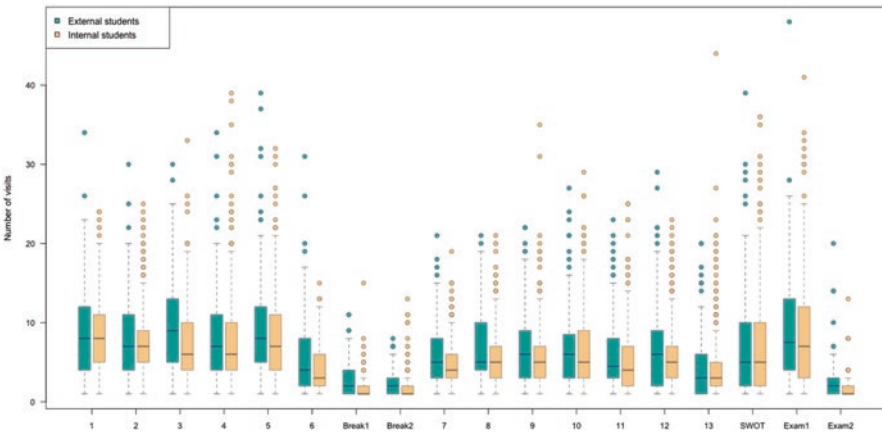
Fig. 8.7 Weekly course site visits by gender

Fig. 8.8 Weekly course site visits by study mode

between the two different study modes, with the exception that students who are externally enrolled typically have higher engagement counts for course site visits than their internal counterparts, as expected. The external students also appear to be more consistent in their engagement behaviour than students who are internally enrolled.

The engagement behaviours of mature age and school leaver students are comparatively shown in Fig. 8.9. Mature age students are typically more engaged than their school leaver counterparts, with greater variability observed in the course site counts between mature age students than school leavers. This trend in engagement patterns persists when separately considering internal and external students by age group (Figs. 8.10 and 8.11). Among the internal students, the mature age and school
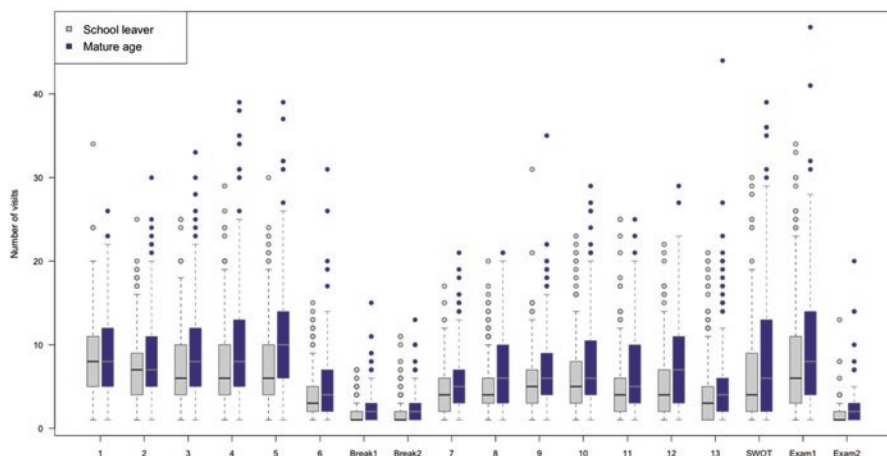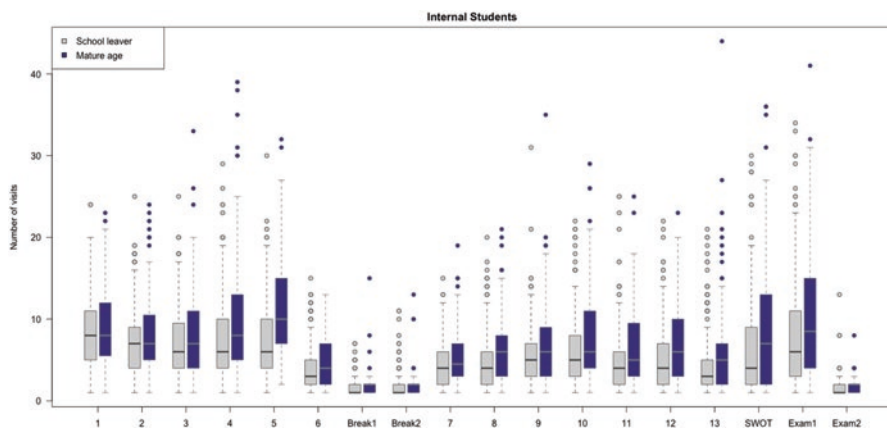

Fig. 8.9 Weekly course site visits for by age group

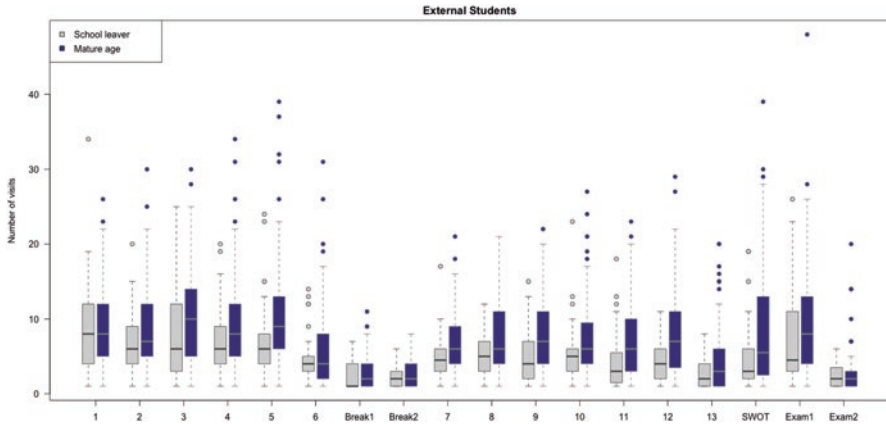**Fig. 8.10** Weekly course site visits for the internal students by age group

Fig. 8.11 Weekly course site visits for the external students by age group

leaver students have similar engagement behaviour for the early teaching weeks with the mature age students typically having higher engagement from approximately Week 5 onwards. By contrast, external mature age students are consistently higher in course site visits than external school leaver students across the entire semester.

To gain further insight into the nature of student demographics and the effect they may have on engagement with course online resources, a treemap of online course site visits was constructed using student gender and study mode as the grouping variables (Fig. 8.12).

Each of the four grouping variable combinations is represented with a rectangular block outlined in black, the size of which is determined by the number of students in that group. Each block is further divided into tiles representing weeks, with the size of each tile driven by the number of course visits for the corresponding week. The tiles are then coloured according to the average number of course site visits.

From Fig. 8.12 it can be seen that the blocks for the internally and externally enrolled females are considerably larger than the blocks for their male counterparts, reflecting the gender composition of the student cohort. Additionally, the blocks for the internal students are consistently larger in size than for the external students, again in keeping with the student enrolment distribution for this course.

Figure 8.12 also shows that Weeks 1–5 and Week 17 (i.e., exam week 1) occupy the greatest areas in the treemap for internal and external females as well as internal male students. This indicates that these are the teaching weeks for which these students are most actively accessing the online course site. In contrast, external male students appear to peak later in the teaching semester, with Weeks 1, 3, 16 and 17 as the most significant weeks for engagement. This indicates that external male students tend to increase their engagement significantly in the weeks immediately preceding the exam, with comparatively low engagement early in the semester.
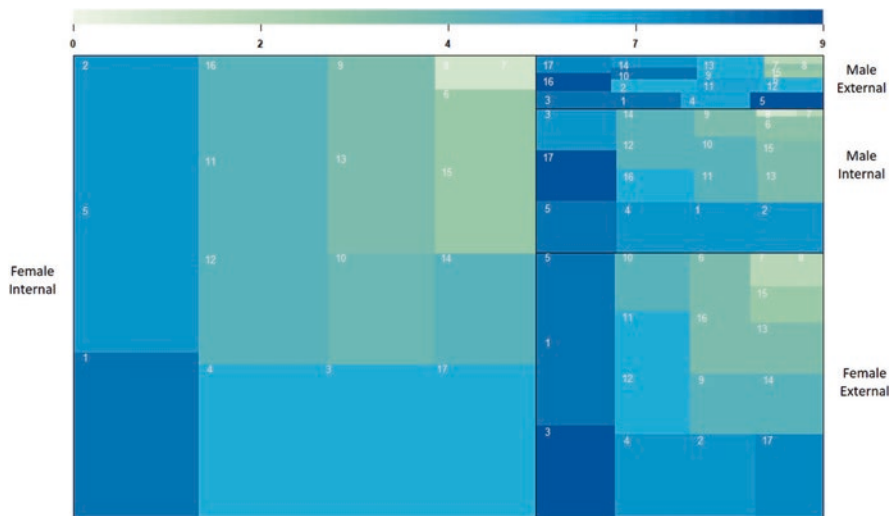
Fig. 8.12 Treemap of the weekly course site visits by student age, gender, and study mode

In terms of tile colour, Fig. 8.12 shows the typical number of course site visits per student appears to be highest in the first 5 weeks in the semester across all blocks. Among the internal female students, Week 1 corresponds to the greatest typical number of visits whereas Week 3 had the greatest number of visits for the external female students. In contrast, male students typically increased their engagement in the weeks leading up to the examination period, namely Week 16 for the external students and Week 17 for the internal students. Patterns of engagement therefore appear to be strongly associated with gender and study mode.

