WHAT IS THE SUCCESS OF HIGH SCHOOL IT EDUCATION? AN INVESTIGATION INTO HIGHER EDUCATION STUDENTS' KNOWLEDGE OF SPREADSHEET APPLICATIONS

[1]**Bakó M.**, [2]**Ráthonyi G.**

University of Debrecen
[1,2]Department of Agricultural Informatics 4032 Debrecen, Böszörményi 138.
bako.maria@econ.unideb.hu
rathonyi.gergely@econ.unideb.hu

Abstract: *Teaching Computer Science in higher education is imperative, even though today's students have been born into a world where technology is an essential part of everyday life. To efficiently master modern, business, technical and scientific knowledge and to proficiently produce quality results in a work environment it is crucial to have high level IT knowledge. In business, Excel is the lingua franca and so knowing how to aptly use spreadsheets is a must for our students.*
The primary objective of the authors was to examine the perceived and actual knowledge of spreadsheet applications of students entering higher education. Accordingly, a questionnaire and a practical assignment have been developed. In the questionnaire, students were asked to provide information concerning their previous IT studies and rate their knowledge of word processing, spreadsheets and database management. During the practical, students were asked to solve an Excel exercise taken from a high school Computer Science final exam at standard level. Out of the 666 registered students on our electronic education system (Neptun) at the beginning of the year, 557 took part in this survey, and following data cleansing and processing, 513 were considered in the results.
Looking at the results of the practical, the most significant proportion of students, 142 of them have performed between 0-10%. A total of 260 students have achieved less than 20% performance and 434 people, which is nearly 85% of students have accomplished less than 40%. Compared to the results from the self-evaluation questionnaire it is very poignant that the actual scores differ quite significantly (in both directions) from the perceived knowledge of the students.

INTRODUCTION

Students today — starting from kindergarten all the way up to higher education — can be considered to be part of generation Z. They were born into a world where computers, mobile phones and video games are an integral part of everyday life as we see it. The English literature has a more technical term to describe this phenomena, these people are digital natives, who speak the language of computers, video games and the digital language of the internet just as well as their mother tongue. Hence it begs to ask the question, do these students need a Computer Science education, can we teach them anything new, or does their intuitive knowledge cover the programs and applications that the job market could ask of them? Looking at job opportunities available for our students doing Economics and Agricultural Business degrees, we see that at least intermediate level computing skills as well as the usage of application softwares, creation of statistics and

graphics is required almost everywhere. Based on the 2012 National Curriculum we should expect students to possess this knowledge by the end of high school. To clearly document this, we have prepared a survey for the first year students at the Faculty of Agricultural and Food Sciences and Environmental Management and the Faculty of Economics and Business of the University of Debrecen and asked them to complete a self-evaluation on how well they think high school has taught them to use application softwares. We wanted to compare the self-evaluations against the actual knowledge of our student. In business, Excel is the lingua franca, so being proficient with spreadsheets is essential for our students. Hence we have asked our first years to complete the Excel questions from the high school Computer Science final exam at standard level of October 2015. In this article we will present the results of this survey, which is a representative sample of over 500 students from different degrees.

DIGITAL COMPETENCY AND THE 2012 NATIONAL CURRICULUM

The members of generation Z were born into the world of technology, so they do not know what the world was like without internet. For them, computers, smart phones and social networks are no longer just part of everyday life, but have become a necessity. At first, the previous generations have only wanted to exploit the potentials of Web 2.0, nowadays they use them as tools, but generation Z practically lives inside of it (PÁL and TÖRŐCSIK 2013).

In 2005, an American research showed that people between the age of 8-18 spend an average of 6.5 hours a day consuming different electronic mediums. When the researchers have started investigating the specifics of these results they have realised that this time actually counts for 8.5 hours. This 2 hours difference can be attributed to the fact that most teens multitask. What does it exactly look like, when we see our watching television whilst also scrolling on their phone? Research has shown that young people do not only surf the internet, but listen to music, chat to their friends whilst also doing their homework and running file sharing programs in the background. (WALLIS, 2006) Unfortunately, not enough time has yet passed to draw a conclusion on how this will affect learning patterns and abilities, relationships, creative outlets or simply our understanding of the world around us. (SÁGVÁRI, 2011).

