SA12122

Determining Student Performance in Online Database Courses

George Garman
The Metropolitan State College of Denver

ABSTRACT

Using data collected from 2002 to 2011, this paper analyzes the determinants of scores earned by students in a beginning online database course at a large urban educational institution. A backward stepwise regression model is used to find the significant explanatory variables on each of four dependent variables. The paper finds that a student's GPA is the most important determinant while declared major, the semester the course was taken, and year the course was taken have some degree of significance on the dependent variables. Gender was found to not be a significant determinant.

Keywords: online, database, assessment, stepwise regression

INTRODUCTION

Metropolitan State College of Denver is a modified open enrollment college that has been given the mission of educating all those who seek higher education studies. To effectively fulfill this mission, the College must continue to accept new applicants. During the past five years, the College has embarked upon an initiative to deliver many of its courses online. This initiative allows students to work and learn from home and avoid the trappings of an overextended campus. The College also benefits as it can attract new students from both within the Denver metropolitan area and from other areas throughout the state. Online delivery of courses also allows the College to keep pace with its competitors who are also delivering courses online.

This paper examines some of the determinants of the scores students earn in an online beginning database course. Some potential determinants include the grade point average the student brings into the class, the declared major of the student, the gender of the student, the semester the student enrolled in the course, and the year in which the student took the course. The research provides value to the faculty who are designing courses and to the department in its quest to develop an effective curriculum. The next section describes the methodology used to assess students the database course. The second section of this paper provides a review of the literature. The third section describes the methodology used to create the final models. The following section presents the results along with a discussion of the results. The final section presents the conclusions.

LITERATURE REVIEW

Alstete, and Beutell (2004) sampled ten online business courses and provided a statistical analysis of the grade performance. The findings indicate that grades, age, work experience, and discussion board grades are significantly related to overall course performance. However, the standardized test scores and organization position levels are not related to the performance in the online courses. Dutton (2002) analyzed the factors that lead to higher grade performance in online computer programming classes. Dutton found a positive significant effect for homework assignment completion and prior computer experience. There was a significant negative relationship between hours worked and grades. Wojciechowski and Palmer (2005) studied the predictors of student grades in an online business course at a community college. They found that the two variables that became the best predictors were attendance at a class orientation session and the student's grade point average.

ONLINE COURSE ASSESSMENT

The beginning database course covers both the theoretical foundations of database theory and the basic elements of the SQL programming language. Students are supported by a required textbook, an SQL guide that provides a self-guided approach to database programming, an extensive set of online tutorial slides, and access to an instructor. Students have access to an Oracle 11g database housed on the College's mainframe with permissions to create a variety of

Oracle objects under their own schemas. They can access this database from any on-campus computer lab or from any off-campus computer that has internet access.

Student work is assessed by three different methodologies. All students are required to complete three examinations that are proctored on campus by the College's Testing and Assessment Center. The examinations have a strict time limit and are closed book and closed notes. This assessment provides an assurance student work is not cooperative and is the justification for weighting examination scores highly in the determination of the grades. Students' are assessed by a series of online quizzes. These quizzes contain exclusively objective questions with the answers found in the textbook. Although the quizzes are open book, they are restricted by time limits. Students are instructed to become familiar with the material before starting the quiz since time limits will not allow them to successfully search for all of the answers. The final assessment is a series of projects that are submitted and graded. The projects are assigned early in the course and students have ample time to complete them. Collaboration among students on any of the quizzes or projects is strictly forbidden. The students never meet one another as a group. Except for the three visits to the Testing and Assessment Center, there would be no other time that the student would need to physically be on campus. It's probably unlikely that students would collaborate. However, since the quizzes remain open for a period of five days, it is possible that students who take the quizzes early may share information with other students. Likewise, covert collaboration on the projects may occur.

METHODOLOGY

Data have been collected from students enrolled in Metropolitan State College of Denver's CIS 3060 Database Management Systems online course sections from the Spring Semester of 2002 through the Fall Semester of 2011. Students who withdrew from the course for any reason before the end of the semester were removed from the study. Also, students who did not formally withdraw but who did not complete the course (as measured by completing the three required examinations) were removed from the study. After all of the eliminations, the study contains 385 observations. All sections of the course were taught by the same instructor.

All of the students in the sample were undergraduates with a class standing of either a junior or a senior. The course requires and enforces a prerequisite of an introductory principles of information systems course and a business problem solving course. The course is a core course required for all CIS majors although students throughout the college are welcome to enroll if they meet the prerequisites. Sixty-four percent of the students were declared majors in CIS. Sixty-eight percent of the students in the sample were male and twenty-four percent of the students enrolled in the course during the accelerated summer session.

