CARNEGIE

NACCAP Financial Aid Optimization Event

Financial Aid Optimization Overview

FRIDAY, JUNE 23, 2023

Presenter

Nick Wagner

VP, Research + Data Science

+ 17 Years Experience in Higher Education
+ Predictive Modeling + Quant Background
+ Competitive + Considerate Shepherd
+ Michigan

A FEW STARTING DETAILS

CARNEGIE

+ 60 minute interactive discussion
+ Ask questions at any time
+ Start thinking about how you allocate financial aid to impact enrollment goals

TODAY'S GOAL(S)

CARNEGIE

+ Building an understanding of Financial Aid Optimization

+ A basic overview of predictive models that are helpful in achieving enrollment and institutional success

+ Identifying special considerations for Christian institutions and smaller colleges and universities

CARNEGIE

1

What is Financial Aid Optimization?

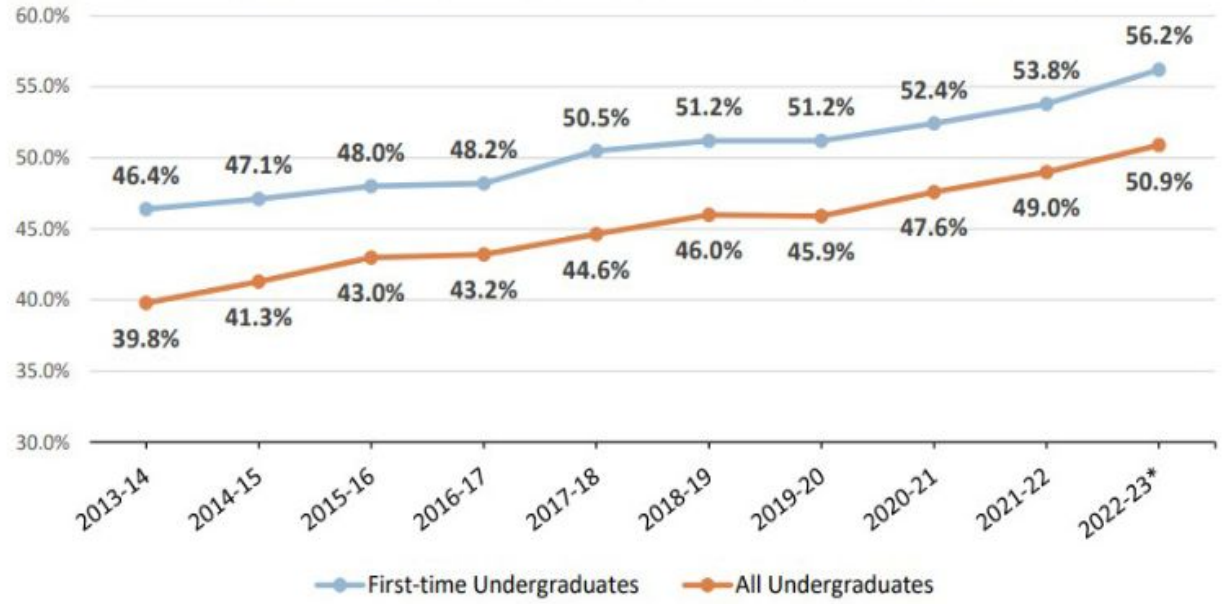
Term 1: Financial Aid Optimization

The use of predictive models to understand the ecosystem of student decision making.

Understanding the relationship between net cost (or institutional aid) and yield ensures that institutions can develop effective and sustainable systems for student recruitment in order to meet critical institutional goals.

Why is This Work Important

Figure 1: Average Institutional Tuition Discount Rate, by Student Category

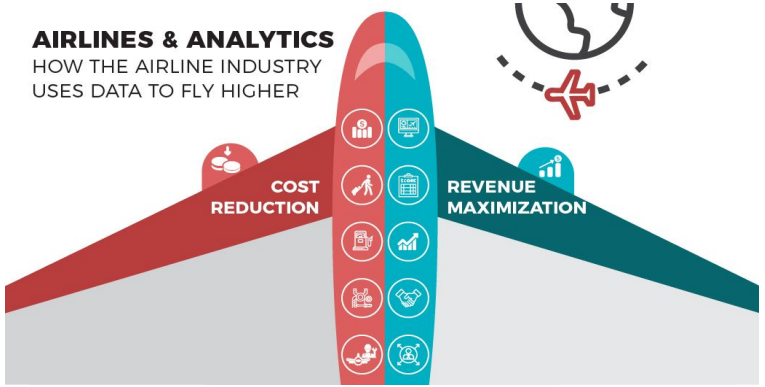


Source: NACUBO Tuition Discounting Study, data as of April 2023.
*Preliminary estimates.

