

Analyzing Performance of Higher Education Institutions for Predicting Student Success

Rohan Doshi, Samhita Karnati,
Kevin Liu, and Avinash Nayak

Guiding Questions

1. Which observed variables are most predictive of student outcomes?
2. Relative to schools with similarly constructed student populations, which schools best prepare their students for eventual success?
3. How can we identify common latent structures across different categories of institutions that are indicative of a school's ability to prepare its student body for success?

Data Preprocessing

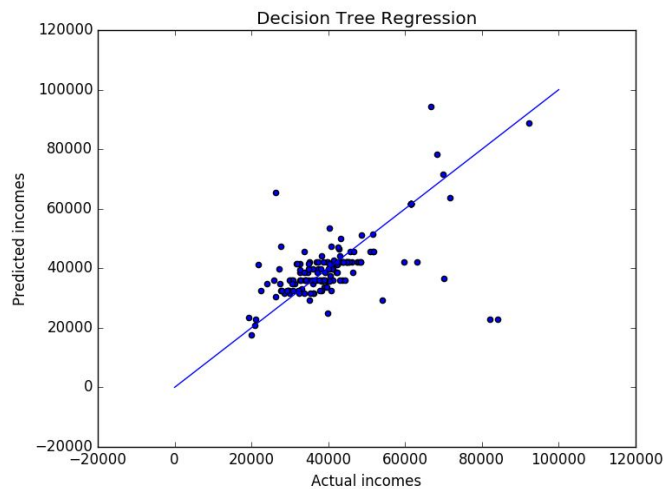
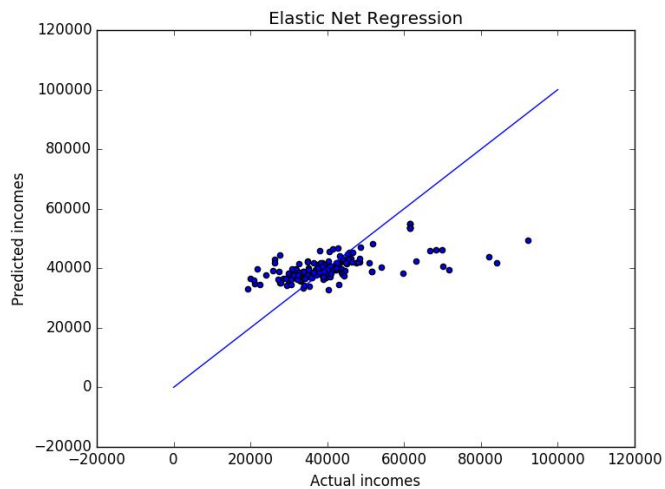
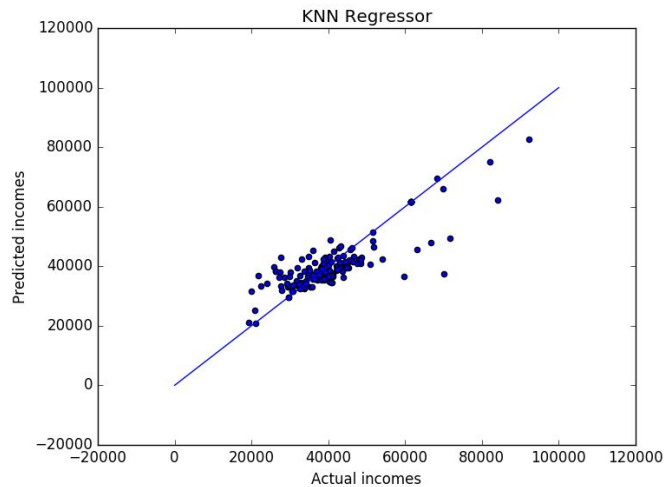
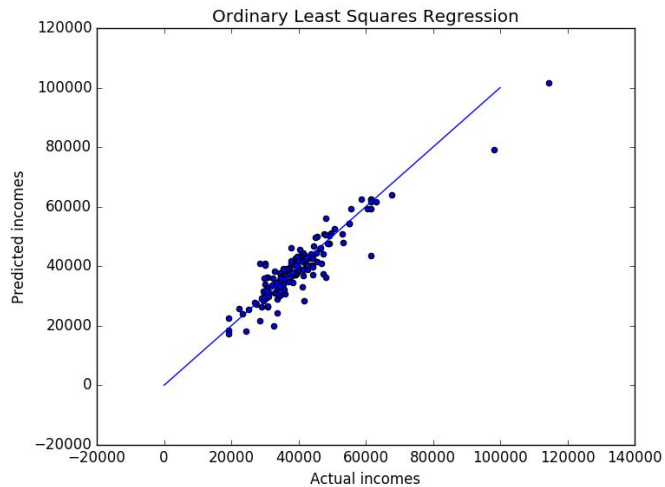
1. Data merging
2. Removing categorical and co-linear features
3. Removing schools that would not be useful
4. Imputing data using k-nearest neighbors
5. Normalizing data