The question asked in the introduction, do these students need a Computer Science education? has also been debated by Gabriella Baksa-Haskó in her article "Preliminary IT knowledge of Economics students". She believes the answer is yes, as it is similar to being fluent in our mother-tongue not excusing us from having to learn correct grammar. We take this idea further. First year primary school students are first taught how to read and write, and later comprehensive reading is introduced. Many children start primary school by already knowing how to fluently read and even how to write to an extent. In these cases, the teacher is responsible for differentiating between competency levels, making sure that all students are kept from being bored, and that children

develop the necessary skills for text comprehension, since being able to read a text does not necessarily imply also understanding it. We believe that Computer Science is quite similar since here we are looking to extend on digital competency.

The question naturally arises, how do we understand digital competency and literacy? These two definitions have changed and evolved over the years as a result of technology advancing and expanding. (BAWDEN, 2001; BAWDEN, 2008; KARVALICS, 2012; VÁRALLYAI and HERDON 2011).

In the 2006 recommendation of the European Parliament and Council (EPC 2006/962/EC) we can find a detailed description of those skills and attitudes that are essential to the development of digital competency. „Digital competence requires a sound understanding and knowledge of the nature, role and opportunities of IST in everyday contexts: in personal and social life as well as at work. This includes main computer applications such as word processing, spreadsheets, databases, information storage and management, and an understanding of the opportunities and potential risks of the Internet and communication via electronic media (e-mail, network tools) for work, leisure, information sharing and collaborative networking, learning and research. Individuals should also understand how IST can support creativity and innovation, and be aware of issues around the validity and reliability of information available and of the legal and ethical principles involved in the interactive use of IST." (EPC 2006).

In Hungary, developing digital competency, knowledge of application softwares at intermediate level, the acquisition, interpretation and consumption of data and active participation in electronic communication were considered an integral part of the 2012 National Curriculum for Computer Science. In 2017, an IT Communication Uniform Reference Framework (IKER) was established in Hungary, which fits the European Digital Competence Framework and identifies the following areas of digital competency:

- Collection, use and storage of data,
- Digital, web-based communication,
- Creating digital content,
- Problem solving, practical applications,
- IT and communication technology security

Furthermore, a digitally literate individual is competent at securely using IT and communication technology, as well as search, distribute and create digital data. (IKER, 2017)

Since in business it is indispensable to be proficient in Excel, we have reviewed in detail the relevant parts of the 2012 National Curriculum Framework for Computer Science that are still in effect today. We have found that spreadsheet applications are taught in Year 10.

The Curriculum Framework places great emphasis on the students learning how to use spreadsheets to solve everyday problems and tasks that may arise in other subjects. By the end of Year 10, students should be familiar with a range of functions, financial, statistical, mathematical calculations, and are expected to solve questions on investments and loan related problems. Additionally, they are presumed to be familiar with sorting, filtering and visualisation of data.

Our years of educational experience has shown that our students are not in possession of all these competencies when entering university. To clarify this, we have designed a survey for the first year students at the Faculty of Agricultural and Food Sciences and Environmental Management and the Faculty of Economics and Business of the University of Debrecen. In this survey, students had to evaluate their level of knowledge regarding application softwares and then we have asked them to complete a series of Excel questions as part of a practical.

MATERIAL AND METHODS

In this survey, those first year students were considered from the Faculty of Agricultural and Food Sciences and Environmental Management and the Faculty of Economics and Business of the University of Debrecen who study Computer Science in their first semester. The questionnaires and the practicals were completed in the first week of the semester. Out of the 666 registered students on our electronic education system (Neptun) at the beginning of the year, 557 took part in this survey, and following data cleansing and processing, 513 were considered in the results.

This survey consisted of a paper based questionnaire and an Excel practical. The primary objective of the questionnaire was to assess the students' IT knowledge. We wanted to apperceive how many hours per week students spent on studying Computer Science in high school, whether they chose it as a subject in their final exams and if they had an ECDL qualification. We also put emphasis on learning how students rated their own knowledge in the different areas of Computer Science. For every area students had to self-evaluate how confident they felt about their abilities. We included 3 text editing, 8 spreadsheet and 4 database management categories, as we believe that these three areas are the most important and most commonly taught in higher education.

