Correlation between achievement in the certificate and achievement in subsequent degrees

Anne Wendt and Kerri Spooner

Auckland University of Technology
Internal Report

Objective and background

Auckland University of Technology (AUT) AK1311 programme, the Certificate in Science and Technology, is used as an entry qualification for AUT undergraduate degrees in Computer Science, Mathematical Sciences, and Engineering, in particular AK1271 (Bachelor of Mathematical Sciences), AK3697 (Bachelor of Computer and Information Sciences), and AK3719 (Bachelor of Engineering Technology). The objective of this research is to validate the certificate's usefulness for students as an entry qualification for those degrees.

The main question that this report tries to answer is: **Is a student's performance in the certificate an indicator of their success in a subsequent degree?**

The following research questions shall be answered:

1. Is there a correlation between students' average grade in the certificate and their average grade in the degree?
2. How do certificate students perform over time in their subsequent degrees?
    a. How does an 'A'-grade student perform in their first, second, and third year of a degree?
    b. How does a 'B'-grade student perform in their first, second, and third year of the degree?
    c. How does a 'C'-grade student perform in their first, second, and third year of the degree?
3. How do certificate students perform in subsequent degrees compared to non-certificate students?
    a. Does this differ by degree?
    b. Does this differ by year of enrolment?
4. Are there any certificate papers that are predictive for subsequent success?
5. Is there a difference between people from a native English background and people who are not native speakers of English?

In this report, the programme AK1311 (Certificate in Science and Technology) is called *certificate* and abbreviated as *Cert*. The three other programmes are called *degree*, using the abbreviations *Math.* for the Bachelor of Mathematical Sciences, *Comp.* for the Bachelor of Computer and Information Sciences, and *Eng.* for the Bachelor of Engineering Technology.

Description of the dataset

Student data was gathered from AUT's Arion database. A complete list of papers taken and grades achieved by students who were enrolled in any of AK1311 (Cert.), AK1271 (Math.), AK3697 (Comp.), or AK3719 (Eng.) from 2011 to 2014 was available. Each row in the list contained the enrol year, student ID, student name, programme code, programme name, campus, paper code, paper name, and achieved paper grade in letter format. Each entry could be uniquely identified by combining student ID and paper code.

In total, there were 903 students enrolled in the certificate and 3,775 students enrolled in a degree from 2011 to 2014. Out of this, 394 students were enrolled in both programmes. Thus, 509 students were only enrolled in the certificate but not in a degree, and 3,381 were only enrolled in a degree without being enrolled in the certificate beforehand.

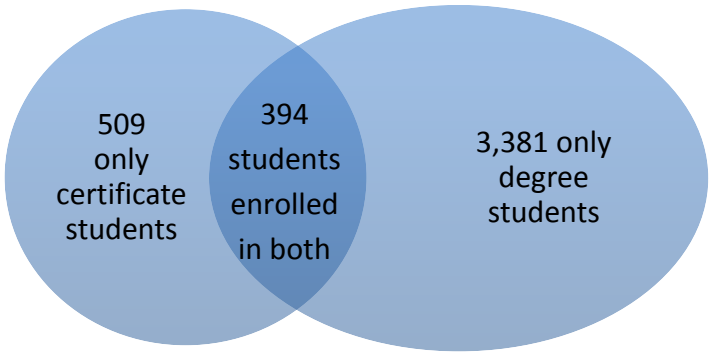

Figure 1: Student numbers for the certificate and the degree courses

After completing the certificate, 394 students progressed to one of the three degrees. Table 1 shows how many certificate students continued in which degree, split by year and degree. It shows how many certificate students were enrolled in a specific degree in the specified year. The aggregated values are not the sums of the columns and rows, because students who progressed through a degree over several years were included in each year separately. Furthermore, there were nine students who changed between degrees, and students who dropped out during the year, which further influenced the aggregated figures.

Table 1: Student numbers progressing from the certificate by year and degree

Year/degree	AK1271 (Math.)	AK3697 (Comp.)	AK3719 (Eng.)	All degrees
2011	3	36	12	**51**
2012	12	75	50	**136**
2013	12	120	90	**219**
2014	19	179	134	**331**
All years	**25**	**226**	**155**	**394**

The numbers for 2011 are relatively lower than for the other years, because the 2011 numbers only include students who did the certificate in the first semester and their degree in the second semester of 2011. This is because no data was available about enrolment in the certificate in 2010. The table data is visualised in Figure 2.

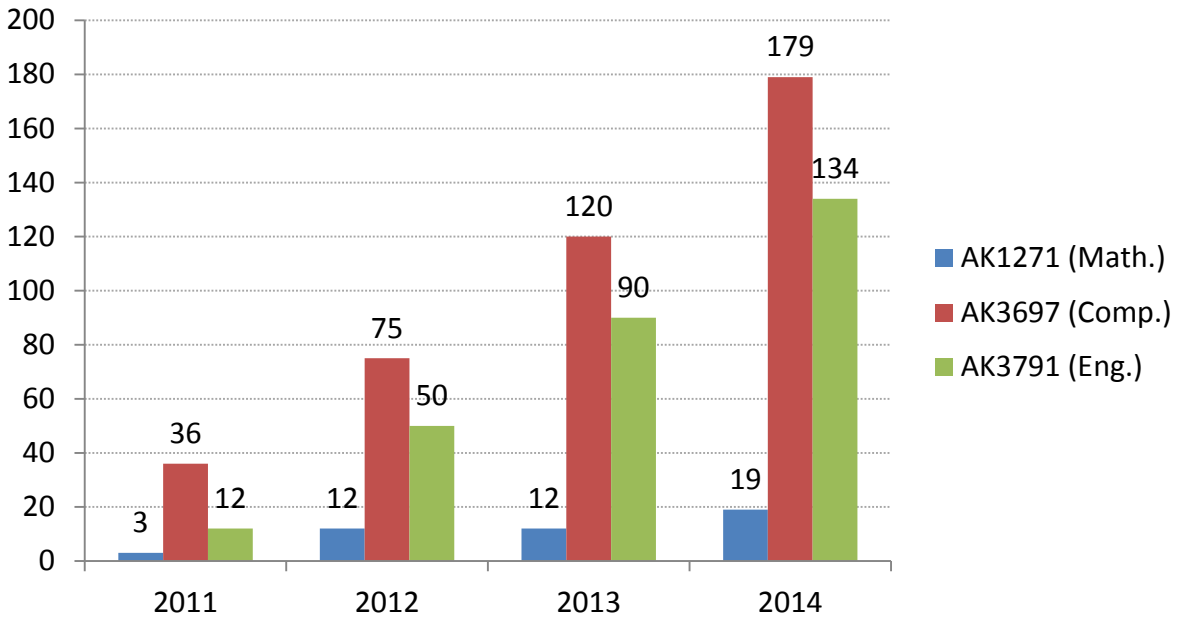

Figure 2: Student numbers progressing from certificate split by year and degree

The majority of students coming from the certificate enrolled in the Bachelor of Computer and Information Sciences, which is followed in popularity by the Bachelor of Engineering Technology. Only very few students progressed to the Bachelor of Mathematical Sciences. It can be noticed that the overall number of students progressing from the certificate to a degree steadily rose over time.

Furthermore, Table 2 was available as input data. It shows how many students staircased straight from the certificate into a degree. Since there were also some students who had at least one semester of no enrolment between their certificate and their degree, these figures differ from the ones in Table 1.