Additional insight into engagement patterns can be gained by categorising weekly student engagement counts into quartiles such that the lowest 25% of students are classified as being in Quartile 1, the lowest quartile of engagement, and the top 25% are classified as Quartile 4. This provides an opportunity to analyse student engagement trajectories across the teaching semester. As an example, Figs. 8.13 and 8.14 show the engagement quartiles trajectories for a random sample of 20 internal students and 20 external students, respectively. Within each figure, the light yellow circles in a given week indicate that the corresponding student was in the lowest quartile of engagement, Quartile 1, whereas the darker shaded circles show students in the highest quartile of engagement, Quartile 4, for that week. Gaps in the engagement trajectories indicate that the students did not engage at all that week. It can be seen that there are notably more yellow circles in the sample of internal students as compared to the external students, indicating these students were more often less engaged with the online course resources than their external peers. Figures 8.13 and 8.14 show that students with low engagement tended to remain in the lower quartiles of engagement across the entire semester while students in the highest quartile of engagement tended to have consistently high engagement. This is particularly
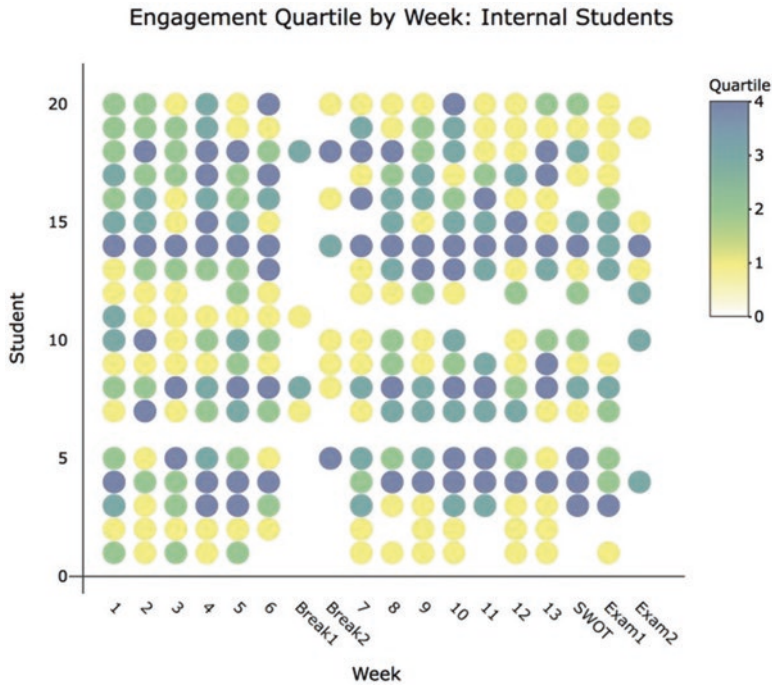
Fig. 8.13 Engagement quartile trajectories for the internally enrolled students across the semester

apparent among the external students (Fig. 8.14), with the exception of the mid-semester break and the exam weeks where engagement generally decreased.

We now turn our attention to students who visited the course website and examine their patterns of engagement with Moodle books containing weekly topic resources. Overall proportions of students visiting topic resources each week are depicted in Fig. 8.15. As was the case with course site visits, a very high proportion of students (over 90%) were visiting topic resources in the first few weeks, with the proportion decreasing progressively but staying relatively higher in weeks with summative assessment (Weeks 3–5 and then 7–10). Further, the proportion of students visiting the course website who then navigated to topic resources was quite high (around 80%) in weeks leading up to the final examination.

Among students who engaged with the weekly course topic resources, the frequency with which students accessed course resources is shown in Fig. 8.16. Engagement frequency was typically greatest in the early teaching weeks, with notably fewer visits to the topic resources during the mid-semester break and Week 13. As was seen with the frequency of course site visits (Fig. 8.6), the frequency with which students accessed topic resources was greatest during the swotvac week and the first week of the exam when students are revising for their final examination. Engagement patterns according to study mode, gender, and age (not shown) were very similar.
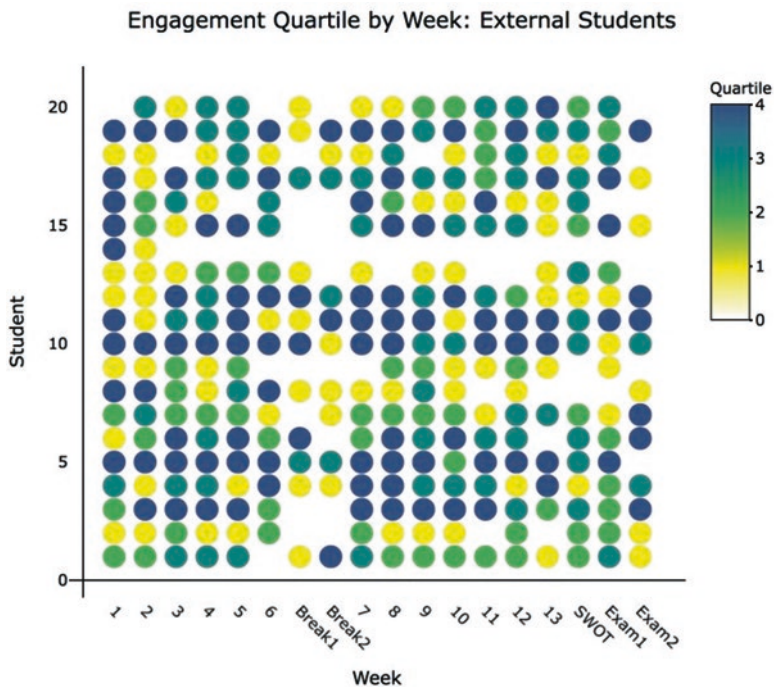
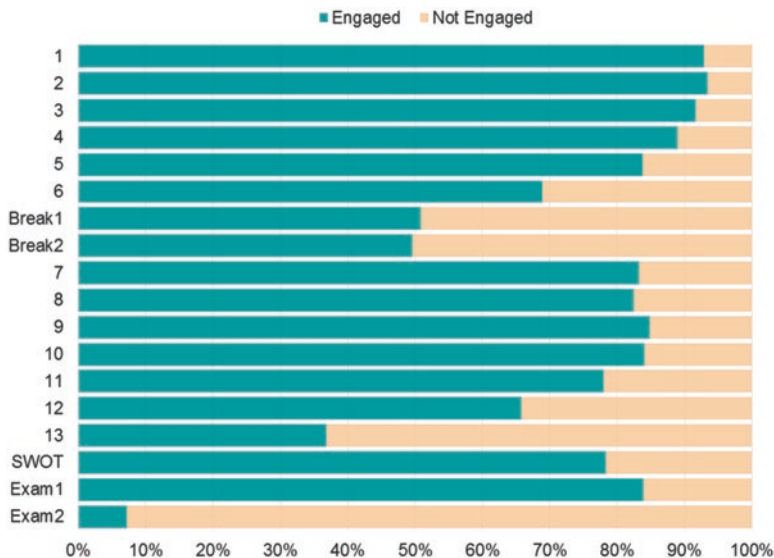**Fig. 8.14** Engagement quartile trajectories for the externally enrolled students across the semester

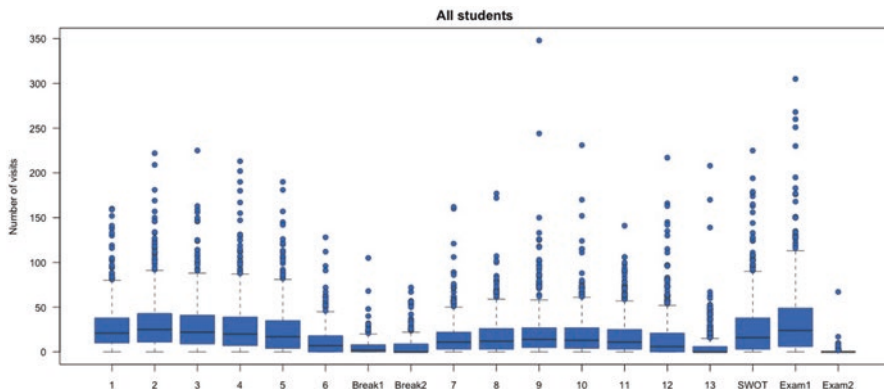**Fig. 8.15** Student engagement with course topic resources

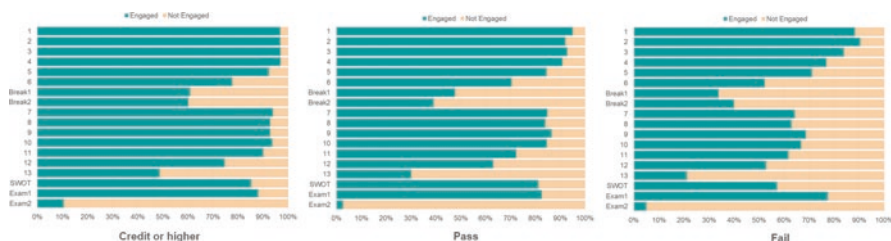**Fig. 8.16**   Student engagement with course topic resources by week

Fig. 8.17   Student engagement with course topic resources by final grade

When students' final grades in the course were taken into account, notable differences were observed in the patterns of online engagement. Figure 8.17 shows the proportions of students engaging with topic resources broken down by final grades. Here, final grades have been divided into three categories, namely 'Credit or higher', 'Pass' and 'Fail'. The general pattern of a high proportion of students visiting resource pages in the early weeks and weeks with scheduled summative assessment is apparent for all three grade categories. Unsurprisingly, the group with the most consistent pattern of engagement are students who eventually earned at least a Credit for the course. In contrast, students who failed the course tended to disengage from course resources as the weeks progressed, many seemingly returning right before the final exam.

This can also be seen in the weekly engagement frequencies for each grade bracket, shown in Figs. 8.18, 8.19 and 8.20. Students who received a credit or higher as their final grade had consistently higher engagement than students who received a passing grade. Moreover, students who failed the course overall had the lowest engagement with course topic resources each week.

