Developing a Dashboard to Compare Predicted and Actual Graduation Rates

2019 California Association for Institutional Research Conference
November 7th, 2019

Cinnamon L. Danube, UC Merced
Tongshan Chang, UC Office of the President
Corinne Townsend, UC Merced

UNIVERSITY
OF
CALIFORNIA

How can IR help ensure that students at our institutions graduate?
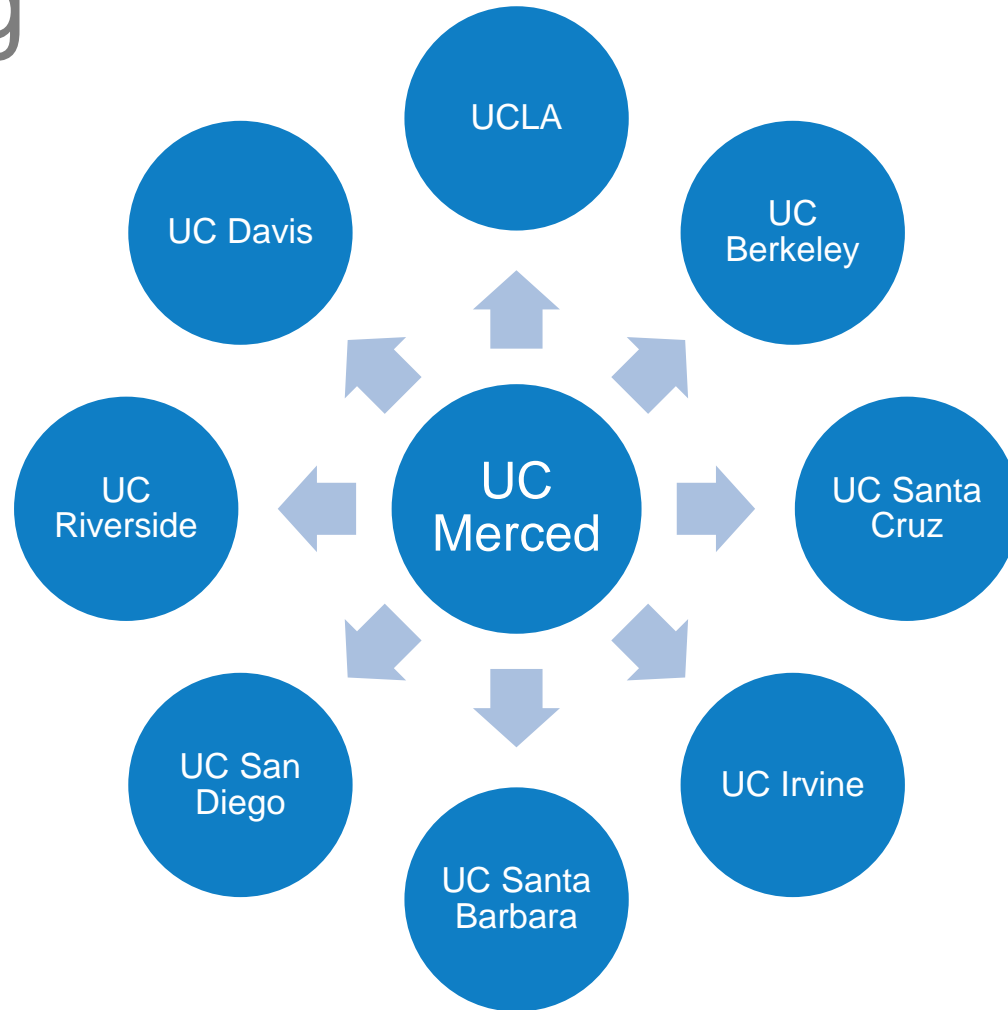
Session learning outcomes:

You will learn one approach for developing predicted graduation rate models for your campus.

