# Institutional Intelligence: Using Predictive Analytics for Student Success

Swagata Kundu, Carlos Rivers and Dr. Shonda Gibson
Texas A&M University-Commerce

Introduction

Predictive analytics encompasses a variety of statistical techniques from predictive modeling, machine learning and data mining that analyze current and historical data to make predictions about future or otherwise unknown events. It can be applied in higher education to predict student enrollment, retention, attrition and calculate financial aid outlay. Predictive modeling is the use of known results to develop a model that can be used to predict values for new data. Modeling provides results in the form of predictions that represent a probability of a target variable based on a set of input variables.
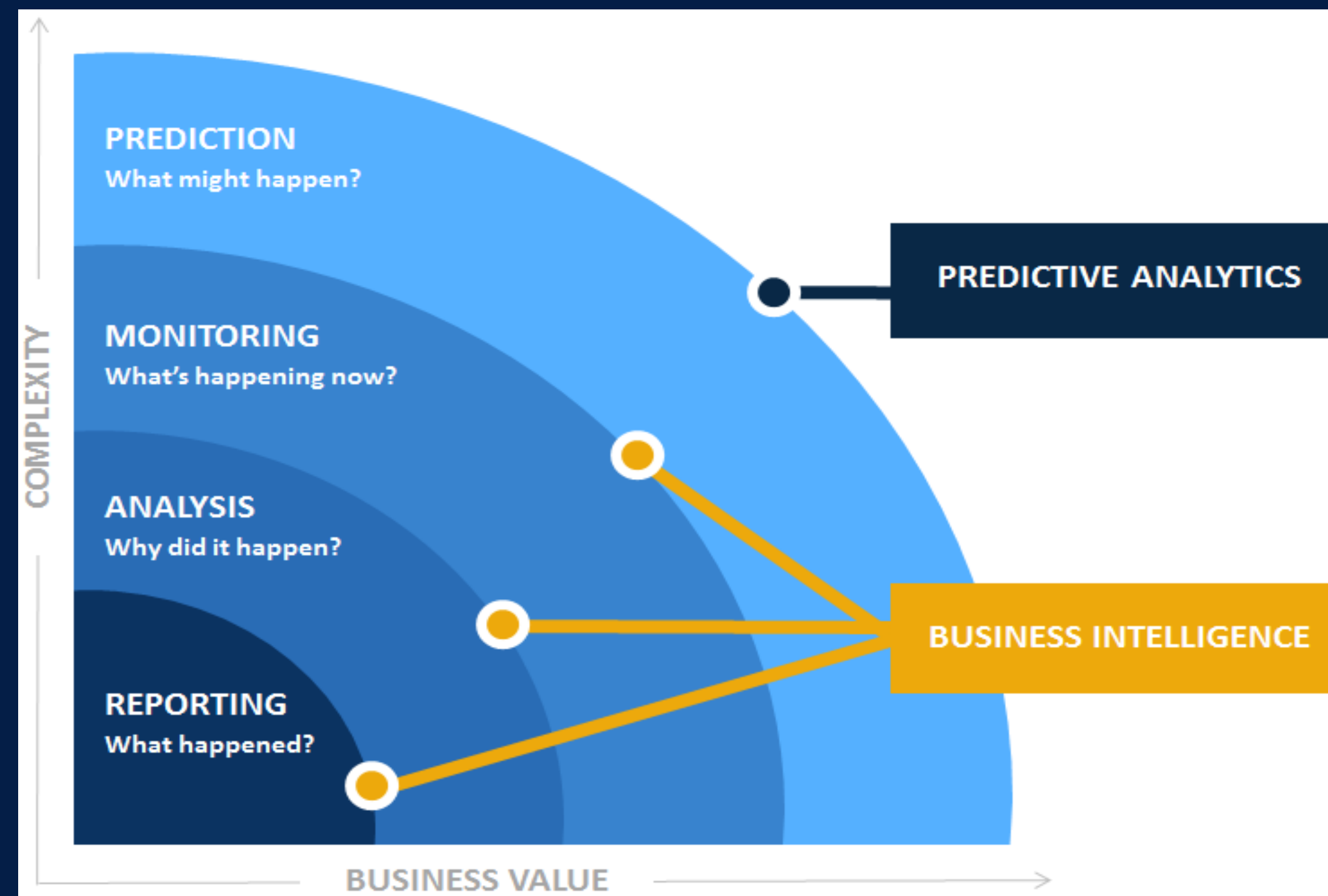
Purpose and Focus of the Study

