Learner Analytics

Predicting Student Success: Assessing the Strength of the Relationship between Student Engagement and Attainment

Hannah Maltby/Dan Swain/Sami Karamalla-Gaiballa

Planning Support Office, University of Manchester

January 2017

The research examines whether high levels of student contact, as measured by learner analytics measurements, are linked to better student academic performance. A number of measurements linked to the electronic 'footprint of a student' were used to measure levels of engagement including accessing digital learning resources (Blackboard) and physical learning spaces (computer clusters and library facilities). In addition other factors that also impact on attainment (e.g. entry qualifications, gender, academic subject and ethnicity) were considered in the regression modelling to isolate the impact of student engagement.

Students study in a variety of ways involving physically attending the University and accessing electronic resources and the learner analytics measurements used in the research only partially covers the plethora of ways that students can engage in learning. The research identified that students that had high levels of engagement with Blackboard and spent more time in the library were more likely to gain a good degree and there were distinct patterns across Academic Schools and programme years. The analysis suggests that computer cluster log-ins and Alan Gilbert Learning Commons access were weaker learner analytics measurements and could not be used to reliable predict attainment. Many of the datasets were not available over sufficient historical periods to facilitate a thorough analysis of all the datasets across the range of students attending the University and further research is required to evaluate how the learner analytics measurements can be used to identify students that struggle academically and take periods of leave of absence.

Contents

Introduction

The University of Manchester is committed to providing an outstanding student experience and amongst other initiatives, has implementing the 'My Student' project which aims to increase the availability of educational data about individual students through the provision of learner analytics. This research project is part of the evidence base for the application of learner analytics and investigates the relationship between student engagement measurements and the key student outcomes of attainment and retention. The analysis is an initial scoping exercise to review the strength of a variety of learner analytics measurements and this includes a review of the availability and accessibility of the data sources. It provides the analysis to underpin the development of a rich dataset which academic advisors can use to inform their personal interventions with students.

The HEI sector has experienced an increase in the availability of educational data and this is providing the potential for learner analytics to 'become a powerful means to inform and support learners, teachers and institutions in better understanding and predicting personal learner performance' (Greller and Drachsler, 2012)[1]. Exploiting the large datasets being collated on student activity through retrieval technologies holds substantial promises for use in education (Johnson et al, 2011)[2]. The potential of learner analytics must be understood alongside the potential dangers as the statistical quantitative analysis produces statistical norms for assessing students and any divergence from these norms requires further investigation to understand the reasons (Greller and Drachsler, 2012). When used effectively learner analytics should provide gains in terms of 'establishing acts of automatic decision making for learning paths' (ibid) which will save time for more personal interventions. The definitions of learner analytics below highlight the common theme of using student driven data to support teaching and learning.

> *'… the collection, analysis, use, and appropriate dissemination of student-generated, actionable data with the purpose of creating appropriate cognitive, administrative, and effective support for learners' (Oblinger 2012 cited in Slade and Prinsloo 2013: 1512)*[3]

> *'Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs'(Ferguson 2012: 305)*[4]

Learner analytics has been adopted across numerous Higher Education (HE) institutions that provide various levels of data to staff and students. This research is an initial scoping exercise to ascertain the value of various measurement of student engagement and their value as predictors of academic performance. Learning analytics is a fast-growing area of Technology-Enhanced Learning (TEL) research. It has roots in business intelligence, web analytics,

---

[1] Greller, W. and Drachsler, H. (2012) Translating learning into numbers: A genetic Framework for learner Analytics. *Educational Technology & Society, 15(3), 42-57.*
[2] Johnson, L., Adams, S. and Cummins, M. (2012) The NMC Horizon Report: 2012 Higher Education Edition, The New Media Consortium, Austin, Texas.
[3] Slade, S. and Prinsloo, P. (2013) 'Learning Analytics: Ethical Issues and Dilemmas', *American Behavioural Scientist*,57(10),1510–1529.
[4] Ferguson, R. (2012) 'Learning analytics: drivers, developments and challenges', *International Journal of Technology Enhanced Learning*, 4(5/6) pp. 304–317.

educational data mining and recommender systems (Ferguson 2012: 304)[5]. In terms of academic origins, it is an intersection of various methodologies, theories and research areas including social network analysis, latent semantic analysis, and dispositions analysis (Ferguson 2012 cited in Slade and Prinsloo 2013: 1511; Gašević et al 2015:65).

The main purpose of learner analytics is the monitoring of students' performance which allows for early targeting of those students who are struggling. This is achieved by both provision of proactive feedback as well as predictions of students' learning success. Certain institutions use a data-mining algorithm which has a signalling system (high risk-red, moderate risk-orange and no risk-green) to identify poor performing students (Picciano 2014:35[6]; Dawson, Gašević, Siemens, & Joksimovic, 2014 cited in Gašević et al 2015: 65[7]; Rubel and Jones 2016: 145-6; Arnold & Pistilli, 2012 cited in Gašević et al 2015:65). Further benefits of learner analytics are improvements in regard to efficient workings of institutions, for instance, learning analytics can be used for cross-institutional analysis (Rubel and Jones 2016: 145-6[8]). As institutions learn more and more about students learning patterns they can make better informed choices as well as respond faster to identified needs of students (Oblinger, 2012 cited in Slade and Prinsloo 2013: 1513[9]). Some argue (Soby 2014: 90[10]) that analytics can '…potentially help transform education from a standard "one-size fits- all" delivery system into a responsive and flexible framework, crafted to meet student academic needs and interests'. There is also a potential to increase institutional transparency (Long and Siemens 2011 cited in Rubel and Jones 2016: 146).

Learner analytics offers the opportunity to apply a vast quantity of data to help understand the behaviour of students but there are a number of challenges when working with learner analytics data.

- **Technological**: there is a need to understand how to work with large datasets effectively (Ferguson 2012: 312-313). In society today there are datasets so large that they are beyond the ability of typical software tools to manage them (Manyika et al., 2011 cited in Ferguson 2012: 306-7). EDUCAUSE Centre for Applied Research survey (Bichsel, 2012 cited in Slade and Prinsloo 2013: 1513-4) indicated that the greatest concern related

---

[6] Picciano, A. G. (2014) 'Big Data and Learning Analytics in Blended Learning Environments: Benefits and Concerns', *Journal International Journal of Interactive Multimedia and Artificial Intelligence- Special Issue on Multisensor User Tracking and Analytics to Improve Education and other Application Fields,* 2(7), pp.35-43.
[7] D. Gašević, S. Dawson and G. Siemens (2015) 'Let's not forget: Learning analytics are about learning', *TechTrends*, January/February, 59 (1), pp.64-71.
[8] Rubel, A. and Jones, K. M. L. (2016)'Student privacy in learning analytics: An information ethics perspective', *The Information Society*, 32 (2), pp. 143–159.
[9] Slade, S. and Prinsloo, P. (2013) 'Learning Analytics: Ethical Issues and Dilemmas', *American Behavioural Scientist*,57(10),1510–1529.
[10] Søby, M. (2014) 'Learning Analytics', *Nordic Journal of Digital Literacy*,9(2), pp. 89–91.

to the growing use of learning analytics is the financial cost of implementation rather than issues related to privacy (Slade and Prinsloo 2013: 1513-4).

- **Ethical**: issues related to informed consent, privacy and the deidentification of data are problematic and complex (Slade and Prinsloo 2013: 1515). One of the key challenges for learning analytics users is to develop an acceptable ethical guidance framework (Ferguson 2012: 312-313).

The quote from Ferguson below indicates that a number of current areas of policy and practice are coming together that has create a growing demand for learner analytics and along with the Teaching Excellence Framework (TEF) means it is crucial to understand how student learn to provide effective academic support.

> *'A review of key reference points for the field shows that a combination of the availability of big datasets, the emergence of online learning on a large scale, and political concerns about educational standards has prompted the development of this field. Learning analytics are distinguished by their concern for providing value to learners, whether in formal, informal or blended settings. They are employed to understand and optimise both learning and the environments within which it takes place.' (Ferguson 2012: 314)*

The project deliverables were requested and refined through consultation with senior PSS staff (Emma Hilton-Wood, Head of Academic Policy) and the results will be fed back into the teaching and learning agenda at the University. The work will be closely aligned with the HEFCE funded project examining 'Learning Gain' which is being supported by the Planning Support Office and complements this work through specific targeted research that examines key areas of the teaching and learning experience of students. The work is realted to a number of the themes underpinning the TEF specifically as institutions will be required to demonstrate a commitment to individual support through high levels of student retention and attainment.

Document Structure

The initial section below details the methodology used within the research and examines some of the data limitations that have been imposed on the research. The key lessons learn through the research are then detailed which includes reflections on data quality and accessibility, and an executive summary of the main research findings. The next section details the descriptive analysis based around a series of line charts profiling student engagement across the learner analytics measurements categorised into attainment groups. Subsequently the report includes a summary of the regression analysis that highlights which learner analytics measurements are significant predictors of attainment. The final section details a case study of the relationship between retention and the learner analytics measurements.

Methodology

Sample

The analysis used four of the University's Academic Schools (Social Sciences, Mathematics, Psychological Sciences and Physics & Astronomy) as exemplars for learner analytics analysis. The research followed the academic careers students of that fit the following criteria[11]:

- began their first degree programme within the academic year 2012/13;
- graduated within the academic year 2014/15; and
- had not repeated a year of programme

Analysis

The learner analytics analysis focused on data within the following areas:

- Alan Gilbert Learning Commons (AGLC) - *number of swipe access entries*
- Blackboard- *number of logins*
- Library Access - *number of swipe access entries*
- Library Time – *Time spent in library*
- Library Loans – *Physical loans*
- Computer clusters- *number of logins*

The six learner analytics measurements detailed above were examined in the context of degree attainment: both in terms of the binary good and lower classifications and the full degree classification categories[12] in order to identify which engagement measurements are associated with attainment. The descriptive analysis section includes data visualisation of the cumulative average monthly usage trends between those who achieved a good or lower degree and also across the full classification system. The learner analytics data visualisations display the monthly average usage of student cohorts classified by attainment and is based on cumulative monthly totals to highlight usage trends across the 3 year data period. Regression analysis is then used to control for the impact of other characteristics such as gender, domicile, age and quality of entry qualifications (tariff score) to determine the significance of each particular area of student engagement on attainment.

[11] The exception to the criteria is the School of Maths which was used as a pilot School initially and the data period is 2011/12 to 2013/14.
[12] Students gaining a 3rd or pass have been drawn together into one category due to small ample sizes in each classification.

The analysis was primarily based around attainment but a section considers the relationship between retention and the learner analytics measurements. However, it must be noted that the non-continuation population is based on a low number of students therefore only the School of Social Science (SOSS) was used within the retention analysis as this had a sufficiently large sample population to produce reliable results. The research uses the HESA non-continuation population and this approach involves tracking a cohort of new entrants over one year. The analysis considers continuation rates and does not take progression into account, therefore, does not examine the students who have not progressed from year 1 to 2. In this document non-continuation students are defined through being absent from the University of Manchester. The analysis presented highlights non-continuation patterns over programme year one. The methodology differs from the HESA definition as HESA identify students who drop out of HE altogether. The non-continuation student population as defined by HESA can be summarised as:

- Undergraduates
- Full-time
- First year of degree or foundation year
- Students are defined as non-continuation if they leave in the year after the 1st December in their first academic year.
- Includes students who transferred to other institutions

The analysis is a scoping exercise to examine the strength of each learner analytics measurement in relation to attainment. Research has shown that there are a range of factors linked to attainment that should be considered within the context of the learner analytics data including gender[13], ethnicity[14], and entry qualifications[15]. The authors attended a number of conferences linked to learner analytics and consulted on different methods of incorporating other factors linked to attainment into learner analytics data. The recommended method is to provide contextual data alongside the learner analytics data so that academic advisors can understand the range of other factors that may impact on attainment alongside student engagement as measured by learner analytics. Included in Appendix 4 is a mock-up of a student profile detailing the type of information that would be useful for academic advisors assessing learner analytics data and this will need to be refined through consultation with IT services to identify what data is available and with system users to understand their requirements.

The project will play an important role in establishing the evidence base for the use of learner analytics services in support of educational practice and learner guidance at the University of Manchester. Learner analytics aims to present and contextualise educational data allowing it to be used to assess the

[13] Leman, P. J. (1999). The role of subject area, gender, ethnicity and school background in the degree results of Cambridge undergraduates, The Curriculum Journal, 10:2, 231-252.
[14] Leslie, D. (2005). Why people from the UK's minority ethnic communities achieve weaker degree results than whites. Applied Economics, 37, 619-632
[15] Hoare, A., Johnson, R. (2010). 'Widening participation through admission policy – a British study of school and university performance'. Studies in Higher Education, 1-21.

performance of learners. The project will explore the link between student engagement proxies and key measurements of student academic outcomes and retention. It will provide the analysis to underpin the development of a rich dataset which academic advisors can use to inform their personal interventions with students.

Data Limitations

- When this research was initiated historical data was only available for a limited period across some of the data sources (see appendix 1). The Alan Gilbert Learning Commons (AGLC) only opened in October 2012 but then closed for a period and re-opened permanently in Jan 2013 so access data is only available from this point. Computer cluster access data has only be stored since 2012/13 and Blackboard data from 2011/12. This has limited the research in relation to examining students that do not progress without any interruptions through their academic studies. These students are important in the context of understanding how learner analytics can be used to identify students that may be struggling academically and further work is required to investigate how learner analytic relates to their academic performance and retention.
- The data is based on degree classification which represents a relatively broad measurement of attainment and further work is required to examine the relationship between course unit marks, end of year marks, final degree mark and the learner analytics measurements.
- The analysis examined the relationship between learner analytics measurements and attainment across 4 Academic Schools and cohorts of students progressing from entry to graduation. The analysis would benefit from a wider range of Academic Schools being considered from all subject areas including medicine and humanities. The scope of the research was limited by resources, the accessibility of the data and the size of the datasets.
- The data is quantitative in nature and does not capture the qualitative aspects of student engagement across the learner analytics measurements. For example, students may access computer clusters for a range of academic and non-academic purposes which is not captured in the data. The computer cluster data indicates that students are on campus accessing resources but the purpose and extent of the access is not captured.

Learning Points – Data Quality and Accessibility

- Many of the data sources used in this research were not easily accessible and the data storage methods will need to be reviewed to facilitate the data being used as a source of learner analytics. For example, the most recent (April 2013 onwards) library physical loan data is stored in the Library Management System's built-in reporting tool, OBIEE (Oracle business Intelligence). Any older data related to physical loans is stored in the data warehouse in a static database.
- Blackboard data represents a huge dataset and this research has not fully explored the potential that this data has for measuring student engagement. The data is not used uniformly across programmes and this makes it difficult to apply a systematic learner analytics methodology to

the data. The analysis below uses Blackboard log-ins as a measurements of engagement but other quantitative and qualitative measurements are available that may enhance the learner analytical value of the data.

- The size of the datasets has made it difficult to analyse the data and resulted in four Academic Schools being chosen as case studies. Blackboard in particular in a huge dataset as it captures all student activity within the Blackboard system and any Learner analytics system will require the capacity to store and process large volumes of data.
- Many of the datasets used in this report are not regularly used across the institution and have not been subject to data quality testing. Further investigation is required to establish the validity of the data sources. For example, time spent in the library is defined by student swiping in and out of the library but further work is required to establish whether there are occasions when student may not swipe out of the library (e.g. closing times) and the impact this may have on data quality. A number of student records also showed that students had visited the library but there was not time allocated to the visits. This may be due to student simple returning books or it may highlight data quality issues.

Executive Summary

- Students are taught and study in a variety of ways the University with a range of study opportunities linked to physically attendance at the University (e.g. seminars, lectures, library usage and AGLC usage) and a range of electronic resources that students can access on-line either on or off campus (e.g. Blackboard, podcasting lecture capture, electronic library facilities and web-based research). The plethora of learning opportunities means it is difficult to capture the full learning experience of students. The learner analytics analysed below should be viewed as capturing aspects of students' learning but will not provide a full picture of their engagement with the University. For example, it is currently not possible to accurately capture student attendance at timetabled teaching provision (e.g. lectures, laboratories etc.) or electronic downloads from library databases which are important aspects of many students' academic studies.
- The data shows that students have different learner analytics engagement patterns across Academic Schools which will reflect the demands of the different programmes and variations in students' studying styles. The data presented does not capture the reason behind the trends across the Academic Schools but it may results from Schools providing different levels of on-line resources to students and also variations in teaching and assessment styles.
- The descriptive analysis is based around trend charts over first degree year of programmes 1 to 3 that highlights peaks in academic engagement in the academic cycle linked to periods when academic exams and assessments are focused (end of semesters 2 and 3). The charts also indicate periods across the summer when student engagement is low (July- August).
- Across the different learner analytic measurements Blackboard access was the strongest predictor of academic attainment with student achieving a good degree having much higher levels of Blackboard access compared to those that obtained a lower degree, and this trend was significant when other variables where taken into account within the regression modelling. The majority of students access Blackboard for purely academic

purposes and are engaging with the variety of academic learning material available whilst students engage with the other learner analytics measurements for non-academic purposes.

- The analysis suggests that computer cluster log-ins and AGLC access levels are not strong learner analytics measurements in relation to attainment. These learner analytic sources show that students are physically attending the University but it does not necessarily capture students engaging in academic related activities. For example, students may access computers on campus for non-academic purposes such as using the internet for personal non-academic reasons (e.g. social media).
- The analysis shows that time spent in the library has potential as a learning analytics measurement but there were distinct patterns across Academic Schools with library time and attainment being more strongly linked to attainment in the Schools of Social Sciences (SOSS) and Physics & Astronomy compared to Maths and Psychology. Students can access a wealth of on-line library electronic resources without having to physically attend a University library. Given the wealth of on-line library electronic resources available to students, using this as a learner analytics measurement represents an important part of students' study experience and learner analytics data related to this area is required to provide a fuller picture of how students use library facilities.
- Attainment is influenced by a variety of factors beyond the learner analytics measurements analysed below and these must be considered within the context of the data. For example, International students have lower attainment levels than UK and EU students, and there are large variations across countries in attainment levels which is one of the attainment trends that will influence the interpretation of learner analytics data.
- The analysis examining retention and learner analytics suggests that students at risk of leaving the University early can be identified using learner analytics measurements but the case study is limited to SOSS and further analysis of other Schools is required to fully explore the relationship.
- The analysis is an initial scoping exercise to evaluate the relationship between a set of learner analytics measurements and attainment. The analysis can be used to inform future work in relation to information provided to academic advisor supporting students. The findings are based on aggregated data and it is important to understand that students can achieve academic success in a variety of ways using the teaching provision and academic support resources provided by the University. Learner analytics measurements provide a framework to initiate discussions with students regarding their academic progress but should be used as a starting point to explore how students are learning and engaging with the University.

Analysis

The information below details trends across the learner analytics measurements referring to data visualisations and further regression analysis is provided in the next section that highlights how the learner analytics measurements relate to a range of key variables linked to attainment. In the section below descriptions of the main trends in relation to the learner analytics measurements is provided and on the next page the relevant data visualisation are displayed.

Alan Gilbert Learning Commons

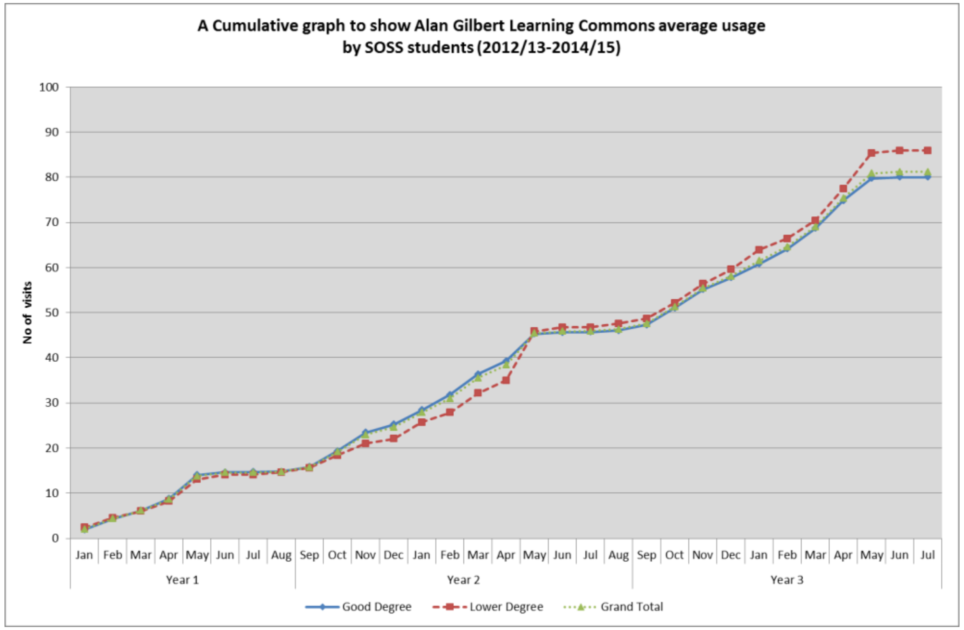
This section of analysis displays the relationship between the number of swipe access entries into the Alan Gilbert Learning Commons (AGLC) and degree attainment. Please note that data only begins in January 2013 as continuous data was not available until this point. The AGLC opened in October 2012 but closed for a period re-opening in January 2013. The AGLC offers 24/7 studying opportunities and has a range of individual and group studying facilities. The data suggests that student engagement as measured by the amount students access the AGLC is not a strong learner analytics measurement in relation to attainment. Students may access the AGLC for a range of purposes which will primarily be linked to study but could also relate to other activities such as non-academic computer usage.