When defining each categories of interest, we aimed to create groups based on the exercise types present in Computer Science final exams such that when evaluating the results, we could efficiently compare the students' estimated knowledge to their actual knowledge.

From the above survey we only present those 5 questions on spreadsheet applications in this article that are related to the questions from the high school Computer Science final exam at standard level from 2015 October.

After completing the questionnaire we asked our students to solve some Excel questions to the best of their abilities. Students had 60 minutes to attempt these questions, and everyone has spent at least 20 minutes on the practical. The practical assignment was made up from the Excel questions in high school Computer Science final exam at standard level from October 2015. We wanted to test how students could carry out different operations and functions in Excel, so we slightly modified some of the exercises. Firstly, students were not given the data in TXT but in Excel format. Whilst this is a simple transformation to make, we wanted to eliminate the possibility of a student not answering any of the questions due

to not being able to read in the data. Furthermore, since we wanted to test whether student could correctly apply Excel functions, to avoid reading comprehensive problems, in places we have inserted points of explanations.

For example in the practical given we had a the following question: The prices of gasoline and diesel do not always increase and decrease simultaneously. Using a formula, write out into the column titled 'asynchronously changed' the text "asynchronous" in those cases when the price of one increased and the other decreased (i.e. when their product is <0)! Otherwise, nothing should appear in the cell!

In the original question "i.e. when their product is <0" was not included. Here we have tried to avoid any student failing to use the IF function and avoid answering this question due to not being able to formulate a condition. We have made four such refinements.

We extended the practical with the following question: In columns G and H, calculate the price of gasoline and diesel in Euro. When creating the formula, use the cell reference L1! With the help of this question, we wanted to investigate if students know what an absolute reference is, and if they could use it.

We wanted to make sure that every student could solve the practical to the best of their knowledge so we drew their attention to the formatting tasks at the end of the practical, which are in nature very similar to Word formatting. Hence, if they have not used Excel before, they could have a look at these questions and possibly still solve some of them. We have also offered that those students who achieved a score above 80% could endorse their results as part of their first assignment grade.

When marking the practical, we have considered the original marking guide. By solving the practical perfectly, a total of 30 marks could be obtained.

RESULT AND DISCUSSION

The paper based questionnaire had questions for 8 categories on spreadsheet applications from which 5 were covered in the practical. We discuss these 5 below. When processing the collected data, we compared the self-evaluation grades with the practical results for each student.

In the first category, Spreadsheet Management 1 we focused on formatting and performing basic operations. We asked our students to estimate their level of knowledge of these.

In the Excel practical, there were a maximum of 9 marks that could be gained in this category. 2 marks for correctly performing two subtractions, 2 marks for a division and converting the result to percentage, 1 mark for setting up the formatting HUF/litre, 1 mark for indentation and 3 marks for border, fill and font style.

On Figure 1, the x-axis shows the estimated knowledge of students (on a scale of 1 (none at all) to 5 (confident)) and the y-axis their actual results in percentage, that is, how many marks they got in this category out of 9. The size of the circles is simultaneous to the number inside of it, it shows the number

of students falling into the same category. More precisely, it shows for example the number of students who predicted their knowledge to be 3, but actually achieved 0%. We opted for this type of chart since this way we do not have to group people into categories such as 80%-100% as that would lead to loss of data. This way, for any given prediction we can show the exact number of students who achieved the same marks as they estimated they would.
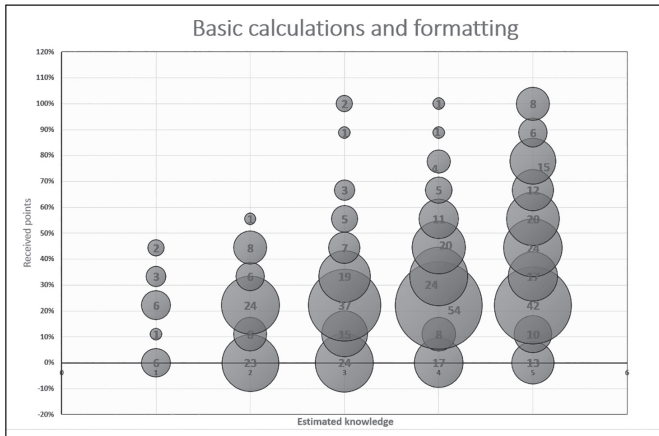

