

## **Using Excel Macros to Analyse Moodle Logs**

*Dr. Andreas Konstantinidis*

*Centre for Technology Enhanced Learning, King's College London, UK,  
andreas.konstantinidis@kcl.ac.uk*

*Dr. Cat Grafton*

*Learning & Teaching Team, School of Arts & Humanities, King's College London, UK,  
cat.grafton@kcl.ac.uk*

### **Abstract**

Learning analytics enables tutors to gain useful insights on the behaviour of students in an online learning environment. This information can then be utilised to customize the educational space, optimize the learning resources and activities, and personalize the student experience. This paper presents our approach to analysing the data of users' behaviours that are recorded in the Moodle logs. Currently, the Moodle logs manager suffers from functional limitations and uninspiring visualizations. Our method utilises the possibility of downloading the logs in Microsoft Excel format and provides a simple and effective offline solution. The method we have developed is based on Excel macros and visual basic. Tutors can experiment with different combinations of metrics such as total page views, unique users, unique actions, IP addresses, unique pages, average session length and bounce rate. Furthermore, the software allows the definition of date ranges and the selection of individual or groups of students. The complicated processes of analysing and combining data are carried out in the background, enabling tutors to focus on the pedagogic implications and invest in practical, realistic scenarios through informed decision-making. Future work includes transferring the offline functionality to an online Moodle plugin and increasing system intelligence to allow the production of meaningful and actionable suggestions with regards to set target goals.

### **Keywords**

learning analytics, Moodle, Excel, macros, visual basic

## **Introduction**

Analytics marries large data sets, statistical techniques, and predictive modelling. It could be thought of as the practice of mining institutional data to produce actionable intelligence (Campbell, DeBlois & Oblinger, 2007). More specifically, learning analytics collects and analyses the behaviours of students in the online learning environment and looks for correlations between those behaviours and the resulting or desired learning outcomes. In other words, 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 (a definition set by the First International Conference on Learning Analytics and Knowledge in 2011).

Moodle is increasingly used in schools, universities and companies (Dierenfeld & Merceron, 2012). Moodle analytics allows institutions to accumulate information which can be used for analysing students' behaviour within a virtual learning environment (Romero, Ventura & García, 2008). Through these analytics, it would be possible to evaluate students' online behaviour, explore what this behaviour can tell us about how students learn online and identify various departmental pedagogical disciplinary practices (Martín-Blas & Serrano-Fernández, 2009).

Moodle systems record user behaviours in the form of logs. Unfortunately, the way the data is presented is unstructured, isolated, abstract, and therefore meaningless. Consequently, non-technical users have great difficulty in utilizing this data to arrive at meaningful conclusions and actionable information which would allow them to customize their applied pedagogy and personalize the student experience.

This paper is structured as follows: In the next section we present related work on the subject of learning analytics in Moodle and clarify the differentiation of our approach. Following this, we expand on the description of our method and examine the application of Microsoft Excel macros, visual basic and the developed

visualizations. In the final section, we discuss conclusions and lessons learnt and elaborate upon the future steps of our work.

## Related work

This section presents related research on Moodle analytics and justifies the method of our approach. Romero & Ventura cite a number of researchers who have focused on analysing usage data in Moodle. On the other hand, Graf et al. (2011) and Leony et al. (2012) present a number of tools that can be integrated with Moodle, which will enable users to explore the logs. However, most of these solutions focus on complicated and specialised data mining techniques which prove incomprehensible to the general academic audience.

For example Zafra & Ventura (2009) apply genetic algorithms to Moodle logs in order to predict the grades of students. Alternatively, Dierenfeld & Merceron (2012) use pivot tables in Microsoft Excel 2010. Others, like Jawaheer et al. (2012), invest in third-party proprietary systems such as Sawmill. Sawmill is a web server log analysis software, which can be customized for the Moodle logs. Through database queries it can be used to extract and aggregate data from Moodle.

