
Managing “At Risk” students and pass rates with SPSS

Christo Potgieter

Wintec
Tristram Street
Hamilton
christo.potgieter@wintec.ac.nz

Wilfred Greyling

Wintec
Tristram Street
Hamilton
wilfred.greyling@wintec.ac.nz

Bruce Ferguson

Wintec
Tristram Street
Hamilton
bruce.ferguson@wintec.ac.nz

This quality assured paper appeared at the 1st annual conference of Computing and Information Technology Research and Education New Zealand (CITRENZ2010) incorporating the 23rd Annual Conference of the National Advisory Committee on Computing Qualifications, Dunedin, New Zealand, July 6-9. Samuel Mann and Michael Verhaart (Eds).

Abstract

Both the Tertiary Education Commission and the Ministry of Education are now moving fast to lower ITP funding for courses with low pass rates. It is imperative that academics understand student progression towards individual outcomes and pass rates of courses. Pro-active management with powerful tools for improved student outcomes will be very advantageous in future. Following work by McCarthy (2005) and Scott (2005) on student retention, completion and progression, Potgieter, Ferguson and Robertson reported during 2009 on several experiences regarding student outcomes. The authors have since developed and now demonstrate a sophisticated tool for general use by proactive programme managers, specifically a statistical model to predict success of the pass/fail categories for IT students. This paper provides opportunities for further study and reference for reflective practice by programme Managers at ITPs.

Keywords

Completions, Pass Rates, SPSS, At Risk students

Introduction

In recent years, the Tertiary Education Commission (TEC) has been promoting specific expectations regarding student outcomes (Tertiary Education

Strategy 2007, Tertiary Education Strategy Monitoring Information 2008, TEC to publish educational performance information, 2010). During March 2010, the newspaper NZ Herald contained two government announcements indicating changes to impact on tertiary institutions: Funding to be aligned to student outcomes and limitations to student loans. TEC expects of institutions to report on performance and initiatives in their business plans and so it is essential that academics improve their understanding of student progression towards individual outcomes and corresponding pass rates as part of their proactive management practices.

At WINTEC, executive views mention a five-year aspiration for reputation of "Quality and Student-Centeredness" (Our Direction, 2009). One of the eight key strategic themes is "Quality and Outcomes", stating amongst other things, that the institution pursues "High levels of student satisfaction and completion rates". Further commitment is reflected in the institution's Academic Direction (2010) –

"Student retention, especially in the early stage of their study programme is a key strategy in improving student completions. We will improve this by having a more explicit and integrated school and support service approach to retention and ensuring we have the processes in place to monitor and report on progress made."

Institutional practice is therefore that attendance monitoring is closely monitored by the Student Experience unit during especially the first half of each semester in order to minimize early fall-out. While attendance does not necessarily imply successful

course completion, it appears to be recognized widely as an important contributing factor, as is also evident from research by Potgieter and Ferguson (2009) regarding international students at one institution.

In addition, Programme Managers of the School of IT assess progress of students after especially the first two assessments and discuss possible action with tutors and Student Experience services for individual support. Finally, the tutors of the School of IT observe behaviour of students and of course academic performance to identify those with significant weakness in participation, response indicating comprehension and, of course, passing assessments! Action is taken with and/or without other role players.

These and other activities became known internally as the Management of 'At Risk' students. From 2009, staff immediately focused heavily on class attendance, but only gradually increased management activity. With attendance monitoring running through the semester (albeit with slight reduction after first six weeks), staff from the School of IT wondered about the focus periods during a semester regarding academic performance.

The proposition was that the final outcome of student performance (pass/fail and final mark at end of semester) is already clear halfway through the course (semester). This implies that special intervention per student could be futile later on in the semester. It also implies that action is required very early in the semester, and therefore very close scrutiny of academic progress is required over the first half of the course.

The added benefit of knowledge about the likely final outcome for a student early on during the semester is

also that one could avoid "last minute crisis" attempts to help students pass, if there is evidence of very poor likelihood of success and therefore very low return on investment. It sets some sort of "point of no return".