Feature Importance

Variable	Meaning	Feature Importance Coefficient
UGDS_AIAN	Prop. of UGs with American Indian Heritage	$3.66 \cdot 10^9$
UGDS_ASIAN	Prop. of UGs with Asian Heritage	$3.60 \cdot 10^9$
UGDS_2MOR	Prop. of UGs with Mixed Heritage	$3.52 \cdot 10^9$
UGDS_BLACK	Prop. of UGs with Black Heritage	$3.49 \cdot 10^9$
UGDS_UNKN	Prop. of UGs with Unknown Heritage	$3.47 \cdot 10^9$
UGDS_NHPI	Prop. of UGs with Pacific Islander Heritage	$3.47 \cdot 10^9$
UGDS_HISP	Prop. of UGs with Hispanic Heritage	$3.42 \cdot 10^9$
UGDS_NRA	Prop. of UGs who are NonResident Aliens	$3.41 \cdot 10^9$
UGDS_WHITE	Prop. of UGs with Caucasian Heritage	$3.39 \cdot 10^9$
PCIP29	% of Degrees in Military Tech/Applied Sciences	$1.22 \cdot 10^9$
NOPELL_ENRL_ORIG_YR6_RT	% without Pell Still Enrolled After 6 Years	$8.08 \cdot 10^8$
PCIP41	% of Degrees in Science Technologies	$5.37 \cdot 10^8$

Supervised Learning

Regression Type	r^2	RMSE
Ordinary Least Squares	0.76	4005.25
K Nearest Neighbors	0.42	6251.03
Ridge	0.61	5087.40
Lasso	0.67	4716.40
Elastic Net	0.37	6514.98
Decision Tree	0.04	8042.44
Random Forest	0.35	6628.65



Ordinary Least Squares

Best schools

Rank	Institution Name	Residual
1	Westwood College-Los Angeles	12259.81
2	Fairfield University	11905.85
3	Savannah State University	11050.12
4	Prescott College	8524.98
5	Auburn University	6719.69
6	Pacific Oaks College	6449.84
7	Rose-Hulman Institute of Technology	6331.67
8	Silver Lake College of the Holy Family	5702.42
9	Daemen College	4529.33
10	University of Colorado Boulder	4446.70

Worst schools

Rank	Institution Name	Residual
1	Rutgers University-Newark	-12387.28
2	Pennsylvania State University-Penn State Lehigh	-11728.43
3	Pennsylvania State University-Penn State York	-11050.88
4	Rutgers University-Camden	-9411.65
5	Swedish Institute a College of Health Sciences	-8643.17
6	Florida National University-Main Campus	-8355.56
7	Baker College of Flint	-8024.05
8	SUNY at Purchase College	-7550.39
9	South University-The Art Institute of Dallas	-7488.18
10	Oklahoma City University	-7364.77

K Nearest Neighbors

Best schools

Rank	Institution Name	Residual
1	Rose-Hulman Institute of Technology	24200
2	University of Oklahoma-Health Sciences Center	21940
3	New York School of Interior Design	15020
4	Bryan College of Health Sciences	14800
5	Butler University	11240
6	CUNY Bernard M Baruch College	11240
7	Michigan Technological University	10020
8	Pace University-New York	8340
9	The University of Texas at Dallas	8220
10	Daemen College	7140

