U.S. colleges data analysis

Pranav Joshi

C.S.E at IITGN (student)

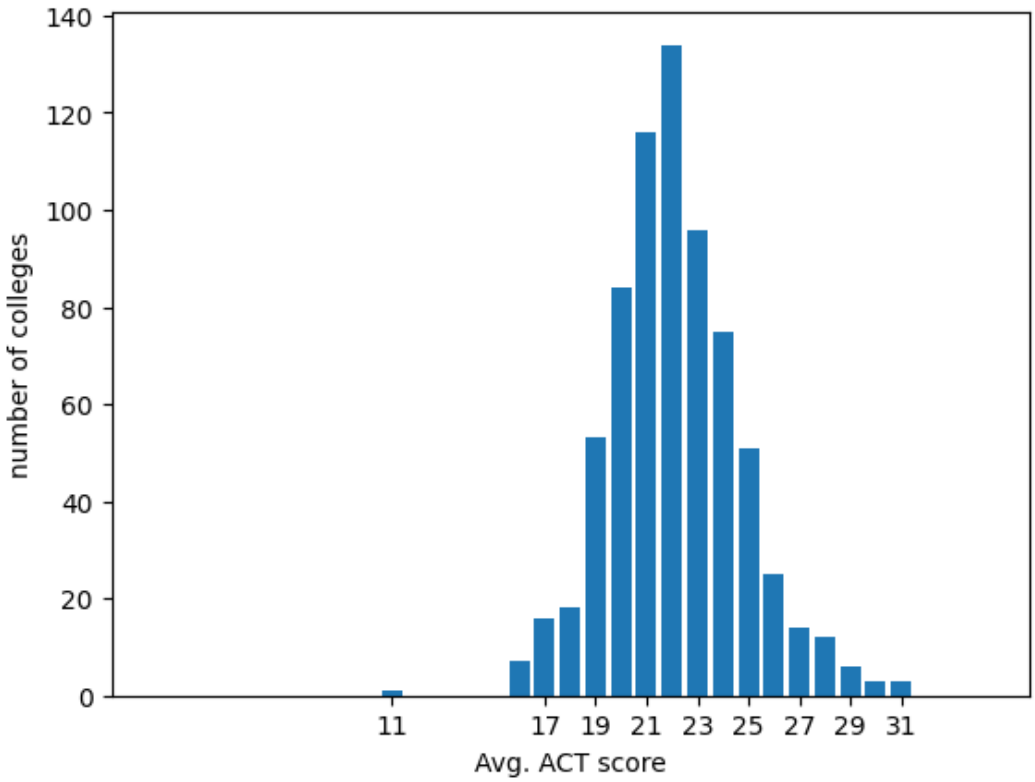
Roll no.: 22110197

Introduction:

The U.S. colleges data-set contains information like average ACT scored, graduation rate,etc. This document is a data narrative based on this data-set.

Distribution of colleges based on different factors.