To identify contemporary challenges faced by institutions in higher education and allow institutional stakeholders to address these challenges to make better informed business decisions that can improve organizational efficiency and effectiveness using predictive analytics. To explore the use of predictive analytics in higher education to improve student access and success and to examine how historical data can be used to identify persistence trends over time.

Methodology

Two student populations were explored: For first-time, full-time freshmen (FTFT-FR), we analyzed persistence patterns from our predictive analytics application. We identified the top variables most predictive of student success, and then measured persistence across two variables (HS Percentile and Max ACT Composite). Using Excel and pivot table analysis, we mined seven years of historical student data and performed a what-if analysis to compare what might have happened if we had used high school percentile during the last seven years instead of awarding scholarships based on test scores. On a second study, we analyzed persistence curves of traditional transfer students across two variables (Credits Transferred In and Prior Term GPA). For CBE transfer students, using Excel and pivot tables, we calculated retention rates and time-to-degree indicators as a point of comparison.

Findings

Is High School Percentile a better indicator of Student Persistence over test scores?

### Top 10 predictors of persistence for first time full time students at A&M

Previous Scholarships awarded by Test Scores

Scholarship	ACT/SAT Scores	Scholarship Amount
Presidential	28/1270	$2,500 per year/$1,250 per semester
Academic Excellence	25/1180	$1,500 per year/$750 per semester
Blue & Gold	23/1050	$1,000 per year/$500 per semester

We found that high school class percentile was ranked in the top 5 most powerful predictors for FTFT-FR. Performing a quick query for both SAT and ACT predictors, no test score predictor appeared until a distant 34th position in the predictive ranking.

Persistence probabilities for first-time full-time students by high school class percentile
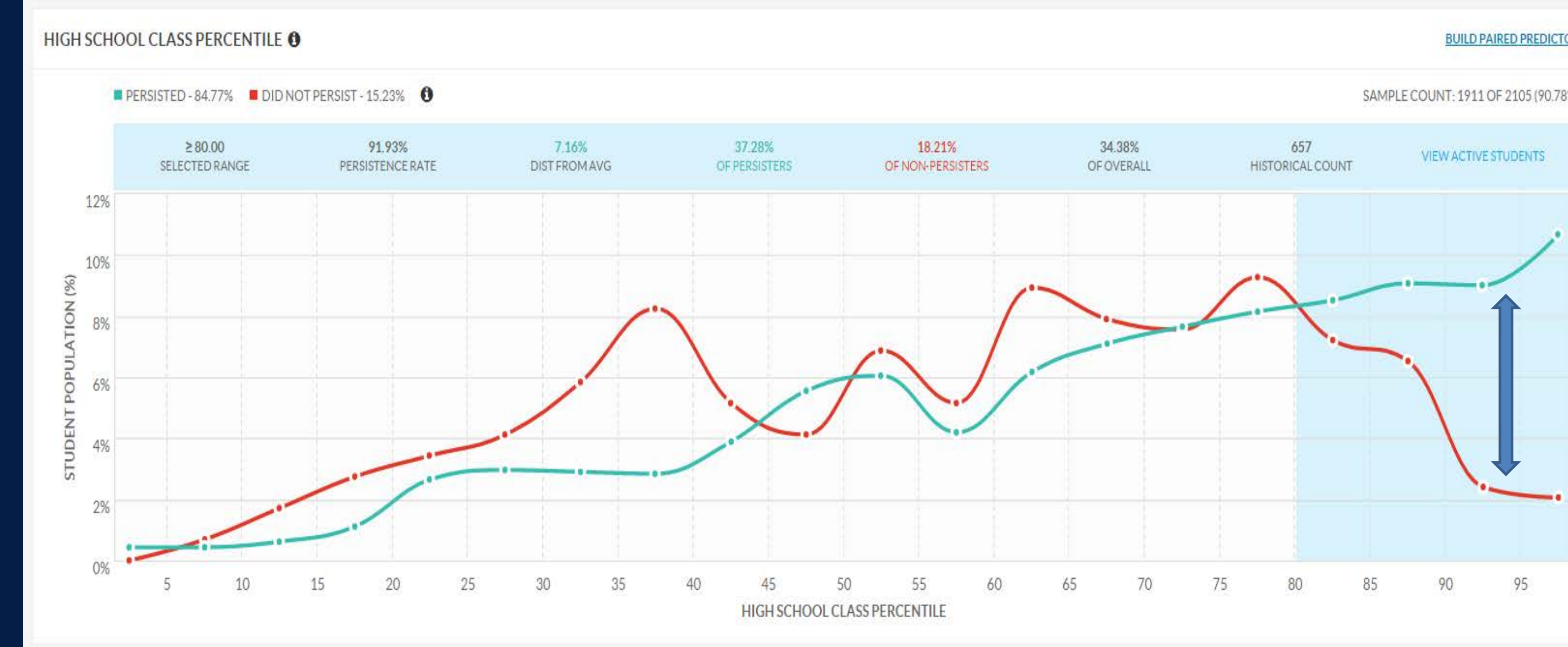
At or above the 80th percentile, this predictor showed an inverse relationship: persistence dramatically increases and non-persistence decreases. Students at or above the 80th percentile persist 7.16 percentage points above the average. There seems to be a strong correlation between persistence and high school class percentile at our A&M-Commerce for this student population.

