

Benefits from Analyzing Video Lecture Logs with Leading Business Analytics Tools

Linda ALKSNE¹, Anita JANSONE¹, Zane BĒRZKALNE²

¹Liepāja University, Lielā Street 14, Liepāja, LV-3401, Latvia
²Rīga Technical University, Kalku Street 1, Rīga, LV-1658, Latvia

`alksne.linda@outlook.lv, anita.jansone@liepu.lv,
zane.berzkalne@rtu.lv`

Abstract. By investing time and resources in learning analytics, students, faculty, and administrators can improve learning and course outcomes, and develop more engaged and effective teaching and learning techniques. Learning analytics provides feedback to individual students and faculty, but can also surface trends across schools or other scopes of interest. This paper is dedicated to Learning analytics. Authors have analyzed logs from Universitat Politècnica de València video lecture recording and management system. UPV is using The Paella Player which is an Open Source Javascript video player of the Opencast Community. Logs have been analyzed with two different data and analytics reporting tools – Power Bi and Tableau from different point of views. Also the results from two different study years have been compared. Paper includes results from both business analytics tools, conclusions about the tools and also results that can be used for improve and benefit multiple audiences from university.

Key words: business intelligence tools, learning analytics, video lectures.

1. Introduction

Every time a student interacts with their university – is that going to the library, logging into their virtual learning environment or submitting assessments online – they leave behind a digital footprint. Learning analytics is the process of using this data to improve learning and teaching. Using the increased availability of big datasets around learner activity and digital footprints left by student activity in learning environments, learning analytics take us further than data currently available can. Videos, including those captured from traditional lectures, those recorded specifically by lecturers for online teaching, and other educational videos also are considered as learning analytics. The data can show to what extent students are engaging with the materials. For example one study has even suggested the optimum length for a video clip, based on the amount of time most users will watch a clip for before moving on (Lester, 2018).

In this paper we analyze results from video lecture watching logs retrieved with two different Business intelligence tools Power Bi and Tableau. Power BI and Tableau are the two front runners in the category if you search for a Business Intelligence or data visualization tool.

Power BI is a business analytics service by Microsoft. It aims to provide interactive visualizations and business intelligence capabilities with an interface simple enough for end users to create their own reports and dashboards (Microsoft, 2016).

Tableau is a product from Tableau Software. Tableau is a powerful and fastest growing data visualization tool used in the Business Intelligence Industry. It helps in simplifying raw data into the very easily understandable format (Guru99, 2019).

In this paper, we are interested in information that can be discovered from logs about students' action intensity and behavior while they have available online video recordings from all the lectures of the semester.

2. Related work

Advances in computation have created numerous opportunities for collecting and analyzing data about learners and their contexts. The ubiquity of online platforms, learning management systems, social media, digital student records, and networked devices has led to an increase in the capture of rich data trails that can provide insight into the learning process. The data generated in educational contexts is often large, complex, and heterogeneous, making it difficult to understand even with advanced data analysis capabilities. While it is now relatively easy to capture rich data about students and their behaviors, instructors, students, and school administrators typically have a difficult time processing and interpreting such data, as they have a limited understanding of necessary data mining and processing techniques.

If educators had access to effective visualizations of educational data, they could potentially use them to provide formative feedback, and to improve the instructional materials. Similarly, providing access to these tools for students could also promote the development of their metacognitive skills, and enable them to choose their own learning paths (U.S. Department of Education, 2016).

Generally, data analysis is done for two major purposes, understanding and prediction, which are relatively independent. Thus, a predictive model (e.g. probabilistic or based on a neural network) does not necessarily need to be understood by a human. Visual representations as an effective way to provide material for human's perception and reasoning play have a crucial role. However, when it is necessary to make sense of very large or complex data, purely visual methods are insufficient. (Andrienko, 2014)

The ever increasing availability of digital data brings unprecedented possibilities to analyze different educational facets. Learners, teachers, designers and managers have been mobilized to investigate how data can be used to support learning and teaching. This intense interest has given rise to tools and techniques in the research field of learning analytics (Baker and Inventado, 2014).