This study identifies four independent variables each to be regressed against an identical set of potential dependent variables. The final course average (FinalAver) is a weighted average of all of the assessments required for the course. The exams are weighted as 75% of the final average, the assignments are weighted as 16.67% of the final average, and the projects are weighted as 8.33% of the final average. Grade assignments were consistent throughout all semesters includes in the study with ninety percent and above receiving an A, eighty percent to nine percent a B,

seventy percent to eighty percent a C, and sixty percent to seventy percent a D. Any final average that fell below sixty percent resulted in the student receiving an F for the course.

The average of the assignment scores (AssignAver) is the average of the ten assignments each being worth ten points. The project average (ProjectAver) is the average score earned by the student on the four completed projects. The exam average (ExamAver) is a simple average of the three examination scores. As previously indicated, any student who does not complete all of the three examinations is treated as a *de factor* withdrawal and removed from the study. However, if a student completes all three examinations and fails to submit any or all of the required assignment or project material, that student remains in the study.

A set of six independent variables were submitted for inclusion in the final model. Each student's cumulative grade point average (GPA) was retrieved from the transcript and, for consistency, the GPA the student had earned when he or she entered the course was used. Each student's declared major was also retrieved from the transcript and recorded as a binary variable (CISMajorBinary) with a one recorded if the student had a declared CIS major and as a zero otherwise. The student's gender (GenderBinary) was recorded as a one if the student was a male and zero if the student was a female. Binary variables were used to record the semester in which the student took the course. If the student took the course in the Spring semester the variable (SpringBinary) received a value of one. If the student took the course in the Fall semester, the variable (FallBinary) received a value of one. The courses offered over the summer would be determined by both the Spring and Fall binaries being equal to zero. Finally, a variable for the calendar year (Year) in which the student took the course was used.

The four equations tested are represented below:

(1) $\text{FinalAvg} = B_0 + B_1\ \text{GPA} + B_2\ \text{CISMajorBinary} + B_3\ \text{GenderBinary} + B_4\ \text{SpingBinary} + B_5\ \text{FallBinary} + B_6\ \text{Year}$

(2) $\text{ExamAvg} = B_0 + B_1\ \text{GPA} + B_2\ \text{CISMajorBinary} + B_3\ \text{GenderBinary} + B_4\ \text{SpingBinary} + B_5\ \text{FallBinary} + B_6\ \text{Year}$

(3) $\text{ProjectAvg} = B_0 + B_1\ \text{GPA} + B_2\ \text{CISMajorBinary} + B_3\ \text{GenderBinary} + B_4\ \text{SpingBinary} + B_5\ \text{FallBinary} + B_6\ \text{Year}$

(4) $\text{QuizAvg} = B_0 + B_1\ \text{GPA} + B_2\ \text{CISMajorBinary} + B_3\ \text{GenderBinary} + B_4\ \text{SpingBinary} + B_5\ \text{FallBinary} + B_6\ \text{Year}$

The final model for each of the four equations was developed from a stepwise regression. A backward elimination stepwise regression model was selected that initially includes all of the independent variables as candidates for inclusion in the final model and then eliminates the variables that don't demonstrate significance. The stepwise regression eliminates variables with a threshold alpha: 0.1 or F: 4 for removal.

RESULTS

A preliminary ordinary least squares regression was performed on each of the four equations to provide an analysis of the variable inflation factors. The variable inflation factors ranged from 1 to 1.8 which is well below the value of 5 threshold for suspected multicollinearity. Therefore, all six of the defined independent variables were used in the stepwise regression model as candidate independent variables. Table 1 below summarizes the results of the four regression equations.

Table 1
Regression Results

(Stepwise Regression Backward Elimination Alpha: 0.1 or F: 4 to remove)
(N = 385)

		Dependent Variable			
		(1) FinalAver	**(2) AssignAver**	**(3) ProjectAver**	**(4) ExamAver**
Constant		726.30	56.95	16.20	868.60
GNP (Range 0 to 4)	Coefficient	10.07	9.70	19.30	9.33
	T-Statistic	12.48	8.03	7.23	10.91
	P-Value	0.000	0.000	0.000	0.000
CISMajorBinary *(CIS Major = 1)*	Coefficient	1.87			2.30
	T-Statistic	1.89			2.15
	P-Value	0.059			0.032
GenderBinary (Male = 1 Female = 0)	Coefficient				
	T-Statistic				
	P-Value				
SpringBinary (Spring Semester = 1)	Coefficient		-3.40		
	T-Statistic		-1.97		
	P-Value		0.049		
FallBinary (Fall Semester = 1)	Coefficient		-3.90		
	T-Statistic		-2.10		
	P-Value		0.037		
Year (2002 to 2011)	Coefficient	-0.34			-0.41
	T-Statistic	-2.16			-2.46
	P-Value	0.032			0.014
R-Squared		29.68%	15.60%	12.00%	24.88%
Standard Error		9.13	13.70	30.20	9.67