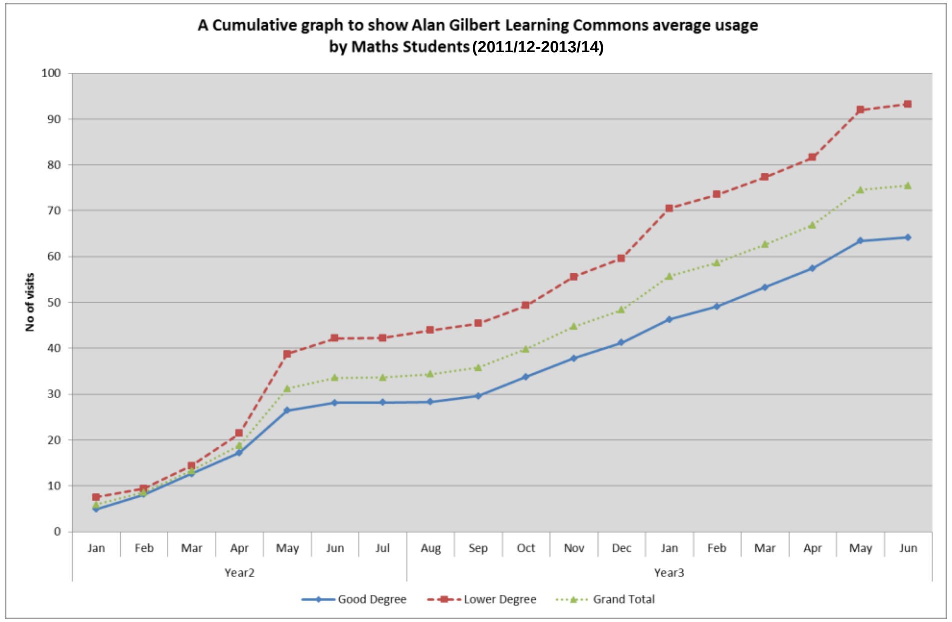
Good/Lower Classification

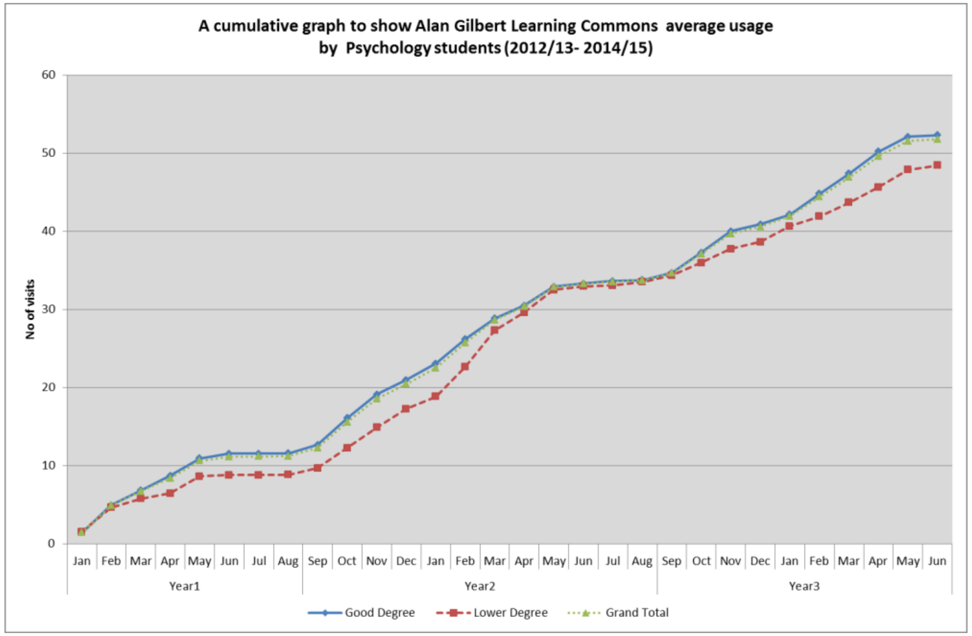
The profile of AGLC access across the degree classifications cohorts of good and lower degrees are fairly similar within SOSS (see chart below) particularly within the year 1 of programmes. For most of the second year of study those that achieved a good classification had a higher usage of the AGLC until May in which those with a lower classification had a slightly higher average number of uses which continued until the end of year three. The aggregated data suggest that overall students that gain a lower degree used the AGLC five more times than those who attained a good degree during their three years of study.

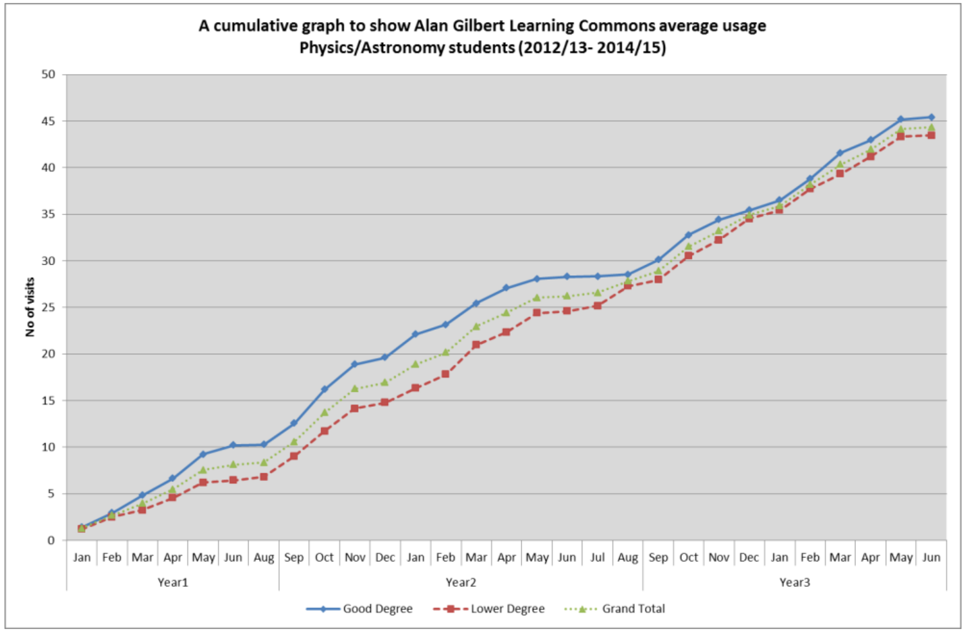
In terms of the School of Maths there were different usage patterns between the two cohorts with lower degree students utilising AGLC much more than their counterparts. The gap between the two groups emerged in April/May in the second year of study and was maintained until January in the 3rd year when it widened. The good degree cohort tended to have fairly consistent access patterns whilst the lower students had a number of high access months (e.g. April year 2 and January year 3).

The Schools of Psychology and Physics & Astronomy had very similar AGLC usage patterns across lower and good degree students. Across both schools good degree students tended to use the AGLC more in their first and start of second years whilst toward the end of second year and into third year the gap in access levels reduced and was minimal at the end of year 3.


A Cumulative graph to show Alan Gilbert Learning Commons average usage
by SOSS students (2012/13-2014/15)
No of visits
0
10
20
30
40
50
60
70
80
90
100
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul
Year 1
Year 2
Year 3
Good Degree
Lower Degree
Grand Total


A Cumulative graph to show Alan Gilbert Learning Commons average usage
by Maths Students (2011/12-2013/14)
No of visits
0
10
20
30
40
50
60
70
80
90
100
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun
Year2
Year3
Good Degree
Lower Degree
Grand Total


A cumulative graph to show Alan Gilbert Learning Commons average usage
by Psychology students (2012/13- 2014/15)
No of visits
0
10
20
30
40
50
60
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun
Year1
Year2
Year3
Good Degree
Lower Degree
Grand Total


A cumulative graph to show Alan Gilbert Learning Commons average usage
Physics/Astronomy students (2012/13- 2014/15)
No of visits
0
5
10
15
20
25
30
35
40
45
50
Jan Feb Mar Apr May Jun Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun
Year1
Year2
Year3
Good Degree
Lower Degree
Grand Total

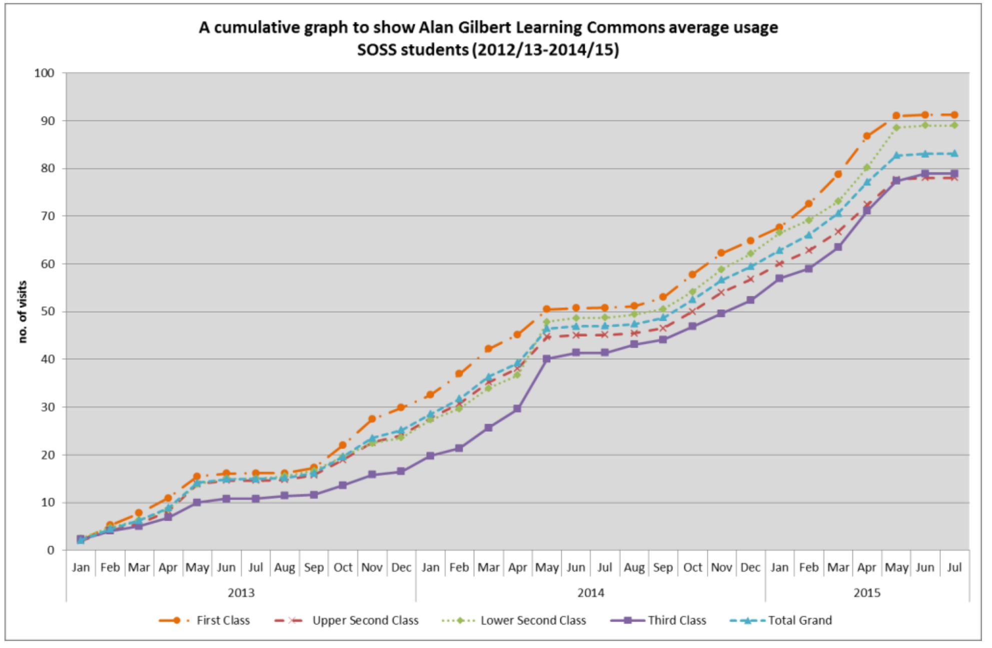
Full Classification[16]

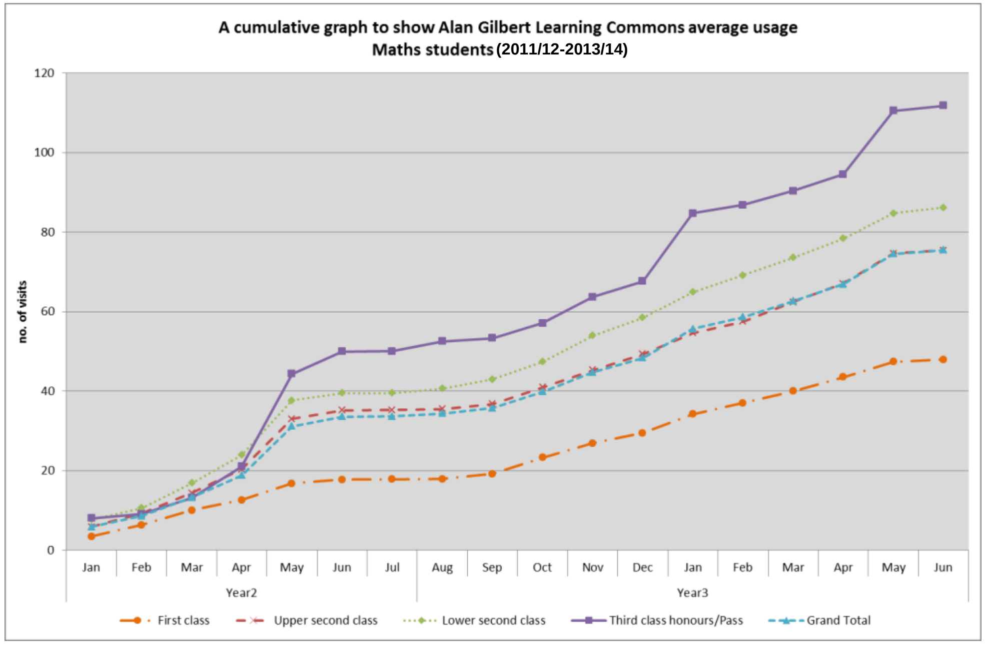
The chart below displaying AGLC access across the four categories of degree classification within SOSS shows that those who achieved a first class degree used the AGLC the most throughout the three years data period with a total average number of uses of 92. Those who achieved a third class degree had the lowest number of across the three years, with an average total number of 73. Interestingly, those who achieved a 2:1 and 2:2 held consistent patterns for AGLC usage within the first year of study. However, within the second year of study, 2:1 students held a slightly higher usage until May in which lower second class students then had a higher cumulative usage. Overall, the students that achieved a 2:1 or third/pass had very similar levels of overall AGLC access.

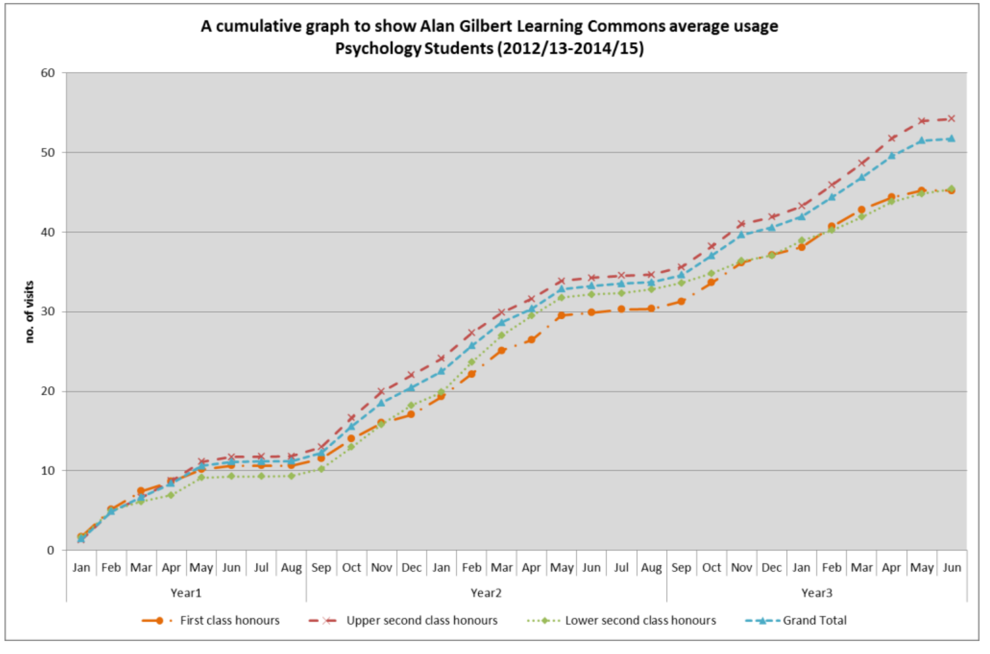
In the School of Maths 3rd/pass students had the highest overall levels of AGLC usage and 1st class students were the students with the least access. Differences in AGLC access between the degree classification groups first began to appear in April of the second year of study and increased again in January in the third year of study. The 2:2 Students used the AGLC more than those that achieved a 2:1 or 1st. These patterns of AGLC usage suggest that the relationship between attainment and AGLC access for Maths students is not a strong positive relationship.

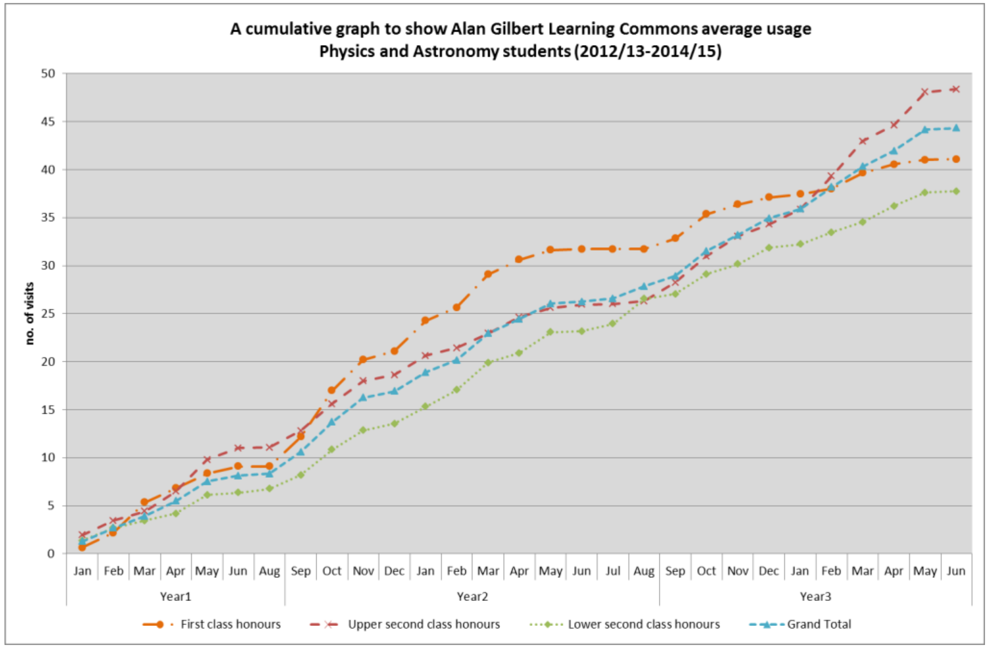
The Physics and Astronomy students that achieved a 1st had high levels of AGLC access until January in year 3 when 2:1 students had a large increase in usage and finished year 3 with the highest overall usage. Physics and Astronomy students that achieved a 2:2 had consistently lower AGLC usage than those that obtained a 1st or 2:1. The School of Psychology had clear patterns of AGLC engagement across the different attainment groups with students that achieved a 1st having the highest levels followed by 2:1 and 2:2 students, and this pattern was evident across the 3 year data period.

[16] The Psychology and Physic and Astronomy charts do not include 3rd/Pass categories as both samples were below 10 and not sufficiently robust data to include in the analysis as the data is based on the access behaviour of only a couple of students that accessed the AGLC from this sample


A cumulative graph to show Alan Gilbert Learning Commons average usage
SOSS students (2012/13-2014/15)
no. of visits
2013
2014
2015
First Class
Upper Second Class
Lower Second Class
Third Class
Total Grand


A cumulative graph to show Alan Gilbert Learning Commons average usage
Maths students (2011/12-2013/14)
no. of visits
Year2
Year3
First class
Upper second class
Lower second class
Third class honours/Pass
Grand Total


A cumulative graph to show Alan Gilbert Learning Commons average usage
Psychology Students (2012/13-2014/15)
no. of visits
Year1
Year2
Year3
First class honours
Upper second class honours
Lower second class honours
Grand Total


A cumulative graph to show Alan Gilbert Learning Commons average usage
Physics and Astronomy students (2012/13-2014/15)
no. of visits
Year1
Year2
Year3
First class honours
Upper second class honours
Lower second class honours
Grand Total

Blackboard logins

The following analysis details the relationship between the number of Blackboard log-ins and degree attainment. The data suggests that Blackboard is a strong learner analytics measurement and this may relate to students accessing Blackboard for academic purposes related to range of activities detailed below.

Blackboard Data Overview

Blackboard data potentially provides a huge learner analytics dataset that contains qualitative and qualitative data sources. It is important to note that there are various issues surrounding the use of Blackboard data with the main issue being processing the huge amounts of data linked to blackboard access and student activity in the system. There are various types of Blackboard activities that could potentially be used to measure student engagement including:

1. User log-ins
2. Overall activity as measured by the number of pages accessed in Blackboard
3. Number of unique pages accessed
4. Blackboard test activity
5. Unit survey activity
6. Assignment submissions
7. ‘My grades’ access
8. Blog access
9. Wiki access
10. Discussions board access
11. Number of downloads
12. Lecture files downloaded
13. Course reading material downloads

The analysis below only contains information on ‘log-ins’ Blackboard as opposed activity within the system which would give a more detailed picture of the students’ usage of Blackboard. Although there is data available concerning count and the type of pages students access within the system (i.e. the type of resource the student is opening) this dataset proved too large to process at this juncture.

The different forms of blackboard activity as detailed above are not always easy to extract from the blackboard data. Some of the fields require identification of text strings within fields. The system has not been configured to produce clean easily accessible learner analytics data. A wealth of qualitative text data will also be available for some students depending upon how they use Blackboard and this could be used as another value data source (e.g. discussion boards).

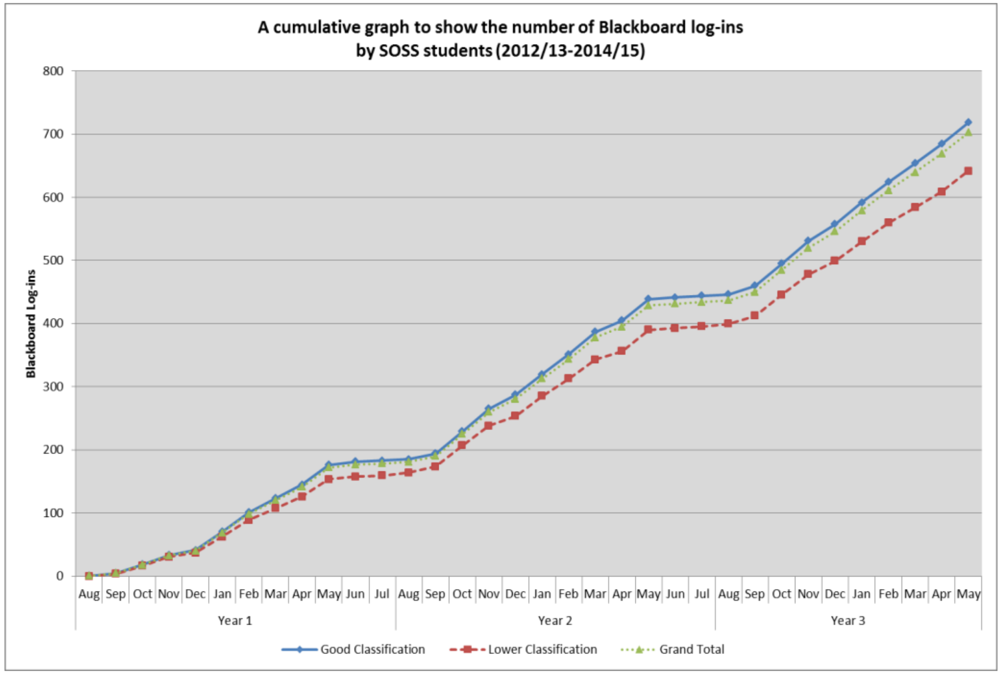
Another key consideration is the variance in Blackboard usage between degree programmes. There is no uniform structure in which particular resources must be made available by Blackboard administrators, therefore, usage between programmes, and, indeed modules can have a large amount of variance. For example, some modules have mandatory assessments posted on blackboard which students must complete and on other courses may have access to discussion boards which are utilised throughout the semester. In this way, fair comparison is not possible between programmes or modules due to this variation.

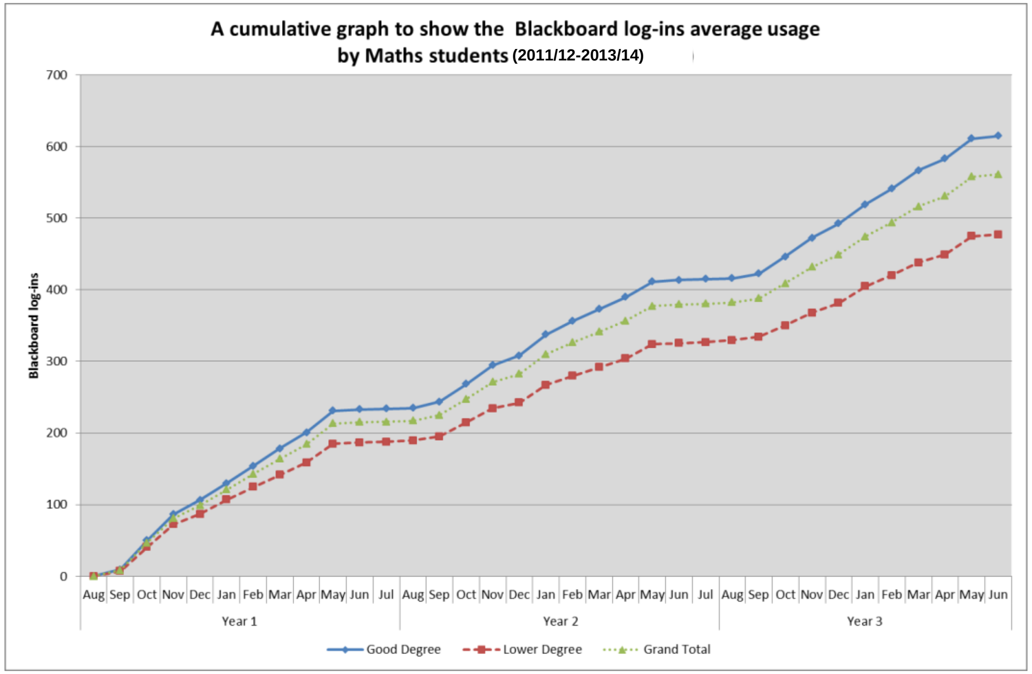
Good Classification

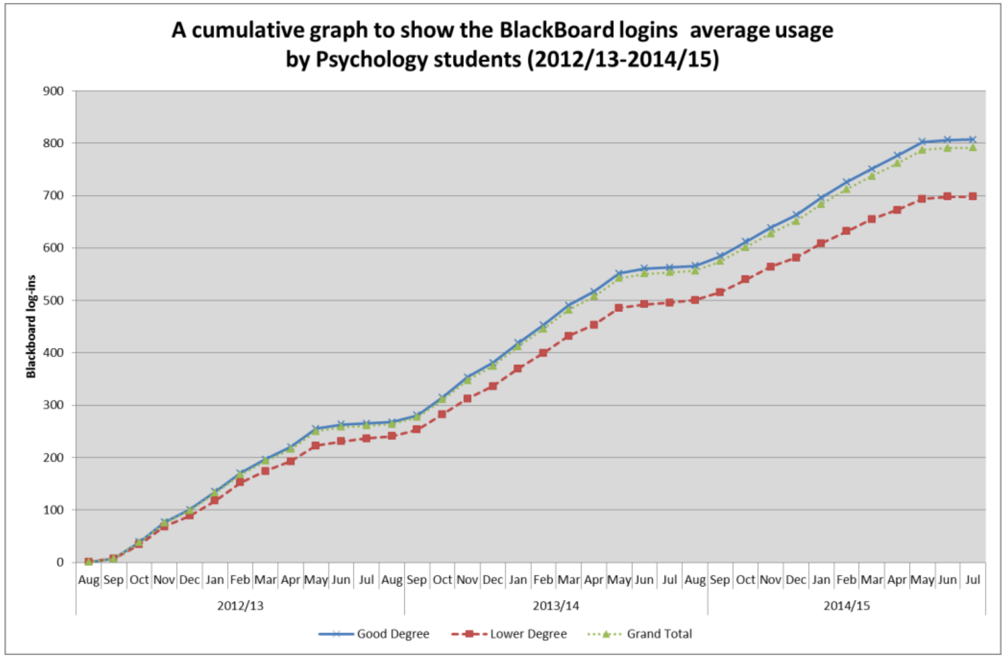
The figures below show that across the Schools of Psychology, SoSS and Maths both good and lower degree students followed the same trends in terms of Blackboard log-ins consistent month-on-month usage levels from Sept/Oct through to April/May. The charts show that Students that achieved a good degree classification consistently had higher levels of log-ins throughout all three years in three of the Academic schools with the exception being Physics and Astronomy.