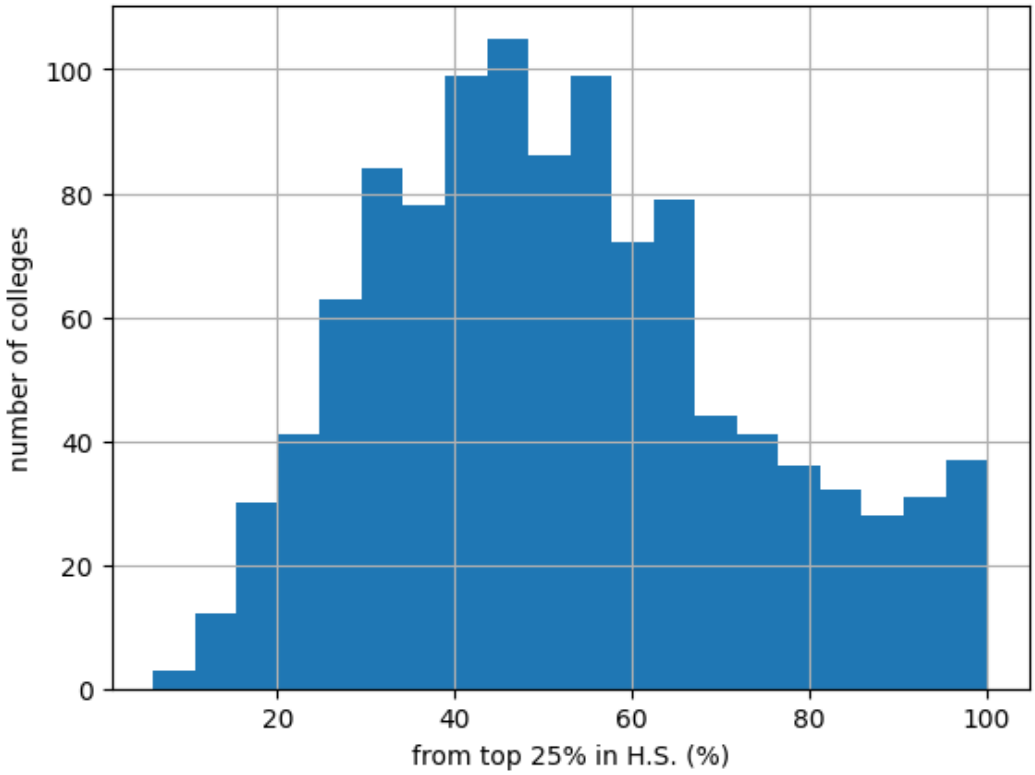
ACT scores



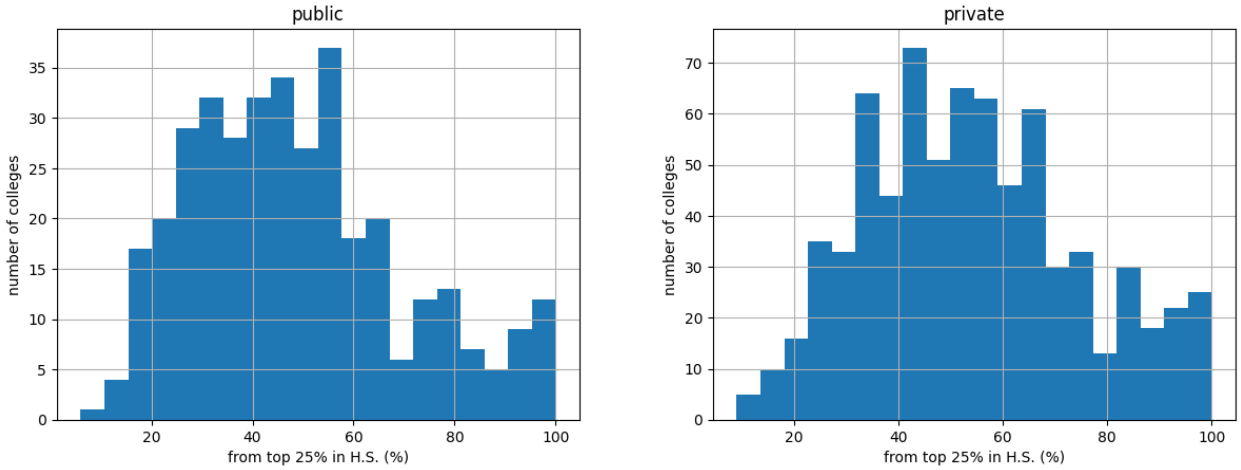
Note: The Avg. ACT score data had all values rounded to integers.

From the graph, 22 is the most probable average ACT score, and there are almost no colleges with average rating below 16.

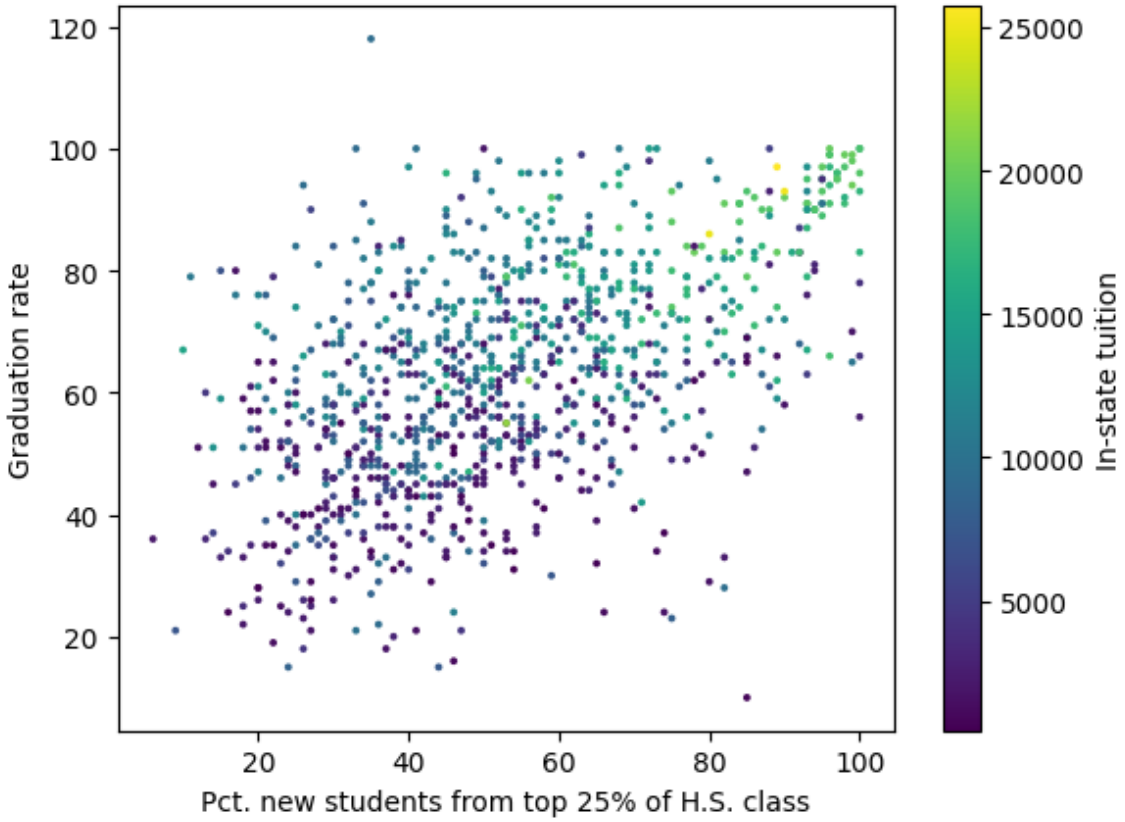
%age students from top 25% of H.S. class



The graph has an unusual peak at 100%. This leads to the conclusion that there are two types of colleges, ones who will let in less proficient students too and ones which wont. Although this suggests that the two types are public colleges and private colleges, that isn't the case, as both public and private colleges also have similar graphs.

