

College Critics

Victoria Velasquez, Hartland Brown, Rami
Alqumaibit

Description

Our project is using a DataSet of “College Scorecards” that were filled out by thousands of universities from 1996-2013. The scorecards contain information about graduation rates, acceptance rates, average debt, average income of graduates based on degree types, and even (morbidly) percentage of death of graduates in different income tiers. Our hope is to find the answers to questions like

- What kind of correlations exist between level of debt and level of income?
- Is there a correlation between students who graduate quickly and income level?
- Do two year universities provide a comparable difference between debt and income as four year universities?

Prior Work

- No work has been done on cleaning up the dataset.
- A lot of work was put into collection dozens of different types of information for each university.

Datasets

- Dataset: US Department of Education: College Scorecard
 - 18 Datasets from 1996-2013 (each year is a separate dataset)
- Can be found here:
<https://www.kaggle.com/kaggle/college-scorecard/data>
- Provided by the US Department of Education on Kaggle.com
- Dataset is downloaded on Hartland's machine.

Proposed Work

- Cleanup NULL values for each university, using previous years as a guide.
- Remove Universities that had too many empty scorecards.
- Combine datasets into a single usable organized by year.
- Look through all columns to determine if there is any other interesting questions to be asked.

Tools

- Tools we will use:
 - Python
 - Pandas
 - Numpy

Evaluation

- Effectively check for Confidence and Support in all relationships
- Use online resources and University Websites to crosscheck results
- Use department of Ed. website to find additional scorecards and information.

<https://collegescorecard.ed.gov/>