Customized Course Advising: Investigating Engineering Student Success with Incoming Profiles and Patterns of Concurrent Course Enrollment

SungJin Nam
University of Michigan
3236 USB, 204 Washtenaw
Ann Arbor MI 48109-2215 USA
+1 (734) 615-4455
sjnam@umich.edu

Steven Lonn
University of Michigan
1401B Duderstadt Ctr, 2281 Bonisteel
Ann Arbor, MI 48109-2094 USA
+1 (734) 615-4333
slonn@umich.edu

Thomas Brown
University of Michigan
439 West Hall
Ann Arbor MI 48109-1107 USA
+1 (734) 763-3520
tommybro@umich.edu

Cinda-Sue Davis
University of Michigan
3236 USB, 204 Washtenaw
Ann Arbor MI 48109-2215 USA
+1 (734) 615-4455
csdavis@umich.edu

Darryl Koch
University of Michigan
273B Chrysler Cntr
Ann Arbor MI 48109-2092 USA
+1 (734) 647-7125
koch@umich.edu

ABSTRACT

Every college student registers for courses from a catalog of numerous offerings each term. Selecting the courses in which to enroll, and in what combinations, can dramatically impact each student's chances for academic success. Taking inspiration from the STEM Academy, we wanted to identify the characteristics of engineering students who graduate with 3.0 or above grade point average. The overall goal of the Customized Course Advising project is to determine the optimal term-by-term course selections for all engineering students based on their incoming characteristics and previous course history and performance, paying particular attention to concurrent enrollment. We found that ACT Math, SAT Math, and Advanced Placement exam can be effective measures to measure the students' academic preparation level. Also, we found that some concurrent course-enrollment patterns are highly predictive of first-term and overall academic success.

Categories and Subject Descriptors

H.4.2 **[Information Systems Applications]**: Types of Systems - *Decision Support*

General Terms

Measurement, Documentation, Design, Human Factors.

LAK '14, March 24 - 28 2014, Indianapolis, IN, USA

http://dx.doi.org/10.1145/2567574.2567589

Keywords

Learning Analytics, Data Analysis, Course Advising, Data Mining

1. INTRODUCTION

Every college student registers for courses from a catalog of hundreds or even thousands of offerings each term. Selecting the courses in which to enroll, and in what combinations, can dramatically impact each student's chances for academic success, progression towards degree, and retention within the field and the institution. Learning analytics has the potential to leverage historical student data to better inform students' decisions and ultimately improve their understanding of how course content and learning outcomes might better serve their ultimate institutional and long-term career objectives.

First implemented in 2008, the Science, Technology, Engineering and Mathematics (STEM) Academy identifies talented diverse incoming students with interest in STEM fields who, for reasons of socioeconomic class, first generation college student status, lack of high school rigor, race, or gender, might not be successful in pursuing a STEM degree [1]. The STEM Academy provides students with a highly coordinated and holistic support system during the critical transition years between high school graduation and the declaration of a STEM concentration by the third undergraduate year. Components of the program include an intensive pre-first year summer program, academic coaching, limited financial support, mentoring, study groups, and community building.

One of the primary objectives of the STEM Academy is to graduate students with an overall grade point average (GPA) of 3.0 or above. There are a number of predictors of overall GPA at completion of the baccalaureate degree [2, 3]. These may include, but are not limited to, standardized test scores, first and second year college GPA, first-generation college status, and family income. What have not been clearly identified are the course combinations and sequences that first and second year students, during this transitional phase from high school to college, should take. Nor has the impact on cumulative GPA of enrolled students taking some courses at community colleges or online and then transferring the credit been studied.

The STEM Academy's academic advisers, as well as general engineering academic advisers, help students select courses that will hopefully lead them on the pathway towards success in their chosen academic career. Until recently, academic advising has been largely based on advisors' intuition, prior experience, and institutional knowledge. With newly created analytic tools and reports, advisors can begin to infer how a particular student may perform in following semester, by comparing the data from previous students who were in similar condition, such as Advanced Placement (AP) credits, previous performance in core courses, or demographic data. However, a coherent system is needed to bring these disparate data sources together in order to better inform the course-selection process for students and their advisors alike.

A course recommendation system can "allow advisors and students to make plans for future semesters, equipped with data on courses or even majors in which past students with similar programs, grades, and course histories had found success" [4]. Degree Compass was one of the first such systems to use predictive analytics techniques to rank courses using grade and enrollment data, delivering information that could potentially help a student progress more readily through a chosen program. While Degree Compass had initial overall success, the predictive model of a C grade (2.0 GPA) or better lacks the specificity needed for the highly competitive undergraduate environment in which the STEM Academy is situated. Furthermore, examining course prediction in isolation obscures the variety of course combinations students register for in a given academic term; particular combinations can be academically hazardous to individual students, depending on their academic background. It is therefore necessary to examine course-taking behavior in concurrent patterns as well. For both academic advisors and students, predictive models utilizing rich and meaningful information, can greatly inform students' course-planning decisions, particularly in the critical transition years from secondary to higher education. Understanding these patterns and student characteristics that can accurately predict academic success is a first step before construction of a concurrent course recommendation system can begin. This paper describes our investigation in service of this goal.