Taking related work, such as the aforementioned research, into consideration we decided to design our approach around three principles:

- Specialised procedures such as data analysis, synthesis and modelling should be carried out in the background.
- Utilize the ability to download Moodle logs in Excel format, in order to simplify the process of exporting data from Moodle and importing it for analysis.
- The system interface should be simple, highlighting useful information, suggesting conclusions, and supporting decision making through meaningful data visualization. Visualizing information is a form of knowledge compression (McCandless, 2012).

The main goal of our approach is to enhance the awareness of academics and students by providing them with an overview of their behaviours in the online course. Raising awareness builds self-reliance, confidence, and responsibility (Whitmore, 2009). This new awareness can then be used to optimize the online course space, enrich the student experience and generate new discussions around individual, group or average student routines and patterns of behaviour. It is also important to explore how learning analytics can be exploited to provide an alternative assessment system through formative feedback and engagement and its relationship to the learners' context and learning design (San Diego et al., 2012).

## Description of our approach

This section expands on the details of the Excel macros approach. The paper explains the process of bringing data from Moodle to Microsoft Excel and introduces the main visualizations used.

The screenshot shows the Moodle logs manager interface for the course '6AAEC060 12-13 SEM1 1 TWENTY FIRST CENTURY AMERICAN FICTION: POSTMODERNISM AND BEYOND'. The interface includes navigation links, a search bar, and a table of logs. The table has columns for Time, IP address, User full name, Action, and Information.

Time	IP address	User full name	Action	Information
Fri 26 April 2013, 8:35 PM	193.82.8.7	Laura Prober	turnitool view	User viewed assignment 'Coursework submission'
Fri 26 April 2013, 8:35 PM	193.82.8.7	Laura Prober	turnitool view	User viewed assignment 'Coursework submission'
Fri 26 April 2013, 8:35 PM	193.82.8.7	Laura Prober	turnitool view	User viewed assignment 'Coursework submission'
Fri 26 April 2013, 8:35 PM	193.82.8.7	Laura Prober	course view	6AAEC060 12-13 SEM1 1 TWENTY FIRST CENTURY AMERICAN FICTION: POSTMODERNISM AND BEYOND
Fri 26 April 2013, 3:43 PM	197.75.127.16	Michael Miller	url view	Jameson
Fri 26 April 2013, 3:43 PM	197.75.127.16	Michael Miller	course view	6AAEC060 12-13 SEM1 1 TWENTY FIRST CENTURY AMERICAN FICTION: POSTMODERNISM AND BEYOND
Thu 25 April 2013, 4:49 PM	197.75.127.16	Michael Miller	forum view	Consumerist Culture
Thu 25 April 2013, 4:49 PM	197.75.127.16	Michael Miller	discussion	
Thu 25 April 2013, 4:49 PM	197.75.127.16	Michael Miller	forum view	Consumerist Culture

Figure 1: The Moodle logs manager (anonymised)

## Bringing data into the spread sheet

Firstly, it should be noted that Moodle logs are intended as reports of site activity, and not learning behaviour. The functionality of the Moodle logs manager is limited (Figure 1). For instance you can only focus on specific dates, or all dates. There is no functionality for defining date ranges. This would be useful if you'd like to focus on a date range before an exam period or during an assignment. Furthermore, although you can focus on specific users, the data presented is segmented and disconnected from the general overview, making utilisation difficult. Finally, visualizing the data in a table is not the optimum approach for providing instant access to important insights of user behaviours.

Moodle allows you to download the logs of user actions and behaviour in three different types of format: text, ODS and Excel. Our approach is a simple two-step process of users downloading the log file in Excel format and loading it into the learning analytics spread sheet. Functionality in the spread sheet has been developed using macros and visual basic.

After loading the Moodle log file into the learning analytics spread sheet, a pop-up box of the unique users recorded in the log appears (Figure 2). This can be used to focus on the data recorded for a single user or a group of users. In addition, users can choose to focus on students alone, by deselecting administrators and tutors. Once users have been selected, the data is analysed in the background and presented through a series of graphs and options.

