Managing "At risk" business students: Statistical analysis of student data profiles

Christo Potgieter	Wilfred Greyling	Vida Botes
Wintec	Wintec	Wintec
christo.potgieter@wintec.ac.nz	wilfred.greyling@wintec.ac.nz	vida.botes@wintec.ac.nz

Abstract

In this paper a multivariate model is developed to predict success or failure, helping to identify potentially at-risk students, based on the first two assessments. A robust discriminant or regression function will allow identification of students who obtained scores on these assessments that would put them at risk of failing the course. This paper uses stepwise discriminant analysis and multiple regression analysis to explore the data profiles for two closely related degree-level qualifications in Business Studies at the same institution. Evidence for the Business Studies qualifications suggests that a robust prediction model is possible. Therefore one could reflect on implications for more efficient teaching and assessment policies and practices. In line with the theme of the NZABE conference to question fundamentals", one could ask whether exams are really required?! Although various kinds of changes seem possible, there are many pitfalls that still need to be explored.

Key words: Student outcomes, completions, pass rates, examination, predictors, predictive efficiency

Literature review

The Tertiary Education Commission (TEC) has clearly stated that tertiary institutions have to improve student outcomes (Tertiary Education Strategy 2007, Tertiary Education Strategy Monitoring Information 2008, TEC to publish educational performance information 2010). Recently the NZ Herald (2010) announced that government funding for institutions would be adjusted on the basis of student outcomes. Tertiary institutions, it has been reported, need to set targets in their business plans and report on achievement, specifically student pass rates and progression.

At WINTEC, the executive refers to a five-year window period to pursue and build a reputation for "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 will pursue "high levels of student satisfaction and completion rates". The institute's commitment is reflected in its 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."

Apart from attendance monitoring as a specific strategy among many to manage "at risk" students, vocational tutors and Programme Managers generally act on results from assessments. It is normal practice for tutors to pay more attention to students who fail assessments. It is generally believed that the final outcome of student performance (pass/fail and final mark at end of semester) can still be changed after the early assessments. Special intervention per student after the first assessment then seems indicated and beneficial. In order to get the benefits of early intervention, learners' academic progress in the first half of the course ought to be scrutinised very closely.

Of course, if reliable information is available about student performance over the first half of a programme, sustained intervention could be planned, avoiding "last minute crisis" attempts to support students to pass. The assumption is that it is still possible for the student to pass in spite of having failed a significant portion of the assessments in the semester! One could be tempted to argue on the basis of personal experience that attempting to change the outcome for any of these at-risk students is a lost cause – "If you fail at the start, you fail at the end". If evidence of such fatalistic outcomes exists, a case might be formulated against interventions because they might be poor return on investment and a waste of resources.

A significant danger is that tutors may adopt the view, at a certain stage of the semester ("generally when students have crossed the point of no return"), that students are "doomed to fail" because, as tutors, they are no longer able to influence the outcome. However, the notion of a "self-fulfilling prophecy" would only be relevant if there was actually still a chance of success, i.e. if students who fail initially, succeed in the end. While there would always be exceptions to the rule, the question is how many students fit this category. With this information the return on investment for interventions might be explored, which is important, considering financial pressures experienced by tertiary institutions.

So, it would be useful to have some insight into patterns of student progress, and how initial assessments allow us to predict the likely outcomes. However, focusing on historical data only would be limiting and reductionist, reflection on current intervention practices makes sense. This is done in a section close to the end of the paper. At this stage it seems prudent to explore existing data. Since most tutors implement some form of student intervention, assessment results and final outcomes already reflect the effect of historical (likely current) practices. These practices need to be captured too.

Research questions

Before formulating research questions, the assessment structures of all the modules in the two qualifications were studied. Modules with assessment structures that deviated from the pattern of two initial formative assessments and a final examination mark were then excluded. Thus, the criterion for inclusion was that a module should have at least two initial assessments and a final examination scores. With the targeted modules identified, the following research questions were formulated:

1. To what extent do the first two assessments in a module allow us to predict final course outcome?
2. How accurately can one predict the final outcome after each of the first and second assessments?
3. How significant in size is the number of students whose final results differ from initial assessment scores?
4. Is the level of incorrect predictions from a specific statistical model acceptably low?
5. To what extent do the initial two assessments discriminate between the selected two qualifications?

