University of Huddersfield Repository

Stone, Graham

Library impact data: investigating library use and student attainment

Original Citation

Stone, Graham (2015) Library impact data: investigating library use and student attainment. In: Library Analytics and Metrics: Using data to drive decisions and services. Facet, London, pp. 51-58. ISBN 978-1-85604-965-8

This version is available at http://eprints.hud.ac.uk/id/eprint/20260/

http://eprints.hud.ac.uk/

Investigating library use and student attainment: the Library Impact Data Project at the University of Huddersfield

Graham Stone, Information Resources Manager, University of Huddersfield

Introduction and background

In 2009 Computing and Library Services at the University of Huddersfield formed a working group to investigate non/low usage of library resources with respect to distinct customer groups, which had originally been investigated as part of an earlier equality impact assessment.

The project looked at three main indicators:

- Book loans using data from the Library Management System

- Access to e-resources using click-throughs from the e-resource system

- Access to the library building using statistics from the gate entry system

The results of this analysis showed that, for all three indicators, non/low-usage ranged from 30% to 50% over a four year period. At this point it was suggested that there would be potential to investigate the relationship between the usage and final student grades. After negotiation with colleagues in Student Services it was agreed to combine this data with final grades for full time undergraduate students. Data for attainment and usage between 2005/6 and 2008/9 was also examined.

After eliminating potential anomalies such as distance learners, postgraduates, part time students, and courses with low numbers where anonymity could not be guaranteed the team began to see a relationship between usage and attainment, for both e-resources usage and library borrowing.

At this point however, the data had not yet been tested for statistical significance and it was therefore not yet known if the experience at Huddersfield was a function of the sample data used, rather than a true reflection of a relationship existing in the wider population (White and Stone, 2010).

The Library Impact Data Project

In late 2010, the University of Huddersfield, along with seven partners: University of Bradford, De Montfort University, University of Exeter, University of Lincoln, Liverpool John Moores University, University of Salford and Teesside University were awarded funding as part of the Jisc Information Environment Programme 2009-11 (Jisc, 2011) for the Library Impact Data Project (LIDP).

The original project, running from February to July 2011, aimed to use the original framework of the non/low-usage research at Huddersfield to support the hypothesis that:

'There is a statistically significant correlation across a number of universities between library activity data and student attainment'

It is important to note that the project acknowledged that the relationship between the two variables is not a causal relationship and that other factors will also influence student attainment from the outset of the project. Ultimately, the projects goal was to use the results to assist the wider higher education community by creating a better understanding of the link between library activity data and student attainment and therefore to encourage greater use of library resources and contribute to improvements in student attainment.

Phase 1 of LIDP was also required to create a number of blog posts throughout the project; the team chose to do this in a series of tagged blog posts from the project blog (University of Huddersfield, 2014; Stone, Ramsden and Pattern, 2011a). This approach proved invaluable and was adopted for the second stage of the project too.

Towards the end of 2011, the project team was approached by Jisc to put together a proposal for a second phase. Phase 2 of the project aimed to build on the original project by digging deeper into the data. The team was asked to test to see whether there was a relationship between library usage and other variables such as demographic data and discipline. Essentially to start looking into possible causal factors in the data and to see which factors carry the most weight.

Analysing the data

The first phase of LIDP looked at 33,074 students across the eight university partners. The project anticipated potential issues with data collection due to the size of the sample and the number of partners. Therefore, a minimum requirement of two out of the three indicators of library use was requested alongside final degree classification. It was felt that this would reduce the risk to the project and that partners would at least be able to provide one set of data to be used.

Phase two concentrated on 2,000 full-time undergraduate students based at the main Huddersfield campus who were awarded a final grade in July 2011. For these students additional data including demographic data, discipline and final results were extracted from Huddersfield's student record system. It also included new measures of library usage: this additional data built upon that collected in the original study and is shown in Table 1.

