

# REPORT ON PATTERNS IN INCLUSION AND ATTRITION

Final version

30 08 2023



Co-funded by the  
Erasmus+ Programme  
of the European Union

This project has received funding from the European Union's Erasmus+ Programme under grant agreement No. 101004062. The content of this document represents the views of the author only and is his/her sole responsibility. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains.

## TABLE OF CONTENT

<b>1</b>	<b>INTRODUCTION .....</b>	<b>3</b>
<b>2</b>	<b>TENDENCIES IN FINDINGS ON STUDENT ATTRITION .....</b>	<b>3</b>
2.1	Variation in attrition between study fields and study programmes .....	4
2.2	First year and bachelor students .....	4
2.3	Study performance before university .....	5
2.4	Motivation for applying to university .....	5
2.5	Study performance at university .....	5
2.6	First attempts to target students at risk of dropping out .....	6
2.7	Summary .....	6
<b>3</b>	<b>PRELIMINARY DATA ANALYSIS ON INCLUSION AND COMPOSITION OF THE STUDENT MASS .....</b>	<b>7</b>
<b>4</b>	<b>SUMMARY AND CONCLUDING REMARKS.....</b>	<b>8</b>

*The content of this report represents the views of the author only and is his/her sole responsibility. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains.*

# 1 INTRODUCTION

This paper is the final report from the Circle U. working group on Learning Analytics. The content of the paper is based on input from administrative staff working with analysis and reporting on student and educational data from the University of Pisa, King's College London, and Aarhus University. This means that the content is limited to results from analysis made by the universities' administrative staff and not research projects.

Prior to writing the present paper, the working group published a White Paper on Legal and Ethical aspects of using Learning Analytics in university administration, which is available on the Circle U. website<sup>1</sup>. The working group also held a workshop on the topic at the Circle U. Week for the Future of Higher Education and Research, which was hosted and organized by Université Paris Cité and took place from 23 to 27 January 2023.

The content of this paper is based on discussions and the sharing of experiences at online meetings within the working groups and on written input from the three universities in the working group. The task given was to make a report on patterns in inclusion and attrition, i.e. findings on patterns on inclusion and attrition based on the preliminary use of Learning Analytics. The working group has worked with a broad understanding of the concept Learning Analytics and has therefore included examples of descriptive, explanatory, and predictive analysis.

Descriptive analytics is the analysis of historical data that answers the question "What happened?". Explanatory analysis seeks to find explanations and to answer the question "Why did it happen?", and lastly predictive analytics uses existing data to predict what is likely to happen in the future, thus answering the question "What might happen?".

We will not aim to make conclusions about general patterns in attrition and inclusion at universities across countries in this paper, as the report is only based on input from three different universities. Nor will we make final recommendations on how universities can ensure inclusion and reduce dropout. Instead, we will present and share for a broader audience which tendencies in attrition have been identified by the universities in the working group – and where these tendencies are similar and different. We will also present a preliminary analysis made on inclusion.

To conclude that the tendencies and correlations described in this paper are general patterns across universities and countries would demand a larger dataset from various universities. Therefore, we encourage readers to take note of the preliminary findings in this paper, collect more wide-ranging data and try to make general and final conclusions that can lead to identifying general ways to reduce attrition and promote inclusion at universities.

The discussions in the working group prior to making this paper made it clear that the participating universities had focused more on analysing patterns in attrition than in inclusion. Therefore, the following paper will focus mostly on describing patterns in attrition. The final section will describe how the universities are beginning to focus on analysing inclusion and the composition of the student mass.

## 2 TENDENCIES IN FINDINGS ON STUDENT ATTRITION

In the following sections of the paper, we present the trends in the findings on attrition that the different universities in the working group have identified by conducting data analysis on student and educational data.

Before describing the trends in the analysis, it is worth defining what the term student attrition means. Student attrition, sometimes referred to as student dropout, is when students withdraw from university (or further education) completely or when students withdraw from a specific educational programme and continue on

another educational programme. The underlying premise of conducting analysis and interventions regarding attrition or dropout is that the final purpose is to reduce it, especially if we consider dropout from further education. However, it is worth mentioning that dropping out can be positive for the individual student if the student continues on an educational path that suits them better.