Advances in computational tools and methods have also created an opportunity to assess students' understanding of open-ended tasks such as engineering design or programming. Researchers and educators have used process data from students' interactions with educational software to characterize their design strategies. Likewise, using non-intrusive logged data from students' interactions with educational tools and platforms, researchers have aimed at identifying behavioral patterns of students that could lead to certain situations impacting student performance and retention (e.g., succeed in the course, drop out of the course) (Vieira et al., 2018)

Learning analytics approaches typically exploit data generated by users during normal interactions with e-learning technologies. Modern Learning analytics tools also

should consider the content of resources in addition to the interaction with them. In the next generation of personalized learning environments, it is essential to provide resources tailored to the learner's need while integrating interactions, skills and competencies with the mapping of knowledge of disciplines (Siemens, 2013).

Multimedia classification and indexing are two tasks required to organize and store resources so that they can be quickly retrieved (Dessi et al., 2018).

One thing that all of us usually do in lectures or watching video lectures is taking notes. This activity, however, is not well supported when learning from watching lecture videos. The conventional video interface does not allow users to quickly locate and annotate important content in the video as notes. To take notes, a user needs to identify important information like facts, numbers, and formulas while comprehending the constantly streaming information from video. Moreover, when writing notes, the user needs to deal with extraneous video control tasks such as pausing the video to avoid missing video content or rewinding to catch up with the video progress. That information we can get from logs (Piolat, 2005).

3. Methodology

Authors have analyzed data from two logs of UPV video lecture capturing system supplemented our quantitative findings.

First one of logs is from year 2016/2017 and the other one is from year 2018/2019, both from first semester (September - January).

There are two of registers actions and footprints in the logs. Actions are actions on the player, so they go after the user doing something. They are mostly understandable. Footprints are periodic information that the player sends with the part of the video viewed. Footprint events are not fired by the user, because user can close the window and then the log would lost all the last interaction, so there is logic behind that stores the data at time intervals. They go from inpoint to outpoint in seconds. Inpoint and outpoint are the number of seconds within the video in which the action happened. For instant events (like PLAY) both will be the same.

“Course_id” is the course (UPV changes semester beginning February) and “mediapackage” is the id of the video. The course id is the series_id. So the mediapackages (videos) belong to a course (series_id).

4. Results

The first data authors have analyzed is a video watching session, which represents a single instance of a student watching a particular video. Also average video watching session is defined which is average engagement time. It is used as the length of time that a student spends on a video (i.e., video watching session length) as the main proxy for engagement. Engagement time is a standard metric used by both free video providers such as YouTube and enterprise providers such as Wistia. However, its inherent limitation is that it cannot capture whether a watcher is actively paying attention to the video or just playing it in the background while multitasking (Guo et al., 2014).

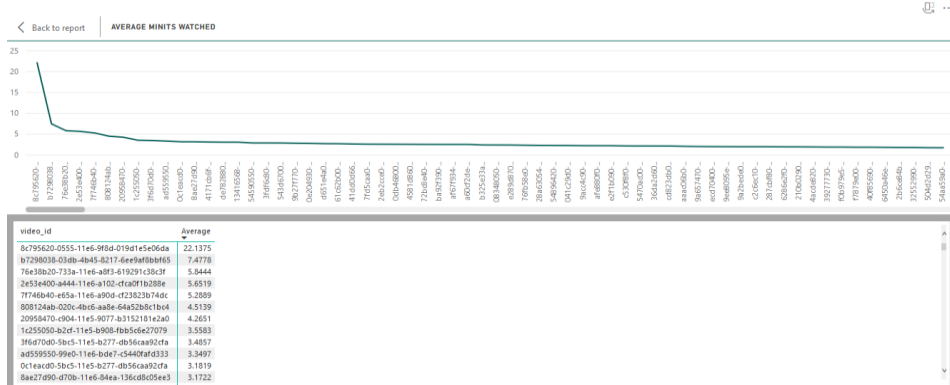


Figure 1. Average minutes watched (video_id)

As we can see in Fig.1 we can just look at the graph in Power BI and we can also use function to see data that stands behind the graph. We now have very good information that proves one of the rules how to produce video lectures to engage students and deliver the maximum amount of information (Alksne, 2016) Full classroom lecture content should be conveyed in 5-10 minute summary videos (Guo et al., 2014). To search for deeper connection, there must be more information about the video.

Fig.2 is a collection of visuals that show various aspects of the data authors have used to create the model in Power BI Desktop. A collection of visuals, in one Power BI Desktop file, is called a report. A report can have one or more pages, just like an Excel file can have one or more worksheets.