Table 2: Number of students staircasing from the certificate into degrees

Year	Degree	Sub-total	Total
2012	Bachelor of Engineering (Honours)	17	152
	Bachelor of Engineering Technology	39	
	Bachelor of Computer and Information Sciences	69	
	Bachelor of Mathematical Sciences	9	
	Bachelor of Business	7	
	Other AUT Bachelor Programmes	3	
	AUT Certificate Programmes	8	
2013	Bachelor of Engineering (Honours)	24	160
	Bachelor of Engineering Technology	47	

	Bachelor of Computer and Information Sciences	54	
	Bachelor of Mathematical Sciences	5	
	Bachelor of Business	10	
	Other AUT Bachelor Programmes	7	
	AUT Certificate Programmes	13	
2014	Bachelor of Engineering (Honours)	23	193
	Bachelor of Engineering Technology	56	
	Bachelor of Computer and Information Sciences	69	
	Bachelor of Mathematical Sciences	8	
	Bachelor of Business	16	
	Other AUT Bachelor Programmes	13	
	AUT Certificate Programmes	8	

Further data that were available was a list of students enrolled in the certificate from 2011 to 2014 whose native language was not English.

Methodology

Microsoft Excel was used for pre-processing, data handling (splitting, merging, accumulating), and data visualisation. The Weka Explorer was used for computational data analysis such as linear regression, information gain calculation, and decision tree.

First, irrelevant data such as student name, programme name (while keeping the programme code), paper name (while keeping the paper code), and campus were deleted from the list. Then, new columns were added which contained a rounded grade and the grade converted to a number. The rounded grade is simply the original grade without any + or - amendments. For example, a B+, a B, and a B- were all rounded to a B. The number grade was converted based on the GPA system from 0 to 4[1]. Table 3 shows the conversion of grades that was used.

Table 3: Grade conversion

Original grade	Meaning	Pass/fail	Rounded grade	Number grade
A+, A	Pass with distinction	Pass	A	4
A-	Pass with distinction	Pass	A	3.7
B+	Pass with merit	Pass	B	3.3
B	Pass with merit	Pass	B	3
B-	Pass with merit	Pass	B	2.7
C+	Pass	Pass	C	2.3

[1] https://registrar.princeton.edu/student-services/transcript/gpa.pdf

C	Pass	Pass	C	2
C-	Pass	Pass	C	1.7
D	Fail	Fail	D	1
CO	Conceded pass	Pass	C	2
S	Aegrotat pass	Pass	C	2
DNC	Did not complete	Fail	D	0
W	Withdrawn	Fail	D	0

For **RQ1**, a student's average grade in the certificate was compared to their average grade in a subsequent degree. For each student, their average grade in the certificate and in the degree was calculated. The values were then put into relation to each other using the linear regression functionality of Weka Explorer. In addition, it was compared if students had passed or failed in the certificate and degree, respectively.

For **RQ2**, a student's average grade per year (and per degree if they changed it) was computed. The results were then compared to see if students performed better or worse over time. For better comparison, students were also split by their average grade in the certificate. It was also checked for how long students were enrolled in a specific degree.

For **RQ3**, the average grades of students who had enrolled in the certificate previous to commencing a degree were compared to the average grades of students who had not enrolled in the certificate. For this, the list of all degree students was split into two groups which were certificate students and non-certificate students. For each group, the average grade was calculated. This analysis was done as an aggregate over all degrees and years as well as split by degrees and years. In addition, the proportions of students coming from the certificate and not coming from the certificate who had passed in the degree were computed and compared.

For **RQ4**, all certificate students' grades in papers were put into relation to their average grade in subsequent degrees. This created a matrix which combined all students with all papers. If a student had taken a paper, the cell value was their grade in that paper. If the student had not taken the paper, the value was set to 0. The first column was the student's ID, followed by 22 columns containing their different paper grades; and the last column was their average degree grade. For the paper grades, three different input types were used which are the original grade, the rounded grade, and the pass/fail grade as defined in Table 3. For the average degree grade, the original, the rounded, and the pass/fail grade were used. Then the Information Gain Evaluator in Weka Explorer was applied to determine which certificate papers predicted the degree grade best. The information gain method shows the change in information entropy if a variable is applied to a class. It is defined as

$$InfoGain(Class, Attribute) = H(Class) - H(Class|Attribute)$$

with H being the entropy, or measure of unpredictability of the information content. Thus, information gain shows the amount of mutual information between the attribute and the class value. Based on the findings from the Information Gain analysis, the five best features

(papers) for each combination were selected and then used to predict the degree grades using an ID3 decision tree algorithm in order to compare their validity.

For **RQ5**, students were divided into two groups, namely students whose first language was English, and students whose first language was not English. Their average grades in the certificate and in their subsequent degree were then grouped by course and year and compared with each other. They were also split by average certificate grade in order to compare their performance.

Results

RQ1 (correlation between average grades in certificate and degree)

Using linear regression, the following relation was computed:

$$average\ degree\ grade = 0.54 * average\ certificate\ grade + 0.54$$

The correlation coefficient between both values is 0.4903. This correlation coefficient indicates a weak to moderate relationship between the average certificate grade and the average degree grade of a student. Figure 3 shows the relation between the average grade in the certificate (x-axis) and the average grade in a degree (y-axis). The line of regression is plotted in blue. The upwards scatter of the data from left to right combined with the correlation coefficient indicate a weak-moderate positive relationship between average grade on the certificate and average grade on the degree.

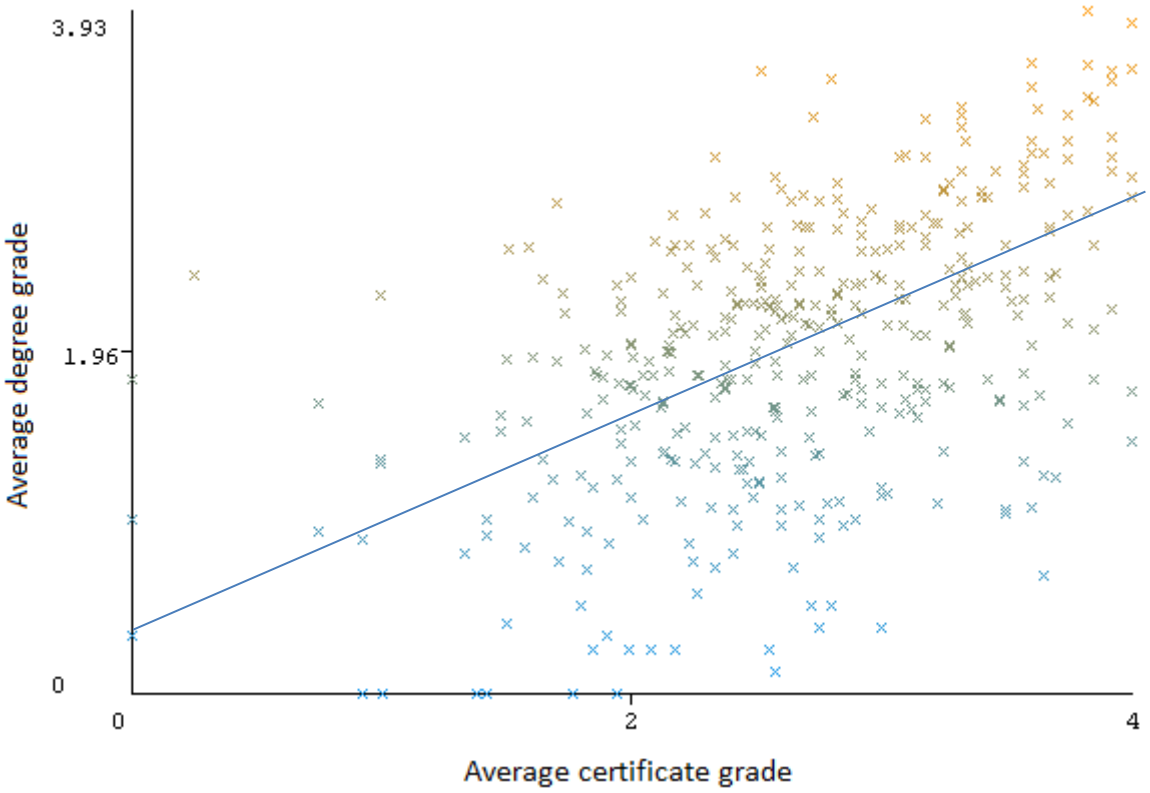

Figure 3: Relationship between average certificate grades and average degree grades

In addition, 86.8% of students (342 students) who passed their certificate papers also passed their papers on the degree. The average grade of the 342 students was 2.79 [B-] in the certificate and an average grade of 2.16 [C+] in their subsequent degree.