It is also informative to consider how often students engage with weekly resources according to study mode, gender and age. Median visits to topic resources across all weeks are shown in Fig. 8.21. Key observations are as follows:
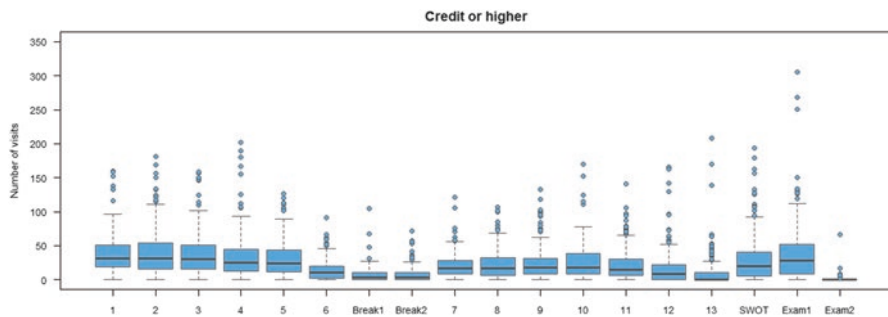
Fig. 8.18 Weekly engagement with topic resources for students who received a credit or higher as their final course grade
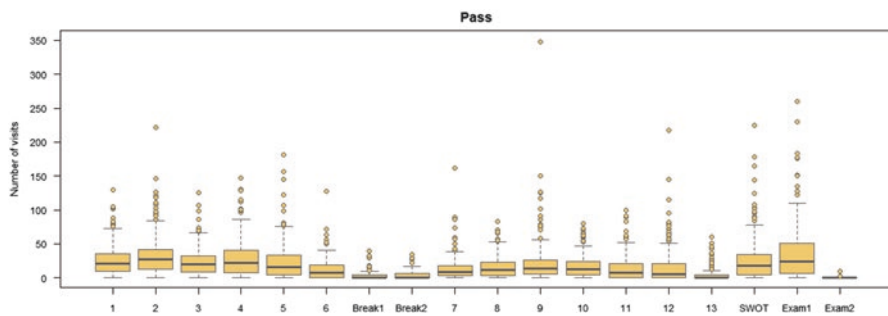

Fig. 8.19 Weekly engagement with topic resources for students who received a pass as their final course grade
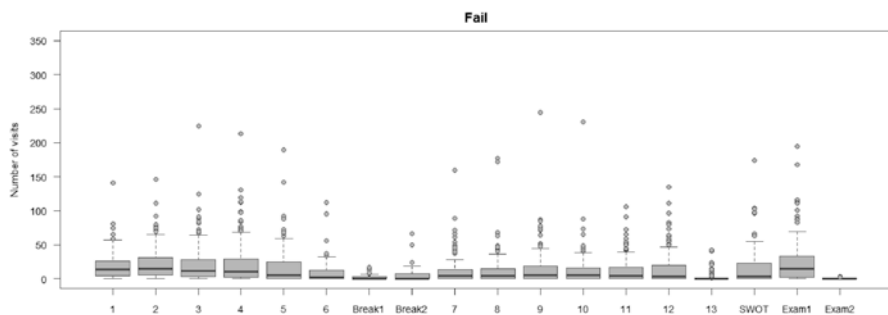

Fig. 8.20 Weekly engagement with topic resources for students who received a fail as their final course grade
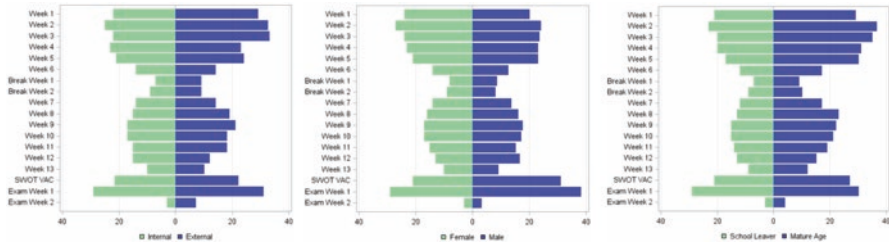
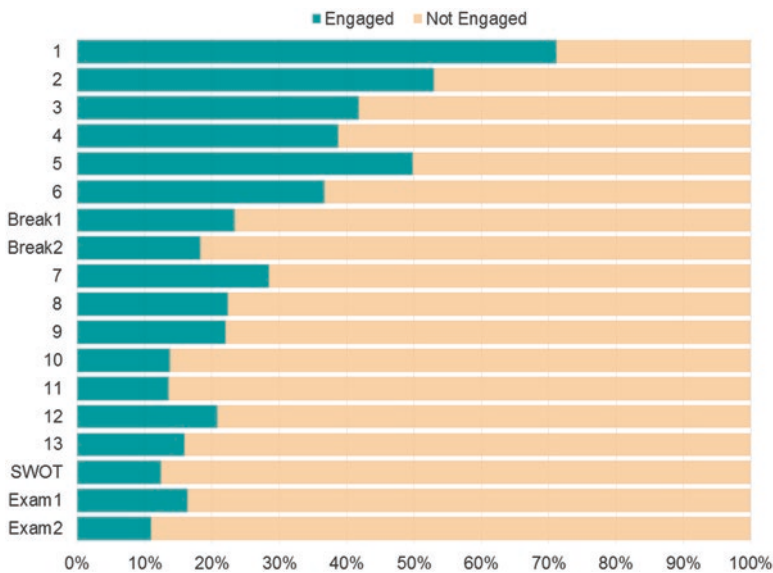Fig. 8.21 Median visits to topic resources by study mode, gender, and age

Fig. 8.22 Student engagement with social forum

- In most weeks, a typical student visited topic resources 20–30 times;
- External students tended to visit course resources more frequently than internal students, particularly in the first 3 weeks of the semester;
- Female and male students visited topic resources at the same rate in all weeks except for swotvac and the first exam week. Male students tended to be more active in those weeks (30–40 visits for a typical male student compared with 20–30 visits for a female student);
- Mature age students tended to visit the topic resources more often than school leavers, particularly during the first five teaching weeks.

Finally, we consider patterns of engagement with course social forums. Overall, forums did not appear to generate much student activity as the median number of visits per week to the forums was typically between 2 and 4. As illustrated in Fig. 8.22, the proportion of students visiting the forums was decreasing quite

dramatically from week to week, starting at approximately 75% in Week 1 and going down to about 10% in the second exam week. The week that stands out the most is Week 5 as the proportion of students visiting the forums in that week (50%) was higher than immediately before and after (40%). This increase in forum engagement may have been due to the mid-test being held in Week 5.

With regards to the number of times students accessed the forum, Fig. 8.23 also shows that the frequency of forum access decreased substantially as the semester progressed. There are a number of students however who accessed the forum an unusually large number of times each week relative to their peers. This suggests that despite generally low levels of engagement, there are some students who appear to benefit from the social forum. Overall however, Figs. 8.22 and 8.23 show that the proportion of students accessing the forum decreases over time with the students who do access the forum progressively reducing their level of engagement with this course resource over the semester. A similar pattern was observed when study mode, gender, and age were taken into account.

As in the case of visits to course resources, we also examine patterns of engagement with forums according to the final grade earned for the course, see Fig. 8.24. Again, for students who eventually failed the course, the proportion visiting forums
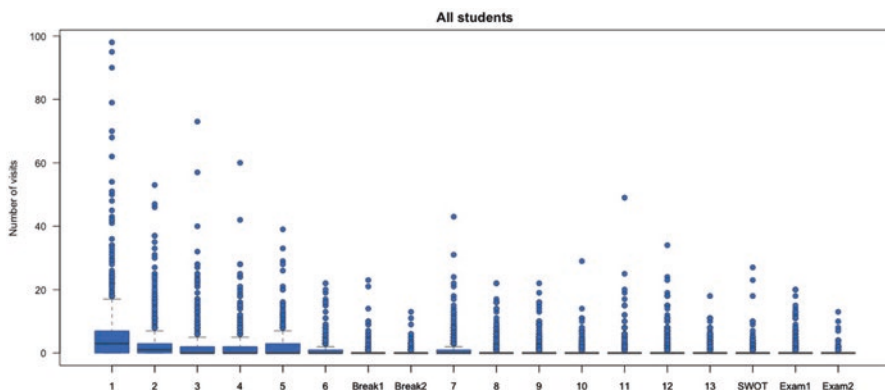

Fig. 8.23 Weekly student engagement frequency with the social forum over the course of the semester
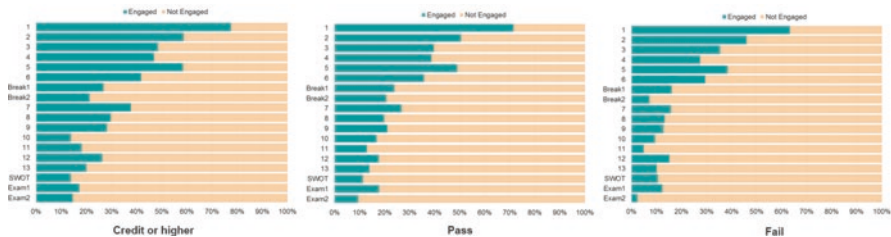

Fig. 8.24 Student engagement with social forum by final grade

was consistently much lower than for the other two grade categories, and decreasing at a faster rate from week to week following the mid-break. Similar patterns can be seen for the weekly frequency of engagement among students who did engage with the social forum each week, shown in Figs. 8.25, 8.26 and 8.27.
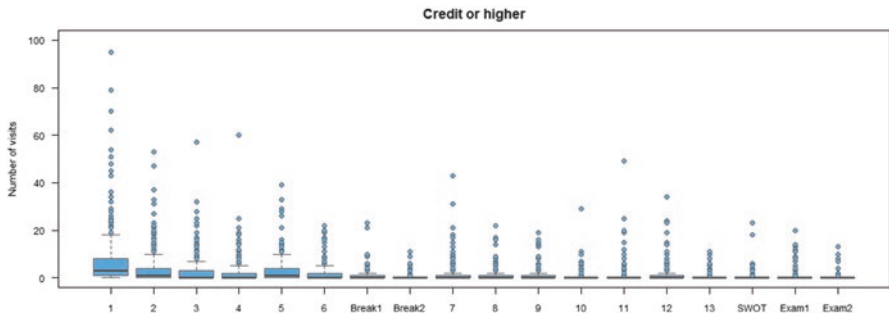

