



## College Search Data

December 3, 2018

### Overview

The College Search App helps Rhode2College Scholars explore schools that they may want to attend based on test score compatibility and school characteristics like region and area of study. This document describes the data underlying the search results found in the app.

### Data Sources

Data are aggregated from public sources and primarily come from the U.S. Department of Education's College Scorecard, which includes school location and size, SAT™ score ranges for admitted students, admission and completion rates, average net annual cost for full-time, first-time undergraduates who receive Title IV aid, average earnings 10 years after enrollment, and available degree programs.<sup>1</sup> Note that many of these data fields are available only for Title IV recipients or students who receive federal grants and loans.<sup>2</sup> For schools that do not provide SAT™ score ranges in the College Scorecard, we look up SAT™ score ranges from the College Board's BigFuture website.<sup>3</sup> We map the 38 unique field of study codes used by the College Scorecard into nine areas of study from the Integrated Postsecondary Education Data System (IPEDS) to provide a succinct list of areas of study for Scholars to use in their search.<sup>4</sup>

### Filtering

The College Scorecard provides information on 7,175 historic and/or currently operating schools. We retain 6,715 schools that are currently in operation. We drop “distance-only” schools that do not require students to be physically present at a school and typically involve online education. Of the 6,292 remaining schools, 3,027 are predominantly associate degree or undergraduate degree granting institutions with more than 100 undergraduate students. This is the set of schools that we use in the app.<sup>5</sup>

### Additional Yearly Income Calculation

The additional yearly income calculation is intended to provide Scholars with an idea of how much they might earn after attending a given school and how this compares to the net price of the degree. Earnings are important because many low-income students take out student loans to fund their degree, and thus should consider whether attending a given school will generate

---

<sup>1</sup> <https://collegescorecard.ed.gov/data/>

<sup>2</sup> <https://collegescorecard.ed.gov/assets/FullDataDocumentation.pdf>

<sup>3</sup> <https://bigfuture.collegeboard.org/>

<sup>4</sup> <https://nces.ed.gov/ipeds/cipcode/browse.aspx?y=55>

<sup>5</sup> We exclude schools that are predominantly certificate-granting institutions and schools that are predominantly graduate schools. <https://collegescorecard.ed.gov/assets/FullDataDocumentation.pdf> (see Page 21).



enough additional income in the future to comfortably pay off their debt. Moreover, there is a large range of earnings outcomes across schools for admitted students with similar SAT™ scores. Having this information will help students make informed decisions.

We compute the *Additional Yearly Income* for students at school  $S$  using the following equation:

$$\text{Additional Yearly Income}_S = \text{College Earnings}_S - \text{High School Earnings}$$

where *College Earnings* is measured as average earnings 10 years after enrolling at the school, and *High School Earnings* is measured as average earnings 10 years after high school graduation among Rhode Island students who did not attend college. Note that *Additional Yearly Income* is not cumulative, rather it represents additional income observed in the tenth year after the student's enrollment at the school. *College Earnings* are available from the College Scorecard and come from de-identified tax records for federally-aided students aggregated by institution. *High School Earnings* come from de-identified National Student Clearinghouse<sup>6</sup> records for Rhode Island high school graduates who do not have a subsequent college enrollment, joined to de-identified wage records from the RI Department of Labor and Training.

## Imputed Earnings

Some schools in the College Scorecard data are missing average earnings 10 years after enrollment, but have values for average earnings 6 or 8 years after enrollment.<sup>7</sup> We impute missing earnings 10 years after enrollment using the predicted earnings from an OLS regression of the form:

$$\text{Earnings10yr}_S = \beta_0 + \beta_1 \text{Earnings8yr}_S + \beta_2 \text{Earnings6yr}_S + \beta_3 X_S + \varepsilon_S$$

where  $S$  is a predominantly associate or degree granting undergraduate school and  $X_S$  is a vector of school characteristics.<sup>8</sup> *Earnings8yr<sub>S</sub>* includes both non-missing earnings and imputed earnings that we calculate from an OLS regression of *Earnings8yr* on *Earnings6yr* analogous to the above equation.

---

<sup>6</sup> <https://studentclearinghouse.org/>

<sup>7</sup> <https://collegescorecard.ed.gov/data/>

<sup>8</sup> School characteristics are degree type, public/private, region, size, and admission rate. We include interactions between region and admission rate and between school characteristics and average earnings (at both 6 and 8 years).



## Annual Loan Payment

We use the average net annual cost of attending each school from the College Scorecard to calculate a student's expected annual loan payment 10 years after enrolling at the school. For students at school  $S$ , the average monthly loan payment 10 years after enrolling is calculated using the following expression:

$$\frac{rP_S(1+r)^N}{(1+r)^N - 1}$$

where  $P_S$  is the principal (the average net annual cost times the number of years for the degree, either 2 or 4) for school  $S$ ,  $r$  is the monthly interest rate (which we assume to be one twelfth of the current annual interest rate of 5.05 percent), and  $N$  is the term of the loan in months (which we assume to be 120).<sup>9</sup> The annual loan payment is this average monthly loan payment multiplied by 12.

## PSAT™ Conversions

Rhode2College Scholars have initially taken only the PSAT™ and not the SAT™. The app provides the ability to search for colleges using a PSAT™ score to filter colleges that students are likely to get into. The app displays college search results with approximate PSAT™ ranges of admitted students for each college. However, only SAT™ score ranges, not PSAT™ score ranges, for admitted students are available from the College Scorecard and BigFuture. To predict which schools students will likely be admitted to based on their PSAT™ score, we convert PSAT™ scores to SAT™ scores using the average score improvement from 10<sup>th</sup> grade PSAT™ to 11<sup>th</sup> grade SAT™ tests reported by the College Board, and then use that converted SAT™ score to filter and display schools that students are likely to get into based on the school's historic SAT™ score range for admitted students.<sup>10</sup> To convert SAT™ ranges to approximate PSAT™ ranges in the displayed search results, we reverse this calculation.

---

<sup>9</sup> We observe first-year retention rates in the College Scorecard. For schools with non-missing retention rates, we calculate principal as the average of one year of cost and the total cost of the degree, weighted by the share of students who drop out after one year.

<sup>10</sup> <https://collegereadiness.collegeboard.org/pdf/student-level-sat-suite-growth-estimates.pdf> (see Table 9).