Measure	Notes	Years
Number of items borrowed		Three
Number of library visits		Three
Hours logged into library PC	The way the system records this means that '1 PC hour' indicates that the student was logged into the computer at least once during a single hour on a single day	Two
Hours logged into e-resources	As for hours logged into library PC	One
Number of PDF downloads		One
Number of e-resources accessed	Individual e-resources are determined by Huddersfield's systems and range from individual journal subscriptions to large journal platforms	One
Number of e-resources accessed 5		One

or more times		
Number of e-resources accessed 25 or more times		One
Percentage of e-resource usage occurring on-campus	Using total number of e-resource logins	One

Table 1: Dimensions of usage (Stone and Collins, 2013)

Information on the specific methodology used for phase 1 of the project were released as part of the LIDP toolkit (Stone, Ramsden and Pattern, 2011b), this methodology was further refined in phase 2 (Stone and Collins, 2013, Collins and Stone, 2014). In addition the data from phase 1 was released as an anonymised dataset under an Open Data Commons licence (Pattern, 2011).

Findings from the original project

Phase 1 of LIDP successfully demonstrated that there was a positive relationship between both book borrowing and e-resource usage and final degree result, "Thus, the more a book or e-resource is utilised; the more likely a student is to have attained a higher level degree result." (Stone and Ramsden, 2013) This research concurred with that of Wollongong (Cox and Jantti, 2012) and Minnesota (Oakleaf, Stone, Pattern, Bowles-Terry, Peterson, Nackerud and Fransen, 2012). Interestingly, the new data suggested that breadth of reading, which was indicated by the number of different e-resources used, might be a particularly important factor in degree success.

Interestingly, the new data suggested that breadth of reading, which was indicated by the number of different e-resources used, might be a particularly important factor in degree success.

The hypothesis was supported by all institutional partners that could supply borrowing and e-resource data. However, it was found that there was no statistical significance between gate entry and attainment. In Huddersfield's case this was because Student Services were also located in the Library at the time of the study meaning that a student may be entering the building for a number of different reasons, the same was true of the partner institutions where cafes, lecture rooms and social spaces were also part of the library space.

The project was aware of that both e-resource usage, e.g. EZproxy or Athens log-ins, and loan figures do not guarantee that a resource has been read and understood, however, as an indicator they could be benchmarked across the different institutions.

Phase 1 also acknowledged that, "The amount of data used to prove a relationship is very large, and thus is more susceptible to demonstrating a relationship." (Stone and Ramsden, 2013) The project recommended that future studies examined data at school or course level.

Demographics

Phase 2 began by looking at demographic data (Table 2) to test whether there was a statistical significance between this and undergraduate usage (Stone and Collins, 2013).

Demographic	Categories
Age	Mature (aged 21 or over on entry)
	Non-mature
Gender	Men
	Women

Ethnicity	Asian
	Black
	Mixed
	Chinese
	Other
Country of domicile (the place where students live when they are not at university)	New EU
	Old EU
	China
	Rest of world

Table 2 Demographic data examined in phase two

Throughout the project, the work of Cohen (1992) was used in classifying effect sizes:

- .1 – small effect

- .3 – medium effect

- .5 – large effect

The data showed that there were differences in use, for example:

- Mature students tend to have higher e-resource usage

- Younger students are more likely to visit the library

- Women show higher usage than men for resources use, but visit the physical library less

- Black and Asian students visit the library more often than white students and have higher PC usage and a higher proportion of their e-resource use occurs on-campus.

- Chinese students borrow fewer items than UK students and also use fewer e-resources.

The study found that, in general, the effect sizes were small; however they did indicate that there is a relationship between demographic factors and library usage and that this research supported the findings at Wollongong (Cox and Jantti, 2012).

Discipline

Phase 2 of LIDP also looked at the relationship between discipline and library usage. To do this the project grouped the 105 full-time undergraduate courses at Huddersfield into six categories, with smaller sub-categories (table 3). This allowed comparison across the 6 categories and within each category; however sub-categories could not be compared between categories. In some cases there were no sub-categories due to the particular disciplines taught at Huddersfield

High-level group	Subject groups	Number of courses included
Science	Science	3
Health	Nursing	5
	Health	7
Computing and engineering	Computing	13
	Engineering	6

Arts	Music	5
	Architecture	2
	Fashion	7
	2D Design	3
	3D Design	4
Humanities	English	2
	Drama	2
	Media & Journalism	6
Social sciences	Business, management and accountancy	22
	Law	2
	Behavioural sciences	9
	Social work	3
	Education	4

Table 3: Discipline variables (Collins and Stone, 2014)

The project found that the social science group was a significantly higher user than the other groups and arts were the lowest users for e-resource usage and PDF downloads (this perhaps reflects the way arts students use the library, particularly the physical resources).