Worst schools

Rank	Institution Name	Residual
1	Naropa University	-20060
2	Florida National University-Main Campus	-16760
3	California College of the Arts	-16720
4	Metropolitan State University of Denver	-15960
5	Eastern Connecticut State University	-14940
6	Hendrix College	-14400
7	Prescott College	-13000
8	Thomas University	-10320
9	Westmont College	-9480
10	New York College of Health Professions	-9300

Elastic Net

Best schools

Rank	Institution Name	Residual
1	University of Oklahoma-Health Sciences Center	20867.00
2	Rose-Hulman Institute of Technology	19831.42
3	Bryan College of Health Sciences	16874.28
4	New York School of Interior Design	16860.56
5	Washington and Lee University	13324.51
6	Washington Adventist University	10727.73
7	Michigan Technological University	10599.53
8	Butler University	10319.31
9	Fairfield University	8142.88
10	CUNY Bernard M Baruch College	7997.13

Worst schools

Rank	Institution Name	Residual
1	Naropa University	-19025.82
2	Caribbean University-Ponce	-15868.10
3	Caribbean University-Vega Baja	-15846.24
4	Universidad Metropolitana	-15624.29
5	Hendrix College	-14336.91
6	American University of Puerto Rico	-11460.36
7	Baker College of Flint	-9918.07
8	Grace College and Theological Seminary	-9868.67
9	University of Alaska Southeast	-9678.56
10	Lancaster Bible College	-9667.06

Decision Tree

Best schools

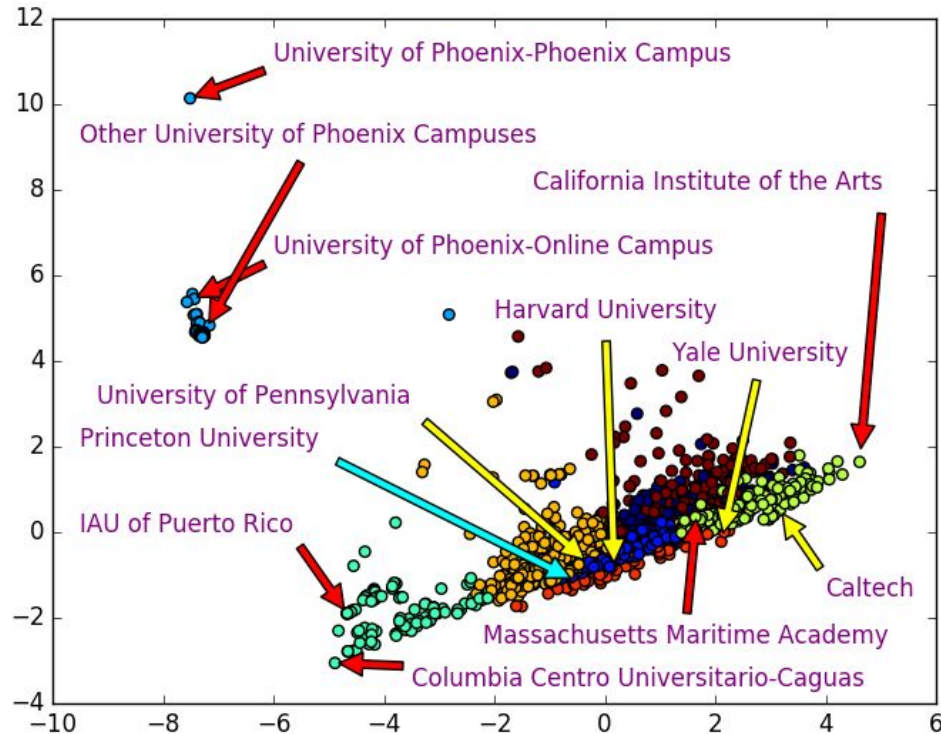
Rank	Institution Name	Residual
1	Rose-Hulman Institute of Technology	31900.00
2	Butler University	21200.00
3	Bryan College of Health Sciences	18541.90
4	University of Oklahoma-Health Sciences Center	17803.13
5	Washington Adventist University	10862.69
6	Pace University-New York	10403.13
7	CUNY Bernard M Baruch College	9521.54
8	The University of Texas at Dallas	8821.54
9	Silver Lake College of the Holy Family	8820.00
10	Stony Brook University	8721.54