The project also offered the opportunity to consider the potential of the statistical package and its tools to assist us in designing prediction models for future use by management. The software package SPSS provides a powerful framework, is used widely and is available locally; for these reasons, the package became the vehicle for analysing student data. This research project was an opportunity to identify patterns in the student data that would promote wider institutional self-reflection. For this reason, consistency with the premises of *exploratory* case studies (Yin, 1993) were considered, prompting exploring programme-related and systemic issues.

Independent skilled observation is important for objective research so possible personal interest ought to be recorded. The authors of this paper have personal interest in the data as managers and have some training but no expertise in the advanced statistical methods used. For this reason data analysis and interpretation are written (like case studies often are) in the first person as personal dialogue showing reasoning of the participants.

Data processing

Our aim was to use existing organisational records and data sets to develop a statistical model for predicting learner success on the basis of three independent variables (i.e. Assessments 1 and 2, as well as Qualification). We argued that these variables could inform decisions to support learners at different stages of their courses in the qualifications included in the study. Put differently, we reasoned that some or all of these variables could be used as predictors of learner success. Once the predictive efficiency of the variables in the model was established, we would be in a position to target vulnerable learners for support to retain them in courses.

In summary, our statistical analysis explored to what extent the three predictors (Assessments 1 and 2, as well as Qualification) explained the variance on the classification variable, Pass/Fail, in a discriminant analysis, or the variance on the dependent variable Final Mark in a multiple regression analysis. Our purpose was to develop a statistical model that would allow us to provide at-risk learners with appropriate and timeous support.

Student performance on the School of Business's degree level programmes BBS and GradDip was analysed. These students are distributed over a wide range of modules within the two qualifications. We labelled both the qualifications (BBS = 1; GradDip = 2) and each module in the database so that in the longer term we could focus on student performance on configurations of modules within these qualifications. All of the GradDip students already have degrees. Most students on GradDip are from overseas (mostly India) while a significant number of international students (average about 40%) are registered for BBS – for purposes of this paper, we do not distinguish between domestic and international students. We dealt with the full cohorts registered for the two qualifications to establish whether specific patterns would emerge for the full cohort from the analysis.

We extracted student performance data from the institute's student database for 2009, covering semesters one and two, as well as Winter/Summer Schools. The latter were implemented in 2009 mostly for students from China under pathway agreements. The data were processed, which meant that we retained only those student data sets where scores for Assessment 1, Assessment 2 and Final Result (Pass or Fail) had been recorded.

Analysis

To respond to the research questions, two statistical procedures were used: a stepwise discriminant analysis and a stepwise linear regression. The stepwise discriminant analysis was deemed appropriate because the predictors were used to predict membership of two discrete categories (Pass=1 Fail=0). In this procedure, the variables are entered in sequence. When a specific predictor is entered, its contribution to the model is assessed. If the predictor explains a significant proportion of the variance on the classification variable, it is retained in the model. When a variable is retained in the model, the variables that are already included in the model are again assessed to see whether they continue to explain a significant proportion of the variance. If a variable in this emerging configuration of variables no longer explains a significant proportion of the variance, it is discarded. The end result is that this iterative process will yield the smallest, yet most efficient set of predictor variables.

The same iterative process applies to the stepwise linear regression. However, the difference is that this particular procedure yields a set of variables to predict scores on a continuous variable (in this case, Final Mark).

Findings

First, we report the relevant descriptive statistics for the variables. In Table 1 are the descriptive statistics for the continuous variables measured in the assessments and the final mark:

Table 1: Descriptive Statistics for group

	Mean	Std. Deviation	N
Fin_M	64.827	10.4930	1790
Ass1_M	61.479	17.4199	1790
Ass2_M	68.493	14.5925	1790

Interestingly, the means for the scores on the two assessments and final mark vary from 61.5 % to 68.5 %. The means are clustered together within a narrow 6% range which suggests that both the assessment practices in these programmes and student outcomes have been consistent across assessment opportunities.