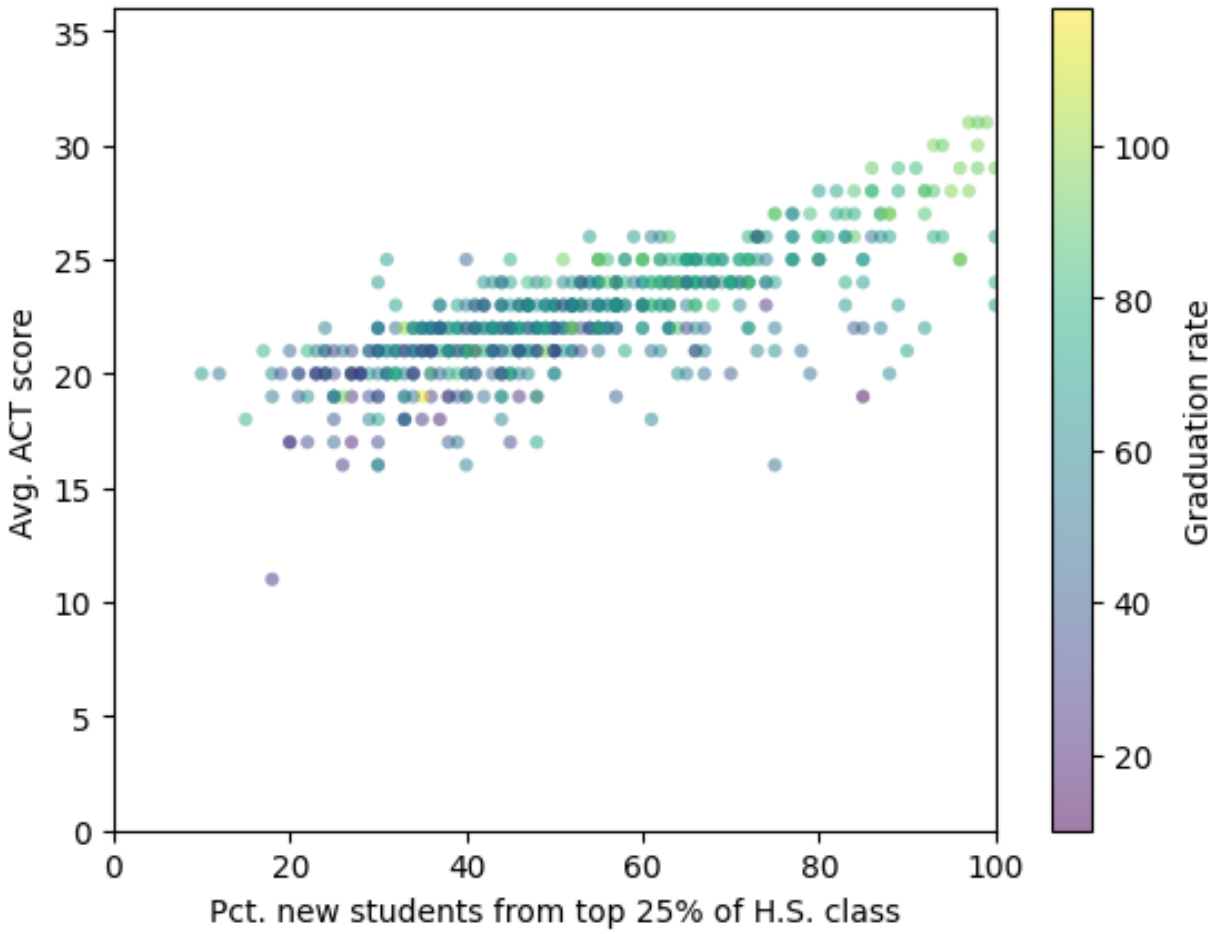


These two types also differ in other factors, like graduation rate, tuition fees and avg. ACT scores. For example:



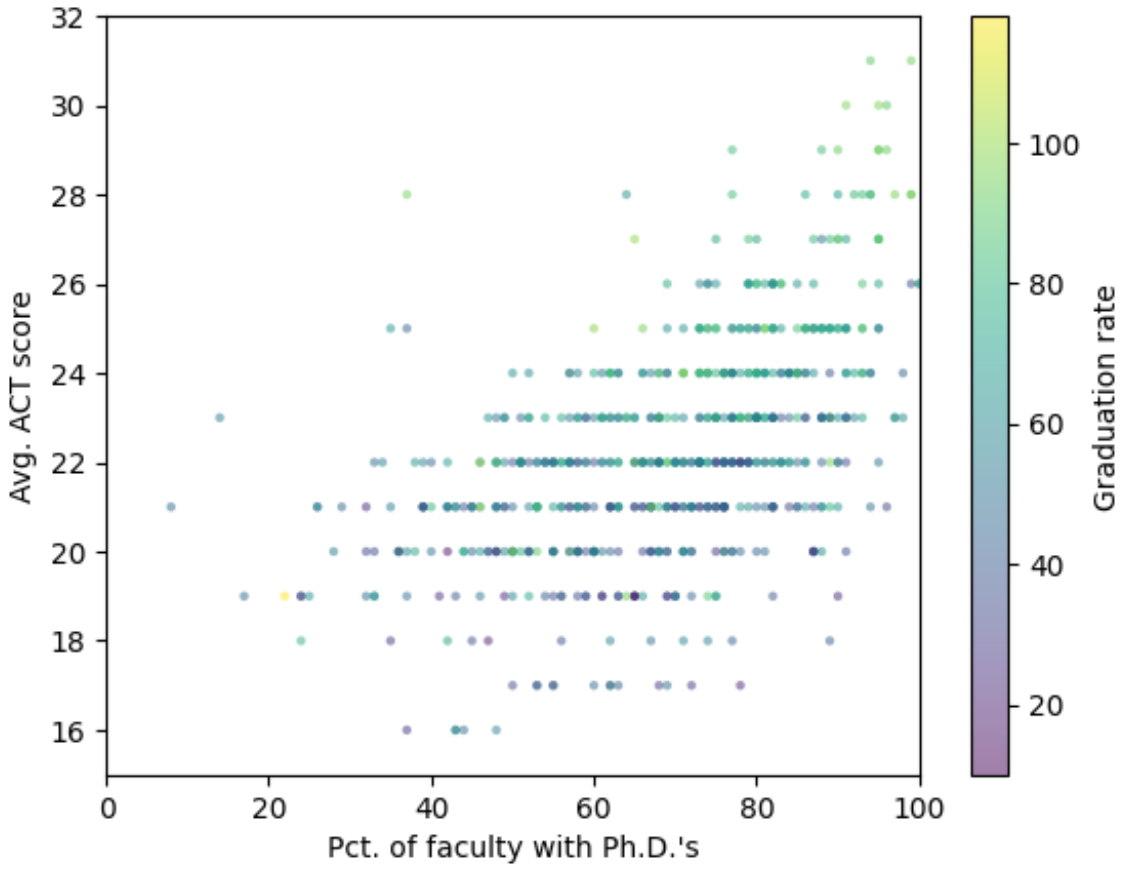
In this scatter plot, we can see two regions with high density of colleges. The upper right one is made of the “prestigious” colleges which have high tuition fees and high avg. ACT scores.

