

CS 4460 - P5
Minyoung Kim | Emily Yates
Data set: colleges.csv

Analytic Tasks

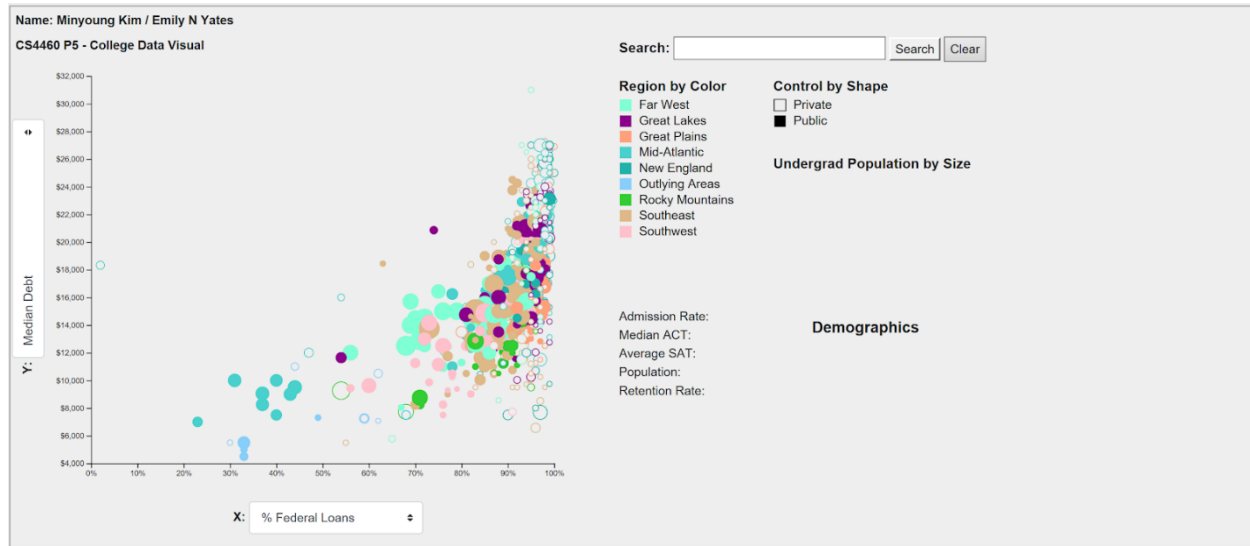
1. Filter
 - By clicking on a specific region and/or public or private, the user can find data cases satisfying specific conditions and hide others
2. Retrieve Value
 - By clicking on a specific data case in the graph, a user can find the attributes of that data case such as the admission rate, median ACT, average SAT score, population and retention rate in the side table.
3. Find Extremum and Anomalies
 - By analyzing the scatter plot and using the filtering tools or changing the axes, a user can find data cases possessing an extreme value of an attribute such as a college having an extremely high average cost or extremely low percent of federal loans.
4. Determine Range
 - When changing the axes, the range changes to fit the data set and a user can click on points on either end of the data set to find the span of values within a given attribute.
5. Characterize Distribution, Cluster, and Correlate
 - By analyzing the scatter plot, a user can be determined the general characteristics of the distribution, identify clusters of data cases, and recognize any correlation between attributes within any specific conditions.
 - The bar graph of college demographics also allows users to see a generalized distribution within a specific college.