In Table 2, we report the classification table produced by the stepwise discriminant analysis. It is clear from the table that the three variables allow us to classify the 83.3 % of the cohort of students correctly within either the pass or the fail categories. Thus, 299 out of 1790 students were classified incorrectly when the current prediction model, based on Assessments 1 and 2, as well as Qualification, was used:

Table 2: Classification Results

	PF category	Predicted Fail	Predicted Pass	Total
Actual	Fail	74	10	84
	Pass	289	1417	1706
%	Fail	88.1	11.9	100.0
	Pass	16.9	83.1	100.0

(a. 83.3% of original grouped cases correctly classified)

Ideally one would want a more efficient model to predict membership of pass/fail so that false positives and false negatives may be reduced. The false negatives (n = 289) are predicted negative, but turn out positive. Perhaps the implication is that existing support measures and delivery strategies aimed at promoting throughput are indeed effective. The false positives (n = 10) are predicted positive, yet turn out negative. In this study these were a negligible proportion of the whole.

Prediction model

The findings suggest a relatively efficient start in selecting predictors for a prediction model. The findings indicate that model 2, using the two assessments as predictors, explains 52.5 % of the variance on the dependent variable **Fin_Mark.** The findings are reported in Table 3:

Table 3: Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.609[a]	.371	.371	8.3244
2	.725[b]	.525	.525	7.2351

a. Predictors: (Constant), Ass1_M
b. Predictors: (Constant), Ass1_M, Ass2_M
c. Dependent Variable: Fin_M

Model	Change Statistics					Durbin-Watson
	R Square Change	F Change	df1	df2	Sig. F Change	
1	.371	1054.496	1	1788	.000	
2	.154	579.898	1	1787	.000	1.249

c. Dependent Variable: Fin_M

In addition, the ANOVA results in Table 4 show that the F value signals a significant difference between the P/F groups if they are compared on the basis of the three variables included in the prediction model:

Table 4: ANOVA[c]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	73072.060	1	73072.060	1054.496	.000[a]
	Residual	123900.744	1788	69.296		
	Total	196972.804	1789			
2	Regression	103428.166	2	51714.083	987.903	.000[b]
	Residual	93544.638	1787	52.347		
	Total	196972.804	1789			

a. Predictors: (Constant), Ass1_M
b. Predictors: (Constant), Ass1_M, Ass2_M
c. Dependent Variable: Fin_M

Beta and T values

In table 5, the standardized Beta coefficients indicate how each variable contributes to the prediction model. The Beta coefficient signals the number of standard deviations a dependent variable (in our case, **Fin_mark**), will change when the value of a predictor variable changes by a standard deviation. Thus, **Ass2M** and **Ass1M** have a significantly greater impact than **Qualification.** In fact, variable

Qual has been excluded from the model because its contribution to explaining the variance on **Final Mark** is negligible. For the t values, the higher the value the more significant the impact of the variable in predicting learner scores on the dependent variable, **Fin_Mark**.

Table 5: Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	T	Sig.
1	(Constant)	42.272	.722		58.555	.000
	Ass1_M	.367	.011	.609	32.473	.000
2	(Constant)	27.325	.883		30.959	.000
	Ass1_M	.270	.011	.449	25.482	.000
	Ass2_M	.305	.013	.424	24.081	.000

a. Dependent Variable: Fin_M

Coefficients[a]

Model		95.0% Confidence Interval for B		Correlations		
		Lower Bound	Upper Bound	Zero-order	Partial	Part
1	(Constant)	40.856	43.688			
	Ass1_M	.345	.389	.609	.609	.609
2	(Constant)	25.594	29.056			
	Ass1_M	.249	.291	.609	.516	.415
	Ass2_M	.280	.330	.594	.495	.393

a. Dependent Variable: Fin_M

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	(Constant)		
	Ass1_M	1.000	1.000
2	(Constant)		
	Ass1_M	.857	1.167
	Ass2_M	.857	1.167

a. Dependent Variable: Fin_M

Excluded Variables[c]

Model		Beta In	t	Sig.	Partial Correlation
1	Ass2_M	.424[a]	24.081	.000	.495
	Qual	-.002[a]	-.085	.932	-.002
2	Qual	.011[b]	.660	.509	.016

a. Predictors in the Model: (Constant), Ass1_M
b. Predictors in the Model: (Constant), Ass1_M, Ass2_M

Table 5: Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	T	Sig.
1	(Constant)	42.272	.722		58.555	.000
	Ass1_M	.367	.011	.609	32.473	.000
2	(Constant)	27.325	.883		30.959	.000
	Ass1_M	.270	.011	.449	25.482	.000
	Ass2_M	.305	.013	.424	24.081	.000

c. Dependent Variable: Fin_M

Correlations

Table 6 presents correlations between the dependent and independent variables. These results show moderate correlations between the two predictor variables and final mark. The correlation between qualification and final mark is a negligible 0.025.