There is, of course, very severe danger in this view in that students could become "doomed to fail" when tutors strongly believe in their inability to influence outcome after a certain point. One should take note of this significant problem of "self-fulfilling prophecy" and one needs to work on avoiding that possibility. However, the fact remains that if evidence shows very low probability of success, one ought to be realistic about why intervention attempts are still being made, the investment with low/no return and implications for students.

But first it would be useful to understand patterns of student progress and have some prediction abilities. This involved analysing historical data without consideration of the interventions that were made. It shows what the typical progression of students is with current intervention practices. Obviously one can therefore not consider the impact of interventions, but the argument is that at this stage one does not even know what student progression typically is! Implications of not considering specific interventions are revisited at the end of the paper.

Planning empirical explorations

Before the empirical stage, the following research questions were formulated:

1. Are the first three assessments indicators of the likelihood of final course outcome?

2. Is final course success best described by pass/fail or the final mark itself (described as the "*best outcomes parameter*")?
3. How accurately can one predict the final outcome after each of the first, second and third assessments?
4. Re exceptions: Are the number of students recovering from initial failure or failing after initial success, significant?

Resulting Null-hypotheses (H_0):

- 1) Re overall model and "*best outcomes parameter*"
 - a) Model with A1, A2 and A3 predicts final mark satisfactorily
 - b) Model with A1, A2 and A3 predicts pass/fail satisfactorily
- 2) Re only first two assessments
 - a) Model with A1 and A2 predicts "*best outcomes parameter*" satisfactorily
 - b) Model with only A1 predicts "*best outcomes parameter*" satisfactorily

One will also want to know whether exceptions to overall model prediction (incorrect predictions by the model) are acceptably low. The statistical procedure will be Discriminant Analysis with the following analyses to support it: Multivariate Analysis of Variance (MANOVA) to explore the interactions among predictor variables and Analysis of Variance to establish whether a statistically significant difference obtains for the variables, split on the basis of the pass-fail categories (Tredoux & Durheim, 2002).

But it was also the intention to learn about statistical tools to create something that could be used as an instrument in future. So SPSS was used to provide a richer framework for future use and to create reference material for colleagues. Finally, the intention was to learn about the organisation to make first observations about current practices for improvement. This is part of a philosophy of reflective practice. The project therefore has similarities to *exploratory* case studies (Yin, 1993) but is also *intrinsic* (Stake, 1995) since the authors have a personal interest in the situation.

Exploring empirically

The School of IT's degree level programmes BInfoTech and GradDip IT were analysed. These students share all but one of their classes, so it gives us an opportunity to compare student cohorts. Importantly, most students on GradDip are from overseas (mostly India) and all of them have degrees, of which a significant proportion are IT degrees. BInfoTech also has a significant number of international students (average about 25%), which contributes to research regarding outcomes for international students.

Data was extracted from the institute's student database for 2009, semesters one and two, as well as Summer School. (Summer School is used for students following a pathway from the diploma (DipICT) into the degree and for one of the intakes of GradDip IT).

Unwanted columns were removed, such as student identification and assessment methods. The file contained 1480 records. Data was cleaned data up by removing a considerable number of lines:

1. 1110 lines remained after removing fields for Final Mark that are empty or not Pass/Fail (Empty fields occur in cases of very early student withdrawals, students changing enrolment, errors in enrolment administration and credit transfer for specific modules).
2. 621 lines remained after removing records with no mark for any one of the first three assessments (empty fields occur when there was an administrative error, or when students changed enrolment late in the semester).

This left us with a total of 621 records, which happen to be from 204 different students, which is about 95% of the 215 students in the full data set.

Analysis: About the SPSS calculations

The sample was big enough for assessing one year. Table 1 shows some high level descriptive statistics, some of which are of concern:

1. The mean assessment 1-3 marks are nearly the same and all just above the pass level of 50% - this means just under half of the students fail during assessment.
2. However, the mean final mark is significantly higher and with much smaller standard deviation than that of the individual assessments.