Understanding Institutional Measures of EM Success

- What is being measured: Headcount, net tuition revenue, discount rate, do they all matter? What is known or most important?
- If it is NTR, financial aid optimization and the modeling behind it lends itself to awarding rules or specific packaging policies.
- If headcount is a priority, NTR can become compromised (spend more = more enrollments), but which students will be most impacted (in terms of their enrollment probability) by additional $?

First Exercise: Who uses Predictive Models?



How do we get there? Enrollment Predictive Modeling

Success relies on the coupling of data science and University enrollment knowledge

What is the purpose of enrollment modeling?

- Aims to capture trends hidden behind a 17/18-year-old's idiosyncratic decision making processes
- Help to manage large expenditure of institutional aid; plan strategic growth
- Identify factors and their relative importance to driving enrollment
- Monitor the enrollment cycle and make targeted and informed pivots
- Not an exercise in pure data mining; blends art and science
- Not a hypothetical exercise; serious implications

Analyze Historical Behaviors

Identify trends, account for subpopulation behavior, quantify hits and misses

Build Predictive Model

Exploratory analysis, test various machine learning models, optimize model for accuracy

Predict Future Outcomes & Monitor

Monitor enrollment cycle, proactively adjust

Not as Simple as it Seems

Enrollment modeling comes with a distinct set of challenges

- Imbalanced class data
- Relatively few records
- Risk of large year-over-year admit changes
- Relatively idiosyncratic decisions

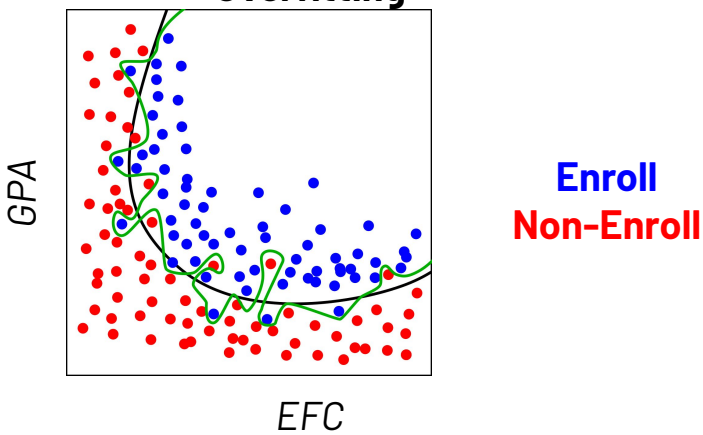
Context is everything

- What data is trustworthy?

Common pitfalls:

Spurious Relationships

Overfitting



What's in a Model?

$$Y = f(X) + e$$

Fixed But Unknown Function ← $f(X)$ | e → **Reducible and Irreducible Error**

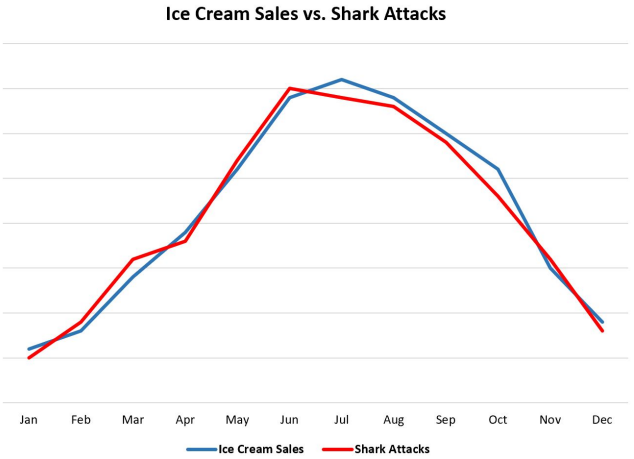
Goal: identify $f^*(X)$ that best approximates $f(X)$ while minimizing reducible error

Modeling helps to answer key questions

- What factors matter in a student's enrollment decision?
- Conversely, which factors do not?
- Which factors are related to each other? (e.g. local students ~ low income)
- How certain are we that a given factor is statistically significant?