Fig. 8.25 Weekly engagement frequency with the social forum for students who received a credit or higher as their final grade
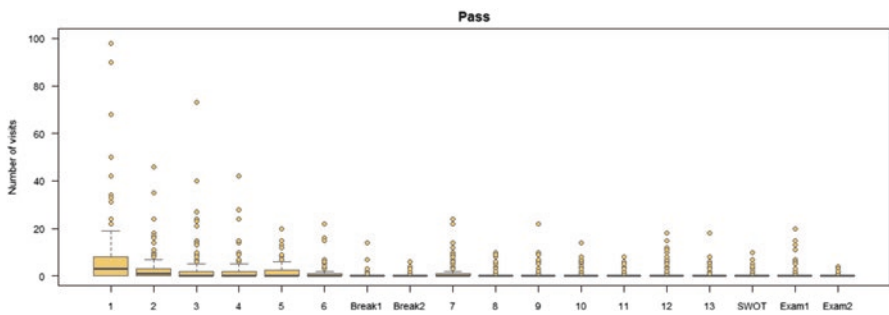

Fig. 8.26 Weekly engagement frequency with the social forum for students who received a pass as their final grade
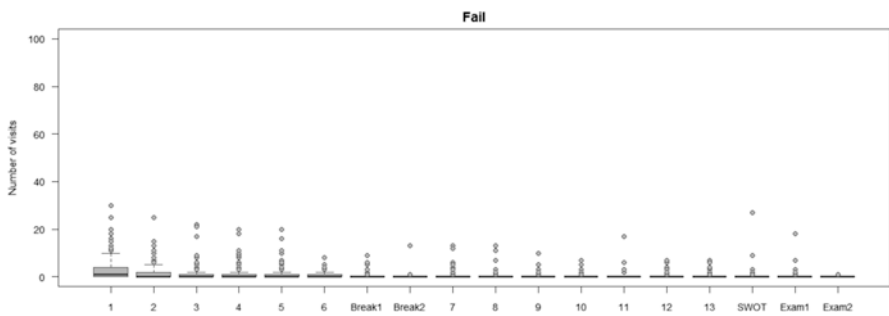

Fig. 8.27 Weekly engagement frequency with the social forum for students who received a fail as their final grade

Clustering Weekly Engagement Patterns

To further investigate the patterns indicated by the initial exploratory analysis, clustering was employed as a means to uncover potential structure within the weekly engagement data. Given the available demographic information about the students, it is natural to investigate whether intrinsic groupings can be detected across the weekly engagement.

Of the available clustering methods, the majority can be divided as belonging to either the group of *hierarchical* or *partitional* clustering methods. Hierarchical clustering is preferred when there is a perceived hierarchy in the data and is preferred for smaller data sets. Partitional clustering, which assumes there are central features around which the data can be clustered in non-overlapping partitions with the number of centres defined a priori, however boasts the advantage of being computationally inexpensive.

In this study the sample size was sufficiently small to allow for the investigation of clusters using agglomerative hierarchical clustering techniques as a first step before trialling partitional clustering methods. With regards to hierarchical clustering, it is not unreasonable to expect the presence of a hierarchical structure across the weekly engagement data, as from the exploratory analysis in section "Patterns of Student Engagement with Course Virtual Environment" it is apparent that student patterns of interaction with the course site and resources varied across the semester. In order to determine the best level of interpretability of the clusters we also investigated the use of partitional clustering via the *k*-means algorithm as well as a hybrid hierarchical *k*-means clustering approach as a way to overcome the limitations of both clustering methodologies.

For the hierarchical clustering we used a linkage algorithm and explored the use of single, complete, and average linkage, as well as Ward's distance to minimise within-cluster variance. Suitability of the clusters produced was evaluated by taking into account the cophenetic distances of each method, as well as cluster interpretability. For the partitional *k*-means algorithm, we assessed the number of clusters *k* to request a priori, by considering the output of the "elbow" method (Thorndike, 1953) combined with the silhouette method (Rousseeuw, 1987) to achieve an appropriate balance between cluster cohesion and separation. Typical values of k ranged between 2 and 4 and all values of k within the recommended ranges were trialled here. We also explored the use of a hybrid hierarchical k-means clustering method (Hasan & Duan, 2015).

We ultimately selected the hybrid hierarchical *k*-means approach in order to capture the expected hierarchy in weekly engagement patterns. This approach also overcomes the limitation of the *k*-means algorithm, namely identifying a priori the number of clusters in the data. An additional attractive advantage of retaining a hierarchical clustering approach is that a complete hierarchy of clusters are generated and as such is considered more informative than *k*-means (Hasan & Duan, 2015) while *k*-means is often preferred as it produces tighter clusters. For the hierarchical clustering method, we trialled the use of single-, complete-, and

average-linkage as well as Ward's method to minimise within-cluster variance. Ward's method produced the most interpretable clusters in this case, with groupings that reflected learning events across the semester and for this reason the results of using Ward's method are presented in what follows.

Weekly engagement as measured by the number of visits to the course website is shown in Fig. 8.28, as well as visits to topic resources and the forum (Figs. 8.29 and 8.30 respectively). As the clustering method is a hybrid hierarchical k-means algorithm, the resulting output is the traditional dendrogram arising from hierarchical clustering with the groupings formed along the dendrograms corresponding to clusters generated by the k-means algorithm.
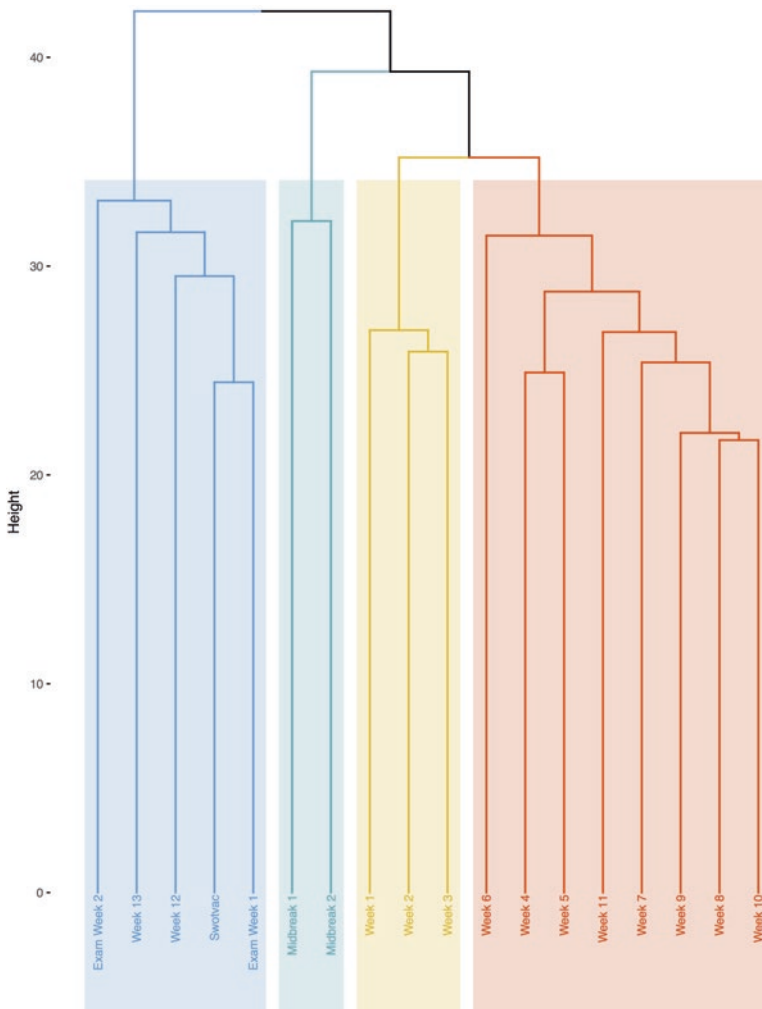

Fig. 8.28 Weekly course site visits
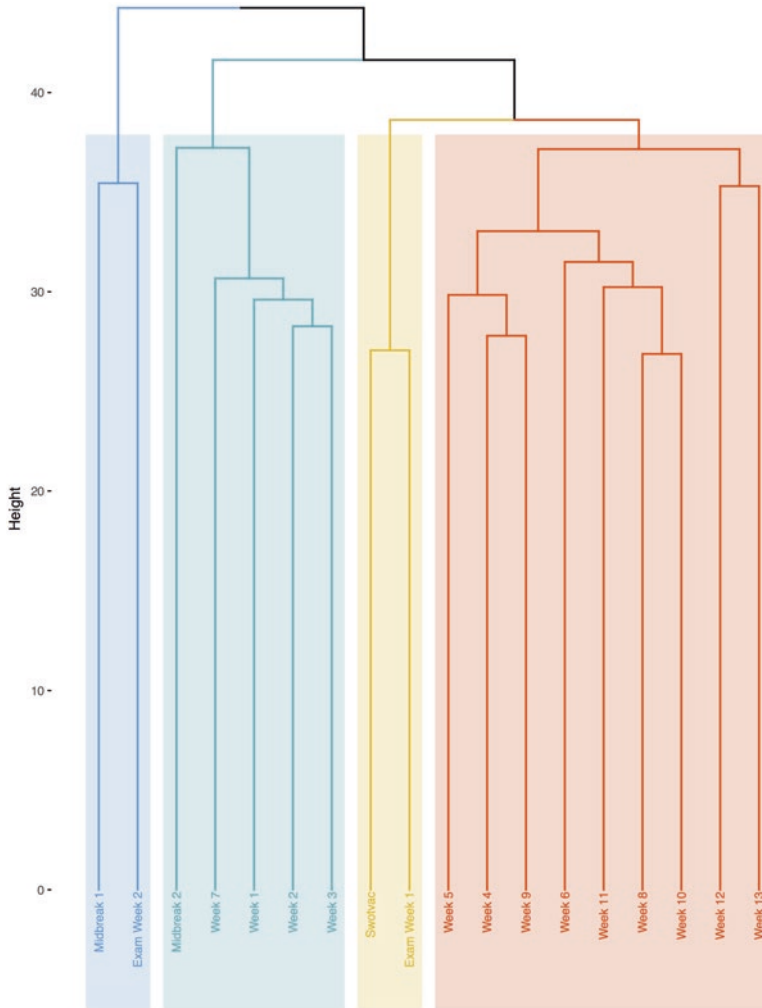
Fig. 8.29 Weekly visits to topic resources

From Fig. 8.28 the pattern that emerges emulates the distinct periods a semester will naturally fall in to as well as the specific assessment structure for the selected course.