As mentioned in the introduction, Learning Analytics can consist of descriptive, explanatory, and predictive analyses. As the universities in the group have more experience with the first two kinds of analysis, these will dominate the following section, but the universities' preliminary use of predictive Learning Analytics on attrition will also be presented.

## **2.1 VARIATION IN ATTRITION BETWEEN STUDY FIELDS AND STUDY PROGRAMMES**

First, all three universities see differences in dropout rates between different study fields and study programmes. This is perhaps no great surprise; nevertheless, this pattern can be seen across all three universities and, at some universities, it seems to be the most significant factor for predicting student attrition.

All three universities monitor dropout rates according to study programme with descriptive statistics. Methods for studying the differences between study fields and study programmes with explanatory and predictive analysis differ at the three universities, even though the patterns identified are similar.

One example is a machine-learning model to predict the risk of student dropout from the university. The purpose of the analysis was to identify the most significant factors for identifying students at risk of dropping out. Two different models were tested, and the first model included different variables such as the student's qualifying institution and average grade from qualifying exams. Of all the different variables in the model, study programme was identified as the most important factor for predicting attrition.

Another example is using regression analysis to test how different variables correlate with attrition. Here, study programme was also identified as an important factor in predicting attrition, as students enrolled at eight specific study programmes were twice as likely to drop out compared with all other student programmes.

Overall, this shows that attrition is a bigger issue on some study programmes than on others.

Based on this comparison between the three universities, we cannot identify specific study programmes that have bigger issues with high dropout rates. Our main point is rather that it is relevant to consider study programme as a factor when trying to improve retention rates. It also raises the question of why different programmes – with different study environments – have either high or low dropout rates. To answer this question, further investigation and analysis must be made.

## **2.2 FIRST YEAR AND BACHELOR STUDENTS**

All three universities in the working group monitor first-year dropout rates by level of degree (Bachelor's or Master's) with descriptive statistics. Consistent patterns can be seen across the universities.

At all three universities, we see that first-year dropout is significantly higher among first-year Bachelor's students than among Master's students. We cannot conclude that this is a general pattern across universities, but it seems highly plausible that university students struggle more with considerations about degree choice and whether to leave university during their first years at university and less during their Master's degree. It also seems worthwhile for universities to focus on the inclusion of new students if their aim is to reduce dropout.

## **2.3 STUDY PERFORMANCE BEFORE UNIVERSITY**

Among the universities in the working group, different types of analysis were conducted on students' performance prior to starting university and the effect this had on attrition.

It has been analysed how attrition differs when comparing students with different grade point averages from qualifying exams. Here the tendency is that students with a lower grade point average from their qualifying exams have a higher tendency to drop out compared with students with higher grades from qualifying exams. A machine-learning model has also been used to predict the risk of dropping out and, of all the factors in the model, including grades from qualifying exams, the most important and driving factor was study programme. This does not mean that study performance prior to university does not affect attrition – just that it is not the most important factor.

There are also examples of analyses of types of high school diploma and grades from qualifying exams. The type of high school diploma was identified as a relevant factor when predicting attrition, but there was not a general tendency for students with lower grades from qualifying exams to be more likely to drop out than students with higher grades.

On the basis of these preliminary findings, we cannot identify general patterns of how study performance prior to university affects attrition at the universities in the working group.

## **2.4 MOTIVATION FOR APPLYING TO UNIVERSITY**

One factor that was only tested by one university in the working group is admission priority. When a student applies for admission at a university in Denmark, he or she must prioritize between the study programmes they apply to. One would expect that students who were accepted onto their number one priority would be more motivated than students who were accepted onto a programme with lower priority. A correlation between students' admission priority and attrition can be identified, and it confirms this hypothesis: there is a tendency towards lower dropout rates among students that were accepted onto their number one priority than students who were accepted onto a programme with a lower priority.

This is not a pattern that was tested at the other universities in the working group and therefore it remains unknown whether this is a general tendency across all three universities.

Because admission procedures differ between different countries and perhaps between different universities in the same country, admission priority might not be a relevant factor at all universities, but motivation may well be.