The value of the student's GPA upon entering the class is the overwhelmingly significant variable in all four of the models. For equation 1 using the final average as the dependent variable, the regression coefficient approximately ten which indicates that a one point increase the GPA results in a ten point increase in the final average. This makes perfect sense as a one point increase in the GPA and a ten point increase in the average are both the measures of precisely one letter grade. Computer Information Systems majors perform slightly (but significantly) better on examinations than non major. This is interesting because when students have few or no time constraints on the assignments and projects, non CIS majors perform at a level identical to the CIS majors. In the more restrictive examination environment, CIS majors perform significantly better than the non CIS majors.

The year in which the student took the course is inverse and significant for the performance on examinations. Assuming both the grading and the difficulty level of the examinations have remained constant, for some reason students are not performing as well now as they did in the beginning years of the study. The semester binaries for both the Spring and Fall are significant and inversely related to the performance on the assignments. This suggests that students perform better on the assignments in the summer than they do during the regular semesters. Since the summer semester is half as long as a regular semester, the course progresses at twice the pace. Perhaps the intensity of the summer semester is somehow responsible for the higher summer scores. Gender is not significant in any of the models.

CONCLUSIONS

The results are rather consistent with previous studies analyzing the predictors on student grades earned in online courses. By far the most significant predictor is the grade point average that the student brings into the class. The enormous impact of grade point average on the final grade is self-evident – the better students do well. The better students are likely more motivated and interested and will perform well regardless of the course delivery method. Computer information systems majors perform better on examinations than non majors but don't perform significantly differently on the class assignments and projects. Students in the accelerated summer session performed better on the course assignments, exclusively objective questions, than did students during the regular Spring and Fall semesters. The year in which the student enrolled in the course was inverse and significant for both the final average and the performance on examinations.

Some area of interest still exist to be explored. The model may benefit by the addition of some explanatory variables. Some studies, Wojciechowski and Palmer (2005) for example, find that age is a significant explanatory variable on the grade earned. Since the database course is taught in the conventional classroom setting as well as online, the difference in the explanatory variables between classroom and online sections of the course would be of interest.

REFERENCES

Alstete, J. and Beutell, N., (2004) Performance indicators in online distance learning courses: a study of management education, *Quality Assurance in Education*, 12(1), pp.6 – 14.

Cheung, L and Kan, A. (2002) Evaluation of Factors Related to Student Performance in a Distance-Learning Business Communication Course. *Journal of Education for Business*, 77(5), 257-263.

Didia, D. and Hasnat, B. (1998) The Determinants of Performance in the University Introductory Finance Course, *Financial Practice and Education*, 8(1), 102-107.

Dutton, John, Dutton, Marilyn, and Perry, Jo, (2002) How Do Online Students Differ From Lecture Student?, *Journal of Asynchronous Learning Networks*, 6(1), 1-20.

George Garman (2007), Evaluating Reading Skills as a Predictor of Success in Online Courses, *The International Business and Economics Research Conference Proceedings*, October 2007, Las Vegas, NV. (distributed on CD ROM)

George Garman (2010), A Logistic Approach to Predicting Student Success in Online Database Courses, *American Journal of Business Education*. 3(12), 1-5.

James, David and Chilviers, Clair (2001), Academic and Non-Academic Predictors of Success on the Nottinghom Undergraduate Medical Course: 1970-1995, 35(11), 1056-1064.

Krieg, R.G. and Uyar, B. (1997), Correlates of Student Performance in Business and Economic Statistics?, *Journal of Economics and Finance*, 21(3), 65-74.

Krieg, R.G. and Uyar, B. (2001) Student Performance in Business and Economics Statistics: Does Exam Structure Matter?, *Journal of Economics and Finance*, 25(2), 229-241.

Monroe, S, Moreno, A., and Segall, M. (2011) Student Performance Determinants in a Business Statistics Course at a Large Urban Institution, *The Academic and Business Research Institute Conference Proceedings*, October, Las Vegas, NV. (distributed on CD ROM)

Parker, Angie, (1999) A Study of Variables that Predict Dropout from Distance Education, *International Journal of Educational Technology,* [online] 1(2).

Wojciechowski, A. and Palmer, L. D., (2005) Individual Student Characteristics: Can Any Be Predictors of Success in Online Classes?, *Online Journal of Distance Learning Administration*, 8(2), 1-21.