Remember: modeling is never perfect

- Correlation is not causation
- Model can only ever explain portion of outcome



What Makes a Good Model Great

Data, Data, Data
- Quality and quantity of student-level data

Model Scope & Assumptions
- Not all students are the same, and neither are their enrollment decisions
- Remove outliers (e.g. dual enrolled students) whose relationship to yield is substantially unique

Feature Creation & Selection
- Parse out information-rich variables from others containing statistical "noise"
- Avoid unstable variables: minor changes may result in major yield implications

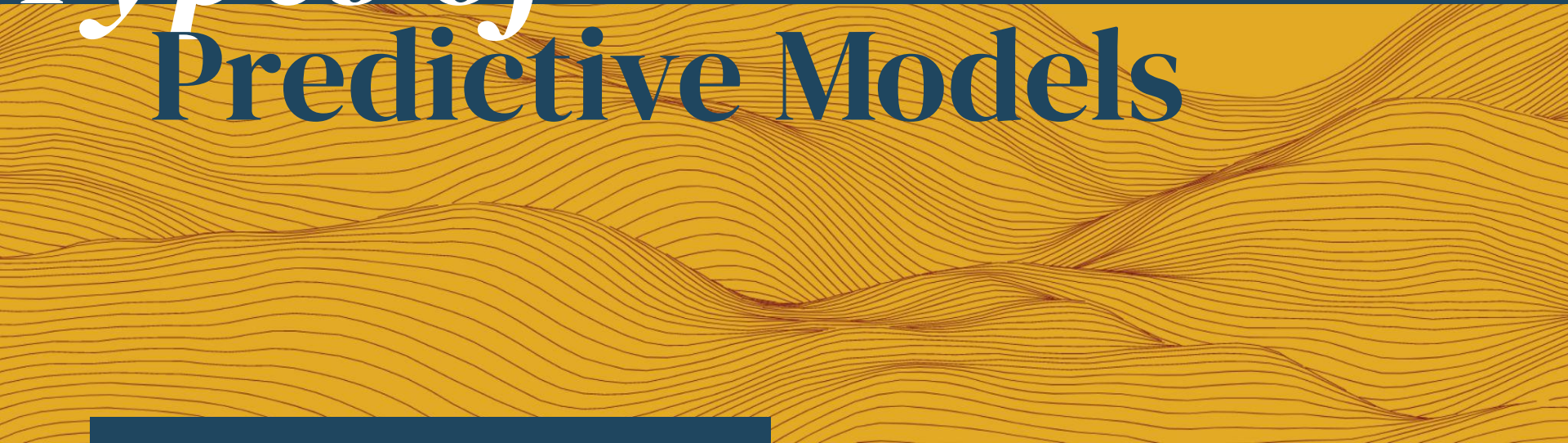
Robust Accuracy Metrics
- Bias vs variance trade-off
- Cross-validated predictive accuracy drives model selection (R^2, AIC, AUROC, etc.)

Reasonable Predictions
- Quantitative results versus qualitative experiential intuition
- What if? simulations