Based on the findings from one university, the takeaway point is that high motivation for entering the specific university programme seems to lead to higher retention and lower risk of dropout. It would be relevant to conduct further analyses and investigate this at other universities as well to see if this is a general pattern, as it seems very plausible.

## **2.5 STUDY PERFORMANCE AT UNIVERSITY**

As mentioned, different machine learning models to predict student dropout have been tested at some of the universities in the working group. One model included factors such as performance in previous semesters, pass rates, and login behaviour in the Learning Management System (LMS). The result from the tested model was that pass rate from previous semesters at university was the most important factor for predicting dropout. It was seen that failing exams increased the risk of dropping out.

Study performance at university has also been tested with regression models, and these found similar patterns. Here the results indicated that number of credits earned during first year correlates with first-year dropout rates. As the number of credits decreased, the likelihood of dropout increased.

It is highly likely that performance at university affects attrition risk but, to draw specific conclusions, further analysis on the subject would need to be conducted.

## 2.6 FIRST ATTEMPTS TO TARGET STUDENTS AT RISK OF DROPPING OUT

In the sections above, we have mostly described patterns seen in descriptive analysis, regression analysis and few predictive analyses based on machine learning models. The focus was to identify specific factors that seem to correlate with attrition.

In the following section, we describe how some of the universities in the group have also worked with predictive data models to identify students at higher risk of dropping out.

Two of the universities have tested predictive data models that enable us to identify and target specific students at higher risk of dropping out. The advantage of this kind of model is that it is possible to identify and thereby offer specific and personalized help for students at risk of dropping out. The downside of this type of model is that it can be a black-box where it is not possible to identify clearly which factors are leading to a high-risk score. There is also a potential risk of stigmatizing students identified by the models and therefore increasing the risk that they drop out.

At one university, a predictive model was built as a proof of concept, but the model was not put into production as it was not possible to identify the specific features and factors in the model that led to a high-risk score. Consequently, it was not possible to recommend appropriate student interventions.

Another university also attempted to use these models in the spring of 2020, at a time when the university was expecting a challenge with attrition and extension of studies due to Covid-19. The model included many of the variables mentioned above that correlate with attrition. The following variables were included:

- 1 Historical attrition rate for the relevant degree programme
- 2 Date of commencement of studies (First-year student: Yes/No)
- 3 Number of ECTS credits earned
- 4 Number of courses for which the student was registered in the previous semester
- 5 Result of the last exam in the previous semester
- 6 Number of exams passed in the current semester
- 7 Number of absences from exams in the current semester

Based on the model, a list of high-score students was extracted and the student counsellors at the university used the information to offer counselling to the identified students.

The model was later tested against actual student performances, and the dropout among the students who the model identified was significantly higher than among those students the model did not identify. Based on this, it seems that the model was good at predicting students at risk of dropping out, but it is of course relevant to further study the effectiveness of the counselling interventions based on data from these types of models.

## 2.7 SUMMARY

The previous sections show that there are several similarities in factors correlating with student attrition between the three different universities. This is not to say that the described patterns can be generalised to

other universities, but we encourage further analysis on the subject, and the patterns presented can be used as a starting point for further investigation on the topic.

One similarity is that attrition rates seem to differ between study fields and study programmes, which means that attrition is a bigger issue on some programmes than on others. We also see a tendency towards higher dropout rates among first-year and Bachelor's students. Another similarity is that factors relating to the individual student, such as study performance, seem to correlate with higher risk of attrition, and that predictive models can be used to identify students with high dropout risk. This means that it is possible to focus retention initiatives on specific groups of students and that the knowledge can help identify where to focus when trying to increase student retention.

### **3 PRELIMINARY DATA ANALYSIS ON INCLUSION AND COMPOSITION OF THE STUDENT MASS**

As mentioned in the introduction, data analysis on inclusion is a relatively new topic at the three universities, but as the topics of inclusion and diversity are rising on the agendas of the universities, all three universities in the working group have initiated data analysis on these subjects.