10.4% of students (42 students) passed their papers on their certificate with an average grade of 2.15 [C+] and failed in their papers in their subsequent degree.

There were five students who failed their papers in both programmes. Six students who had failed their papers in the certificate had passed their papers in the degree with an average grade of 1.8 [C-]. Figure 4 shows these numbers in a diagram.

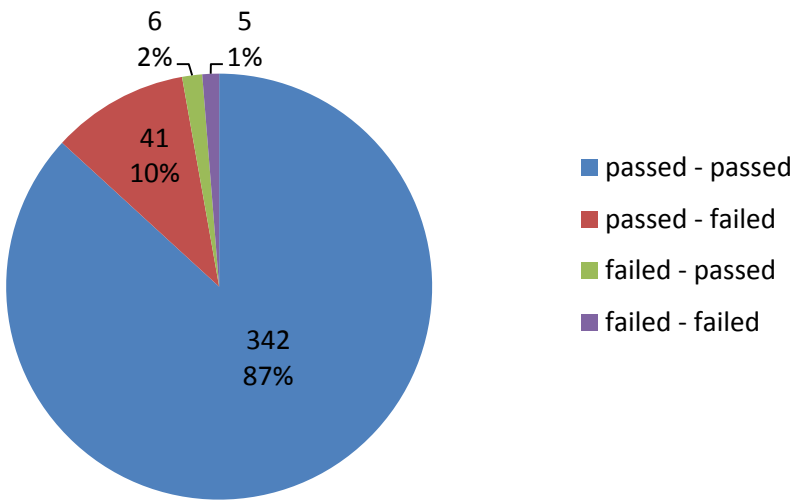

Figure 4: Distribution of students passing and failing the certificate and degree

RQ2 (certificate students' performance over time)

Figure 5, Figure 6, and Figure 7 show the average degree results for certificate students enrolled in AK1271 (Math.), AK3697 (Comp.), and AK3719 (Eng.), respectively, over time.

