University of Huddersfield Repository

Stone, Graham and Collins, Ellen

Library Impact Data Project Phase II: the data strikes back

Original Citation

Stone, Graham and Collins, Ellen (2012) Library Impact Data Project Phase II: the data strikes back. Taking Stock, 21 (1). pp. 6-7. ISSN 0966-6745

This version is available at http://eprints.hud.ac.uk/id/eprint/13536/

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

Library Impact Data Project Phase II: the data strikes back

Graham Stone and Ellen Collins

1,142 words (including references)

The Library Impact Data Project (LIDP) brought its initial findings to the NAG Conference in September 2011 [1]. Inspired by earlier work undertaken at the University of Huddersfield [2, 3], the project aimed to investigate the hypothesis that:

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

Phase 1 of the project was funded through the JISC Activity Data programme and ran from February – July 2011 with 7 partners; University of Bradford; De Montfort University; University of Exeter; University of Lincoln; Liverpool John Moores University; University of Salford and Teesside University. The project successfully demonstrated a statistically significant relationship across a number of universities between library activity data (specifically the number of items borrowed and logins to e-resources in the library) and student attainment.

Phase I outputs

LIDP blogged throughout the project using a number of themed blog posts [4]. After consultation with the partners, anonymised data was made available in Excel, comma separated and plain text and contains final grade and library usage figures for 33,074 students studying undergraduate degrees at the eight partner universities under an Open Data licence [5]. The final output from the project was a toolkit [6], which provides instructions for libraries on how to extract their own data in order to benchmark against the data described above. The toolkit discusses the extraction of the data and gives tips for statistical analysis and suggestions for further investigation. Phase I of the project also produced a number of journal articles and conference papers and was recently referenced in the SCONUL response to the Higher Education White Paper "Higher Education: Students at the Heart of the System" [7].

Introducing Phase II

In December 2011, JISC approached Huddersfield to submit a proposal for Phase II of LIDP in order to deepen our understanding of the statistically significant relationship between students' degree outcomes and their usage of library resources. Phase II runs from January – July 2012 and will use additional data, including percentage-level data about student attainment, to identify predictors for student outcomes, including negative outcomes such as non-completion, with a particular focus on engagement with the library and library services.

As with Phase 1, the work concentrates on full-time, undergraduate students' final grade, the course title and several variables relating to library usage:
  * the number of books borrowed
  * the number of times e-resources were accessed
  * the number of times each student entered the library

However Phase II looks at the percentage score, in addition to the class of degree and incorporates this with the following variables:

- Gender
- Age
- Country of domicile
- Declared disability
- Retention
- VLE usage
- School
- Course
- Library PC usage

The project will also use more granular breakdowns of the library usage variables including:
- Number of items borrowed:
  - Per term (first, second and third)
  - Day (8am-8pm) or night (8pm to 8am)
- Number of books borrowed:
  - Per term (first, second and third)
  - Day (8am-8pm) or night (8pm to 8am)
- Number of times student entered the library
  - Per term (first, second and third)
  - Day (8am-8pm) or night (8pm to 8am)
- Number of times e-resources accessed
  - Per term (first, second and third)
  - Each hour
- Types of e-resources accessed (for 2010-11 only)
  - Total number of unique resources accessed
  - Number of resources accessed 5 or more times

Data analysis

Phase I found the hypothesis that "There is a statistically significant correlation across a number of universities between library activity data and student attainment" to be correct. Phase II does not plan to re-test this hypothesis, but will build on phase I through a number of work packages:

1. **Demographic factors and library usage**. To test to see whether there is a relationship between demographic variables (gender, ethnicity etc.) and all measures of library usage, and to see which factors carry the most weight in such a relationship.
2. **Retention vs non-retention.** To test to see whether there is a relationship between patterns of library usage and retention – patterns will include issues such as increasing or decreasing intensity of usage and time of usage.
3. **Value added**. UCAS entry data and library usage data will be used to establish whether use of library services has improved outcomes for students. We will compare students with high UCAS entry points and high and low library service usage, and then compare students with low UCAS entry points and high and low library service usage. This work package is subject to availability of UCAS data for every student in the sample.
4. **VLE usage and outcome**. Subject to data availability, this data will be tested to see whether there is a relationship between VLE usage and outcome.

The aim of these tests is to understand the relationship between behaviour, demographics and outcome, so as to improve the library's ability to identify students that may need additional support.

Qualitative case studies

Phase II will be holding a number of workshops to expand on the findings from the quantitative

analysis in order to explore why the relationships identified in this work might exist. They will also explore questions raised during Phase I of the project, and will give an understanding of what students consider to be low usage to compare with the definition used in previous analysis. After preliminary findings from the quantitative work, groups in Human and Health Sciences, Computing and Engineering and the Business School have been identified alongside two groups representing international students and students with a declared disability.

Outputs
Phase II of LIDP will create a revised toolkit including a set of models which will allow other universities to understand how they can explore their own data. As in Phase I, all data will be anonomised and released under an Open Data Licence.

The LIDP blog and Twitter hashtag (#lidp) will be used throughout the project to document progress.

Acknowledgements

References
1.   Stone, G. (2011) The Library Impact Data Project. In: National Acquisitions Group Conference, 7-8 September 2011, Manchester. http://eprints.hud.ac.uk/11389/
2.   Goodall, D. and Pattern, D. (2011) Academic library non/low use and undergraduate student achievement: a preliminary report of research in progress. *Library Management*, 32 (3), 159-170. doi:10.1108/01435121111112871. http://eprints.hud.ac.uk/7940/

3.   White, S. and Stone, G. (2010) Maximising use of library resources at the University of Huddersfield. *Serials*. 23 (2), 83-90. doi:10.1629/2383. http://eprints.hud.ac.uk/7811/

4.   Library Impact Data Blog: http://library.hud.ac.uk/blogs/projects/lidp/

5.   Pattern, D. (2011) *Library Impact Data Project Data.* http://eprints.hud.ac.uk/11543/

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

7.   SCONUL response to the Higher Education White Paper "Higher Education: Students at the Heart of the System": http://www.sconul.ac.uk/news/he_whitepaper/hewhitepaper_response.doc