All three universities collect personal information on the students in connection to admission, such as information on educational background, age and gender. To function as a university, it is necessary to collect and process a range of personal data to ensure that public sector tasks related to data-supported quality assurance can be carried out (see the White Paper "Legal and Ethical aspects of using Learning Analytics from a University perspective"). Therefore, information such as educational background, age and gender is collected. The collected information can be used to monitor and analyse the composition of the student mass at the university, and discussions in the working group show that this is also the case at the universities in this working group.

All three universities monitor the student mass with regard to age, gender, type of high school diploma, citizenship, region of residence (among other things) and present this information in data reports. This means that educational leaders and other relevant staff across the university can access the information and use this to analyse student composition and to carry out their various decision-making processes.

The universities differ slightly in whether they collect information about family background, such as parents' income and educational level, and information about student disabilities.

At one university, data on parental income and education level is collected as part of the regulatory reporting process and used to analyse opportunities for widening participation for traditionally under-represented groups. At this university, equality, diversity, and inclusion data has been widely monitored and analysed for several years, but more emphasis is placed on descriptive rather than predictive analytics. Guardrails to prevent bias in predictive models is being considered as the university seeks to move into predictive analytics.

At another university, data about family background is not currently collected, but the university will start collecting information about parents' educational level in a questionnaire by the 2023-2024 academic year for newly enrolled students. This is because the information is required by some international ranking agencies. This may lead to new possibilities for analysis on inclusion and diversity.

At the last university, information about family background is not collected, but the information is available on an aggregated level from the Ministry of Higher Education and Science. Therefore, student intake, the number of students and student dropout can be examined by background information such as gender, age, origin, parent's educational background and grade from qualifying exam. As inclusion is now on the university agenda, it has led to new projects and mapping of the current diversity of the student population. The purpose

is to shed light on how diverse the population is on different parameters, and how it has changed over time, thereby being able to identify which initiatives are required to increase inclusion.

The initial result of the project shows that the student's educational level is equally as dependent on the educational level of their parents as it was 40 years ago, but, as parents' educational levels have slightly improved over time, so has the educational level of their children. This means that, even if the general educational level has increased over time, social inheritance is still strong.

As the analysis on diversity and inclusion is relatively new at some of the universities, it is not possible for us to identify similarities and differences between the three universities in this paper. Again, we cannot point to general patterns. Instead, the above section is included to show that the topic is on the agenda at the universities and therefore is also starting to be the subject of data analysis in the universities' administrative departments.

## 4 SUMMARY AND CONCLUDING REMARKS

In this paper, the Circle U. working group on Learning Analytics has presented examples of analysis on and tendencies in student attrition and inclusion at three different universities. The purpose of the paper was to give the reader insights into what the three university administrations have been working on and tested in connection to analysis on attrition and inclusion, where the universities can identify similar trends, and where they differ. The paper did not include an actual analysis on the subject, as the data quantity is too small, but the paper can be used as inspiration for further analysis on the subjects.

For attrition, we identified similar trends across the three universities according to differences in dropout rates between Bachelor's and Master's students, differences between types of study fields, and difference between higher- and lower-performing students. We see differences in analysis results about student performance before entering university, while the link between motivation for applying for university and retention was only tested at one university. Two out of three universities have also tested predictive models that can identify specific students with high dropout risk in order to offer them counselling.

For inclusion, we have described the universities' preliminary analysis on inclusion and diversity, but, given that the universities have conducted analyses of different scopes and lengths on these subjects, we are unable to point out differences and similarities at present. This needs to be further elaborated when further analysis has been made.

Overall, the mapping and discussions in the working group show that the universities have focused primarily on student retention. They are also beginning to focus more on topics of diversity and inclusion, which is also reflected in the analysis conducted at the individual universities. The activities conducted as part of the bridging school and university project in Circle U. – which included the development of a mentoring programme and student-driven bridging activities – show promise, and their importance is underlined by the findings in this report.

Based on this paper, it is not possible to conclude that the tendencies and correlations described in the paper are general patterns across universities and countries, as making such conclusions would demand a larger dataset from various universities. Instead, we encourage readers to pick up where we have left off and use the preliminary findings in this paper as inspiration for further analysis that can lead to identifying general ways to reduce attrition, or increase retention, and secure inclusion at universities.