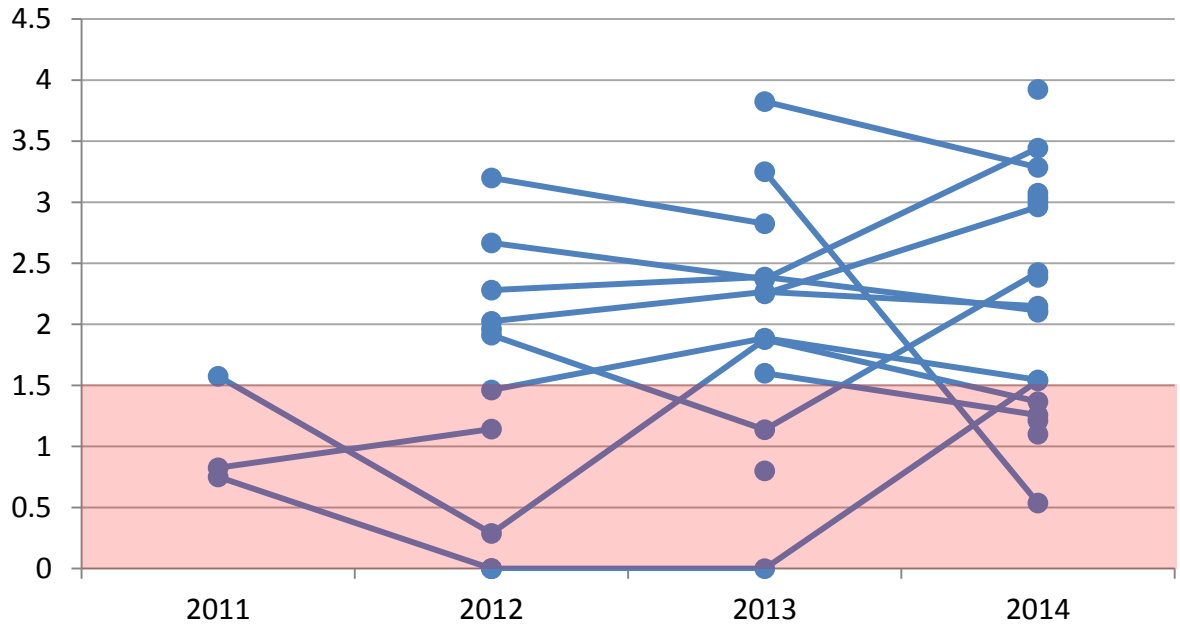

Figure 5: Performance of certificate students in AK1271 (Math.) over time

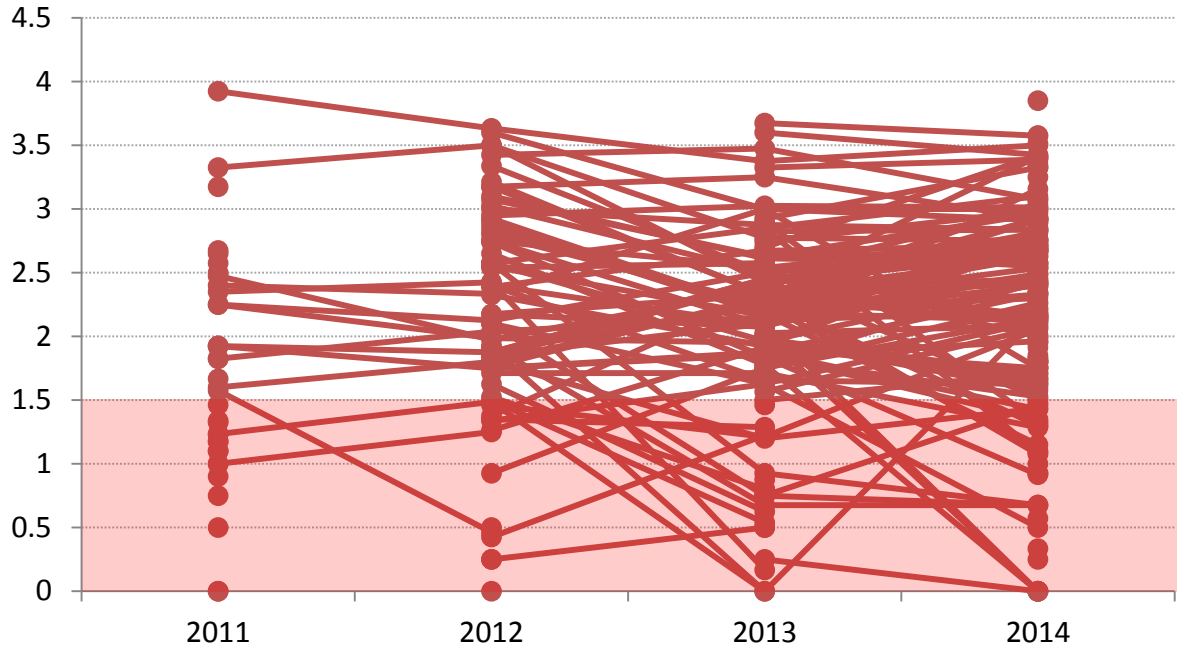

Figure 6: Performance of certificate students in AK3697 (Comp.) over time

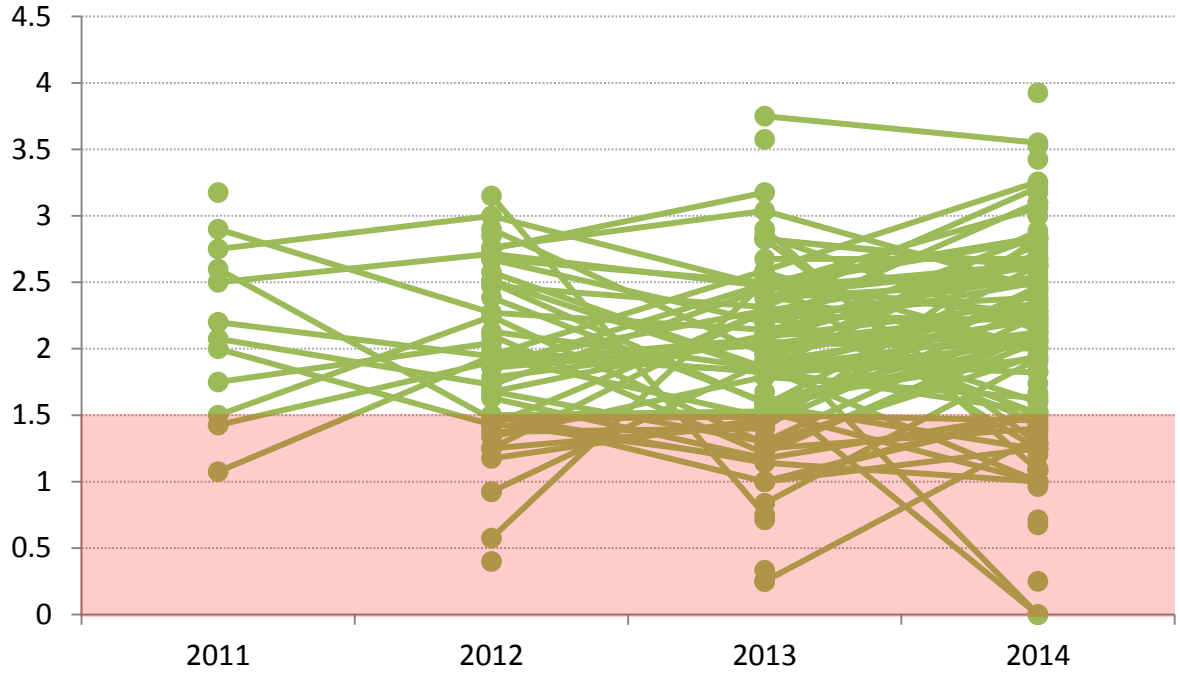

Figure 7: Performance of certificate students in AK3719 (Eng.) over time

Since the numbers of students enrolled in AK3697 (Comp.) and AK3719 (Eng.) were too large to be analysed properly in one diagram, the students were split by their average certificate grade. Table 4 shows a comparison of these results. The red overlay shows fail grades.

Table 4: Comparison of certificate students' success in the degrees, split by average certificate grade

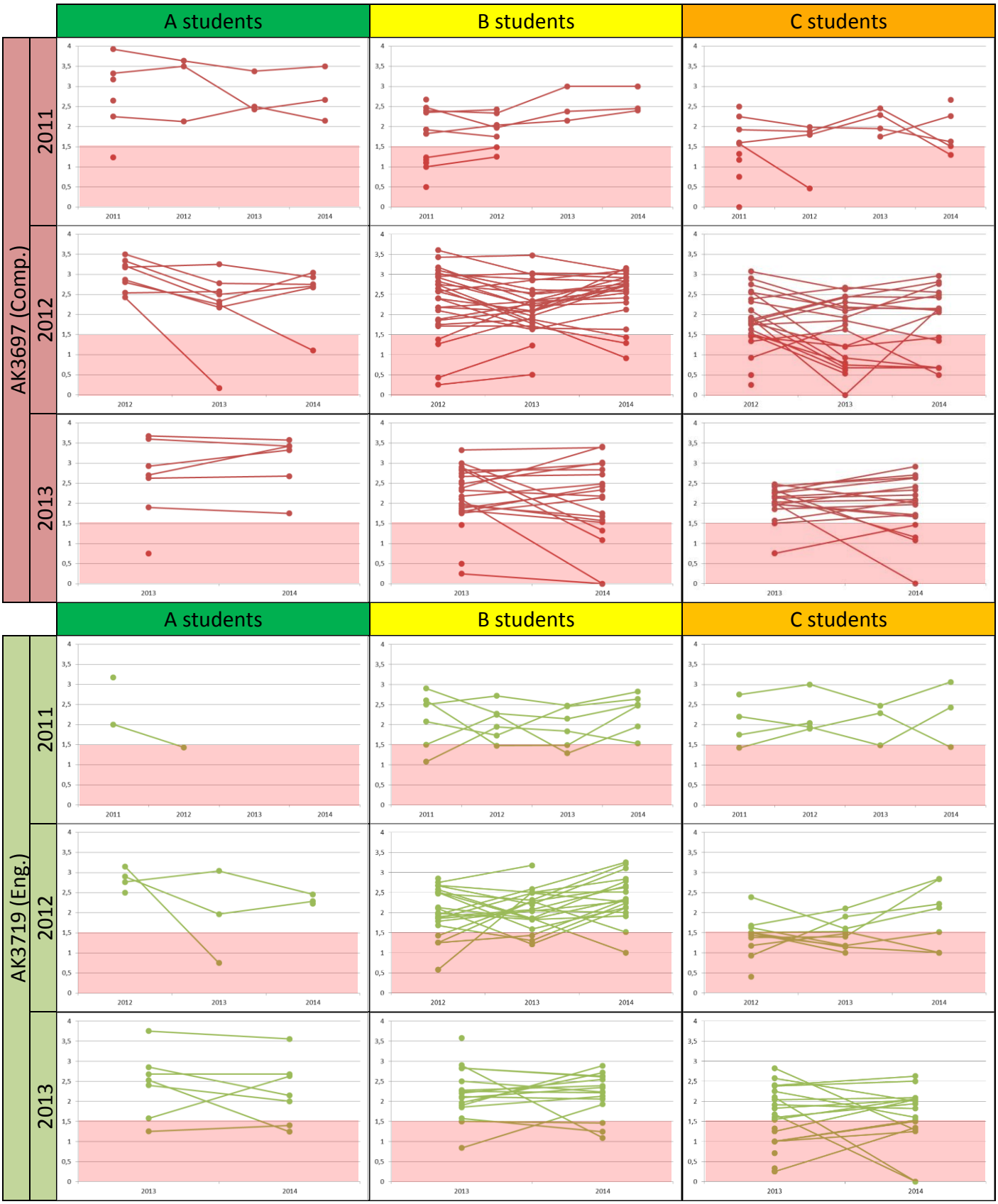

Figure 8 shows how many students were enrolled for how long in a specific degree. An enrolment of one year was coloured red. These students dropped out of their degree very quickly. Two years are marked in orange. These students dropped out before completion of their degree; however, they did not give up so quickly. An enrolment of three years is marked green. These students usually finished their degrees. Four years are marked in yellow. These are individual cases. For example, some students could have started their studies in the middle of the year (meaning they were enrolled for six semesters but spread over four years), or they were enrolled part-time and thus took longer to complete their degree, so this is not generally an indicator of failure. All students who are still enrolled but had not yet completed their degree in 2014 are marked blue.

