

# Modeling and Visualizing First to Second Year Retention

Resche Hines, PhD, Assistant Vice President for Institutional Research & Effectiveness  
Angela Henderson, Director of Institutional Research & Effectiveness



## Abstract

A model based on historical first to second year retention was developed to examine factors which significantly impact retention at a small liberal arts university. Findings from the retention model were imported into data visualization software (Power BI) to create an interactive report that allows users to examine the likelihood of retention based on a number of variables found to be significant predictors.

## Objectives

- To develop and implement a first-time in college (FTIC) first-to-second year retention model which informs the effective use of resources,
- to determine retention risk groups within the FTIC population, and
- to provide a visually engaging system to facilitate discussion around data driven decision making for student success.

## Population

- FTIC cohorts for Fall 2011, 2012, 2013, and 2014 as training group (n=3,158) with Fall 2015 FTIC students (n=983) as predicted group.

## Variables

- Dependent: retention to 2nd fall
- Predictor variables:
  - Demographic: student age, gender, ethnicity, residency
  - Academic: academic index score, undeclared major, first term earned credits, first term college GPA
  - Social: athletics or Greek life
  - Economic: total family financial contribution, unmet financial need, tuition discount received

## Methodology

- Simultaneous multiple regression on predictor variables to determine the best predictors of retention
- Logistic regression analysis to assess whether resulting predictor variables significantly predicted retention.
- Cut value set to baseline retention probability to maximize correct classification
- Scatterplot of Cook's influence to examine outliers and ROC curve to examine goodness-of-fit.
- Predicted values, including probabilities, group membership, and analogs of Cook's influence calculated and added to dataset.
- Visual binning to create risk groups based on the predicted probabilities.
- Binning cut points based on mean and +/-1 std. deviation.

## Results

5 variables found to be predictors of retention: first term GPA, out of state, unmet financial need, tuition discount received, and Greek life

- Twenty-four outliers with Cook's influence > .025 were identified and excluded.

### Significant Predictor Variables of Retention

Variable	B	S.E.	df	Sig.	Exp(B)
Out of State	-.462	.103	1	.000	.630
Unmet Need	-.256	.100	1	.010	.774
Discount Received	3.201	.566	1	.000	24.566
1 <sup>st</sup> Term GPA	1.193	.063	1	.000	3.296
Greek	.894	.125	1	.000	2.444
Constant	-5.197	.594	1	.000	.006

- All five predictor variables considered together significantly predicted retention ( $X^2 = 648.433$ ,  $df = 5$ ,  $n = 3,134$ ,  $p < .001$ ).

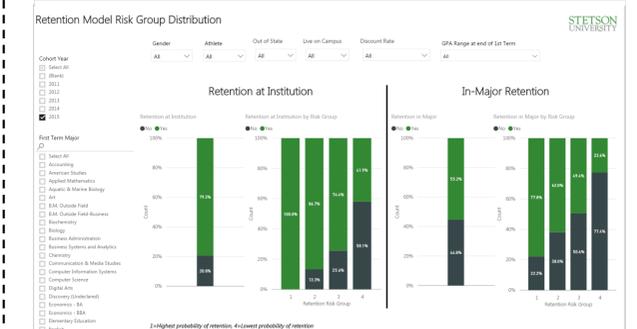
71.1% of 2011-2014 FTIC students who were not retained were classified correctly by the model

## Fall 2015 Predictions

- Visual binning was used to place Fall 2015 FTIC students into predicted retention probability risk groups based on the model
  - 1: 96.2% or higher prob. of retention
  - 2: 77.7% to 96.1% prob. of retention
  - 3: 59.6% to 77.6% prob. of retention
  - 4: Less than 59.5% prob. of retention

- Risk groups data shared with campus through an interactive retention model visualization.

### Visualization of Retention Risk Model Data



- Predictions by risk group were compared to actual retention of Fall 2015 FTIC cohort

### Fall 2015 FTIC Predicted & Actual Retention Rates

Risk Group	n	Predicted* retention rate	Actual retention rate
1	18	100.0%	100.0%
2	676	83.7%	86.7%
3	164	69.5%	74.4%
4	124	37.1%	41.9%
Total	982	75.8%	79.2%

1= highest probability of retention;  
4=lowest probability of retention.  
\*predicted as of July 1, 2016

- Risk groups 2, 3, and 4 retained at slightly higher rates than predicted by the model.
- This may be due to interventions increased stemming from the visualization of the model and identification of students within each risk group.

### Contact Information

Resche Hines, Stetson University,  
[rhines@stetson.edu](mailto:rhines@stetson.edu)  
Angela Henderson, Stetson University,  
[ahender1@stetson.edu](mailto:ahender1@stetson.edu)