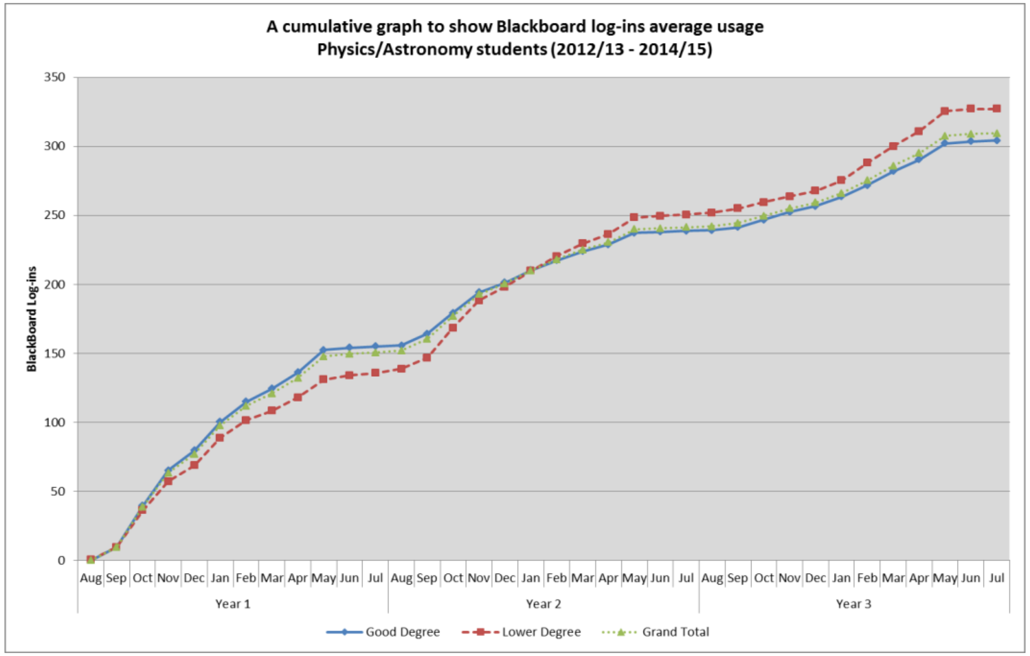
In terms of the School of Maths, there were large differences in Blackboard usage between good and lower degree students with good degree students on average logging into Blackboard 138 more times than lower students across the three data period years. The gap in relation to the number of log-ins increased across year 2 and was relatively stable in year 3 until the final few months. The same trend also occurred within the School of Psychological Sciences with a slightly smaller difference in overall log-ins between cohorts (n=100) but the gap between the good and the lower classifications emerging in year 2 and widened in the final few months of year 3.

The Blackboard log-in trends in the School of Physics and Astronomy was not consistent with the other three Schools. In year 1 good degree students were more likely to access Blackboard but this trend was reversed in year 2 when the lower degree student had much higher levels of access. Overall, the lower degree students had higher Blackboard access levels but across year 3 the levels were relatively similar across both groups.


A cumulative graph to show the number of Blackboard log-ins
by SOSS students (2012/13-2014/15)
Blackboard Log-ins
Year 1
Year 2
Year 3
Good Classification
Lower Classification
Grand Total


A cumulative graph to show the Blackboard log-ins average usage
by Maths students (2011/12-2013/14)
Blackboard log-ins
Year 1
Year 2
Year 3
Good Degree
Lower Degree
Grand Total


A cumulative graph to show the BlackBoard logins average usage
by Psychology students (2012/13-2014/15)
Blackboard log-ins
2012/13
2013/14
2014/15
Good Degree
Lower Degree
Grand Total


A cumulative graph to show Blackboard log-ins average usage
Physics/Astronomy students (2012/13 - 2014/15)
BlackBoard Log-ins
Year 1
Year 2
Year 3
Good Degree
Lower Degree
Grand Total

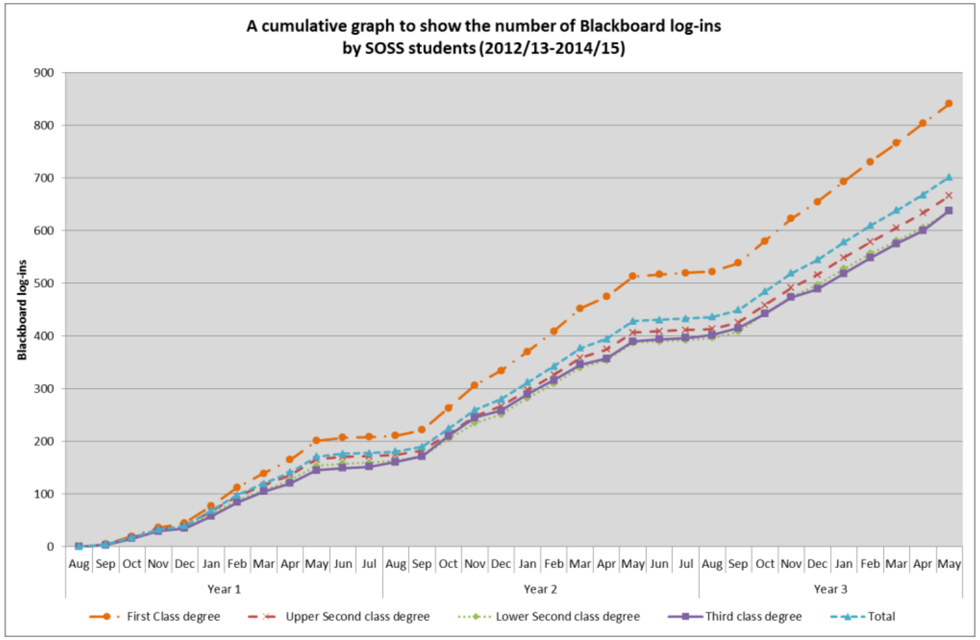
Full Classification

The full classification figure for SOSS (see chart below) shows that 1st class students have consistently higher Blackboard usage than the other degree classifications groups. The other three classifications held similar levels of usage particularly within year one and the first semester of year two, however, from this point onwards those who achieved a 2:1 classification had a slightly higher cumulative usage compared to those that achieved a 2:2 or 3rd/pass.

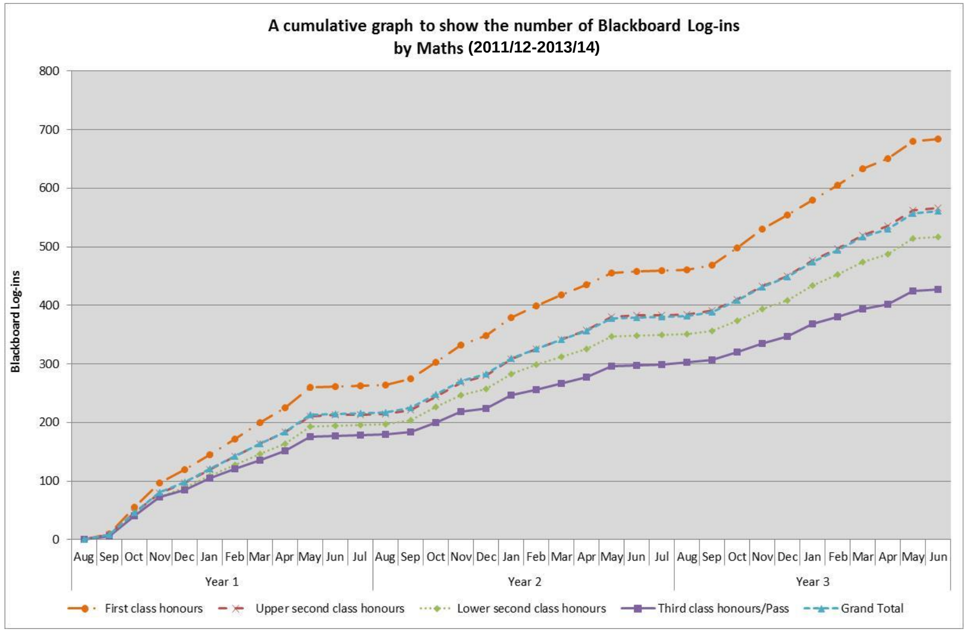
Within the School of Mathematics, Blackboard usage differences between attainment cohorts are pronounced with a general pattern of students that have higher levels of attainment using Blackboard more. The range of Blackboard log-ins over the 3 years was 189 to 1,340 per student. Analysis of trends across the full degree classification cohorts shows that students awarded a 1st class degree had the highest levels of log-ins followed by those achieving a 2:1, 2:2 and the lowest levels were within the student cohort awarded 3rd/pass . The students achieving a 1st class degree had much higher levels of log-ins particularly within the 2nd and 3rd programme year.

Surprisingly, those who achieved a third class degree within the School of Psychological Sciences had the second highest number of Blackboard log-ins but again this must be caveated with the small number of students within this cohort (n=2). Otherwise this School follows a similar pattern to Mathematics and SOSS but with higher level of log-ins compared to the other 3 Academic Schools. Students that achieved a 1st had the highest levels of Blackboard log-ins and they had particularly high levels of usage in year 3. Students that gained a 2:2 had the lowest levels of usage and the gap in usage between this cohort and the other degree classification groups emerged in year 2 and widened across year 2 and 3.

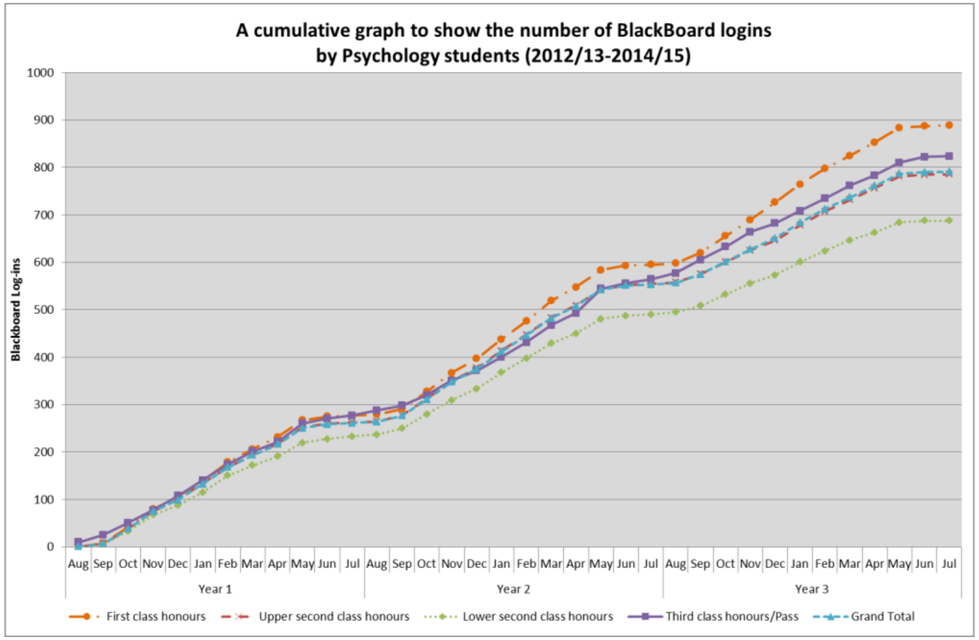
Physics and Astronomy students did not follow the Blackboard log-in and attainment patterns that were apparent across the other three Schools. Students that achieved a third/pass had the highest levels of Blackboard log-ins and the other three degree classification cohorts had very similar levels of Blackboard log-ins. At the start of the 2 year the 3rd/pass students had a peak in Blackboard usage and the gap in usage levels between this group and the other degree classification cohorts was then sustained until the end of the 3rd year.



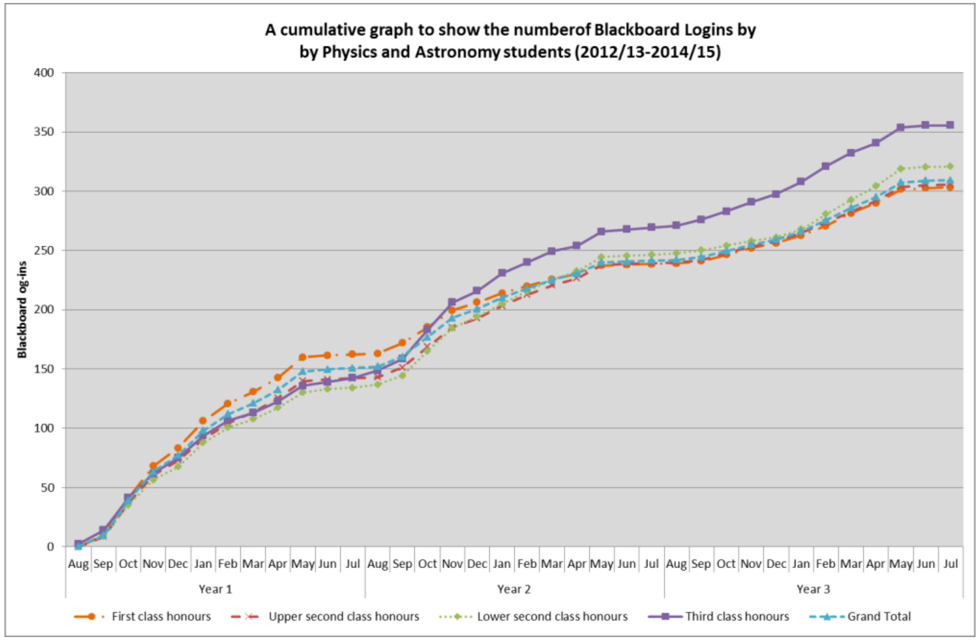
A cumulative graph to show the number of Blackboard log-ins
by SOSS students (2012/13-2014/15)
Blackboard log-ins
Year 1
Year 2
Year 3
First Class degree
Upper Second class degree
Lower Second class degree
Third class degree
Total



A cumulative graph to show the number of Blackboard Log-ins
by Maths (2011/12-2013/14)
Blackboard Log-ins
Year 1
Year 2
Year 3
First class honours
Upper second class honours
Lower second class honours
Third class honours/Pass
Grand Total



A cumulative graph to show the number of BlackBoard logins
by Psychology students (2012/13-2014/15)
Blackboard Log-ins
Year 1
Year 2
Year 3
First class honours
Upper second class honours
Lower second class honours
Third class honours/Pass
Grand Total



A cumulative graph to show the numberof Blackboard Logins by
by Physics and Astronomy students (2012/13-2014/15)
Blackboard og-ins
Year 1
Year 2
Year 3
First class honours
Upper second class honours
Lower second class honours
Third class honours
Grand Total

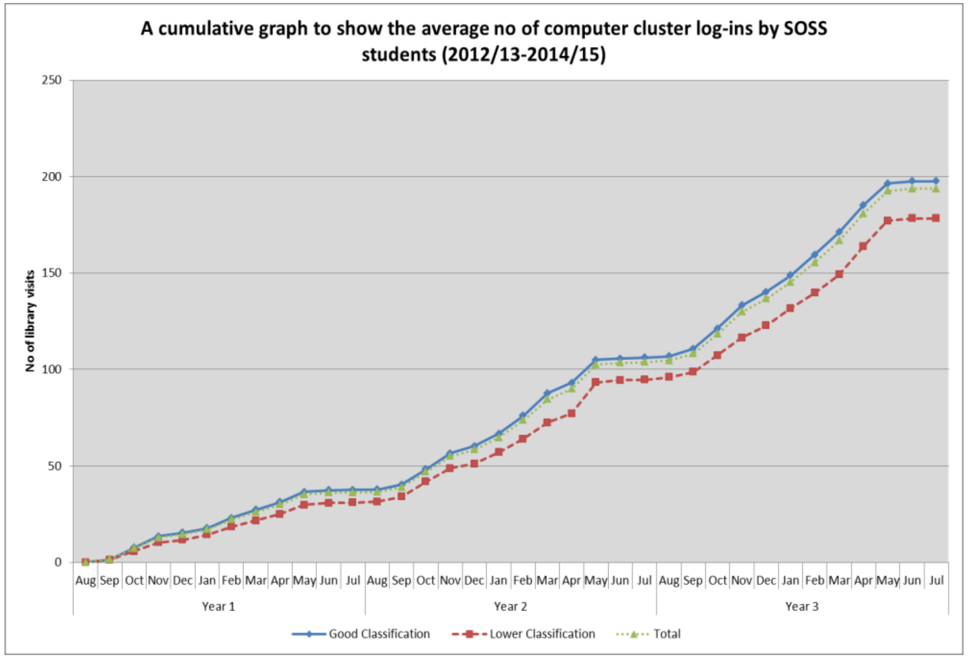
Computer Cluster Logins

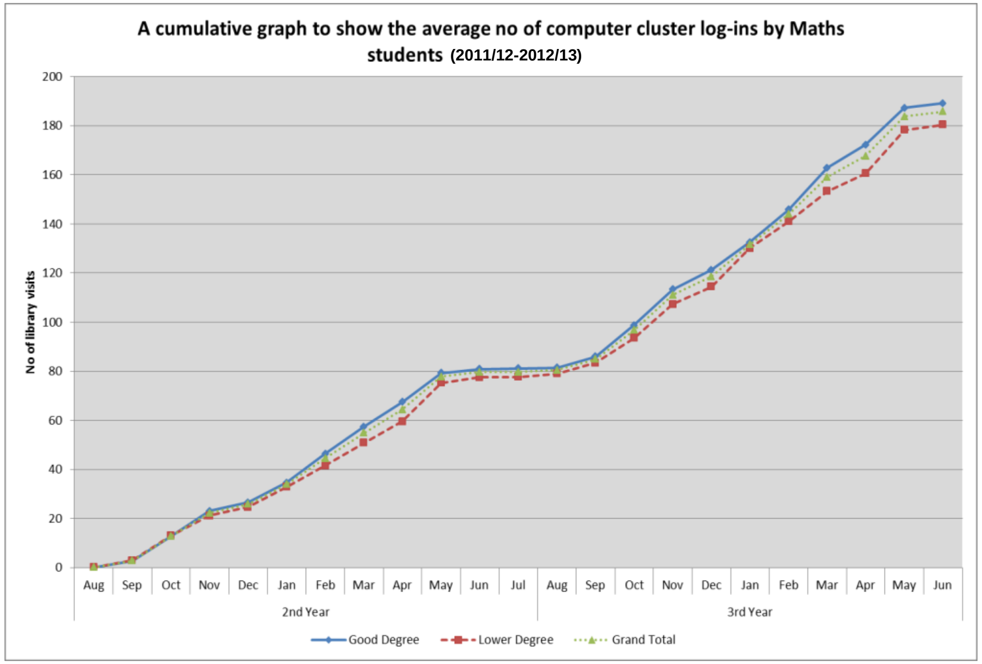
The analysis below details the relationship between computer cluster log-ins and attainment. It will include both good/lower degree and full classification analysis. The data only relates to log-ins and data was not available to evaluate the type of computer activity that students engaged in therefore it was not possible to assess whether students were engaged in academic or non-academic activities.

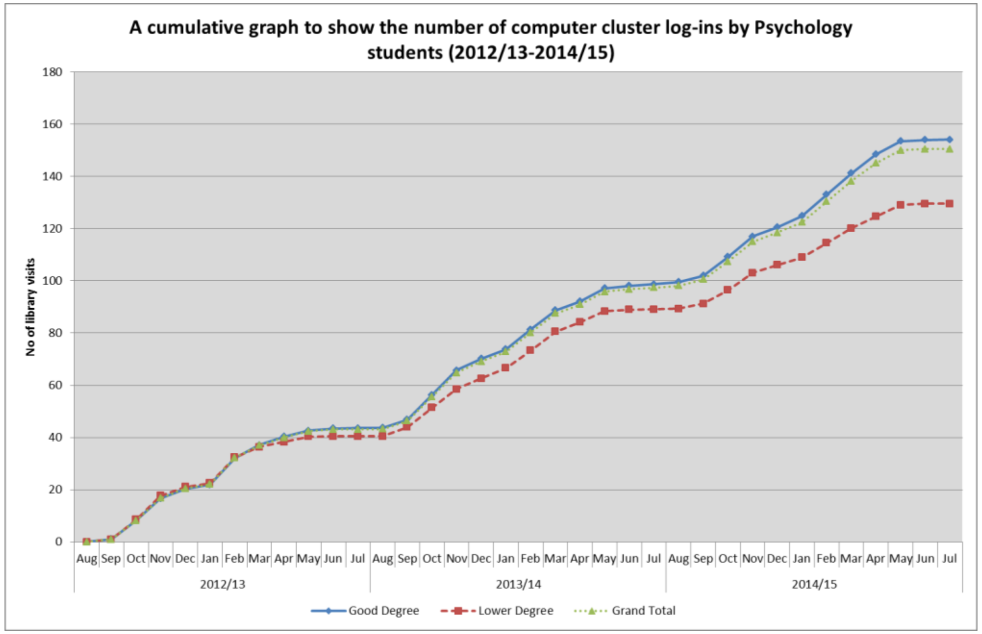
Good/Lower Classification

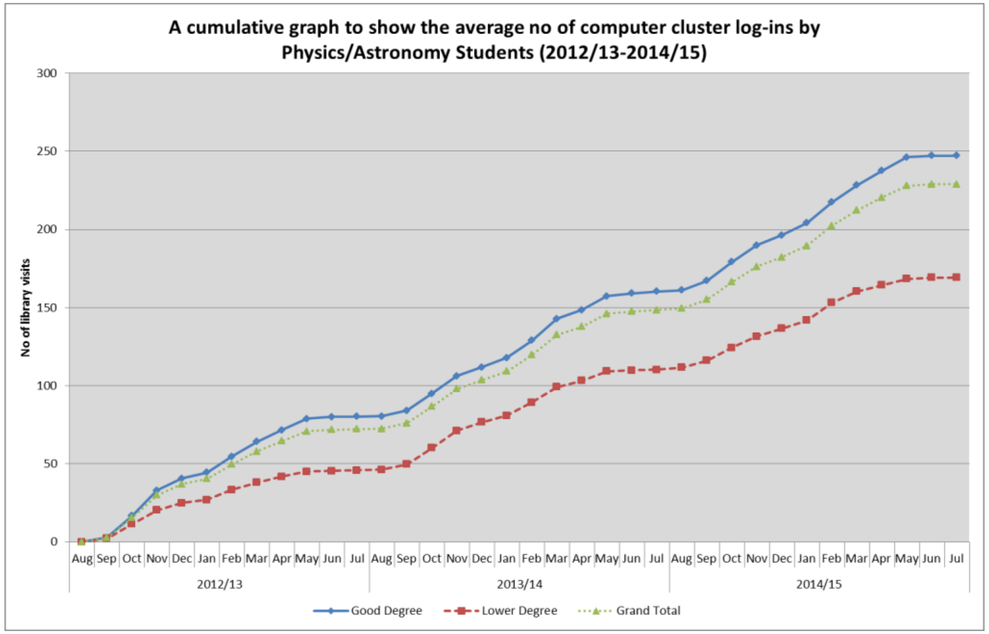
In terms of SOSS, the chart shows that students that obtained a lower degree used computer clusters less than the students that obtained a good degree. Students that achieved a good degree used the computer cluster approximately twenty more times than those who achieved a lower degree over the three year period. Both populations show similar trends month-on-month with fairly consistent usage across September to May and higher usage in year 3.