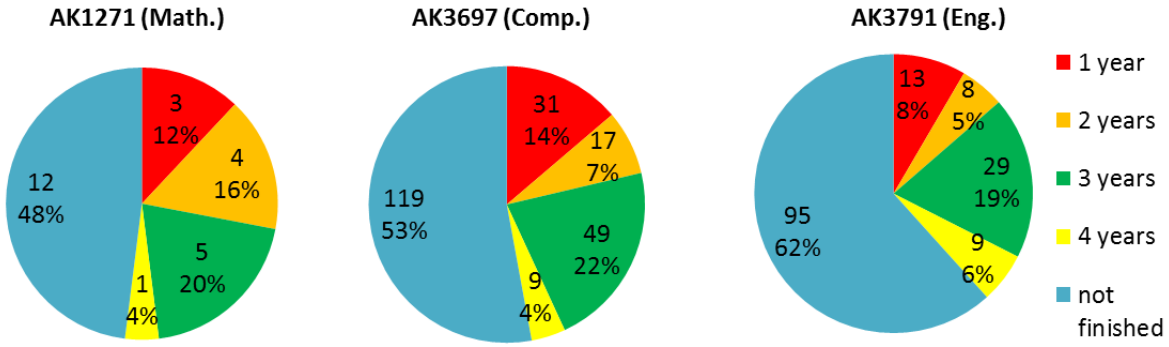

Figure 8: Duration of enrolment of certificate students by degrees

Figure 9 shows for how long certificate students were enrolled in a specific degree. This excludes all students who have not finished yet (the blue group from Figure 8).

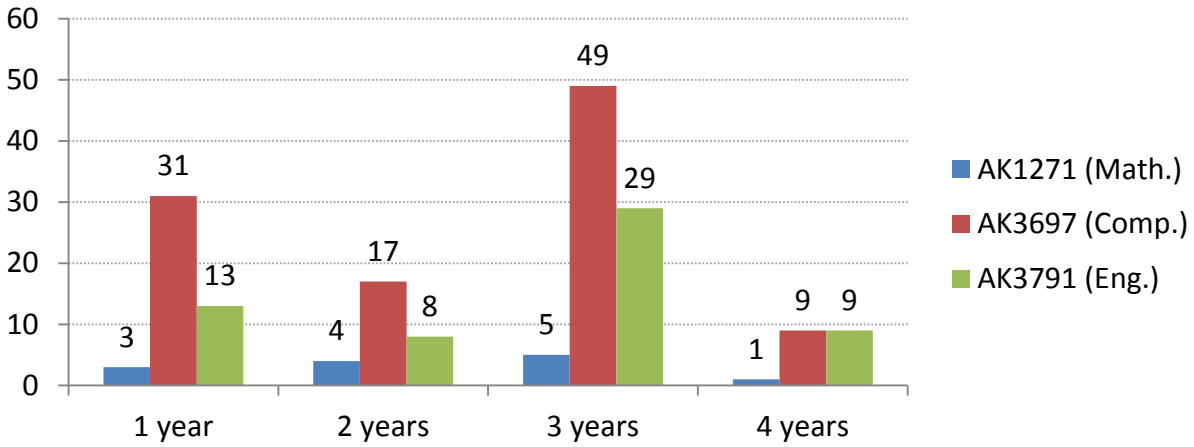

Figure 9: Duration of enrolment of certificate students by years

The highest number for all three courses can be found for three years of enrolment. This means the majority of students completed their degrees. With the exception of AK1271 (Math.), the drop-out rate is higher at the end of one year than after two years. Only very few students were enrolled for four years.

Figure 10 shows how many certificate students were enrolled in each degree for how long, excluding all those students who were categorised as not finished yet (marked in blue in Figure 8).

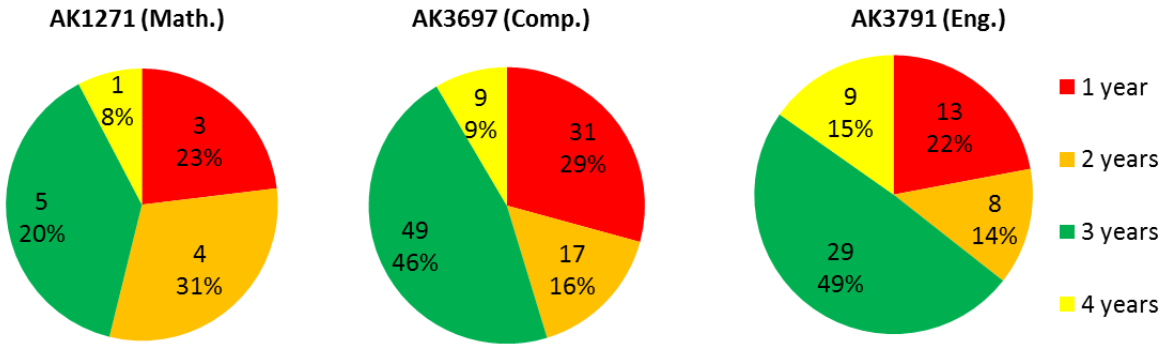

Figure 10: Duration of enrolment of certificate students by degrees excluding not finished students

The majority of certificate students on Comp. and Eng. degrees stayed enrolled long enough to finish their degrees. $\frac{6}{13}$ Math. students stayed enrolled long enough to finish their degrees.

Combining Comp. and Eng. certificate students the large majority of the students (71.1%) stayed enrolled long enough to finish their degrees. (The Math. results are not included as the numbers are too low they will distort the true picture).

There was no data available that could confirm that these students actually graduated. We confidentially say a large percentage of certificate students complete their degree.

RQ3 (performance of certificate and non-certificate students)

The overall average grade of students enrolled in AK1271 (Math.), AK3697 (Comp.), or AK3719 (Eng.) between 2011 and 2014 who had also enrolled in AK1311 (Cert.) before commencing their degree was 1.98[2]. In this group, 348 out of 394 students (88.3%) passed their papers in their degree. The overall average grade of students enrolled in AK1271 (Math.), AK3697 (Comp.), or AK3719 (Eng.) who had *not* enrolled in AK1311 (Cert.) before commencing their degree was 2.20. In this group, 2,960 out of 3,381 students (87.5%) passed their papers in the degree.

Table 5 shows the average grades of Cert. and non-Cert. students split by year and degree. Each cell contains two values, which are certificate students and non-certificate students respectively.

[2] Note that this value is different from the one presented for RQ1 because this includes *all* Cert. students.

Table 5: Average grades split by year, degree, and whether the student was enrolled in the certificate

degree / year	AK1271 (Math.)		AK3697 (Comp.)		AK3719 (Eng.)		All degrees	
	cert	*no cert*	*cert*	*no cert*	*cert*	*no cert*	*cert*	*no cert*
2011	1.05	2.24	1.73	2.32	2.16	2.15	1.79	2.26
2012	1.41	1.97	2.16	2.38	1.92	2.28	2.00	2.32
2013	2.21	2.20	2.03	2.38	1.88	2.31	1.98	2.34
2014	2.23	2.28	2.21	2.21	2.00	2.47	2.17	2.29

Figure 11 shows these student grades split by year and degree. It visualises the data presented in Table 5. A darker colour was used for Cert. students, and a lighter colour was used for non-Cert. students.