Within the main categories, behavioural sciences were the highest user in the social sciences group. Business had higher usage than law, social work and education, but borrowed fewer items. Lawyers are extremely low users of library resources.

The full set of results from the study into discipline differences can be found of the LIDP project blog (University of Huddersfield, 2014).

Retention
The project also looked at the relationship between retention and library usage. Students who dropped out in the first two terms were excluded from analysis, and a cumulative measure of usage for the first two terms of the 2010-11 academic year was examined, this allowed all users (particularly first years) time to use the libraries resources and to establish usage patterns, otherwise a student who used nothing and dropped out in week two would skew the data. Thus students included in the study were at the university in the first two terms, and have all had the same opportunity to accumulate usage.

The study found that there is a significant relationship between e-resource usage and book borrowing and student retention. This does not mean that non/low use leads to dropping out, but that there is a relationship. Therefore non/low use may be an indicator; in the same way that a drop in attendance can be a sign of possible retention issues.

Further research

At Huddersfield
Over the course of the project, statistical significance was shown for the relationship between library usage and student attainment and retention. Differences were also recorded for certain groups. However, these results are merely an indicator; they do not tell us the reasons for the differences.

Further work is required at Huddersfield to understand these differences. As part of both phases of the project a number of focus groups were held, in particular a focus group with students from computing and engineering helped the project group to understand the behaviour of students in this area. In some cases the data needs to be broken down further, for example, measuring age in only two groups does not go very far in explaining any possible behaviour.

However, the study did give some useful intelligence on some of the bigger groups, must notably discipline and country of domicile. Work at Huddersfield has started to concentrate on some of these areas where non/low use was found. If targeted successfully there could be a rise in use of library resources, which in turn could be a contributing factor to better student attainment. The strategy for engagement is to approach both staff and students. In depth subject knowledge of the Academic Services Team will be used to identify whether marketing and promotion or information literacy sessions are required for the various cohorts. The results certainly imply that a one-size fits all approach is inappropriate for information literacy sessions and that library analytics can enhance the understanding of student behaviour.

The project concluded by suggesting that, like the first stage of LIDP, it would be useful to replicate phase 2 with data from a wider range of universities.

The wider impact of LIDP

During both phases of LIDP, data extraction and processing took as long as four months, clearly, if this is to become a regular exercise rather than a project, there is a need to automate this process as much as possible. Towards the end of the second phase of LIDP, Huddersfield and Mimas collaborated on a library analytics survey to understand any potential demand for a data analytics service, which could enhance business intelligence at an institutional level to support strategic decision-making and whether there was appetite for a shared service approach to process the raw data and to provide analytics tools and data visualisations back to local institutions. The survey received 66 replies from library staff, including many library directors, 96% of those who replied confirmed that they would want automated provision of analytics demonstrating the relationship between student attainment and library usage within in their institution, with 94.6% wanting to benchmark their data with other institutions. Furthermore, 87.7% were interested in the richer data that was used as part of the second phase of LIDP, e.g. discipline, age, year, nationality and grade.

The key strategic drivers for the use of library analytics identified by the library survey were, perhaps unsurprisingly:

1. Enhancing the student experience
2. Demonstrating value for money
3. Support research excellence

A subsequent meeting of representatives from the LIDP and Copac projects, Jisc, SCONUL and RLUK decided that there was sufficient evidence demonstrating the need, and desire, for a shared analytics service. This resulted in the funding of the Library Analytics and Metrics Project (LAMP) (Jisc, 2013). The project is a partnership between Jisc, Mimas and the University of Huddersfield.