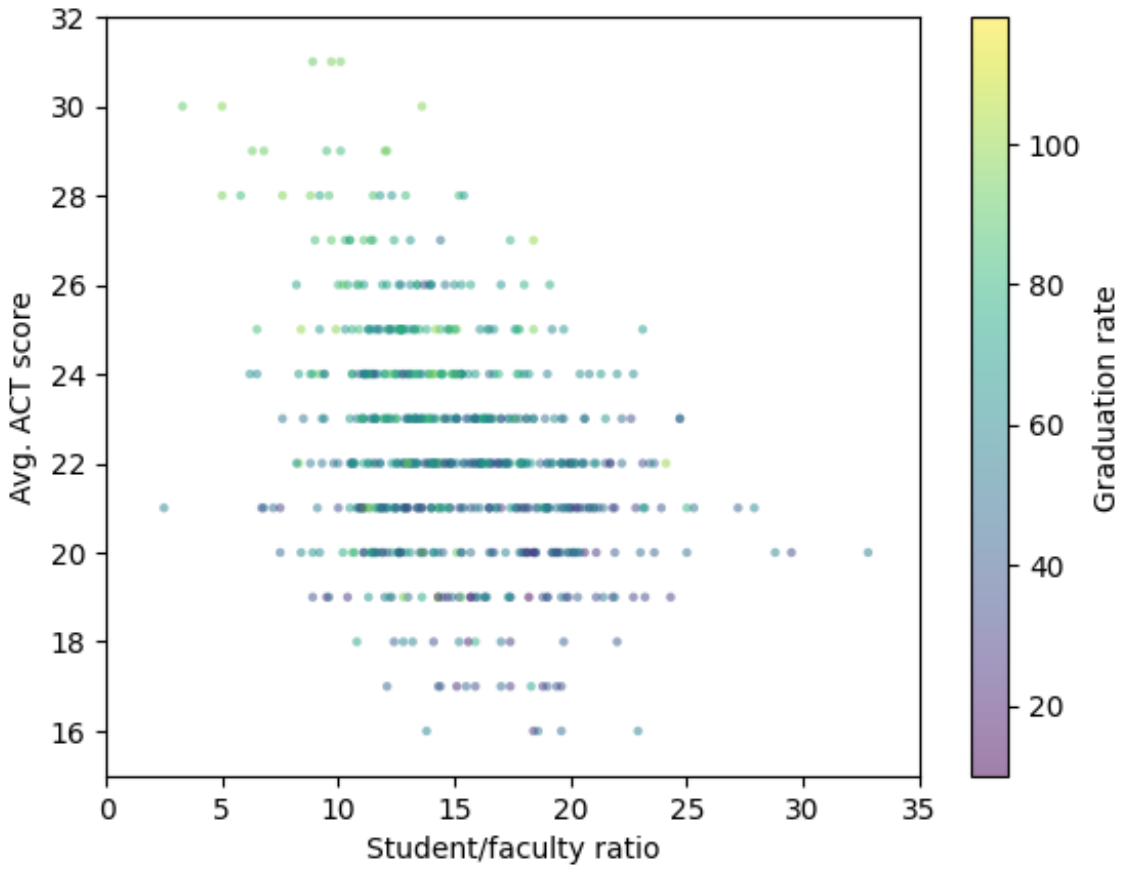
The avg. ACT scores being high is no surprise, as avg. ACT score is linked to percentage of students from top 25% in high school, as described by this scatter plot:



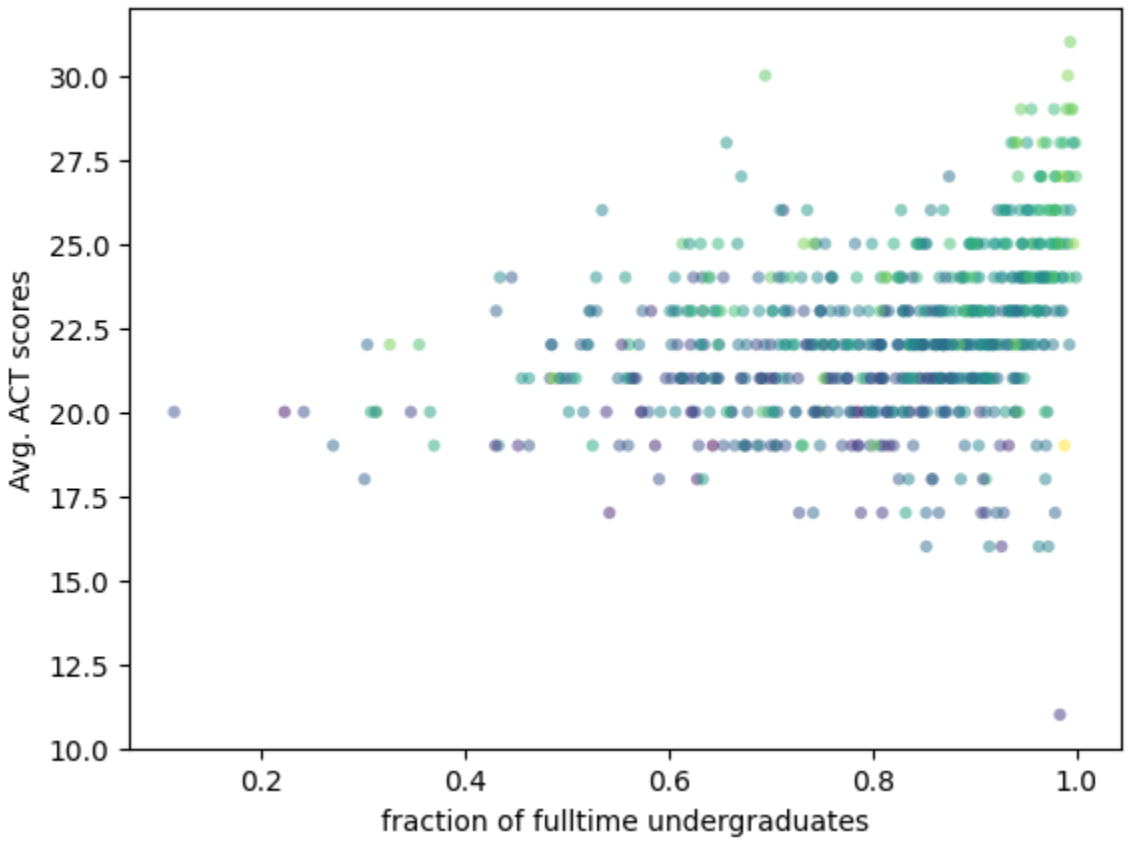
Effect of different factors on graduation rate.