2. RESEARCH QUESTIONS

The overall goal of the Customized Course Advising (CCA) project is to determine the optimal term-by-term course selections for all engineering students based on their incoming characteristics and previous course history and performance, paying particular attention to concurrent enrollment. This paper describes the preliminary data mining investigation that must precede any implementation of a course recommendation system in order to identify the key metrics and refine the predictive models. The overarching research questions guiding this investigation are:

- What scholastic characteristics of first-year engineering students are most predictive of future academic success?
- What concurrent course-enrollment patterns are most predictive of future academic success?

Uncovering the relationship between first-year engineering students' academic background and their concurrent course enrollment is an important and critical component of the development of a robust and dynamic course recommendation system that is able to consider the totality of a students' course load, academic trajectory, and ultimate goals within the institution.

3. BACKGROUND

Many students plunge into college level work with no difficulty. Other students can struggle and seem to need carefully adjusted course loads and sequences that will maximize academic success as indicated by course grades and overall GPA while still maintaining a reasonable time to degree (and may have positive effects on students' self-efficacy and self-regulation as well). This transitional phase for any given student, if needed at all, can range from the first semester freshman year up to and including the first semester sophomore year. The challenge lies in (1) identifying the student who needs this transitional period and (2) determining the combination of courses (A) within a given semester, (B) from semester to semester, and (C) transfer credit from other academic institutions that will optimize student success while still remaining on track to graduation.

In prior work, members of our research team have worked with the STEM Academy to develop Student Explorer, an early warning system (see [5], [6]). While early warning systems leverage learning analytics techniques to identify students in need of additional academic support (e.g., [7]), course recommendation systems are designed to be useful for all students regardless of perceived need. Both types of systems share the goal to ultimately reduce the time and effort involved in the "feedback loops" between students, instructors, and academic advisors [8]. Early attempts at course recommendation systems relied on social input (e.g., [9]), predating automated input from student academic history data which utilizes learning analytics to identify actionable points of decision-making, in this case, which courses, and in what combination, to enroll in a given term. Ultimately, the goal of a well-designed course recommendation system should be to provide useful data that allows the student (and their advisor, if applicable) to make intelligent decisions for their own course of study, supporting notions of self-advocacy and self-regulated decisions.

4. METHODOLOGY

The sample for the analyses was drawn from the data warehouse of a large American Midwestern university. According to the Carnegie Classification of Institutions of Higher Education (http://classifications.carnegiefoundation.org), this institution is a large, public, primarily residential four-year research university with very high research activity and a majority undergraduate enrollment. The university enrolls approximately 28,000 undergraduate students, more than 80% of whom are enrolled full-time, have entering test scores in the top fifth of all American baccalaureate institutions, and fewer than 20% of whom transfer in from other institutions. The engineering college enrolls approximately 5,800 undergraduate students, 21% of the total undergraduate population.

4.1 Data Sources

Academic history and demographic data from the institutional data warehouse was drawn from undergraduate students who matriculated into the engineering college between the fall 2001 and fall 2012 semesters (N=12,836). Nearly three-quarters (75%) of the students were male and 11.5% were international (non-US) students. The average incoming Math ACT score was 31 (out of 36) and the average incoming Math SAT score was 709 (out of 800) (ACT and SAT are college readiness assessments taken in high school as part of the college application process). These students enrolled in 6,877 unique courses throughout their undergraduate tenure (some students have not yet graduated).

The warehouse data included records such as students' course enrollment information, credit for courses taken at other institutions (transfer credit), achievement and placement test scores, and final degree information. In this paper, the analyses primarily focused on analyzing enrollment information, which included course name, course section, course grade, semester GPA, and enrolled term. Students' preparation level when entering university, such as ACT and SAT scores, Advanced Placement (AP) credits, and placement tests as well as final degree status, such as number of years took to graduate, kind of achieved degree, and final GPA, were also included in the analyses. The scope of the analysis was limited to enrollment data and incoming academic measures, while setting aside demographic information. Demographic information was excluded to limit the complexity for this preliminary study.

4.2 Procedure

The analysis started by identifying the first semester math course students selected. Often times, proficiency in math readiness assessments is considered as an indicator of success in university engineering courses [10]. In this context, the first semester math course can be considered as an indicator of undergraduate engineering students' academic preparation level, in terms of equivalent credit and subjective confidence in math. For example, incoming students can satisfy course requirements with AP, IB, or other A-Level Credit exams, and take higher-level courses in the first semester. This leads to a wide variety of first semester course combinations.

The grade of 3.0 was set as a standard of academic success. This standard, a grade point average of 3.0 on a 4.0 scale, was set because it is (1) the minimum GPA required by most engineering employers for internships and permanent positions, and (2) it is a goal of the STEM Academy. Whether students finish their degree as originally entered was also considered as an indicator of success. Although students entered the university as a part of the engineering program, not all of them graduate with an engineering degree. So if a student graduated from the College of Engineering with a final GPA of 3.0 or better, we considered this student as successfully graduated as an 'engineer'.

There are six math courses that engineering students can select in their first semester. This variety is the widest among all subjects in core-required courses that freshmen engineering students take, such as engineering, chemistry, and physics. Grouping students by their first term math courses was a simple and easy way to group students and achieve various profiles from many students.

5. RESULTS

5.1 Scholastic Characteristics of Engineering Students

In order to predict which students are likely to be successful in college, the analysis started from profiling the incoming characteristics of students, which include high school GPA, high school GPA percentile, ACT math score, and SAT math score. Groups were compared using an analysis of variance (ANOVA) and post-hoc Tukey.