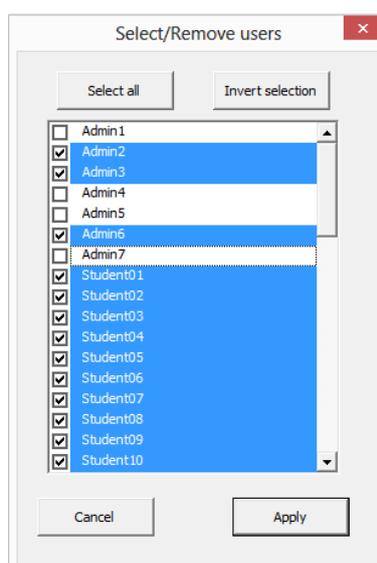
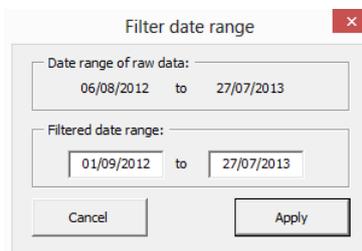


Figure 2: Selecting users for analysis (anonymised)

## A description of the interface and visualizations

Once the data has been analysed, relevant information populates the available graphs and fields. At the top of the spread sheet, the name of the course space and three buttons appear. The “Get Data” button allows users to load another Moodle log file. The “Select/remove users” button can be used to refocus the data that is presented by adding or removing users. Lastly, the “Filter dates” button enables the definition of date ranges (Figure 2).



**Figure 3: Date filter for data analysis**

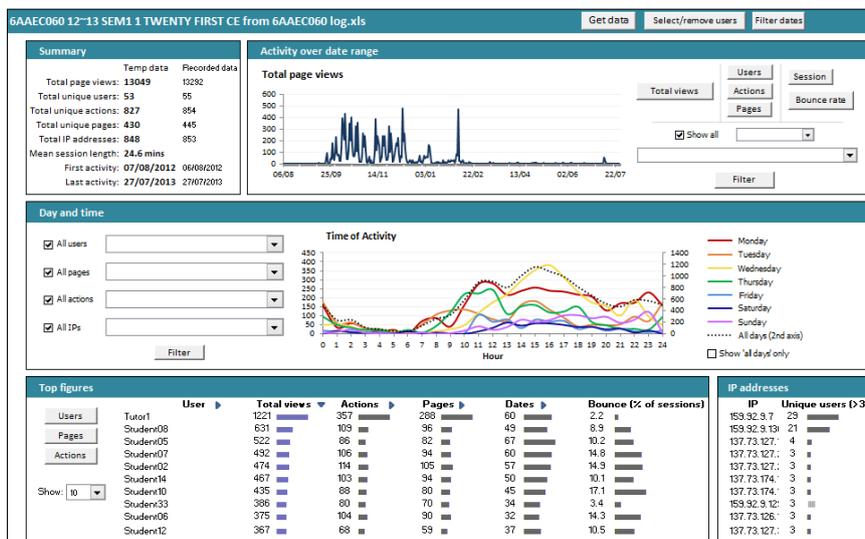
The “Summary” section presents a quick overview of the data in the log in two columns of data. The “Temp data” column refers to the data in the specified date range and selected users. Alternatively, the “Recorded data” column refers to all the data that has been recorded in the log. Table 3 describes the metrics which are presented in this section and used throughout the spread sheet.

Table 3: The “Summary” section metrics

<b>Total Page views</b>	This measures the number of times any page within the course space has been accessed.
<b>Total unique users</b>	This is the number of unique users that have accessed the course space at least once.
<b>Total unique actions</b>	This is the number of unique actions that have been carried out in the course space at least once. The nature of an action depends on the context of the activity or resource that is being used. Typical Moodle actions are View, Add, Delete etc.
<b>Total unique pages</b>	This is the number of pages that make up the structure of a course space.
<b>Total IP addresses</b>	These are the unique IP addresses which have been recorded to have accessed the course space at least once.
<b>Mean session length</b>	This refers to the average amount of time that users spend inside the course space. It should be noted that Moodle logs do not record the logging in and logging off of users. Therefore, this metric is an approximation and is based on the automatic log off time, as set by the administrators. For example, if the log off time is set to 90 minutes, and the user performs two actions with a gap of more than 90 minutes between them, then the second action is considered part of a new, different session. Because this metric is an approximation, it is perhaps more meaningful to focus on its fluctuations over time, rather than on the value of the metric itself.