Table 1: Descriptive Statistics

	Mean	Std. Deviation	N
Fin_Mark	66.234155	15.3496691	621
Ass1M	53.336634	27.2994170	621
Ass2M	51.405990	27.8364583	621
Ass3M	50.165354	26.8254017	621

Before exploring possible fundamental reasons for a significant difference between final mark and the first three assessments, it is important to verify the data again. It was found that there is good reason to remove more records from the sample:

1. Courses for which marks for only two assessments are recorded (e.g. Special Topic with only a presentation and report).
2. Courses not using all of the first three assessment fields for results (i.e. using some of the remaining five fields).
3. Courses where assessment marks are not recorded in percentages (appears to be the practice of a particular tutor).
4. Students missing results for some assessments (ie final mark calculated on fewer assessments than for peers, for whatever reason).
5. Combinations of the above caused by situations where marks of several assessments were combined before loading it.

The file size then drops by 107 to 513 lines, which is still from over 90% of the students and therefore an acceptable sample size.

Table 2: Descriptive Statistics

	Mean	Std. Deviation	N
Fin_Mark	66.274951	14.9166168	513
Ass1M	58.714191	24.9770898	513
Ass2M	55.877875	25.5781867	513
Ass3M	55.691481	24.5680650	513

It is noticeable from Table 2 that the mean for Assessments 1-3 is now higher, but it remains notably far below the mean of the final mark. Also noticeable is that standard variation of Final Mark remains much smaller than that of Assessments 1-3. One would have to pose the question why this happens. Anecdotal evidence from browsing the student data base indicate that for many courses, the first three assessments contribute significantly less than 40% of the final mark and that marks for assessments increase later in semester. This matter ought to be investigated further.

Analysis: Towards testing hypothesis

What does the data say about the possibility of predicting outcome for the student? Stepwise multiple regression was conducted using final mark as dependent variable, with the following predictors: Assessment 1, Assessment 2, Assessment 3, Qualification and Semester. Significance of the model is within boundaries:

Table 3: Model Summary from SPSS

Model	R		Adjusted R Square	Std. Error of the Estimate	Change Statistics				
	R	Square			R Square Change	F Change	df1	df2	Sig. F Change
1	.551 ^a	.304	.302	12.4586503	.304	222.954	1	511	.000
2	.515 ^b	.370	.375	11.7901720	.074	60.500	1	510	.000
3	.534 ^c	.403	.399	11.5641649	.025	21.129	1	509	.000
4	.540 ^d	.409	.405	11.5075318	.007	6.013	1	508	.015
5	.546 ^e	.417	.412	11.4427391	.008	6.778	1	507	.009

a. Predictors: (Constant), Ass3M
b. Predictors: (Constant), Ass3M, Ass1M
c. Predictors: (Constant), Ass3M, Ass1M, Sem
d. Predictors: (Constant), Ass3M, Ass1M, Sem, Ass2M
e. Predictors: (Constant), Ass3M, Ass1M, Sem, Ass2M, Qual
f. Dependent Variable: Fin_Mark

Results for the stepwise procedure showed that sequencing for entering and retaining predictor variables in the model is as follows: Assessment 3, Assessment 1, Semester Assessment 2 and Qualification. These are the variables that contribute most to the model explaining final results from variables, ie the variables that most closely predict the final outcome for a student.

It seems that a tentative start may have been made in selecting predictors for a model - the proposed model explains 41.2% of the final mark. Note that the strongest model for prediction contains Assessment 3, Assessment 1, Semester, Assessment 2 and Qualification - Assessment 2 therefore falls to the side in favour of Semester. This observation is revisited during reflection at the end of the paper.

Findings on the analysis of variance (ANOVA, table 4) relate to overall significance of the predictor model,

answering the question whether the four predictors allow one to predict learner scores on each dependent variable. F values are significant at probability level of 0.000 which signals that these variables allow its use for predicting final marks of students.

