
Challenges
Solutions
Prezi

Kara Larkan-Skinner

Frances Frey

Using Predictive Models to Improve Student Success

Prezi

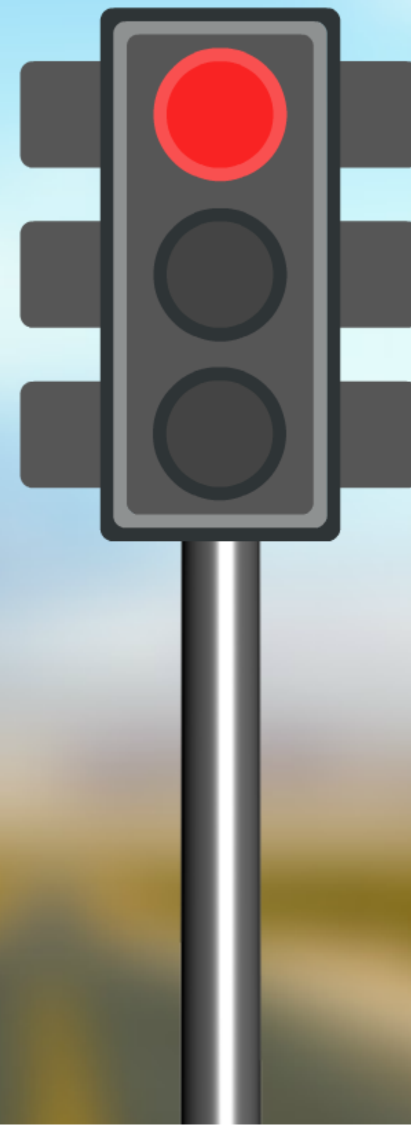

Challenges
Solutions
Prezi

Low Graduation Rates

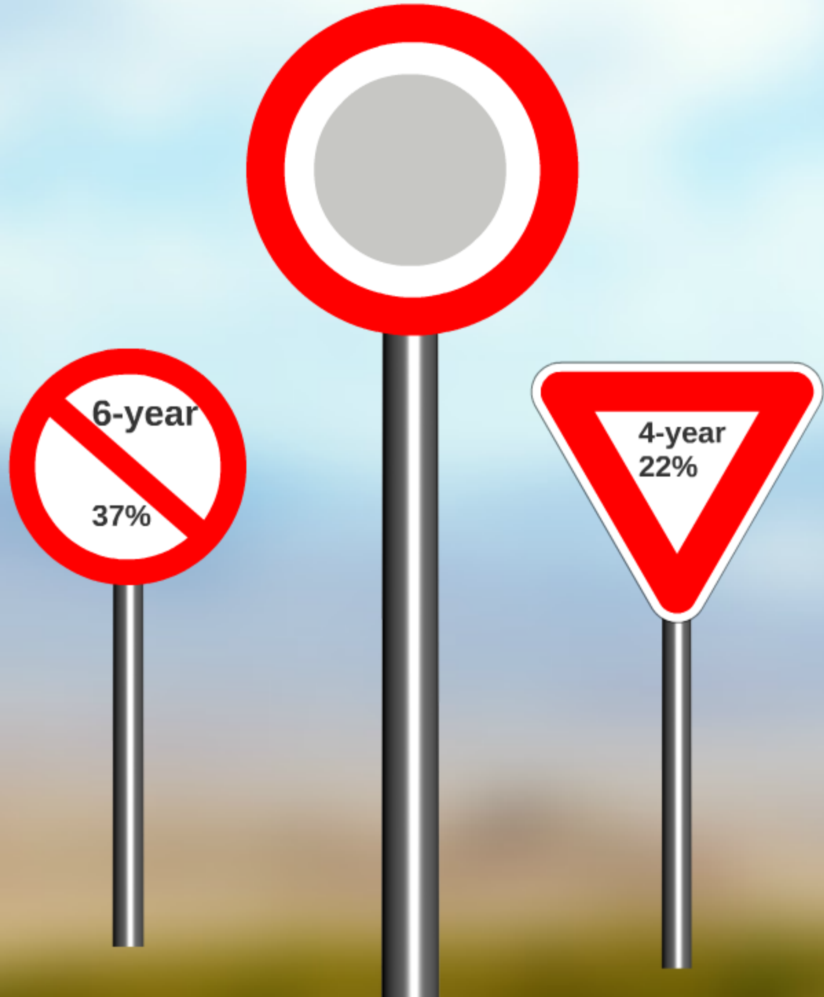
High Drop/Fail Rates

High Student Debt

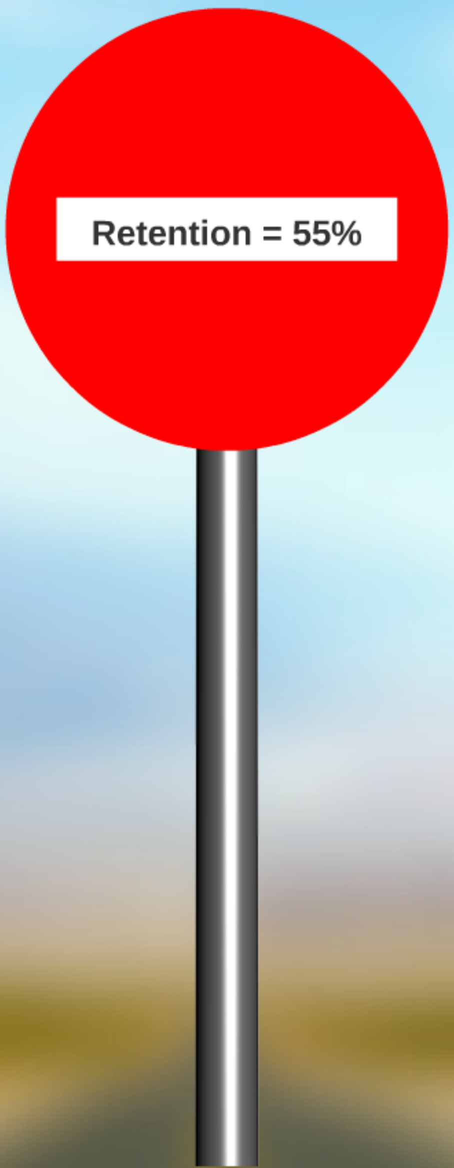
High % of At-Risk Students



Low Graduation Rates



Low Fall-to-Fall Retention



High Risk Population

Incoming Student Characteristics

90% of Students At-Risk:

Average SAT = 918

Average ACT = 19

50% First Generation

72% Hispanic

98% on Financial Aid

70% Pell Eligible

35-40%
Losing State Aid
1st Semester
When GPA Drops
Below 2.5

Need a way to identify most at-risk students before they lose aid and leave.


Challenges
Solutions

- Collaboration with Student Success office
- 1st academic year challenges
- Strategic Use of Resources
- Leverage Data to Improve Success
- Use predictive analysis to determine students most at risk before they even step foot on campus

Predicting At-Risk Students

Logistic Regression

- Dichotomous
- Probability score
- Predict group membership
- Aligned with goals

Initial Attempt to Predict Retention

- Unable to predict using pre-entry variables
- First term GPA highly predictive of retention
- Generated new research question

