

Problem Statement Worksheet (Hypothesis Formation)

What are the primary factors that have contributed to the closing of 4-year undergraduate colleges and universities in the US in the last decade?

1 Context

Roughly 400 colleges closed between 2012 and 2017. Given the financial strain facing colleges in the face of the coronavirus, it could be helpful to be able to predict which schools may be in danger of closing, and to be able to take steps to mitigate the factors that contribute to schools closing.

2 Criteria for success

Identify reasonable/explainable factors?
Model predicts closure on train/test data with 85% accuracy?
(Seems low, but I don't know what to expect)

3 Scope of solution space

The plan is to focus on 4-year, primarily undergraduate institutions. I anticipate that focus will be on private schools, which lack the level of federal and state government funding provided for public schools.

4 Constraints within solution space

There have been decreasing numbers of high school graduates, which means fewer prospective students for colleges to recruit. It may not be possible (practically and financially) for colleges to increase enrollment.

There may be economic factors that affect students' ability and willingness to attend college, particularly in an uncertain global environment.

5 Stakeholders to provide key insight

College administrators - presidents, provosts, deans, finance and financial aid offices, etc.

The Higher Learning Commission, the US Department of Education, and individual colleges Institutional Data Offices are important resources for data collection and questions.

6 Key data sources

US Dept of Ed College Scorecard, particularly the CURROPER column, for tracking schools that have closed.
Statista 2019 dataset "Number of higher education institutions in the United States from 1980 to 2017"