Within the School of Mathematics the levels of learner analytics engagement, as measured by the computer cluster log-ins, shows that students achieving a good degree had slightly higher usage levels than those gaining a lower degree but the gap was marginal and the difference emerged towards the end of the year 3. A similar trend was found within the School of Psychological Sciences in which good degree students logged into computer clusters 24 more times than their counterparts across the three year data period. Differences between good and lower Physics and Astronomy students was more pronounced with good degree students having much higher levels of computer cluster access than lower degree students. Differences between the two cohorts are apparent from the beginning of the data period although the attainment gap widens over time with larger differences emerged within the third year of study.


A cumulative graph to show the average no of computer cluster log-ins by SOSS students (2012/13-2014/15)
No of library visits
Year 1
Year 2
Year 3
Good Classification
Lower Classification
Total


A cumulative graph to show the average no of computer cluster log-ins by Maths students (2011/12-2012/13)
No of library visits
2nd Year
3rd Year
Good Degree
Lower Degree
Grand Total

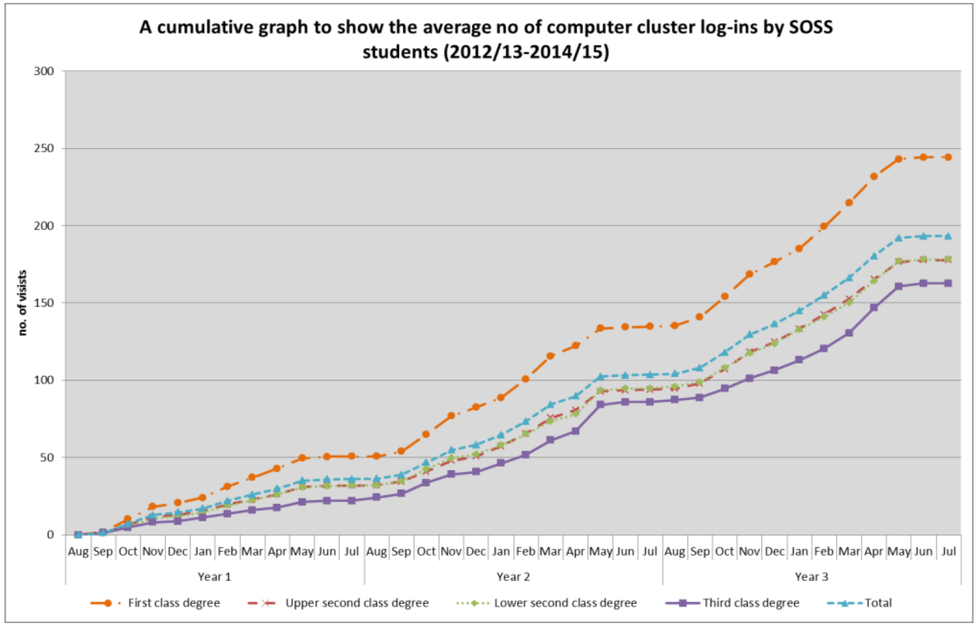

A cumulative graph to show the number of computer cluster log-ins by Psychology students (2012/13-2014/15)
No of library visits
2012/13
2013/14
2014/15
Good Degree
Lower Degree
Grand Total


A cumulative graph to show the average no of computer cluster log-ins by Physics/Astronomy Students (2012/13-2014/15)
No of library visits
2012/13
2013/14
2014/15
Good Degree
Lower Degree
Grand Total

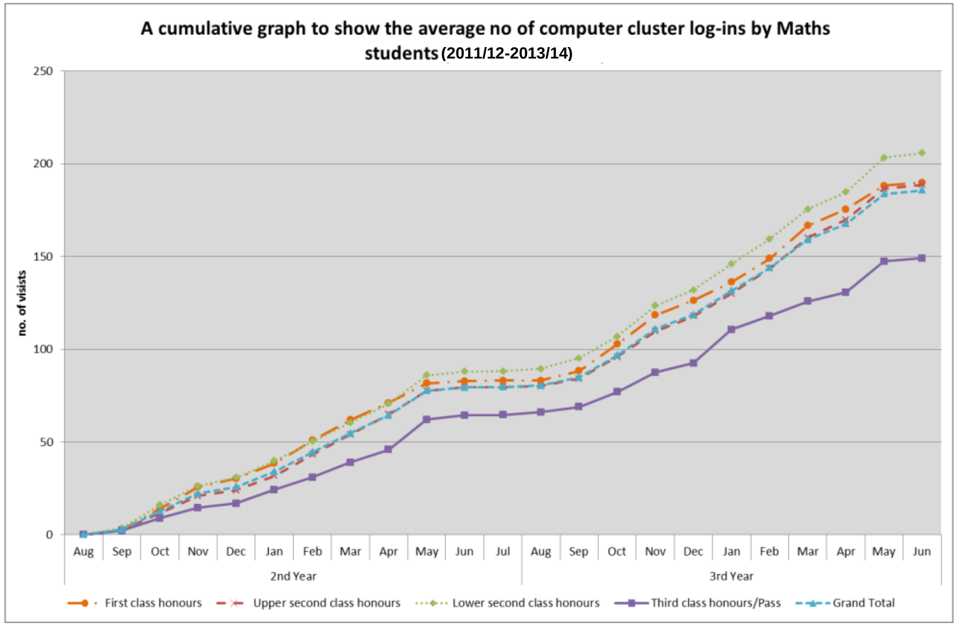
students categorised by the full degree classification system shows that the highest computer cluster log-in levels were amongst the students gaining a 2:2 with those achieving a 1st and 2:1 having very similar levels. The data clearly show that students gaining 3rd/pass had much lower levels of computer cluster log-ins compared to their counterparts.

Examining computer cluster data within Psychology (see chart below), it is clear that those who achieved a 3rd/pass degree had a much lower level of usage than the other degree classification cohorts however this cohort only included two students. Patterns of computer cluster usage within 1st class students and those who achieved a 2:1 were similar throughout the three years of study, cumulating in n=156 and n=154 average total uses respectively. Those who achieved a 2:2 also followed a similar pattern throughout the majority of years 1 and 2 although differentiations in usage occurred particularly within the third year of study, resulting in a cumulative usage of 128 total average computer log-ins per student.

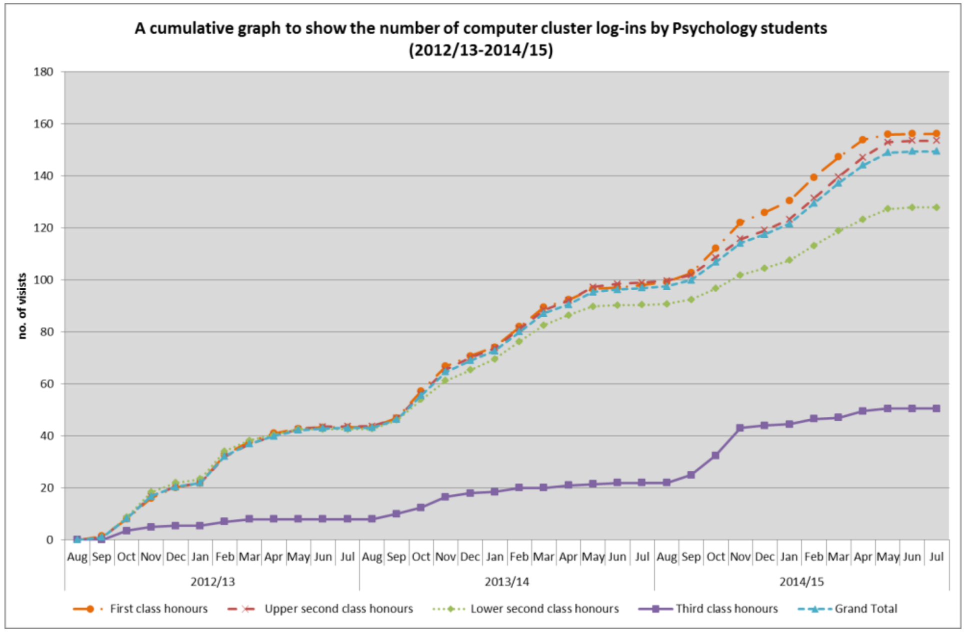
Those who achieved a 2:1 within the School of Physics and Astronomy used the computer clusters slightly more than those who achieved a 1st degree although the difference here is small (n=12). There was however a clear gap between the 1st and 2:1 student cohorts and those that achieved 2:2 and 3rd/pass degrees.



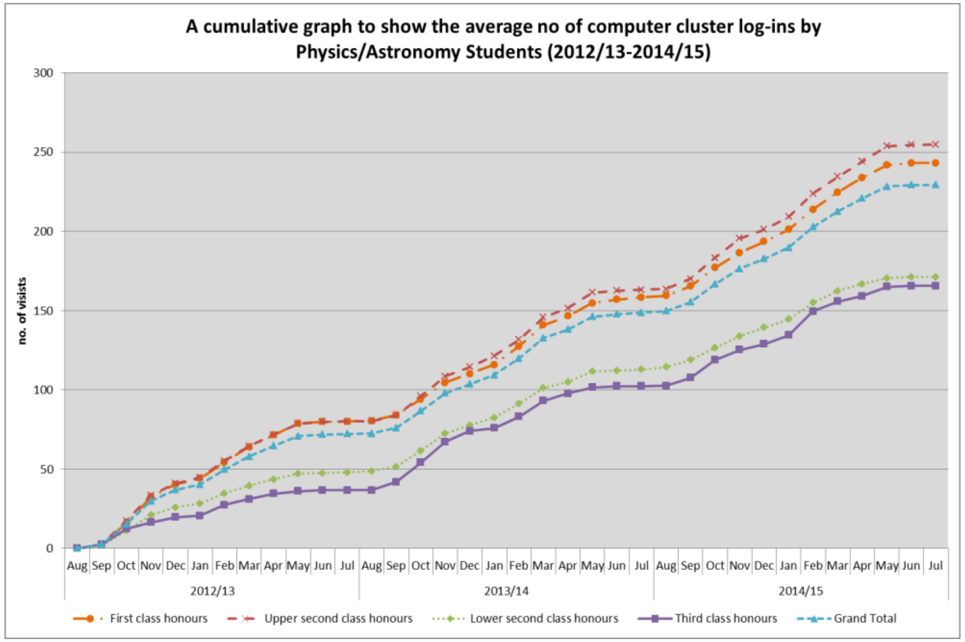
A cumulative graph to show the average no of computer cluster log-ins by SOSS students (2012/13-2014/15)
no. of visists
Year 1
Year 2
Year 3
First class degree
Upper second class degree
Lower second class degree
Third class degree
Total



A cumulative graph to show the average no of computer cluster log-ins by Maths students (2011/12-2013/14)
no. of visists
2nd Year
3rd Year
First class honours
Upper second class honours
Lower second class honours
Third class honours/Pass
Grand Total



A cumulative graph to show the number of computer cluster log-ins by Psychology students (2012/13-2014/15)
no. of visists
2012/13
2013/14
2014/15
First class honours
Upper second class honours
Lower second class honours
Third class honours
Grand Total



A cumulative graph to show the average no of computer cluster log-ins by Physics/Astronomy Students (2012/13-2014/15)
no. of visists
2012/13
2013/14
2014/15
First class honours
Upper second class honours
Lower second class honours
Third class honours
Grand Total

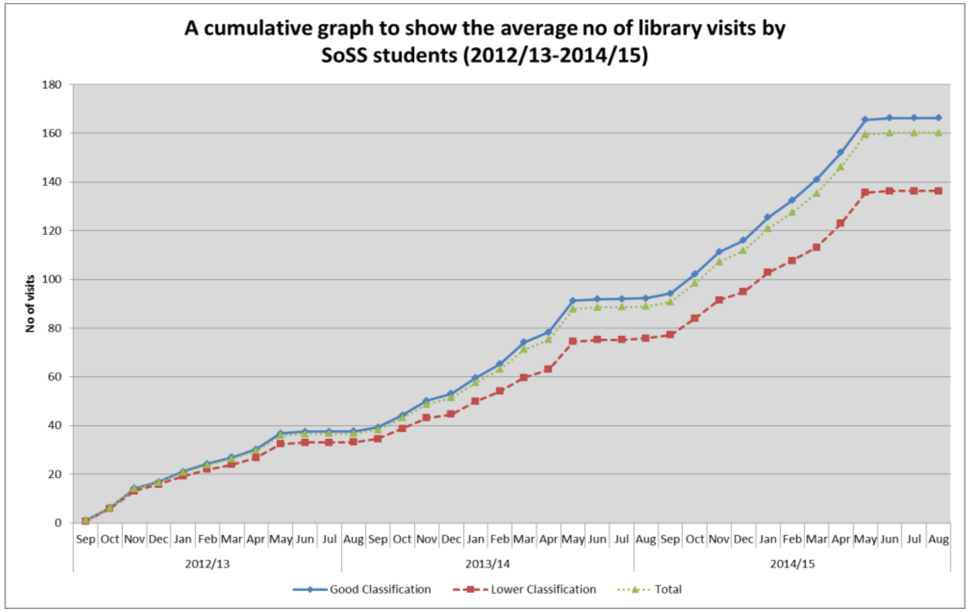
Library Access

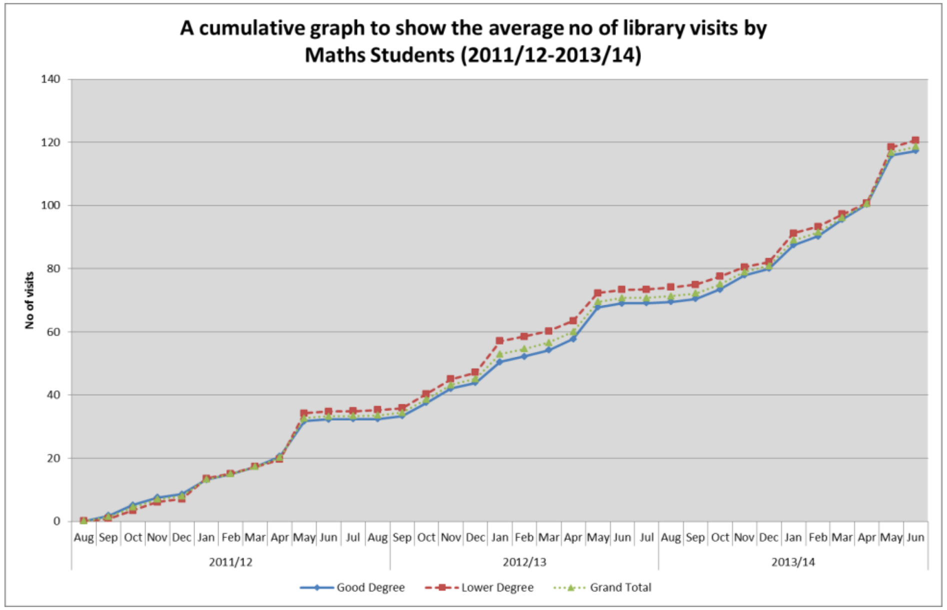
This section will detail the relationship between quantity of library access and degree attainment. It will firstly compare those with a good degree classification to those with a lower degree classification and then compare the four degree classifications cohorts.

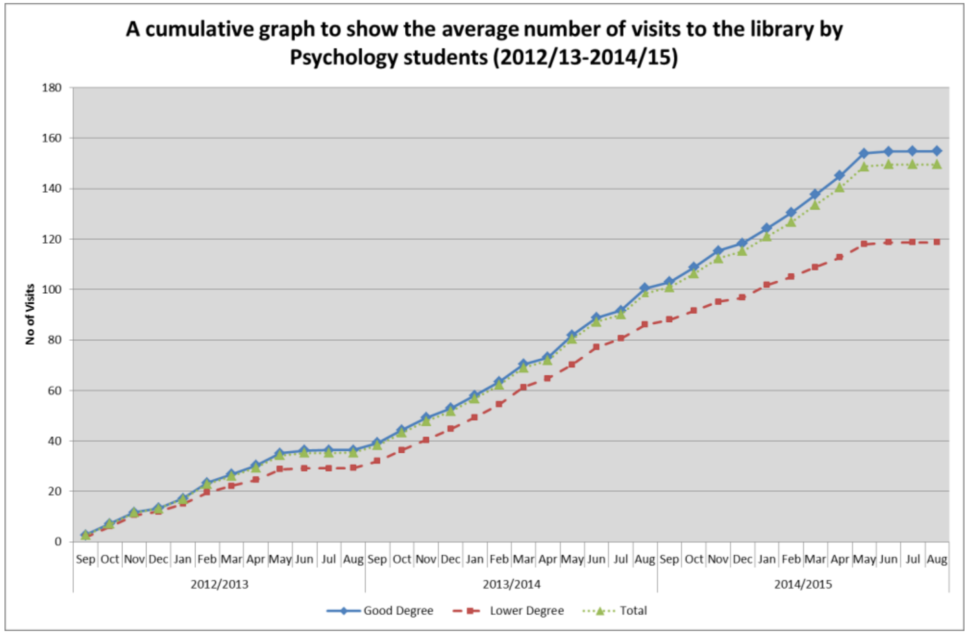
Good/Lower Classification

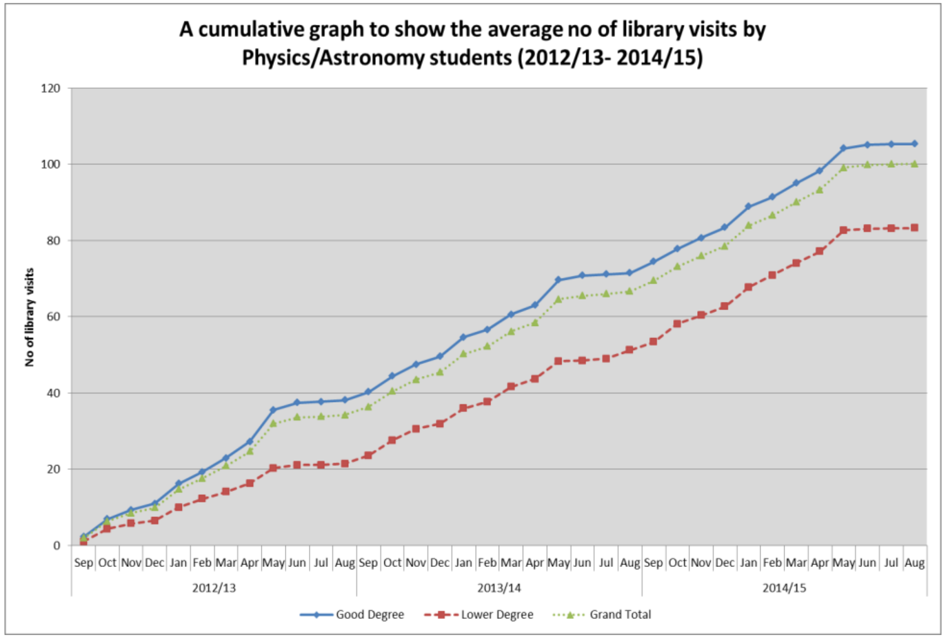
In terms of differences between those who achieved a good and lower degree classification the charts show that good degree students had a higher usage of the library throughout the three year period with the difference between the two populations increasing each year within Physic & Astronomy, SOSS and Psychology. In Physics and Astronomy the gap in access levels emerged within year 1 and was maintained across years 2 and 3. In The School of Psychology the gap between the two attaiment groups was evident in years 1 and 2 but grew considerable in year 3. Over the three year data period, on average a SoSS good degree student had 30 more uses of the library than the lower degree students with the gap in access widening in years 2 and 3.

Students from the School of Maths had much smaller differences in library access between the good and lower degree cohorts but overall the lower degree students very marginally having higher access levels. Stduents from the School of Psychology and SoSS had a higher levels of library usage compared to Maths and Physics & Astronomy. The data indicates that in general levels of library usage are linked to better attainmnet levels but there are distinct patterns across Academic Schools with the School of Math having either library access for lower rather than good degree students.