Design Overview

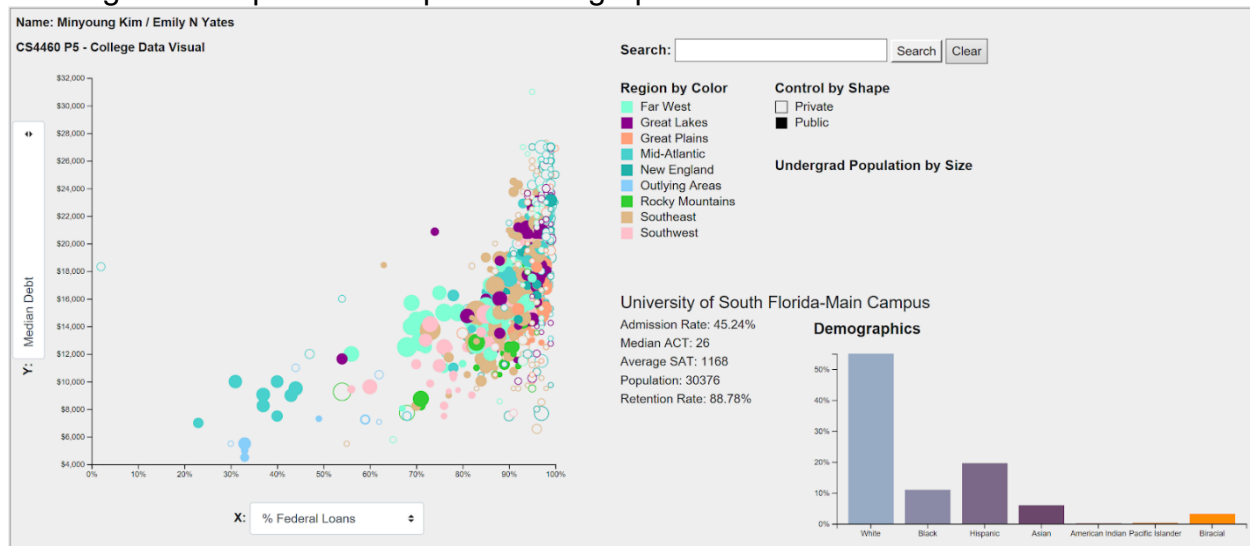
The main focus on of our info visualization was communication and comparison with support for both browsing and searching. The main objective of the scatter plot was to allow users to compare data cases. The added capabilities of filtering by region, public or private college, and changing the attributes on both the x and y axes just allow even more comparison in different conditions to occur. Even when communicating discrete details, the scatter plot still compares data sets. An example is, “which public college in the southeast has the highest median debt?”. This analytical question desires a finite answer but still involves comparison. The scatter plot however does also allow the user to browse the data set. Without having a specific question or hypothesis, the user can view overall trends, see clusters of data cases, and identify anomalies and extrema using the scatter plot. Users can compare public schools to private schools, one region to another, or one attribute to another. A user can also compare one school to the trend of the rest of the data. An example is, “Is Georgia Tech on the higher end of average cost among southeast colleges?”. The search bar tool supports users looking for an answer to a specific question about one school. An example is, “Where does Tulane lie on the graph comparing employment rates 8 years after graduation and average cost among colleges in the far west?”. The search bar allows easier identification of a certain college. The ability to click on a data point and the table to generate specific statistics and a bar graph of demographics is an implementation of details on demand. A user can use this info vis to answer the question of like “What is the average SAT score of UC Berkeley?” and “What is the dominant ethnicity at University of Arizona?”.

User Interface

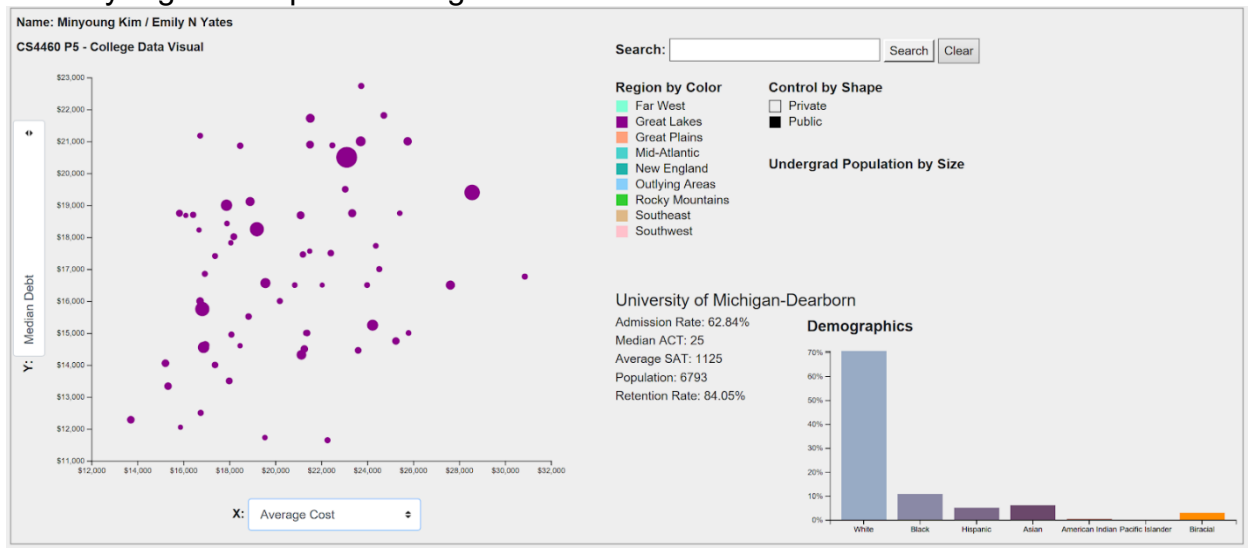
Main Screen



Clicking on one specific data point in the graph.



Filter by region and public college.



Searching for the specific college.