As is evident from the previous graph, graduation rate increases with the initial proficiency of the students (their ACT scores). What about other factors? A lot of these factors contribute to the prestige of universities where bright students with high scores go, that's why I am showing the effect of these factors, alongside the avg. ACT scores, so that we can look at all colleges with similar ACT scores and then see how different conditions affect graduation rate for only those colleges.

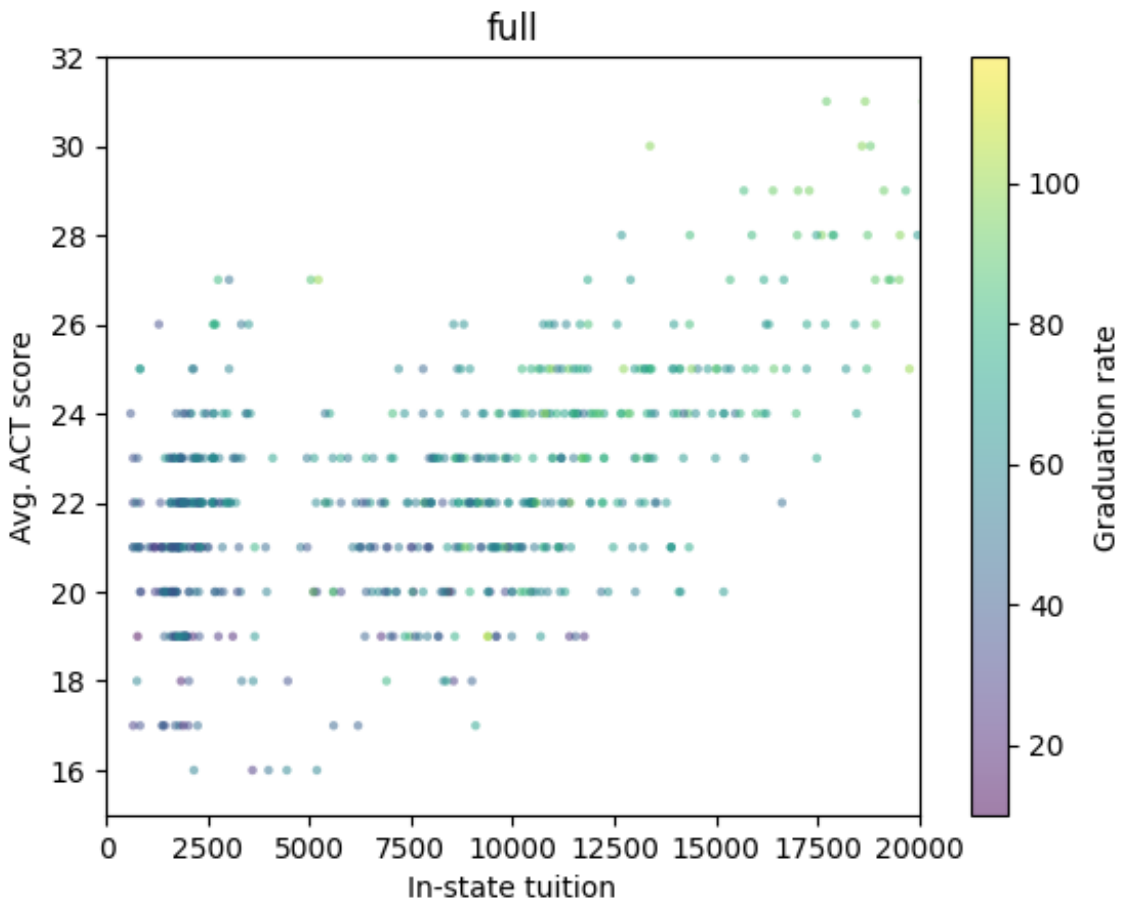

Avg. ACT score
Pct. of faculty with Ph.D.'s
Graduation rate
32
30
28
26
24
22
20
18
16
0
20
40
60
80
100
100
80
60
40
20



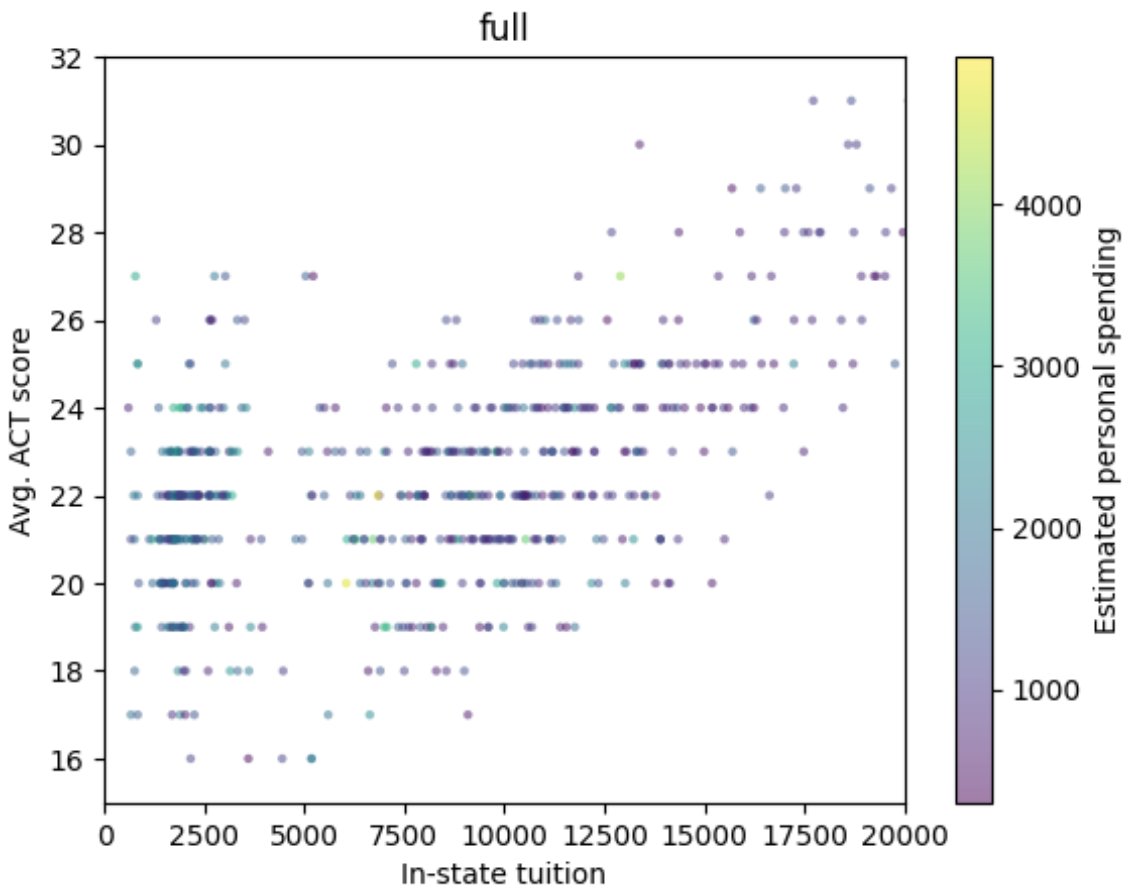
These is no strong correlation between percentage of faculty with Ph.D.’s or Student/faculty ratio with graduation rate for a constant avg. ACT score.