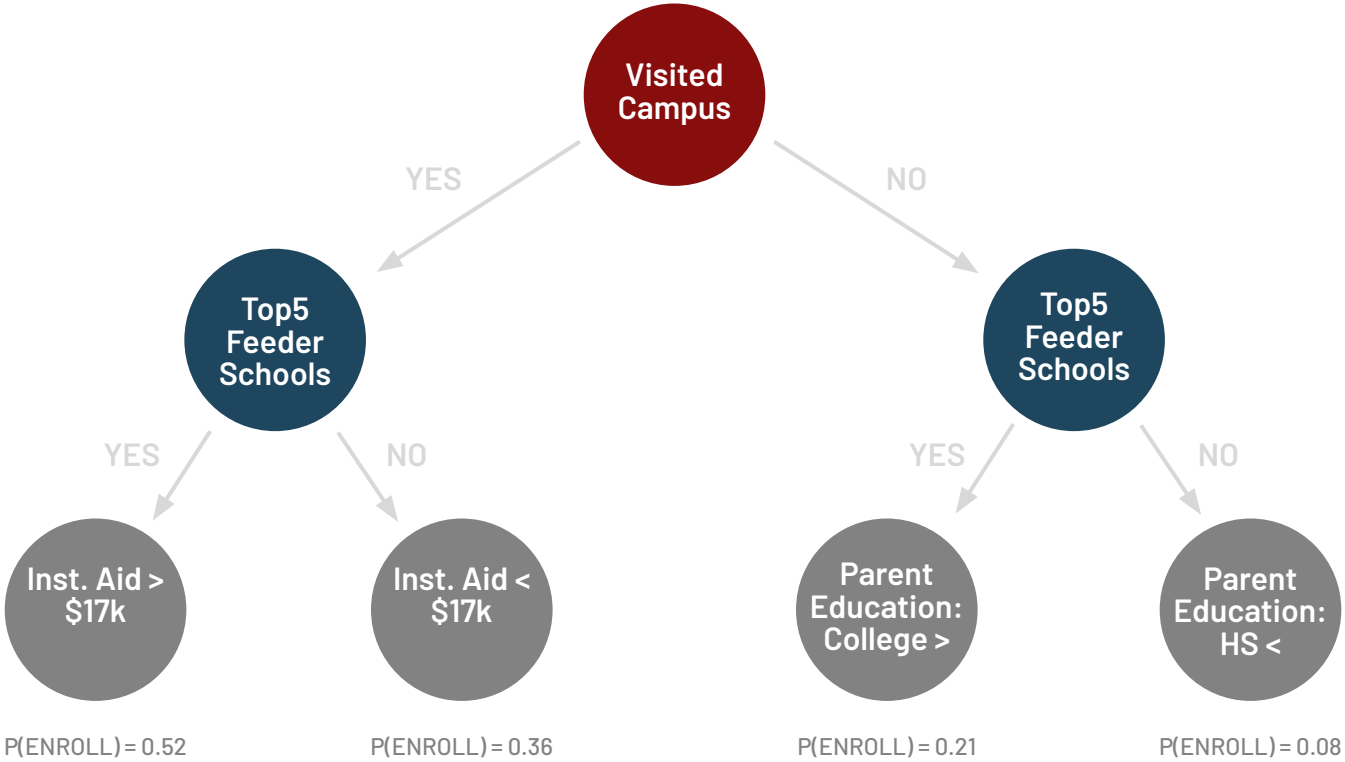
Realistic Expectations
- Incorporates future expectations not captured in historic outcomes

2

Types of Predictive Models



Types Of Models: Decision-Tree Approach

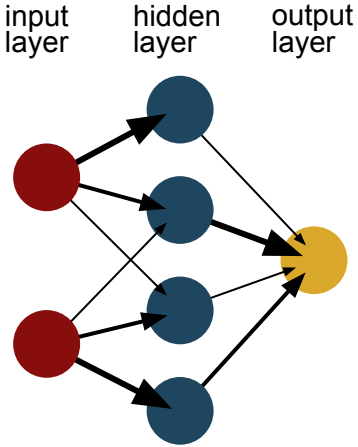


Types Of Models: Neural Networks

Long History (1940s & 50s) but have become prominent with availability of large data sets.

- Neural networks are a class of machine learning models inspired by the human brain's structure and functioning.
- They consist of interconnected nodes called neurons organized in layers: an input layer, one or more hidden layers, and an output layer.
- Neural networks are capable of learning complex patterns and relationships in data, making them powerful for various tasks like classification, regression, and pattern recognition.
- Deep neural networks have multiple hidden layers, allowing them to learn hierarchical representations and solve more complex problems.

A simple neural network



Types Of Models: Logistic Regression

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta \square x \square$$

Historically successful model for over 30 years

- Plus, well documented success of logistic regression in other industries

Practically, what does the model tell us?

