College Navigator: R Shiny App

Khanh Do, Joyce Gill, Matthew Billings-Chiu

2025-05-08

Introduction and Research Motivation

This project explores higher education data from U.S. colleges to uncover trends related to **admissions selectivity, student body diversity, and similar colleges**. The goal is to build an interactive Shiny app to help users — such as students and parents compare and identify colleges based on customizable filters like state, admission rates, SAT scores, demographics, and more. Our dataset is from the U.S. Department of Education College Scorecard's Most Recent Institution-Level Data File. It contains over 3,300 columns and 6,600 rows ofinstitutional data. Technical Documentation containing a glossary of all relevant information regarding the College Scorecard Institution-Level Data. Data Source - Most Recent Institution-Level Data

Feature 1

Users will be able to select an institution and view its racial composition through a Plotly pie chart, using variables like `UGDS_WHITE`, `UGDS_BLACK`, `UGDS_HISP`, `UGDS_ASIAN`, and more. The chart will display raw percentages when hovering over each slice, giving a quick and clear breakdown of the student body.

Feature 2

The app will feature sliders for SAT average and admission rate (in ranges), along with filters for state and institution `CONTROL` type (Public = 1, Private = 2 or 3), Locale (by rural & town, suburb & city) . Based on the user's selections, the app will return a list of colleges that meet the criteria in the form of a scatter plot with x-axis is the SAT average and y-axis is the Admission rate, where on hover the point, it shows the information, on click, user can click on the school website link. User will also have the option to tick whether they want to include criteria SAT, Admission rate. For Locale and IS_PUBLIC, they can also not tick any box or all boxes. The sliders for SAT and Admission rate would have a start and end pointer to show the range in which they want to filter for.

Feature 3

The app will include a plot that allows users to filter colleges based on median student debt, median earnings, undergraduate enrollment, admission rate, and ACT average (calculated from subject scores using data manipulation). After selecting a college of interest, the app will use clustering techniques to identify and visualize similar institutions, helping users explore comparable schools based on these key financial and academic attributes.

One Relevant Finding

We used our app to guide our college selection process based on specific criteria: SAT average range (1000-1200), admission rate (0.6-0.7), private institution, and rural or town location. Using Feature 2, we narrowed

our list to options such as Abilene Christian University, Maryville College, Concordia College at Moorhead, and Northwest Nazarene University. Next, we used Feature 1 to evaluate racial diversity, and identified Abilene Christian University as one of the more diverse schools. Finally, Feature 3 allowed us to find institutions similar to Abilene Christian University, further supporting our search and expanding our list of potential colleges.

Level of Difficulty

Our app consists of three distinct, fully interactive features, each offering insights into U.S. colleges from different perspectives: racial composition, admissions data, and clustering of similar institutions. Based on its scope, complexity, and design, we argue that our project meets the criteria for A-level diffculty for the following reasons:

1. The app integrates a large, real-world dataset with over 3,300 columns and 6,600 rows, requiring substantial filtering, reshaping, and preprocessing. We used a wide range of dplyr, tidyr, and conditional logic functions to make the dataset efficient and responsive. Across features, we also implemented multi-step reactive pipelines that accomodate multiple user-selected filters, PCA preparation, and K-means clustering.

2. Each tab of the app features a well-structured, informative, and aesthetically pleasing visualizations. We customized titles, tooltips, color schemes, and layout to ensure that the visualizations are not only accurate but also engaging and easy to interpret.

3. Our app prioritizes usability and functionality. Users can filter colleges based on detailed criteria, explore interactive plots, click to view institution-specific details, and visit official websites. They can also visualize racial demographics or identify similar colleges baed on selected characteristics. This goes well beyond basic dropdown menus, offering a meaningful, personalized exploration experience.

4. Our code is clean, modular, and reactive, making the app efficient and scalable.

5. We implemented robust error handling to address edge cases with dynamic UI feedback. We also integrated PCA, clustering, and plotly – topics that required independent research beyond what was covered in class.