The figure at the bottom is using a paired predictor plot tool comparing High School Class Percentile (X) *with* Max ACT Composite Score (Y). The top right quadrant is telling us that a combination of both predictive indicators is the best indicator of persistence. The top left quadrant shows that Max ACT Composite by itself is 5.19 percent below the average persistence for this student population. The bottom right quadrant demonstrates that High School percentile by itself is slightly above the average historical persistence for these students. We ran the same analysis multiple times with SAT predictive indicators as well, displaying similar results.

Max ACT vs. High School Class Percentile
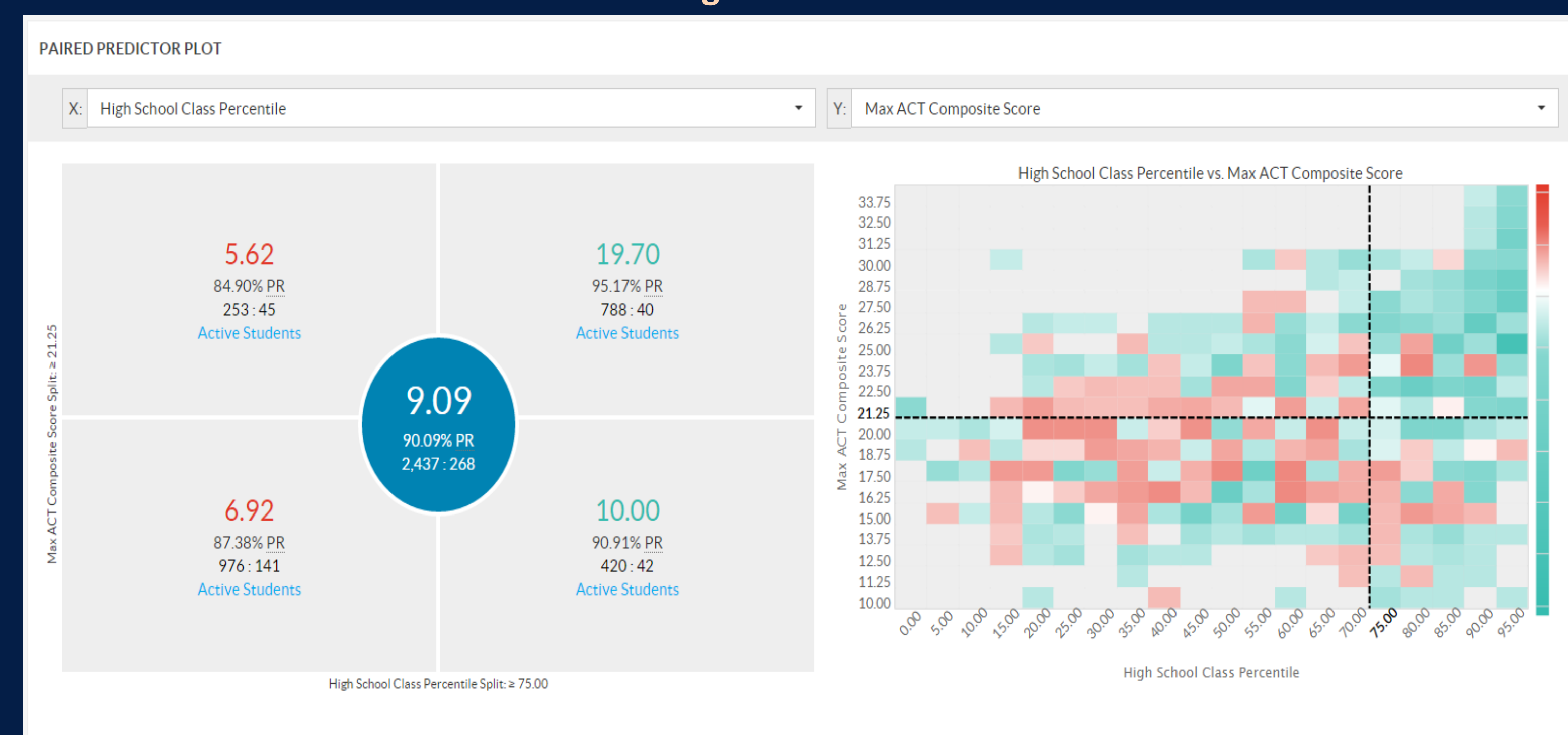
- High school class percentile by itself is slightly above the average historical persistence for first-time, full-time freshmen.
- The heat map on the right shows that the highest concentration of persistence is in the bottom right quadrant at or above 75 percentile or above.

## Findings from What-If Analysis on 7 years of historical data:

- We could have awarded scholarships, on average, to six percent more students in each first-time freshmen cohort.
- We would have awarded scholarships, on average, to 45 percent more minority students in each first-time freshman cohort, with a particularly large boost in funds for new Hispanic students.
- Our retention of first-time freshmen would have been at least three percentage points higher after three years.

A new policy was implemented to better serve our growing and diverse student population. A couple of external charts added below, show how minority students are outperform by white students in every single test score metric in Dallas ISD. However, enrollment shows an important demographic shift in the state, were minorities now represent 95% of all enrolled students. Our data analysis shows that using high school percentile would be a more inclusive policy for a majority of our incoming students.

New Scholarships awarded by H.S Class Percentile

| Scholarship	High School Percentile	Amount
Presidential	Top 10%	$2,750 per year/$1,375 per semester
Academic Excellence	11%–25%	$1,500 per year/$750 per semester