You will learn how to use the model results to develop a dashboard to engage campus leaders to use the results to develop a strategic plan for student completion.
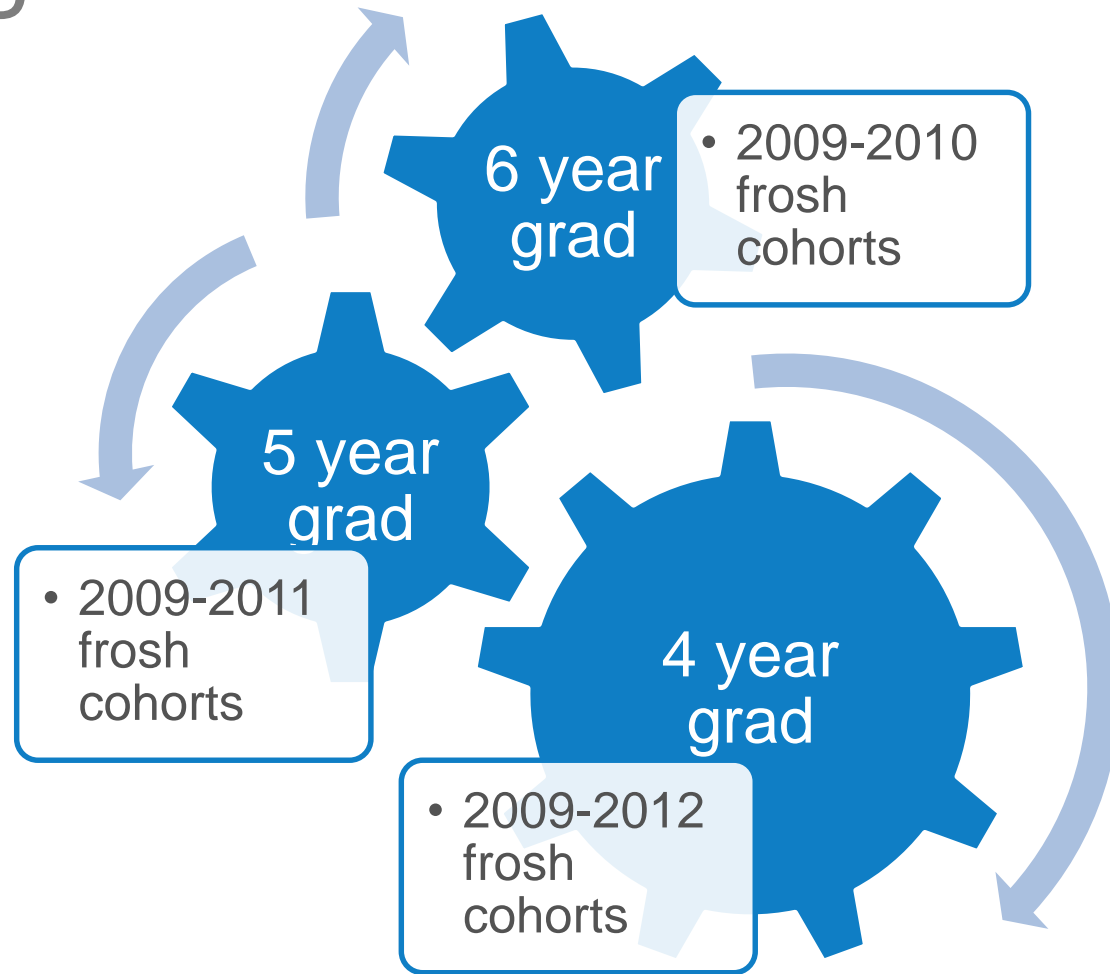
Modeling Methodology - Inputs

Residency	First Generation Status	Race/Ethnicity
Pell Grant Status	Gender	High School GPA
Test Scores	Pre-UC Units	Major Discipline

UNIVERSITY OF CALIFORNIA

Modeling Methodology – Model Building

UNIVERSITY
OF
CALIFORNIA

Modeling Methodology – Model Building

- 2009-2010 frosh cohorts (6 year grad)
- 2009-2011 frosh cohorts (5 year grad)
- 2009-2012 frosh cohorts (4 year grad)