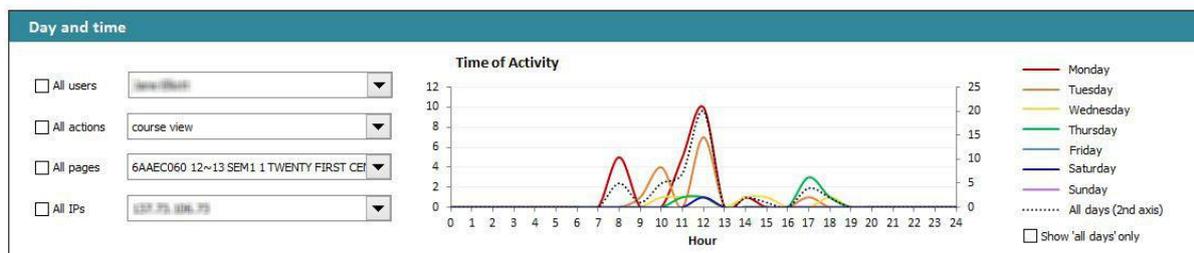
The next section is called the “Activity over date range” section. This section presents the same data as the “Summary” section, but in graphical format and with relation to unique days within the specified date range. Furthermore, it introduces a new metric called “Bounce rate”. The “Bounce rate” refers to the percentage of sessions in which a user enters the course space, but leaves after viewing just one page. With regards to the objective of the course space, the “Bounce rate” can be used to determine the effectiveness of the central or entry page. An entry page with a low bounce rate can mean that users are enticed to continue on deeper into the

course space. In addition, this section includes a filtering mechanism which enables focus on activity relating to a specific user, page, action or IP address. For example, the flow of popularity for a specific resource over time can be presented.



**Figure 4: The main Learning Analytics interface in Microsoft Excel (anonymised)**

The “Day and time” section presents the activity relating to users, actions, pages and IP addresses, with a focus on the time of day. In other words, you can see which day of the week the course space is more active and specifically the flow of activity within a 24 hour timescale. The graph also presents the sum of all days and times. This section also provides filtering functionality, which enables focus on data relating to specific users, actions, pages and IP addresses. The drop-down menus used for filtering can also be applied in combination with one another. For example, you can visualize which day and time a particular student prefers to perform a certain action, on a specific activity from a singular IP address (Figure 5).



**Figure 5: Narrowing the data to a specific user, action, page and IP address pattern (anonymised)**

The next section is called the “Top figures” section. This section sorts users, pages and actions based on activity. Therefore, it is possible to immediately recognize the most or least active students, as well as the most or least popular resources in the course space. More specifically, users can be sorted based on their total page views, unique actions, unique page views or unique dates. In a similar fashion, pages and actions can be sorted based on total views, unique users and unique dates.

The final section is the “IP Addresses” section. This section presents a list of IP addresses which have been used by more than three different users. Popular IP addresses could possibly belong to a specific computer lab within a campus. This information can be used to organize group activities and collaboration.

## Conclusions and future work

Paul Otellini, president and CEO at Intel, was asked in 2010: “What is going to be obsolete next?” Otellini responded: “Ignorance”. According to Phil Long and George Siemens (Norvig, 2012), “the most dramatic factor

shaping the future of higher education is something we can't actually touch or see: big data and analytics." Learning analytics enables us to add context to numbers, correlate different data sets and make the invisible, visible. Using activity logs as a means to investigate learner behaviour involves a complex process. However, sensible use of analytics will enable institutions to make significant improvements in student learning services. Furthermore, among other things, learning analytics can be used to (JISC e-Learning Blog, 2012):

- Identify students at risk, in order to provide positive interventions designed to improve retention
- Provide recommendations to students, in relation to learning resources and activities
- Detect the need for, and measure the results of, pedagogic improvements
- Optimize, customize and personalize course offerings
- Identify tutors who are performing well, and tutors who need assistance with teaching methods

Future work of our project will focus on four areas: a) academic feedback, b) a student version, c) Moodle integration, and d) intelligent agents.