Specifically, Weeks 1–3 (second cluster from the right) form an individual component of behaviour as might be expected with students anecdotally being motivated within the first few weeks of semester; the second component (rightmost cluster) indicates the in-between behaviours that in part reflect the assessment structure of the course. Specifically, the students had a major assessment piece in Week 5 while Week 6 was a quiet week from an instructional viewpoint. From the clustering we can see Week 6 forms an unusual week in terms of classification as it is the

leftmost member of the cluster in the dendrogram which typically indicates the outlier position in a data set. Similarly, Week 11 forms an outlier as well in that there are no assessment items in that particular week. On the other hand Weeks 3 and 4, as well as Weeks 7–10 all contain online quizzes. The mid-break cluster is intuitive in its grouping as it is natural students would access the course website differently during this time as there is no instruction taking place. It is interesting to note that Week 12 falls in the fourth cluster (first cluster on the left) and collectively with the run-up to the exam period. Week 12 marks a practical exam for the students before exam review commences (Week 13) and then the final exam period. In particular, Swotvac and Exam Week 1 are closest together within this larger block which would reflect preparation for the final exam itself.

When considering visits to topic resources (Fig. 8.29) a somewhat different pattern emerges. In the leftmost cluster, we can see Midbreak 1 and Exam Week 2 have been grouped together – these are the two non-instructional/non-assessment times in the course and it is natural they will appear together in a cluster. On the other hand, it is reasonable to expect visits to topic resources during Midbreak 2 as students had an assessment item in Week 7 and it is possible they were focused on quiz preparation. In these weeks, visits to topic resources matched the pattern of visits in Weeks 1–3, while Swotvac and Exam Week 1 (second cluster to the right) formed their own grouping which would reflect final exam preparation via the online topic resources. The remainder of the weeks (rightmost cluster) behave in a similar fashion to one another. There are sub-groupings within this cluster however the overall generic visitation behaviour is collectively distinguishable from the other three clusters.

Forum visits (Fig. 8.30) tell a different story once again. The leftmost cluster contains groupings specifically in Weeks 8 and 11, then Weeks 13 and Swotvac. Without knowledge of the content of the quiz it could be possible that week 8 indicated a challenging topic, however with Week 11 it could be postulated the upcoming practical exam (Week 12) was the driver for forum interaction. The second cluster from the left appears to contain the low-activity weeks of the course from the perspective of instruction and/or assessment. The cluster second from the right seems to have captured the forum visits over the first month of the course. The standout cluster in this case is the rightmost cluster in which both Weeks 5 and 12 are grouped together and both contained a large assessment item and Exam Week 1 also joined this grouping, indicative of the upcoming final examination.
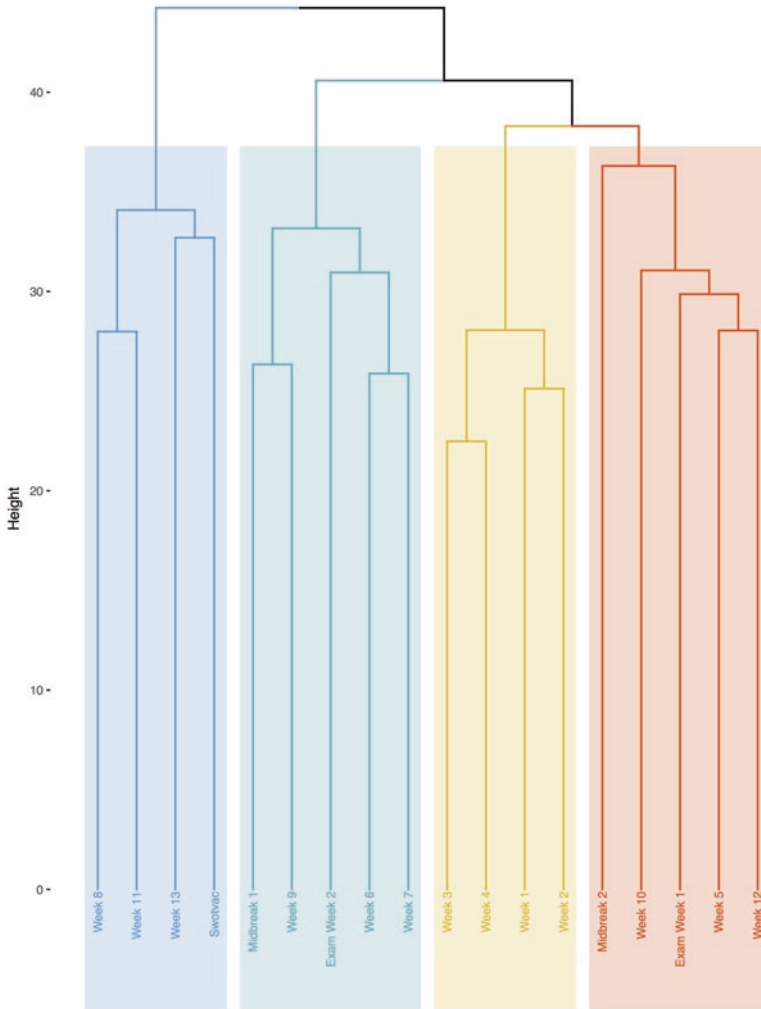
Fig. 8.30 Weekly forum visits

Clustering Weekly Engagement Patterns by Cohort Classification

In this section we further investigate clusters for student engagement on a finer level of detail by classifying students by:

- Gender;
- Student age;
- study mode; and
- Grade category based on the final grade achieved in course.

To compare clustering patterns according to these classifications we use tangle-grams capturing student engagement in terms of course site visits, topic resources visits, and forum visits. For student age, we again categorised students as being either school leavers or mature age; study mode reflected whether the student took the course internally or externally, while grade category allowed for three possibili-ties namely that a student achieved a final grade of credit or higher, any other type of pass or a fail grade for the course.

The tanglegram shown in Fig. 8.31 depicts the difference in course site visit behaviour when comparing school leavers and mature age students. The dashed lines on the outermost edges of the dendrograms indicate unique nodes whereas solid lines indicate common subtrees. Thus when comparing school leavers to mature age students, starting from the top of the tanglegram we can see both tang-legrams displayed similar behaviour in terms of engagement with the course site in SwotVac (SV) and Exam Week 1 (EW1) whereas behaviours differed in Week 13 (Wk 13) and Exam Week 2 (EW2) as these nodes appear in different subtrees of their respective tanglegrams. Specifically, for school leavers visits to the course site were similar in Week 13 (Wk13) and Exam Week 2 (EW2) whereas for mature age students visits to the course site in Week 13 were closer in behaviour to the SwotVac and Exam Week 1 behaviours, while Exam Week 2 was closer to the visitation behaviour in Week 12 (Wk 12). The behaviours in Weeks 9 and 10 (Wk 9 and Wk 10 respectively) are identical in both age groups, however Weeks 7 (Wk 7) and 8 (Wk 8) differ considerably between the two groups. A similar interpretation carries through to the bottom of the tanglegram where the two-week mid-break behaviours (MB1 and MB2) are identical in this case. It is particularly interesting to note the behaviour around the final exam for which school leavers behave similarly in SwotVac and Exam Week 1 whereas from the clustering it appears mature age stu-dents also include Week 13 as part of the exam preparation behaviour.

As an interesting contrast to the tanglegram in Fig. 8.31 is the tanglegram for course site visit behaviour by gender (Fig. 8.32). Recalling that dashed lines in the tanglegram indicate unique nodes, it appears that there is no commonality of behav-iour in engagement, as measured via course site visitation, between males and females throughout the semester. Starting from the top of the dendrogram, females tend to treat the first 3 weeks of semester equally in terms of course visitation whereas male students treat Weeks 1 and 2 similarly but by Week 3 engagement behaviour has shifted. At the other end of the semester when comparing exam prep-aration via visits to the course site, female students treat SwotVac and Exam Week 1 with equal importance whereas male students treat Exam Week 1 in the same vein as Week 5 – for which there was a major assessment item — however do not include SwotVac as part of that preparation.