High correlation between first semester GPA and graduation GPA ($r = 0.763$, $p < 0.001$) showed that first semester performance is very important for engineering students. Choosing which courses to take in the first semester may become a critical decision with students' academic performance throughout the graduation, and selecting the first semester math course can be a major determinant of the overall first semester performance.

5.1.1 High School Records, ACT Math Score, and SAT Math Score

Of the six math courses offered in the fall semester that freshmen engineering students typically select, four of them, Calculus I, Calculus II, Multivariable Calculus, and Differential Equations, belong to the engineering core required courses. Two other math courses, Pre-Calculus and Applied Honors Calculus II are not included in the core courses but are popular choices. Pre-Calculus is recommended for students who do not have AP credit and the result of math placement test advised them as need to take additional preparing course for the later calculus sequences. Applied Honors Calculus II is an alternative to Calculus II, designed for students who scored a 4 or 5 (out of 5) on the Advanced Placement (AP) Calculus exam.

Records of high school performance were not meaningful for predicting the academic performance of students. The correlations between students' high school records and first semester performance or graduation GPA were significant, but the absolute values for coefficients were below 0.1 (Table 1). There were significant differences between students' high school GPA between starting math courses [$F(5, 5779) = 64.66$, $p < 0.001$]. A post hoc Tukey test showed that the high school GPAs of Pre-Calculus students were significantly higher than those of students starting with Multivariate Calculus ($p < 0.01$) and Differential Equations ($p < 0.001$). Students' high school GPA percentile had a similar pattern [$F(5, 5779) = 67.89$, $p < 0.001$].

Engineering students' ACT and SAT math scores, however, were positively correlated with first semester GPA and graduation GPA, ranging from 0.364 to 0.412 (Table 1). Also, students' ACT and SAT math scores, on average, increased by the difficulty of starting math courses (Table 2). ACT Math scores of each starting math courses were significantly different [$F(5, 4841) = 342.4$, $p < 0.001$]. A post-hoc test showed that scores are significantly different except between Applied Honors Calculus II and Multivariate Calculus, and Differential Equations ($p < 0.001$). Students' SAT math scores were all significantly different [$F(5, 4286) = 225$, $p < 0.001$]. SAT Math scores were significantly different except between Multivariate calculus and Differential Equations ($p < 0.001$), meaning that ACT or SAT math subject scores are more meaningful predictors than high school records for assuming students' academic preparation level for engineering.

Table 1. Correlation Between Academic Performance in University and Incoming Characteristics

	High School GPA	High School GPA Percentile	ACT Math	SAT Math
First Semester GPA	-0.047***	-0.070***	0.391***	0.364***
Graduation GPA	-0.062***	-0.080***	0.412***	0.391***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 2. Incoming Characteristics: High School Performance, ACT Math scores, and SAT Math score by Starting Math Courses

First Semester Math Course	High School GPA	High School GPA Percentile	ACT Math	SAT Math
Pre-Calculus	3.60 (SD = 0.65)	91.51 (SD = 16.76)	26.69 (SD = 2.36)	605.34 (SD = 56.08)
Calculus I	3.56 (SD = 0.92)	90.02 (SD = 22.90)	29.66 (SD = 2.79)	679.48 (SD = 63.70)
Calculus II	3.74 (SD = 0.57)	94.16 (SD = 14.12)	31.48 (SD = 2.52)	704.90 (SD = 53.29)
Applied Honors Calculus II	3.77 (SD = 0.47)	95.33 (SD = 11.30)	32.25 (SD = 2.44)	723.78 (SD = 47.85)
Multivariate Calculus	3.32 (SD=1.32)	83.48 (SD = 33.18)	32.98 (SD = 2.25)	738.09 (SD = 48.12)
Differential Equations	2.83 (SD=1.72)	71.00 (SD = 43.01)	32.84 (SD = 2.48)	746.40 (SD = 49.70)

5.1.2 AP Calculus Exam Scores

The AP Calculus exam is one way that students can show their proficiency in math. Depending on what kind of AP Calculus exam the student passes, they can waive corresponding courses and proceed to more advanced math course. AP Calculus AB (AB) is a calculus course taken after the Pre-Calculus course in high school. AP Calculus BC (BC) is also a calculus course, but covers additional topics from AB. If students have scores of AB (4 or 5), or 4 of BC exam, it is equivalent of Calculus I. If students have score of 5 from BC exam, it is equivalent of Calculus II. Most students who have credit from AP Calculus exam take Calculus II or higher courses as their first semester math course. Table 3 shows the success rate (percentage of students who received a 3.0 or better) on first semester GPA, graduation GPA, and starting math courses by AP Calculus exam scores.

The AP scores largely determine enrollment numbers. For example, students with Calculus AB typically enroll in Calculus II and Applied Honors Calculus II, while Calculus BC students typically enroll in Multivariate Calculus and Differential Equations courses.

Students' first semester GPA and final graduation GPA was significantly different by AP Calculus exam scores [$F(3, 3202) = 124.8$, $p < 0.001$]. A post-hoc Tukey revealed that the differences for semester GPA between AB and BC tests were significant ($p < 0.01$). The semester GPAs of BC 5 students were significantly higher than students with other AP Calculus scores ($p < 0.01$), while AB 5 and BC 4 students were not significantly different. A similar pattern was observed with final graduation GPA [$F(3, 3202) = 107.6$, $p < 0.001$], such that students with BC 5 scores were significantly higher than AB 4 and BC 4 students ($p < 0.01$). Students scoring an AB 5 were not significantly different with students with a 4 or 5 on the AP BC exam.