Table 6: Correlations

		Fin_M	Ass1_M	Ass2_M	Qual
Pearson Correlation	Fin_M	1.000	.609	.594	.025
	Ass1_M	.609	1.000	.378	.044
	Ass2_M	.594	.378	1.000	-.012
	Qual	.025	.044	-.012	1.000
Sig. (1-tailed)	Fin_M	.	.000	.000	.144
	Ass1_M	.000	.	.000	.032
	Ass2_M	.000	.000	.	.299
	Qual	.144	.032	.299	.
N	Fin_M	1790	1790	1790	1790
	Ass1_M	1790	1790	1790	1790
	Ass2_M	1790	1790	1790	1790
	Qual	1790	1790	1790	1790

Results

Using stepwise regression analysis and stepwise discriminant analysis, the basic elements of a model for predicting learner performance have been identified. This finding is supported by the significance of the F value reported above (F =987.903, p = 0.000, Adjusted R^2 = 0.525). The predictors found to be significant in the model were (see percentages in tables): Assessment 2 scores (Beta=0.449, p=0.000) and Assessment 1 scores (Beta=0.422, p=0.000), with Qualification excluded from the model due to its negligible predictive efficiency. The Beta value explains the extent that the classification variable will change when the predictor value changes by a standard deviation. The large t-values in the table show that changes in these variables impact significantly on Final Mark.

Considerations

We felt that several interesting observations could be made about the statistical findings reported above. For example, although GradDip and BBS share several classes, the patterns of student performance do not reveal significant differences when these qualifications are compared. This is interesting considering the GradDip students have already completed several years of undergraduate study. This is also valuable to know since each qualification has a significant proportion of overseas students from two international regions in particular, namely India and China.

Although the means of the independent variable (Final Mark [65%]) and the assessment variable means [69% and 62%] occur within a narrow band, we have little evidence to suggest that a dramatic improvement in student performance did (or did not) occur. We could begin to define specific metrics-based outcomes. For example, we could argue that a significant improvement would be if the group scores improved by the standard deviation for the first assessment. We could argue that the mean of 61.5% is the starting point, and if we add the standard deviation of 17.5, we have a target: a mean of 79%. Or if we take half the standard deviation of 8.8, we could perhaps set a more achievable objective of 70.5% [61.5 + 8.8 = 70.3 rounded off to 70.5]. We could then direct our strategies at achieving SMART objectives (Specific, Measurable, Attainable, Relevant, Time-bound). We have to note though that students who withdrew halfway after possibly failing early on, were not included in the analysis. Thus, their performance data are not reflected in the coded data set we had extracted for analysis.

A selection of modules which did not follow the same pattern of assessments was excluded from the analysis. These included modules where two assessments were not administered; these included self-study modules

For future projects

While we analysed the findings, we deemed the following questions worthwhile to investigate in future:

1. Were statistical methods and software used in suitable fashion, considering the complexity of the methods and tools and the implications of conclusions?
2. Would conclusions be the same for years other than the 2009 sample?
3. Would a module-based analysis be meaningful and what would such an analysis yield? For example, are there any modules where the first two assignments are not useful predictors of the final outcome?
4. Do courses with traditional high failure rates, perhaps ones such as Accounting and Law, also display a pattern similar to the overall trend?
5. BBS is a multi-year multi-specialization degree programme – is the pattern the same at each level of year of study?
6. Do international students by any chance progress differently from domestic students (for yet unidentified reasons)?
7. Is the pattern for international students under pathway arrangements (i.e. starting here with second and third year modules) any different?
8. Would the statistical findings be any different if we only looked at students doing a course for the first time (i.e. excluding students repeating modules after having failed)?
9. Is there possibly a core group of students causing most assessment and module failures and how should they be supported?
10. Do GradDip students with an earlier degree in a business field perform different from other students?
11. Does the pattern of progression for students change over the study period, especially when doing the multi-year degree?
12. How does this compare to other qualifications and institutions, including analysis and publication by Potgieter with various co-authors during 2009 and 2010?

