Predicting 6-Year Graduation and High-Achieving and At-Risk Students

**By Dmitri Rogulkin, Sr. Research Analyst**   **May 2011**

The Office of Institutional Research, Assessment and Planning has conducted studies about factors that impact first-time full-time freshmen students' progress during their first year in college. Preparation for college, first-term GPA and participation in student support groups were found to be the most influential on first-year retention. In addition, students who do not return for their second fall semester are unlikely to graduate. Only 6% of the 321 students in the fall 2004 cohort who dropped out after the first year returned to Fresno State and earned a degree within six years.

The focus of this analysis is on students who continue into their second year. What factors are the most influential in predicting their graduation? Who are the at-risk students? Who are those that might benefit from interventions designed to improve their odds of graduating? Data mining algorithms were used to answer these questions. The sample included 1,922 students from the fall 2004 cohort who continued their enrollment into fall 2005. 58% of these students graduated from Fresno State within six years, while 12% persisted into their seventh year.

The Decision Tree Model[1] showed that staying on track, which was defined as reaching sophomore level no later than the end of the third semester, and cumulative GPA after the first year were the most influential factors in predicting six year graduation (Figure 1). Most of the students (85%) were able to reach sophomore level by their third semester. Their graduation rate was 67%. Only 12% of those who were not on track graduated within 6 years.

For those who stayed on track, graduation rates varied based on the cumulative GPA. Students with the lowest GPA after the first year (less than 1.9) had a 35% probability of graduating, while the best performing students (GPA 3.4 or higher) graduated at an 84% rate. Students with a cumulative GPA from 2.9 to 3.4 graduated at a 74% rate. Those with a GPA range between 2.4 and 2.9 had a 59% rate. Students whose cumulative GPA was close to the academic probation mark (GPA 1.9 to 2.4) graduated at a 47% rate.

Figure 1: Decision Tree Predicting 6-Year Graduation
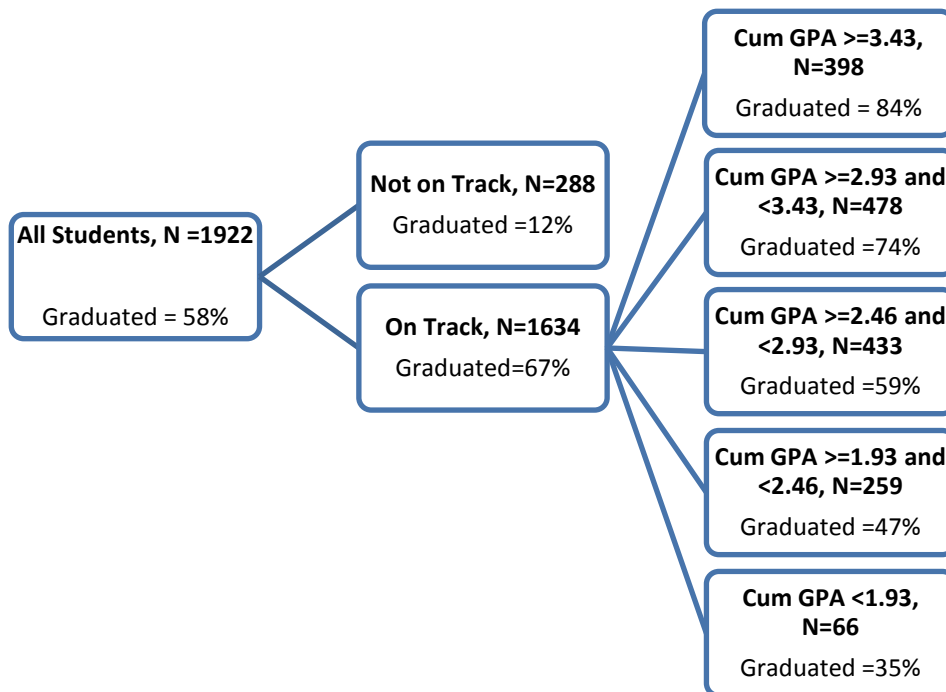
Clustering

The data mining clustering algorithm was used to identify profiles of students who were at-risk of not graduating, those who had high odds of graduating, and those who might potentially benefit from some sort of support to improve the likelihood of graduating. Five clusters of students emerged (Table 1).

Cluster #1: Best prepared for college, best performing students.