Figure 1: Basic calculations and formatting in the dimension of estimated knowledge and received marks
Source: own composition

If everyone actually possessed the same knowledge they assumed they did, the circles should line up on the diagonal. Since in this question we have uniquely asked about formatting and basic operations, it is incredibly clear that there were barely 18 students out of 513 who self-assessed to a knowledge of level 1, but only 6 of them achieved 0%. Those who assessed themselves to level 2 or 3 should have been able to comfortably solve this problem, but still 23, and 24 students (respectively) achieved 0%. It is surprising to see that there were 30 individuals who rated themselves level 4 or 5, but still showed 0% performance. The biggest circles for any level can be found around 22%, meaning that our students gained 2 out of 9 marks. Out of 513, only 11 students could accomplish 100% on this question, and out of these 8 rated themselves level 5.

If we look at the solutions, we must draw the conclusion that most of our students have opted to solve the easiest, subtraction question. Many have tried setting the border, fill and font style, but in most cases their final results were very far from the requirements and so only a small portion of the marks could be awarded.

Spreadsheet Management 2 focused on basic functions, absolute and relative referencing. The question's relevant material contained a AVERAGE, a MAX and a MIN function and a question focusing on absolute referencing. There were 5 marks in total for this category. As Figure 2 shows, the actual results in this category are closer the the students' self-evaluated knowledge than in the previous one. Alongside the large circles at 0%, there are also big ones around 60%. Most of these students could calculate the functions but could not handle the absolute reference. The row corresponding to 100%

also has prominent circles, and we notice that there is one in every column. On this figure, it is easy to differentiate the rows 0%, 60% and 100%, meaning that either the student was not able to solve anything in this category, maybe was successful in drawing up the functions, or handled the whole category perfectly. Like before, we meet some extremes, 3 students who rated themselves as level 1 but achieved 100% and 10 students who rated themselves as level 5 and achieved 0%.

These results further highlight the fact that students' evaluation of themselves differs quite significantly of their actual knowledge. It is important to note, that self-declaration questionnaires may show distorted results as we are prone to over or underestimate our own performance.
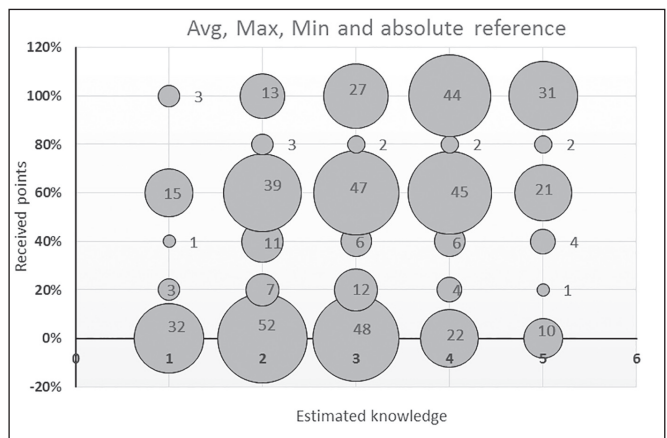

Figure 2: Basic functions and absolute references in the dimension of estimated knowledge and received marks
Source: own composition

Spreadsheet Management 3 of the questionnaire focused on conditional functions, whilst Spreadsheet Management 4 focused on the creation of charts.

In the practical assignment the functions IF and COUNTIF were presented, and a more complex bar diagram had to be made on a separate worksheet with a title, legend and the vertical axis partitioned from 300 to 500 along with the axis titles.
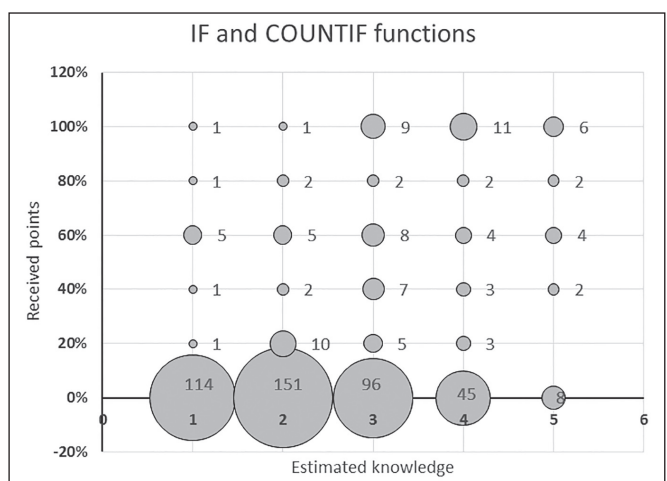

