

CS4460 – d3 lab7 description

Name: Ting Hao Wu

1. Dataset chosen: colleges.csv

As shown in figure 1, I chose nineteen columns to create scatterplots, bar chart, and information box for my lab7 visualization. The columns are: Name, Control, Region, Locale, Admission Rate, ACT Median, SAT Average, Undergraduate Population, % White, % Black, % Hispanic, % Asian, % American Indian, % Pacific Inlander, % Biracial, Average Cost, Retention Rate, Median Debt, and Median Family Income.

Figure 1: picked columns

Name	Control	Region	Locale	Admission	ACT Medi	SAT Averi	Undergrad	% White	% Black	% Hispanic	% Asian	% American Indian	% Pacific I	% Biracial	Average C	Retention	Median De	Median Family Income
Abilene C Private		Southwest	Mid-size C	0.4894	24	1087	3647	0.7069	0.076	0.1206	0.0099	0.0036	0.0005	0.0373	39611	0.7941	20698.5	75873.5
Adams Sta Public		Rocky Mtn	Remote To	0.526	20	999	2227	0.5123	0.0768	0.3026	0.0112	0.0153	0.0063	0.04	17759	0.5392	11000	34243
Adelphi U Private		Mid-Atlant	Large Subu	0.6814	22	1098	5000	0.5296	0.1062	0.1336	0.0756	0.0016	0.002	0.022	41559	0.8112	21325	54115
Adrian C Private		Great Lake	Distant To	0.5628	22	1007	1616	0.7605	0.1002	0.0217	0.0056	0.0037	0.0006	0.0384	40199	0.5811	14250	71008.5
Adventist U Private		Southeast	Mid-size C	0.1336	19	884	2064	0.4748	0.1434	0.2016	0.0562	0.0039	0.0044	0.0199	21358	0.7143	15236	49657.5
AITB Colle Private		Great Plain	Mid-size C	0.6887	0	0	1007	0.6306	0.0497	0.0586	0.0189	0.006	0.003	0.0218	22469	0	13125	53783
Alabama A Public		Southeast	Mid-size C	0.8989	17	823	4051	0.0279	0.9501	0.0089	0.0022	0.0012	0.001	0	18888	0.6314	19500	29039
Alabama S Public		Southeast	Mid-size C	0.5125	17	830	5354	0.0161	0.9285	0.0114	0.0015	0.0009	0.0007	0.0064	17400	0.6219	15854.5	24029.5
Albany Co Private		Mid-Atlant	Small City	0.634	27	1195	1075	0.6633	0.0381	0.0484	0.1284	0.0028	0.0019	0.0084	40718	0.8097	24930.5	80225
Albany Sta Public		Southeast	Small City	0.1992	19	887	3566	0.0443	0.91	0.0095	0.0017	0.0014	0.0003	0.002	19025	0.697	24500	29794
Albertus M Private		New Engla	Mid-size C	0.6821	0	830	1278	0.4476	0.277	0.1354	0.0141	0.0023	0	0.0141	41520	0.9256	26313	50062.5
Albion C Private		Great Lake	Fringe Tow	0.5639	25	1143	1284	0.8178	0.0343	0.0335	0.0156	0.0031	0.0008	0.0288	45690	0.8056	22500	85027.5
Albright C Private		Mid-Atlant	Small City	0.6231	23	1033	2333	0.6215	0.1582	0.0883	0.0257	0.006	0.0004	0.0146	46170	0.7146	24562.5	66577.5
Alcorn Sta Public		Southeast	Remote Ru	0.8291	18	851	3157	0.0326	0.9328	0.006	0.0016	0.0003	0.0025	0.0114	20330	0.6838	19696.5	21250
Alderson E Private		Southeast	Distant To	0.4321	22	989	1032	0.7403	0.1609	0.0397	0.0116	0.0068	0	0.0107	33057	0.5761	9500	48013
Alfred Uni Private		Mid-Atlant	Remote To	0.698	25	1106	1905	0.6751	0.0819	0.0688	0.0163	0.001	0	0.0262	38830	0.7463	21500	69334
Allegheny Private		Mid-Atlant	Distant To	0.6487	26	1182	2126	0.802	0.0437	0.0616	0.0254	0.0005	0	0.0405	50044	0.8545	27000	85933
Alma Coll Private		Great Lake	Distant To	0.6791	25	1100	1385	0.8657	0.0217	0.0303	0.0166	0.0087	0	0.0245	43157	0.7885	21500	74312
Alvernia U Private		Mid-Atlant	Small City	0.7762	20	976	2169	0.7045	0.1415	0.0742	0.0143	0.0005	0.0005	0.0161	39999	0.7105	22500	53135
Alverno C Private		Great Lake	Large City	0.7887	20	950	1833	0.5428	0.1768	0.1811	0.0475	0.012	0.0016	0.0316	30496	0.7173	25500	38705
American J Private		New Engla	Mid-size C	0.6779	19	885	1513	0.4382	0.2538	0.0826	0.0152	0.004	0.0073	0.0443	40803	0.6482	20726	49259
American U Private		Mid-Atlant	Large City	0.4312	28	1258	6829	0.5671	0.064	0.1012	0.0638	0.0018	0.0004	0.0499	55159	0.8821	19500	86431
Amherst C Private		New Engla	Large Subu	0.1428	32	1434	1785	0.4174	0.1165	0.1283	0.1255	0.0011	0	0.056	59060	0.9762	10453	85922
Anderson U Private		Great Lake	Small City	0.5545	23	1053	1862	0.811	0.0575	0.0269	0.007	0.0032	0.0027	0.0043	37937	0.734	23000	72723
Anderson U Private		Southeast	Small City	0.6502	24	1079	2381	0.8202	0.0962	0.0328	0.0101	0.0067	0.0004	0	35340	0.734	19500	58707.5