In Calculus II, the higher students scored in the AP test, the higher the success rates were. Students' course performance was significantly different by AP exam scores [$F(3, 1620)=62.27$, $p < 0.001$]. A post-hoc test revealed that the performance of students with a BC 5 score was significantly higher than other AP scores ($p < 0.05$). While the average grade of AB 5 and BC 4 students were not significantly different, both were significantly different from students who scored an AB 4 ($p < 0.001$). The average grades of Applied Honors Calculus II students were significantly different by AP exam scores [$F(3, 401) = 16.19$, $p < 0.001$], while the over 90% of students with AB 5, BC 4, and BC 5 earned 3.0 or better grade. A post-hoc test showed that the performance of students with AB 4 was significantly lower than students with other AP Calculus scores ($p < 0.001$). Students in Multivariate Calculus showed significant difference of average grade by AP exam score [$F(3, 987) = 3.617$, $p < 0.05$]. However, no pairs of AP exam scores showed significantly different results. Lastly, students in Differential Equations had the similar pattern as students in Applied Honors Calculus II. Average course grades were significantly different by AP exam scores [$F(3, 182) = 3.855$, $p < 0.05$]. Performance of students in this course with AB 4 was significantly lower than students with BC scores of 4 or 5.

Table 3. Success Rate (Percent of Students Earning 3.0 or Better) of First Semester GPA, Graduation GPA, and Starting Math Courses by AP Calculus Scores.

AP Calculus Exam Type and Scores	First Semester GPA	Graduation GPA	Calculus II	Applied Honors Calculus II	Multivariate Calculus	Differential Equations
AB, 4 N=639	62.75%	69.64%	47.30% N=520, M=2.74, SD=0.80	60.61% N=66, M=3.00, SD=0.79	70.96% N=31, M=3.21, SD=0.71	59.09% N=22, M=3.11, SD=0.60
AB, 5 N=1092	83.88%	87.17%	74.47% N=803, M=3.23, SD=0.70	92.42% N=211, M=3.53, SD=0.56	86.67% N=30, M=3.50, SD=0.58	85.42% N=48, M=3.43, SD=0.64
BC, 4 N=386	81.61%	82.64%	78.43% N=255, M=3.26, SD=0.67	92.94% N=85, M=3.53, SD=0.46	77.78% N=36, M=3.18, SD=0.83	100% N=10, M=3.63, SD=0.31
BC, 5 N=1089	89.53%	89.72%	97.83% N=46, M=3.60, SD=0.42	95.35% N=43, M=3.66, SD=0.69	85.12% N=894, M=3.46, SD=0.62	91.51% N=106, M=3.54, SD=0.53

Note: For math courses, N is number of students with the corresponding AP score who enrolled in the specified course. Success rates for each course are the percentage of the course N.

5.2 Concurrent Course Enrollment Patterns

Students were enrolled in 3,186 unique course combinations for their first semester across the 10 years in our dataset. Although this number is quite large, the variability of course combination can be abstracted based on the required courses.

In order to declare an engineering major, every engineering undergraduate student needs to finish their 55 credit hours of core requirement courses, including chemistry, engineering, physics, math, and courses for intellectual breadth (courses from outside of the engineering college) with a minimum GPA of 2.0.

Engineering core courses consist of Introduction to Engineering and Introduction to Computers and Programming. For chemistry, there is General Chemistry lecture, two general chemistry labs, and a Structure and Reactivity lecture and lab. Physics core courses are General Physics I lecture and lab, and General Physics 2 lecture and lab. Core math courses are four math courses: Calculus I and II, Multivariate Calculus, and Differential Equations. Pre-Calculus is a requirement for Calculus I, but it does not count as a core course. Applied Honors Calculus is an equivalent of Calculus II. For intellectual breadth, students are required to take 16 credit hours from outside of the engineering college. This wide range of courses generates copious unique combinations in concurrent course enrollment.

5.2.1 Course Combination Selection Pattern

Table 4 shows the top 5 popular course combinations for each first-term math course. This table explains 24.35% of total course combinations. Although most of engineering students' course selections, especially for the first semester, are limited within a certain range, courses that count as intellectual breadth, such as Introduction to Psychology or Introduction to Sociology, add to the wide variety of small-percentage combinations.

By comparing the subtotal frequencies of combinations by first-term math courses, it can be observed that the concurrent course enrollment patterns are more fixed, as the level of math courses gets lower. Students who start with higher-level math courses select a wider variety of courses from core-required courses or others. Statistically, when comparing each pair of math courses, the proportion of students who take one of the top five most popular course combinations is significantly higher for the lower level math course ($z > 3$, $p < 0.01$), with the exception of the pairs of Calculus II-Applied Honors Calculus II and Multivariate Calculus and Differential Equations which are not significantly different.