The original LIDP was a very successful partnership between the eight collaborators, showing that a number of very different institutions could retrieve and share data in a very short pace of time. In

addition, LIDP also made international contacts with Wollongong and Minnesota and the three projects found themselves on very similar paths. The LAMP project is also liaising with both Wollongong and Minnesota with a view to finding common themes. One such theme is 'a striking variety in terminology' (Poll, 2012, p.122), no two universities seem to agree on ethnic groups or subject classification and there are also differences in the way attainment is measured, e.g. degree classification verses Grade Point Average.

Conclusion

Phase 1 of LIDP fully supported the original project hypothesis that that there was a positive relationship between book borrowing and e-resource usage and final degree result for all eight project partners. However, the project concluded that this was not a causal relationship and that other factors would have an influence.

The second phase went on to look at demographic and discipline differences at Huddersfield and found that there were significant differences among groups of students for both area of study. In addition the phase two found that there was also a statistical significance between the number of items borrowed, the hours logged into e-resources and the number of PDF downloads and retention. Like the earlier study the relationship was not one of cause and effect, however, it could be a predictor and therefore a useful warning signal in addition to course tutor feedback, lack of attendance at lectures etc. Finally, the second phase also revealed that there was a statistical significance between attainment and the number of e-resources accessed. This suggests that breadth of reading might be an important factor in degree outcome.

It is hoped that the work of LIDP and now LAMP can foster further international collaboration and start to tease out a set of best practice guidelines and standards to give a robust mixed methods approach to library analytics.

Acknowledgements

References

Cohen, J. (1992), A power primer, Psychological Bulletin, 112 (1), 155-159.

Collins, E. and Stone, G. (2014) Understanding patterns of library use among undergraduate students from different disciplines. In press.

Cox, B.L. and Jantti, M. (2012). Capturing business intelligence required for targeted marketing, demonstrating value, and driving process improvement. *Library and information science research*, 34 (4), 308-316. http://dx.doi.org/10.1016/j.lisr.2012.06.002

Jisc. (2011). *Jisc Activity Data*. Retrieved from
http://www.jisc.ac.uk/whatwedo/programmes/inf11/activitydata.aspx

Jisc (2013) *Library Analytics and Metrics Project*. Retrieved from http://jisclamp.mimas.ac.uk/

Oakleaf , Megan, Stone, Graham, Pattern, David, Bowles-Terry, Melissa, Peterson, Kate, Nackerud, Shane and Fransen, Jan (2013) Do or Do Not...There is No Try: The Quest for Library Value. In *Association of College & Research Libraries Conference*, 10-13 April 2013, Indianapolis, Indiana. Retrieved from http://eprints.hud.ac.uk/16421/

Pattern, D. (2011) *Library Impact Data Project Data*, University of Huddersfield, Huddersfield. Retrieved from http://eprints.hud.ac.uk/11543/

Poll, R. (2012), Can we quantify the library's influence? Creating an ISO standard for impact assessment, *Performance Measurement and Metrics*, 13 (2), 121 – 130. http://dx.doi.org/10.1108/14678041211241332

Stone, G. and Collins, E. (2013). Library usage and demographic characteristics of undergraduate students in a UK university. *Performance Measurement and Metrics*, 14(1), 25-35. http://dx.doi.org/10.1108/14678041311316112

Stone, G. and Ramsden, B. (2013). Library Impact Data Project: looking for the link between library usage and student attainment. *College and Research Libraries*, 74(6), 546-559. Retrieved from http://crl.acrl.org/content/74/6/546.abstract

Stone, G., Ramsden, B. and Pattern, D. (2011a). *Looking for the link between library usage and student attainment.* Ariadne (67). Retrieved from http://www.ariadne.ac.uk/issue67/stone-et-al/

Stone, G., Ramsden, B. and Pattern, D. (2011b). *Library Impact Data Project Toolkit.* Manual. University of Huddersfield, Huddersfield. Retrieved from http://eprints.hud.ac.uk/11571/

University of Huddersfield (2014) Library Impact Data Project. Retrieved from http://library.hud.ac.uk/blogs/projects/lidp/

White, S. and Stone, G. (2010) Maximising use of library resources at the University of Huddersfield. *Serials* 23 (2), 83-90 http://dx.doi.org/10.1629/2383