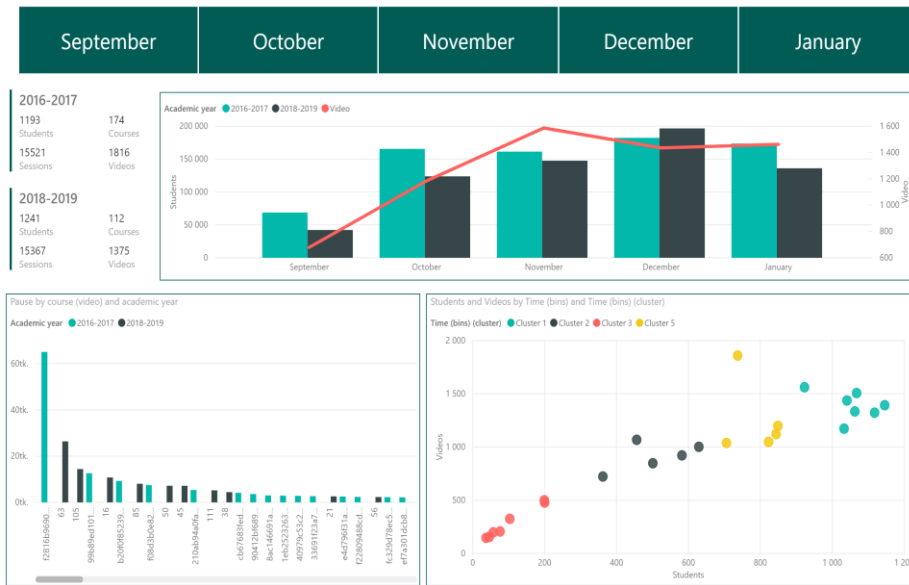


Figure 2. Power Bi report

The authors were able to find information about the time when students are watching video lectures – is it through the week or more on weekends. Is it equal through the semester or students are more active when exams come in the end of semester.

As mentioned above it was interesting to find out about pausing the video lecture – is there some connection between course id and the pausing frequency for taking notes for example.

PowerBI Desktop is an application that lets you connect to, transform, and visualize your data. With Power BI Desktop, you can connect to multiple different sources of data, and combine them (often called modeling) into a data model that lets you build visuals, and collections of visuals you can share as reports.