Table 4: ANOVA results for models

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	34606.413	1	34606.413	222.954	.000 ^a
	Residual	79316.381	511	155.218		
	Total	113922.794	512			
2	Regression	43028.625	2	21514.313	154.770	.000 ^b
	Residual	70894.169	510	139.008		
	Total	113922.794	512			
3	Regression	45854.270	3	15284.757	114.296	.000 ^c
	Residual	60060.524	509	133.730		
	Total	113922.794	512			
4	Regression	46650.594	4	11662.649	88.069	.000 ^d
	Residual	67272.200	508	132.426		
	Total	113922.794	512			
5	Regression	47538.102	5	9507.620	72.613	.000 ^e
	Residual	66384.693	507	130.936		
	Total	113922.794	512			

a. Predictors: (Constant), Ass3M
b. Predictors: (Constant), Ass3M, Ass1M
c. Predictors: (Constant), Ass3M, Ass1M, Sem
d. Predictors: (Constant), Ass3M, Ass1M, Sem, Ass2M
e. Predictors: (Constant), Ass3M, Ass1M, Sem, Ass2M, Qual
f. Dependent Variable: Fin_Mark

The standardized Beta coefficients (Table 5) indicate how each variable contributes to the prediction model. The Beta coefficient signals change to dependent variable (Final Mark) from change in a predictor variable. It can be seen that Assessment 3 and Assessment 1 have more impact than other variables. For the t values, the higher the value the more significant the impact of the variable in predicting

learner scores on the dependent variable, Final Mark. Again Assessment 3 and Assessment 1 stand out as having the most significant impact.

Table 5: Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
	B	Std. Error	Beta	Std. Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1 (Constant)	17.638	1.361			12.957	.000	14.952	20.324					
	.335	.022	.501		14.952	.000	2.91	3.79	.501	.501	.501	1.000	1.000
2 (Constant)	42.122	1.473			28.606	.000	30.229	45.015					
	.212	.025	.382		9.305	.000	1.85	2.81	.551	.381	.326	.722	1.385
3 (Constant)	36.230	1.931			18.762	.000	32.437	40.024					
	.232	.024	.382		9.472	.000	1.84	2.80	.551	.387	.320	.722	1.389
4 (Constant)	35.046	1.981			17.687	.000	31.155	38.938					
	.183	.025	.307		7.448	.000	1.36	2.92	.522	.314	.254	.685	1.460
5 (Constant)	31.013	2.506			12.375	.000	26.090	35.937					
	.210	.026	.347		8.159	.000	1.60	2.61	.551	.342	.278	.643	1.556

The question now is how well these variables predict a Pass/Fail, as well as the so-called incorrect predictions. Outcomes are predicted correctly for 75% of course enrolments – see Table 6. A pass is forecast but then the student fails, fortunately for only 1.6% of course enrolments.

Table 6: Incorrect predictions

		PREDICTED			Total	
		Fail	Pass			
ORIGINAL	Count	Fail	48	8	56	11.0%
		Pass	120	337	457	89.0%
	%	Fail	85.7	14.3	100.0	
		Pass	26.3	73.7	100.0	
%	Fail	9.4%	1.6%	100%		
(Overall)	Pass	23.4%	65.7%	100%		
		32.8%	67.3%			

25% of outcomes predicted incorrectly (23.4% + 1.6%).
 Remarkably 23.4% predicted to Fail but then they eventually Pass.
 Note that only 1.6% predicted to Pass who then unfortunately Fail.
 Interestingly 89% of students who continued to the end, Pass.

It is also clear that 23.4% pass after a fail was predicted from the model. This change from fail to pass partially reflects an increase in the marks occurring after the third assessment.

Regarding previous reference to a possible “best outcomes parameter” (ie. Final mark vs Pass/Fail), one sees that the model explains 41.2% of the final mark and 75% of the Pass/Fails. It would appear the model is stronger in its ability to predict Pass/Fail rather than the final mark. The ability to predict Pass or Fail for three quarters of the students is noteworthy.

Analysis: Hypotheses and Research Questions

While the above interpretation is useful and interesting, a scientific interpretation revisiting hypotheses formulated earlier on is required:

1. Model with A1, A2 and A3 predicts final mark satisfactory: YES (41.2%).
2. Model with A1, A2 and A3 predicts Pass/Fail satisfactory: YES (75%).
3. Model with A1 and A2 predicts Pass/Fail (“Best outcomes parameter”) satisfactory: NO, because it lacks the most important predictor – Assessment 3
4. Model with only A1 predicts Pass/Fail (“Best outcomes parameter”) satisfactory: NO, because it lacks the most important predictor – Assessment 3.