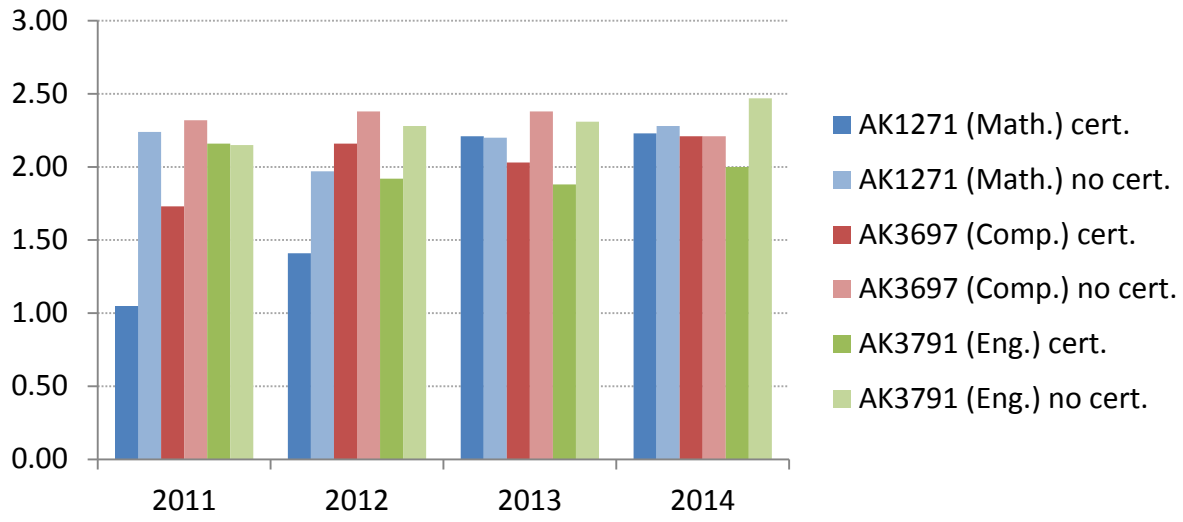

Figure 11: Average degree grades split by year, degree, and whether the student was enrolled in the certificate or not

Table 6 shows the differences between the average grades of certificate and non-certificate students. The colour scale indicates how much better or worse the certificate students perform in comparison with the non-certificate students. Red indicates a very high difference (certificate students are much worse than non-certificate students) and green indicates that their grades are very close to each other. Figure 12 shows these differences in a diagram.

Table 6: Grade differences between certificate and non-certificate students

Year/degree	AK1271 (Math.)	AK3697 (Comp.)	AK3719 (Eng.)	All degrees
2011	-1,19	-0,59	0,01	-0,47
2012	-0,56	-0,22	-0,36	-0,32
2013	0,01	-0,35	-0,43	-0,36
2014	-0,05	0,00	-0,47	-0,12

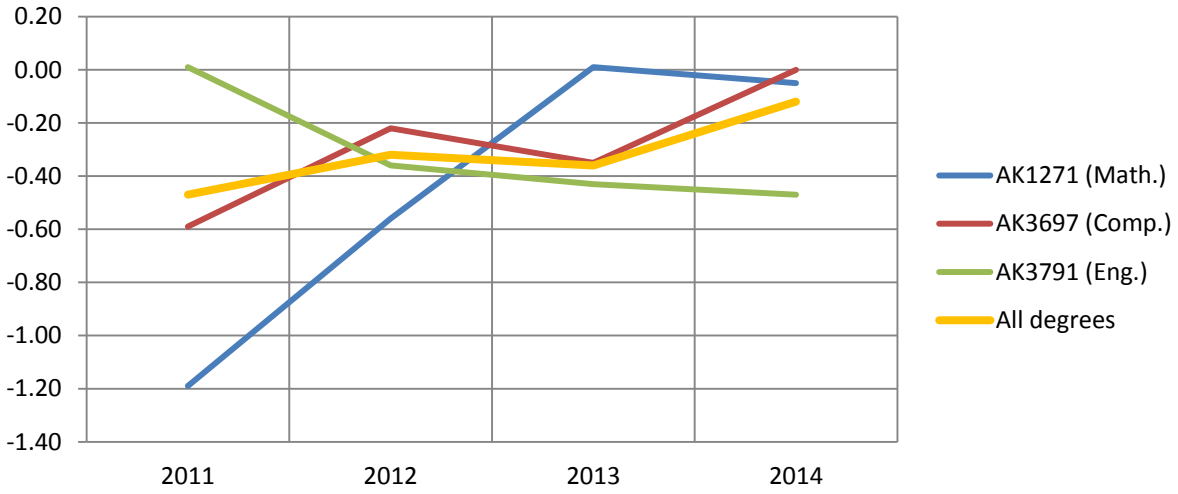

Figure 12: Differences between certificate and non-certificate students by degree over time

It can be noticed that Cert. students who enrolled in the Math. degree performed better over time compared to non-certificate students, whereas the gap between certificate and non-certificate students enrolled in the Eng. degree increased slightly over time. The results of Comp. students fluctuated, which also influenced the overall performance value, since the majority of students were enrolled in that degree.

RQ4 (predictive certificate papers)

Table 7 shows which papers predicted the degree grade best. Given are the top five paper codes for each combination with the respective information gain value (for calculation method, please refer to Methodology section). The paper codes are highlighted with colours to better identify them.

Table 7: Papers from the certificate that predicted the degree grade

	Original degree grades	Rounded degree grades	Pass/fail grades
Original paper grades	404024 with 0.30494 154011 with 0.29859 724001 with 0.2721 714287 with 0.26056 404504 with 0.20966	404024 with 0.20029 724001 with 0.15849 154011 with 0.1508 714287 with 0.12744 404504 with 0.11026	404024 with 0.08047 154011 with 0.05898 724002 with 0.04641 714287 with 0.04421 154102 with 0.04014
Rounded paper grades	404024 with 0.16677 154011 with 0.15557 404504 with 0.13971 724001 with 0.13071 714285 with 0.11919	404024 with 0.10978 154011 with 0.08552 404504 with 0.0853 724002 with 0.07479 714286 with 0.07456	404024 with 0.053887 154011 with 0.03896 724002 with 0.03183 404504 with 0.02615 144702 with 0.02291
Pass/fail grades	724001 with 0.06685 154011 with 0.05986 724002 with 0.0542 714285 with 0.05149 714286 with 0.04405	724002 with 0.04794 724001 with 0.04178 714286 with 0.03487 154011 with 0.03211 144702 with 0.02887	154011 with 0.024642 404504 with 0.02156 144702 with 0.016226 404024 with 0.012615 724002 with 0.011818

It was found that throughout all different combinations the five most predictive papers were

- 154011 (Academic Literacies for Computing & Mathematical Science) with nine instances
- 404024 (Foundation Programming) with seven instances
- 404504 (Foundation Logic Skills) and 724002 (Foundation Physics B) with six instances each
- 724001 (Foundation Physics A) with five instances

The information gain was greatest for mapping original paper grades into original degree grades. However, this may be due to over-fitting. Selecting the five most predictive papers from that combination as features and applying the ID3 decision tree algorithm to the dataset yielded only a classification success of 7.6%. On the other hand, the information gain was lowest for mapping pass/fail certificate grades into pass/fail degree grades. Selecting the five most predictive papers for this combination as features and applying the same decision tree algorithm to the dataset yielded a classification success of 69.2%. This means that the single papers are most predictive for the original detailed grades, but a classification can be performed best when generalising the grades to pass and fail. All results for classifying the degree results based on the paper grades are shown in Table 8:

Table 8: Classification results for predicting the degree grade from certificate paper grades

Prediction	Original degree grade	Rounded degree grade	Pass/fail degree grade
Original paper grade	7,6%	22,9%	48,1%
Rounded paper grade	13,7%	37,4%	62,3%
Pass/fail paper grade	19,6%	43,3%	69,2%

It can be noticed that predicting the pass or fail of a student can be achieved much better than predicting the exact grade. Also predicting from a pass/fail grade achieves better results than predicting from other grade types.

RQ5 (performance of native and non-native speakers of English)

Table 9 shows the difference between students from a native English background and students who speak English as a second language. The average grades of students who progressed from the certificate to a degree were split by year and course. Each cell contains two values: the left one is the average grade for native English speakers, and the right one is the average grade for non-native English speakers.