There is a slight dependence of graduation rate on the fraction of students who are full-time undergraduates.

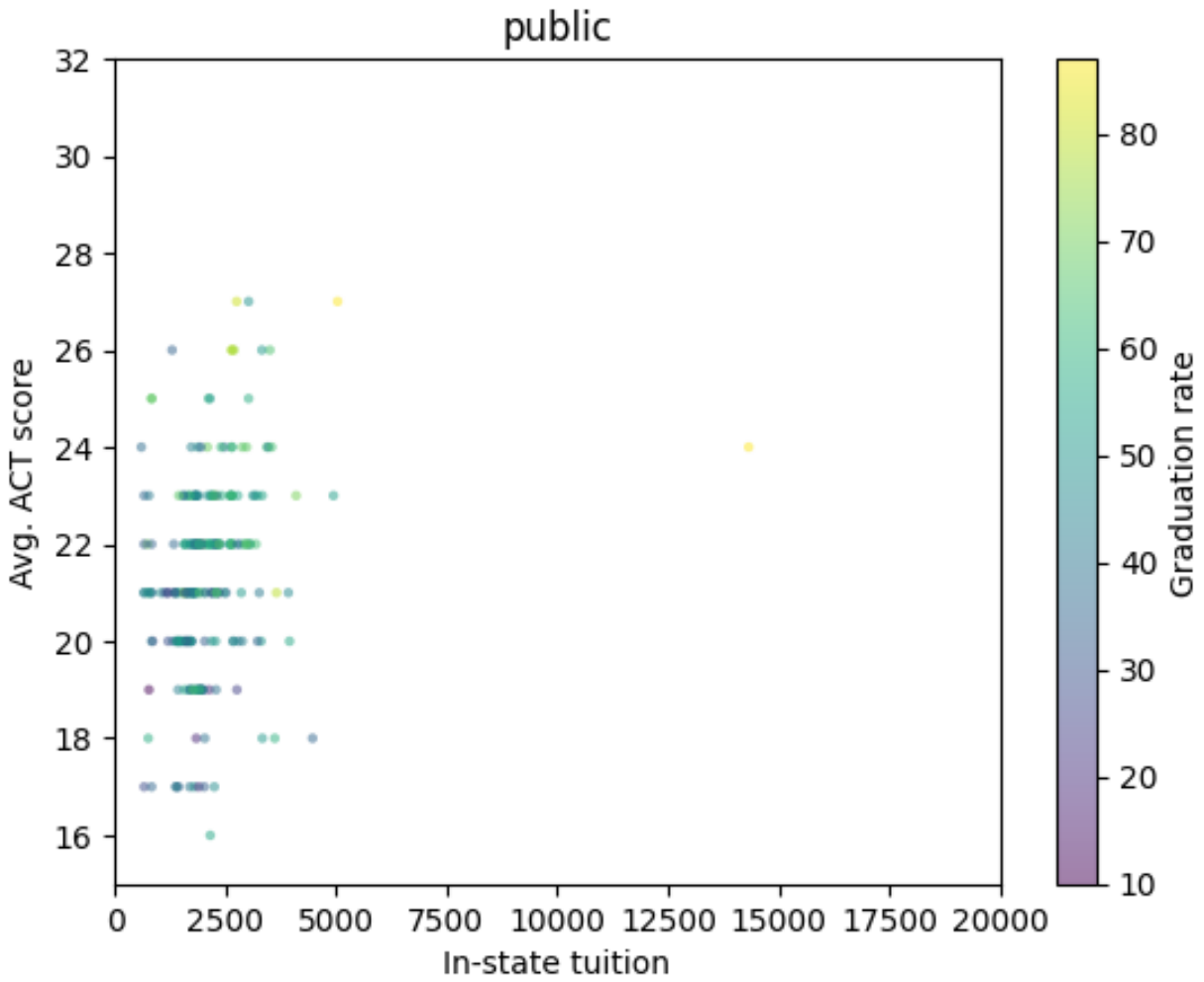


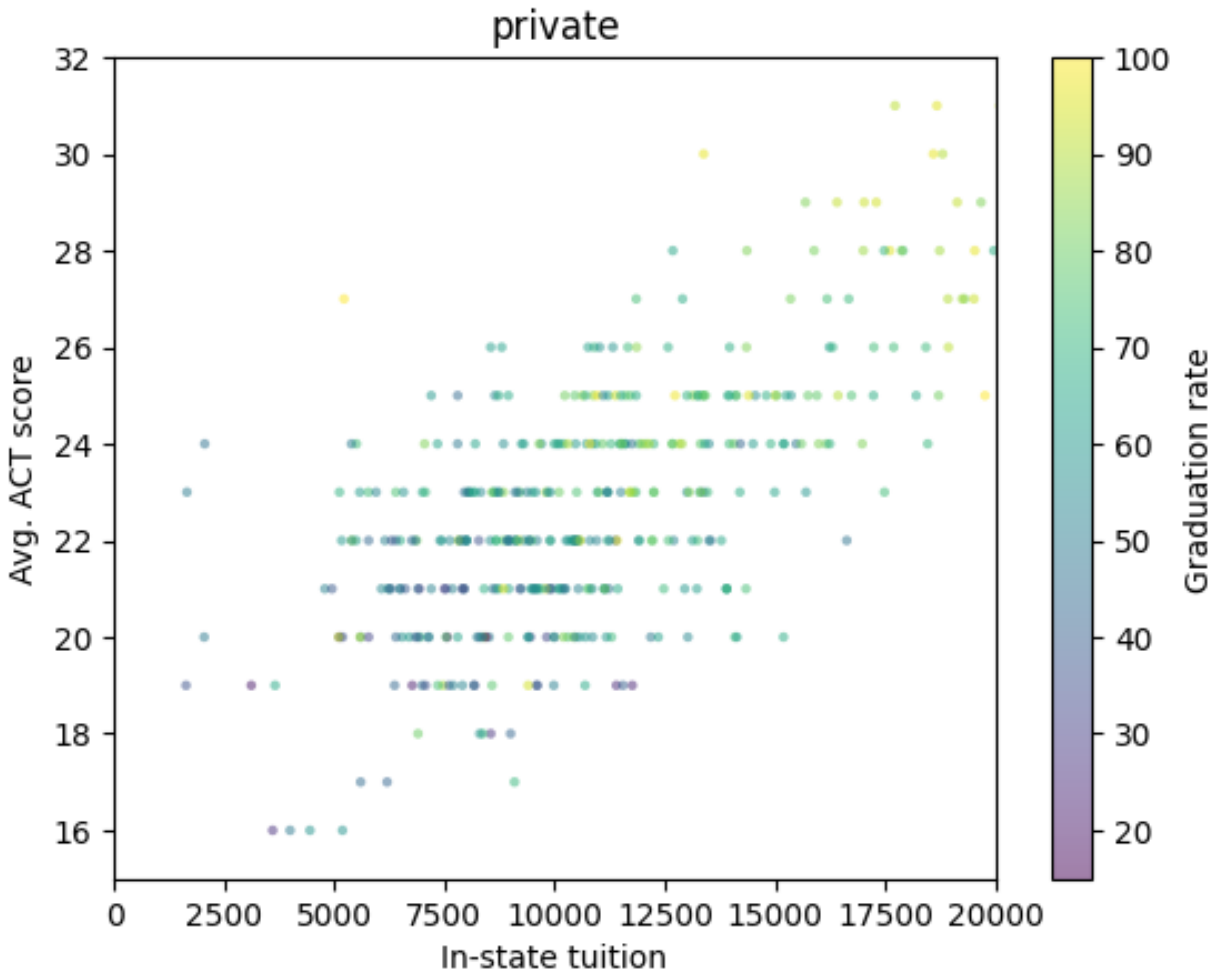
There is a strong dependence of graduation rate with tuition fees. This is probably because students who have paid a lot of fee amount to study in college did so because they were already interested in the material (either directly, or for the benefits it brings to their lives). To show that they aren't just