Figure 3: If and Countif functions in the dimension of restimated knowledge and received marks
Source: own composition

We consider these two categories together since we can see from the corresponding Figures 3 and 4 that the students' results are very similar in both cases. For both categories a maximum number of 5 marks culd be achieved. The figures clearly show that very few could solve the questions for these categories. Big circles are only present on the x-axis, meaning that they have achieved 0% on the practical. Apart from one category (IF, COUNTIF, (5, 20%)) there is a couple of students (between 1-14 individuals) who completed the practical in accordance to their self-evaluation. Again, it would have been the most optimal if the circles took place on the diagonal, and so for example students who rated their knowledge level 3 should have achieved around 60%. On these figures it is significantly visible how the estimated knowledge of students is not in accordance with the actual results.
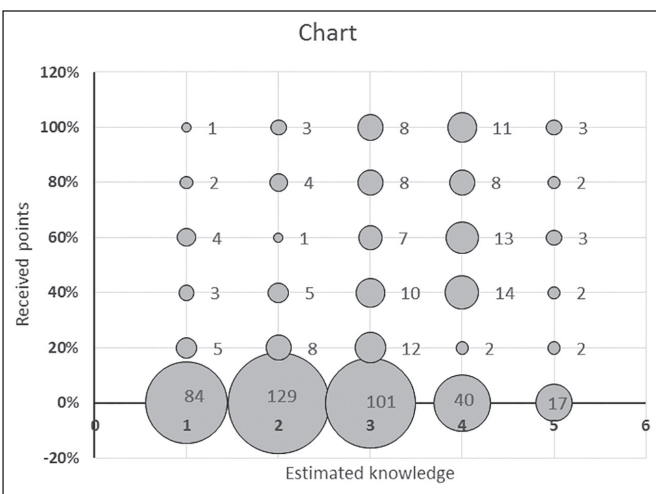

Figure 4: Charts in the dimension of estimated knowledge and received marks
Source: own composition

Spreadsheet Management 5 focused on search functions and nesting functions. A total of 6 marks could be obtained, from which 2 were awarded for successfully putting the ROUND function around an AVERAGE function, and 4 marks for correctly using the pair of INDEX and MATCH.
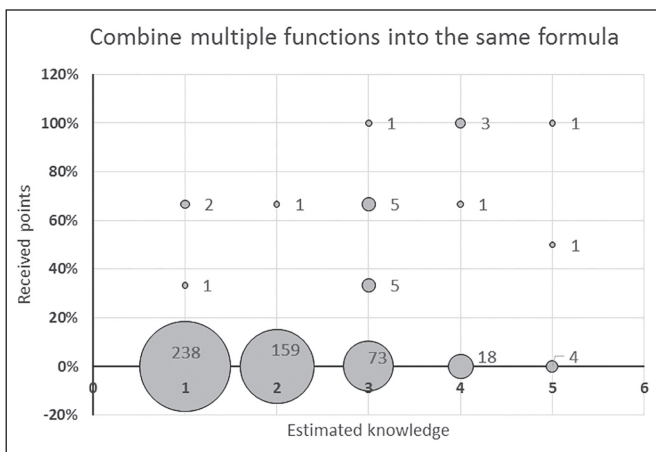

Figure 5: Combining multiple functions in the dimension of estimated knowledge and received marks
Source: own composition

On Figure 5 we can see that only 5 students out of 513 managed to complete this section to 100%. From this, one person rated their knowledge to level 3 and another to level 5, and 3 of them to level 4. 2 students used an auxiliary cell to solve the question on the pair of INDEX and MATCH functions, whilst 3 students managed to get the correct answer without one.

