Extended Abstract - Correlating ACT Test Scores to Performance in First Year Engineering and Math Courses

Joni Torsella, Jeff Kastner, Gregory Bucks, Kathleen Ossman,
Rodney Roseman, and F. James Boerio
University of Cincinnati, Joni.Torsella@uc.edu,
Jeffrey.Kastner@uc.edu, Gregory.Bucks@uc.edu, Kathleen.Ossman@uc.edu,
Rodney.Roseman@uc.edu, James.Boerio@uc.edu

Abstract **- The College of Engineering and Applied Science at the University of Cincinnati recently enhanced the engineering curriculum by requiring all first-year engineering students to take three new classes. The first class, Engineering Foundations, introduces students to the various fields of engineering through hands-on laboratory exercises. The other two classes, Engineering Models I & II, provides students with an introduction to MATLAB® and using computational methods for solving calculus based engineering problems. This work describes an effort to correlate ACT scores to a student's performance in first-year STEM classes. The advantage of the ACT scores is that these scores reflect a student's competence in English, Math, and Reading. A typical statistical analysis was done between datasets where correlation values and trend lines were studied. The data was split into statistical clusters to further quantify the relationship between the two variables being correlated. The initial analysis began by forming clusters based on student performance in the Engineering Models and Calculus I classes. The clusters were formed specifically using the final grade in the Engineering Models I class and the first exam grade and final grade in Calculus I. After each cluster was formed, the average and standard deviation of the ACT scores were found for each cluster. The general data trend was that students who tended to struggle in both classes had the lowest ACT scores. It was also found that the ACT English score had the greatest variation between clusters while ACT Math had the least variation.**

Index Terms - First Year Engineering Courses**,** ACT, Correlation Analysis, Spearman, Cluster Analysis, kmeans

INTRODUCTION

The College of Engineering and Applied Science (CEAS) at the University of Cincinnati (UC) recently enhanced the engineering curriculum by including three classes required by all first-year engineering students. Engineering Foundations and Engineering Models I are taken during the first semester while Engineering Models II is taken during the second semester. The Engineering Foundations class introduces students to the various fields of engineering through hands-on laboratory exercises while the Engineering Models I & II classes provide students with an introduction to MATLAB® and computing methods for solving calculus based engineering problems. All three classes take place in a new 10,000 square foot Learning Center which opened at the beginning of the Fall semester in 2012 and is dedicated to the education of first-year students in the College of Engineering and Applied Science (CEAS). This facility features two state-of-the-art classrooms where 1 hour lecture sections are held and three "project rooms" where 2 hour recitation sections meet. In the project rooms, students work in small groups to solve problems involving applications of calculus in engineering.

Engineering Models I and II are taken concurrently with Calculus I and II and are intended to help students understand the applications of calculus to engineering problems. It has been well documented in the literature that success in engineering is closely related to success in calculus. [1]-[2] Our own research is consistent with the literature and has shown that students that earn a grade of C^+ or better in their first calculus course have a 75% probability of graduating from CEAS.

PROCEDURE

Prior to the beginning of each semester, it is important for each instructor to understand the skill set of their students. This will ensure that the instructor does not expect more than their students can produce or reteach concepts that students have already mastered in high-school. Some useful metrics that provide some understanding of a UC students' skill set is their ACT scores and the UC Math Placement test (MPT). This work describes an in-depth correlation analysis between students' performance on the ACT and their first-year STEM classes. The advantage of the ACT scores is that these scores reflect a student's competence in English, Math, and Reading. It has also been shown that students who have engineering within their high-school curriculum tend to score higher on the ACT math and Act composite [3]. ACT scores are also a key contributor when trying to predict a student's success in engineering [4]. At the same time, one challenge with the ACT metrics is that they are measured about a year before the student steps foot onto campus. Another challenge with the ACT math metric is that many engineering students perform well on that section of the ACT test. This leads to a smaller spread of that specific data.