This group had 298 students and practically all of them (95%) completed a degree at Fresno State. More than half (54%) did it within four years. These students were the best prepared for college. Their average high school GPA was 3.84, and only a few needed English or Math remediation. Their average SAT, ELM, and EPT scores were the highest among all five clusters. Most of the students (75%) were not PELL grant eligible, which is a proxy for identifying low/high income status, and were continuing generation students (56%). Two-thirds were female and the majority (59%) was white students. After the first year of college, this cluster had the highest average cumulative GPA (3.50). On average, these students earned more units in their first year (42) than any other group. All of them stayed on track.

Cluster #2: Average preparation, best graduation rates.

This was the largest cluster by size (N = 526). Almost all of them graduated (99%), with two-thirds completing a degree during the fifth and six years. Compared to the first cluster, this group had a lower average high school GPA and lower standardized tests scores. Almost three quarters of this cluster needed English remediation and 60% needed Math remediation. However, these students managed to do well academically. Their average cumulative GPA was 3.10 and they earned an average of 29 units during the first year. All of them were sophomore students after three semesters of enrollment. A low percentage of these students repeated courses during the first two years of college. 10% of these students received support from the campus Educational Opportunity Program (EOP). A majority of this group was first-generation and underrepresented minorities. 42% were PELL grant eligible or low-income students.

Cluster #3: Least prepared, but potential for success.

This was the second largest cluster (N = 435). They were least prepared for college based on their standardized tests scores and remediation status. All of them needed English and Math remediation. Despite that, 41% of this cluster managed to graduate within six years with another 18% enrolled into their seventh year. This group had the highest proportion of first-generation students (86%) and underrepresented minorities (57%). Almost eight out of ten in this cluster were PELL grant eligible. Also, these students were much more likely to be female (73%) and come from the Fresno Service Area (77%). 37% of the group (largest among all five clusters) were EOP students. The academic performance of the group was below average among the five clusters (cumulative GPA 2.54 and 30% not on track). However, they graduated and persisted at considerably higher rates than the "at-risk" cluster, which had comparable preparation and performance characteristics. It may be that academic and financial support that this group potentially received helped students to improve their odds of graduating.

Cluster #4: High test scores, average performance, but below average graduation

This cluster had the second highest average scores on standardized tests after Cluster #1. Most in this group did not need Math or English remediation. Even though the tests scores were high, the high school GPA (3.24) was below average. Students were likely to stay on track (85%). They earned, on average, a 2.64 cumulative GPA and 28 units at the end of the first year. However, this cluster had the second lowest graduation rate (37%) and their seventh year persistence rate (12%) was considerably lower than the rates in Cluster #3 and Cluster #5.

The demographic profile of this cluster was similar to Cluster #1, except that this group had a higher proportion of male students (56%) and those coming from outside of the Fresno Service area (61%). Both percentages were the highest among all clusters. About a quarter of this group dropped out from Fresno State after two years. From the study conducted using the National Student Clearinghouse database (see IRAP brief "After 3 Years of College, Why Leave Without a Degree"), we know that those who left after two years were likely to transfer to a four-year college. So, considering these last facts, it appears that Cluster #4 students, though performing relatively well, did not find a fit with the Fresno State campus or Fresno area and left, likely transferring to another four-year university.

Cluster #5: At-Risk students, lowest graduation rates.

There were 340 students in this cluster. Only 14% of them graduated within six years, while 27% were still enrolled in their seventh year. Like Cluster #3, these students had the lowest average high school GPA. However, their standardized tests scores and remediation levels were just slightly below the rates of Cluster #2 which had an almost perfect graduation outcome. The academic performance of this group was the lowest among all clusters. Their average cumulative GPA was only 2.4 and, like Cluster #3, 30% of these students did not stay on track. Also, they had the highest proportion of students who repeated two or more courses during the first two years in college (37%). Overall, the most repeated courses were MATH 45, ENGL 1, PLSI 2, HIST 11 and HIST 12. There were some similarities between this cluster and Cluster #3 in demographics. Both groups had large proportions of first-generation students (71%) and underrepresented minorities. However, the "At-Risk" cluster had fewer students who were PELL eligible and more males compared to Cluster #3. A much smaller percentage of students were in EOP than in Cluster #3. This seems like a group that needs more university support to improve their graduation rates.