A cumulative graph to show the average no of library visits by SoSS students (2012/13-2014/15)
No of visits
2012/13
2013/14
2014/15
Good Classification
Lower Classification
Total


A cumulative graph to show the average no of library visits by Maths Students (2011/12-2013/14)
No of visits
2011/12
2012/13
2013/14
Good Degree
Lower Degree
Grand Total


A cumulative graph to show the average number of visits to the library by Psychology students (2012/13-2014/15)
No of Visits
2012/2013
2013/2014
2014/2015
Good Degree
Lower Degree
Total

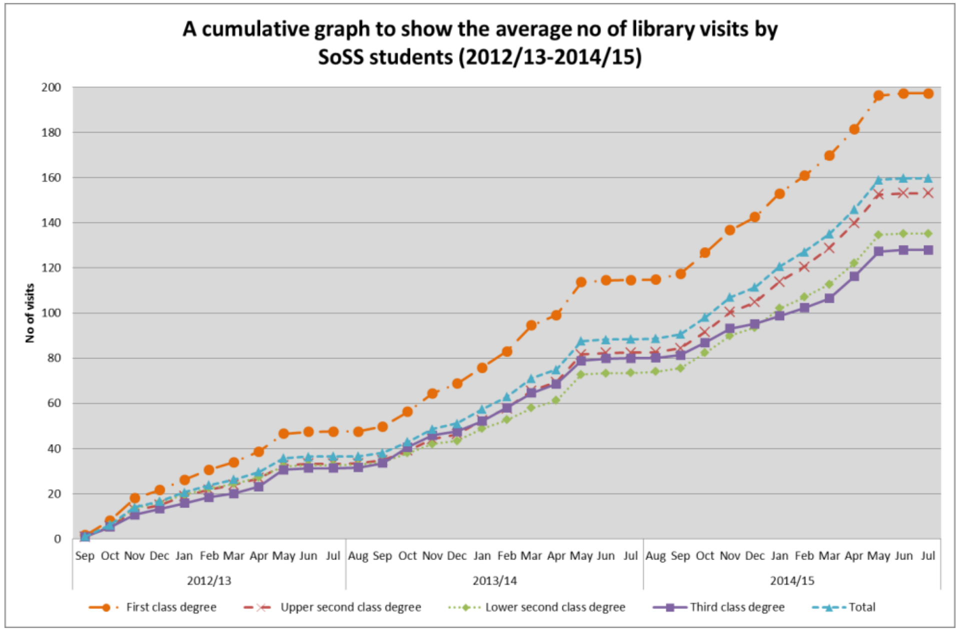

A cumulative graph to show the average no of library visits by Physics/Astronomy students (2012/13- 2014/15)
No of library visits
2012/13
2013/14
2014/15
Good Degree
Lower Degree
Grand Total

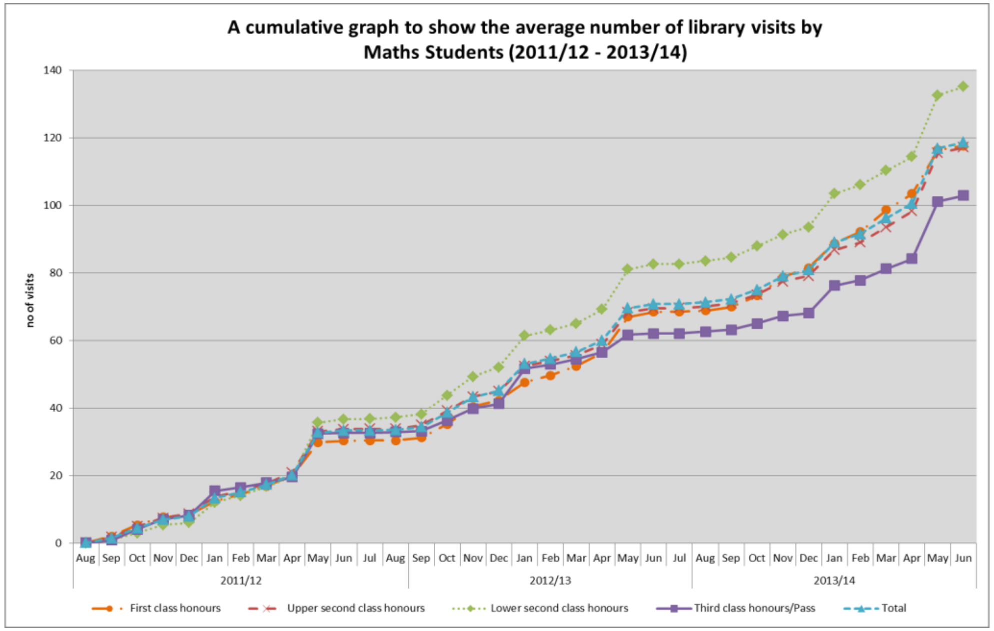
Full Classification

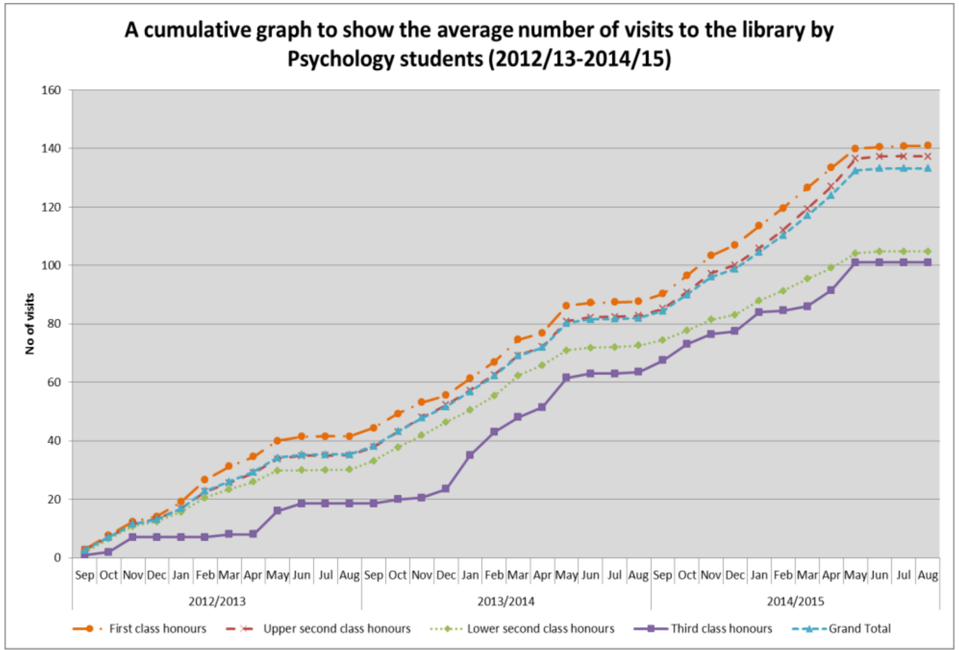
The charts below display library access data split into the four degree classifications cohorts. The SOSS data visualisation below suggests that 1st degree students had the high levels of library access throughout the three year period compared to students gaining lower degree. In SOSS library access levels were similar between the other three degree classifications within their first year of study and then towards the end of the second year 2:1 students increased their usage compared to 2:2 and 3rd/pass students. It is important to note at this point that the 3rd/pass cohort is disproportionately smaller (n=17) than the other classification groupings and therefore data may not be reliable. However, it is clear that there are large differences between those achieving a first class degree and the other degree classifications. Overall, students that achieved a 1st degree on average had 89 more library accesses than those who achieved a 3rd/pass.

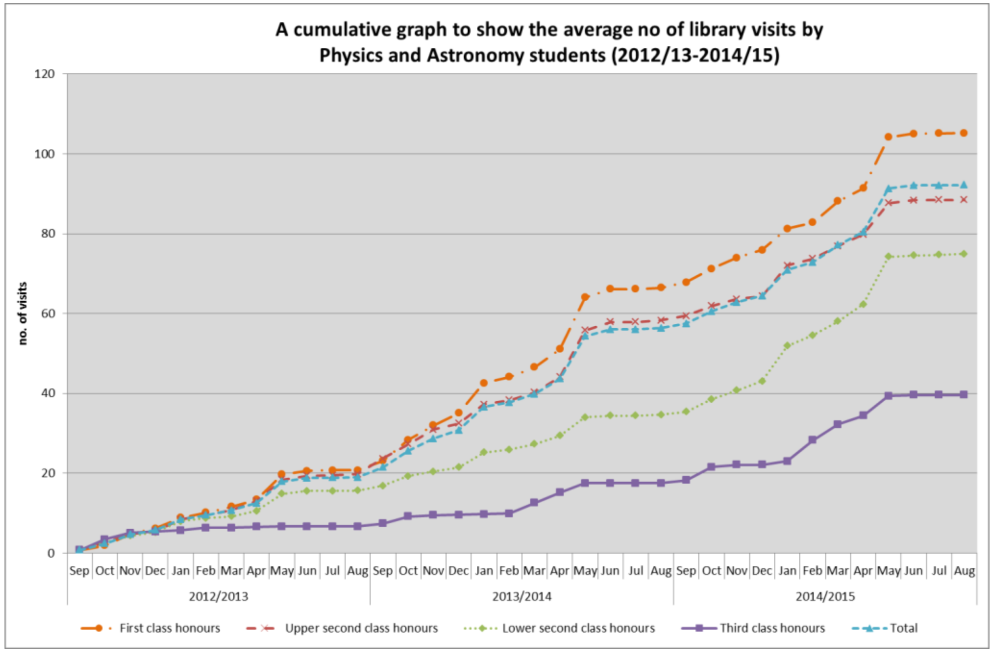
The School of Maths chart below shows that despite cohesive trends within the first year of study, those with a 2:2 degree actually had the highest average number of library visits overall (n=135) with both 1st and 2:1 students having n=118 and n=117 visits respectively. Differences began to emerge mid-year in the second year of study. In terms of Psychology, it is important to note that the total count of students graduating with a third class honours degree was only n=2 and therefore findings in this category may be unreliable however this cohort did appear to consistently use the library facility less than any other cohort with the 2:2 students not too dissimilar in terms of overall number of visits. Those who achieved a 1st class degree and those who achieved a 2:1 had fairly similar trends and overall number of visits.

The Physics and Astronomy cohort differed from the other Academic Schools as those who achieved an 2:1 had the highest library usage across the data period and their average total uses of the library was around 25 visits more than those who achieved a 1st class degree. Furthermore, the 3rd/pass students had a slightly higher usage than those who achieved a 2:2 although the difference is smaller.


A cumulative graph to show the average no of library visits by
SoSS students (2012/13-2014/15)
No of visits
2012/13
2013/14
2014/15
First class degree
Upper second class degree
Lower second class degree
Third class degree
Total


A cumulative graph to show the average number of library visits by
Maths Students (2011/12 - 2013/14)
no of visits
2011/12
2012/13
2013/14
First class honours
Upper second class honours
Lower second class honours
Third class honours/Pass
Total


A cumulative graph to show the average number of visits to the library by
Psychology students (2012/13-2014/15)
No of visits
2012/2013
2013/2014
2014/2015
First class honours
Upper second class honours
Lower second class honours
Third class honours
Grand Total


A cumulative graph to show the average no of library visits by
Physics and Astronomy students (2012/13-2014/15)
no. of visits
2012/2013
2013/2014
2014/2015
First class honours
Upper second class honours
Lower second class honours
Third class honours
Total

Time Spent in Library

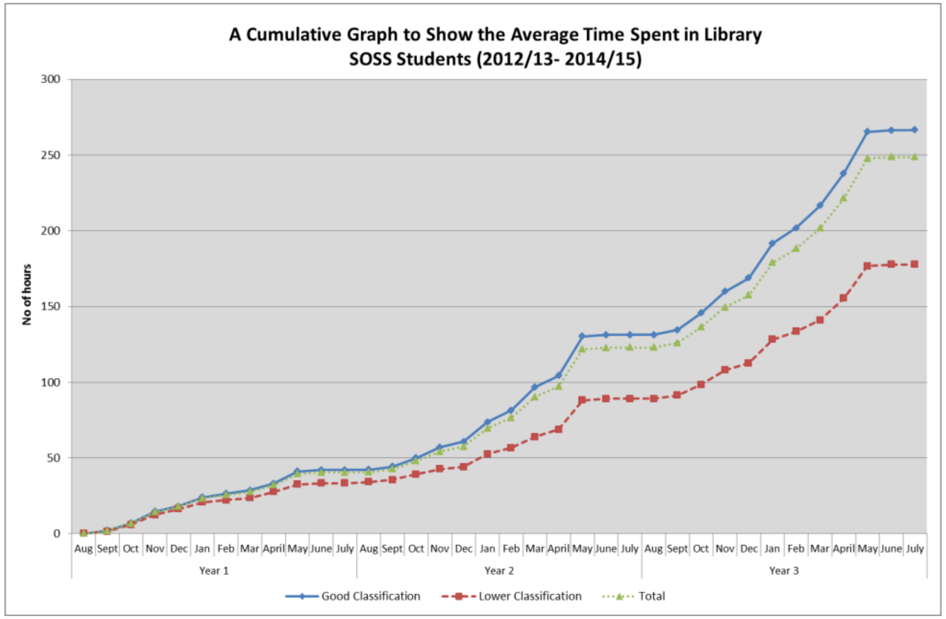
The analysis below details the average number of hours students spend in the library over a three year period categorised by using both the good/lower degree classification and full degree classification split. Time spent in the library is capture by students swiping in and out of University libraries and not all the libraries have swipe access control which will impact on reliability of the data particularly related to certain subject areas (e.g. Martin Harris Library does not have swipe access control). The data is dependent on student swiping in and out of the library and further investigation is required to establish the validity of the data and if there are occasions where students may not swipe out of the library (e.g. closing times) which could cause data quality issues.

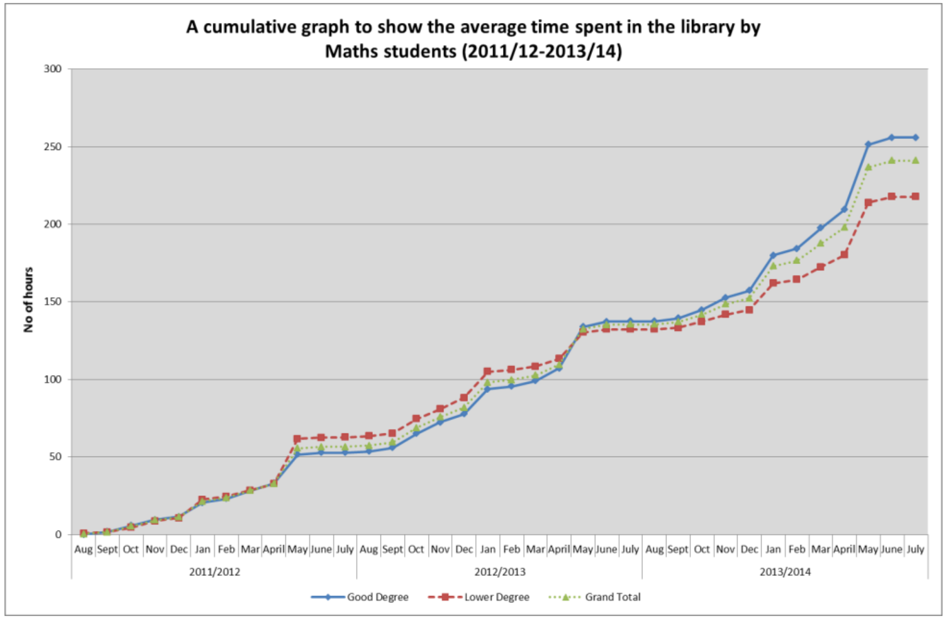
Good/Lower Classification

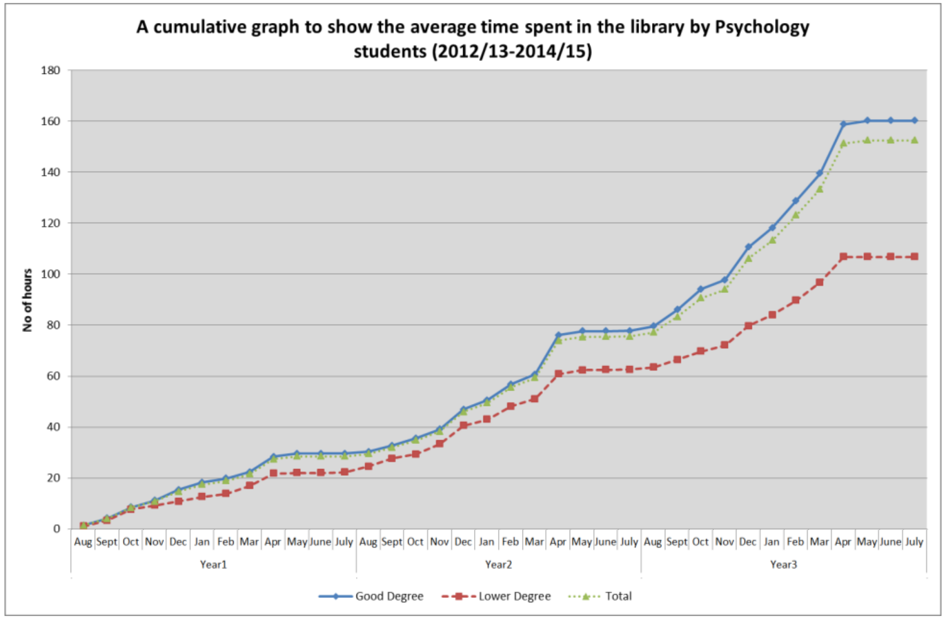
In SOSS chart below there is an obvious difference between good and lower degree students which begins to widen from around December in the second year of study with good degree students spending much more time in the library. Both degree populations follow similar trends in usage, including particularly high usage in the months of April and May.

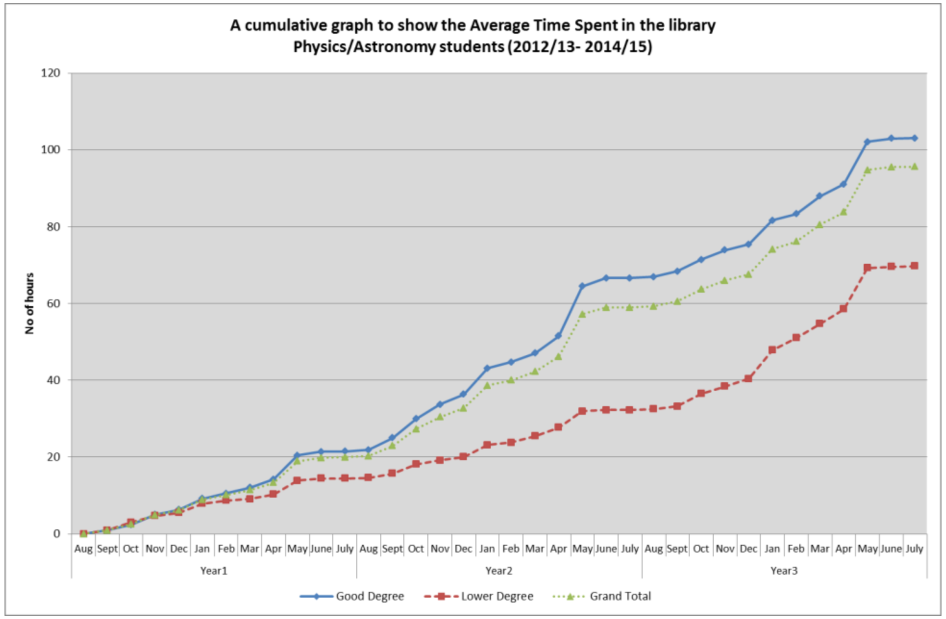
Compared to SOSS students from the School of Math had a much smaller library time gap between good and lower degree students. The cumulative level of library time for good degree students is higher than lower degree students from the second half of the third year of study but before this point levels of library time were very similar. In the School of Psychology, there was a large difference in library time between good and lower degree students that emerged in May of the second year of study. Overall, good degree students accumulating 53 more hours within the library than their counterparts.

The Physics and Astronomy data trends follow the patterns evident in the SOSS and Psychology charts with good degree students having much higher levels of library time. The gap in library time between good and lower degree student initially emerged in May in year 1 and widened considerable across year 2 and was maintained at a consistent level across year 3. The good degree students had period of intensive library usage (e.g. May year 2 and Year 3) whilst these peaks were not evident amongst the lower degree students.


A Cumulative Graph to Show the Average Time Spent in Library
SOSS Students (2012/13- 2014/15)
No of hours
Year 1
Year 2
Year 3
Good Classification
Lower Classification
Total


A cumulative graph to show the average time spent in the library by
Maths students (2011/12-2013/14)
No of hours
2011/2012
2012/2013
2013/2014
Good Degree
Lower Degree
Grand Total


A cumulative graph to show the average time spent in the library by Psychology
students (2012/13-2014/15)
No of hours
Year1
Year2
Year3
Good Degree
Lower Degree
Total


A cumulative graph to show the Average Time Spent in the library
Physics/Astronomy students (2012/13- 2014/15)
No of hours
Year1
Year2
Year3
Good Degree
Lower Degree
Grand Total

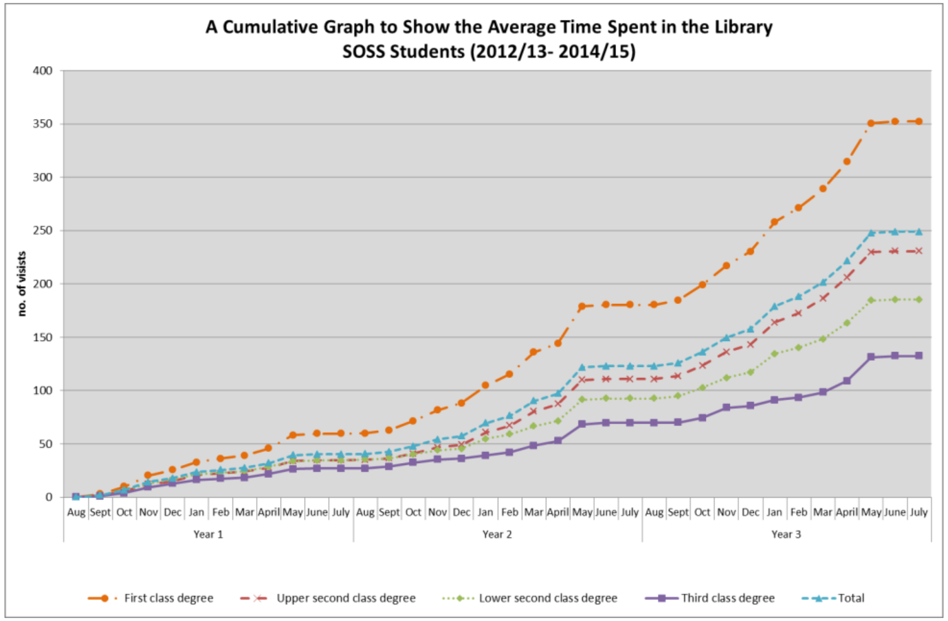
Full Classification

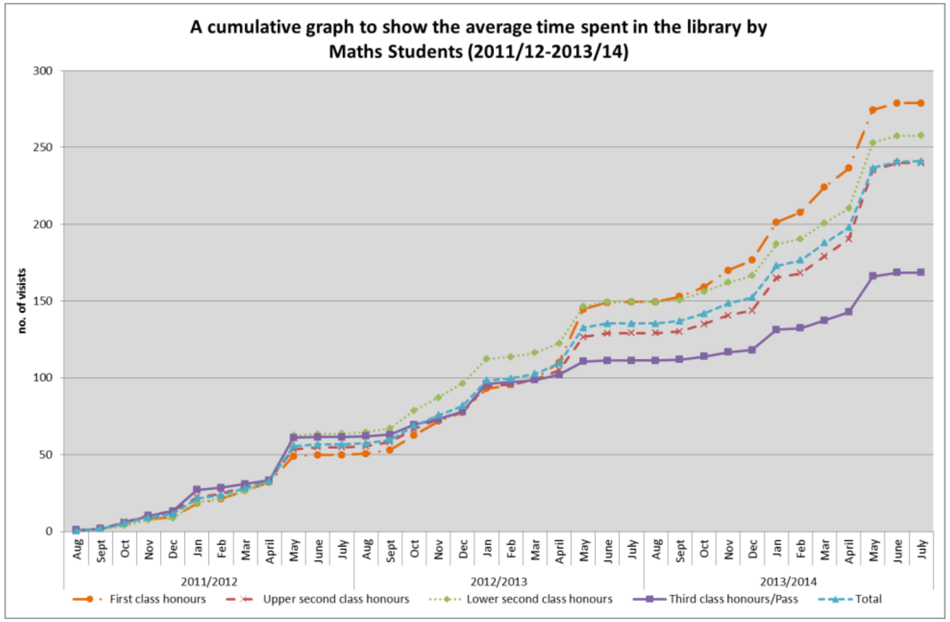
Examining library time access split by the full degree classification groups within SOSS (see chart below) shows that students that gained a 1st class degree had a much higher average number of hours in the library in comparison to all other classifications with an average of 552 hours as opposed to students who achieved a 3rd/pass that had the lowest total hours of 132. 1st class degree level students had a much higher usage from November of the first year of study, although particularly strong differences developed within the exam period of May in year 1. 2:1 students completed their degree with an average of 231 hours as opposed to 185 hours completed by 2:2 students.