Also, exceptions to overall model prediction (incorrect predictions by the model) are acceptably low, considering good statistic fit, ability to forecast for three quarters of students and the fact that the majority of incorrect predictions still turn out positive for students when they pass rather than fail. (This

matter, however, relates to the noticeable increase in final marks, which needs to be investigated).

Returning to the research questions posed earlier:

1. Are the first three assessments indicators of the likelihood of final course outcome? Yes, they are generally good indicators.
2. Is above final course success best described by pass/fail or the final mark itself (described as the "*best outcomes parameter*")? Pass/Fail best describes success, being accurate for 75% of students with the first three assessments.
3. How accurately can one predict the final outcome after each of the first, second and third assessments? Prediction is relatively weak until Assessment 3.
4. Re exceptions: Are the number of students recovering from initial failure or failing after initial success, significant? The number is relatively low and in favour of the student.

One could therefore conclude that Multivariate Analysis of Variance (MANOVA) of the first three assessments could accurately predict Pass/Fail for three quarters of the students. This, of course, assumes that the currently unexplained increase of marks over rest of the semester continues in future.

However, from the above it appears that the ability to predict final marks and/or Pass/Fail without Assessment 3 is rather weak. Since Assessments are recorded in order of occurrence, the timeline for more accurate prediction extends accordingly. In practice it means predictions early on in the semester are fairly inaccurate, especially since there is a noticeable

increase after the first three assignments towards the final mark.

Possible improvements

There are several opportunities to improve the model by reviewing data and parameters:

1. Investigation of the increase in marks after first three assessments – it might be very explainable, considering integration of learning, student development and intervention.
2. The presence of two or fewer assessments for some courses ought to be followed up – apart from the obvious academic consideration, are changes required to data processing with SPSS?
3. Comparison with other years, especially considering the sample year contained industrial action by staff, possibly impacting on patterns.
4. Since Semester makes an impact in the model, it ought to be explored for explanation, even with the relatively small Summer School.
5. Explore the possibility of different patterns of progression during a semester as the students progress over a multi-year study programme.
6. For generalisation the range of qualifications ought to be increased in the organisation (across academic fields) and/or inter-institutional.
7. Considering the high proportion of International students as a possible significant sub-group, classifying origin for predictor is advisable.
8. Some international students already hold partial or full IT degrees, which could influence the pattern of marks, so classification might be advantageous.
9. Consideration of interventions per course taken during the semester – this might be the cause of increase of marks towards final mark.

10. Improvements to future extracts from the students database for sampling to exclude non-relevant enrolments.
11. Formally reviewing that SPSS analysis has been done satisfactorily from the perspective of statistical practice.
12. Capturing results as quickly as possible in the semester, in order to enable earlier prediction for use during planning of intervention.
13. Addressing the practice, by tutors of a few courses, to use other than percentages for some of the assessments on a course
14. Considering how to approach qualifications using competency-based instead of performance based assessments.
15. Using historical data and SPSS, performing prediction for each student and comparing to a list of "At Risk" students drafted by academic staff.
16. Reviewing assessment schedules to do and record two assessments early in the semester, in order to inform intervention decisions.
17. Building mechanisms (procedures, software, manuals, etc) to simplify data processing and interpretation for Programme Managers.
18. Exploring simpler calculations with Excel to overcome or avoid the complications of SPSS and student administration systems.
19. Literature exploration regarding influential matters from the perspective of learning in the ICT discipline.

In summary, there is a range of opportunities available for incremental refinement of the current approach. The above list is reflective of possibilities from the perspective of information systems, general and ICT-specific teaching and learning, organization development, academic management, research methodology and so on.

Conclusion

The concept of using a sophisticated statistical tool on unverified historical data in very limited time range is challenging. Data is often messier than anticipated, statistical analysis is complex and organization practices not focussed to add value. This is a good learning experience to plan more mature projects.