- Predicts (0,1) the probability of a student becoming a net deposit
- Reveals relative influence of student-level attributes on that prediction

Why is logistic regression a good choice for yield modeling?

- Highly interpretable, allows for clear insights into student population that can drive actionable change
- Robustness (low variance) helpful for imbalanced and low N-size data
- Flexible and complex. We don't assume static yield for given subpopulation. Instead 10-15 different variables produce individualized prediction

Why not other models?

- We have a strong, proven hypothesis about yield relationships in a logistic regression setting. More exotic models sacrifice interpretability for potentially minimal accuracy gain

General Data Request To Perform The Work

- ~ 3 Years of Historical Data
- First Year and Transfer (Admit Status)
- Admission
  - CRM - academic, geodemographic, behavioral, student status
- Financial Aid
  - SIS - FAFSA data, actual awards (are records purged for non-enrollees?)
- Enrollment
  - SIS - Actual enrollment (validation and retention)

Feature Creation and Selection

Goal: Identify bundle of statistically significant variables which accurately predict enrollment

STEP 1: Comprehensive variable creation
- Recoding all available data

STEP 2: Data exploration and identification of trends
- Dashboard
- Calculate subpopulation yield and correlations amongst variables and with net deposit behavior

STEP 3: Variable selection
- Intentional removal of unstable (low N) and spurious variables
- Automatic via stepwise selection (maximize information gain)
- Manual intervention based on experience, including interaction terms

Variable selection blends mathematical results and experiential intuition
- Iterative, exploratorative, comparative

Financial Aid Strategy Design

A Typical Case

- The institution needs to have a financial aid awarding strategy in place by October for this admissions cycle.
- The admit pool for the current year is not yet known, but there are some assumptions about size and composition.
- T&Fs, room and board are likely to increase but is also unknown.
- The institution wants to meet all of their institutional goals for the incoming class in the most efficient manner as possible, in this case with the priority being as much NTR as possible.
- The institutional goals for the incoming class are:
    - 500 new first year students
    - Avg HS GPA of 3.5
    - 100 students in the College of Arts & Behavioral Sciences

- Design a financial aid awarding strategy that accomplishes these goals.

Financial Aid Strategy Design

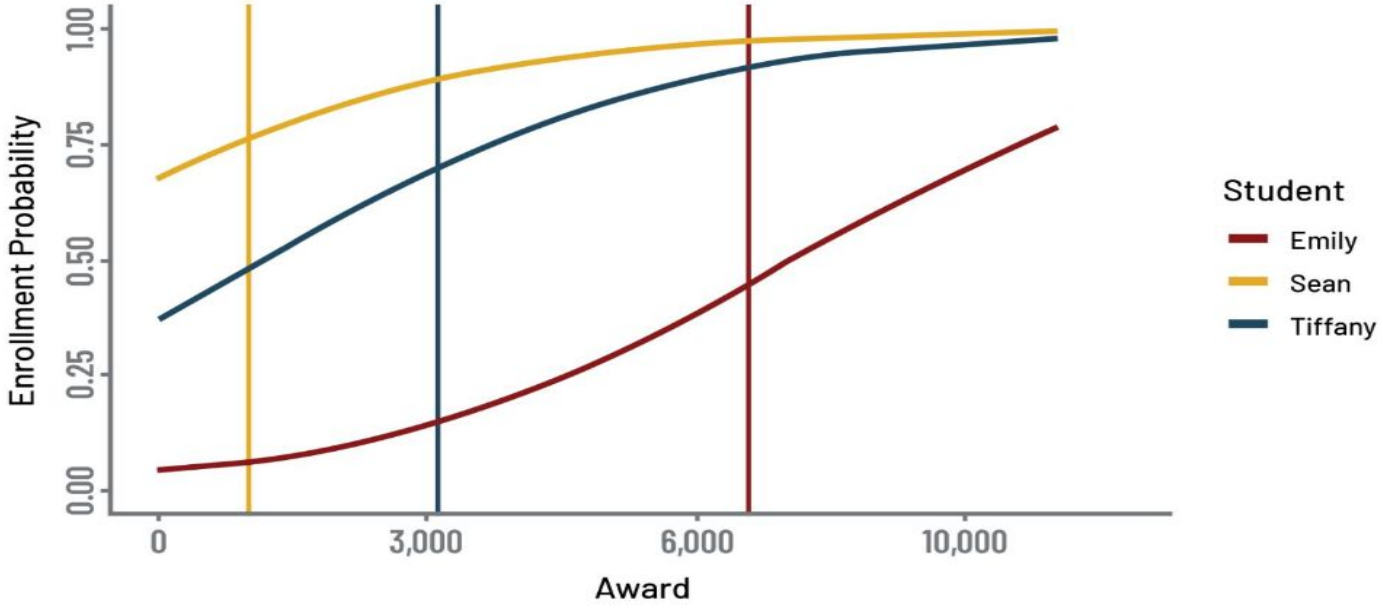
A couple different forms:

- Individualized awarding strategy
  - An equation is created from the model that can be used to award students on an individual basis based on significant factors (HSGPA, need, major, scholarship amount, etc.)

- Grid-based awarding strategy
  - Usually a grid where the x axis represents academic preparation ranges (HS GPA) and the y axis represents different ranges for demonstrated financial need

All of this strategy design can be run with different assumptions on the admit pool (total number, composition, etc.) depending on how it is trending.

Enrollment Probability as a Function of Aid



3

Exercise: (Hypothetical) Factors + Directionality

Model Outputs

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta \square x \square$$

Outputs include coefficients, significance levels, various goodness of fit metrics

Reveals the mathematics behind the model

Factor (x)	**Coefficient (β)**	**Significance**
Total Gift Aid	0.00002	0.000001 (***)
Academic Rating: 5	-0.82918	0.000016 (***)
Has Test Score	-0.54265	0.000745 (***)
Male Students	-0.24214	0.114301 (*)
School: Humanities	-0.41313	0.091135 (**)
First Generation	-0.65008	0.004338 (***)
Constant	0.40699	0.338017 (*)
Illustrative Example Using Fall 2022 Model		

Model Outputs: Average Marginal Effects

Factor	Directionality	Effect (holding other variables constant)
Total Gift Aid	+	2% per $10k gift aid
Academic Rating: 7	-	-4%
Has Test Score	-	-3%
Male Students	-	-3%
School: Humanities	-	-2%
First Generation	-	-6%
Campus Visit	+	+12%

CARNEGIE

4

Special Considerations

for Christian institutions and smaller colleges and universities.