Again, on this figure the biggest circles take place along the x-axis, altogether 492 students achieved 0%, but only less than half thought that they had no knowledge on this category.

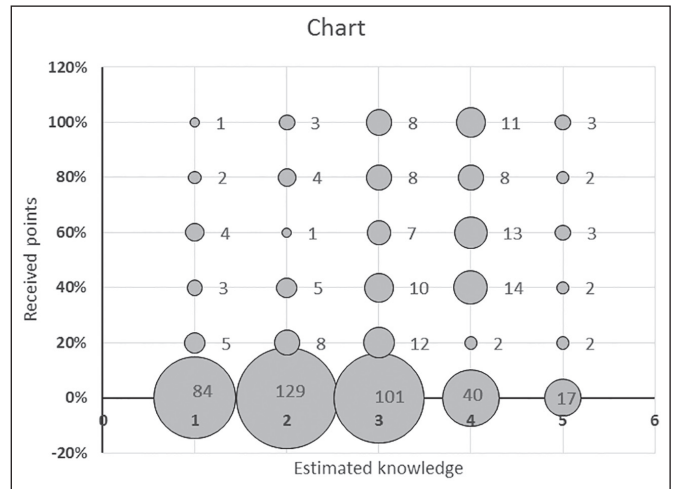Figure 6 contains a summary of all our results.

Figure 6: Aggregated results in the dimension of estimated knowledge and received marks
Source: own composition

The average presumed knowledge of our students is shown as a function of their actual results, grouped into categories. Regardless to the estimated knowledge the largest circles are around the x-axis. The most prominent results are made up of the 142 students accomplishing between 0% and 10%. Altogether 434 individuals, that is almost 85% of the students participating in this survey achieved below 40%, from which 260 students were below 20%. To get 40%, 12 marks were necessary, which could have been easily earned by doing the subtractions, formatting problems and correctly using the AVG, MAX, MIN functions. We must note, that there was only one student who completed the practical to 100%, and he estimated his knowledge to (3;4;3;4;3). We found three surveys where the students have estimated their knowledge to level 5 in each category. We have investigated their results, and reached the conclusion that their marks greatly differ. There was a 15%, a 33% two 60% and one 97% solution, the last losing a mark by missing an essential component of a chart.
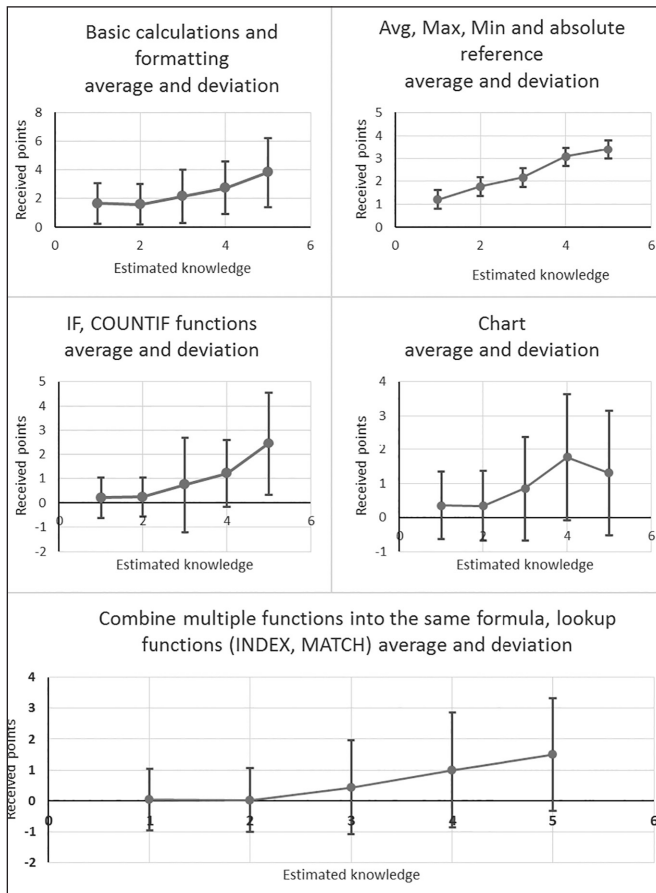
*Figure 7: Averages and deviations of the different assignments in the
dimension of estimated knowledge and received marks
Source: own composition*