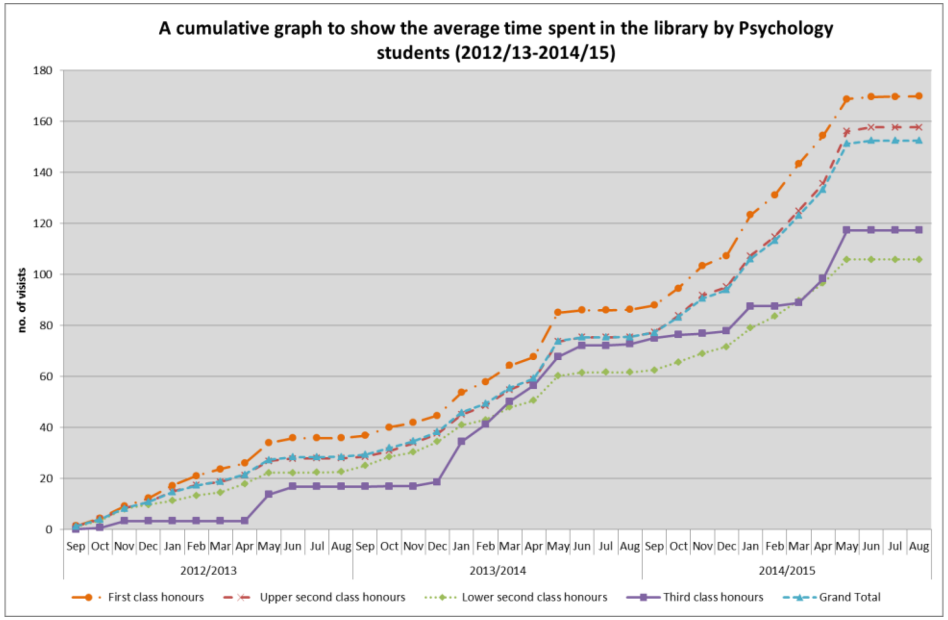
Within the School of Mathematics the students that obtained a 1st class degree spent the most time in the library followed by 2:2 and then 2:1 students with the 3rd/pass students having the lowest levels of library time. The classification groups had very similar levels of library time until April in the second year of study when those gaining a 3rd/pass started to spent less time in the library than their counterparts.

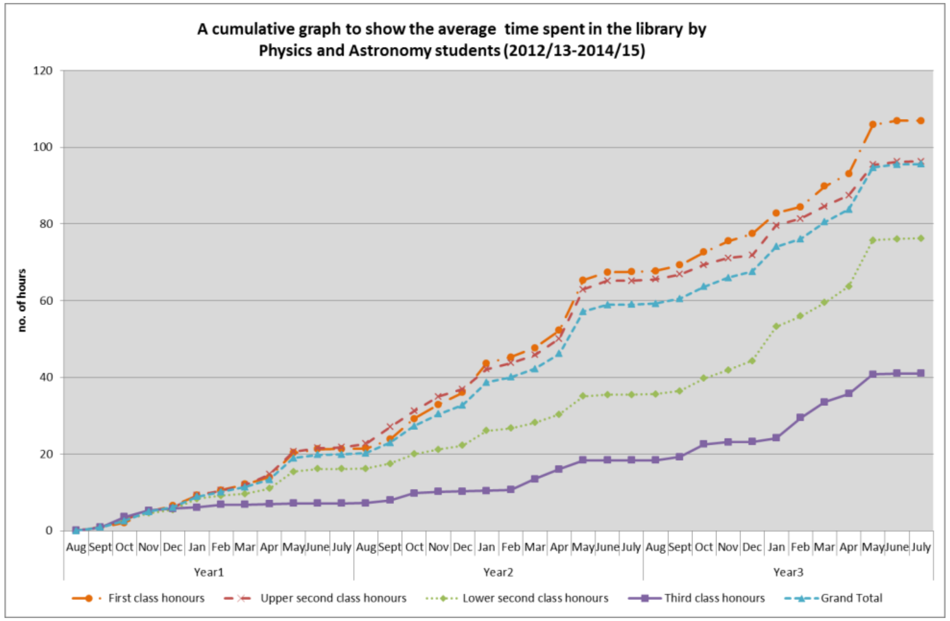
In the School of Psychology there is a very clear pattern of 1st and 2:1 students having much higher levels of library time compared to those achieving a 2:2 or 3rd/pass. Students that gained a 2:2 had the lowest levels of time in the library and the 3rd/Pass group had fluctuating levels of average monthly access which may be caused by the small sample size in this degree classification group.

The Physics and Astronomy chart below shows that the level of library time across 1st and 2:1 graduates is very similar with students that achieved a 1st having slightly more library time over the 3 years of study particularly in the last few months (March to May year 3). The 3rd/pass students had very low levels of library time but this is a small sample (n=10) and the data should be treated with caution of the 3rd year.


A Cumulative Graph to Show the Average Time Spent in the Library
SOSS Students (2012/13- 2014/15)
no. of visists
Year 1
Year 2
Year 3
First class degree
Upper second class degree
Lower second class degree
Third class degree
Total


A cumulative graph to show the average time spent in the library by
Maths Students (2011/12-2013/14)
no. of visits
2011/2012
2012/2013
2013/2014
First class honours
Upper second class honours
Lower second class honours
Third class honours/Pass
Total


A cumulative graph to show the average time spent in the library by Psychology
students (2012/13-2014/15)
no. of visists
2012/2013
2013/2014
2014/2015
First class honours
Upper second class honours
Lower second class honours
Third class honours
Grand Total


A cumulative graph to show the average time spent in the library by
Physics and Astronomy students (2012/13-2014/15)
no. of hours
Year1
Year2
Year3
First class honours
Upper second class honours
Lower second class honours
Third class honours
Grand Total

Library Loan Data

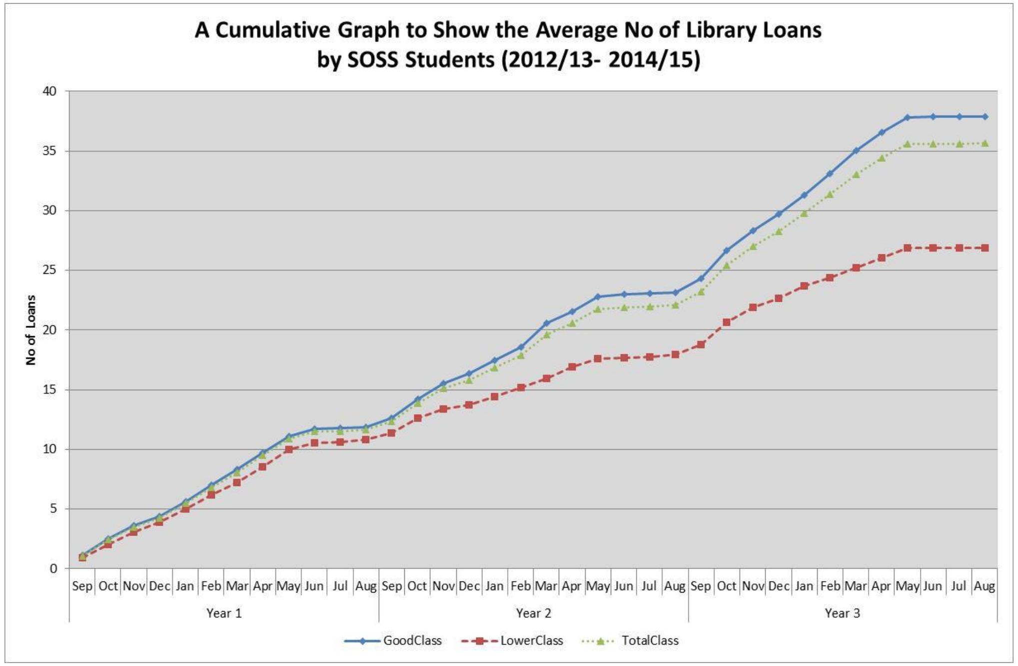
The charts below show the average cumulative count of physical library loans for students throughout their three years of study categorised by attainment groups. The data is based on physical loans therefore there is a data gap in relation to usage of electronic library database downloads which at the time of writing was not available. Electronic library download data forms a substantial part of many students access to academic material and usage of electronic library resources will vary across Academic Schools reflecting the different demands of teaching and learning.

Library loan data is not easily accessible and it is stored in two locations. The most recent Library loan data is stored in the Library Management System's built-in reporting tool, OBIEE (Oracle business Intelligence) and data is stored from April 2013 onwards in this system. Any older data related to physical loans is stored in the data warehouse in a static database. This created problems for collating data for the research presented below and it is not known whether there is any data quality checks have been undertaken on the data. The number of loans for students from the School of Psychology was very low for physical loans and it is not known whether this is a result of students mainly using electronic databases or data quality issues. When the data for the School of Maths was combined across the two datasets the data produced contained very low numbers and did not follow the patterns observed in the other learner analytics charts. It was not possible to verify the quality of the data and it has been excluded from the analysis.

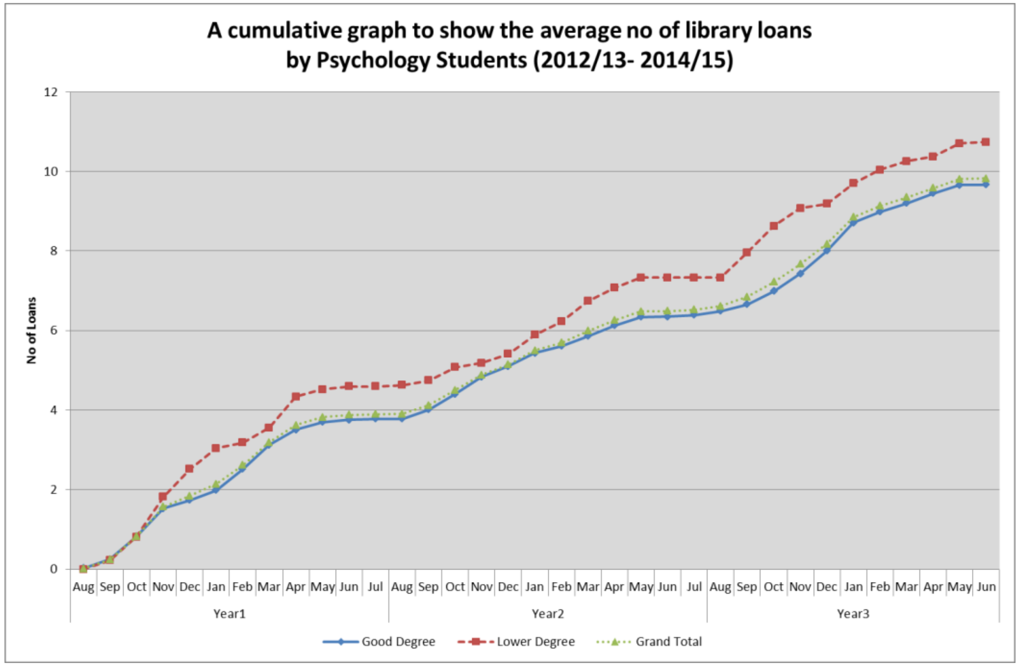
Good Classification

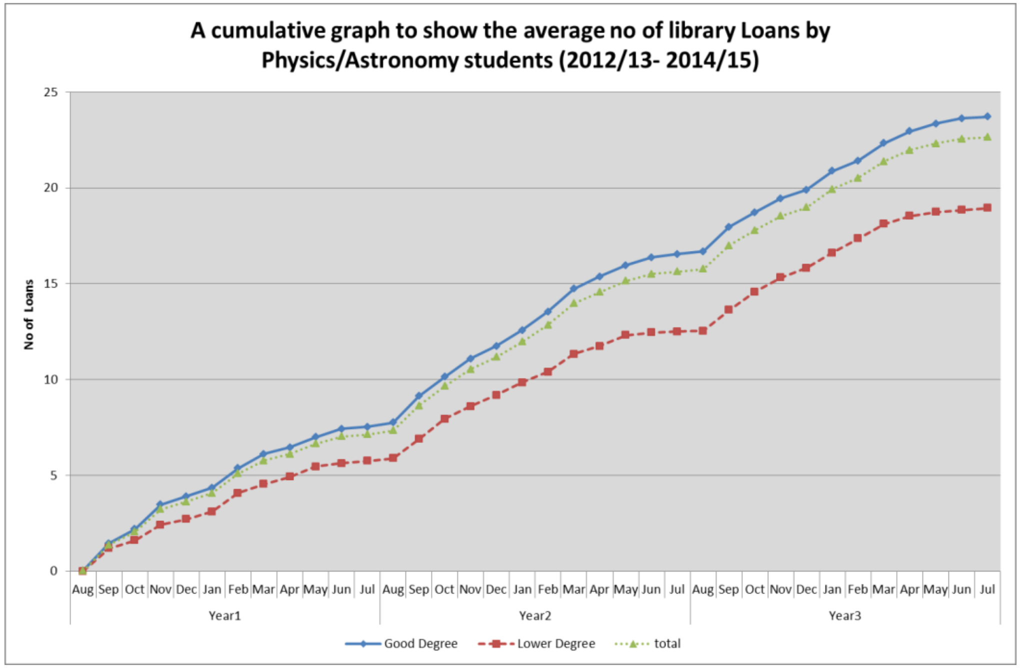
In SoSS there is little difference between the number of library loans by good and lower degree students within their first year of study. Differences started to emerge around December within the second year of study as good students increased their amount of loans at a greater rate than their counterparts. In total good students are likely to loan on average 12 more books than lower degree students with the loan gap reaching its peak within April of the third year of study.

In the School of Psychology the lower degree students had consistently higher levels of physical loans across the three year data period although the gap between good and lower degree students loans fluctuated. In the School of Physics and Astronomy good degree students consistently had higher levels of physical library loans across the data period and the gap between the two attainment groups increased across the 3 academic years.



School of Mathematics data not available due to data quality issues (see above)





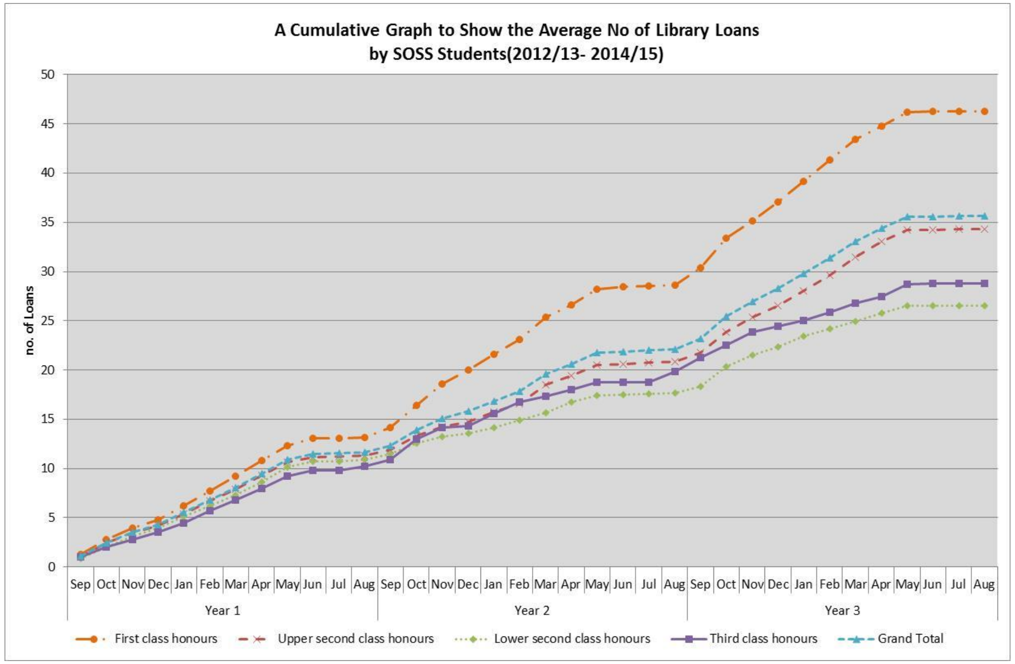
Full Classification

When examining the SOSS physical loan data categorised by full degree classification, it is clear that those with a 1st class degree have taken out a far greater number of library loans than their counterparts, particularly within the second and third years of study, with a total of 46 loans in comparison to the overall average of 36. Other classifications were more cohesive, with those achieving a 2:2 degree having the lowest count of average total library loans (n=27) and the 3rd/pass students had similar levels (n=29). Those who achieved a 2:1 degree were largely on par with the two lower classifications cohorts until around December in the third year of study and then the number of loans increased more rapidly.

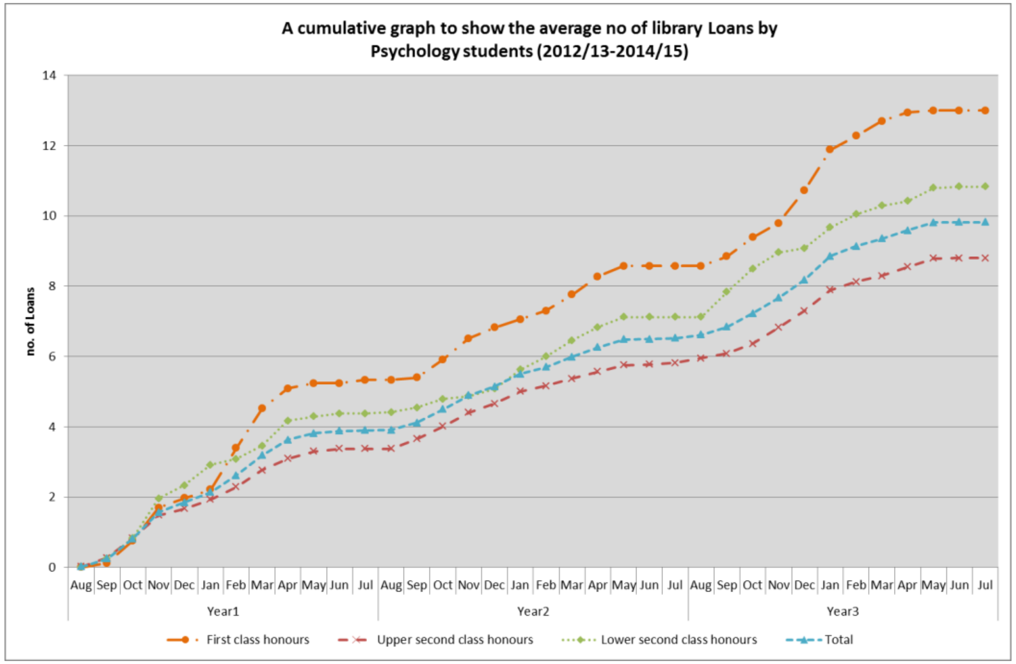
The School of Physics and Astronomy had very clear patterns of loans with those in the 1st and 2:1 classifications having much higher levels of library loans than the 2:2 and 3rd/pass cohorts. The gap in levels of library loans between the good and lower degree groups emerged within the 1st year of study and steadily increased over the data period. The students achieving a 2:1 had slightly higher loan levels than those that achieved a 1st.

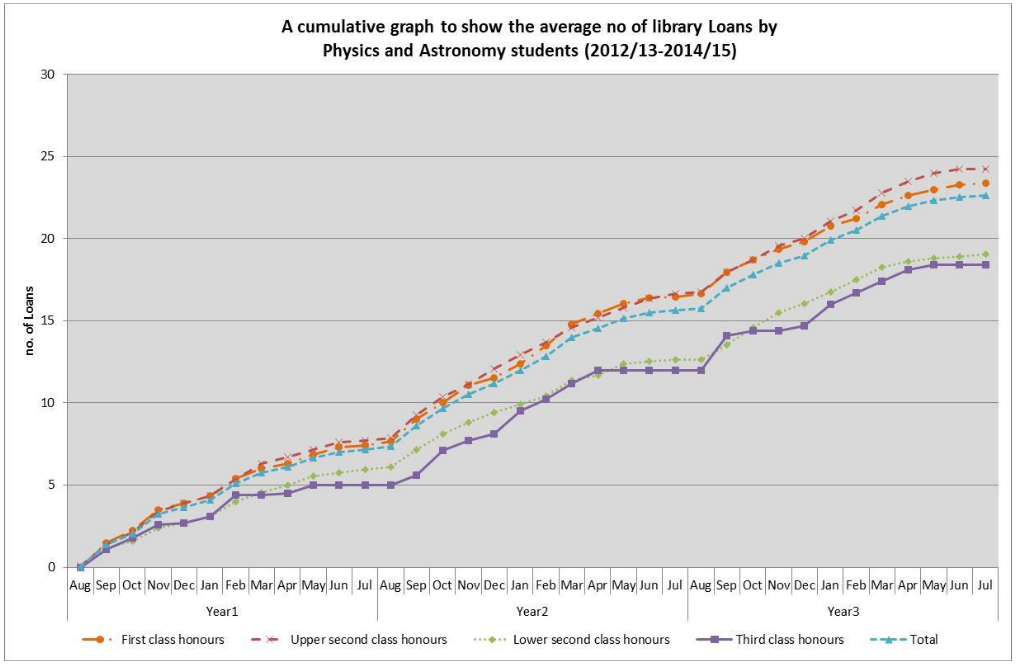
The library loan patterns within the School of Psychology suggest that students achieving a 1st had the highest levels of loans followed by 2:2 and then 2:1 students. 1st degree students had much higher rates of library loans in the 1st year and the gap between the loan levels increased over years 2 and 3. Data is not included for 3rd/pass students as the small sample produced data that was not reliable.

The descriptive analysis above shows that Blackboard and library engagement measurements have the potential to be used as learner analytics measurements to track academic attainment and there are distinct patterns across Academic Schools which needs to be taken into account when modelling and interpreting the learner analytics data. The regression analysis below takes the analysis above further by examining the relationship between attainment and the learner analytics taking into account other key variables that impact on attainment (e.g. entry qualifications).



School of Mathematics data not available due to data quality issues (see above)





Regression Modelling

Regression modelling is a form of analysis which can give a prediction of a dependent variable when controlling for numerous independent variables. The dependent variable used in the analysis below is dichotomous and in this case is the binary degree classification: good and lower degree. The regression models determine whether a variable is significant by establishing whether it makes an impact on attainment despite controlling for other variables.

Regression models were run to ascertain the relationship between each learner analytics measurements (n=5) across each of the Academic Schools (n=4) and across levels of engagement in each of the 3 academic years and the total engagement (n=4). This resulted in 80 models and these are summarised in the table below detailing whether the learning analytics had a significant impact on the probability of obtaining a good degree across the course of students' academic studies. It was not feasible to include all the models so examples of the model are provided below that relate to the totals for each learner analytics measurement in the data period for SOSS.

Variables

Below is a list of variables used within the regression model, with degree classification being the dependent variable. This report will highlight the variables with a significant effect on the dependent variable. It was not possible to include all the variables across the regression models as small cohorts in some variables (e.g. disabled students and mature students) caused issues related to the confidence levels within the models therefore relevant variables were removed from the models.

Variables:

- Degree Classification (good/lower degree)
- Age
- Gender
- Disability
- Domicile (UK, EU, Non-EU)
- Tariff

Effects Plots

Effects plots are diagrams produced from regression models to visualise the strength of the effect a variable has on the dependent variable, which is the probability of gaining good degree, and the relationships between categories in a variable. The vertical line through each point is the 95% confidence interval. If the interval is large it is not possible to confidently predict the outcome of the variable on the dependent variable.

Regression – Summary

The table below displays data showing whether the regression models detected significant relationships between the probability of gaining a good degree and individual student activity across the learner analytics measurements. The table splits the data each academic year and also includes the total in the three year data period. The main findings to emerge from the regression models were that levels of Blackboard access was the learner analytics measurements with the strongest relationship to good degree attainment followed by library time and library physical loans. The models suggest that levels of computer cluster access, library visits and Loan Gilbert access are poor learner analytics measurements.

Examining the Blackboard data in more detail reveals some interesting patterns particularly in the School of Physics and Astronomy were in the first year there was a significant ($p<0.05$) positive relationship between Blackboard access levels and the probability of gaining a good degree whilst in year two this relationship was a significant negative relationship. The School of Maths showed the strongest relationship between Blackboard engagement and attainment with Blackboard access levels positively associated with the probability of gaining a good degree across all academic years.