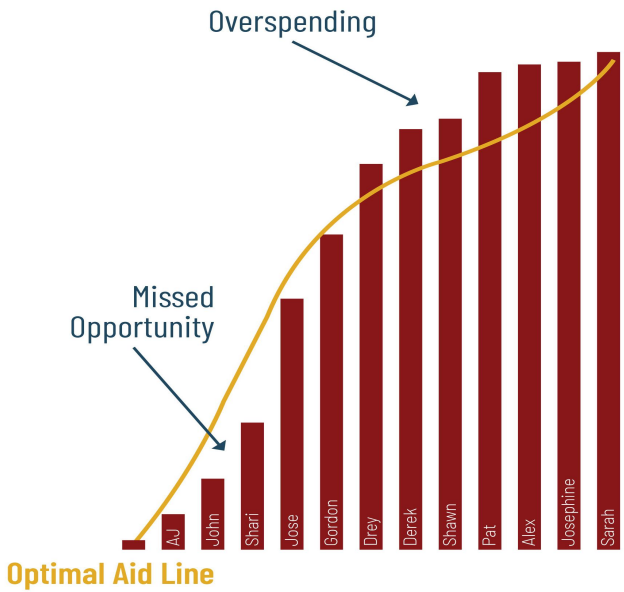
CARNEGIE

Aid Strategy

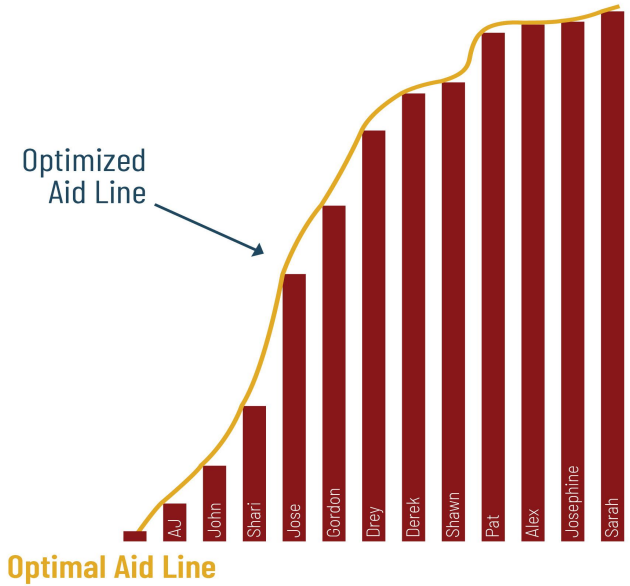
+ Commitments to affordability, prioritizing accessibility that may take precedence over certain goals
+ Alignment with the Mission/Values of the institution to find ways to differentiate awarding
+ Careful thought around stewardship of resources
+ Donor + Alumni + External Organization Engagement

Individualized Awarding

Traditional Approach



Individualized Aid Approach



CARNEGIE

Too many institutions make financial aid strategy decisions with an exclusively budget-driven focus.

Goal-Oriented

Strategies for assigning financial aid should be created with the intent of responsibly balancing;

+ Desired Enrollment
+ Net Revenue
+ Student Profile
+ Financial Aid Budget

This should come as a RESULT of the goals conversation, not the starting point

CARNEGIE

Additional Considerations:

How to plan for success beyond year one

Solution:

Retention Modeling & FAO

CARNEGIE

Retention Modeling

uncovers the distinctive factors impacting retention at your institution and builds a strategy that leverages that data for student success beyond year one. Student success leads to institutional success and vice versa.

Holistic Thinking/Considerations

Factors/variables that could be impacting student retention:

+ Pre-enrollment (academic preparation data, financial aid awards, need, net cost, geodemographic attributes)

+ Post-enrollment student level data (LMS data, term or year gpa, engagement data)

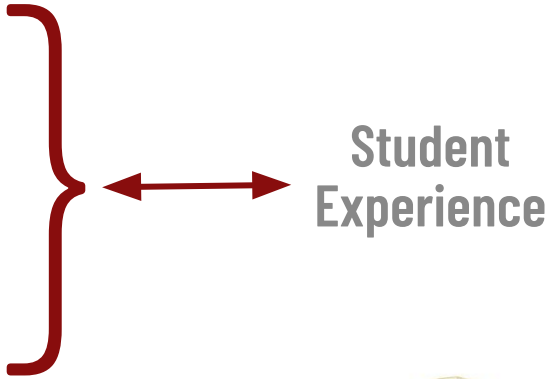
More Considerations

Important Considerations for Student Success Modeling:

+ Financial aid optimization modeling can influence/impact retention modeling
+ Modeling does not stop with enrollment, returning students play a large role in the success and stability of any institution.
+ Use modeling to inform the intervention (think about this in terms of resources, outreach, etc.)
+ Human behavior as it relates to retention/attrition can be complex

Challenges of Disentangling Causality

- Academic difficulty
- Financial stress
- Social-emotional challenges
- Physical & mental health



Explanatory Considerations

NOTE: "Transferring" is not a withdrawl reason, it is the action spurred by a withdrawl reason!

Primary/Secondary Reasons	Academic struggle	Work conflict	Financial struggle	Personal health	Family issue	Social struggle	Campus climate	Major change
Academic struggle								
Work conflict								
Financial struggle								
Personal health								
Family issues								
Social struggle								
Campus climate								
Major change								

- Academic struggle exacerbated by non-academic stressors
- College obligations interfering with immediate need for income
- Lack of affinitity to university can alter perceived value, i.e. "tuition is not *worth* it"
- Strong bonds to family and HS friend networks overide or inhibit relationship building at university

CARNEGIE

4

Questions + Answers

CARNEGIE

1:00 PM EDT

Coming Up Next …

Preparing for the New FAFSA