On Figure 7 we summarise the means and standard deviations, that is, we calculated what was the actual results and and how large was the standard deviation for every estimated level of knowledge for each category. In the first two, those who estimated themselves to be of level 1 achieved higher results on average. It is interesting to note that their average is slightly higher than those who rated themselves level 2. Furthermore, we see similar results for the charting category, since those who estimated themselves to be level 4 reached higher results (1.77) than those who self-evaluated to level 5 (1.31). The standard deviation is very high in all cases part from the first category of basic functions.

CONCLUSION

In this research, which consisted of a paper based questionnaire and an Excel practical, authors examined the spreadsheet knowledge of first year students at the Faculty of Agricultural and Food Sciences and Environmental Management and the Faculty of Economics and Business of the University of Debrecen in 2018/2019 first semester.

We tried to answer important questions, such as can a student be considered well-versed in computers if they can use their smartphones, play games, chat online, or surf the internet or is it necessary to have high school/university students learn how to operate a spreadsheet?

Our results shows that students have inadequate knowledge about MS Excel, which highlight the relevance that we should teach spreadsheet management for our students. Altogether 434 individuals, that is almost 85% of the students participating in this survey achieved below 40%, from which 260 students were below 20% and 142 students accomplishing between 0% and 10%, which means that almost every third students' knowledge are under the basic.

Our results further highlight the fact that students' evaluation of themselves differs quite significantly of their actual knowledge. It is important to note, that self-declaration questionnaires may show distorted results as we are prone to over or underestimate our own performance. We found similar results in connection with the presumed knowledge comparing with other Hungarian surveys (SÜVEGES and SZABÓ 2013; BAKSA-HASKÓ, 2017).

Besides our results, international researches (OKE, 2004, LEBLANK and GROSSMAN, 2008) also emphasis the importance of spreadsheet management (Microsoft Excel). FORMBY et al., (2017) highlight that students' spreadsheet knowledge is essential in order for the likelihood of success in the job market. Therefore it is important to transfer the knowledge to our students which help them to become productive and economically sustainable members of society.

In order to meet the market expectation and to provide a higher level of teaching our students should be at standard level of high school Computer Science final exam when get into higher education. In order to reach this level of knowledge it is essential to place great emphasis on the students learning at high school level to solve everyday problems and tasks that may arise in other subjects. In this case, in higher education, the knowledge of students could be expanded with additional marketable and competitive knowledge that meets today's expectations. Competitive knowledge can be for example business analytics and big data trends and according to DUMBILL, (2012) MS Excel is still the ubiquitous and popular choice for data analysts.

Examining the 2012 National Curriculum we have found that for high school students' one Informatics class is compulsory per week for two years. During this time, there is 30 hours dedicated to learning the programs in Microsoft Office. Therefore, there is around 16 hours devoted to spreadsheet applications such as Excel. So our students are expected to familiarise themselves with spreadsheet applications through 16 classes, with a class a week so that they are comfortable with a range of functions, financial, statistical, mathematical calculations, and are expected to solve questions on investments and loan related problems. Additionally, they are presumed to be familiar with sorting, filtering and visualisation of data.

To learn, practice and truly deepen this knowledge, so that students gain a lasting knowledge of the subject in 16 hours could only be achieved with an extensive amount of homework. However, the probability that a 16 year old student will spend their limited free time practicing spreadsheet applications is negligible. Naturally, this work could be assigned in the form of graded practicals, but such an amount of homework that

would be needed for acquiring this deep knowledge cannot be made mandatory.

Therefore, an increase in the number of weekly classes from one to at least two would be an effective solution, though for achieving real results a minimum of three classes per week would be needed.

Furthermore, the use of spreadsheet applications could be enhanced by introducing them as a tool for computing calculations, creating statistics and visualising data on figures for other subjects.

Examining the National Curriculum also suggests that the use of computers, or more specifically spreadsheet applications is considered to be essential to the development of scientific competencies. However, looking at high school textbooks we saw that they do not incorporate any information on how to solve problems using spreadsheet applications.