2. Analytic tasks:

For this project, I hope I can create a visualization that could help the prospective students and their parents, and anyone who intends to go to college. By using this visualization, I hope users can be able to do a couple of things to build up a better understanding of the colleges. Following are the things I would like the user to do:

- Be able to see the relationship between the SAT/ACT scores and the admission rate. To have an idea that what range of score they want to get.
- Be able to see the relationship between the college average cost and the median household income of the user's family.
- Be able to see the correlation between the SAT/ACT scores and the median household income of the user's family.
- Be able to check out the ratio of race/ethnicity at users' targeting college.
- Easy to check out detailed information (in number and string) such as control, region, undergrad population, costs, incomes, and test scores of users targeting college.

3. Design overview:

The main goal of this visualization is to let the user finds the college information to help

them have a better understanding of their targeting school. Of course, the visualization will be mainly used by students and their parents, and some other people such as elders or office workers that may want to seek higher education and knowledge. College is an important place for people to learn knowledge and to be trained and skilled in different fields. However, there are many factors to be concerned:

- Average cost and the median debt: the most important factor which lets users know if they can afford to go to their targeting school.
- Test score: let users know what scores, SAT, or ACT, they should achieve to get the ticket of the particular colleges.
- Admission rate: let users have an idea of the possibility to get rejected even they pass the threshold scores.
- Location: let users do some trade-off by the distance factor.
- Race/ethnicity: let users know the composition of the ethnicity at the targeting college.
- Is there some trend between SAT/ACT scores and admission rate? When the score gets close to the extreme value, is it affects the admission rate?
- Is there some trend between SAT/ACT scores and family incomes? If a family has a higher household income, would their child have relatively higher scores?

