



# Factors Most Important to University Ranking

Team 2cool4school

Samantha Bernier, Maia Goel,

Haojin Li, Mikey Schwartz

# Research Background

**Question:** Which characteristics of a university are most associated with rankings on the Niche College Ranking list? Of these characteristics, what is relationship between high and low rank?

# Data



2023 Best Colleges in America

- College name, rank



U.S. DEPARTMENT OF EDUCATION  
College Scorecard

- **Categorical Variables:** Region, Public vs. Private, etc.
- **Numerical Variables:** Avg. SAT, Avg. Family Income, etc.

# Approach #1: Numerical Variable Analysis

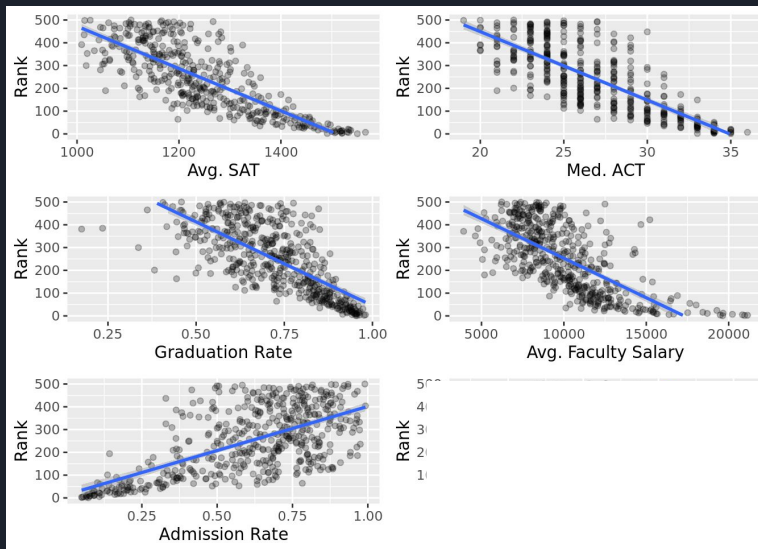
## Methods

1. Create separate linear models between all numerical explanatory variables and Niche rank.
2. Order  $R^2$  of models to determine strongest correlations.
3. Compare scaled slope to determine greatest influence on higher rank.

## Results

- Trend: relationships with higher  $R^2$  also have higher scaled slopes. Most associated variables are also most influential.

| variable            | R-Squared | Scaled Slope | Un-scaled Slope |
|---------------------|-----------|--------------|-----------------|
| Avg. SAT            | 0.6291468 | -115.00      | -0.9277         |
| Med. ACT            | 0.6062324 | -112.90      | -30.0400        |
| Graduation Rate     | 0.5137627 | -103.50      | -735.2000       |
| Avg. Faculty Salary | 0.4692148 | -98.89       | -0.0347         |
| Admission Rate      | 0.4048157 | 91.74        | 387.0200        |



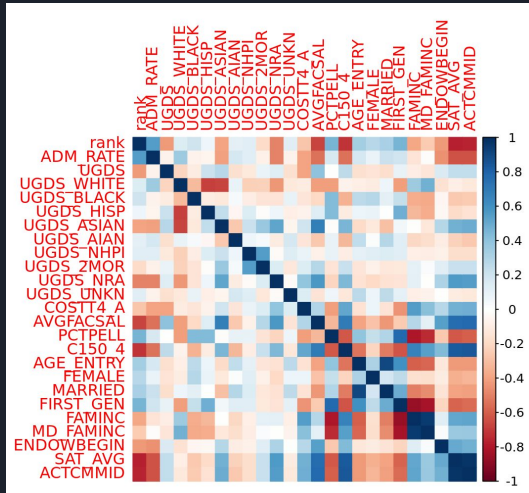
# Approach #2: Stepwise Regression Model

Final Model:

```
rank ~ REGION + CONTROL + CCBASIC + ACCREDITAGENCY + ADM_RATE +  
      UGDS + UGDS_ASIAN + UGDS_2MOR + UGDS_UNKN + COSTT4_A + AVGFACSA  
      AGE_ENTRY + SAT_AVG + FIRST_GEN + UGDS_HISP
```