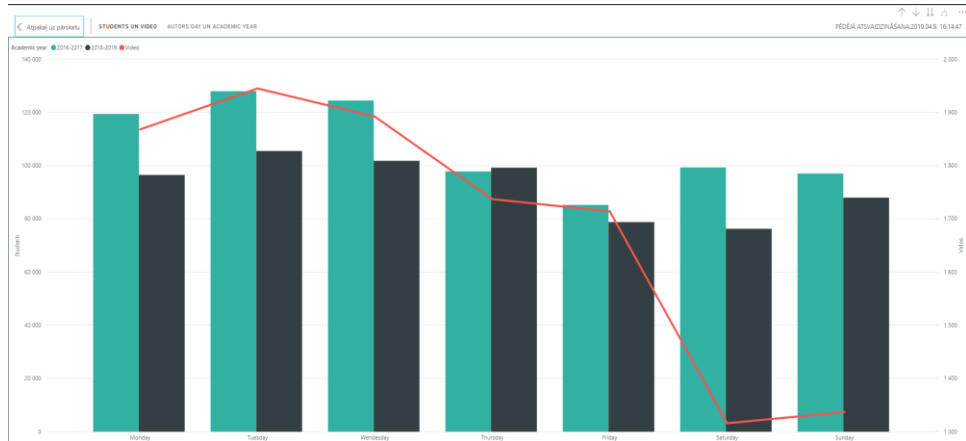


Figure 3. Videos watched through week days

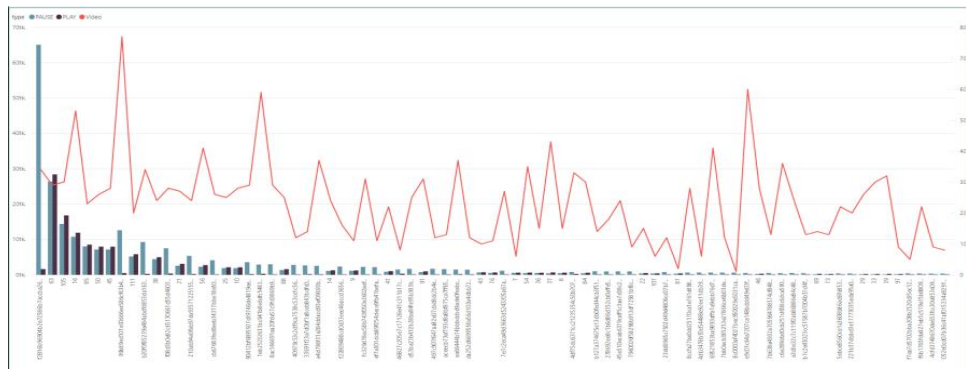


Figure 4. Pausing during the session vs times video has been watched

Working with Power BI Desktop is convenient and intuitive. It helps that there are video materials available and also information from forums, and also every month user receives information about all the updates and improvements. Authors liked that for data analysis with Power BI Desktop it is possible to use different sources during the one project and to combine data from them in the same time.

Analyzing data about video lecture logs there were two .csv files available, they were similar, but names of the columns were different in each file. But you do not have to change anything in the files, but you can do it all in Power BI, where it is possible to edit or customize all available data, to make one file to work with. Very useful and convenient seemed that while editing data with Edit Queries you can see all the applied steps in the panel. You can see all your moves in sequence. But when understanding that some action was unnecessary, it is possible to recall it without touching other actions. Power BI also allows making dimensions that helps to create the graph more with more details. Interactivity between data is also possible to switch off if interaction is not needed between graphs. From data that is available you can create new measures and if you need also make new columns in the table. When working from different work stations without connection through domain or VPN, than files that are used as data source must be in that computer which is used in that time and also Data source settings must be updated.

Overall available data from video lecture logs are suitable for analyzing with Power BI and all the different layer filters in Power BI very relieved preparing this report.

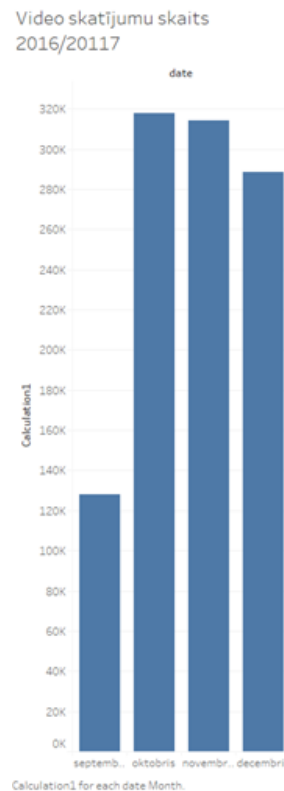


Figure 5. Video watched during the semester

Also graphs with finished work are very dynamic, good looking and modern. It is nice to go through the months in this case and watch how the picture is changing like flowing through the video. Also it is great that you can easily see which data has been used for every graph. Possibility to export pictures to pdf or power point also is very

useful. In Power BI it happens much faster than in Tableau. Also graphs in Tableau are more like Excel style. Not so dynamic.

In Fig.5 there is data about how students are watching lectures during semester. This graph is made with Tableau.