The ACT scores in this work were specifically correlated to three courses during the fall semester (1) Calculus I, (2) Engineering Foundations, and (3) Engineering Models I. The overall performance in Calculus I was compared to ACT scores. For the Engineering Foundations class, the ACT scores were correlated to performance on the lab reports. For the Engineering Models class, the ACT scores were correlated to the performance in the class and the final exam. At the same time, the grades in the three courses were correlated to see how the courses were linked.

Data based on human experiences often lack a high correlation due to the variability involved in human interactions. Certainly, data based on freshman engineering students is no different. While a student may be extremely academically talented, there are other factors affecting their success. For example, this may be their first time away from home, they may not have the discipline to structure their day for success, they may have more financial responsibilities, etc. Furthermore, we believe the time delay between taking the ACT and class performance may affect the strength of the correlation between these two variables. The current data stays within the boundaries of performance based scores and rank, however future work will include having the students fill out a survey which helps explain some of their out of class behavior. The survey will be particularly aimed at time management issues such as do they have a job, are they living at home or in the dorms, do they prepare their own meals, etc.

A statistical analysis was done between datasets where correlation values and trend lines were studied. It was found during this part of the research that the test scores from the first-year engineering classes were not normally distributed. To remedy this, a Spearman correlation analysis was done instead of a Pearson. The idea here is to use rank instead of test score when correlating data. The general conclusion from the Spearman correlation analysis was that there was a significant correlation between performance in Calculus I and the Engineering Models I class. Particularly, the final grade in the Engineering Models I class had a correlation value of 0.52 to the Exam I score in Calculus and a correlation of 0.68 to the Final Grade in Calculus. The correlations in this case are significantly high, but still are based on the students' performance in the class and provide little information for an instructor at the beginning of the class.

Table 1 shows the correlation between the students' ACT scores and their performance in Calculus I and Engineering Models I. The correlations are lower than the class to class comparison, but they are shown to be significant. It is interesting to see that early in the semester the Math Act score has the highest correlation to the Calculus I exam, but by the end of the semester, the Models I grade and Calculus I grade are best correlated to the Total ACT composite. The Calculus I grade is still correlated well to the ACT Math, but the difference between Composite score and Math score was much more different at the start of the semester based on the Calculus I exam correlations. Also, the Engineering Models I class shows the highest correlation to the Composite and English scores and a much lower correlation for the Reading and Math scores.

TABLE I
CORRELATION BETWEEN PERFORMANCE IN CLASS TO ACT SCORES

	ACT			
	Total	Eng.	Read.	Math
Models I grade	0.343	0.335	0.256	0.25
Calc I, Exam I	0.304	0.272	0.19	0.366
Calc I grade	0.342	0.328	0.203	0.341

The data was split into statistical clusters to help better interpret the correlation data while still using quantitative means. The kmeans cluster analysis [5]-[6] was chosen and then performed using MATLAB®. The initial analysis began by forming kmeans clusters on data from the Engineering Models I and Calculus I class. The Engineering Models I class used the final grade and two separate cluster analysis was done for the first exam in Calculus I and the final grade in Calculus.

One challenge with a cluster analysis is the ability for the results to be repeatable. The analysis can be done to any number of clusters, but as the numbers of clusters increases, the repeatability of the cluster groups reduces. Three clusters of the Models I and Calculus I final grade were formed using the kmeans analysis and found to be 100% repeatable. The average and standard deviation of the Models I and Calculus I grades for each cluster is shown in Table 2. In general, the three clusters were mostly split into students who did very good, average, and below average in Calculus I. The cluster with the best Calculus I grade also had the highest average of any cluster for the Models I grade. The other two clusters followed this trend. Cluster 2 had the second highest average Calculus I grade and the second to highest average Models I grade, and cluster 3 had the lowest average Calculus I grade and the lowest Models I grade.

Aside from looking at the average score in a given cluster, it was also necessary to look at the spread in data for a given cluster. For example, cluster 3 has a very large spread in the Engineering Models I grade (see Table 2). The standard deviation for this cluster was 10% while the other two clusters had a spread of only 5% for the average group and 3% for the above average. However, the large standard deviation in the Models I grade for cluster 3 implies that the Calculus I grade was the main reason a student was placed in this cluster. This is a very important result and shows that there are some students struggling in Calculus I that may actually be a good engineer, but unfortunately may never get the opportunity to be an engineer due to their calculus grade.