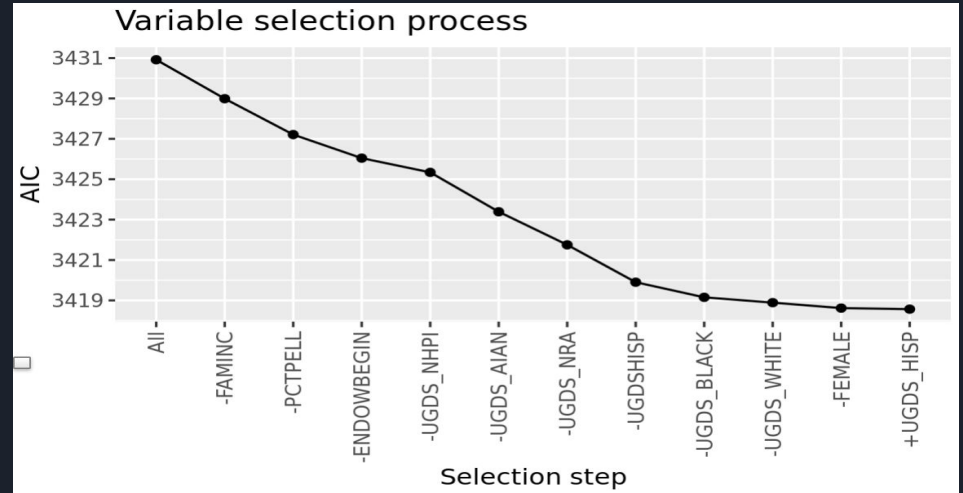
## Manual Selection

Correlation matrix



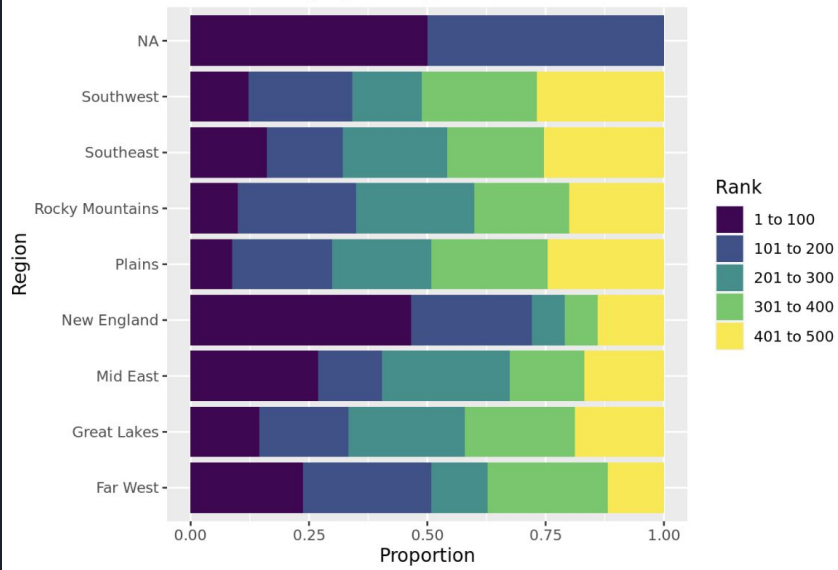
## Automatic Selection

Forward + backward selection by AIC

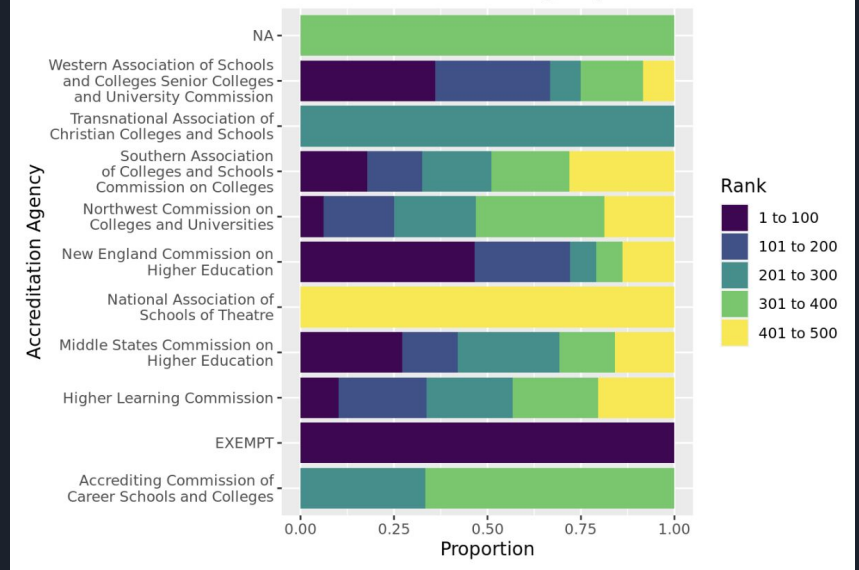


# Categorical Variable Analysis

Rank vs Geographic Region of the United States

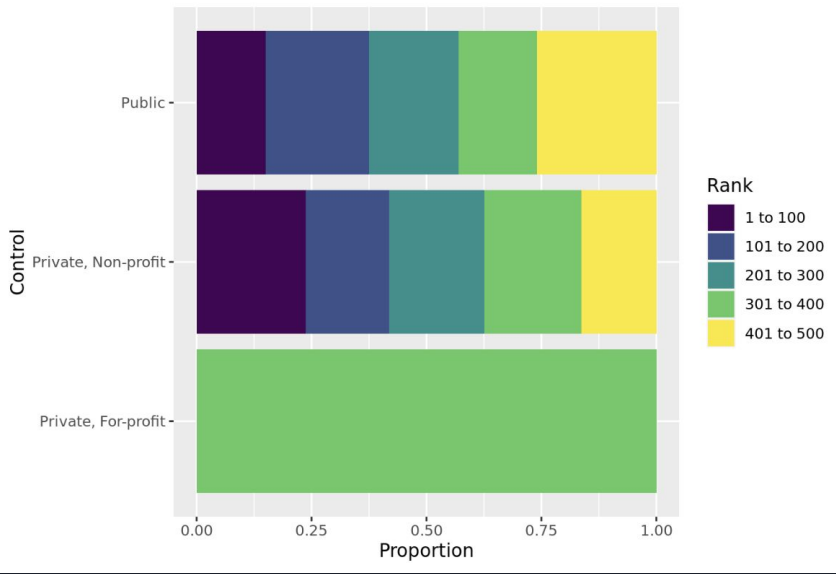


Rank vs Accreditation Agency

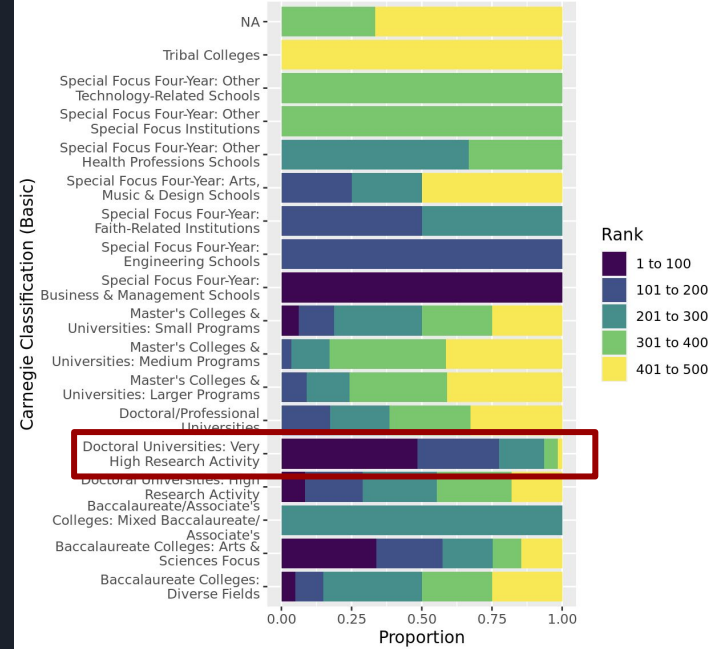


# Categorical Variable Analysis

Rank vs Control



Rank vs Carnegie Classification





# Conclusion and future works

- Most Important: selectivity, resources, and outcomes
  - Five most associated and influential numerical variables: Avg. SAT - Avg. ACT - Graduation Rate - Average Faculty Salary - Admission Rate
- Limitations:
  - no robust calculation of categorical variables individually
  - assumed that all variables had linear relationship so we could create stepwise regression
  - these are all things that could be fixed by learning new methods
- Look at more ranking systems, look at more schools