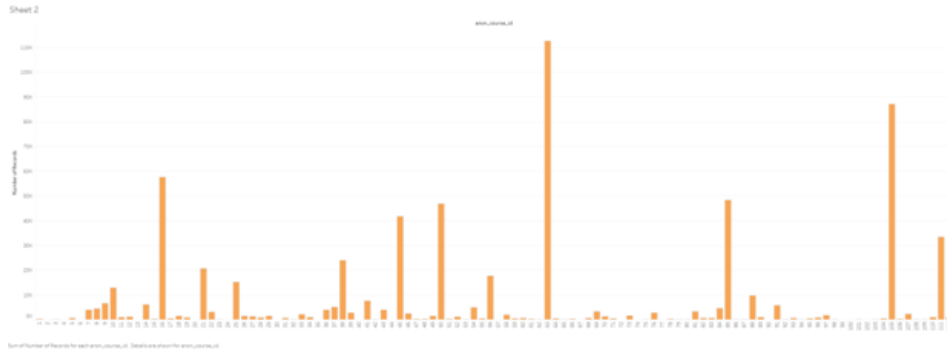


Figure 6. Graph made with Tableau. How many times each video has been watched

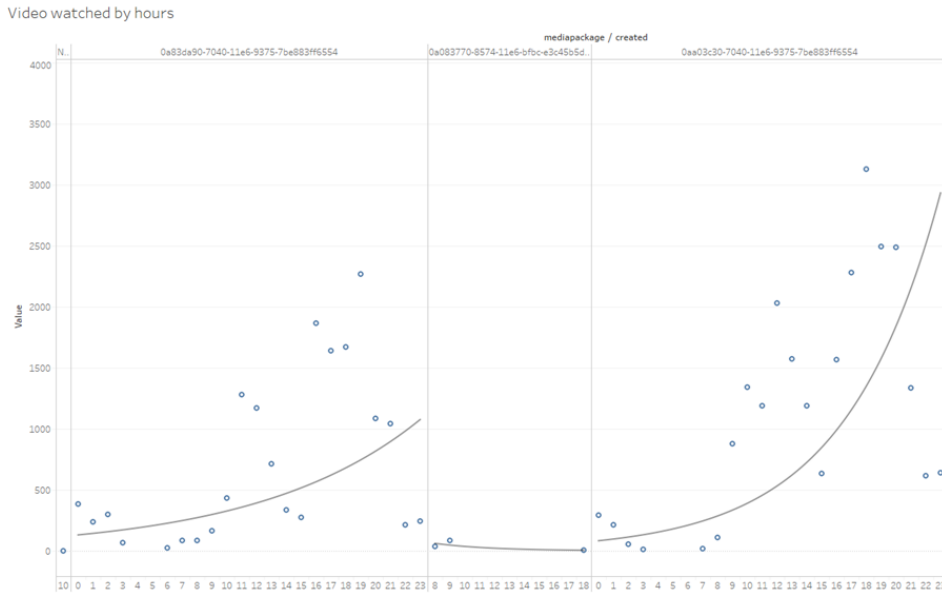


Figure 7. Videos watched during the day, by hours

We can see how many times each video has been watched in Fig.6. Tableau also gives easy platform where to choose data and change them from different views. This is fast way how to get information for university staff about most popular videos and subjects. It could be interesting to compare this information with how many students are in every class. Also for videos which have been watched outstanding many times it would be interesting to see the content of it. Is it very interesting is it very difficult or are in this subject many students.

And also Tableau offers different analyzing options – average lines, trend lines, cluster etc. Exponential trend line is used to see the trend in which hours of the day students watch videos. In Fig.7 there are three different videos analyzed. Which videos have been watched all day long – during studies and university and studies at home and which ones have been watched only few times maybe once during the lecture. University staff can estimate not only how requested or popular videos are but also how busy and loaded servers are during the day.

Fig.8 gives very interesting information about the activity of students – how many times students have been watched videos. For getting this information authors used such an option in Tableau – first to make calculations, in this case counted the videos, and then use this calculation for making the graph. It is very useful tool and easy to use, because you cannot only count but use very different functions for calculations.