Modeling Methodology – Predictive Accuracy

	4 year grad	5 year grad	6 year grad
Pseudo R-squared	16%	12.4%	11.9%

Modeling Methodology – Logistic Regression & Predicted Probabilities

Step 1: obtain coefficients for each predictor variable (Beta weights)

Step 2: apply coefficients to get predictions at student level

- EXP(Y)/[1+EXP(Y)] where Y = a + b1x1 + b2x2 + … + bixi

Step 3: Calculate average predicted probabilities overall and across student groups

Dashboard Development Steps

Clarify goals for end user interaction with the dashboard

Consider data structure and variables to ensure execution of dashboard design goals

Dashboard Goals for User

Use predicted grad rates to set grad targets and/or view progress relative to targets

Use gaps between predicted and actual grad rates to identify intervention targets

Dataset Structure and content

|	pidm	cohort	gender	graduated_in_4_yrs	gender_OPmodel	predicted_prob_4yrgrad
1		200930	F	1	1	.51
2		201030	M	0	0	.60
3		200830	F	1	1	.38
4		200830	F	1	1	.72
5		200830	F	0	1	.59
6		200830	M	0	0	.49
7		200830	M	0	0	.35
8		200830	M	1	0	.40
9		200830	M	0	0	.39
10		200830	F	0	1	.36
11		200830	F	1	1	.61
12		200830	M	0	0	.27
13		200830	F	0	1	.51
14		200830	M	1	0	.26
15		200830	F	0	1	.47
16		200830	F	1	1	.42
17		200830	M	0	0	.33

In Tableau

Calculated average predicted probabilities

Calculated actual graduation rates

Dashboard Demo – 6 year grad figure

Predicted and Actual 6 Year Graduation Rates for Each Cohort by No Breakout and No Filter - No Filter
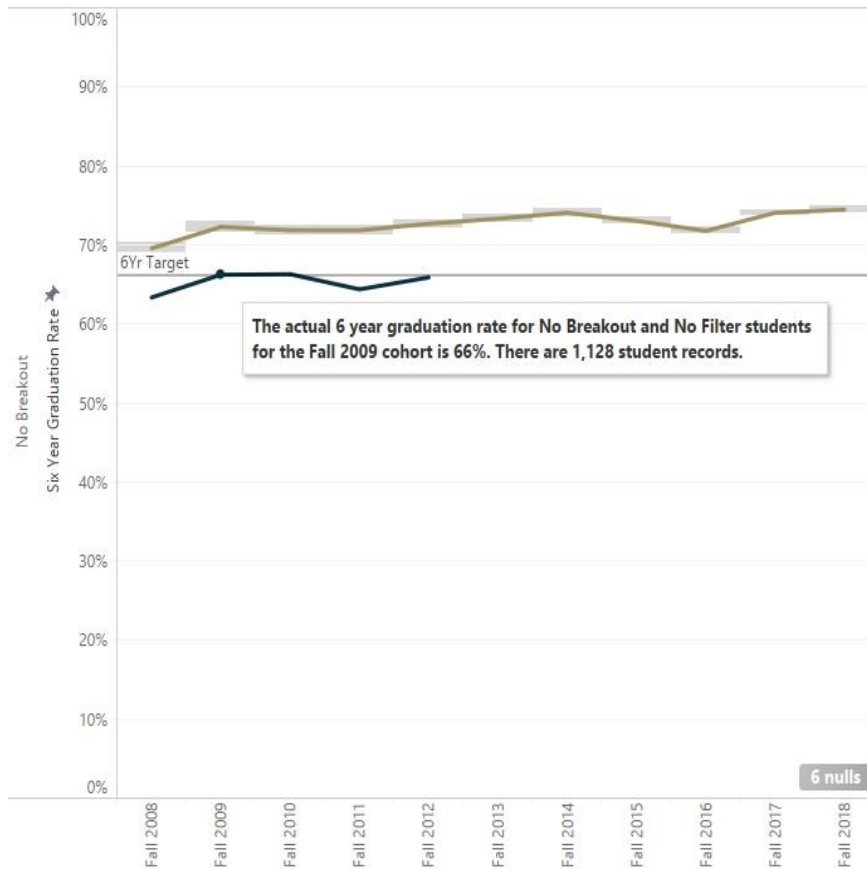
Figure Legend
- ■ Actual Six Year Graduation Rate
- ■ Predicted Six Year Graduation Rate (Average)