Table 9: Comparison of native and non-native English students

course / year	AK1311 (Cert.)		AK1271 (Math.)		AK3697 (Comp.)		AK3719 (Eng.)	
	native	non	native	non	native	non	native	non
2011	1.95	2.04	1.16	0.83	1.69	1.75	2.37	2.02
2012	1.87	1.97	0.14	1.89	2.12	2.13	1.82	2.04
2013	2.25	2.12	2.33	2.03	2.00	2.11	1.80	1.90
2014	2.40	2.19	1.81	2.46	2.15	2.40	1.89	2.26

The date from Table 9 is visualised in Figure 13.

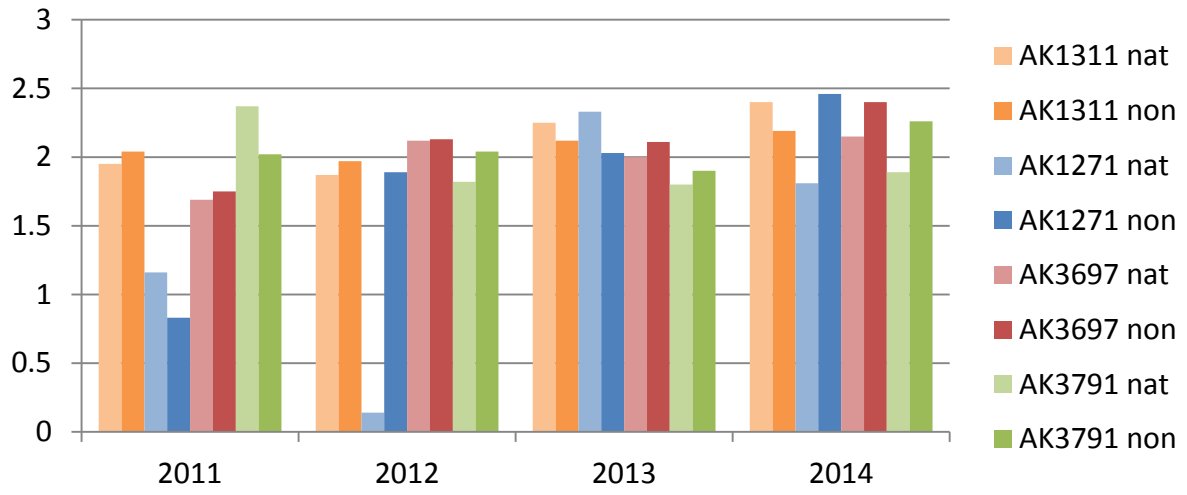

Figure 13: Comparison of native and non-native English students

The differences between native and non-native speakers of English are shown in Table 10. The grades of non-native English speakers were subtracted from the grades of native speakers. A good performance of non-native students is indicated in yellow, a good performance of native students is indicated in blue, and similar performance is indicated in green.

Table 10: Grade differences between native and non-native speakers of English (value = native – non-native)

Year	AK1311 (Cert.)	AK1271 (Math.)	AK3697 (Comp.)	AK3719 (Eng.)
2011	-0.09	0.33	-0.06	0.35
2012	-0.1	-1.75	-0.01	-0.22
2013	0.13	0.3	-0.11	-0.1
2014	0.21	-0.65	-0.25	-0.37

It can be seen that the differences between native and non-native students of English are very small, with the exception of the Math. degree. The reason for this could be that there are only very few students progressing into that course. Figure 14 visualises these grade differences in a diagram.

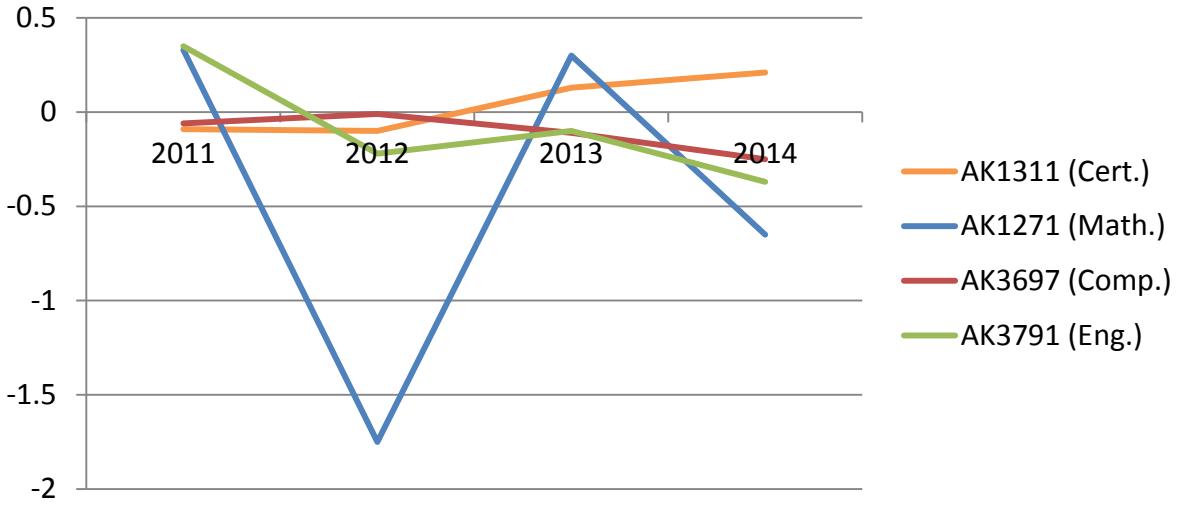

Figure 14: Grade differences between native and non-native speakers of English

It can be noticed that especially in the certificate and the Comp. degree there is no big difference between students from a native and a non-native English background. Figures for the Math. degree fluctuate greatly because there were only three students with a native English background in 2012 who all performed very badly.

The results for comparing the average grades of native English speakers and non-native English speakers were further broken down by the students' average certificate grades in order to find out if there were any differences between them. Table 11 shows these data. If the value is N/A, there were no students enrolled in that specific degree with that certificate grade and that English background in that year.

Table 11: Average grades of native and non-native speakers of English, split by their average certificate grade

		AK1311 (Cert.)		AK1272 (Math.)		AK3697 (Comp.)		AK3719 (Eng.)	
Year	Grade	native	non	native	non	native	non	native	non
2011	A	3.77	3.73	N/A	N/A	3.93	2.59	2.39	N/A
	B	2.82	2.88	1.16	0.83	1.69	1.72	2.90	1.95
	C	2.12	2.14	N/A	N/A	1.25	1.62	1.75	2.13
2012	A	3.68	3.67	0.00	N/A	3.17	2.93	2.05	2.73
	B	2.83	2.85	0.19	2.04	2.04	2.30	1.90	2.16
	C	1.86	2.07	N/A	N/A	1.45	1.83	1.71	1.63
2013	A	3.72	3.65	3.25	3.83	2.12	2.67	1.86	2.41
	B	2.84	3.05	1.88	2.04	2.18	2.23	2.07	1.98
	C	2.17	2.31	N/A	1.53	1.82	1.84	1.36	1.65
2014	A	3.73	3.73	1.70	3.29	2.84	2.96	1.95	2.52
	B	3.01	2.87	1.90	2.29	2.27	2.47	2.01	2.39
	C	2.29	2.20	1.76	2.96	1.81	2.12	1.67	1.93

Table 12 shows the differences between the grades of native and non-native English speakers, split by average certificate grade. Again, the values were computed by subtracting

the average grades of non-native English speakers from those of native English speakers. As in Table 10, colour shading was used to indicate how far the values are apart. Yellow indicates a comparatively good performance of non-native English students, blue indicates a comparatively good performance for native English students, and green indicates a relatively similar performance. It can be clearly seen that the grade differences are marginal. Only the differences for the Math. degree (AK1271) were comparatively high, which is probably due to the (little) number of students enrolled in that programme.