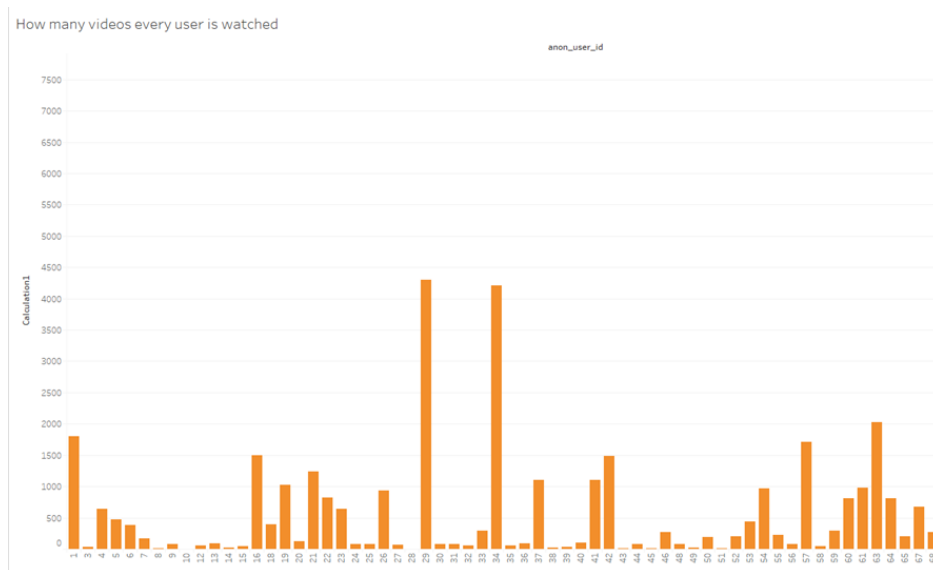


Figure 8. Number of videos watched by each student

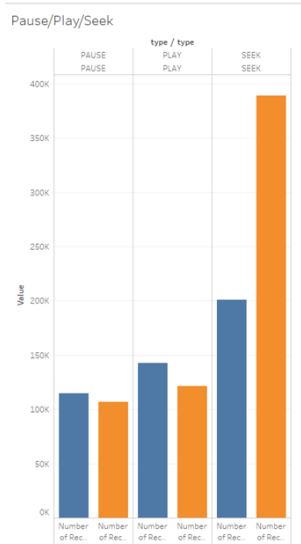


Figure 9. Pause, Play and Seek buttons

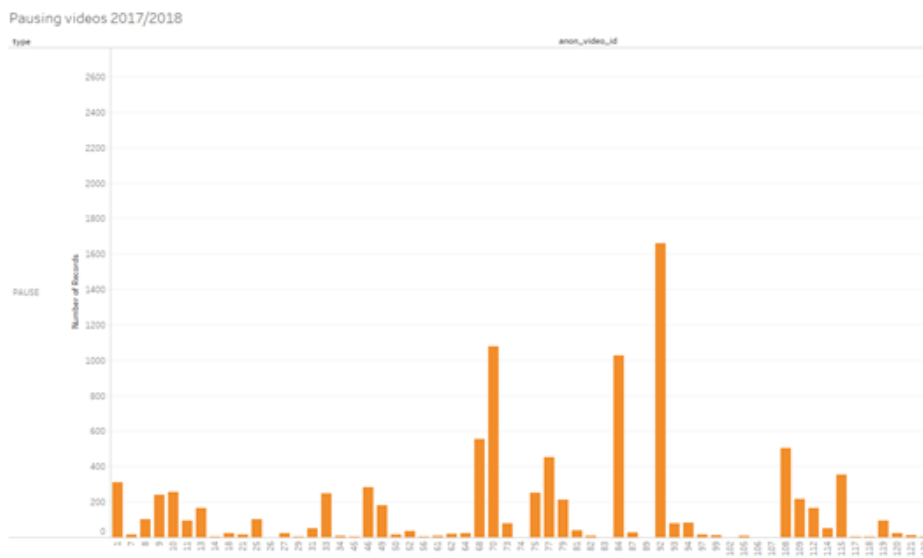


Figure 10. How many times pause has been pressed during each video – semester 2018/2019

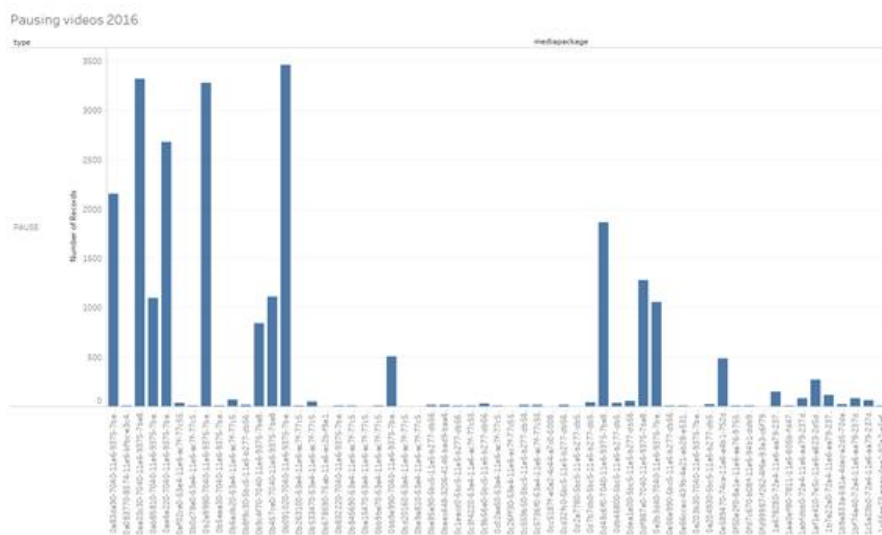


Fig.11. How many times pause has been pressed during each video – semester 2016/2017