The actual 6 year graduation rate for No Breakout and No Filter students for the Fall 2009 cohort is 66%. There are 1,128 student records.

Error Band

95% CI Above & Below Prediction

Breakout Parameter

No Breakout ▼

Filter Parameter

No Filter ▼

Filter Parameter Group

(All) ▼

Instructions:
From the Figure Legend, this figure displays actual (blue line) and average predicted (gold line) 6 year graduation rates. The Target line displays the current target 6 year graduation rate. The grey band displays one possible measure of error for the average 6 year graduation rate prediction (the 95% Confidence Interval around the prediction).

To modify the visualization:
(1) Select a Breakout Parameter from the list if you want to make a side-by-side comparison between students based on a particular characteristic. Select "No Breakout" if you do not want to make a side-by-side comparision.
(2) Select a Filter Parameter to filter the data to include only students based a particular characteristic. Select "No Filter" if you do not want to filter the data.
(3) If you selected a Filter Paramenter, next use Filter Parameter Group to select the specific value(s) of the characteristic of the students you want to include in the visualization. The "All" box should be selected if no filter has been applied.

Dashboard Demo – 6 year grad table

Predicted and Actual 6 Year Graduation Rates for Each Cohort by No Breakout and No Filter - No Filter

Breakout Param..		Fall 2008	Fall 2009	Fall 2010	Fall 2011	Fall 2012	Fall 2013	Fall 2014	Fall 2015	Fall 2016	Fall 2017	Fall 2018
No Breakout	Actual Six Year Graduation Rate	63%	66%	66%	64%	66%						
	Predicted Six Year Graduation Rate (Average)	70%	72%	72%	72%	73%	73%	74%	73%	72%	74%	74%
	6 Year Gap	-6%	-6%	-6%	-7%	-7%						
	Number of Student Records	925	1,128	1,341	1,443	1,495	1,654	1,551	1,790	2,049	2,293	2,217

The Predicted Six Year Graduation Rate (Average) for No Breakout and No Filter students for the Fall 2018 cohort is 74%.

Breakout Parameter

No Breakout ▾

Filter Parameter

No Filter ▾

Filter Parameter Group

(All) ▾

Instructions:

This table displayes actual and average predicted 6 year graduation rates. It also computes a gap between the predicted and actual rate (actual rate minus predicted rate). Zero or positive values indicate that we have met or exceeded the predicted rate. Negative values indicate that we have not met the predicted rate. The table also displays the number of records used in the calculation.

To modify the visualization:

(1) Select a Breakout Parameter from the list if you want to make a side-by-side comparison between students based on a particular characteristic. Select "No Breakout" if you do not want to make a side-by-side comparision.

(2) Select a Filter Parameter to filter the data to include only students based a particular characteristic. Select "No Filter" if you do not want to filter the data.

(3) If you selected a Filter Paramenter, next use Filter Parameter Group to select the specific value(s) of the characteristic of the students you want to include in the visualization. The "All" box should be selected if no filter has been applied.

Summary

Shared one approach for developing predicted graduation rate models for your campus.

Shared how to use the model results to develop a dashboard. Can be used to engage campus leaders develop a strategic plan for student completion.