Again there is very little commonality when comparing students by study mode (Fig. 8.33). Course site visit behaviour tends to behave similarly in SwotVac and Exam Week 1 across both study modes. For external students however, their behav-iour in Week 13 is included as part of this exam period, whereas for internal students Week 12 is treated similarly to the exam period. This could be due to the practical examination that takes place in this week. On the other hand course site visit

behaviour is identical in the first 3 weeks of semester (Wk1–Wk3), in Exam Week 2 (EW2) and in the first week of the mid-term break (MB1).
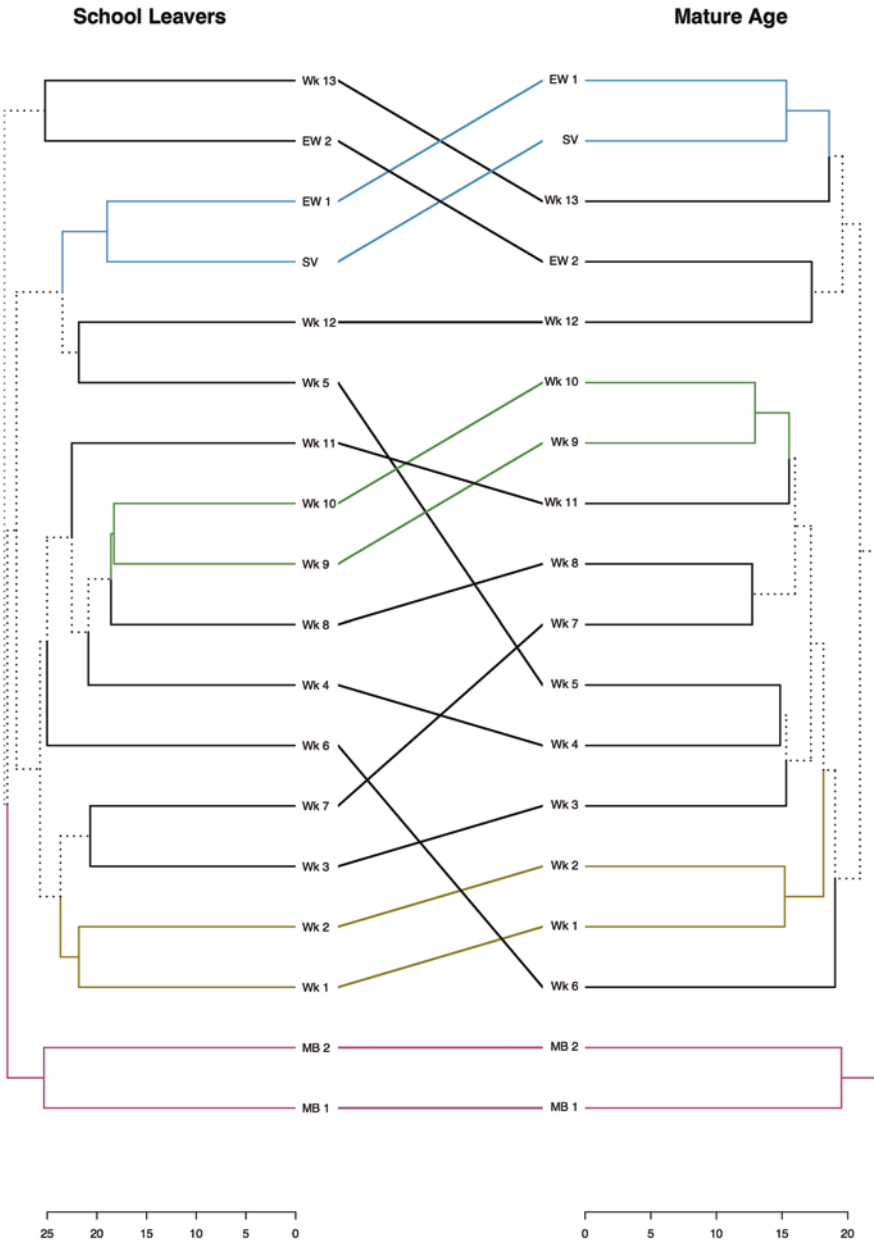

Fig. 8.31 Tanglegram of course site visits by age. Dashed lines indicate unique nodes in the dendrogram

Fig. 8.32 Tanglegram of course site visits by gender. Dashed lines indicate unique nodes in the dendrogram

External   **Internal**
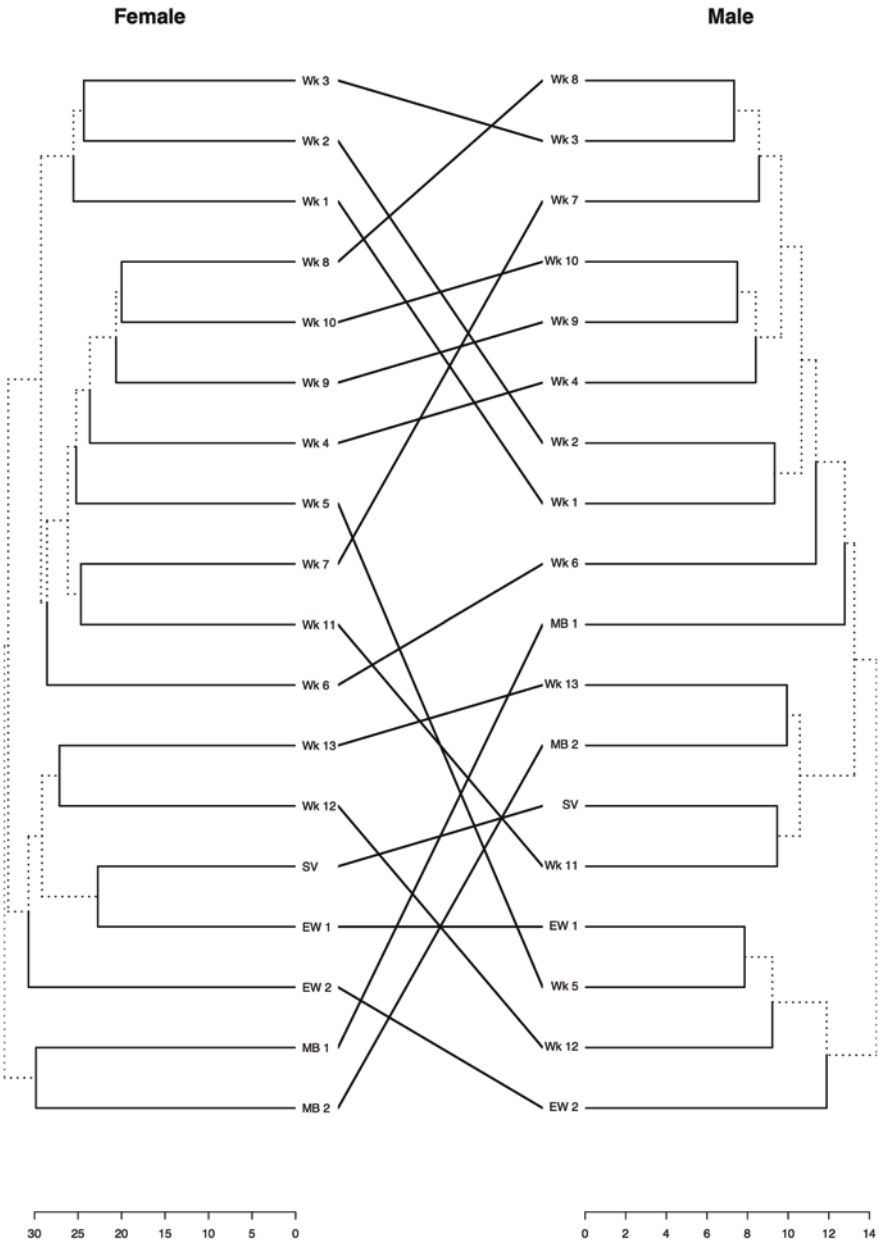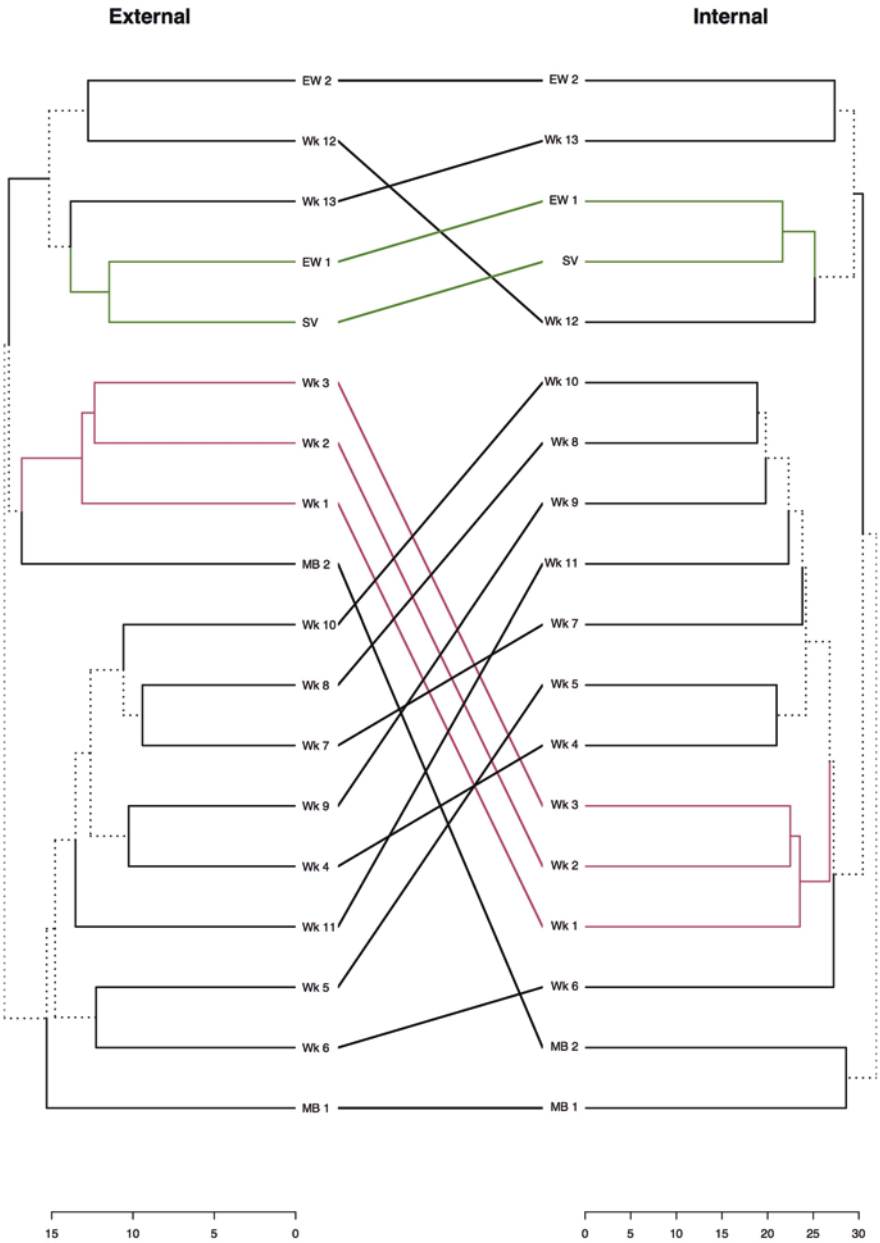

