

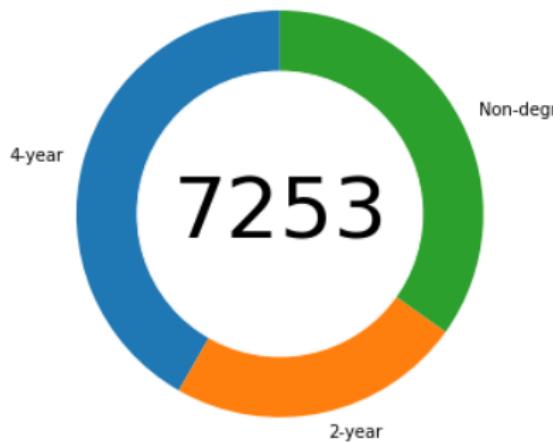
# Predicting school closings from the DOE College Scorecard

Kirsten Regier  
October 5, 2020



# Number of US Postsecondary Institutions

2012-2013



2016-2017

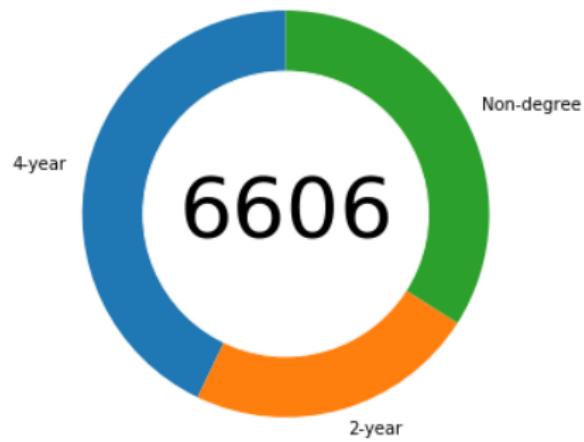


Figure: Distribution of US colleges by degree type for AY 2012-13 (left) and 2016-2017 (right).

Source: U.S. Department of Education, National Center for Education Statistics. (2019). Digest of Education Statistics, 2018 (NCES 2020-009), Chapter 2., <https://nces.ed.gov/fastfacts/display.asp?id=84>

# Closing of US postsecondary institutions



postsecondary  
institutions closed  
between 2012-2013 and  
2016-2017.

**ITT Technical Institute** closed 130+ campuses in 2016, affecting more than 40,000 students and 8,000 faculty

<https://www.insidehighered.com/news/2016/09/07/itt-tech-shuts-down-all-campuses>

**Cincinnati Christian University** stopped "offering accredited degree programs" in Fall 2019 <https://ccuniversity.edu/a-letter-to-our-students/>

# Can we predict which schools are likely to close?

## Why does it matter?

- ▶ Trends predict fewer college-going high-school graduates
- ▶ Rising tuition and student loan debt
- ▶ Financial strain from COVID-19

## Potential stakeholders:

- ▶ Prospective students (and parents)
- ▶ Faculty, staff, and support services
- ▶ College administrators, boards of trustees, and advisors
- ▶ Accrediting agencies, funding agencies and donors
- ▶ Employers looking to hire graduates

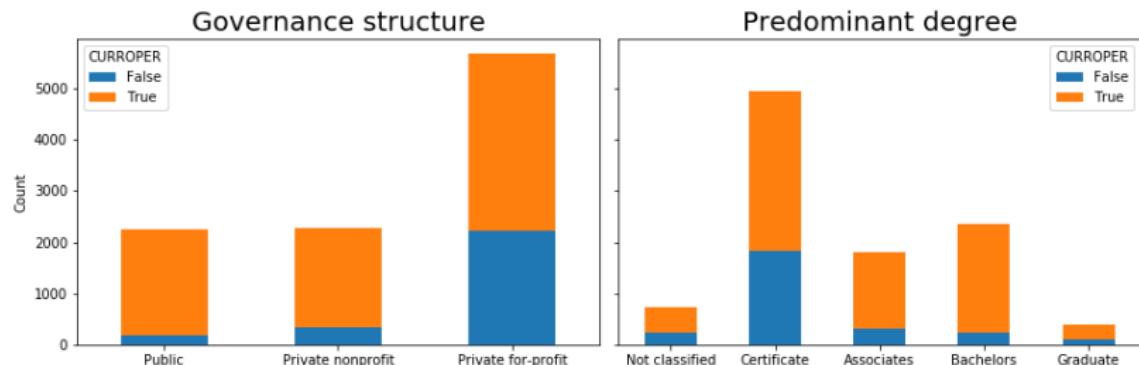
# The Data

## US Department of Education College Scorecard

(<https://collegescorecard.ed.gov/data/>)