Table 12: Grade differences between native and non-native English students, split by their average certificate grade (value = native – non-native)

Year	Grade	AK1311 (Cert.)	AK1271 (Math.)	AK3697 (Comp.)	AK3791 (Eng.)
2011	A	0.03	N/A	1.33	N/A
	B	-0.07	0.34	-0.03	0.95
	C	-0.02	N/A	-0.37	-0.38
2012	A	0.00	N/A	0.24	-0.68
	B	-0.02	-1.84	-0.26	-0.26
	C	-0.21	N/A	-0.38	0.07
2013	A	0.07	-0.58	-0.56	-0.55
	B	-0.21	-0.17	-0.05	0.08
	C	-0.13	N/A	-0.02	-0.29
2014	A	0.00	-1.58	-0.13	-0.56
	B	0.14	-0.38	-0.20	-0.37
	C	0.09	-1.20	-0.31	-0.26

Discussion

RQ1 (correlation between average grades in certificate and degree)

The correlation coefficient is 0.4903, shows there is a weak to moderate relationship between the certificate grade and the degree grade. As it can be seen in Figure 3, the scatter of the data and upwards trend from left to right of the data points indicate a weak positive relationship between average grade on the certificate and average grade on the degree. The majority of the students who had passed in the certificate also passed in the degree, which means that success in the certificate can be seen as an indicator of success in the degree. The linear regression formula[3] means that generally students perform slightly worse in a degree paper than in the certificate though this cannot be said of all students and all papers. From the average grades of people who passed in both programmes it can be concluded that

[3] average degree grade= 0.54 * average certificate grade + 0.54

certificate students who got an average grade of B- or better in the certificate are also very likely to pass in their degree.

Even though it was shown that an average grade of B- or higher in the certificate indicates a pass in the degree, it must be noted that this conclusion is based on data from previous years, and, thus, established teaching methods. Changes that are undertaken to current teaching methods and paper content will affect this result.

RQ2 (certificate students' performance over time)

There was no clear picture when comparing combined students' performance in a degree over time. A rough distinction could be seen when the students were split by their average certificate grade; however, this distinction was not very strong. The results confirm what was already found in RQ1: In general, students who had an A or B average in the certificate also pass their degree papers, whereas students who had a C average in the certificate are more likely to fail. Table 4 shows that a certificate B student has the most stable results and is most likely to improve performance over time.

Regarding the enrolment time, it was found that around three in five students were enrolled for the full time of their degree (three or more years). The other students only stayed enrolled for one or two years, which means that they did not finish their degree. It was also found that most students who did drop out did so after being enrolled for one year rather than for two years. The reason for this could be that after two years, the student has usually completed more than half of their degree, so the motivation to complete the last year as well is higher.

The distribution of students dropping out before completing the degree is similar in all three degrees. For the certificate this means that students should be encouraged to continue their education regardless of which pathway they pursue. This could for example be done by offering courses on how to succeed in further studies (from organisational and academic perspectives), or by facilitating regular meetings of students who progressed from the certificate to the same degree in order to create a supporting environment for them.

The reasons for students dropping out are various. They are not recorded for privacy reasons, so no further analysis can be done on this issue at the moment. In future research, the reasons for students to drop out early could be analysed by interviewing them or doing a survey, and then taking countermeasures accordingly.

RQ3 (performance of certificate and non-certificate students)

When looking at Math., Comp., and Eng. Degrees combined for the period looked at Certificate students performed slightly worse in subsequent degrees compared to students who had not previously enrolled in the certificate (average certificate student degree grade 1.98, average non-certificate student degree grade 2.20). There has been an overall

decrease of this gap over the period covered, meaning there has been an increase in preparation from the certificate for later performance on the degree.

When we look at the degrees separately it shows a slightly different story. First there are 3 cases when the certificate students outperformed the non-certificate students on the degree. Eng. in 2011, Math. in 2013, and Comp. in 2014. Second there is an increase over time in performance of the Math. and Comp. certificate students. Third there is a decrease in performance of the Eng. Certificate students over time.

The results shown in Table 5, Table 6, and Figure 11 suggest that before 2011 the certificate was geared towards Eng. students. From 2011 onwards there has been a move for the certificate to better service all three post certificate degrees – Math., Comp. and Eng. The results suggest the preparation of the Math. and Comp. students is good. Preparation of the Eng. students from the certificate needs to be improved. Engineering certificate papers need to reviewed and modified for relevant content for preparation for Eng.

The percentage of degree students with certificate background who passed in the degree was 0.8 higher than for students without certificate background. This finding can be interpreted as an indicator for the success of the certificate. Since all students were subject to the same grading criteria in their degree, students coming from the certificate are to some extent better prepared than the other students. The reason for this can be found in the different circumstances of students who enrolled in the certificate and of those who did not. The certificate is usually used as an entry into university by people who have not previously achieved the requirements for university entrance. Some of these students have underperformed at school, some are coming to university after a break away from study, or have studied in different school systems around the world where they have not achieved university entrance in New Zealand. On the other hand, people who directly start a degree usually have successfully achieved university entrance from school. People who enrol in the certificate usually see a strong need for doing so. It involves extra cost and time on the way to their goal. This could contribute to a greater commitment to complete their degree opposed to a non certificate student.

RQ4 (predictive certificate papers)

The five papers that were found to be most predictive were Academic Literacies for Computing & Mathematical Science, Foundation Programming, Foundation Logic Skills, Foundation Physics B, and Foundation Physics A. All of these papers, except for Physics B (which is only part of the Eng. pathway), are common for all of the students and, therefore, are most frequently taken, which means that most data was available for these papers. In addition, Foundation Physics A and B are both perceived as difficult by the students, which could be an indication that they reflect future success in a degree relatively well. The mapping from original paper grades to original degree grades appeared to achieve the highest information gain; however, this may be due to over-fitting since the dataset is relatively small.

The best classification result was achieved when trying to predict from pass/fail grades on the certificate to pass/fail grades in the degree. However, without applying any techniques to improve the performance of the selected algorithm, it performs very badly with only around 70% correct classification. In future work, a more detailed analysis of predicting degree grades based on certificate grades can be done. This would include the comparison of several classification algorithms, performance optimisation, and also the comparison of different feature selection algorithms (other than the Information Gain method).

RQ5 (performance of native and non-native speakers of English)

Across the certificate and all three degrees, there are no major differences between the performance of students who are native speakers of English and students who are not native speakers of English. There is high fluctuation in the data for AK1271 (Math.), because there were only very few students who progressed from the certificate to that degree in 2012 and were native English speakers, and all of them performed very badly (most of them actually dropped out so they did not complete their papers). The results for this research question mean that no further action has to be taken in order to support either group.

Conclusions and future work

The overall number of students progressing from the certificate to a degree rose steadily over time. There is a positive relationship between achievement on the certificate and achievement on the degree. The majority of the students who passed in the certificate also passed in the degree. In general, it can be said that a student who graduates from the certificate with an average of B- or better will also pass their papers in a subsequent degree. There is no clear trend of the certificate students' performance in their degree over time. On the other hand, certificate students have a slightly higher pass rate on their degree than non-certificate students, and the general trend of their degree grades is upwards. The results suggest the certificate has been better preparing students overtime though, in particular Math. and Comp. students, though there is a need to re-review papers for better preparing Eng. students. The paper that was found to indicate further success in the degree was *Academic Literacies for Computing & Mathematical Science*, however, since this is a core paper in the certificate, its predictability is still low. There was no considerable difference found when comparing the results of native and non-native speakers of English.

Recommendations:

- Review of engineering papers for addition of relevant engineering content

- Graduating full year Cert students with average grades of C+ should be given extra support if they are given entry onto a degree.

- Students eligible for consideration for stair casing from the Cert onto an engineering, computing or mathematics degree should also have passed all four first semester with a B+ with two of these papers being LSKL401 and MATH401.