### Dallas ISD SAT/ACT results by student Demographics
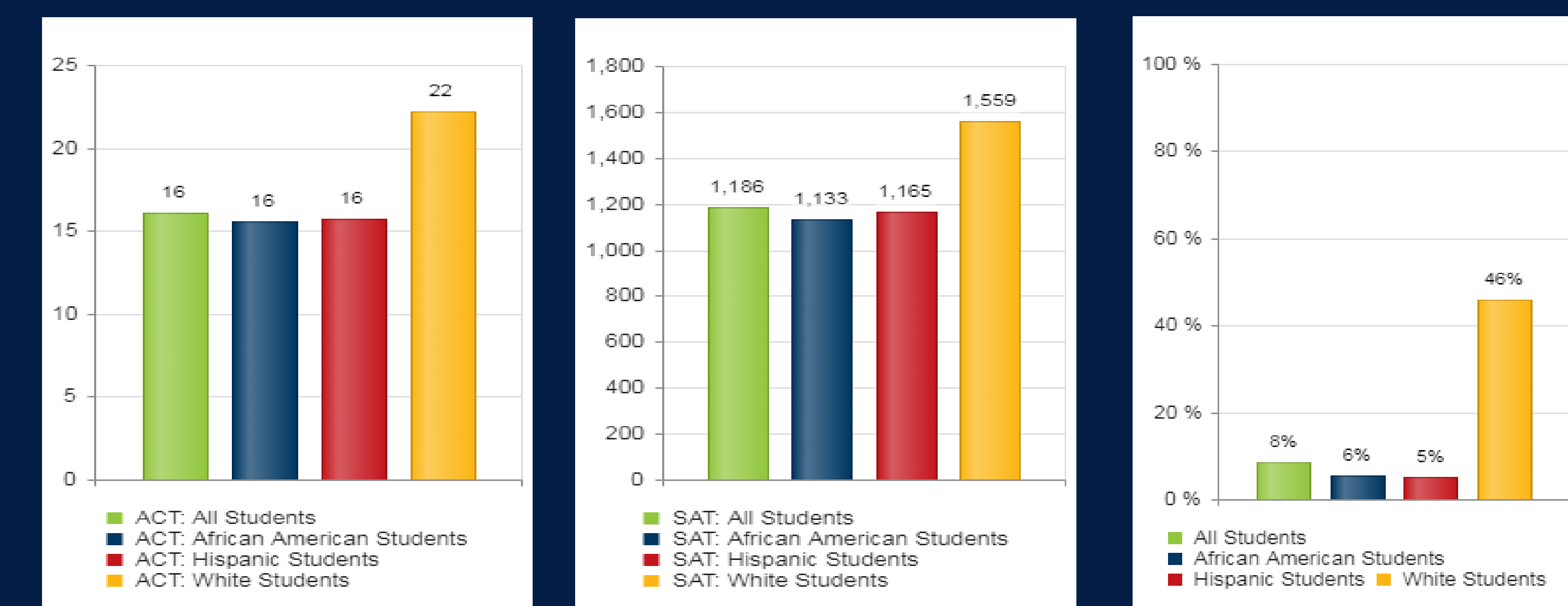
Are Competency-Based Educations Transfers persisting at higher rates than traditional transfers?

- Historically, traditional transfer students who only take 6 or less credit hours (Part-Time) represent more than 52% of all non-persisting students.
- On average CBE students are taking 6 SCHs per term and are being retained at 87%.