Table 4. First Semester Concurrent Course Pattern by Starting-Level Math Courses

						Chemistry					Engineering		Physics		Other
						General Chemistry			Structure and Reactivity				General Physics I		
	%	N	% First Sem. GPA Success	Avg. First Sem. GPA	Num. of Credit	Lec-ture	Lab 1	Lab2	Lec-ture	Lab	ITE	ITC	Lec-ture	Lab	
Pre-Calculus (N=578)	20.93%	121	28.93	2.58	13	X	X	X			X				
	7.61%	44	36.36	2.85	15	X	X	X			X				TEP
	4.15%	24	45.83	2.85	13	X					X				TEP
	2.08%	12	50.00	2.87	15	X					X				ITP
	1.73%	10	10.00	2.50	15	X					X				ITS
Calculus I (N=4418)	9.10%	402	42.79	2.80	13	X	X	X				X			
	8.24%	364	42.58	2.86	13	X	X	X			X				
	5.25%	232	58.19	3.05	15	X	X	X				X			TEP
	3.73%	165	66.67	3.10	15	X	X	X			X				TEP
	2.20%	97	41.24	2.86	13				X	X	X				
Calculus II (N=3208)	6.23%	200	65.00	3.10	13	X	X	X				X			
	5.08%	163	55.21	3.05	13	X	X	X			X				
	4.77%	153	55.56	3.03	13							X	X	X	
	4.05%	130	76.15	3.23	15	X	X	X				X			TEP
	3.68%	118	61.86	3.08	13				X	X	X				
Applied Honor Calculus II (N=849)	4.95%	42	83.33	3.42	13	X	X	X				X			
	4.83%	41	78.05	3.35	13							X	X	X	
	4.48%	38	84.21	3.56	15							X	X	X	TEP
	3.89%	33	87.88	3.36	13				X	X	X				
	3.89%	33	75.76	3.46	15	X	X	X				X			TEP
Multivariate Calculus (N=2315)	4.15%	96	81.25	3.33	13				X	X	X				
	4.06%	94	70.21	3.29	13							X	X	X	
	3.28%	76	72.37	3.22	13	X	X	X			X				
	3.24%	75	88.00	3.46	15							X	X	X	TEP
	2.85%	66	66.67	3.16	13	X	X	X				X			
Differential Equation (N=703)	4.27%	30	93.33	3.79	16							X			MC, POE
	3.27%	23	78.26	3.10	13							X	X	X	
	2.99%	21	61.90	3.00	13				X	X	X				
	2.56%	18	83.33	3.45	13	X	X	X				X			
	2.56%	18	94.44	3.61	13				X	X		X			
Total (N=12071)	24.35%	2939													

Note: Introduction to Engineering = ITE; Introduction to Computers and Programming = ITC; The Engineering Profession = TEP; Introduction to Psychology = ITP; Introduction to Sociology = ITS; Principles of Economics I = POE; Multivariate Calculus = MC

Some courses were limited with certain groups of math courses. For example, popular courses for Pre-Calculus starters were limited to General Chemistry and Introduction to Engineering. Variations were made with whether taking lab courses of General Chemistry, and extra courses from outside of core requirements. For students enrolled in Calculus I, Structure and Reactivity, an advanced chemistry course set, and Introduction to Computing and Programming, which requires prior or concurrent enrollment in Calculus I or equivalent, were a popular combination. Patterns for students enrolled in Calculus II or higher starter groups were similar. Physics courses were also a popular course combination. Students in Calculus II or higher choose the General Chemistry set, the Structure and Reactivity set, or the General Physics I set. And a selection of Introduction to Computers and Programming is added.

5.2.2 Performance Characteristics of Patterns

In Table 4, the most popular course that students selected from outside of the core engineering courses was The Engineering Profession (TEP). This course is 2 credit hour seminar course learning about career paths as an engineer. Since the grade curve of this course is set high (an average grade of 3.82 (SD = 0.49)), taking the TEP course increased students' overall term and final graduation GPAs for most combinations.

For students enrolled in Pre-Calculus, first semester GPA between course combinations were significantly different [F(4, 206) = 2.968, $p < 0.05$]. However, taking TEP with the General Chemistry lecture and Introduction to Engineering did not result in a boost for students' semester GPA. Introduction to Sociology and Introduction to Psychology did not make significant differences either.

The performances of Calculus I combinations were significantly different [F(4, 1255) = 11.5, $p < 0.001$]. Taking TEP course with Calculus I significantly increased students' semester GPA. Students' semester performance was significantly higher when taking TEP with General Chemistry lecture and lab courses and either one of the engineering core courses ($p < 0.01$).

Calculus II combinations' semester GPA were also significantly different [F(4, 759) = 2.752, $p < 0.05$]. However, the semester GPA for students enrolled in a combination of TEP, General Chemistry lecture and lab, and Introduction to Computers and Programming was not significantly higher than a combination without TEP.

For Applied Honors Calculus II, combinations were not significantly different (F(4, 182) = 1.339, $p = 0.257$]. Although the semester GPA was significantly different for students enrolled in Multivariate Calculus combinations [F(4, 402) = 3.706, $p < 0.01$] and Differential Equations [F(4, 90) =4.092, $p = 0.01$], no meaningful comparison of TEP between combination was found.

Combinations between General Chemistry and General Physics I, with Introduction to Computers and Programming were not significantly different. Among popular combinations of Calculus II and Differential Equations courses, Introduction to Computers and Programming was concurrently taken with General Chemistry or General Physics I. Semester GPAs were not significantly different in both starting math course groups. A similar pattern was observed between students enrolled in General Chemistry and Structure and Reactivity. In Calculus I, Applied Honors Calculus II, and Multivariate Calculus, the most popular combinations included pairing the math course with either Introduction to Engineering or Introduction to Computers and Programming. Average semester GPAs were not significantly different for students enrolled in these three first-term math course groups.