Implications & Conclusion

Staying on track is an important factor in students' progression towards a degree. This suggests a potential intervention that would involve identifying, contacting and providing support to students who appear to be off track based on units earned and enrolled in during their first year. There are two clusters of students who are less prepared for college and graduate at lower rates. They are predominately first-generation, under-represented minority, low-income students. One of these clusters has a substantially higher graduation rate than the other and a larger percentage of the higher achieving group is supported by EOP. From our previous studies, we know that several support interventions helped students improve their academic performance and increased their retention rates. This suggests that the lower achieving group may have greater potential to succeed with more university support. However, this study did not include all types of university support that students may have received. Therefore, we will further investigate to determine if these students participated in other university support programs.

There is also a group of students who are better prepared and better performing academically, but do not complete a degree at Fresno State. The majority of these students are male and they are coming from outside of the Fresno Service area. We do not know why they are leaving. However, based on our previous research they fit the profile of students who are transferring to another school. Therefore, it is likely that they are transferring and earning a degree elsewhere rather than dropping out. There may be nothing Fresno State can do to retain these students if the reason for their leaving is lack of fit to this campus or the Fresno area. However, another one of our studies raised the question of whether students are challenged enough academically. Given that these are well-prepared students, this could be an issue. This is another question that we will attempt to explore further.

Addendum

1.   Did fewer students in the lowest achieving cluster participate in the university support programs?

Yes. Only 20% of students in the at-risk cluster #5 were involved in support programs. In contrast, among the cluster of similarly prepared students that had a substantially higher graduation rate (#3), half participated in support programs (Table 2). Students were most likely to participate in EOP and CAMP. Other programs included in the analysis were the Health Career Opportunity Program, Student Support Services, and University Migrant Services but very few students were involved in these. Data on participation in some of the other support services such as the Mentoring Institute, Supplemental Instruction, Learning Center, Career Services, and Health and Psychological Services are not available in PeopleSoft, so cannot be included in our analyses.

Implication 1: With more support from the university, a larger percentage of students in the at-risk group may graduate. However, given the cost of programs such as EOP and CAMP, it may be beneficial to determine what elements of these programs are most beneficial to students and whether they can be incorporated into other programs on a larger scale. In addition, students must be willing to utilize support services.

Implication 2: Student support exists in many forms at Fresno State, but without centralized data collection, the extent of support students are receiving cannot be determined. Neither can these programs be included in analyses that may help reveal beneficial practices. A more robust data collection system is needed.

2. Did students in the out of service area cluster who did not complete a degree at Fresno State transfer out?

Yes. A substantially higher proportion of leavers in the out of service area cluster #4 transferred out compared to the other clusters with low graduation rates. Almost three-quarters (72%) of the leavers in cluster #4 transferred out. 30% of the leavers transferred to another four-year university (Table 3).

Implication: This is a group that the University is likely to lose for which there may be no intervention. Further research into their entering characteristics and experience on campus would be necessary to determine that. However, it is likely that most of these students are graduating from another college and therefore are finishing their education rather than dropping out.

Other Related Research and Briefs:

▸ After 3 Years of College, Why Leave Without a Degree?
▸ When Do Students Who Drop Out Leave Our University?
▸ First Generation Students: First in the Family, Less Likely to Succeed in College
▸ First-Year Leavers: More Students Transfer Than Dropout
▸ First-Year Student Retention and Attrition
▸ SupportNet Assessment
▸ Does Service Learning Help Students Succeed?
▸ Honors Program Assessment Summary Report and Appendix
▸ Are We Helping Freshmen Succeed? Results of Fresno State's First Year Experience and Academic Success Course
▸ Are Fresno State Seniors Academically Challenged Enough?