Looking at scientific subjects one by one, we can deduce that spreadsheet applications could be used excellently for the education of Mathematics and Physics. In Mathematics, we use real sets of data to create statistics, and in Physics we deduce approximations from the results of tests and measurements.

Analysing Biology, Geography and Chemistry textbooks we find that spreadsheet applications could be mainly used for summarising data and illustration purposes. In these subjects, we can use real statistical results to demonstrate the magnitude and gravity of existing environmental problems. By letting students process and illustrate the data themselves, they become active participants of the class, rather than staying passive listeners of the teacher's observations.

Last but not least, a differentiated range of tasks and solutions alongside the demonstrated use of the application helps students familiarise themselves with the deep learning required to achieve lasting knowledge in a range of IT tools and programmes.

ACKNOWLEDGEMENT

REFERENCES

Abonyi-Tóth A., Turcsányi-Szabó M. (2015): A digitális írástudás fejlesztésének lehetőségei. Pátria Nyomda Zrt. 184. p.

Baksa-Haskó G. (2017): A közgazdászhallgatók informatikai előismeretei, Vezetéstudomány/Budapest Management Review; 2017, 48(8-9): 101-111.

Bawden D. (2001): Information and digital literacies: a review of concepts. Journal of Documentation; 2001, 57(2): 218-259.

Bawden D. (2008): Origins and concepts of digital literacy. In: Lankshear, Colin – Knobel, Michelle (ed.): Digital Literacies: concepts, policies and practices. New York. Peter Lang. pp. 17–32.

Dumbill E.: (2012): Planning for big data. O'Reilly Media, Inc. 83. p.

EPT: Az Európai Parlament és a Tanács ajánlása, (2006. december 18.), Az egész életen át tartó tanuláshoz szükséges kulcskompetenciákról, (2006/962/EC)

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32006H0962

Formby S.K., Medlin B.D., Ellington V. (2017): Microsoft Excel®: Is It An Important Job Skill for College Graduates? Information Systems Education Journal (ISEDJ); 2017, 15(3): 56-63.

IKER: http://progress.hu/wp-content/uploads/2017/03/Praktikus-utmutato-az-IKER-referenciakeret-hasznalatahoz.pdf

Karvalics L.Z. (2012): Információs kultúra, információs műveltség – egy fogalomcsalád értelme, terjedelme, tipológiája és története. In: Információs társadalom; 2012, 12(1): 7-43.

LeBlanc L.J., Grossman T.A. (2008): Introduction: The Use of Spreadsheet Software in the Application of Management Science and Operations Research. Interfaces 38(4): 225-227.

Oke S.A. (2004). Spreadsheet Applications in Engineering Education: A Review. The International Journal of Engineering Education. 20(6) 893-901.

Pál E., Törőcsik M. (2013): Irodalmi áttekintés a Z generációról. Pécsi Tudományegyetem 19.p.

Prensky M. (2001): Digital Natives, Digital Immigrants. On the Horizon NCB University Press; 2001, 9(5): 1-6.

Ságvári B. (2008): Az IT generáció, technológia a mindennapokban: kommunikáció, játék és alkotás. Új ifjúsági szemle: ifjúságelméleti folyóirat; 2008, 6(4): 47-56.

Süveges P., Szabó I.E. (2013): Felmérés a BME-re 2012-ben felvett hallgatók pályaválasztásáról és szociális helyzetéről. Budapesti Műszaki és Gazdaságtudományi Egyetem Diákközpont. 42. p.

Várallyai L., Herdon M. (2011): E-Competences and ICT in Europe. In: Herdon M., Rózsa T., Szilágyi R. (szerk.) Agrárinformatika 2011 Konferencia / Agricultural Informatics Conference: Innovativ információtechnológiák az agrárgazdaságban / Innovative information technologies in agriculture, Debrecen, Magyarország : Magyar Agrárinformatikai Szövetség (MAGISZ), pp. 101-108.

Wallis C. (2006): The Multitasking Generation; They're e-mailing, IMing and downloading while writing the history essay. What is all that digital juggling doing to kids' brains and their family life? Accessed date: October 2018. http://content.time.com/time/magazine/article/0,9171,1174696,00.htm