5.2.3 Number of Credit and Semester Performance by Starting Math Courses

Students most often take either 13 credits or 15 credits in the first semester. The institution considers 13-18 credits as full time enrollment, equating to 3-4 core courses and one non-core course.. Since engineering students need to complete their intellectual breadth requirements, many students who take more than 13 credits enroll in courses outside of the engineering college. Students who took more than 13 credit hours had significantly higher semester GPAs (M=3.06 vs. 2.94, SD=0.85 vs. 0.86, t(7395.508) = 7.5383, $p < 0.001$). Between the math courses, semester GPA was significantly different [F(11, 12033) = 164.3, $p < 0.001$], except for students who took Pre-Calculus and Applied Honors Calculus II in their first semester ($p < 0.001$) (Table 5). This finding is partially complicated. Pre-Calculus is a required course, for students who test low on the mathematics placement exam, to enroll in Calculus I. However, Pre-Calculus enrollment does not count for financial aid enrollment since it does not count toward the engineering degree

Table 5. Semester Performance by Starting Math Courses and Credit Hour

First Semester Math Course	More Than 13 Credit Hour	N	% First Semester GPA Success	Average Semester GPA
Pre-Calculus	Yes	339	38.94%	2.80 (SD=0.61)
	No	239	31.8%	2.65 (SD=0.64)
Calculus I	Yes	2963	56.63%	3.04 (SD=0.56)
	No	1455	45.29%	2.85 (SD=0.64)
Calculus II	Yes	2147	69.63%	3.05 (SD=0.55)
	No	1057	60.17%	3.20 (SD=0.62)
Applied Honors Calculus II	Yes	601	82.86%	3.41 (SD=0.47)
	No	246	79.67%	3.32 (SD=0.54)
Multivariate Calculus	Yes	1629	83.79%	3.45 (SD=0.52)
	No	669	75.93%	3.30 (SD=0.58)
Differential Equations	Yes	532	87.47%	3.53 (SD=0.49)
	No	168	75.00%	3.26 (SD=0.68)

completion. Thus, students with financial aid packages enrolled in Pre-Calculus must take at least 15 credit hours.

5.2.4 Comparison among Core Requirement Courses

As presented in section *5.2.1*, concurrent course enrollment patterns were very complex. For example, some students did not take the lab courses with related lecture class in the same semester. Some students enrolled in more than 13 credit hours, and some took a variety of courses not included in the core engineering required courses. This complexity made it difficult to recognize the impact of certain courses on the first semester performance. However, we observed that some courses or course sets are rarely enrolled together in a single semester.

Regardless of the course levels, the popular course combination for the first semester student was taking one course from math subject (12,071 out of 12,836 students enrolled in one of the math courses in the first semester), one from engineering subject, and a course set (lab and lecture courses) of either chemistry or physics (Table 4). Generally, this combination fills the minimum 13 credit hours. This analysis was conducted to determine which course selection gave students more benefit among these courses or disciplinary course set.

A logistic regression model was built to see how the particular choice of core course or course set within the same subject category effects on students' first semester success (whether they earn 3.0 or better GPA) (Table 6). Students were categorized by their starting math courses. In terms of chemistry and physics course sets, it was counted only if a student took both lab and lecture courses in the first semester. With this model, we could investigate which course or course set can be more beneficial to the students by their starting math courses. The exponential of the log odds in Table 6 represent the ratio of students' proportion on being successful in comparing course over base course.

Regardless of starting math course group, every course or course set selection had a significant impact on the first semester performance except one course set pair. Taking an advanced level course or course set within the same subject or taking a physics course set instead of chemistry course set gave students a better chance of success in their first semester (from 1.152 to 2.465 times, $p < 0.05$). Taking more than 13 credit hours also increased the chance of success by 1.744 ($p < 0.001$).

However, when the model was adapted separately by starting math course, only few courses or course sets showed a significant increase in the odd ratios. For Pre-Calculus starters, taking the Structure and Reactivity course set rather than the General Chemistry course set increased the likelihood of getting a semester GPA of 3.0 or better by 5.442 times ($p < 0.01$). For students who are starting with Differential Equations, taking Introduction to Computers and Programming, instead of Introduction to Engineering, increased the odds of success in the first semester by 2.348 times ($p < 0.001$). Also, students who took the Structure and Reactivity course set showed more likelihood of successful semester GPA than those who took General Physics I course set. Lastly, in terms of whether students take more than 13 credit hours in the semester, every starting math group showed the significant improvements if they took more credits, except for Pre-Calculus and Applied Honor Calculus II groups.