Prezi

Building Logistic Regression Model

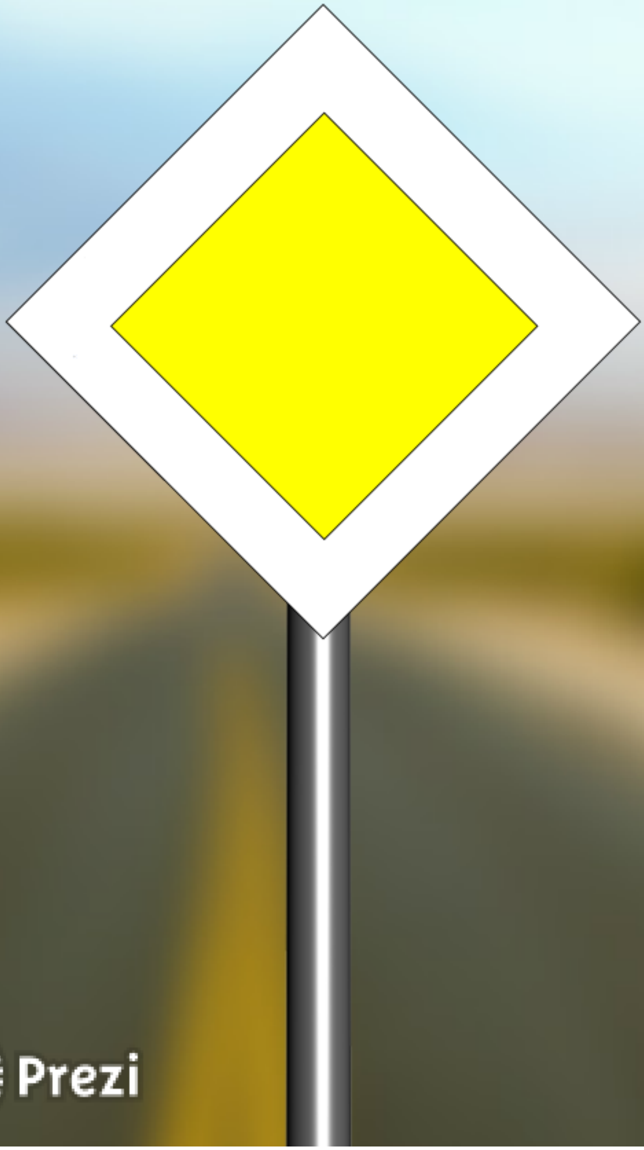
Test Model

Dependent variable

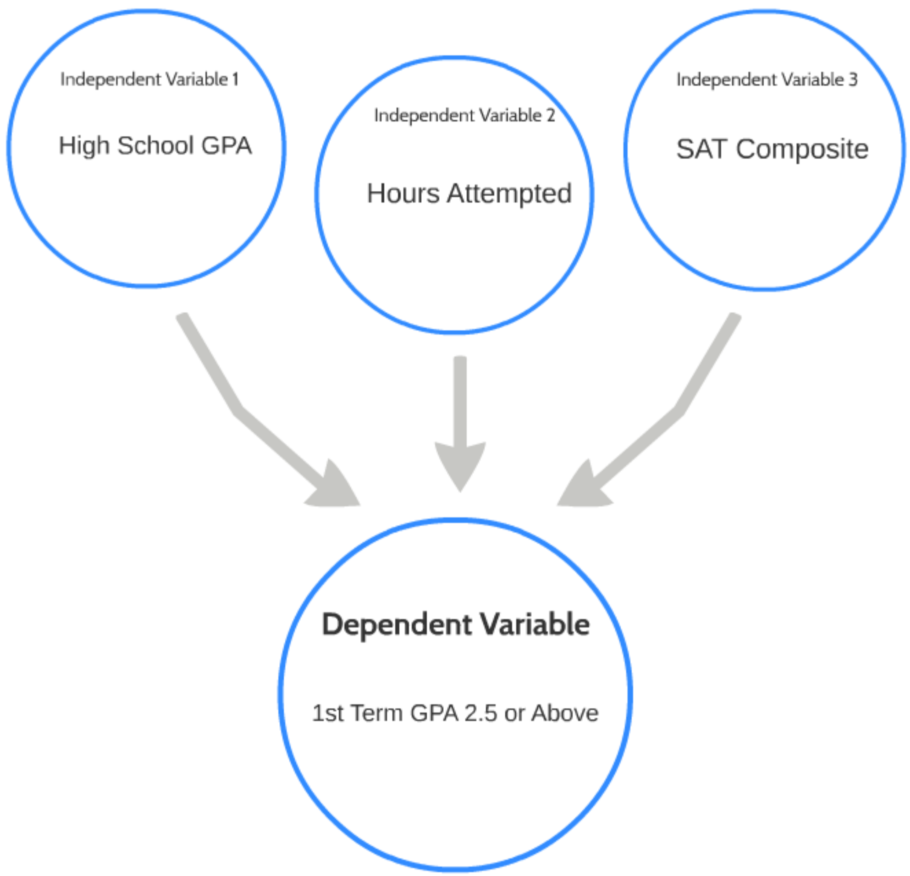
- 1st Term GPA 2.5 or above (yes or no)

Independent variables

- Financial factors
- High school performance factors
- College entrance exams
- Demographics



Final Model



Independent Variable 1
High School GPA
Independent Variable 2
Hours Attempted
Independent Variable 3
SAT Composite
Dependent Variable
1st Term GPA 2.5 or Above

Model Summary
Fall 2012 Data

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
3	345.488[a]	.211	.306

Classification Table[a]

			Predicted		
			Term GPA >2.499		Percentage Correct
Actual			.No	Yes	
Step 3	Term GPA >2.499	No	35	65	35.0
		Yes	20	253	92.7
	Overall Percentage				77.2

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 3[c]	SAT Comp.	.002	.001	3.904	1	.048	1.002
	HS GPA	1.608	.333	23.316	1	.000	4.993
	Hours Enr.	.366	.071	26.829	1	.000	1.442
	Constant	-11.359	1.586	51.294	1	.000	.000

Results:
Model = 77% Accurate Fall 2012 Data;
Model = 78% Accurate Fall 2013 Data