TABLE II
AVERAGE SCORES AND ST DEV OF SCORES USING KMEANS METHOD

	Calc I		Models I	
	Mean	STD	Mean	STD
Cluster 1	85	4.6	94	3.1
Cluster 2	69	5.6	88	5.4
Cluster 3	48	7.8	78	7.1

Once the three main clusters were formed the next step was to pull quantitative data from each given cluster regarding their performance on the ACT. For a given cluster it was found that the cluster with the highest ACT scores (composite and individual) came from cluster 1, which was also the cluster with the highest Calculus I and Models I grade. The students in cluster 3 which had the lowest Calculus I and Models I average also had an ACT Math score 2 points lower. In a more pronounced manner, cluster 3 had ACT English and Reading scores that were 3.3 and 2.4 points lower, respectively than students in cluster 1. This seems to support the argument that students who are well-rounded in their education will ultimately perform better in first-year engineering course.

TABLE III
COMPARISON BETWEEN CLUSTERS AND AVERAGE ACT SCORES

	ACT			
	Total	Eng.	Read.	Math
Cluster 1	High	High	High	High
Cluster 2	-1.3	-1.6	-1.6	-1.2
Cluster 3	-2.0	-3.3	-2.5	-2.1

CONCLUSIONS

A statistical analysis was done on data from 200 students enrolled at UC as first-year engineering students. It was shown that there was a high correlation between Calculus I and performance in the Engineering Models I class. At the same time, a correlation analysis and a statistical cluster analysis seem to show that the students who have higher ACT English and Reading scores seem to standout as the best performers. Identifying these markers will help us better form groups. It also emphasizes the importance of teaching our freshman vocabulary and making sure they are well-rounded learners.

The best way to further the data conclusions will be to have more students. The current dataset was formed from half of the first year students in the College of Engineering, who took the ACT and took Calculus I during the fall semester. The other half of first year students will be analyzed to increase the number of samples. This should provide a total of 400 students per year for this analysis. Other clusters not included in the analysis were the students enrolled in Calculus II during the fall semester.

The results from this analysis will be used to generate a questionnaire for first-year engineering students and then use the data to form teams of students that are most likely to succeed together. All our first-year engineering classes emphasize team work. However, our formation of teams is very ad hoc. The questionnaire and understanding of their Calculus exam I score and ACT scores will help us form teams that can enhance the collaborative learning environments in our first-year engineering classes. The questionnaire/survey will also help us better understand the tools our students have going into the class and how to actively steer the class so they are properly taught. The results and efforts made will then be statistically tracked to see if there is an increased retention and increased performance in future classes.

REFERENCES

[1] Klingbeil, N., Rattan, K., Raymer, N., Reynolds, D., Mercer, R., Kukreti, A., and Randolph, D. (2008). The WSU Model for Engineering Mathematics Education: A Multiyear Assessment and Expansion to Collaborating Institutions. In Proceedings of ASEE Annual Meeting, Pittsburgh, PA, June 22-25, 2008.

[2] Hart, B.G., Holloman, T.L, O'Connor, C.A., (1995), A calculus retention program for students at risk in engineering. In Proceedings of Frontiers in Education Conference, Atlanta, GA, 1995.

[3] Phelps, A., Camburn, E., Durham, J. (2009), Engineering the Math Performance Gap, Careers Conference, A publication by the Center on Education and Work, Univeristy of Wisconsin, Madison, June 2009.

[4] O'Connor, C.A., McAnulty, B.H. (1981) The Predictive Ability of ACT Scores for Student Success at an Engineering School, Measurements and Evaluation in Guidance, v14, n2, p54-60, Jul 1981.

[5] Lloyd, S., (1982). Least squares quantization in PCM, IEEE Transactions on Information Theory, Vol. 28, No. 2, pp 129-137.

[6] Estivill-Castro, V. (2002), Why so many clustering algorithms, ACM SIGKDD Explorations Newsletter, Vol 4, No. 65.