Worst schools

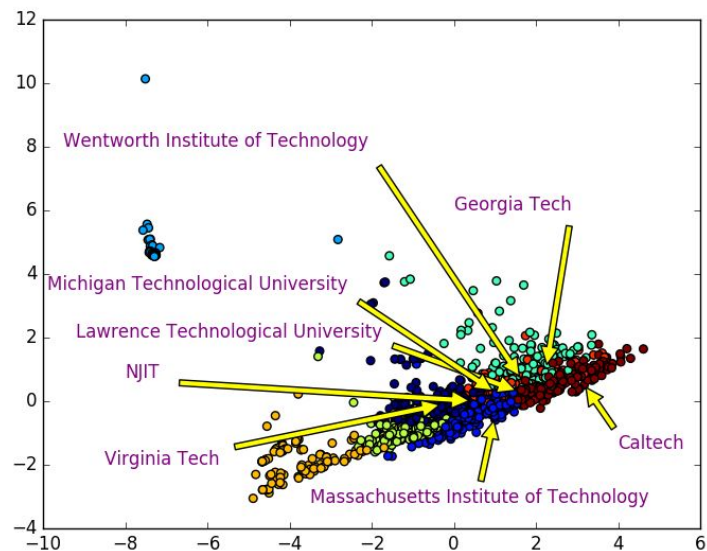
Rank	Institution Name	Residual
1	Eastern Connecticut State University	-30900.00
2	Southern University and A & M College	-26800.00
3	Auburn University	-26000.00
4	Savannah State University	-23600.00
5	Touro College	-20116.67
6	Hendrix College	-18692.31
7	California College of the Arts	-15392.31
8	Northwestern University	-15100.00
9	Valley Forge Christian College	-13338.78
10	Naropa University	-12325.00

Unsupervised Learning

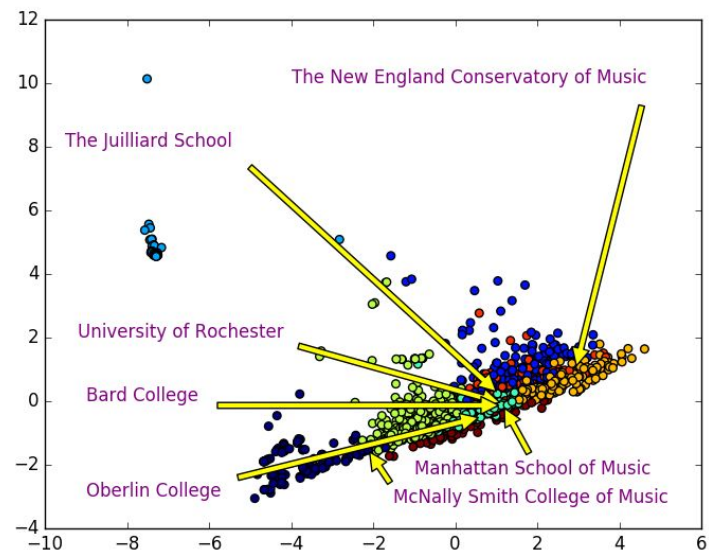
Universities Clustered on Financial Success