Quick Technical Note on Logistic Regression (you can close your eyes and hum for this part)

- With a logistic regression, you are essentially modeling the probability that an individual will fit into one of two groups.
- However, if you use linear regression on a dichotomous variable, you get probabilities > 1.
- To fix this, the dependent variable is transformed
  - First into odds (odds = (Probability of Low Risk)/(Probability of High Risk)), which makes the range 0 to ∞
  - Then into logits(natural log(odds)), which makes the range - ∞ to ∞


OLLU's Model
Practical
Conceptual
Prezi

Created ‘Prediction Template’:

(Step 1)
Logit(probability of Low Risk) = Constant +(B*High School GPA) +(B*Hours)+ (B*SAT)

(Step 2) Convert logit into odds

(Step 3) Convert odds into probabilities
Probability = Odds/((1+Odds))

(Step 4)
Determine your threshold
Reduce number of false positives (i.e., “low risk”)

Prezi

predictive model for retention developed and applied - Microsoft Excel

	A	B	F	G	H	I	J	K	L	M	N	O	P
1	ID	Last	Active Reg	SATC	HS GPA	Formula	ExpNum	ExpDen	Probability	Actual Fall Term GPA	Actual Term GPA >=2.5	Predicted Term GPA >=2.5	Model Accuracy
78			SAMPLE			=-11.359+(1.608*H79)+(0.002*G79)+(0.366*F79)	=EXP(I79)	=1+J79	=J79/K79				
79	1	Smith	15	1300	4.6	4.1278	62.04128186	63.04128186	**98.41%**	4	1	1	1
80	2	Smith	15	1300	4.648	4.204984	67.01952535	68.01952535	**98.53%**	4	1	1	1
81	3	Smith	12	1300	4.99	3.65692	38.74183407	39.74183407	**97.48%**	4	1	1	1
82	4	Smith	14	1260	4	2.717	15.13484952	16.13484952	**93.80%**	3.714	1	1	1
83	5	Smith	13	1230	3.7	1.8086	6.101898791	7.101898791	**85.92%**	3.769	1	1	1
84	6	Smith	16	1210	3.8	3.0274	20.64348968	21.64348968	**95.38%**	3.813	1	1	1
85	7	Smith	14	1190	3.6	1.9338	6.915740186	7.915740186	**87.37%**	0.214	0	1	0
86	8	Smith	15	1160	3.69	2.38452	10.85385157	11.85385157	**91.56%**	3.8	1	1	1
87	9	Smith	12	1150	3.56	1.05748	2.879106491	3.879106491	**74.22%**	2.333	0	1	0
88	10	Smith	13	1140	3.2	0.8246	2.280968196	3.280968196	**69.52%**	3.308	1	1	1
89	11	Smith	13	1140	3.9	1.9502	7.030093459	8.030093459	**87.55%**	3.692	1	1	1
90	12	Smith	13	1140	3.296	0.978968	2.661707941	3.661707941	**72.69%**	1.846	0	1	0
92	14	Smith	15	1130	4	2.823	16.8272568	17.8272568	**94.39%**	3.8	1	1	1
93	15	Smith	13	1130	3	0.483	1.620929905	2.620929905	**61.85%**	3.538	1	1	1
94	16	Smith	13	1110	3.5	1.247	3.479887619	4.479887619	**77.68%**	2.308	0	1	0
95	17	Smith	14	1100	3.7	1.9146	6.784224564	7.784224564	**87.15%**	4	1	1	1
96	18	Smith	18	1100	2.7	1.7706	5.87437693	6.87437693	**85.45%**	3	1	1	
97	19	Smith	14	1090	3.09	0.91372	2.493581424	3.493581424	**71.38%**	3.357	1	1	
98	20	Smith	13	1090	2.33	-0.67436	0.509482385	1.509482385	**33.75%**	0.923	0	0	1
99	21	Smith	18	1080	3.6	3.1778	23.99390884	24.99390884	**96.00%**	4	1	1	1
100	22	Smith	14	1080	3.471	1.506368	4.510319529	5.510319529	**81.85%**	4	1	1	1

SAT GPA HRS Predict Model | FL13NFR | FL13TermGPAActual | Original Model Rough | Sheet5 | Sheet7 | Sheet6 | Final M

Prezi

	A	B	C	D
1	**Retention Alert Model**			
6	**Predictive GPA Model**			
7	HSGPA	3.06		
8	SAT	850		
9	First Sem Reg Hrs	12		
10				
16	**Probability > 2.5**		**69.86%**	
17				
18				
19				
20				