Tableau is very convenient for filtering information straight to change the graphs. There is no need to filter data in excel file to make a graph. It really saves the time. In Fig.10 authors have put together both data sources from both semesters and filtered actions that have been done with videos – play, pause and seek. We can see that from last semester students have been more active in using search option while watching the video. Authors think that it is because technologies have developed also through these two years and searching gives us better result and also it is good information form teachers – if they have to improve keywords or bookmarks for their videos.

5. Conclusions

The main conclusions are:

1. Both data visualization tools – Power Bi and Tableau are suitable to process that kind of big data that has been collected from video lecture logs.
2. Theses interactive data visualization tools – Power Bi and Tableau - are intuitive and without preliminary knowledge it is possible to get difficult graphs, but with more training and knowledge graphs can be very rich with used filters and information can be very detailed.
3. Data from video lecture logs are anonymized but if we had data about subject and courses and teachers then benefit for University staff and

teachers themselves could be used for analyzing lectures, their work and figure out why students choose or do not choose the subject. And also if necessary find the result why students have or do not have good grades in subject.

4. From all these graphs most important information gets service maintainers because we could get very precise data about the workload of video lecture recording and management system.
5. In UPV it is not mandatory to watch video lectures, but demand and usage of video materials is huge.
6. Logs give data about student actions while watching the video lectures, information about seeking during the session is signal for teachers how important bookmarks and keywords are. It is also question for developers – is searching tool advanced enough.
7. When comparing booth tools – Power Bi and Tableau – first ones interface seems more modern and dynamic.
8. Data export from Power Bi works much faster than from Tableau.

References

- Alksne, L. (2016). How to produce video lectures to engage students and deliver the maximum amount of information, 2, 503–516.
- Andrienko, G. (2014). Visual Analytics Tools for Analysis of Movement Data, (December 2007). <https://doi.org/10.1145/1345448.1345455>
- Baker, R. S., Inventado, P. S. (2014). Educational data mining and learning analytics. In Learning analytics (pp. 61e75). Springer.
- Chen, C.-M., Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education*, 80, 108–121. <https://doi.org/10.1016/j.compedu.2014.08.015>
- Colvin, C., Rogers, T., Wade, A., Dawson, S., Gasevic, D., Buckingham Shum, S., Fisher, J. (2015). Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement. Australian Department of Education. Retrieved from http://he-analytics.com/wp-content/uploads/SP13-3249_-Master17Aug2015-web.pdf.
- Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Educational Technology & Society*, 15(3), 58–76.
- Greller, W., Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society*, 15(3), 42–57.
- Guo, P. J., Kim, J., Rubin, R. (2014). P41-Guo, 41–50. <https://doi.org/10.1145/2556325.2566239>
- Guru99. (2019) What is Tableau? Uses and Applications. <https://www.guru99.com/what-is-tableau.html>
- Microsoft. (May 2016). Bring your data to life with Microsoft Power BI. Microsoft.com
- Nguyen, C., Liu, F. (2016). Gaze-based Notetaking for Learning from Lecture Videos. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 2093–2097. <https://doi.org/10.1145/780542.780576>

- Piolat, A., and Boch, F. Note Taking and Learning: A Summary of Research. *The WAC Journal* (2005), 101–113.
- Ritsos, P. D., Roberts, J. C. (2014). Towards more visual analytics in learning analytics. *Proceedings of the 5th EuroVis workshop on visual analytics* (pp. 61–65). .
- Romero, C., Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12–27
- Siemens, G., Gasevic, D. (2012). Guest editorial-Learning and knowledge analytics. *Educational Technology & Society*, 15(3), 1–2.
- U.S. Department of Education (2016). Future ready Learning: Reimagining the role of technology in education - 2016 national education technology plan. Retrieved from <http://tech.ed.gov/files/2015/12/NETP16.pdf>
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500–1509.
- Vieira, C., Magana, A. J., Boutin, M. (2017a). Using pattern recognition techniques to analyze educational data. *Proceedings of Frontiers in education conference (FIE)*, 2017.

Received September 2, 2019, accepted September 17, 2019