Fig. 8.33 Tanglegram of course site visits by study mode. Dashed lines indicate unique nodes in the dendrogram

Fig. 8.34 Tanglegram of course site visits based on students who failed versus those who passed although did not achieve a highly graded pass. Dashed lines indicate unique nodes in the dendrogram

Fig. 8.35 Tanglegram of course site visits based on students who passed versus those who achieved a highly graded pass. Dashed lines indicate unique nodes in the dendrogram

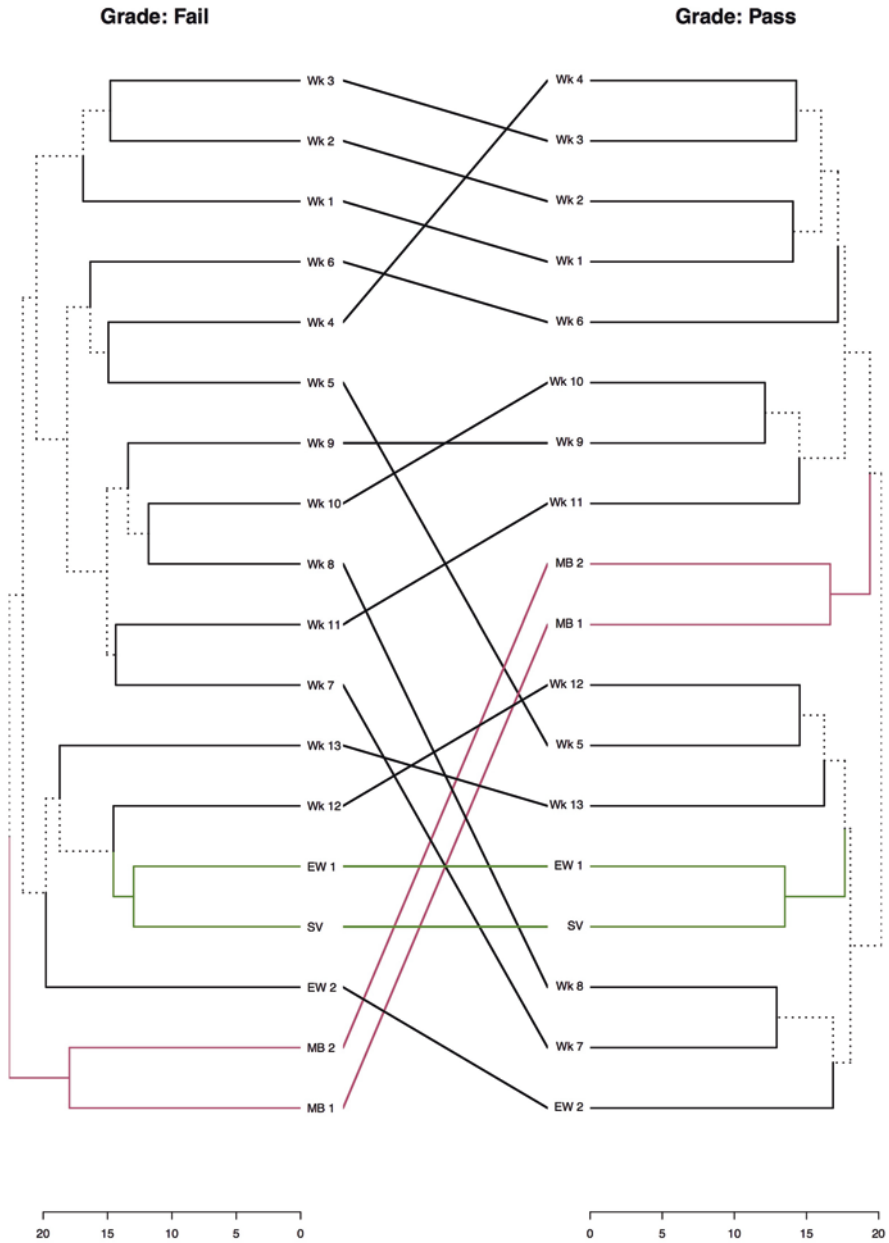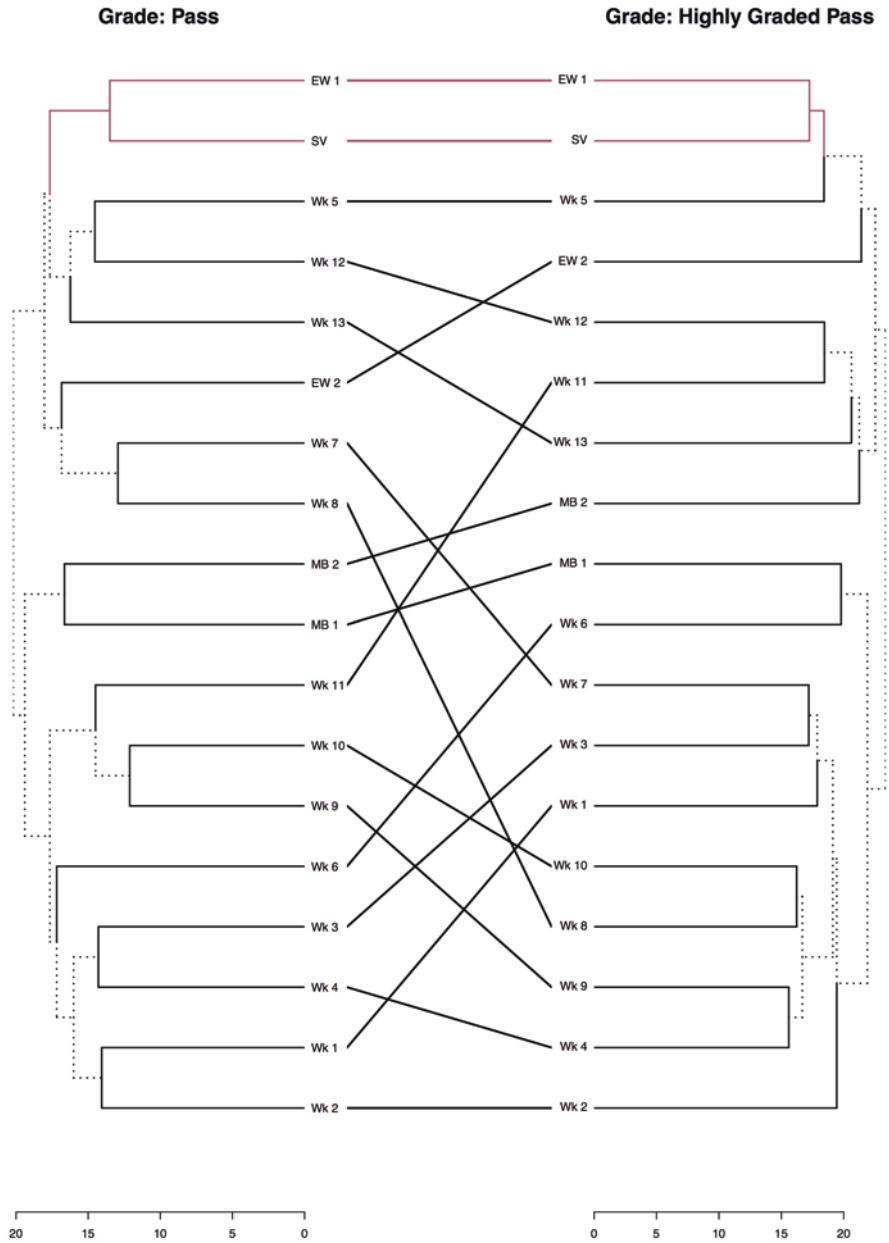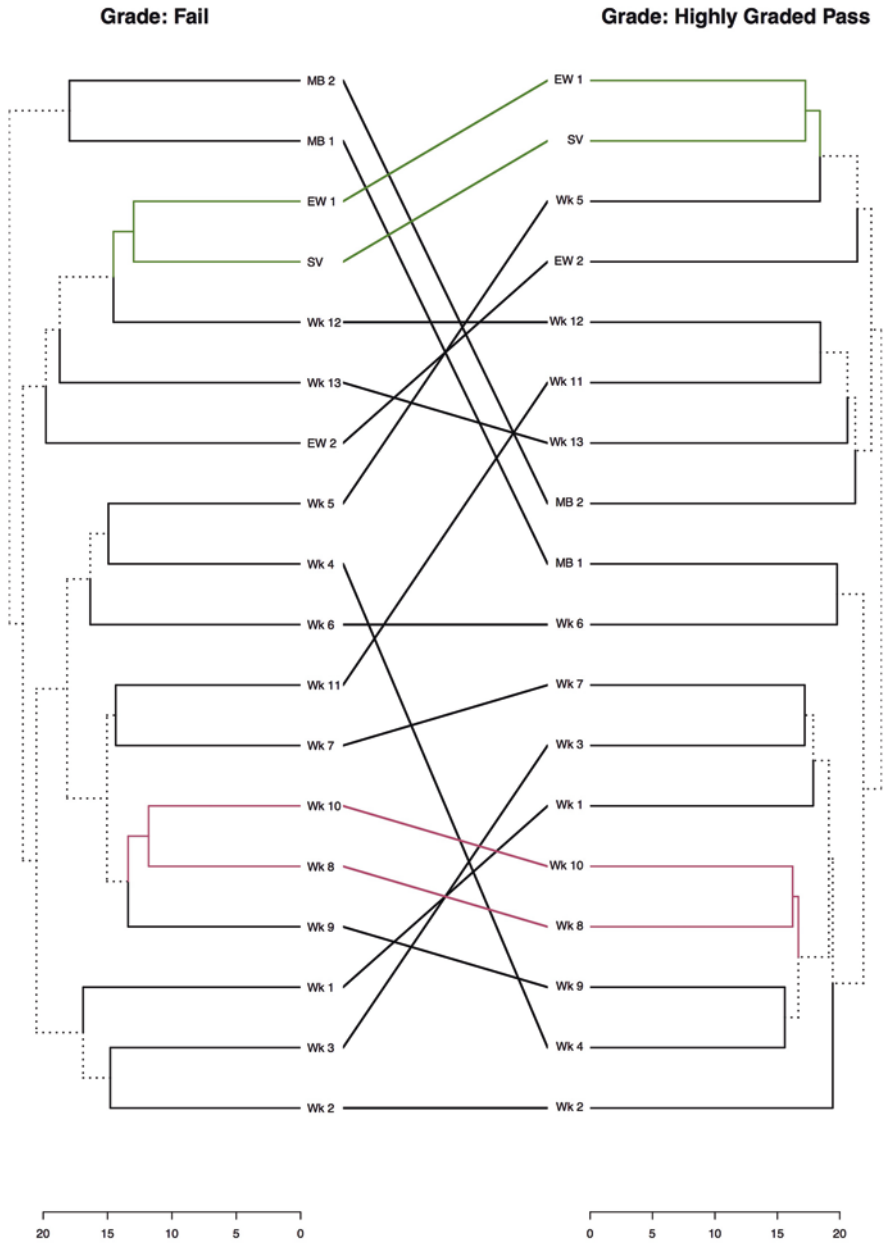**Fig. 8.36** Tanglegram of course site visits based on students who failed versus those who achieved a highly graded pass. Dashed lines indicate unique nodes in the dendrogram