Persistence probabilities for traditional transfer students by credits earned the prior term
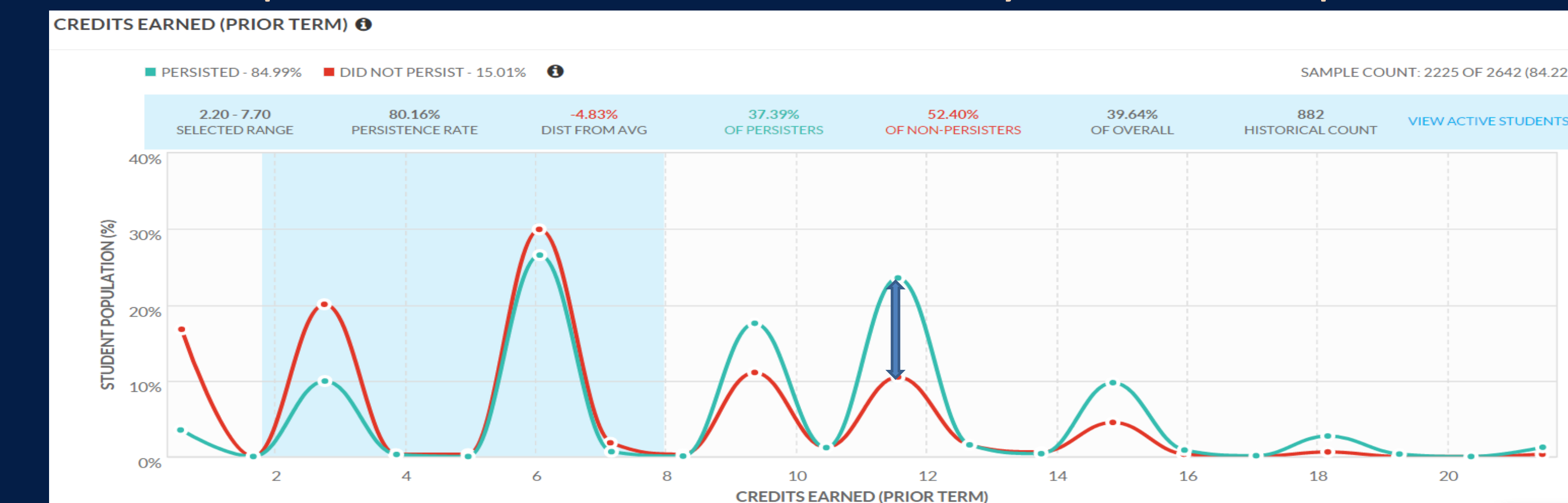
Persistence probabilities for traditional transfer students by Transfer GPA Credits (Cumulative)
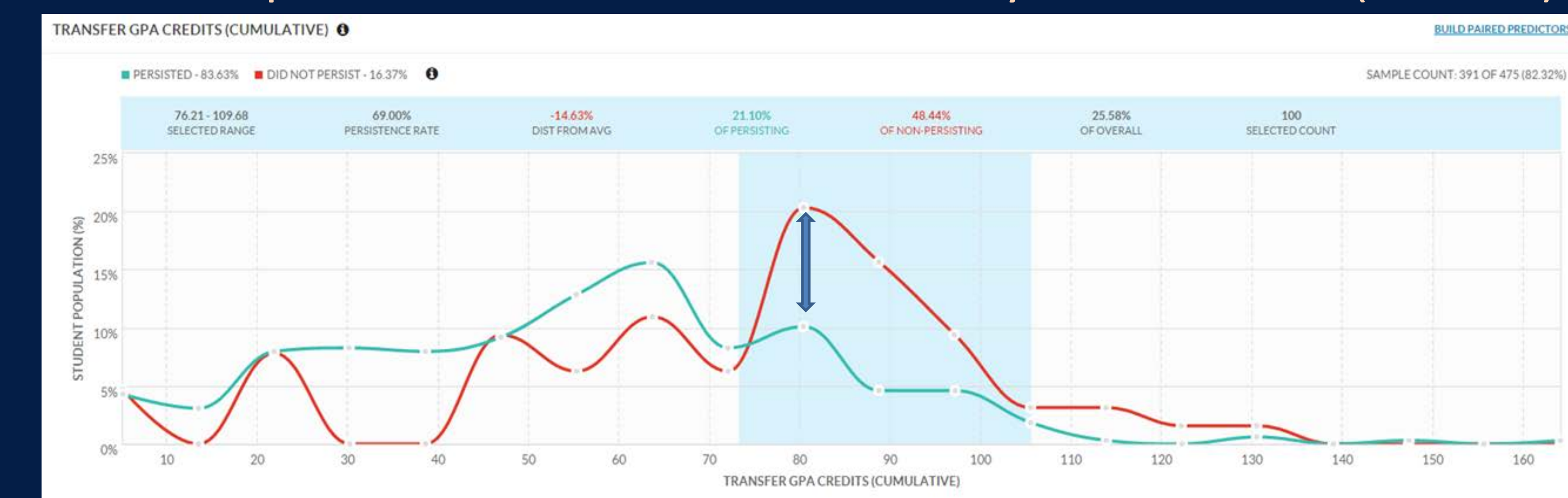
- Historically, traditional transfer students who bring 80 or more credits represent more than 48% of all non-persisting students.
- CBE program graduates transferred in 85 SCHs on average, and graduated in less than a year from the program.

Conclusion

The examples above demonstrate how predictive analytics has changed the landscape in which institutions operate. Data usually has been used to tell us what has happened and improves strategic planning moving forward; however, with predictive analytics we can ask what might happen and act on that knowledge today, when it matters. With predictive analytics one can make more informed decisions about things we already know to be true and discover unknowns to which we need to adapt in order to enhance our abilities to improve student access and success in higher education.