- ▶ **Target variable** - currently operating (CURROPER)
- ▶ 10,200 observations x 32 features
  - ▶ 7441 schools operating in 2013
  - ▶ 2759 closed schools from 2010-2013
- ▶ School identifiers - OPEID, name, location
- ▶ Institutional features - governance structure, predominant degree, highest degree
- ▶ Student features - enrollment, completion rates, median debt
- ▶ Financial info - cost, tuition revenue, institutional expenditure, average faculty monthly salary

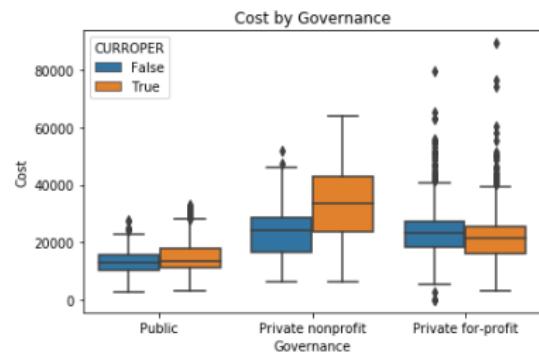
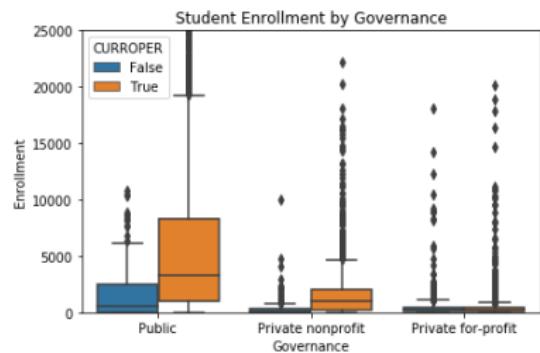
# Features of analyzed schools



Distribution of schools in the data by governance (left), predominant degree (right), and whether they are currently operating (CURROPER).

# Features of analyzed schools

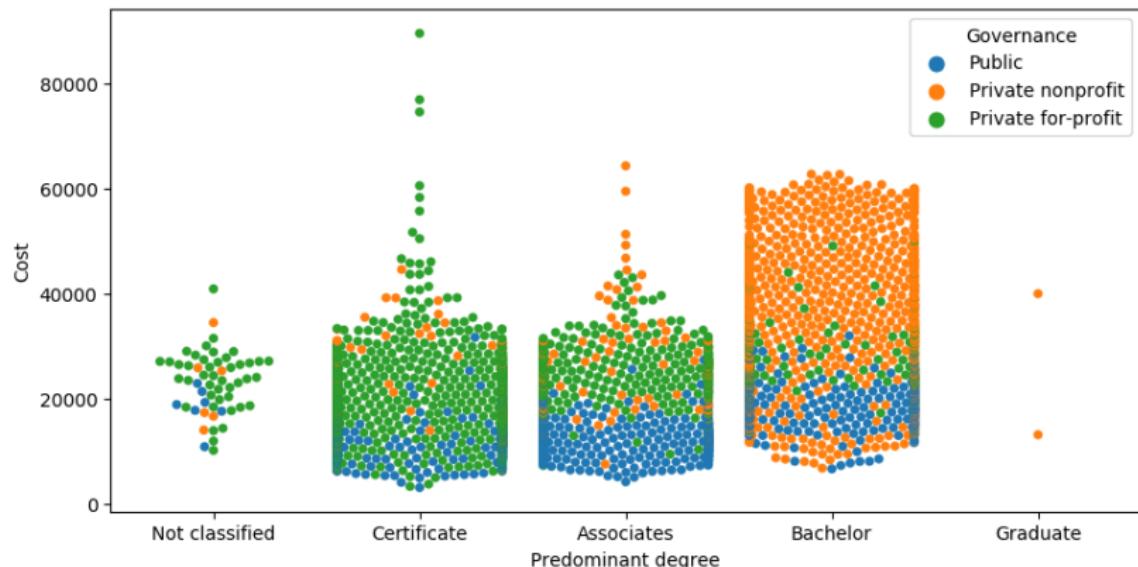
Public schools are larger and less expensive than private schools. Private non-profit schools are slightly larger and more expensive than private for-profit schools, though private for-profit schools show more variation in enrollment and cost than private nonprofit schools.