Table: Multiple regression table to show the levels of significance of learner analytics				
Leaner Analytic	Year 1	Year 2	Year 3	Total
Alan Gilbert LC - Maths	NS	NS	NS	NS
Alan Gilbert LC - Physic	NS	NS	NS	NS
Alan Gilbert LC - Psy	NS	NS	NS	NS
Alan Gilbert LC - SoSS	NS	NS	NS	NS
Blackboard				
Blackboard - Maths	***	**	**	***
Blackboard - Physic	*	^	NS	NS
Blackboard - Psy	*	*	NS	*
Blackboard - SoSS	**	NS	*	*
Computer Clusters				
Computer Clusters - Maths	NS	NS	NS	NS
Computer Clusters - Physic	NS	NS	NS	NS
Computer Clusters - Psy	NS	NS	NS	NS
Computer Clusters - SoSS	NS	NS	NS	NS
Library Access				
Library - Maths	NS	NS	NS	NS
Library - Physic	NS	NS	NS	NS
Library - Psy	NS	NS	NS	NS
Library - SoSS	NS	NS	NS	NS
Library Loans				
Library Loans - Maths	NA	NA	NA	NA
Library Loans - Physic	NS	NS	NS	NS
Library Loans - Psy	NS	NS	NS	NS
Library Loans - SoSS	*	NS	NS	NS
Library Loans				
Library Time - Maths	NS	NS	NS	NS
Library Time - Physic	NS	*	NS	NS
Library Time - Psy	NS	NS	NS	NS
Library Time - SoSS	NS	**	**	**

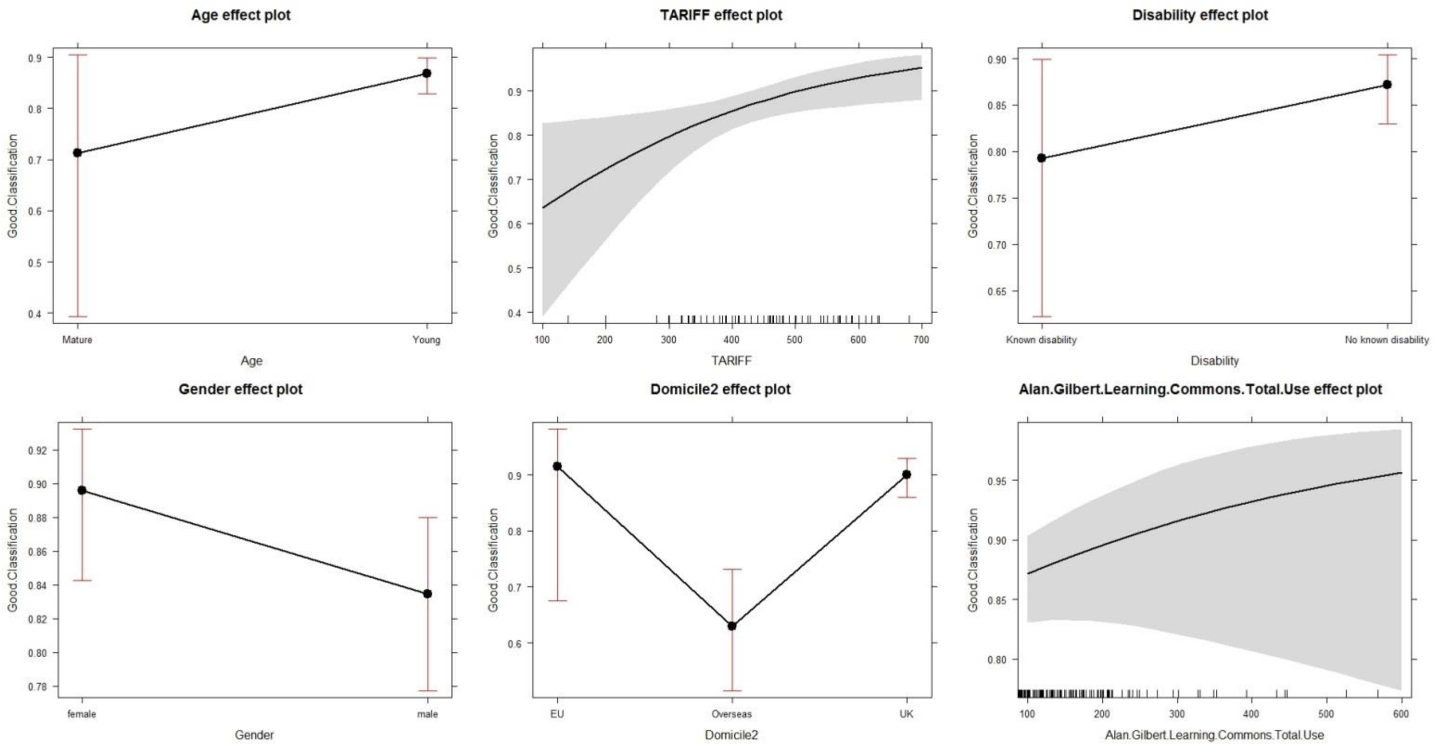
Key: NA – Not Available, NS – Not Significant, '***' $p<0.001$ – positive relationship, '**' $p<0.01$ – positive relationship, '*' $p<0.05$ – positive relationship, '^' $p<0.05$ - negative relationship.

The regression models below display the trends for each of the learner analytics measurements in SOSS. The data is based on the total measurements across the 3 year data period. The data from each of the models is available in appendix 2.

SoSS Regression Model - Effects Plots

Alan Gilbert Learning Commons

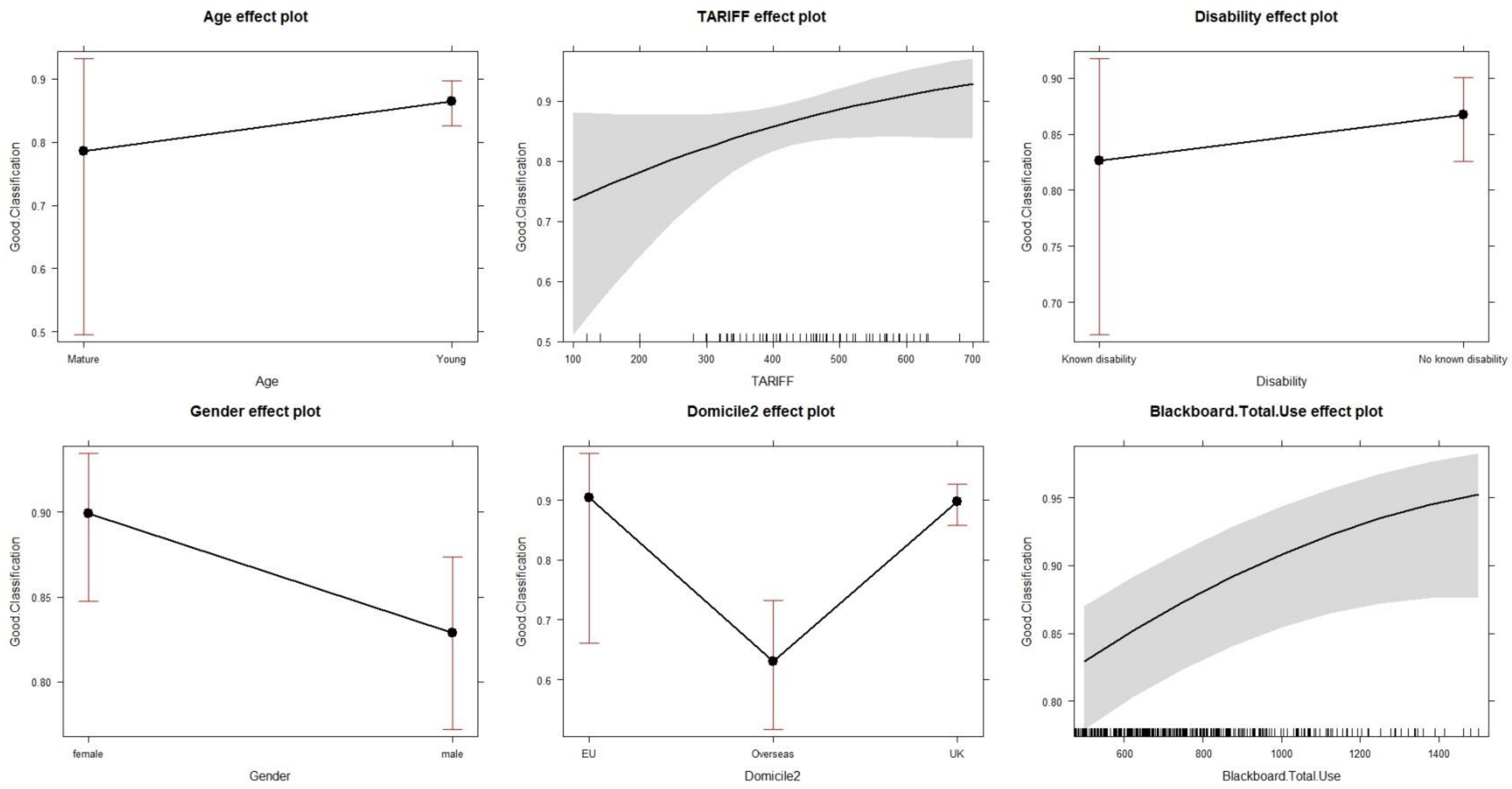
The regression analysis is based on the probability of gaining a good degree and includes as an independent variable the total number of AGLC accesses across the data period. There models suggest there is no significant relationship between AGLC usages and attainment. The only significant factors having an effect on attainment are domicile and tariff score. In the case of domicile, UK students have a significantly higher probability ($p<0.05$) of achieving a good degree as opposed to those from overseas; although there is not a significant difference with UK and EU students. A higher tariff score impacts on the probability of achieving a good degree ($p<0.05$). Regression analysis has also been employed separating the data by year of study (year 1, year 2 and year 3) and there were no significant relationship between attainment and amount of AGLC usage in any of the years.



Blackboard Logins

There was a significant relationship between Blackboard log-ins and the probability of obtaining a good degree ($p<0.05$). Other factors which were significance included gender ($p<0.05$) in which females were more likely to achieve a good degree than males; domicile ($p<0.05$) as UK students were more likely to achieve a good degree as opposed to overseas students. Those with a higher tariff score were not significantly more likely to achieve a good degree when all other variables were taken into account. Within programme years 1 and 3 Blackboard logins had a significant impact on attainment ($p<0.01$) but the relationship was not significance within year 2.

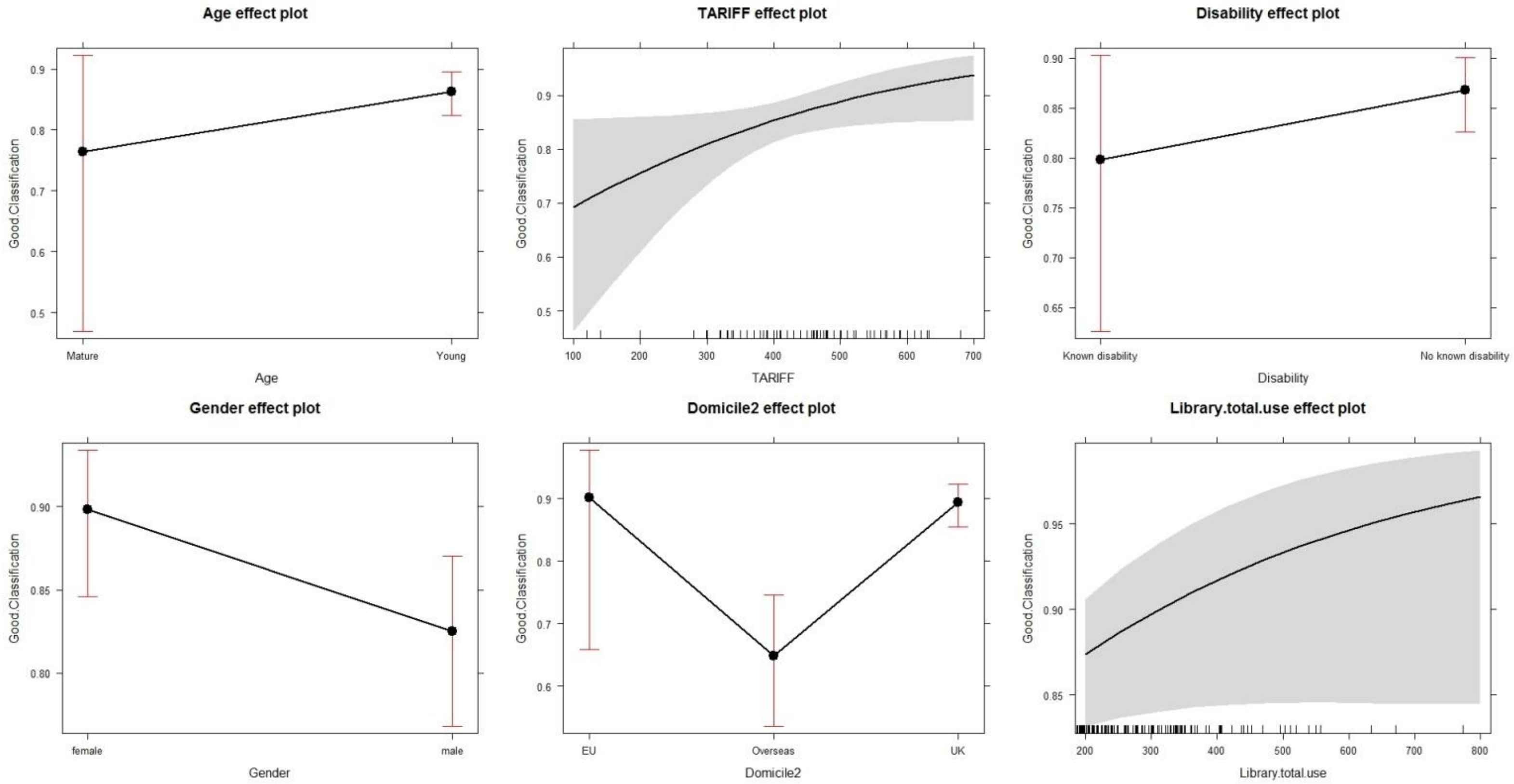
Figure: Effects Plots for SOSS Good Degree Attainment and Blackboard log-ins



Library Access

In terms of how many times students accessed the library, this was not a significant predictor of in relation to the probability of gaining a good degree. Overseas student were less likely to obtain a good degree than UK or EU but this was not a significant difference. The variables with a significant effect on attainment included tariff score ($p<0.05$) and gender (females outperforming males) ($p<0.05$). When separated into years of programme study, library access was not a significant predictor of attainment within years 1, 2 or 3.

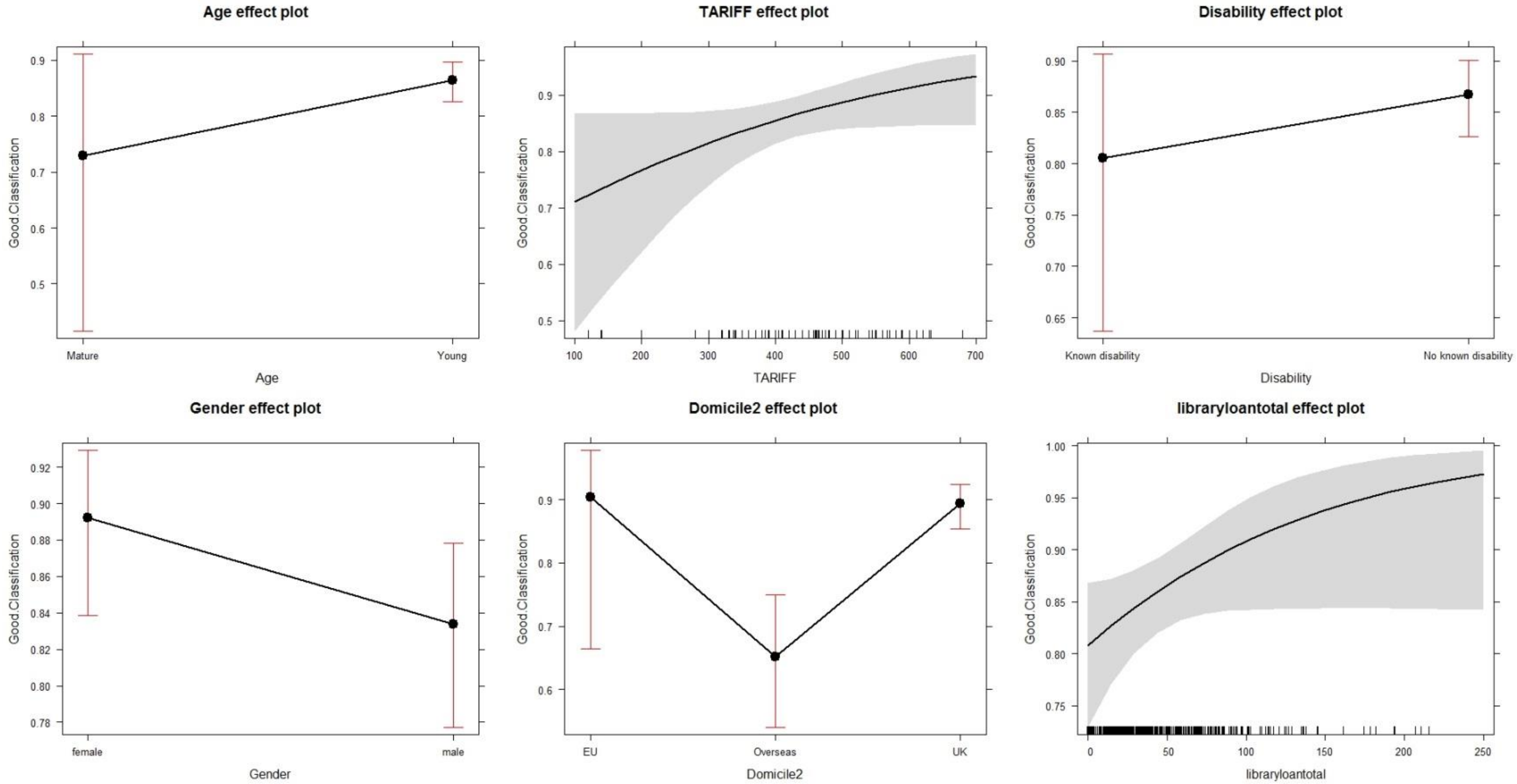
Figure: Effects Plots for SOSS Good Degree Attainment and Library Access



Library Loans

The regression model suggests that the number of library physical loans over the three years of study was not a significant predictor of degree attainment. When models were run to examine the impact of loans within each year of study on attainment and library loans significance association was only found in the first year of study ($p<0.05$). There were no significant differences in attainment across categories in the other variables and tariff was also not found to be a significant predictor of attainment.

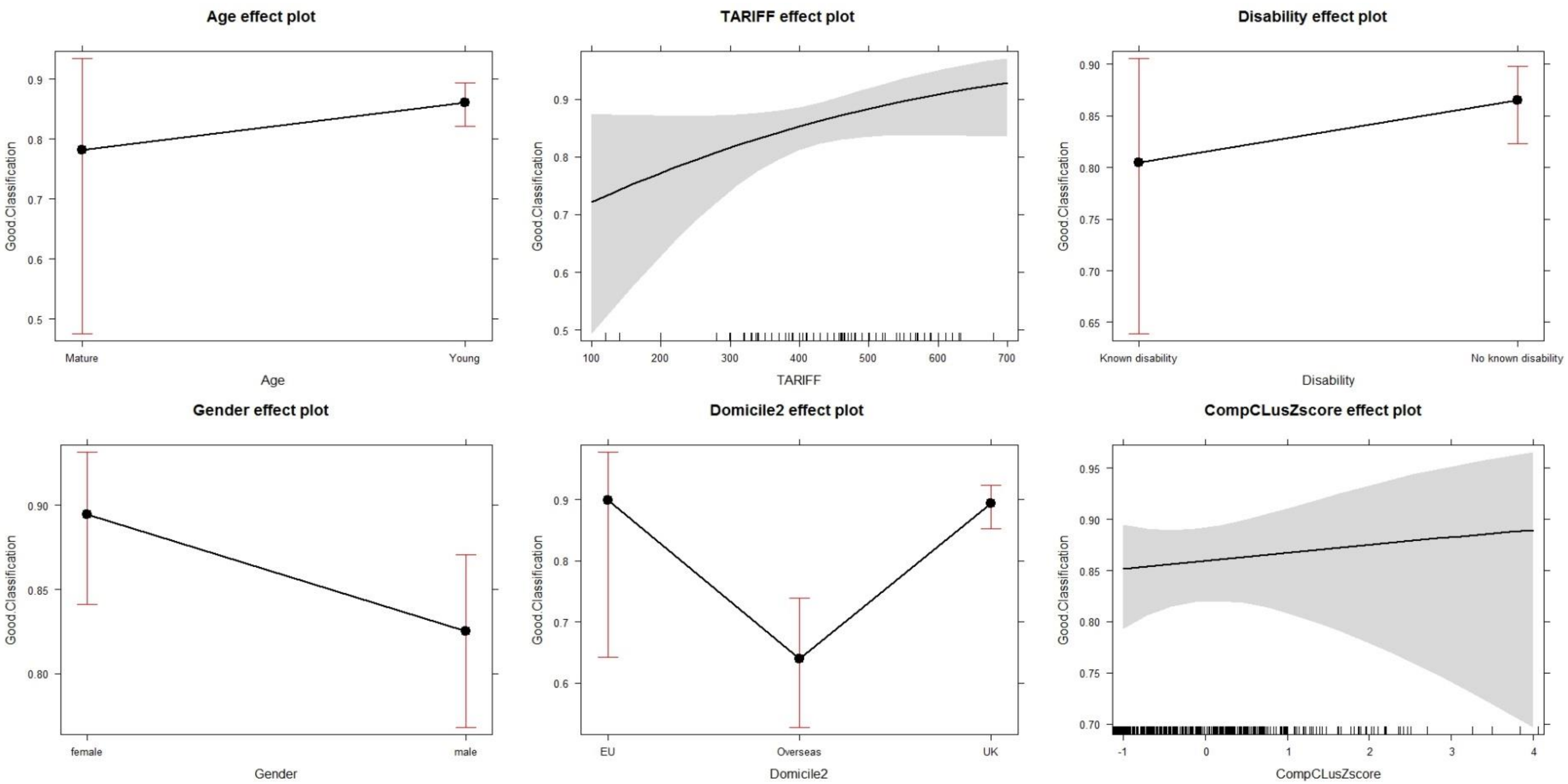
Figure: Effects Plots for SOSS Good Degree Attainment and Library Access



Computer Cluster log-ins

The regression models found that Computer cluster log-ins did not have a significant effect on attainment. In this instance only gender holds some significance ($p<0.05$) with females more likely to achieve a good degree than males. Students with a high tariff score and either UK or EU domiciles compared to overseas were more likely to gain a good degree but these differences were not significant. Computer cluster logins had no significance association with attainment in any of the years of study.