Conclusion

In a general sense, we conclude that the ITP sector could profitably develop statistical models to predict learner success or failure. To this end, we need to identify those predictors that explain the most significant proportion of the variance of end-of-course pass and fail categories. This paper, we reasoned, would be a tentative first step.

Mindful of the limitations of our approach, we soon sensed that the statistical model we were working on to predict success and failure was questionable on two grounds:

Firstly, of the total cohort of 2160 student data sets, we excluded 370 sets which had missing values. To develop a more nuanced approach, we would have to look at the reasons for these missing values, coding the reasons as variables to be included in the data set. Secondly, of those student data protocols retained in the data set, it seems that more than 80% performed at levels consistent with their initial two assessment scores. Given (a), we would have to be cautious to interpret initial assessment scores as consistent with final outcome. Moreover, we need to acknowledge that the strategy of coding non-academic variables could yield non-training-related predictor variables that explain a significant proportion of the variance on the pass/fail category. To improve the predictive efficiency of the model, we would have to re-code the existing data and extend the data set to account for biographical, motivational and socio-economic variables (to name a few).

Although 17.1% of the total number of data sets have been excluded, we could argue that for 79% of the cohort (2160 - [370 excluded + 84 failures]), their initial assessments were consistent with final mark. Moreover, of the 1790 student data sets retained in the data set, only 84 have failed. This reinforces our conclusion that we need to look beyond training-related variables, albeit that the latter should not be excluded, to develop more nuanced accounts of student performance and attrition. Indeed, we may discover that our focus should be on factors other than the academic to improve the effectiveness and efficiency of our interventions.

In summary, we want to argue the case for a comprehensive, nuanced approach to learner support and pastoral care, looking at learners as complex beings whose functioning in training-related and non-training-related contexts require individual and customized attention. These decisions, we argue, have to be informed by reliable and valid information about student worlds of experience and performance. To return to our initial thesis: we should use prediction models to benefit the training of all students – our focus on at-risk students is a deficit model. We have to add value to the training of all students, and good students should become better students, while we also deal effectively and efficiently with those at-risk. If we have reliable and valid assessment practices in place, we may develop models to select those students who sit (or do not sit) for final examinations.

References and Citations

Academic Directions 2010. (2010). Faculty Planning document, WINTEC.

Code of Practice for the Pastoral care of International Students. (2003). Ministry of Education, New Zealand. Accessed 09 March 2009 at http://www.minedu.govt.nz/educationSectors/InternationalEducation/ForProvidersOfInternationalEducation/CodeofPracticeforInternationalStudents.aspx

First Year Experience (FYE) Retention strategy. (2008) Planning document of PERSON-A, Manager: Student Experience, WINTEC

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

Our Direction. (2009). Presentation by Chief Executive of WINTEC, May 2009.

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 programme September 29, 2009 from http://hyperdisc.unitec.ac.nz/naccq09/programme/programme_saturday.html.

Potgieter, C., Greyling, W., Ferguson, B. (2010). Managing "At Risk" students and pass rates with SPSS. Proceedings of the 23rd conference of NACCQ, 6-9 July 2010, Dunedin, New Zealand, pp 227-234.

Potgieter, C., Roberton, 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

Stepwise Discriminant Analysis. http://marketing.byu.edu/htmlpages/books/pcmds/DISCRIM.html ; last accessed on 5 July 2010.

Stepwise linear regression in SPSS. http://www.statisticshell.com/multireg.pdf ; last accessed on 5 July 2010.

Tertiary Education Strategy Monitoring Information Website, accessed 15 August 2008. http://wiki.tertiary.govt.nz/~TESMon/MonitoringInformation/CompletionRetentionAndProgression

Tertiary Education Strategy 2007 - 2012 (2007). Wellington: Ministry of Education, accessed 15 August 2008 at https://staff.wintec.ac.nz/corporatedocs/index.htm

TEC to publish educational performance information. 2010. Accessed 15 March 2010 at 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 course 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