from well off families, in which case they could afford going to an expensive college without having real interest in the material, let's look at the same graph, but with personal spending as the color:



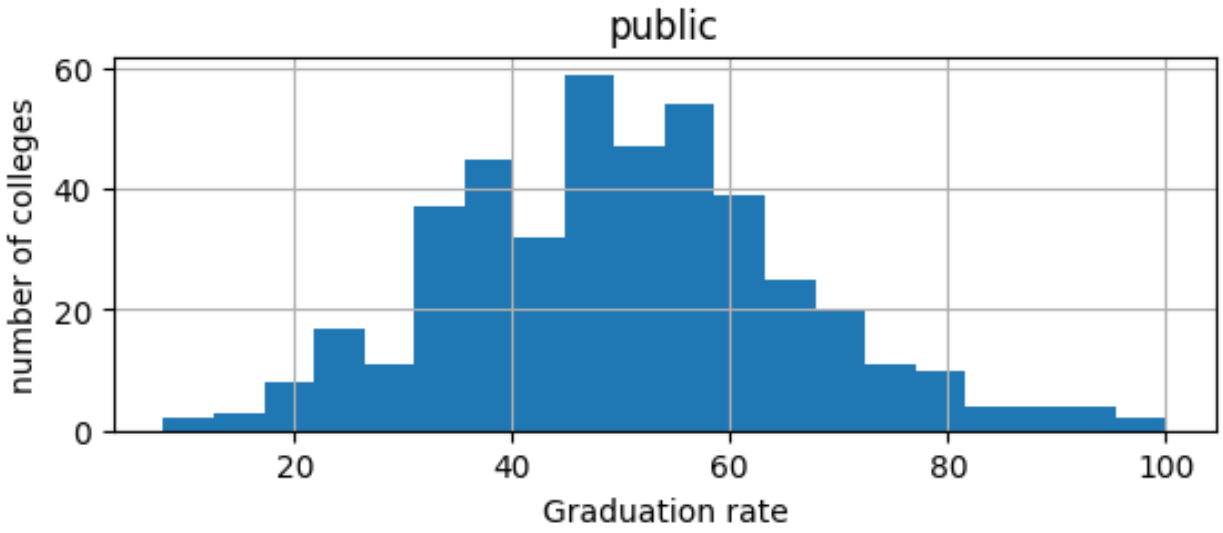
As you can see, the colleges where tuition fees are high, personal spending is low. This shows that the students there don't usually have easy access to large amounts of money, and are instead taking loans for their education, as described in this article [1].