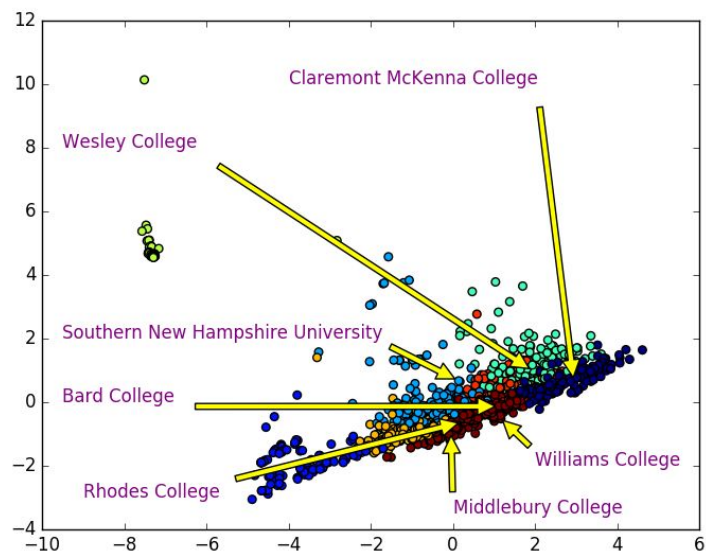
Technical Universities



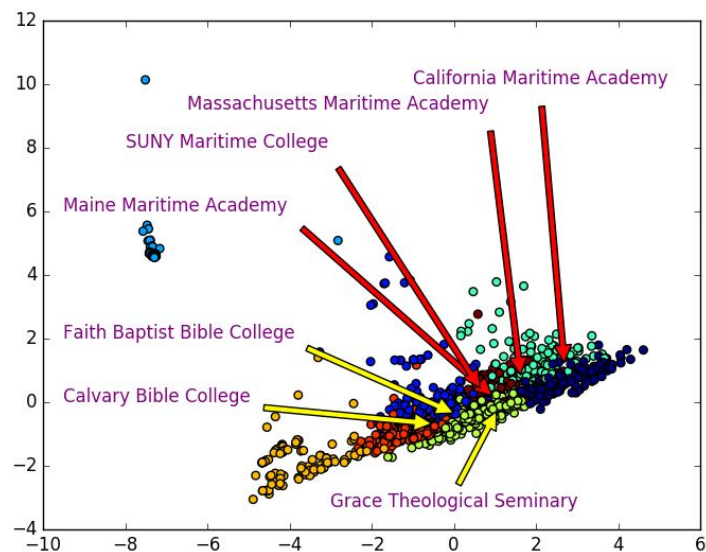
Music Schools



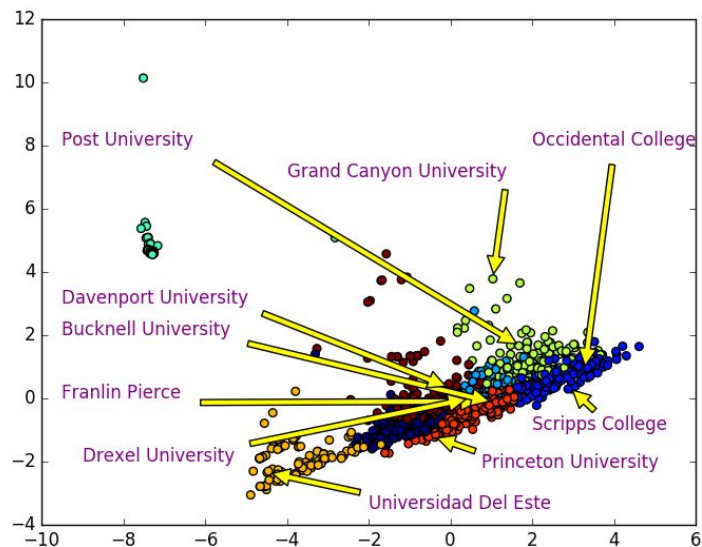
Liberal Arts Colleges



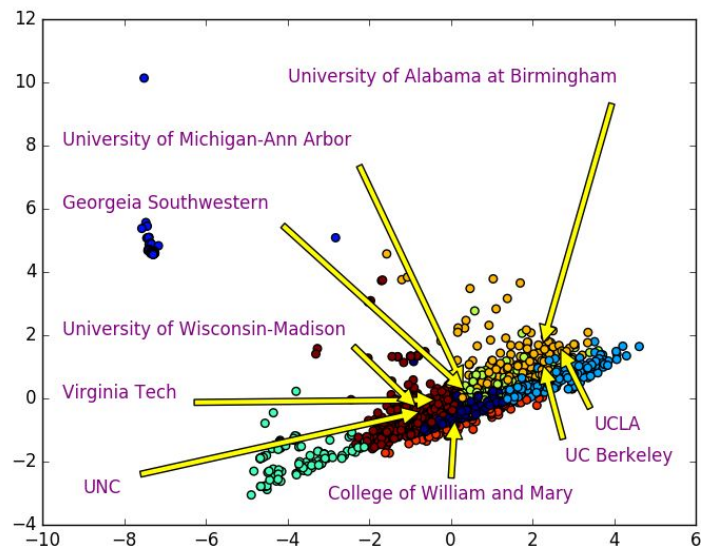
Trade Schools



Private Universities



Public Universities



Supervised vs. Unsupervised Learning