Lessons Learned About Dashboards

- Get feedback from colleagues (not just in IR)
- (Ironically) Increases engagement between campus stakeholders and IR office

  - "How do I find the information I need?" (the firehose problem)
  - Changes questions from "what" to "why"?
  - Demos are important for increasing comfort with self-service
  - Users still want personal connection

Future Directions

- UC system 2030 strategic planning
- Collaborations between system office and campuses can be important

Resources

Logistic Regression

- Percontor workshop - Exploring the power of predictive analytics: A step-by-step introduction to building a student-at-risk prediction model
- The Analysis Factor workshop - Logistic Regression for Binary, Ordinal, and Multinomial Outcomes

Tableau

- Free training webinars
- Community resources and forums

Questions?

Please complete your session evaluation using the CAIR app!

Thank you!

Contact us:
Cinnamon Danube – cdanube@ucmerced.edu
Tongshan Chang - Tongshan.Chang@ucop.edu
Corinne Townsend – ctownsend@ucmerced.edu

Extra Slides

Statistics of Logistic Regression Models (see the notes and broad fields tabs for detailed information about the methodology)

Predictor_Type	Predictor	Estimate	Prob	Std Err
Intercept	Intercept	-0.639	0.000	0.027
Gender	Female (male as reference)	0.519	0.000	0.013
Ethnicity	African American (White as reference)	-0.322	0.000	0.034
	American Indian (White as reference)	-0.336	0.000	0.075
	Asian/Pacific Islander (White as reference)	0.076	0.000	0.016
	Hispanic/Latino(a) (White as reference)	-0.330	0.000	0.020
	Other/Unknown (White as reference)	-0.132	0.000	0.037
First-Generation	Unknown (not first-generation as reference)	-0.067	0.050	0.034
	Yes (not first-generation as reference)	-0.056	0.000	0.016
Pell Status	Yes (not Pell as reference)	-0.215	0.000	0.015
Residency	International (CA resident as reference)	-0.256	0.000	0.029
	Nonresident domestic (CA resident as reference)	0.186	0.000	0.047
High School GPA	>0 - <3.25 (3.40 - < 3.68 as reference)	-0.328	0.000	0.027
	3.25 - <3.40 (3.40 - < 3.68 as reference)	-0.228	0.000	0.028
	3.68 - < 3.95 (3.40 - < 3.68 as reference)	0.222	0.000	0.018
	3.95 - <4.16 (3.40 - < 3.68 as reference)	0.407	0.000	0.019
	4.16 - < 4.29 (3.40 - < 3.68 as reference)	0.604	0.000	0.025
	4.29 - < 4.33 (3.40 - < 3.68 as reference)	0.744	0.000	0.046
	4.33 or above (3.40 - < 3.68 as reference)	0.864	0.000	0.041
	Unknown (3.40 - < 3.68 as reference)	0.263	0.000	0.057
Test Score	400 - < 1000 (1000 - <1200 as reference)	-0.338	0.000	0.021
	1200 - < 1400 (1000 - <1200 as reference)	0.095	0.000	0.016
	1400 or above (1000 - <1200 as reference)	0.014	0.571	0.024
	Unknown (1000 - <1200 as reference)	-0.889	0.000	0.217
Pre UC Units	Pre UC Units (numerical)	0.019	0.000	0.001
Discipline	Arts (Engineering/Computer Science as reference)	0.658	0.000	0.036
	Health Sciences (Engineering/Computer Science as reference)	1.041	0.000	0.073
	Humanities (Engineering/Computer Science as reference)	0.734	0.000	0.032
	Life Sciences (Engineering/Computer Science as reference)	0.395	0.000	0.037
	Math/Physical Sciences (Engineering/Computer Science as reference)	0.255	0.000	0.026
	Professional Fields (Engineering/Computer Science as reference)	0.990	0.000	0.031
	Social Sciences (Engineering/Computer Science as reference)	0.839	0.000	0.023
	Undeclared (Engineering/Computer Science as reference)	0.476	0.000	0.019