Table 6. Odd-Ratios of for the First Semester Success (≥ 3.0 GPA) by Starting Math Courses and Core Requirement Subjects

	Subject	**Engineering**	**Chemistry or Physics**						**Credit Hours**
Odd Ratios	**Comparing Course / Base Course**	**ITC / ITE**	**CHEM II / CHEM I**	**PHYS I / CHEM I**	**PHYS II / CHEM I**	**PHYS I / CHEM II**	**PHYS II / CHEM II**	**PHYS II / PHYS I**	**More Than 13 Credit Hours**
First Semester Math Courses	**Overall (N=12071)**	1.152 ***	1.660 ****	1.736 ****	2.465 ****	1.046	1.485 *	1.4201 *	1.744 ****
	Pre-Calculus	0.611	5.442 **	NA	NA	NA	NA	NA	1.273
	Calculus I	0.922	0.949	1.151	0.578	1.213	0.610	0.503	1.871 ****
	Calculus II	1.142	1.215	1.005	0.989	0.827	0.814	0.984	1.628 ****
	Applied Honor Calculus II	0.921	1.150	1.107	0.736	0.963	0.640	0.665	1.262
	Multivariate Calculus	1.065	1.166	0.891	1.321	0.764	1.132	1.482	1.809 ****
	Differential Equation	2.348 ***	1.498	0.724	0.831	0.483 *	0.555	1.149	2.219 **

Note: Introduction to Engineering: ITE, Introduction to Computers and Programming: ITC, General Chemistry lecture and lab: CHEM I, Structure and Reactivity lecture and lab: CHEM II, General Physics I lecture and lab: PHYS I, General Physics II lecture and lab: PHYS II.

**** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

6. DISCUSSION

6.1 Summary of Our Findings

Our investigation described in this paper was guided by two overarching research questions: (1) what scholastic characteristics of first-year engineering students are most predictive of future academic success? And (2) what concurrent course-enrollment patterns are most predictive of academic success? The high correlation coefficient between first semester GPA and final graduation GPA indicated that first semester performance is very important to engineering students' academic success. Thus, analyzing students' first semester enrollment and success patterns is critical for developing a course recommendation system.

To answer the first question, we compared high school records, ACT Math scores, SAT Math scores, and AP Calculus exam scores with students' academic performance. We found that ACT Math and SAT Math scores were significantly correlated with first semester GPA and graduation GPA, while high school records only showed small correlation coefficients. Also, ACT Math and SAT Math scores were significantly different by students' starting math courses. Semester GPA and graduation GPA were also significantly different by AP Calculus Exam scores. The performance of students with AB 4, the lowest score to earn credit for Calculus I, in math courses were significantly lower than other AP scores across math courses. These results indicate that ACT Math, SAT Math, and AP Calculus scores effectively indicate the academic preparation level of freshmen engineering students, and can be used for predicting student success.

For the second guiding research question, we observed that there is a greater variety of unique course combinations starting math course increases in difficulty level. Also, The Engineering Profession (TEP) survey course has relatively high average grade. Taking the TEP course with other core requirement courses were expected to boost students' semester performance. However, only students who started with the Calculus I course benefited from TEP enrollment.

Students who took more than the minimum 13 required credit hours per semester performed significantly better, except those who started with Pre-Calculus and Applied Honors Calculus II. Students with financial aid packages enrolled in Pre-Calculus must take at least 15 credit hours. This means that the non-engineering requirement courses, which other students' success rates benefit from, may not helping students who begin in low level math courses like Pre-Calculus.

Lastly, we found that some courses or course sets affects students' semester performance significantly. Among students who started with Pre-Calculus math course, those who took Structure and Reactivity course set showed higher likelihood of earning a 3.0 or better GPA in the first semester than students who started with the General Chemistry I course set. Students who started with Differential Equations had a higher chance of success when they took Introduction to Computers and Programming instead of Introduction to Engineering. The Structure and Reactivity course set also helped when compared to the General Physics II.

6.2 Contribution of This Study

This study sheds light on the importance of identifying the scholastic characteristics and concurrent course enrollment of freshmen engineering students. Both are important to help students achieve success in their first semester, which is very highly correlated with the final performance. Suggesting the "right" courses for students based on the incoming characteristics may help them achieve the academic success necessary for pursuing successful STEM careers and graduate STEM degrees.

The findings in this paper help identify the requirements for a predictive model that can give insight for students and advisors during course registration to raise the probability of earning a GPA of 3.0 or better in first semester, and ultimately the foundation of a concurrent course recommendation system.

For example, AP Calculus scores are better predictors of success in the core math courses than GPA or ACT score. So AP Calculus scores, when available, should be treated as the main factor in selecting the appropriate math course for an incoming freshman.

If a student from the engineering college starts with a low-level math course, such as Pre-Calculus or Calculus I, the selection of other courses is limited since they do not meet the pre-requirements of other courses, such as physics course sets. Also, students who started with Pre-Calculus showed that they do not benefit from taking non-required courses.

On the other hand, students who can start in higher-level math courses may widen their range for course consideration. Students with higher AP Calculus score earned higher semester GPA, final graduation GPA, and starting math course performance. While starting in higher-level course may fulfill their academic curiosity and accelerate their time to degree, starting in relatively lower-level class can lead to solidifying the students' preparation for further engineering major courses.

6.3 Limitations

There were some limitations of the study. First, in this initial investigation, there were too many variables to consider with limited time and resources. With more than 3,000 unique course combinations among first-year engineering students, analyzing only the most popular combinations is not enough to reveal every detail of concurrent course enrollment patterns. Second, the scope of the analysis was limited. Even though the results from first semester are important, it does not fully explain students' future academic success or course enrollment. Analyses such as course taking sequences [11], or tracking the semester performance by incoming characteristics can be applied. Third, our analyses were only limited to undergraduate engineering students. The curriculum for engineering students is a relatively fixed sequence of courses. Other disciplines, both STEM and non-STEM such as social sciences or humanities, may have a much more complex structure of requirement courses, more numbers of unique course combinations, and different measurements for academic preparation that need to be applied. Lastly, even though the demographic information was not considered with the analysis of this paper, we believe that integrating the demographic information with incoming characteristics or concurrent course enrollment would be a worthwhile investigation. For example, the location of high school with AP exam results and high school GPA may be related, and could help improve the precision of a course recommendation system. However, care must be taken to ensure that including such information does not recreate an existing inequality (e.g., low female enrollment in STEM courses).