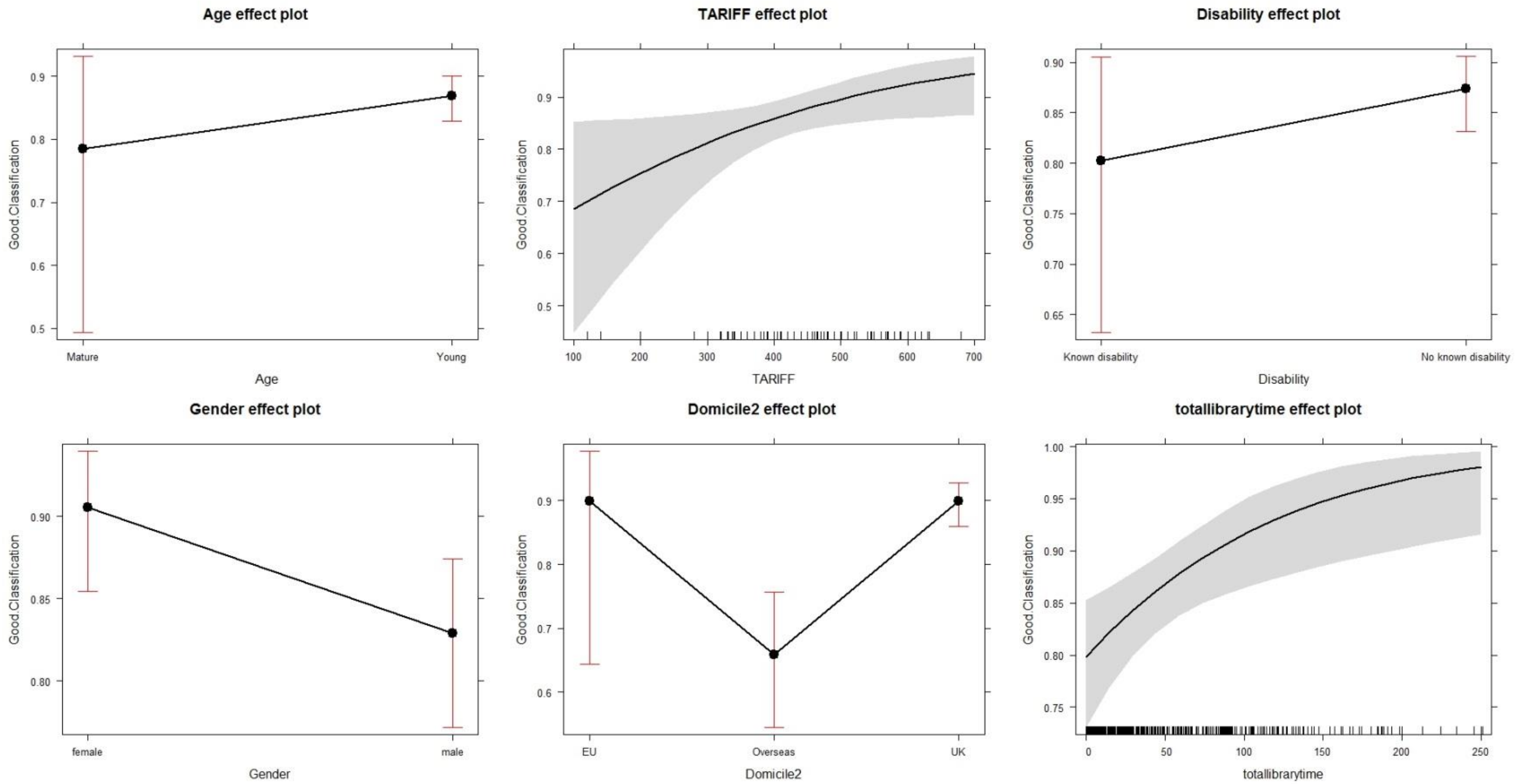
Figure: Effects Plots for SOSS Good Degree Attainment and computer cluster log-ins



Time spent in Library

Time spent in the library over the three years of study was a significant predicator of attainment ($p<0.001$). In fact, it holds the greatest significance within this model, with gender (females outperforming males) and tariff also being significant but to a less degree ($p<0.01$). UK students were significantly more likely to achieve a good classification as opposed to overseas students ($p<0.05$).

Figure: Effects Plots for SOSS Good Degree Attainment and Time spent in library



Student Retention

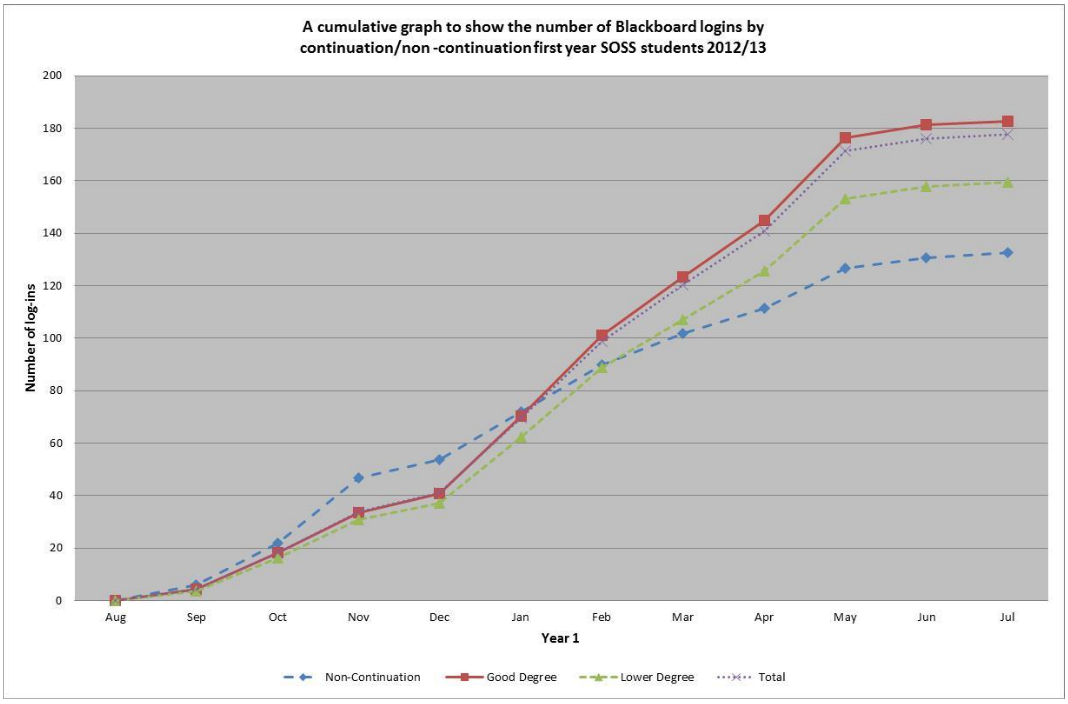
The analysis below examines the relationship between the learner analytics measurements and retention using SOSS as a case study. SOSS Data has been analysed from the academic year 2012/13 when all students were within their first year of university. Only first year data has been used as the number of non-continuation students becomes much lower in years 2 and 3. In SOSS it is important to note that the total sample of non-continuation students is small (n=38). This analysis draws comparisons of student resource engagement using the three categories of:

- Good degree students
- Lower degree students
- Non-continuation students

The monthly average learner analytics measurements for non-continuation students are based on student that left the University over the data period. The average monthly totals are based on the number of students present at the University so as students leave the University the sample of non-continuation students used to calculate the monthly averages reduces. The data is based on information linked to four of the learner analytics measurements: Blackboard log-ins, Library access, computer cluster log-ins and Alan Gilbert visits.

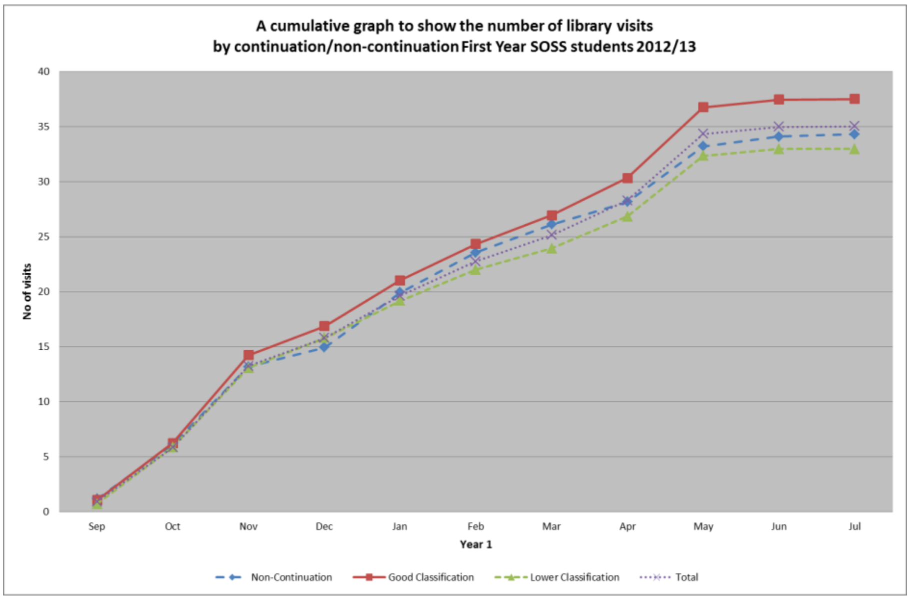
Blackboard Logins

The chart below details the cumulative frequency of Blackboard logins throughout 2012/13. Initially, non-continuation students logged into the system more, on average, than any other population. Until January in which those who achieved a good degree began to use Blackboard more frequently than their counterparts. In February, those who achieved a lower degree increased their usage of the Blackboard system; finishing the year with roughly n=15 more logins than those who did not continue their degrees.



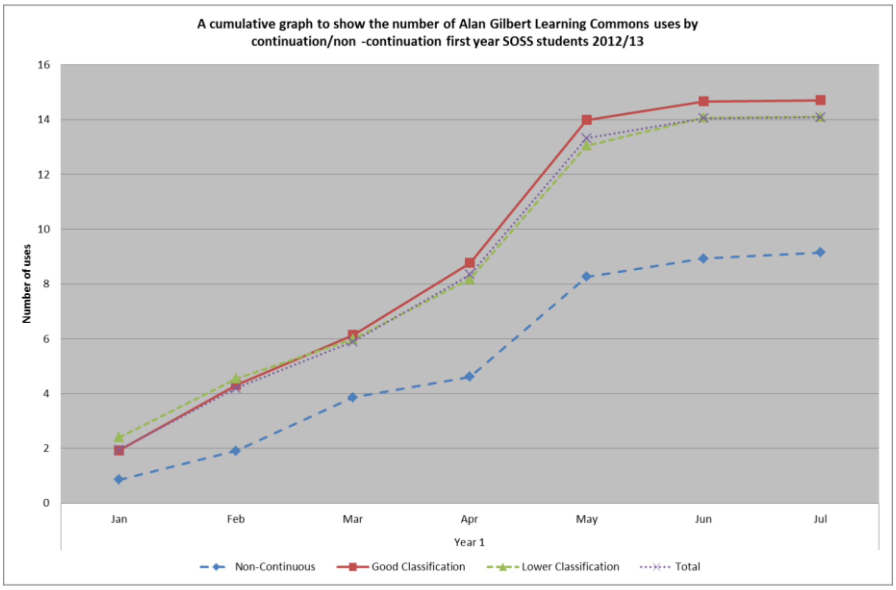
Library Access

In terms of library access, the trends for each population of student were more similar, although non-continuation students had a slightly higher number of uses than those who graduated with a lower degree. Those achieving a good degree had the greatest number of library uses (n=37) than all other populations. All populations followed the trend of an increased usage within the months of November and May.



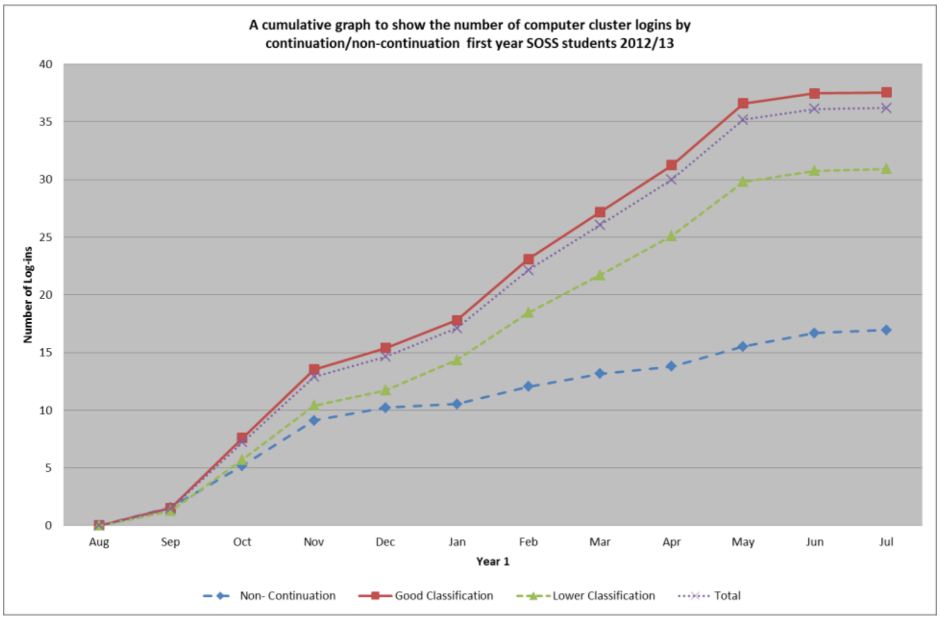
Alan Gilbert Learning Commons

Unfortunately, there is not a complete year of continuous data for the Alan Gilbert Learning Commons as it did not open until January in the academic year 2012/13. In the seven months of data available it is evident that 'good degree' students and 'lower degree' students had a similar amount of uses following the same pattern month on month whereas non-continuation students only used the Learning Commons around n=9 times within this period in comparison to the total average of n=14 times. For all populations there was a particular increase in uses in the month of May.



Computer Cluster Log-ins

When considering computer cluster log-ins, it is apparent that non-continuation students follow a similar pattern of usage until around December, after which there was only a slight increase in usage in comparison to the other two populations. This measurement shows the largest difference in usage between non-continuation and all other students with non-continuation cumulatively using this facility n=17 times throughout the year as opposed to lower classification students who used it n=31 times and good degree students (n=38 times).



The charts above show patterns of retention that suggest that the learner analytics measurements could be effectively at detecting students that are less likely to complete their studies at the University. The data is based on a limited case study but the computer cluster, Blackboard and Alan gilbert data trends show that non-continuation students tend to access these resources much less than the students that complete their studies.

Appendix 1- Data availability

The table below shows the data available for analysis, indicated by an 'X'. Due to the limitations with historical data (in particular usage of the Alan Gilbert Learning Commons and computer cluster log-ins) the sample data was taken from the cohort of students beginning in the academic year 2012/14, graduating in 14/15. They all must have completed the course without any repeated years in order to obtain comparable data. The School of Maths was initially conducted as pilot case study and is based on 2011/122 to 2012/13 data.

Academic Year	Library Access	Blackboard	Computer Cluster log-ins	Alan Gilbert	HESA Data
08_09	X				X
09_10	X				X
10_11	X				X
11_12	X	x			X
12_13	X	x	X		X
13_14	X	x	X	X	X
14_15	X	X	X	X	X
15_16	X	X	X	X	X

Appendix 2- Full details of generalised linear model (SOSS)

Blackboard Logins

```
glm(formula = Good.Classification ~ Age + TARIFF + Disability + Gender + Blackboard.Total.Use + Domicile2, family =
binomial(logit), data = newnewla)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.5463   0.3297   0.4702   0.5971   1.3822

Coefficients:
                                    Estimate Std. Error z value Pr(>|z|)
(Intercept)                       -0.2963777  1.1867512  -0.250   0.8028
Age[T.Young]                       0.5569902  0.6813462   0.817   0.4137
TARIFF                             0.0025820  0.0015359   1.681   0.0927 .
Disability[T.No known disability]  0.3183267  0.4639188   0.686   0.4926
Gender[T.male]                    -0.6132416  0.2910364  -2.107   0.0351 *
Blackboard.Total.Use               0.0014204  0.0005806   2.446   0.0144 *
Domicile2[T.Overseas]             -1.7072378  0.8342186  -2.047   0.0407 *
Domicile2[T.UK]                   -0.0689121  0.8260218  -0.083   0.9335
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 381.72  on 428  degrees of freedom
Residual deviance: 342.43  on 421  degrees of freedom
  (144 observations deleted due to missingness)
AIC: 358.43

Number of Fisher Scoring iterations: 5
```

Alan Gilbert Learning Commons

```
glm(formula = Good.Classification ~ Age + TARIFF + Disability + Gender + Domicile2 + Alan.Gilbert.Learning.Commons.
Total.Use, family = binomial(logit), data = newnewla)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.4132   0.3638   0.4550   0.5732   1.4233

Coefficients:
                                         Estimate Std. Error z value Pr(>|z|)
(Intercept)                             -0.674458   1.239906  -0.544   0.5865
Age[T.Young]                             0.971562   0.698010   1.392   0.1640
TARIFF                                   0.004048   0.001624   2.492   0.0127 *
Disability[T.No known disability]        0.573258   0.469798   1.220   0.2224
Gender[T.male]                          -0.530193   0.296960  -1.785   0.0742 .
Domicile2[T.Overseas]                   -1.838630   0.860683  -2.136   0.0327 *
Domicile2[T.UK]                         -0.173485   0.856789  -0.202   0.8395
Alan.Gilbert.Learning.Commons.Total.Use  0.002358   0.001754   1.344   0.1789
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 367.61  on 416  degrees of freedom
Residual deviance: 329.57  on 409  degrees of freedom
  (156 observations deleted due to missingness)
AIC: 345.57

Number of Fisher Scoring iterations: 5
```

Library Access

```
glm(formula = Good.Classification ~ Age + TARIFF + Disability + Gender + Domicile2 + Library.total.use, family = bi
nomial(logit), data = newnewla)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.4278   0.3550   0.4629   0.5940   1.3254

Coefficients:
                                   Estimate Std. Error z value Pr(>|z|)
(Intercept)                       -0.219746   1.191407  -0.184   0.8537
Age[T.Young]                       0.664427   0.673470   0.987   0.3239
TARIFF                             0.003155   0.001543   2.045   0.0408 *
Disability[T.No known disability]  0.507424   0.476029   1.066   0.2864
Gender[T.male]                    -0.625256   0.291948  -2.142   0.0322 *
Domicile2[T.Overseas]             -1.603524   0.826561  -1.940   0.0524 .
Domicile2[T.UK]                   -0.084259   0.817960  -0.103   0.9180
Library.total.use                  0.002344   0.001232   1.903   0.0571 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 381.72  on 428  degrees of freedom
Residual deviance: 344.98  on 421  degrees of freedom
  (144 observations deleted due to missingness)
AIC: 360.98

Number of Fisher Scoring iterations: 5
```

Computer cluster logins

```
glm(formula = Good.Classification ~ Age + TARIFF + Disability + Gender + Domicile2 + ComClustTot, family = binomial
(logit), data = Dataset1)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3802   0.3853   0.4745   0.5559   1.3714

Coefficients:
                                    Estimate Std. Error z value Pr(>|z|)
(Intercept)                        0.3829108  1.1807730   0.324   0.7457
Age[T.Young]                       0.5453575  0.7138280   0.764   0.4449
TARIFF                             0.0026835  0.0015498   1.732   0.0834 .
Disability[T.No known disability]  0.4405323  0.4671928   0.943   0.3457
Gender[T.male]                    -0.5855707  0.2938977  -1.992   0.0463 *
Domicile2[T.Overseas]             -1.6096428  0.8466842  -1.901   0.0573 .
Domicile2[T.UK]                   -0.0619533  0.8359179  -0.074   0.9409
ComClustTot                        0.0004286  0.0009546   0.449   0.6534
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 366.91  on 414  degrees of freedom
Residual deviance: 336.11  on 407  degrees of freedom
  (162 observations deleted due to missingness)
AIC: 352.11

Number of Fisher Scoring iterations: 5
```

Library time

```
glm(formula = Good.Classification ~ Age + TARIFF + Disability + Gender + Domicile2 + totallibrarytime, family = bin
omial(logit), data = Dataset1)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2984   0.3168   0.4538   0.6032   1.3496

Coefficients:
                                   Estimate Std. Error z value Pr(>|z|)
(Intercept)                       -0.433681   1.193437  -0.363  0.71631
Age[T.Young]                       0.598589   0.680227   0.880  0.37887
TARIFF                             0.003436   0.001578   2.178  0.02943 *
Disability[T.No known disability]  0.532195   0.477614   1.114  0.26516
Gender[T.male]                    -0.680070   0.295851  -2.299  0.02152 *
Domicile2[T.Overseas]             -1.526009   0.843636  -1.809  0.07047 .
Domicile2[T.UK]                    0.001777   0.834145   0.002  0.99830
totallibrarytime                   0.010085   0.003517   2.868  0.00413 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 381.72  on 428  degrees of freedom
Residual deviance: 338.95  on 421  degrees of freedom
  (148 observations deleted due to missingness)
AIC: 354.95

Number of Fisher Scoring iterations: 5
```

Library Loans

```
glm(formula = Good.Classification ~ Age + TARIFF + Disability + Gender + Domicile2 + libraryloantotal, family = bin
omial(logit), data = Dataset)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.3608  -0.5851  -0.4604  -0.3485   2.4140

Coefficients:
                                   Estimate Std. Error z value Pr(>|z|)
(Intercept)                        0.346795   1.220581   0.284   0.7763
Age[T.Young]                      -0.861870   0.693775  -1.242   0.2141
TARIFF                            -0.002916   0.001543  -1.890   0.0588 .
Disability[T.No known disability] -0.462141   0.473664  -0.976   0.3292
Gender[T.male]                     0.498721   0.291116   1.713   0.0867 .
Domicile2[T.Overseas]              1.615546   0.826842   1.954   0.0507 .
Domicile2[T.UK]                    0.111683   0.816738   0.137   0.8912
libraryloantotal                  -0.008542   0.004528  -1.886   0.0592 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 381.72  on 428  degrees of freedom
Residual deviance: 344.87  on 421  degrees of freedom
  (148 observations deleted due to missingness)
AIC: 360.87

Number of Fisher Scoring iterations: 5
```

Appendix 3 - Student Profile

Learner Analytics are important in terms of understanding students' levels of engagement and academic performance but it is also necessary to consider other student characteristics and academic information that can influence attainment levels. The analysis above has shown that levels of Blackboard log-ins are a strong predictor of attainment but other factors such as domicile, gender and ethnicity have been shown to be influence attainment. Academic advisors when viewing learner analytics data will require a full student profile to be able to make an informed decision regarding the needs of students. The profile below indicates the type of information that would be useful for academic advisors assessing students' learner analytics and attainment but this will need to be refined through consultation with IT services to identify what data is available and with system users to understand their requirements.

Fred Trueman			
STUDENT PROFILE			
Student ID	98765476	*E-mail University*	fred.trueman@manchester.ac.uk
Gender	M	*E-mail Other*	Ftrue@hotmail.com
Entry Age	18	*Module*	07954 673421
Ethnicity	White		
Domicile	English		
ACADEMIC CAREER			
Course Title	BSc Photography and Film Studies	*Academic Advisor*	Professor Plum
Mode of Study	Full-Time	*Current Registration Status*	Fully registered 2015/16
Prog Yr	2	*Currently Repeating Yr*	No
ENTRY QUALIFICATIONS			
Highest Qual on Entry	Lvl 3 A/AS Level		
UCAS Tariff	460		
All Entry Qualifications			
Type	*Year of Award*	*Grade*	*Subject*
A level	Jul-14	*A	History
A level	Jul-14	A	Maths
A level	Jul-14	A	English
A level	Jul-14	B	Chemistry
Music Grade 8	Aug-12	Distinction	N/A
ACADEMIC MODULE MARKS			
Module Code	*Module Title*	*Mark*	*Yr of Prog*
BIOL10402	Introduction to Experimental Biology	*67.5%*	1
BIOL10511	Biodiversity	*66.0%*	1
BIOL10521	Genes, Evolution and Development	67.8%	1
BIOL10532	Microbes, Man and the Environment	*75.0%*	1
BIOL10602	Field Course in Marine Biology I	*66.0%*	1
BIOL10622	Field Course in Comp and Adaptive Biol	*55.0%*	1
BIOL10632	Field Course in Freshwater Biology	67.9%	1
BIOL10811	Body Systems	*45.0%*	2
BIOL10822	Drugs: From Molecules to Man	*67.0%*	2
BIOL10832	Excitable Cells	*56.0%*	2

Note: Currently draft version of Learner Analytics student Profile. The availability of the required data fields will need to be confirmed along with the required fields for academic advisors to understand the profile of students and academic performance.