For creating this visualization, I used different d3 skills we learned in class to help the users to have better results from the potential analytics I listed above. For the scatter plots, I implemented dropdown menus for users to change the axes of the plot so they can see the different relations and distribution upon the data. I also added a highlight and original data buttons for the user to switch between the plot that highlights the location of the selected college dot and the plot that shows all college dots. Following are the operations:

Figure 2: original plots

Figure 3: axes changed – different options

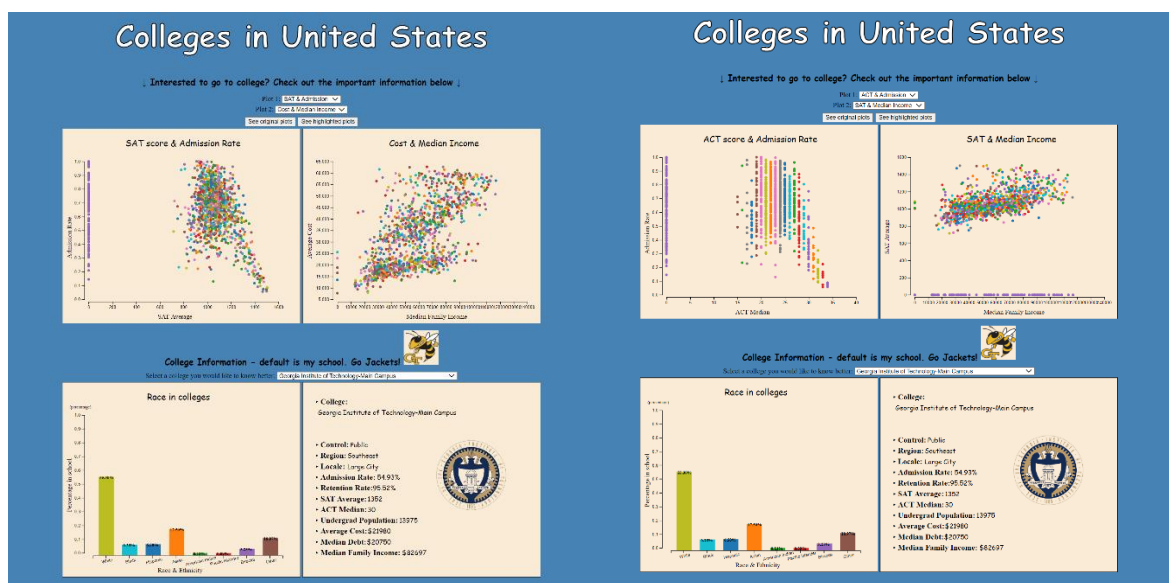
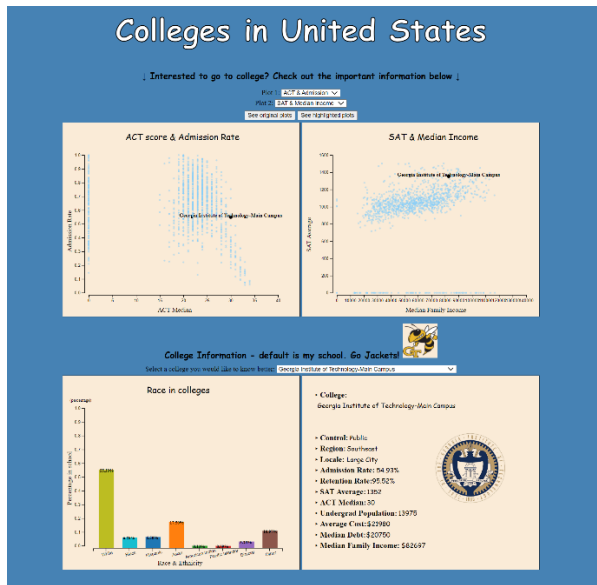


Figure 4: college dot highlighted



There is another dropdown menu that is related to the scatter plots. Users can find and select the college they want. When a college is selected, the highlight function will be triggered which allows users to see where the college is located on the scatter plot. Also, when a college is selected, the other two charts, the bar chart, and the information box will be updated as well. The bar chart shows the ratio of the race/ethnicity at the selected college and it will have a simple animation when the data is updating. The information box, on the other hand, displays all the detailed information about the selected college. Following are the operations:

Figure 5: Georgia Institute of Technology is selected

Figure 6: Stanford University is selected