With regards to academic feedback, we aim to expand our research by allowing tutors to use the system in their classes for the new academic year. Based on their feedback, suggestions and recommendations, we will extend and enhance system functionality. For example, additional functionality could include the ability to compare different academic years, or separate modules in a department. The current software has been built, from the ground up, as a system aimed at allowing tutors an overview of their online courses. A different version directed at students can also be envisaged. This would grant students a sense of progress through the material, nudge them with reminders or behavioural stimuli, and notify them of best or alternative practices with relation to their current or past peers.

Furthermore, our work will examine bringing the developed offline functionality and logic to an online system. Our exploration will focus on integrating the learning analytics dashboard to Moodle as an additional plug-in. The main concerns to this approach include user privacy, hardware limitations and server overhead. Finally, in the long term, research could focus on converting the passive learning analytics visualizations into an active artificially intelligent agent. The agent would be capable of analysing and synthesizing data, in order to facilitate predictions, propositions and informed decision-making.

Moodle has a large and diverse community with over 1 million registered users in 231 countries, making it the most popular open source learning management system. It is therefore part of our commitment, to release the functionality we have and will be developing to the wider research and academic community.

## References

- Campbell, J.P., DeBlois, P.B. & Oblinger, D.G. (2007). Academic Analytics: A New Tool for a New Era, *EDUCAUSE Review*, vol. 42, no. 4 (July/August 2007), pp. 40–57.
- Dierenfeld, H. & Merceron, A. (2012), Learning Analytics with Excel Pivot Tables. In Proceedings of the 1st Moodle Research Conference (MRC2012), Retalis, S. & Dougiamas, M. (Eds), 115-121.
- Graf, S., Ives, C., Rahman, N. & Ferri, A. (2011). AAT – A Tool for Accessing and Analysing Students' Behaviour Data in Learning Systems. In LAK2011 Proceedings of the Conference on Learning Analytics & Knowledge, Banff, Alberta, Canada, February 27 – March 01, 2011). ACM New York, NY, USA .
- Jawaheer, G., Kostkova, P., Sultany, A., Bullimore, A. (2011) Final Report on the Moodle Analytics Project, City eHealth Research Centre & Learning Development Centre, online at: <http://tinyurl.com/Jawaheer>
- JISC e-Learning Blog (2012). Analysing analytics, December 2012, available at: <http://elearning.jiscinvolve.org>
- Leony, D., Pardo, A., De La Fuente Valentín, L., Sánchez De Castro, D. & Delgado Kloos, C. (2012). GLASS: A first look through a Learning Analytics System. In LAK2012 Proceedings of the Second International Conference on Learning Analytics & Knowledge, Vancouver, British Columbia, Canada, April 29 – May 02, 2012).
- Martín-Blas, T., & Serrano-Fernández, A. (2009). The role of new technologies in the learning process: Moodle as a teaching tool in Physics. *Computers & Education*, 52(1), 35–44.
- McCandless, D. (2012). The beauty of data visualization, November 23, available at: <http://www.youtube.com>
- Norvig, P. (2012). The 100.000 student classroom, February 2012, available at: <http://www.ted.com>

- Otellini, P. (2010). Interview, Charlie Rose, February 26, available at: <http://www.charlierose.com/>
- Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368–384.
- San Diego, J.P. et al (2012), Do Moodle analytics have a role to play in learning design, feedback and assessment?. In *Proceedings of the 1st Moodle Research Conference (MRC2012)*, Retalis, S. & Dougiamas, M. (Eds), 61-68.
- Whitmore, J. (2009) *Coaching for Performance: GROWing Human Potential and Purpose*, Nicholas Brealey Publishing; 4th Revised edition edition, ISBN-13: 978-1857885354.