The final set of tanglegrams are presented in Figs. 8.34, 8.35 and 8.36 and depict the differences in course site visits when considering students by the final grade achieved in the course. It is interesting to note that all three groups of students treated SwotVac and Exam Week 1 with equal importance. Students who failed or passed both treated the 2 week mid-term break in a similar fashion. However, students who achieved a highly graded pass distinguished between these 2 weeks by treating the first week of the mid-term break similarly to Week 6. Recalling there was a major assessment item in Week 5, students achieving a highly graded pass treated this week with equal importance to SwotVac and Exam Week 1. In contrast, students achieving a pass treated Week 5 similarly to Week 12 (practical examination) while students who failed did not view Week 5 as a major assessment week in that the tanglegram (Fig. 8.34) shows these students treat this week in a similar vein to Week 4, in which there was no assessment of any kind.

Comparing students who achieved a pass and those who achieved a highly graded pass (Fig. 8.35) those achieving a higher grade treated Week 12 as part of a larger block spanning Weeks 11–13 indicating preparation for the practical examination in Week 12, whereas students who passed with a lower grade displayed different behaviour in Week 11 and behaved as though Weeks 12 and 13 were similar. Students who failed compared to those with a highly graded pass treated Week 12 at a similar level of importance as SwotVac and Exam Week 1 (Fig. 8.36) however the lead-up preparation in Week 11 is missing; in this case students who failed behaved similarly in Weeks 7 and 11 – the third week of online assessment (Week 7) and the week after the final online assessment.

Investigation of engagement behaviour when viewing visits to topic resources and the forum indicated similar differences between the cohort classifiers used here. For space considerations these tanglegrams are shown here, however it is noted that these behaviours permeated across the course site irrespective of the resources utilised.

Conclusions

In this chapter, patterns of engagement were explored for three components of the online course environment, namely the course website, weekly topic resources, and the social forum. Engagement patterns were relatively consistent across the three online components, with the exception of the social forum, where engagement levels declined substantially over the course of the semester.

Across each course component, there is a selection of students with unusually high levels of weekly engagement. Students appear to engage most frequently during the early semester teaching weeks and the weeks leading up to the examination period. Internal and external students differ in their patterns of engagement, with external students consistently engaging at a higher frequency than their internal counterparts.

It was also seen that male and female students do not appear to differ substantially in their engagement behaviour with the exception that a higher proportion of male external students tend to disengage from the course earlier in the semester than their internal counterparts. Mature age students were found to typically have a higher frequency of access for course resources than school leaver students, with greater variability in their engagement counts per week.

When considering engagement patterns by final course grade, it was found that students who receive a credit or higher are engaging with the online course resources at a much greater frequency than students who receive a passing grade. Students who failed the course overall had considerably lower levels of engagement throughout the semester. Another distinction was that students achieving a credit or higher also viewed earlier major assessment items in the same vein as the final exam at the end of the exam period.

With regards to study mode, it was determined that for internally enrolled students and externally enrolled female students, the weeks with the greatest engagement are Weeks 1–5 and the first week of the exam period. For external male students however, the key teaching weeks were Weeks 1 and 3 and the late-semester weeks for SwotVac and the examination period.

When considering the engagement trajectories of individuals over the course of the semester, it was found that high-engaging students tend to remain highly engaged whereas students in the lowest quartile of engagement tend to maintain relatively low engagement each week. Moreover, internal students are more likely to have lower online engagement levels than their external counterparts.

What we are able to learn from learning analytics appears to depend crucially on the structure of a virtual learning environment as well as decisions made by online engagement data custodians. There is a potential disconnect between what data is available and what data is useful, making learning from learning analytics challenging. For the course on which analyses in this chapter are based, the main difficulty lay with the interpretation of weekly counts which were obtained from the virtual learning environment access logs. Without more detailed information of how student activities contributed to these counts, we were able to only gain some insights in relation to student engagement. That being said, based on our analysis, we can observe that external and mature age students appear to engage with virtual learning environments differently. A higher proportion of external students tend to disengage from the online resources, particularly if they are male and school leavers. On the other hand, once they engage, mature age students appear to maintain strong online presence throughout the semester. Our analysis also leads us to conclude that assessment, both formative and summative, is one of the deciding factors whether a first year student engages with online course resources. This suggests the need for differentiation of online support resources and advice given to students on what contributes to study success.

References

Beer, C., Clark, K., & Jones, D. (2010). Indicators of engagement. In C. Steel, M. Keppell, P. Gerbic, & S. Housego (Eds.), *Curriculum, technology & transformation for an unknown future. Proceedings ASCILITE Sydney 2010* (p. 75–76)

Coffrin, C., Corrin, L., de Barba, P., & Kennedy, G. (2014). Visualizing patterns of student engagement and performance in MOOCs. In *Proceedings of the fourth international conference on learning analytics and knowledge LAK'14* (pp. 83–92). ACM.

Ferguson, R., & Clow, D. (2015). Examining engagement: Analysing learner subpopulations in massive open online courses (MOOCs). In *Proceedings of the fifth international conference on learning analytics and knowledge LAK'15* (pp. 51–58). ACM.

Hasan, M., & Duan, Z.-H. (2015). Hierarchical k-means: A hybrid clustering algorithm and its application to study gene expression in lung adenocarcinoma. In H. Arabnia & Q. Tran (Eds.), *Emerging trends in computational biology, bioinformatics, and systems biology* (pp. 51–67). Elsevier.

Henrie, C., Bodily, R., Manwaring, K., & Graham, C. (2015). Exploring intensive longitudinal measures of student engagement in blended learning. *The International Review of Research in Open and Distance Learning, 16*(3), 131–155.

Hughes, G., & Dobbins, C. (2015). The utilization of data analysis techniques in predicting student performance in massive open online courses (MOOCs). *Research and Practice in Technology Enhanced Learning, 10*(10). https://doi.org/10.1186/s41039-015-0007-z

Hung, J.-L., & Zhang, K. (2008). Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching. *MERLOT Journal of Online Learning and Teaching, 4*(4), 426–437.

Kizilcec, R., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge LAK'13* (pp. 170–179). ACM.

Phan, T., McNeil, S., & Robin, B. (2015). Students' patterns of engagement and course performance in a massive open online course. *Computers & Education, 95*, 36–44. https://doi.org/10.1016/j.compedu.2015.11.015

Ramesh, A., Goldwasser, D., Huang, B., Daume, H., III, & Getoor, L. (2014). Learning latent engagement patterns of students in online courses. In *Proceedings of the twenty-eighth AAAI conference on artificial intelligence* (pp. 1272–1278). AAAI Press.

Rousseeuw, P. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Computational and Applied Mathematics, 20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Sclater, N., Peasgood, A., & Mullan, J. (2016). *Learning analytics in higher education. A review of UK and international practice*. Retrieved from https://www.jisc.ac.uk/sites/default/files/learning-analytics-in-he-v3.pdf

Siemens, G., & Gasevic, D. (2012). Guest editorial – Learning and knowledge analytics. *Educational Technology & Society, 15*(3), 1–2.

Thorndike, R. (1953). Who belongs in the family? *Psychometrika, 18*(4), 267–276.

Belinda A. Chiera is a senior lecturer in Statistics and the Program Director for Service Teaching at the School of IT & Mathematical Sciences, University of South Australia. Belinda has a wide and varied range of interests including social network analysis, modelling of high-dimensional networks and enjoys uncovering insights hidden in large databases. She also works on a number of projects with a particular focus on the student wellbeing and data analytics in support of student learning and academic success.

Małgorzata W. Korolkiewicz is a senior lecturer in Mathematical Sciences and the Program Director for Data Science at the School of IT & Mathematical Sciences, University of South

Australia. Her research interests include mathematical finance, risk management, environmental modelling, as well as data analytics and challenges posed by Big Data. She is particularly interested in developing students' statistical thinking and providing them with the skills needed to derive insights from data. She is a high-quality university teacher who in 2011 was awarded the UniSA Award for Teaching Excellence.

Lisa J. Schultz is a Research Associate in the Centre for Industrial and Applied Mathematics at the University of South Australia. She was awarded a PhD in mathematics and statistics in 2016 and has a broad range of research interests, including environmental modelling, student health and well-being, outlier analysis, and data analytics. She has a keen interest in uncovering the hidden stories in large datasets through statistical analysis and is a member of the Society for Learning Analytics Research as well as the International Association for University Health and Wellbeing.