Enrollment (left) and cost (right) of schools by governance structure and whether they are currently operating (CURROPER).

## Features of analyzed schools

Private non-profit schools offer predominantly bachelor's degrees, and are more expensive. Public schools are less expensive than public for-profit schools for all degree types.



Cost by degree type and governance structure.

## Recommendations

Use model to predict status of member institutions.

- ▶ Run 180+ member schools through model.
- ▶ Pay attention to schools that are predicted to be 'closed'.

Use model to predict probabilities of school closings.

- ▶ Model generates probability that each school is open.
- ▶ Current model uses 50% as threshold.
- ▶ Can adjust threshold for classification, or examine the probabilities.

# Data Processing

**Impute missing values using group median**

- ▶ Institution name and federal ID
- ▶ Governance structure
- ▶ Predominant degree

**Filter data by year**

- ▶ Closed schools from 2010-2013
- ▶ Currently operating schools from 2013

Double the number of minority class observations via **resampling** with replacement.

# Modeling

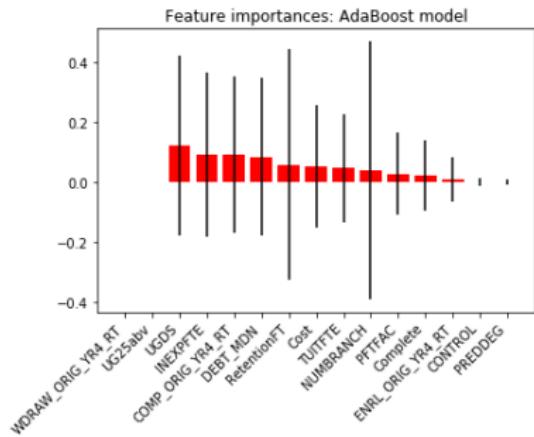
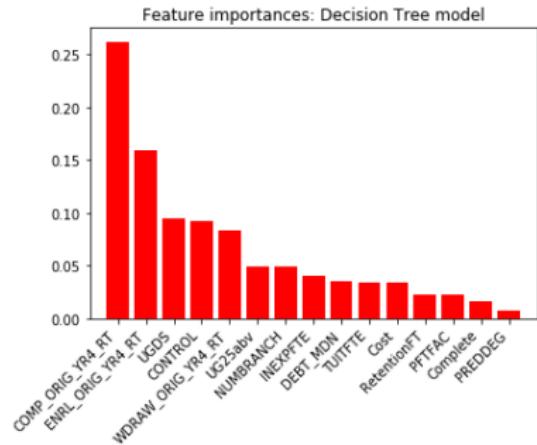
1. Decision Tree model
2. AdaBoost model

## AdaBoost Model Performance Metrics

Currently operating		Predicted		Recall
Actual	No	No	Yes	
		1227	30	0.98
	Yes	69	2163	.97
Precision		.95	.99	
Accuracy				.97

Table: Confusion matrix and evaluation metrics for the final AdaBoost model with duplicated/resampled observations removed.

# Model results



Feature importance levels based on the Decision Tree model (left) and the AdaBoost model (right).

## Model results

	Decision Tree		AdaBoost	
	Factor	Coefficient	Factor	Coefficient
1	% students completed after 4 years at original institution	0.262313	Undergrad enrollment	0.121028
2	% students enrolled after 4 years at original institution	0.158929	Instructional expenditure	0.090679
3	Undergrad enrollment	0.094170	% students completed after 4 years at original institution	0.089959
4	Governance structure	0.092008	Median student debt	0.084749
5	% students withdrawn after 4 years from original institution	0.082807	Retention of full-time students	0.058268

Table: Top 5 predictive features and coefficients

## Weekend price by Resort size (acres of skiable terrain)

All US Resorts

Montana Resorts

Current **weekend** ticket price by acres of skiable terrain for all US ski resorts (left) and for ski resorts in Montana (right). The larger black dot represents Big Mountain Resort.

## Weekend price by Resort size (number of runs)

All US Resorts

Montana Resorts

Current **weekend** ticket price by number of runs for all US ski resorts (left) and for ski resorts in Montana (right). The larger black dot represents Big Mountain Resort.

# Conclusions

- ▶ In Montana - current Big Mountain price is highest; second largest resort with many runs
- ▶ Nationally - current weekend prices are below model predictions.
- ▶ Raising weekend ticket prices to \$86-\$87 would bring in sufficient additional revenue to cover the operating cost of the new chairlift.