End User View

Organizational Impact

- Enrollment Management
- Student Success Division
- 3rd Party Vendor

Prezi

Results

- 10% increase in 2.5+ GPA
- Fall-Fall retention increased 8%
- Estimated $500,000 - $1M savings
- 77-78% model accuracy

Next Steps

Monitor Model

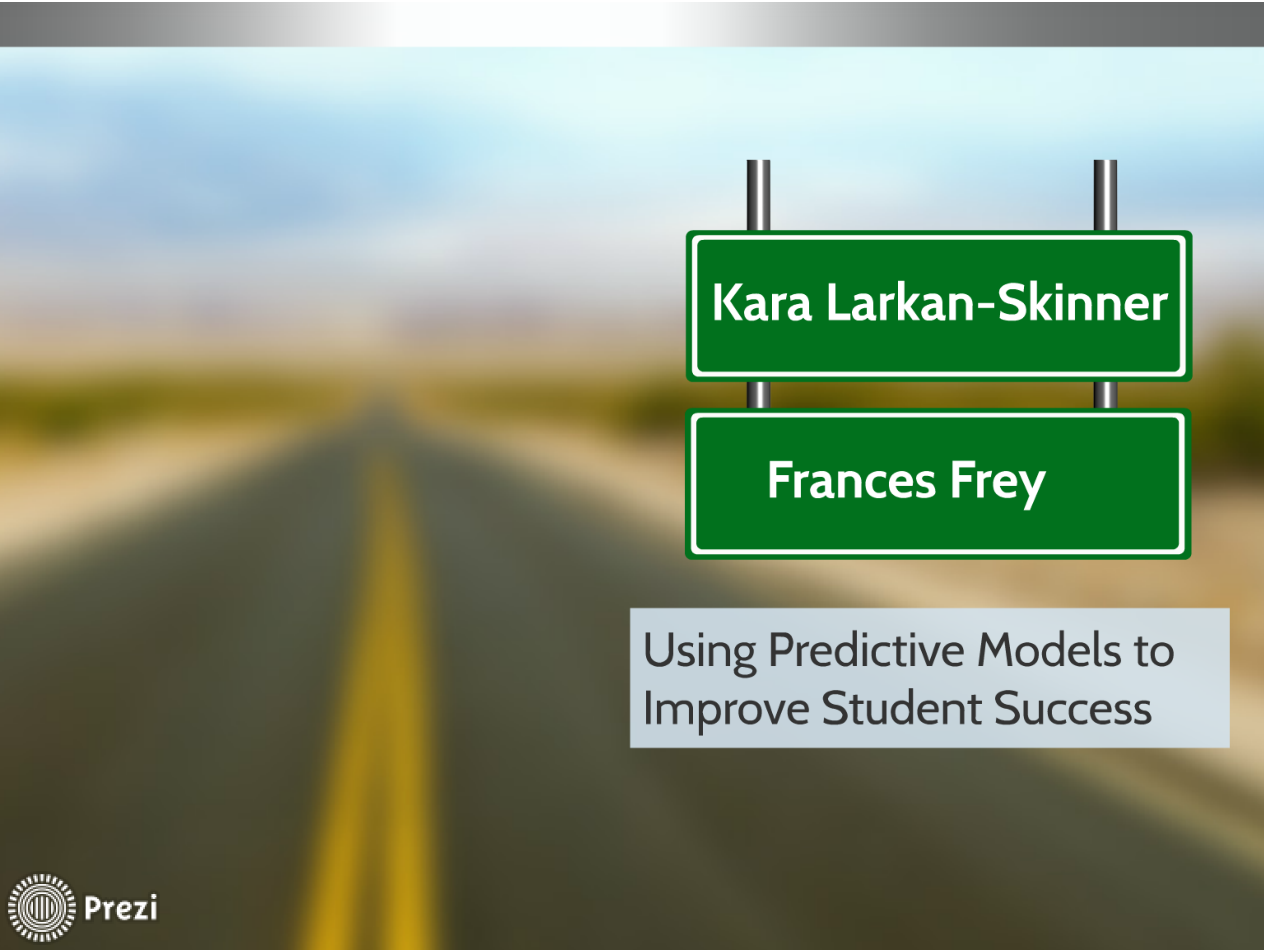
- Long-term impact
- Triage effectiveness

Other Predictions

- Online courses
- Enrollment

Advocate for Policy Change

- Financial aid
- Advising


Kara Larkan-Skinner
Frances Frey
Using Predictive Models to Improve Student Success
Prezi