Table 1: Cluster Profiles

	Variables	States	All	Best Prepared, Best Performing Cluster 1	Average Preparation Best Graduation Cluster 2	Least Prepared, Potential to Succeed Cluster 3	Transfer Out Cluster 4	At-Risk, Lowest Performing Cluster 5
	Size		1922	298	526	435	323	340
	Graduated within 6 years	Yes	1123	**95%**	**99%**	**41%**	**37%**	**14%**
		No	799	5%	1%	59%	63%	86%
	Graduation Outcome	Degree in 4 yr	378	54%	32%	5%	8%	1%
		Degree in 5 & 6 yr	745	42%	66%	36%	30%	13%
		Persist in 7 yr	222	1%	0%	18%	12%	27%
		Left after 2 yr	242	2%	0%	16%	26%	22%
		Left after 3+ yr	335	2%	1%	25%	25%	36%
Performance	Cumulative GPA after 1st year	Mean	*2.83*	*3.50*	*3.10*	*2.54*	*2.64*	*2.40*
		Deviation	0.7	0.4	0.4	0.6	0.6	0.6
	Cumulative units after 1st year	Mean	29	42	29	23	28	24
		Deviation	10	14	6	6	7	6
	Stayed on Track	yes	1634	100%	100%	70%	85%	70%
		no	288	0%	0%	30%	15%	30%
	Count of repeated courses during two years[2]	Zero	1072	88%	75%	37%	46%	33%
		One course	459	10%	19%	30%	29%	30%
		Two Plus courses	391	2%	5%	33%	25%	37%
Preparation	High School GPA	Mean	3.36	3.84	3.43	3.16	3.24	3.19
		Deviation	0.5	0.3	0.4	0.4	0.4	0.4
	SAT MATH	Mean	490	584	474	377	558	468
		Deviation	0.1	0.0	0.1	0.1	54.1	0.3
	SAT VERB	Mean	470	582	452	354	533	444
		Deviation	0.1	0.0	0.1	0.1	59.7	0.3
	ELM Total Score	Mean	43	57	46	31	55	45
		Deviation	0	0.8	0.1	0.0	0.6	0.1
	EPT Total Score	Mean	144	153	146	136	152	144
		Deviation	0	0.8	0.1	0.0	0.5	0.1
	Remedial English	Needed	1156	8%	71%	100%	26%	82%
		Did not need	766	92%	29%	1%	74%	18%
	Remedial Math	Needed	1011	7%	60%	100%	15%	68%
		Did not need	911	93%	40%	0%	85%	33%
	EOP Status	No	1669	99%	90%	63%	99%	86%
		Yes	253	2%	10%	37%	1%	14%
Demographics	Gender	Female	1182	67%	69%	73%	44%	53%
		Male	740	33%	31%	28%	56%	47%
	Ethnicity	White	720	59%	43%	8%	61%	22%
		Hispanic	593	15%	31%	48%	16%	41%
		Asian	353	10%	17%	33%	8%	22%
		African American	110	2%	4%	10%	5%	9%
		Unknown	137	15%	5%	2%	11%	5%
		American Indian	9	1%	0%	0%	0%	1%
	Underrepresented Minorities[3]	Non-URM	1071	68%	60%	41%	69%	45%
		URM	710	17%	35%	57%	21%	50%
		Unknown / Inter	141	16%	5%	2%	11%	5%
	PELL Grant eligibility	Not eligible	1062	76%	58%	21%	81%	46%
		Eligible	860	24%	42%	79%	19%	54%
	First Generation Status[4]	Yes	1142	39%	59%	86%	38%	71%
		No	681	56%	36%	10%	59%	22%
		Unknown	99	6%	5%	7%	3%	5%
	Fresno Service Area[5]	Yes	1204	65%	64%	77%	40%	65%
		No	718	35%	36%	23%	61%	35%

Table 2: Participation in Support Programs*

	Cluster3 (Grad. Rate 41%)		Cluster5 (Grad. Rate 14%)	
Received University Support	224	51%	68	20%
Didn't Receive Support	211	49%	272	80%
Total Students	**435**	**100%**	**340**	**100%**

* Include EOP, CAMP, Health Career Opportunity Program, Student Support Services, University Migrant Services

Table 3: Transferred Out Data

	Cluster3 (Grad. Rate 41%)		Cluster4 (Grad. Rate 37%)		Cluster5 (Grad. Rate 14%)	
Transferred Out	**116**	**64%**	**123**	**72%**	**143**	**64%**
2-year college	98	54%	71	41%	106	48%
4-year college	18	10%	52	30%	37	17%
Didn't Transfer	**64**	**36%**	**49**	**28%**	**79**	**36%**
Total Leavers	**180**	**100%**	**172**	**100%**	**222**	**100%**

Technical Notes

1. The Decision Tree model includes a 6-year graduation outcome (yes / no) for each individual student as a predictable variable and the following input variables: cumulative GPA after the first year, High School GPA, SAT Math and Verbal Scores, EPT and ELM scores, English and Math Remediation Status, EOP Status, Gender, Underrepresented Minority Status, First-Generation Status, PELL Grant Eligibility (as a proxy of high/low income), Fresno Service Area Residence, Staying on Track Status.
2. Count of repeated courses included number of courses that a student repeated during the first two years. Only courses that were repeated due to a failing grade were included.
3. Underrepresented minorities included American Indian, African American, and Hispanic.
4. First Generation students are those with neither of their parents having earned a Bachelor's degree.
5. Fresno Service Area includes Fresno, Madera, Tulare and Kings counties.