Statistical analysis shows that the final outcome for most students is already set by the third assessment, but that predictions after the first two assessments are premature as many students improve dramatically.

This observation of statistics and experience of the projects enables the authors to more comprehensively plan future initiatives to improve student outcomes as expected by stakeholders.

Acknowledgements

Appreciation is expressed to the Dean of Faculty for approval to publish this organisational research.

References

- Academic Directions 2010*. (2010). Faculty Planning document, WINTEC.
- Our Direction*. (2009). Presentation by Chief Executive of WINTEC, May 2009.
- First Year Experience (FYE) Retention strategy*. (2008) Planning document of PERSON-A, Manager: Student Experience, WINTEC
- Code of Practice for the Pastoral care of International Students*. (2003). Ministry of Education, New Zealand. Retrieved March 9, 2009 from <http://www.minedu.govt.nz/educationSectors/InternationalEducation/ForProvidersOfInternationalEducation/CodeofPracticeforInternationalStudents.aspx>
- McCarthy, C (2005) *Encouraging Student Retention: a study of student retention practices*. Proceedings of the 18th Annual NACCQ Conference, accessed 12 Mar 2009 at www.naccq.ac.nz/conference05/proceedings_04/mccarthy_retention.pdf.
- NZ Herald *Tertiary education funding to be performance linked* (2010). Retrieved March 15, 2010 from http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10630881
- NZ Herald *Time-limit plan for student loan test* (2010). Retrieved March 15, 2010 from http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10631773
- Potgieter, C., Ferguson, B. (2009) *An Analysis of Completion and Retention of a Graduate Diploma*. Proceedings of New Zealand Applied Business Education Conference, Rotorua, 28-29 Sep 2009.
- Potgieter, C, Ferguson, B. (2009). *Managing International Students Attendance with Consideration of Completion and Satisfaction*. Proceedings of the 22nd conference NACCQ, 10-13 July 2009, Napier, New Zealand. pp 87-90.
- Potgieter, C., Ferguson, B. (2009). *Analysis of completion and retention of a graduate diploma*. Presentation at 22nd conference NACCQ, 10-13 July, 2009, Napier, New Zealand. Retrieved September 29, 2009 from http://hyperdisc.unitec.ac.nz/naccq09/programme/programme_saturday.html.
- Potgieter, C., Robertson, G. (2009). *Student success: Overcoming quality systems and data problems*. Presentation at 22nd conference NACCQ, 10-13 July, 2009, Napier, New Zealand. Retrieved September 26, 2009 from http://hyperdisc.unitec.ac.nz/naccq09/programme/programme_saturday.html.
- Student Experience Journey*. (2009). Planning document of Jo Shortland, Manager: Student Experience, WINTEC, 2009
- Our Directions*. (2009). Staff presentations by Chief Executive, WINTEC.
- Scott, D. (2005) *Retention, completion and progression in tertiary education in New Zealand*, in Journal of Higher Education Policy and Management, Vol. 27, Issue 1, pp3-17, accessed 15 August 2008 at http://www.educationcounts.govt.nz/publications/tertiary_education/retention_completion_and_progression_in_tertiary_education_in_new_zealand
- Tertiary Education Strategy Monitoring Information Website, Retrieved August 15, 2008 from <http://wiki.tertiary.govt.nz/~TESMon/MonitoringInformation/CompletionRetentionAndProgression>
- Tertiary Education Strategy 2007 - 2012* (2007). Wellington: Ministry of Education, Retrieved August 15, 2008 from <https://staff.wintec.ac.nz/corporatedocs/index.htm>
- TEC to publish educational performance information*. 2010. Retrieved March 15, 2010 from <http://www.tec.govt.nz/About-us/News/Media-releases/Media-release-TEC-to-publish-educational-performance-information/>
- Tredoux, C., Durheim, K. (2002). *Numbers, Hypotheses & Conclusions. A courses in statistics for the social sciences*. UCT Press: Cape Town.
- Yin, R. (1993). *Applications of case study research*. Newbury Park, CA: Sage Publishing.
- Stake, R. (1995). *The art of case research*. Newbury Park, CA: Sage Publications