This preliminary study mainly focused on data exploration. It may not answer every question that students or academic advisors may encounter or prove the causal relationship between the variables and students' performance yet. Further analysis that considers a wider range of students with more variables would help to build a tool that can actually be used by an advisor.

6.4 Future Research

In future investigations, we plan to expand the scope of the analyses to include additional semesters and additional final degrees, beginning with non-engineering STEM degrees. Furthermore, analyzing the sequence of courses may be necessary for expanding the scope to additional semesters. For example, students may want to compare the possible outcomes between taking Introduction to Engineering in this semester and Introduction to Computers and Programming in following term, and vice versa. Another example can be comparing when students enroll in chemistry or physics lab courses non-concurrently with the corresponding lecture course.

Investigating when students enroll in particular courses can also potentially affect academic success. The time "off" between the fall semester and winter semester (winter vacation) is relatively shorter than between winter semester and fall semester (summer vacation). The difference in these gaps may be correlated with the performance of courses that are related to each other. For example, taking Calculus I and Calculus II in fall-winter semester sequence may result in a higher success rate compared to the case of winter-fall enrollment, since the shorter time between courses may increase knowledge transfer between courses. Also, taking courses in the summer semester may further increase success rates as the summer semester class may shorten the learning time gap that can occur during a long summer vacation. Particularly with students who enroll in courses like Pre-Calculus, it may be interesting to investigate differences in future academic performance controlling for whether students finish their required math courses more quickly than others. Finally, we plan to study the effect of taking courses at external institutions. Many students complete required core courses at other institutions, such as local community colleges. The effect of taking core required courses at other institutions may be related to future course performance, such as the next level physics course for particular engineering majors (e.g., Mechanical Engineering vs. Electrical Engineering).

7. CONCLUSION

This research is critical to not only building a course recommendation system, but also better supporting students through their chosen source of study and balancing that against the wealth of knowledge that the institution has at its disposal regarding those course selections. Providing students more information about details of successful characteristics may raise the ethical questions, such as replicating gender, ethnicity, or socioeconomic differences in choosing courses. But it is also true that students, properly scaffold, may understand those trends and instead leverage the prior course and student success history to maximize their own outcomes. As these analyses become easier to conduct and automate, the institution may be compelled to properly release these findings such that student success can be optimized for every student, regardless of demographics and incoming test scores [12].

8. ACKNOWLEDGEMENTS

This project was funded by a grant from the University of Michigan Learning Analytics Taskforce. The authors would like to thank the members of the USE Lab for their helpful feedback on this paper.

9. REFERENCES

[1] Davis, C. S., St. John, E., Koch, D. & Meadows, G. (2010). Making academic progress: The University of Michigan M-STEM Academy. Proceedings of the joint WEPAN/NAMEPA Conference, Baltimore, Maryland.

[2] French, B., Immekus, J. & Oakes, W. (2005). An examination of indicators of engineering students' success and persistence. Journal of Engineering Education 94(4), pp 419-425.

[3] Richardson, M., Abraham, C. & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychological Bulletin, 138(2), pp. 353-387.

[4] Denley, T. (2012). Austin Peay State University: Degree Compass. EDUCAUSE Review Online. Available: http://www.educause.edu/ero/article/austin-peay-state-university-degree-compass

[5] Lonn, S., Krumm, A. E., Waddington, R. J., and Teasley, S. D. (2012). Bridging the gap from knowledge to action: Putting analytics in the hands of academic advisors. Paper presented at The 2nd International Conference on Learning Analytics and Knowledge. Vancouver, BC, Canada.

[6] Krumm, A. E., Waddington, R. J., Lonn, S., & Teasley, S. D. (In Press). A learning management system-based early warning system for academic advising in undergraduate engineering. In (J. Larusson & B. White, Eds.) Handbook of Learning Analytics: Methods, Tools and Approaches. New York: Springer-Verlag.

[7] Arnold, K. E. & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. Paper presented at The 2nd International Conference on Learning Analytics and Knowledge. Vancouver, BC, Canada.

[8] Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. Paper presented at The 2nd International Conference on Learning Analytics and Knowledge. Vancouver, BC, Canada.

[9] Farzan, R., & Brusilovsky, P. (2006). Social navigation support in a course recommendation system. In Adaptive Hypermedia and Adaptive Web-Based Systems (pp. 91-100). Springer Berlin Heidelberg.

[10] Veenstra, C. P., Dey, E. L. & Herrin, G. D. (2008). Is modeling of freshman engineering success different from modeling of non-engineering success? Journal of Engineering Education, 97(4), 467-479.

[11] d'Aquin, M. & Jay, N. (2013). Interpreting data mining results with linked data for learning analytics: Motivation, case study and directions. Paper presented at The 3rd International Conference on Learning Analytics and Knowledge. Leuven, Belgium.

[12] Willis III, J. E., Campbell, J. P. & Pistilli, M. D. (2013). Ethics, big data, and analytics: A model for application. EDUCAUSE Review. Available: http://www.educause.edu/ero/article/ethics-big-data-and-analytics-model-application