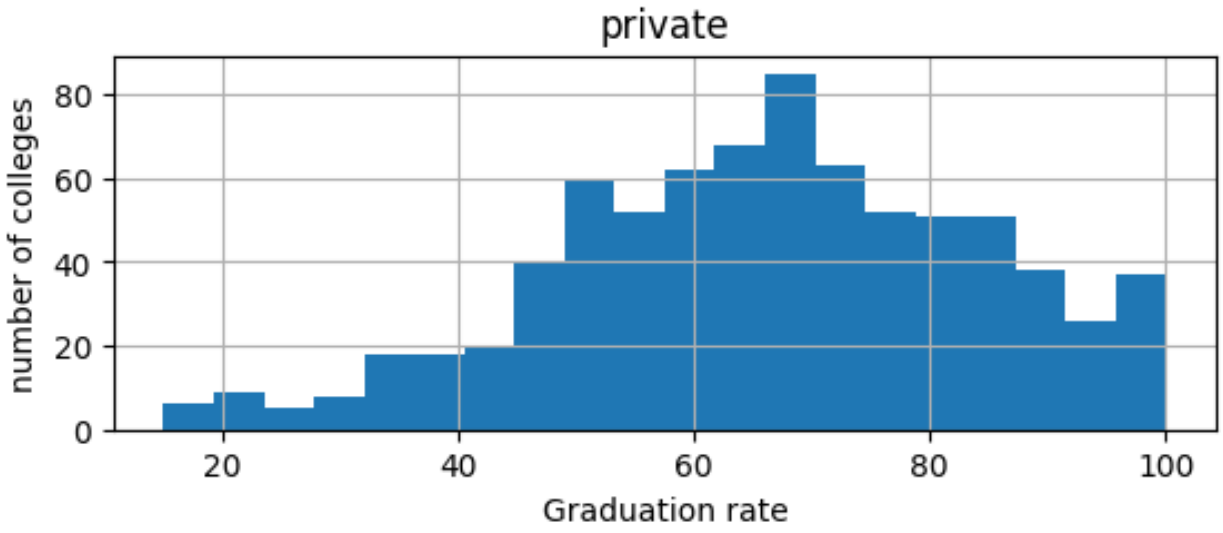
Also, there are two clusters in the plot. These in-fact are the public and private colleges. The left cluster contains public colleges mostly, which provide education at much lower costs.


public
Avg. ACT score
In-state tuition
Graduation rate
32
30
28
26
24
22
20
18
16
0
2500
5000
7500
10000
12500
15000
17500
20000
80
70
60
50
40
30
20
10



This further suggests that students who do not have to pay massive tuition fees are less committed to that particular course. This doesn't take into account all the students who transferred, or ones who just couldn't afford a loan. Since public colleges are less expensive, a lot of students get admitted in them for the first year, and later drop when they realise that college isn't for them, or get transferred. On average, only about half of the students admitted to a public college graduate,



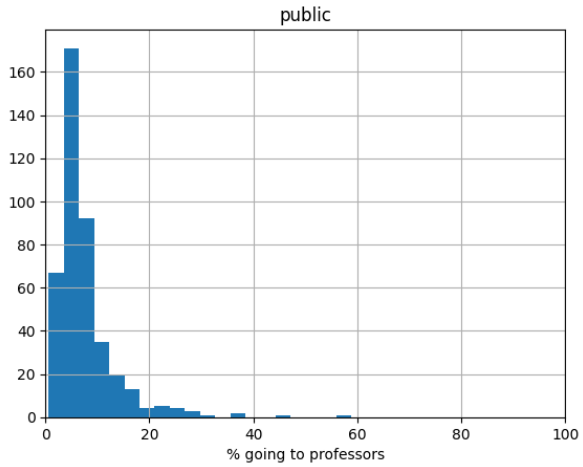


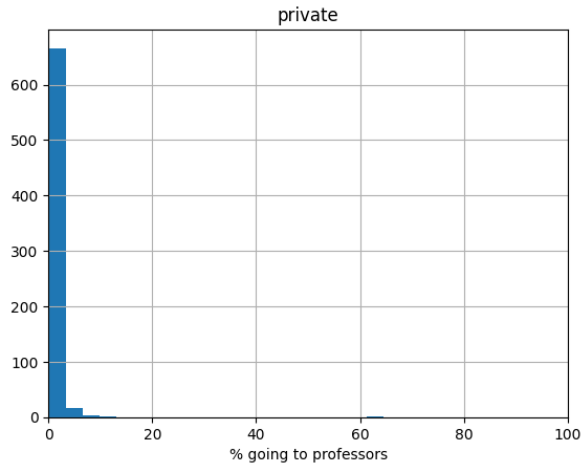
and about 3/4 of admitted students graduate in private colleges.

This [2] article goes in more detail about the same issue.

Why is college so expensive?

A good guess, after looking at the tuition fees, would be that most of the money goes to professors. But actually, the amount of money going to professors' salaries is on average 3.7% of the total tuition fees (summed over all the students). This average is 7.6% for only the public colleges, 1.3% for only private.





Most of the money goes into maintenance of the college's infrastructure.

Summary

- 22 is the most probable average ACT score, and there are almost no colleges with average rating below 16
- There are two types of colleges (not public and private), ones that let in less proficient students too and ones which wont.
- Avg. ACT score is proportional to percentage of students from top 25% in high school
- These is no strong correlation between percentage of faculty with Ph.D.'s or Student/faculty ratio with graduation rate for a constant avg. ACT score.
- There is a slight dependence of graduation rate on the fraction of students who are full-time undergraduates.
- There is a strong dependence of graduation rate on tuition fees.
- Students in expensive colleges don't usually have easy access to large amounts of money, and instead take loans for their education.
- On average, only about half of the students admitted to a public college graduate, and about 3/4 of admitted students graduate in private colleges.

Libraries used

- matplotlib python library
- numpy python library
- pandas python library

References

[1]: Student loans in the United States, Wikipedia

URL: https://en.wikipedia.org/wiki/Student_loans_in_the_United_States

[2]: College grad rates are ‘bad data’, *Daniel de Vise,* January 26, 2012

URL: https://www.washingtonpost.com/blogs/college-inc/post/college-grad-rates-are-bad-data/2012/01/26/gIQAfmdKTQ_blog.html